UC Berkeley
UC Berkeley Electronic Theses and Dissertations

Title
Empirical Analyses in Agricultural and Resource Economics

Permalink
https://escholarship.org/uc/item/3t6599jx

Author
Stevens, Andrew William

Publication Date
2017

Peer reviewed|Thesis/dissertation

eScholarship.org Powered by the California Digital Library
University of California

Empirical Analyses in Agricultural and Resource Economics

by

Andrew William Stevens

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Agricultural and Resource Economics

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Peter Berck, Chair
Professor David L. Sunding
Chancellor's Associate Professor Solomon M. Hsiang

Spring 2017

Empirical Analyses in Agricultural and Resource Economics

Copyright 2017
by
Andrew William Stevens

Abstract

Empirical Analyses in Agricultural and Resource Economics

by

Andrew William Stevens

Doctor of Philosophy in Agricultural and Resource Economics

University of California, Berkeley

Professor Peter Berck, Chair

Agriculture has played a profound and unique role in humanity's development. We are dependent on agriculture for the vast majority of our food supply, and have so far been successful at increasing agricultural production to meet rising demand. At the same time, agriculture is the largest and most direct way that humans have altered our planet's landscape and natural environment. Indeed, over half of all land in the United States is used for some agricultural purpose. In this dissertation, I explore three different aspects of human-environmental-agricultural interdependency in the United States. In the first chapter, I study agricultural workers' wage-responsiveness under different environmental conditions on California blueberry farms. In the second chapter, Fiona Burlig and I study the effect of human social networks on agricultural technology adoption in three upper-Midwestern states. Finally, in the third chapter, I study how the location of ethanol refineries in the US Corn Belt affects crop choice decisions and nutrient runoff. Each of these chapters highlights a different interaction between human economic systems (wages, social networks, renewable fuel policy), environmental conditions (temperature, nitrogen application/runoff), and agricultural enterprises (specialty crop labor productivity, adoption of fertilizer, crop rotations). I utilize a similar empirical strategy in each chapter, employing fixed effects and other panel data techniques to control for time-invariant determinants of productivity, technology adoption, and optimal crop choice, respectively. This dissertation highlights the benefits of panel data methods in agriculture, especially in the modern era of abundant micro-level data.

In the first chapter, I study how agricultural laborers' productivity responds to changes in the piece rate wage they are paid: a wage paid per unit of output rather than per unit of time. Specifically, I exploit quasi-experimental variation to estimate the elasticity of labor productivity with respect to piece rate wages by analyzing a high-frequency panel of over 2,000 blueberry pickers on two California farms over three years. To account for endogeneity in the piece rate wage, I use the market price for blueberries as an instrumental variable. I find that picker productivity is very inelastic on average, and I can reject even modest elasticities of up to 0.7. However, this average masks important heterogeneity across outdoor working

conditions. Specifically, at temperatures below 60°F, I find that higher piece rate wages do in fact induce increases in labor productivity. This is suggestive evidence consistent with a model where at moderate to hot temperatures, workers face binding physiological constraints that prevent them from exerting additional effort in response to higher wages. This insight has important implications for understanding how climate change will affect the agricultural labor sector.

In the second chapter, Fiona Burlig and I use historical data and a natural experiment to study the effect of social networks on agricultural technology adoption. We present a model of the effects of social network size on information and technology take-up and test its implications using a unique natural experiment in the mid-20th century US Midwest. We find that social network expansions, in the form of mergers between congregations of the American Lutheran Church, led to increased rates of agricultural technology adoption among farmers. In counties that experienced a merger, the number of farms using nitrogen fertilizer increased by over 7% and the total fertilized acreage increased by over 13% relative to counties without a merger. We provide evidence that these effects are driven by increased information sharing between farmers as a result of these congregational mergers.

In the third chapter, I study how the location of ethanol refineries within the US Corn Belt affects farmers' land use decisions. Ethanol production in the United States, driven by federal renewable fuel policy, has exploded over the past two decades and has prompted the construction of many ethanol refineries throughout the US Corn Belt. These refineries have introduced a new inelastic demand for corn in the areas where they were built, reducing basis for nearby farmers and effectively subsidizing local corn production. I explore whether and to what extent the construction of new ethanol refineries has actually increased local corn acreage. I also explore some environmental effects of this acreage increase. Using a thirteen year panel of over two million field-level observations in Illinois, Indiana, Iowa, and Nebraska, I estimate a net increase of nearly 300,000 acres of corn in 2014 relative to 2002 that can be attributed to the placements of new ethanol refineries. This increase comprises approximately 0.75% of the total 2014 corn acreage within my dataset. Furthermore, this effect is separate from the general equilibrium effect of ethanol policy increasing aggregate demand for corn. Back-of-the-envelope calculations suggest that over 21,000 tons of the nitrogen applied to fields in my sample in 2014 can be attributed to refinery location effects. Essentially all of these observed effects occur only in areas within 30 miles of an ethanol refinery, suggesting that refineries have meaningful localized impacts on land use and environmental quality such as nitrate runoff.

Contents

Contents i

List of Figures ii

List of Tables iv

Acknowledgements v

1 Temperature, Wages, and Agricultural Labor Productivity 1
- 1.1 Introduction . . . 1
- 1.2 Theoretical Framework . . . 4
  - 1.2.1 A model of optimal effort under piece rate wages . . . 4
  - 1.2.2 Previous literature on piece rate wages . . . 8
- 1.3 Context: California Blueberries . . . 10
- 1.4 Data Sources and Description . . . 13
  - 1.4.1 Employee-level production figures . . . 13
  - 1.4.2 High-frequency temperature readings . . . 19
  - 1.4.3 State-level market prices . . . 22
  - 1.4.4 Additional summary statistics . . . 25
- 1.5 Empirical Strategy . . . 28
- 1.6 Results . . . 30
- 1.7 Discussion . . . 47
  - 1.7.1 Wage effects . . . 47
  - 1.7.2 Direct temperature effects . . . 49
    - 1.7.2.1 Existing literature on temperature and other environmental conditions . . . 49
    - 1.7.2.2 Temperature effects on California blueberry pickers . . . 51
  - 1.7.3 Implications for California blueberry growers . . . 51
- 1.8 Conclusion . . . 53

2 Reap What Your Friends Sow: Social Networks and Technology Adoption 54
- 2.1 Introduction . . . 54

2.2 Context . . . 56
2.2.1 Church and congregational mergers . . . 59
2.3 Theoretical framework . . . 61
2.4 Data . . . 62
2.4.1 Church data . . . 62
2.4.2 Agriculture data . . . 64
2.5 Empirical strategy . . . 67
2.6 Results . . . 70
2.7 Discussion . . . 77
2.7.1 Placebo test . . . 77
2.7.2 Alternative explanations . . . 78
2.8 Conclusion . . . 80

3 Fueling Local Water Pollution: Ethanol Refineries, Land Use, and Nitrate Runoff. 81
3.1 Introduction . . . 81
3.2 Model of Optimal Crop Choice . . . 84
3.3 Data . . . 86
3.3.1 Cropland Data Layer . . . 86
3.3.2 Common Land Unit . . . 88
3.3.3 Ethanol Refineries . . . 89
3.3.4 Summary Statistics . . . 91
3.4 Econometric Methods . . . 92
3.5 Results . . . 94
3.6 Conclusion . . . 101

Bibliography 102

A A Model of Optimal Piece Rate Wages 116

List of Figures

1.1 A Model of Optimal Worker Effort . . . 7
1.2 Farm Locations . . . 10
1.3 Average Productivity by Picker . . . 14
1.4 Average Productivity by Day . . . 15

1.5 Time of Production . . . 17
1.6 Hours Worked . . . 17
1.7 Days Worked . . . 18
1.8 Time-Weighted Average Temperatures . . . 20
1.9 Temperature at Time of Production . . . 21
1.10 Market Prices and Piece Rate Wages . . . 23
1.11 Daily Blueberry Production . . . 24
1.12 Average Productivity Across Observable Variables . . . 26
1.13 How Frequently Does the Minimum Wage Bind? . . . 27
1.14 Effect of Temperature on Worker Productivity . . . 31
1.15 Effects of Daily Temperature on Daily Productivity . . . 38
1.16 Effects of Daily Temperature on Labor Supply . . . 40
1.17 Effect of Temperature on Worker Productivity Without Possible Shirkers . . . 43
1.18 Effect of Temperature on Worker Productivity Without Transient Workers . . . 45
1.19 Graphical Summary of Findings . . . 48

2.1 Fertilizer Applied in US Corn Production . . . 57
2.2 Spatial Distribution of Lutherans . . . 58
2.3 Lutheran Church Mergers . . . 60
2.4 Treated, control, and excluded counties . . . 64
2.5 Parallel Pre-Trends in Farms Using Fertilizer . . . 69
2.6 Change in farms using fertilizer, 1959 to 1964 . . . 72
2.7 Randomization Inference – Number of Farms Using Fertilizer . . . 76

3.1 Growth of US Ethanol Production and Refineries, 2002–2014. . . . 82
3.2 Model prediction of the effect of distance to nearest ethanol refinery on probability a field is planted to corn . . . 86
3.3 2014 Cropland Data Layer (CDL) with ethanol refinery locations. . . . 87
3.4 Detail of Cropland Data Layer (CDL) and Common Land Unit (CLU) data: Iowa, 2014 . . . 89
3.5 Ethanol refinery locations in 2002 (black) and 2014 (gray) . . . 90
3.6 Distributions of field-level distance-to-nearest-refinery in 2003 and 2014. . . . 91
3.7 Unconditional probability of growing corn: specification (1) . . . 95
3.8 Probability of growing corn conditional on growing corn in the prior year: specification (2) . . . 96
3.9 Probability of growing corn conditional on growing soy in the prior year: specification (3) . . . 97
3.10 Probability of growing corn conditional on growing something other than corn or soy in the prior year: specification (4) . . . 98
3.11 Effect of distance to nearest ethanol refinery on corn acreage . . . 99

List of Tables

1.1 Summary Statistics 16
1.2 Effects of Wage and Temperature on Worker Productivity 35
1.3 Piecewise Linear Effect of Temperature on Worker Productivity 36
1.4 Testing Market Price as an Instrument for Piece Rate Wage 37
1.5 Effects of Wage and Temperature on Labor Supply 39
1.6 Effect of Piece Rate Wage on Worker Productivity by Temperature 41
1.7 Effect of Piece Rate Wage on Worker Productivity by Temperature: OLS 42
1.8 Piece Rate Wage Effect Without Possible Shirkers 44
1.9 Piece Rate Wage Effect Without Transient Workers 46

2.1 Pre-treatment Summary Statistics 66
2.2 Test of Differential Trends 68
2.3 Impact of Congregational Mergers on Fertilizer Use – Farms 71
2.4 Impact of Congregational Mergers on Fertilizer Use – Fertilizer and Lime 73
2.5 Impact of Congregational Mergers on Fertilizer Use – Corn 73
2.6 Impact of Congregational Mergers on Land Use and Irrigation 74
2.7 Impact of Congregational Mergers on Inputs – Placebo Tests 78
2.8 Impact of Congregational Mergers on Capital 79

3.1 Summary Statistics 92
3.2 Probability of a field being planted to corn 100

Acknowledgments

I am thankful for the many people who have supported me during my graduate career, and who have directly offered feedback and support for the research contained in this dissertation. In particular, I am eternally indebted to the faculty and graduate students of the Department of Agricultural and Resource Economics at the University of California – Berkeley. My colleagues epitomize scholarly excellence, and have become some of my closest friends.

In relation to the first chapter of this dissertation, *Temperature, Wages, and Agricultural Labor Productivity*, I have several specific people to acknowledge. I am deeply grateful to PickTrace, Inc. for sharing data for this project, and particularly appreciative to Austin and Harrison Steed, as well as several anonymous growers, for answering countless questions about growing blueberries. I also thank Michael Anderson, Max Auffhammer, Tim Beatty, Kendon Bell, Marc Bellemare, Peter Berck, Fiona Burlig, Meredith Fowlie, Sol Hsiang, Larry Karp, Jeremy Magruder, Matthew Neidell, Jeff Perloff, Louis Preonas, Gordon Rausser, Howard Rosenberg, Elisabeth Sadoulet, Leo Simon, David Sunding, Becca Taylor, Sofia Berto Villas-Boas, Brian Wright, and David Zilberman for helpful comments.

In relation to my second chapter, *Reap What Your Friends Sow: Social Networks and Technology Adoption*, I am particularly grateful for my co-author, Fiona Burlig. Fiona is responsible for a large share of the heavy lifting that turned an interesting idea into a rigorous and robust research paper. Together, we wish to thank Maximilian Auffhammer, Kendon Bell, Marc Bellamare,Jay Coggins, Alain de Janvry, Benjamin Faber, Meredith Fowlie, James Gillan, Kelsey Jack, Rachid Laajaj, Jeremy Magruder, Edward Miguel, Katherine Morris, Louis Preonas, Gautam Rao, Subrata Kumar Ritadhi, Elisabeth Sadoulet, Megan Stevenson, Rebecca Taylor, Sofia Villas-Boas, and David Zilberman for helpful suggestions. We also benefited from comments from seminar participants at the University of California – Berkeley, the University of Minnesota – Twin Cities, the USDA Economic Research Service, and NEUDC 2014. We are grateful to Daniel Sparks, Shiyang Fu, and Jose Sandoval for providing research assistance. Luther Seminary, the Evangelical Lutheran Church in America, and the Concordia Historical Institute graciously shared data. Fiona Burlig acknowledges generous support by the National Science Foundation's Graduate Research Fellowship Program.

In relation to my third chapter, *Fueling Local Water Pollution: Ethanol Refineries, Land Use, and Nitrate Runoff*, my work has benefited greatly from comments by Peter Berck, Kendon Bell, Fiona Burlig, Louis Preonas, Keith Knapp, Nathan Hendricks, Solomon Hsiang, David Sunding, David Zilberman, Aprajit Mahajan, Jay Coggins, and participants at the 2015 AAEA & WAEA Joint Annual Meeting.

More broadly, I am deeply thankful for the advice and guidance of my dissertation committee. Peter Berck has been an exemplary academic advisor, leading me to fruitful ideas while preparing me for life as a scholar. David Sunding has given me the opportunities to become a better teacher, as well as to see the big picture in the work I do. Solomon Hsiang has provided an important outside perspective on my work, always pushing me to make my research as relevant as possible to important policy discussions.

Several other people deserve special recognition. Sofia Berto Villas-Boas has been one of the most unabashedly positive, energetic, and supportive mentors one could have. Maximilian Auffhammer has been extraordinarily generous with his time and attention. And I could not imagine navigating graduate school over the past years without the constant support of Carmen Karahalios, Diana Lazo, Janna Conway-Hamilton, and Gail Vawter. I owe Kendon Bell, Fiona Burlig, and Louis Preonas a great debt for listening to all my half-baked ideas, and helping me zero-in on interesting questions. Finally, I am blessed with a loving and supportive family without whom I would not be who I am today: thank you Kris, Jon, David, Vannie, and everyone else.

Chapter 1

Temperature, Wages, and Agricultural Labor Productivity

1.1 Introduction

A canonical prediction of economic theory is that high wages increase labor productivity. In settings where workers are salaried or paid by the hour, this is the concept of efficiency wages (Akerlof and Yellen, 1986). In settings where workers are paid in proportion to their output (e.g. under piece rate wages), the theoretical connection between wages and productivity is even clearer.[1] However, it has proven difficult to empirically estimate the responsiveness of labor productivity to piece rate wages, since much of these wages' variation is driven by endogenous characteristics of the production process. In this paper, I provide the first quasi-experimental estimate of the elasticity of labor productivity with respect to piece rate wages. Specifically, I analyze a high-frequency panel of worker-level production data from over 2,000 California blueberry pickers paid by piece rates. Surprisingly, I find that on average, labor productivity is very inelastic with respect to wages.

Piece rate wages are interesting to study because they offer such a direct, clear, and salient link between a worker's effort and reward. In general, optimal labor contracts can be quite complex, as they must effectively incentivize worker effort while simultaneously accounting for issues like risk aversion, asymmetric information, and moral hazard (Hart and Holmström, 1987). However, these complications are less of a concern in settings where a firm can cheaply monitor both worker productivity and product quality. In such cases, theory suggests piece rate wages will outperform other common incentive schemes (Brown, 1990; Prendergast, 1999).[2] Understanding how workers respond to changes in a piece rate wage is

[1] Piece rate wage schemes create "the clearest link between a worker's effort and the reward" (Billikopf, 2008). Specifically, a higher piece rate raises the marginal benefit of worker effort and therefore incentivizes workers to work harder.

[2] Empirical studies have confirmed that, under appropriate institutional circumstances, workers are indeed more productive under piece rate wages than under hourly wages (Lazear, 2000; Shearer, 2004; Shi, 2010; Billikopf and Norton, 1992).

important in sectors where these wages can vary over time, like in specialty agriculture, the auto repair industry, or the growing rideshare market (e.g. Uber and Lyft).[3]

Econometricians face a fundamental challenge when trying to estimate the causal effect of piece rate wages on labor productivity: these wages are inherently endogenous. As an example, consider blueberry picking. When ripe berries are scarce and spread out (at the beginning of the season), average worker productivity is low. When ripe berries are abundant and dense (at the peak of the season), it is easier for workers to pick berries quickly, and average productivity is markedly higher. Because farmers aim to keep their workers' average effective hourly pay relatively stable over time, they set piece rate wages higher when picking is more difficult, and lower when picking is easier.

In order to account for piece rates wages' endogeneity, I adopt a two-pronged identification strategy. First, exploiting the richness of my multidimensional panel data, I econometrically control for environmental factors like seasonality and temperature that directly affect the berry picking production function. Second, I use the market price for blueberries as an instrument for piece rate wages. This price is a valid instrument because it affects a farmer's willingness to raise piece rates (since it alters the value of workers' output), but is otherwise uncorrelated with picker productivity. Furthermore, the market price for California blueberries is set by global demand and global supply. As a result, individual farms are too small to directly affect the market price, and supply shocks at the farm level can be considered orthogonal to aggregate supply shocks.

I find that, on average, labor productivity is very inelastic with respect to piece rate wages, and I can reject even modest elasticities of up to 0.7. This finding contrasts with both canonical economic theory and previous structural estimates: relying on a calibrated structural model of worker effort, Paarsch and Shearer (1999) estimate a labor effort elasticity of 2.14 in the British Columbia tree-planting industry, and Haley (2003) estimates a labor effort elasticity of 1.51 in the U.S. midwest logging industry. Why, then, do blueberry pickers not seem to respond to changes in their wage? One explanation of my findings could be that blueberry pickers respond to average effective hourly wages rather than marginal piece rate wages, similar to how electricity consumers respond to average prices rather than marginal prices (Ito, 2014). This is unlikely, both because piece rate wages are highly salient in the context I study, and because my identification strategy specifically isolates marginal effects from average effects. Instead, I find suggestive evidence that blueberry pickers face some binding constraint on physical effort that is related to temperature.

Specifically, I find that at moderate to hot temperatures, I cannot reject that the piece rate wage level has no effect on labor productivity. However, at temperatures below 60 degrees Fahrenheit (15.6 degrees Celsius), a one cent per pound increase in the piece rate wage increases worker productivity by nearly 0.3 pounds per hour, implying a positive and statistically significant productivity elasticity of approximately 1.6. In other words, blueberry pickers respond to the piece rate wage level at cool temperatures, but seem not to respond to changes in their wage at higher temperatures.

[3]For recent analyses of the rideshare market, see Chen and Sheldon (2015) and Sheldon (2016).

Temperature also affects productivity *directly* in economically meaningful ways. Specifically, I find that blueberry pickers' productivity drops precipitously at very hot temperatures: workers are 12% less productive at temperatures above 100 degrees Fahrenheit (37.8 degrees Celsius) than they are at temperatures between 80 and 85 degrees Fahrenheit (26.7 and 29.4 degrees Celsius). However, I also find negative effects at cool temperatures. Workers are nearly 17% less productive at temperatures below 60 degrees Fahrenheit (15.6 degrees Celsius) than at temperatures in the low eighties. The most likely explanation of this finding is that berry pickers lose finger dexterity at cool temperatures and find it uncomfortable to maintain high levels of productivity. This hypothesis is supported by evidence from the ergonomics literature (Enander and Hygge, 1990), and highlights that temperature's effects on labor productivity depend on the particularities of the relevant production process.

To demonstrate the robustness of my findings, I address several threats to my identification strategy. First, I investigate berry pickers' labor supply on both the intensive (hours worked in a day) and extensive (probability of showing up to work) margins. I show that neither temperature nor wages have a statistically significant effect on these measures. Next, I address the fact that there exists a minimum hourly wage rule in the setting I study. This constraint binds for approximately 15.8% of my observations, raising concerns that workers falling below this threshold have an incentive to shirk or "slack off." I re-estimate my results using only those observations where workers earn more than the minimum wage and see no qualitative change in my findings. Finally, I confront the possibility of adverse selection in my sample by limiting my sample to only the observations from workers who work more than thirty days in a single season.

My results highlight the importance of environmental conditions in outdoor industries. Previous studies have shown, and I confirm, that temperature affects labor productivity directly.[4] However, I am the first to demonstrate that temperature also affects labor productivity indirectly by disrupting the economic relationship between wages and worker effort. As global temperatures rise, my findings suggest that firms in outdoor industries like agriculture and construction will have a reduced ability to effectively incentivize their employees' productivity. This can have large economic consequences. In the $76 billion U.S. specialty crop sector, for instance, harvest labor can account for more than half a farm's operating costs.[5] This is also a setting where piece rate wages are common: in California alone, over 100,000 specialty crop farm workers were paid by piece rates in 2012.[6]

My econometric estimates allow me to make several predictions about how rising temperatures will affect the agricultural labor sector. To do this, I develop a model of a firm

[4]See, for instance, Adhvaryu et al. (2016b), Sudarshan et al. (2015), and Seppänen et al. (2006).

[5]Specialty crops are defined as "fruits and vegetables, tree nuts, dried fruits and horticulture and nursery crops, including floriculture" (United States Department of Agriculture, 2014). In 2012, US specialty crop farms sold over $76 billion of crops (National Agricultural Statistics Service, 2015). Jimenez et al. (2009) find that harvest labor accounts for over 50% of operating costs for a typical south California blueberry farm.

[6]Specialty crop farms in California employed 414,564 workers in 2012 (National Agricultural Statistics Service, 2015), and roughly one in four agricultural workers in the US west are paid piece rate wages (Moretti and Perloff, 2002).

choosing an optimal piece rate wage under some exogenous environmental condition (e.g. temperature).[7] My model produces two interesting sets of comparative statics. First, I show that temperature's effect on the optimal piece rate wage depends on (1) how temperature affects labor productivity directly, and (2) how temperature affects labor productivity's responsiveness to the wage. Plugging my empirical estimates into this model, I find that an optimizing blueberry farm would pay its workers a higher piece rate wage on particularly cool days, *ceteris paribus*. Second, I show that temperature's effect on overall farm profits has the same sign as temperature's direct effect on labor productivity. In the case of California blueberry farms, where cool temperatures have meaningful negative effects on productivity, this suggests that the first-order effect of rising temperatures on profits is likely to be positive. However, in contexts where cool temperatures do not lower labor productivity, the opposite is likely to be true.

The remainder of this paper is organized as follows: in section 1.2, I develop a simple theoretical model of workers' optimal effort under a piece rate wage scheme. In section 1.3, I describe the institutional details of the two California blueberry farms I study in this paper. I then discuss my data and report summary statistics in section 1.4. Section 1.5 outlines my empirical strategy, and section 1.6 reports my results. I discuss my findings in section 1.7, giving particular attention to how rising temperatures are likely to affect the agricultural labor sector. Finally, in section 1.8, I conclude.

1.2 Theoretical Framework

In this section, I model a worker choosing to exert an optimal level of effort under a piece rate wage scheme.[8] Assuming a well-behaved production function, this framework predicts a positive elasticity of productivity with respect to the wage. I also allow worker effort to depend on exogenous characteristics of the production process, including an environmental condition (e.g. temperature). Doing so allows me to predict how changes in the environment will affect labor productivity. Later, in appendix A, I use these results to explore how rising temperatures may affect the agricultural labor sector.

1.2.1 A model of optimal effort under piece rate wages

Consider a setting where workers are employed to harvest some resource. Worker-level output is determined by a production function $f(e, \theta, T)$ that depends on the level of effort e expended by the worker, the abundance (or density) of the resource θ, and an environmental condition T. It is assumed that production is increasing in effort and resource abundance: $f_e > 0$, $f_\theta > 0$, with marginal production decreasing in e: $f_{ee} < 0$. Additionally, marginal

[7] For a more general review of optimal contracts in agriculture, with special attention to transaction costs and risk, see Allen and Lueck (2002).

[8] The practice of modeling labor effort, rather than simply working hours, dates back to at least Robbins (1930).

production is assumed to be increasing across effort and resource abundance: $f_{e\theta} > 0$. Workers are paid a piece rate wage r per unit of output they produce. Workers also bear a cost of providing effort, $c(e, T)$, that is increasing in effort expended: $c_e > 0$. The marginal cost of effort is assumed to be increasing as well: $c_{ee} > 0$.

Workers therefore face the following utility maximization problem, where income enters utility linearly:

$$\max_e \quad rf - c, \tag{1.1}$$

which leads to the first-order condition:

$$rf_e - c_e = 0, \tag{1.2}$$

and the subsequent second-order condition:

$$rf_{ee} - c_{ee} < 0 \tag{1.3}$$

where the inequality follows from the assumptions on f_{ee} and c_{ee}. Equation (1.2) implicitly defines an optimal level of effort to expend as a function of the piece rate wage, resource abundance, and the environmental characteristic: $e(r, \theta, T)$. I now want to sign the following partial derivatives: e_r, e_θ, and e_T. To do so, I first note that in this case, $de/dr = e_r$, $de/d\theta = d_\theta$, and $de/dT = d_T$, since the worker takes r, θ, and T to be exogenous.

I calculate de/dr, $de/d\theta$, and de/dT by totally differentiating the first-order condition in equation (1.2) with respect to r, θ, and T respectively, and rearranging the resulting expressions. I obtain the following:

$$\frac{de}{dr} = -\frac{f_e}{rf_{ee} - c_{ee}} > 0 \tag{1.4}$$

$$\frac{de}{d\theta} = -\frac{f_{e\theta}}{f_{ee}} > 0 \tag{1.5}$$

$$\frac{de}{dT} = \frac{c_{eT} - rf_{eT}}{rf_{ee} - c_{ee}} \Longrightarrow \frac{de}{dT} < 0 \iff c_{eT} > rf_{eT}. \tag{1.6}$$

It is now clear that $e_r > 0$, $e_\theta > 0$, and that the sign of e_T depends on the nature and level of T.[9]

[9] A brief note is warranted about the sign of e_T. Expression (1.6) implies that if the marginal effort-cost of T is large enough relative to the marginal effort-product of T – as one could reasonably expect for "bad" environmental conditions like pollution or very hot temperatures – then optimal effort is decreasing in the environmental condition T. If we take T to represent very hot temperatures, then this condition appears to contradict the canonical finding of climate-change-adaptation models, where an increase in temperature leads to higher use of other inputs, thus mitigating the negative direct effect of temperature on output (Antle and Capalbo, 2010). Effort is different from other inputs in this case, since its cost *also* depends on temperature. In particular, as long as the numerator in expression (1.6) is positive, that is, $c_{eT} > rf_{eT}$, then optimal effort will decrease with T. One might term this phenomenon "negative adaptation."

To determine a worker's optimized output, denoted by X, I simply plug the worker's optimal level of effort into the production function f:

$$X(r, \theta, T) \equiv f(e(r, \theta, T), \theta, T). \tag{1.7}$$

Time is not a choice variable in this model, so the function X can be interpreted as a worker's level of productivity given some piece rate wage r, resource abundance θ, and environmental condition T. I now directly derive expressions for X_r, X_θ, and X_{rr}:

$$X_r = f_e e_r > 0 \tag{1.8}$$

$$X_\theta = f_e e_\theta + f_\theta > 0 \tag{1.9}$$

$$X_{rr} = f_e e_{rr} + e_r^2 f_{ee}. \tag{1.10}$$

The signs for X_r and X_θ are immediate consequences of earlier assumptions on f and expression (1.5). The sign of X_{rr}, however, requires an additional condition. Rearranging equation (1.10) gives the following:

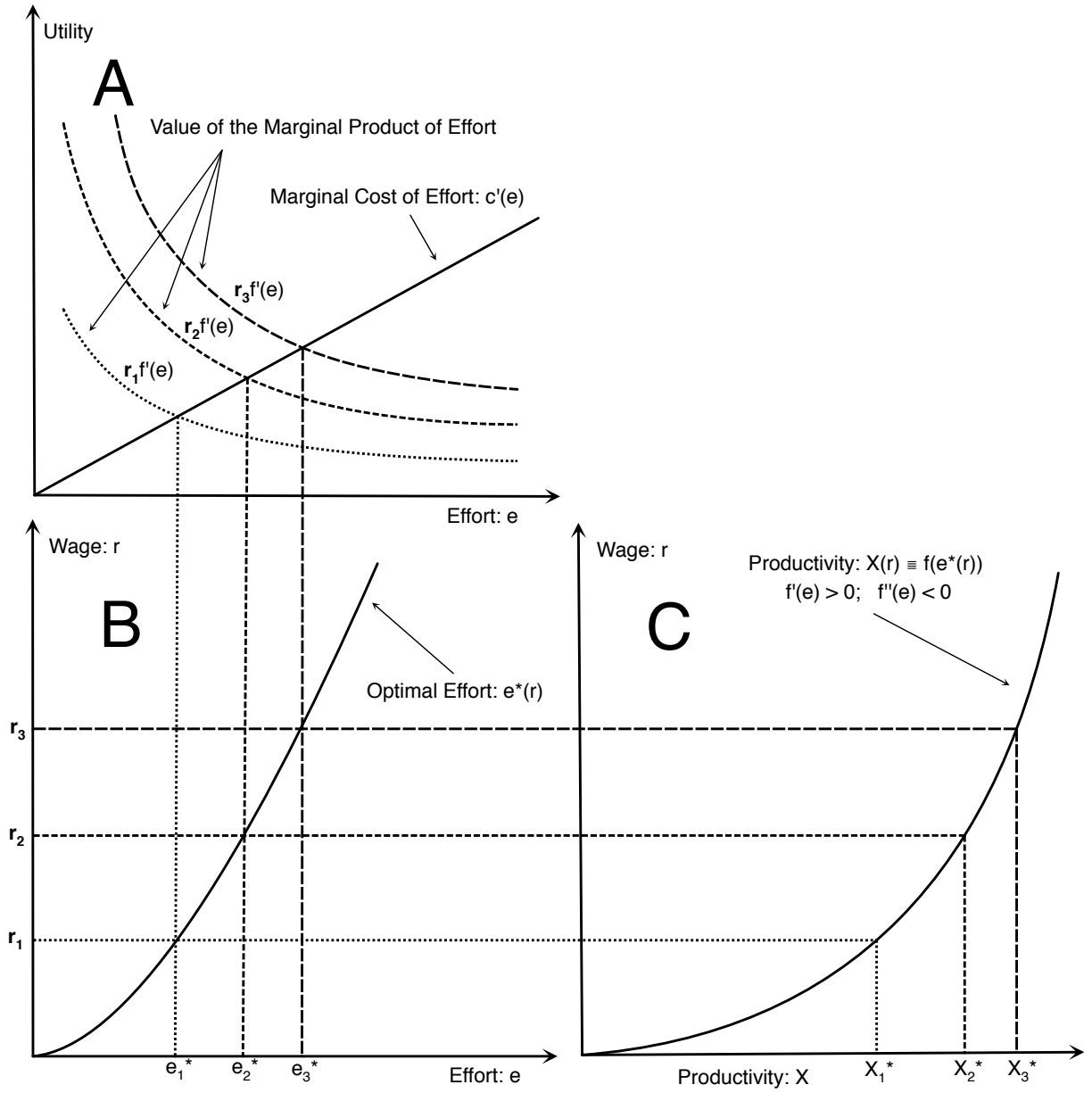
$$X_{rr} < 0 \iff e_{rr} < \frac{e_r^2 f_{ee}}{f_e} < 0. \tag{1.11}$$

That is, as long as marginal effort is decreasing severely enough with the piece rate wage r, laborers' marginal supply of output will also be decreasing in r. This condition is easy to swallow, since worker effort likely faces a finite psychological and/or physiological upper limit (Wyndham et al., 1965).

Figure 1.1 visualizes this model, holding θ and T fixed. Panel A summarizes the first-order condition in equation (1.2), where the value of marginal product of effort equals the marginal cost of effort, at three different piece rate wages (r_1, r_2, and r_3). Panel B translates the results of panel A into a relationship between piece rate wages and optimal effort. Finally, panel C shows how the production function f turns optimal effort e^* into productivity X. The figure is drawn such that condition (1.11) holds. The result, shown in panel C, is worker productivity X as a function of wage r.

Note that equations (1.8) and (1.10) imply that $X_r > 0$ and $X_{rr} < 0$, as long as condition (1.11) holds. That is, my model predicts that there is a positive elasticity of productivity with respect to the piece rate wage, and that this elasticity decreases as the wage increases. I focus on worker productivity X, rather than optimal effort e, because effort is fundamentally unobservable.

Figure 1.1: A Model of Optimal Worker Effort



This figure illustrates the model developed in section 1.2.1, holding resource abundance (θ) and environmental conditions (T) constant. Panel A summarizes a worker's utility maximization process where the optimal level of effort is determined by the intersection of the value of marginal product of effort and the marginal cost of effort. This is drawn at three different piece rate wages: r_1, r_2, and r_3. Panel B translates the information from panel A into an optimal effort curve that is a function of the piece rate wage: $e^*(r)$. Finally, panel C turns the optimal effort curve into a productivity curve by plugging effort into the production function f. The result is productivity as a function of piece rate wage: $X(r)$. As drawn, $X_r > 0$ and $X_{rr} < 0$, as discussed in section 1.2.1. Empirically, I cannot observe effort e, but I can and do estimate X_r.

1.2.2 Previous literature on piece rate wages

There has been relatively little theoretical work done on piece rate wage schemes in the past, partly because their structure is so straightforward, and partly because they are so much less common than salaries or hourly wage schemes. Nonetheless, previous research has highlighted several important aspects of piece rate wages that are relevant to this paper. Prendergast (1999) and Brown (1990) both provide good summaries of when and where piece rates are likely to be effective. Specifically, in cases where firms can cheaply monitor productivity and ensure quality control, piece rates should correctly align workers' incentives with those of their employer, maximizing labor productivity.[10]

Several papers have confirmed the prediction that, under the correct circumstances, piece rate wage schemes better incentivize labor productivity than do more traditional wage schemes. Lazear (2000), studying an auto glass company, finds that a switch from hourly to piece rate wages boosts output per worker by an average of 44%. Shi (2010), studying a tree-thinning company, estimates a more modest effect of 23%. Shearer (2004), studying tree-planters in British Columbia, also finds an effect near 20%. Bandiera et al. (2005) study agricultural workers in the United Kingdom and come to a similar conclusion, noting that piece rates based on individual production eliminate cross-worker externalities found in relative incentive schemes. In a non-causal study from California, Billikopf and Norton (1992) also provide evidence that piece rate wages boost vine-pruners' performance relative to hourly wages. Such increases in productivity under piece rates seem to come from increased worker effort, as Foster and Rosenzweig (1994) demonstrate by measuring workers' net calorie expenditures under different pay schemes.

None of the papers cited above, however, estimate how labor productivity responds to changes in a piece rate wage.[11] Among the most well-known papers that have attempted to do so are Paarsch and Shearer (1999) and Haley (2003). In both cases, the authors calibrate a structural model of worker effort (motivated by Grossman and Hart (1983)) in order to address piece rates' endogeneity. They find positive elasticities of effort with respect to wages, of 2.14 and 1.51 respectively, in line with theoretical predictions (e.g. equation (1.8)).[12]

Other papers have relied on natural experiments or natural field experiments to try and recover the effect of piece rate wage levels on productivity. For instance, Treble (2003) exploits a natural experiment from the 1890s in an English coal mine to derive a near-unit-elastic productivity response. In a more recent setting, Paarsch and Shearer (2009) implement a natural field experiment with tree-planters in British Columbia and estimate a productivity elasticity of 0.39. While the authors note that this estimate is "substantially smaller" than that of Paarsch and Shearer (1999) and Haley (2003), it is unclear wether

[10] Eswaran and Kotwal (1885) address similar questions in a specifically agricultural context.

[11] For yet another example, Dillon et al. (2014) derive a model of worker effort under piece rate wages that incorporates workers' health – as well as workers' perceptions of their health – to assess the labor productivity effects of a malaria testing policy in Nigeria. However, piece rates in their context are static.

[12] Paarsch and Shearer (1999) argue that knowing this elasticity is important for firms, since the optimal (profit-maximizing) piece rate wage is increasing in this elasticity, as shown by (Stiglitz, 1975).

they think this result invalidates the earlier estimates.[13] Finally, Guiteras and Jack (2017) conduct an experiment in rural Malawi to explore how variation in piece rate wages affects both quantity and quality of worker output. The authors find a positive but very inelastic effect of piece rate wages on workers' output.[14]

Despite the theoretical simplicity of a piece rate wage scheme, it is not immune to employees' behavioral responses. Even though a firm may be able to set a different piece rate every day, doing so may foment unrest among employees if the changes are seen as arbitrary (Billikopf, 2008). In other situations, high piece rates may operate as efficiency wages – à la Yellen (1984), Shapiro and Stiglitz (1984), and Newbery and Stiglitz (1987) – especially if a firm is trying to retain high-quality workers (Moretti and Perloff, 2002). An additional consideration is that variable piece rate wages may lead to a less reliable supply of labor on the intensive margin. In other words, piece rate employees may work fewer or more hours depending on the day's wage. Such behavior would be consistent with a reference-dependent preference model like that of Kőszegi and Rabin (2006) where workers have some internal reference point for how much money they intend to earn in a particular day.[15]

Finally, piece rate wages are much more common in seasonal specialty agriculture than in many other industries or settings. Tasks such as picking, pruning, or planting can be easily measured and tracked, making piece rates feasible. In these cases, productive workers can earn considerably higher incomes under a piece rate scheme than under an hourly wage scheme: Moretti and Perloff (2002) find that US agricultural workers paid by piece rate earn 26% more than their hourly counterparts. This number is slightly misleading, and certainly not causal, considering that workers select into particular work in part based on the compensation scheme. Rubin and Perloff (1993) note that piece rate workers tend to be disproportionately young or old: "[a]pparently, prime-age workers find that higher earnings in piece-rate jobs do not compensate for the difficulty of more intensive effort, more variable incomes, and possible greater injury risk or shortened farm-work career" (p. 1042). However,

[13]The authors write both: "Given [our] identification results, this suggests that the values of $\bar{u}$ used by Paarsch and Shearer as well as Haley to identify γ were too low" (Paarsch and Shearer, 2009, p. 487), and "[O]ur results... [are] consistent with previous results obtained by Paarsch and Shearer (1999) as well as Haley (2003)" (Paarsch and Shearer, 2009, p. 493).

[14]Specifically, "Increasing the piece rate by 10 MWK increases the number of units sorted per day by between 0.24 and 0.31 units, relative to a mean of 7.4 units. Going from the lowest piece rate (5 MWK) to the highest piece rate (25 MWK) increases output by between one-half to one unit per day" (Guiteras and Jack, 2017, p. 19).

[15]Chang and Gross (2014) find evidence this sort of behavior in their study of pear packers. In particular, they observe that workers provide different amounts of effort when being paid overtime wages, and that this effect varies with whether overtime pay is expected or unexpected. They also find different effects for differently-skilled workers. Chang and Gross' research builds upon several studies that analyze workers who are free to set their own hours. The most well-known papers in this literature focus on New York City taxicab drivers (Camerer et al., 1997; Farber, 2005, 2008; Crawford and Meng, 2011), but other papers also explore stadium vendors (Oettinger, 1999), bicycle messengers (Fehr and Götte, 2007), and fishermen (Giné et al., 2010). Specialty agriculture may differ from these contexts in important ways, however. For instance, Billikopf (1995) finds that few agricultural workers in California reduce their work hours when paid according to piece rate wages.

these selection issues are irrelevant if the goal is to understand how piece rates affect the productivity of workers who select into such work in the first place.

1.3 Context: California Blueberries

California is the fifth largest blueberry producer in the United States after Washington, Oregon, Georgia, and Michigan; the state grew over 31,000 tons of the fruit in 2015 alone (National Agricultural Statistics Service, 2016). Blueberries' popularity among California's specialty crop farmers is relatively new, and the California Blueberry Commission (CBC) was not established under the state's Food and Agricultural Code until 2010. I study two blueberry farms: an organic farm in San Diego County, and a conventional farm near Bakersfield. In order to protect the farms' identities, I cannot share their exact locations. However, figure 2.4 maps the approximate location of each farm within the state.

Figure 1.2: Farm Locations

This figure maps the approximate locations of the farms studied. In order to protect the farms' identities, exact locations cannot be shared. The San Diego farm grows organic blueberries while the Bakersfield farm grows conventional blueberries. Source: author's spatial approximation.

Harvesting fresh blueberries is a labor intensive process. Berries grow in small bunches and ripen at differing times. This means that a single blueberry bush can be harvested

multiple times each season. However, since each berry-bunch contains both ripe and unripe berries, pickers must harvest fruit carefully by hand. Mechanized blueberry harvesters exist, but they are imprecise and are used primarily for harvesting berries destined for the processing (secondary) market.[16]

Berry-pickers collect fruit in small buckets fastened on the front of their bodies. Once the buckets are full, the workers carry their harvest to a weigh-station at the end of a field row. Workers pour their berries into standardized bins which are then weighed, packed into trucks, and driven to a refrigerated packing plant. Because blueberries are delicate and perishable, they must be refrigerated quickly after being picked. When workers bring their berries to be weighed, a foreman closely watches the process to ensure quality control. If a picker's fruit is intermingled with too many twigs, leaves, or unripe berries, the foreman will warn the picker that their quality must improve to keep their job.

The farms I study both utilize an automated system to track workers' productivity and calculate payroll.[17] Each picker is given a unique barcode that they wear as a badge, and each fruit tray is assigned its own barcode as well. When a picker brings their fruit to be weighed, the weigher scans both the picker's barcode and the tray's barcode to record the tray weight. The picker then receives a receipt of their weigh-in. The farmer likes the barcode system because it is quick, automatic, reliable, provides real-time data, and replaces a cumbersome paper-and-pencil system. Pickers like the barcode system because they are able to witness the fruit-weighing and are thus confident that the farmer is paying them honestly for the fruit they pick.

At the beginning of each work day, around 6:00 or 6:30 a.m., the farmer sets the day's piece rate wage and posts the wage in a public spot for all workers to see.[18] Workers are paid the piece rate for each pound of berries they harvest, and the rate does not change throughout the day. The piece rate does, however, change over the course of the season (mid-April to mid-June each year). As fruit becomes more abundant on the bushes through May and June, picker productivity rises. Farmers therefore generally lower the piece rate wage throughout the season as more and more berries ripen. Anecdotally, farmers say they lower their piece rates "when there's a lot of fruit in the field" with the goal of maintaining a relatively stable effective hourly wage for the average berry picker.[19]

[16]Gallardo and Zilberman (2016) conclude that in order for the current incarnation of mechanical blueberry harvesters to be profitable for fresh market producers, (1) the price gap between fresh market blueberries and processing market blueberries would have to shrink considerably, (2) labor wages would have to rise more than 60%, or (3) yield losses from mechanical harvesting would have to fall by over 60%. None of these changes are likely in the near future.

[17]PickTrace, Inc. (`http://www.picktrace.com/`)

[18]Blueberries cannot be picked when it is raining, or if it has recently rained, since moisture on the berries disrupts the packaging process. Workers will not even bother to show up at the farm if it is raining in the morning.

[19]As long as a farm's average effective hourly wage is somewhat competitive in the agricultural day-labor market, the farm has monopsony power to set its own particular piece rate wage. That is, the piece rate wage is not set directly by the labor market. Fisher and Knutson (2013) highlight the fact that US agricultural labor markets are fundamentally more localized and heterogeneous than many may think.

If any one worker picks a small enough quantity of fruit that their effective hourly wage for the day falls below the legal minimum wage, the farmer pays them according to the hourly minimum wage. In these cases, the farmer often then gives the picker in question additional training and a warning that they may be fired if they do not quickly improve. Anecdotally, the hourly minimum wage is most likely to bind during a new employee's first few days on the job as they develop their skills as a fruit picker. If a worker consistently falls below the minimum wage cutoff, they frequently quit on their own accord or are effectively fired and asked not to return the next day.

Because blueberries are delicate and highly perishable, they are not bought and sold in a central commodity market. Instead, individual producers set short-term contracts with different marketers or buyers to provide a certain quantity of berries in particular packaging at a particular time. These contracts are set on a near-daily basis, and prices can change quickly throughout the season. While there is certainly some quality differentiation within the blueberry market, buyers and marketers view different producers as close substitutes. This means that individual producers have relatively little, if any, market power. I thus take California blueberry contract prices as an accurate reflection of a competitive market price for blueberries in the state.

Blueberry prices in California are highly seasonal: prices are quite high at the beginning of the season in April, and much lower near the end of the season in June. This seasonality in price is largely explained by (1) variation in aggregate production throughout California, and (2) variation in the availability of blueberries from other global producers. In the early spring, the United States imports fresh blueberries at high prices from Mexico or other countries since domestic production is agronomically infeasible. By mid-to-late-June, farms in northern states such as Washington, Oregon, and Michigan begin to produce berries in large quantities, driving down the market price. California blueberry farmers therefore face a relatively short season when it is profitable to harvest and sell their fruit. While blueberry bushes continue to yield berries through June and into July, labor costs are too high relative to market prices at that time for California farmers to justify continued production. To summarize, the California blueberry season begins agronomically, but ends economically.

Organic blueberries regularly command a price premium of around two dollars per pound. While the harvesting process is identical for conventional and organic berries, organic bushes produce fewer berries per bunch. Thus, pickers of organic berries spend more time finding

Whether or not a farm faces a labor shortage depends on local labor market conditions and seasonality rather than on aggregate state or national statistics. Indeed, Fan et al. (2015) note that fewer agricultural laborers are migrants now than at any time in the recent past, meaning local labor conditions can vary significantly across space. In Kern and San Diego counties, where the farms I study are located, the blueberry season (mid-April through mid-June) competes with relatively few other crop harvest seasons (Kern County Department of Agriculture and Measurement Standards, 2016; San Diego Farm Bureau, 2016). Additionally, harvesting conditions on blueberry farms (where pickers spend the day standing upright) are less onerous than on strawberry farms, meaning a blueberry farmer can attract and maintain workers for lower wages than competing strawberry farms (Guthman, 2017). Finally, if farmers' concerns primarily relate to worker retention, Gabbard and Perloff (1997) suggest there is a higher return to extra money spent on benefits or improved working conditions than relative wages.

and harvesting berries than do their conventional counterparts. Additionally, fruit quality is more variable in organic blueberries. This leads to a smaller proportion of berries ultimately reaching market.

1.4 Data Sources and Description

I utilize data from three distinct sources: employee-level production and payroll figures, high-frequency temperature readings, and state-level market prices. I observe over 2,000 fruit pickers on 170 days over three growing seasons at two farms for a total of over 300,000 unique fruit weigh-ins.

1.4.1 Employee-level production figures

As described in the previous section, the farms I study use a digital fruit weigh-in system to track worker productivity and generate payroll data. I utilize data from these weigh-ins to conduct my analyses. In particular, I observe the weigh-in time, the berry picker's unique employee identifier, the field where the berries were picked, and the weight of the picker's harvest. I divide the harvest's weight by the time elapsed since the picker's previous weigh-in to obtain a weight-per-hour measure of worker productivity. For the first weigh-in of the day, I use time elapsed since morning check-in to calculate this measure.

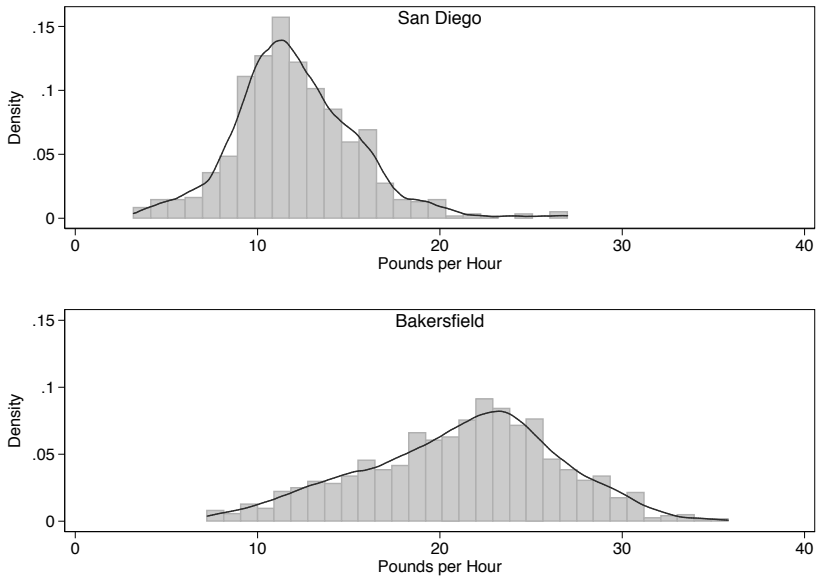
As reported in table 1.1, average productivity pooled across both farms is just over nineteen pounds picked per hour. This number, however, masks significant heterogeneity across farm, day, and worker. At the San Diego farm, which grows organic berries, average productivity is slightly under fourteen pounds per hour, while at the Bakersfield farm, which grows conventional berries, average productivity is over twenty-two pounds per hour. Figure 1.3 plots the distribution of workers' average productivities, while figure 1.4 plots the distribution of each day's average productivity, in both cases separated by farm. These two figures highlight substantial variation in picker skill, as well as in daily productivity.

In southern California and the central valley, where the farms I study are located, temperatures peak in the mid-to-late afternoon. To avoid the hottest part of the day, most pickers begin work as early as 6:00 a.m. and end around 3:00 p.m. This pattern is reflected in figure 1.5: most fruit picking ends by mid-afternoon. The average picker works around eight hours each day, as shown in figure 1.6. Under California law in my sample period (2014–2016), agricultural workers do not earn overtime pay until after working ten hours in a single day. In my data, only the San Diego farm ever lets pickers work more than ten hours in any given day.

Farms employ pickers on a day-to-day basis, either directly or through a labor contractor.[20] Some pickers only work for a day or two, but others work continuously for several

[20]The San Diego farm (smaller, organic) hires pickers directly, while the Bakersfield farm (larger, conventional) uses a labor contractor. Previous research has suggested a farmer will use a labor contractor if they are particularly concerned about having workers when needed (Isé et al., 1996). The same authors also find

Figure 1.3: Average Productivity by Picker

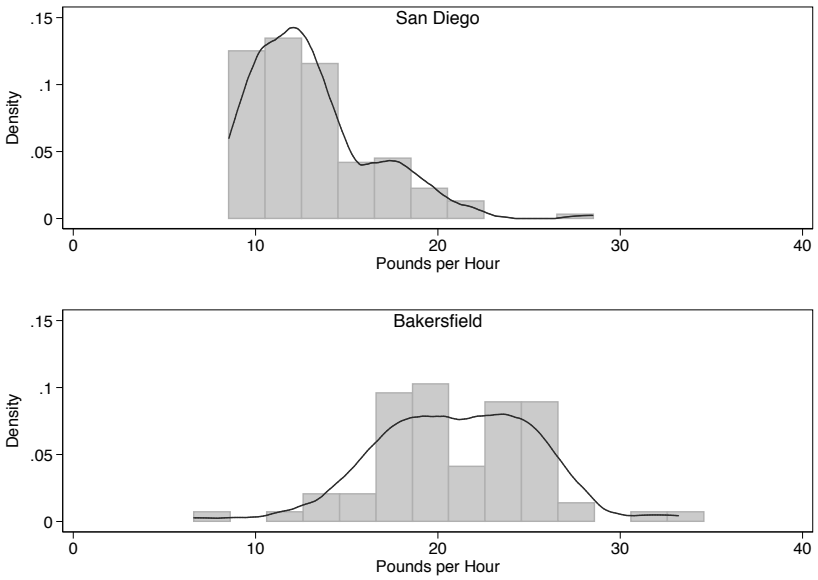


This figure plots the distribution of each picker's average productivity measured by pounds of blueberries picked per hour at two farms over three growing seasons. The San Diego farm (top panel) grows organic blueberries while the Bakersfield farm (bottom panel) grows conventional blueberries. The kernel density estimates use an Epanechnikov kernel. Source: proprietary payroll data.

weeks or months as shown in figure 1.7. A handful of pickers return to their respective farm each year. Indeed, several employees in my data work for a farm in two or all three of the years I study. Unfortunately, I do not observe each worker's initial date of hire, so I am unable to confidently measure lifetime worker tenure on either farm.

suggestive evidence that larger farms are more likely to use a labor contractor.

Figure 1.4: Average Productivity by Day



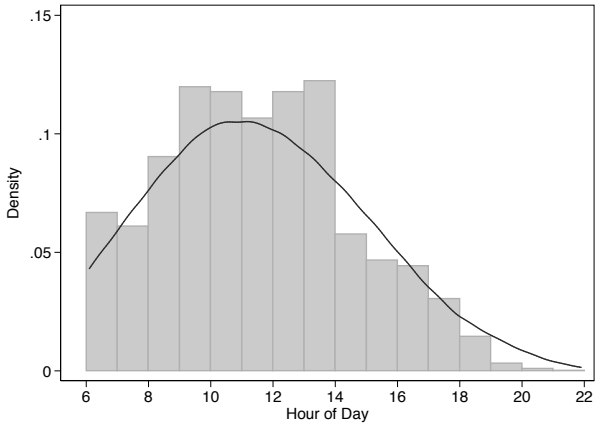
This figure plots the distribution of all pickers' productivity – measured by pounds of blueberries picked per hour and averaged over each day – at two farms over three growing seasons. The San Diego farm (top panel) grows organic blueberries while the Bakersfield farm (bottom panel) grows conventional blueberries. The kernel density estimates use an Epanechnikov kernel. Source: proprietary production data.

Table 1.1: Summary Statistics

	Pooled Sample			San Diego Farm			Bakersfield Farm		
Variable	Observations	Mean	SD	Observations	Mean	SD	Observations	Mean	SD
Panel A: variables that vary by weigh-in									
Worker productivity (lb/hr)	305,980	19.09	9.86	122,549	13.87	8.76	183,431	22.58	8.97
Temperature (°F)	305,980	70.63	9.31	122,549	72.34	9.27	183,431	69.48	9.17
Effective hourly wage ($/hr)	305,980	14.35	7.20	122,549	13.90	9.09	183,431	14.64	5.58
Panel B: variables that vary by worker-day									
Hours worked	47,939	7.71	1.90	23,026	7.71	2.50	24,913	7.72	1.06
Worker tenure by season (days)	47,939	17.75	13.04	23,026	22.07	14.68	24,913	13.75	9.74
Hourly minimum wage binds (proportion)	47,939	0.21	0.41	23,026	0.32	0.47	24,913	0.11	0.31
Panel C: variables that vary by day									
Piece rate wage ($/lb)	229	0.97	0.26	156	1.10	0.20	73	0.69	0.13
Market price ($/lb)	229	5.37	2.35	156	6.14	2.30	73	3.71	1.44
Total production (tons)	229	12.61	12.81	156	5.86	2.96	73	27.03	13.82
Number of workers	229	209.34	120.23	156	147.60	41.00	73	341.27	127.46
Female workers (proportion)	229	0.65	0.17	156	0.72	0.15	73	0.49	0.05
Panel D: number of days and employees									
Unique days in sample	170			156			73		
Unique employees in sample	2,022			655			1,367		

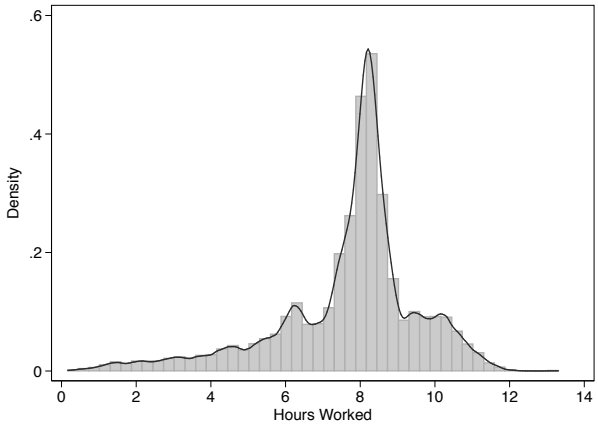
Note: SD: standard deviation. "Hourly minimum wage binds" is an indicator variable equal to one for a worker-day if the worker in question picked enough fruit per unit of time to earn more money under the piece rate wage than under the hourly minimum wage. "Worker tenure by season" is a variable that counts the number of days (inclusive) that a single employee has worked at a given farm in a given year. On a worker's first day of work in a season, this variable equals one. On their second, it equals two. And so on. The San Diego farm grows organic blueberries while the Bakersfield farm grows conventional blueberries. This explains the difference in mean market price faced by the two farms.

Figure 1.5: Time of Production



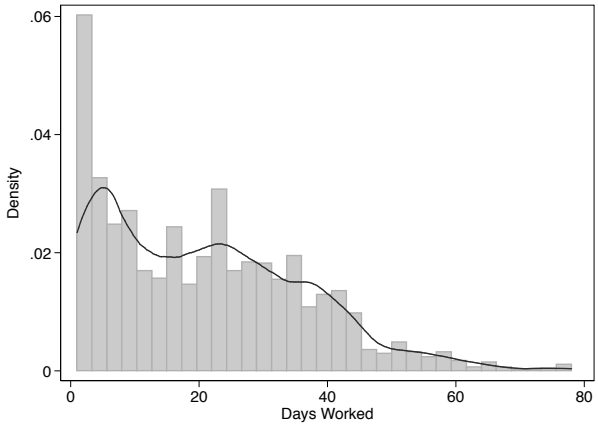
This figure plots the time distribution of blueberry pickers' "picking periods" at two California blueberry farms over three growing seasons. An observation in this distribution is the midpoint in time between when a picker begins filling a single tray of berries and when they weigh that tray. A single picker will have several "picking periods" within a single day. The kernel density estimate uses an Epanechnikov kernel with a bandwidth of two hours. Note that most weigh-ins occur before 3:00 p.m., as farmers and pickers avoid the hottest part of the day. Source: proprietary production data.

Figure 1.6: Hours Worked



This figure plots the distribution of daily hours worked by blueberry pickers on two farms over three growing seasons. An observation is a picker-day. The kernel density estimate uses an Epanechnikov kernel. Under California law at the time of this study, agricultural laborers do not earn overtime pay until after working ten hours in a day. All observations in this figure with more than ten hours worked come from a single farm (San Diego). Source: proprietary payroll data.

Figure 1.7: Days Worked



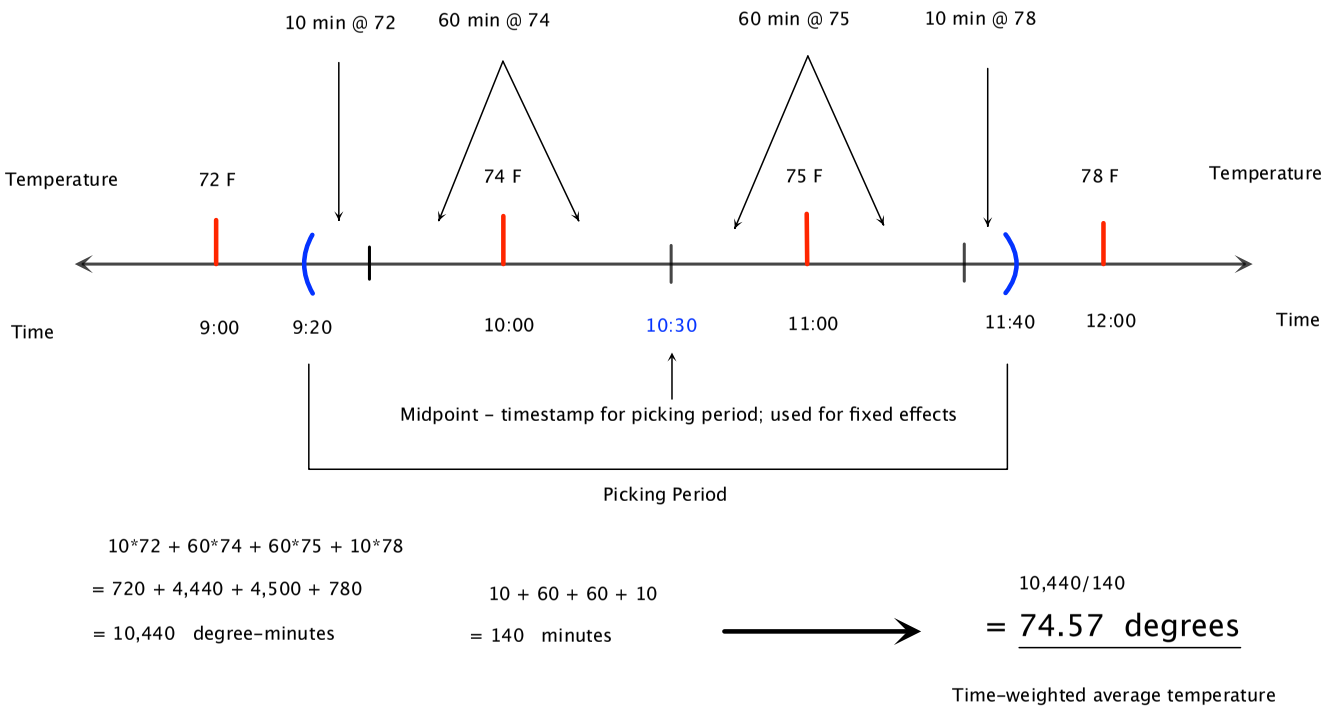
This figure plots the distribution of days worked in a single season by blueberry pickers on two farms over three growing seasons. An observation is a picker-season. The kernel density estimate uses an Epanechnikov kernel. Source: proprietary payroll data.

1.4.2 High-frequency temperature readings

I utilize high-frequency temperature data sourced from the MesoWest database maintained by the University of Utah.[21] In particular, for each year, I find the temperature monitor closest to each farm with hourly or finer temperature readings. In order to protect the identity of each farm, I cannot share the precise locations of these monitors, but they are all between 1.2 and 14.8 miles away from their respective farm. Temperature readings are available at least hourly, with some available at fifteen-minute intervals. I match temperature observations to each "picking period" – the span between two of a worker's sequential weigh-ins – using a time-weighted average of observed temperature. Figure 1.8 describes in detail how I calculate this time-weighted average. By using individual picking periods as my unit of observation, and matching these periods to time-weighted temperature measurements, I am able to exploit variation in temperature throughout each work day that accurately captures the heat exposure faced by outdoor laborers at different points of their shift. Figure 1.9 displays the distribution of time-weighted average temperatures within my data. There are few observations with extremely high temperatures, largely due to the fact that berry pickers usually end work in the mid-afternoon, before the hottest part of the day.

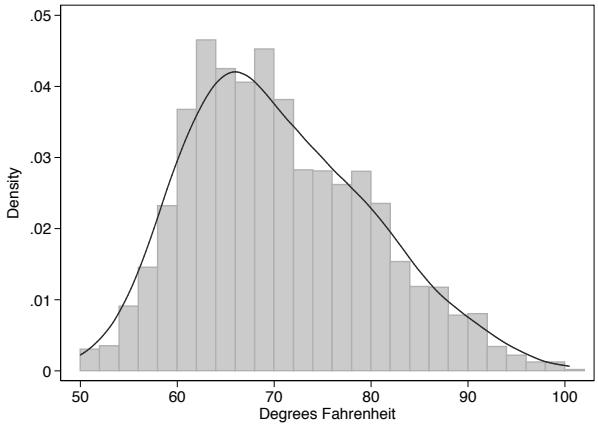
[21] `http://mesowest.utah.edu/`

Figure 1.8: Time-Weighted Average Temperatures



This figure describes the method I use to match temperature observations to a single "picking period" using hypothetical data. In particular, I calculate a time-weighted average of temperature observations during or near the time a picker is actively picking. Source: author's illustration.

Figure 1.9: Temperature at Time of Production



This figure plots the distribution of time-weighted average temperature in degrees Fahrenheit for each "picking period" at two California blueberry farms over three growing seasons. The kernel density estimate uses an Epanechnikov kernel with a bandwidth of three degrees. Sources: proprietary production data, MesoWest (temperature).

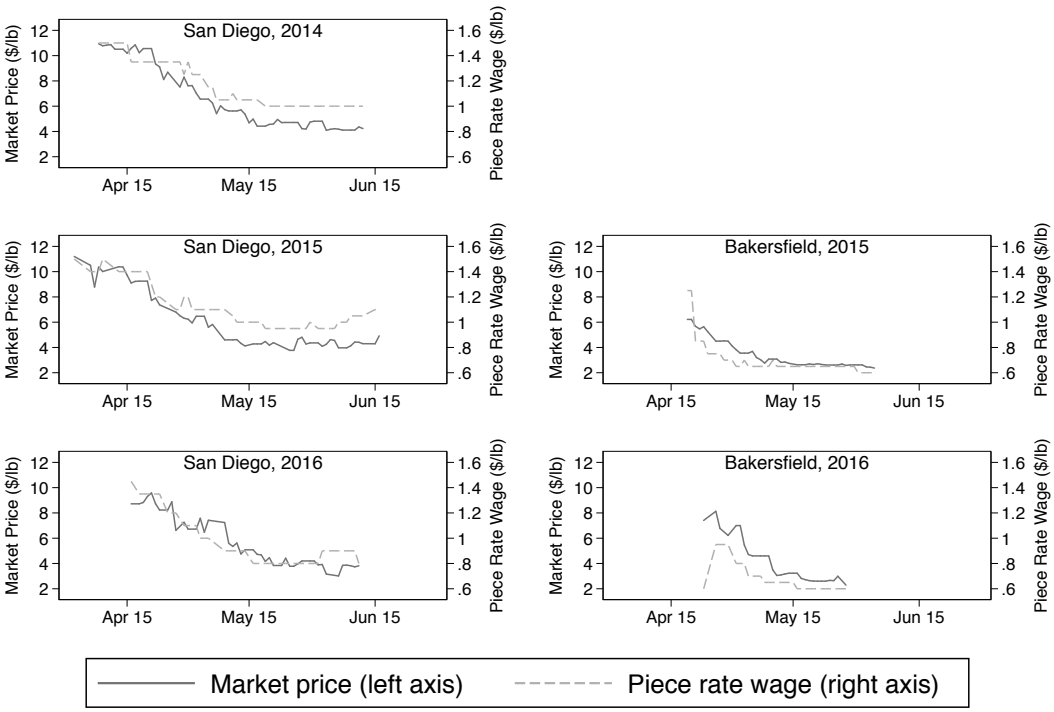
1.4.3 State-level market prices

While I know each farm's daily piece rate wage from the its payroll data, I obtain information on market prices for California blueberries from the Blueberry Marketing Research Information Center (BMRIC) of the California Blueberry Commission (CBC). As an official agricultural commission, the CBC legally requires all blueberry producers in the state to report daily production and sales figures. The CBC then publishes daily summary statistics of these data through BMRIC. Individual blueberry producers are able to access a daily BMRIC report online that summarizes the high, low, and weighted average prices received by blueberry producers throughout the state on the previous day. Separate statistics are provided for conventional and organic blueberries. In order to capture the information a farmer could have accessed on any particular day, I use each day's most recent *previous* BMRIC report as the relevant measure of market prices. Because BMRIC publishes a daily report each weekday except for holidays, the relevant market price data for harvest data collected on a Thursday is from the Wednesday prior. Similarly, the relevant market price data for harvest data collected on a Monday is from the Friday prior. Based on personal conversations, the blueberry farmers I study track these BMRIC reports quite closely throughout the season.

From April to June each year, both market prices and piece rate wages fall as the California blueberry season progresses. Figure 1.10 documents this relationship across the three years and two farms in my dataset. Recall that the San Diego farm grows organic blueberries while the Bakersfield farm grows conventional berries. This distinction accounts for why the two farms face differing market prices in the same year.

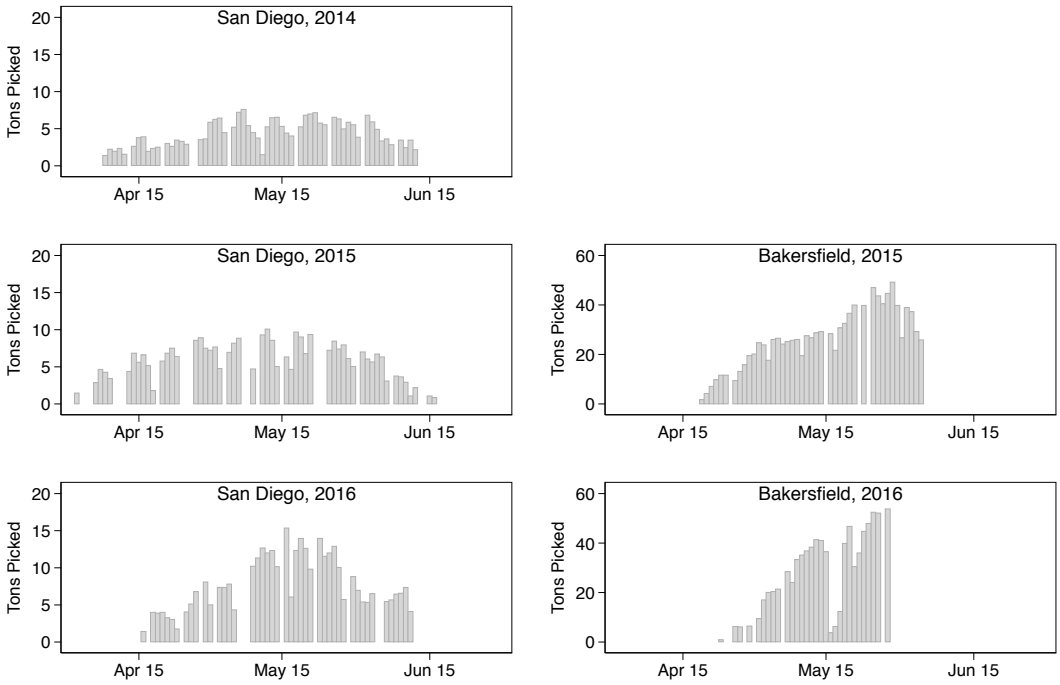
Market prices and piece rate wages are highly correlated over time, due in large part to seasonality in blueberry production. Figure 1.11 plots each farm's daily total production over time for each season. At times of high production, blueberry bushes are likely to be full of easily-pickable ripe berries. This abundance of fruit leads farmers to cut the piece rate as described in the previous section. In order to disentangle the various factors that affect farms' piece rate wages in my empirical exercises, I control both for seasonality in production as well as the field where berries are harvested.

Figure 1.10: Market Prices and Piece Rate Wages



This figure plots the market price of California blueberries and the piece rate wage paid to berry pickers by two farms over the span of three growing seasons. The San Diego farm (left column) grows organic blueberries while the Bakersfield farm (right column) grows conventional blueberries. This explains the difference in the market price faced by the two farms. Sources: Blueberry Marketing Research Information Center, California Blueberry Commission (market prices); proprietary payroll data (piece rate wages).

Figure 1.11: Daily Blueberry Production



This figure plots production at two California blueberry farms over three growing seasons. Production is measured in tons, and observations are daily. Transitory gaps in production are generally weekends (especially Sundays) or rainy days. The San Diego farm (left column) grows organic blueberries while the Bakersfield farm (right column) grows conventional blueberries. Production ends abruptly as competition from northern producers (Washington, Oregon, and Michigan) pushes the market price for blueberries below profitable levels for California producers. Source: proprietary production data.

1.4.4 Additional summary statistics

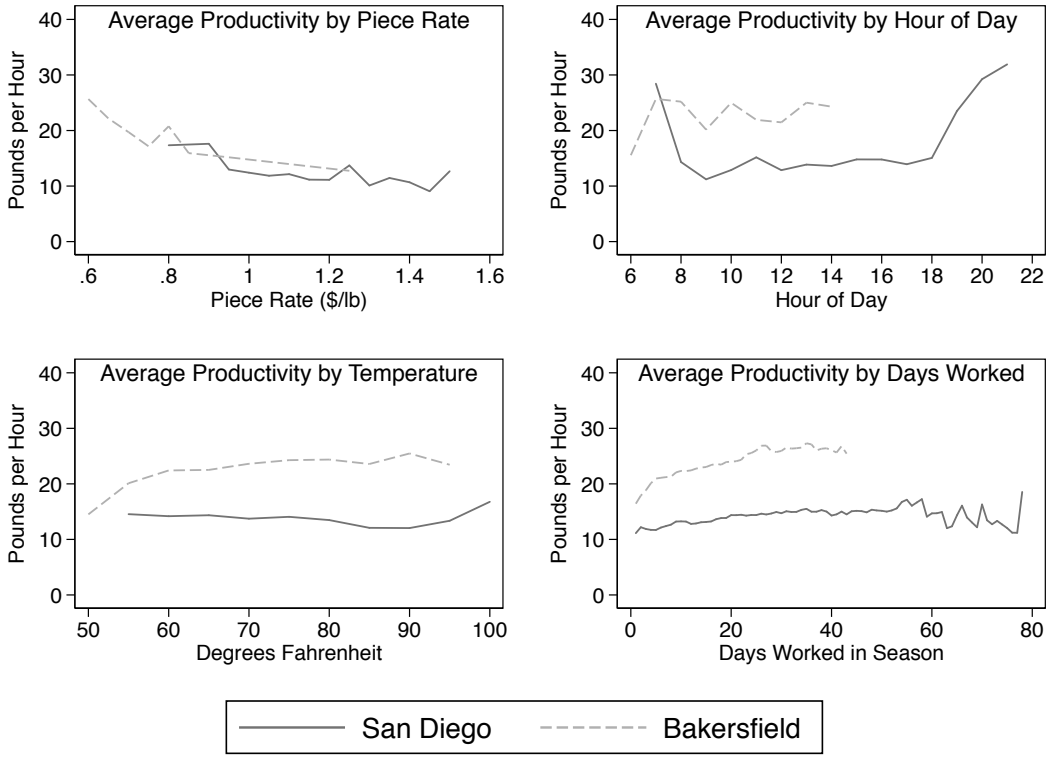
In my subsequent econometric analyses, I estimate the causal effects of piece rate wages and temperature on picker productivity. Figure 1.12, in contrast, plots the naïve relationship between average picker productivity and piece rate wages, temperature, and two other observable characteristics: time of observation and worker tenure by season. First, note that productivity and piece rate are negatively correlated, since farmers lower the piece rate when fruit is plentiful in the fields.[22] Second, note that there are no sharp decreases to average productivity at particularly high temperatures, as one may hypothesize. Finally, note that there is a clear increasing and concave relationship between worker tenure within a season and productivity. In other words, there is learning-by-doing in berry picking, and this learning has decreasing marginal returns over time.

While most employees out-earn the hourly minimum wage under the piece rate system, some fall below this threshold and are paid according to the minimum wage for the day. As Graff Zivin and Neidell (2012) note, if there is not a credible threat that these workers could be fired for their low output, they may shirk and provide less effort than they otherwise would. Figure 1.13 plots the distribution of normalized daily productivity that identifies those picker-days where shirking could be a problem. Observations to the left of one are picker-days where the picker's effective hourly wage is below the minimum wage, and observations to the right of one are picker-days where the picker out-earns minimum wage under the piece rate scheme. A picker with a normalized productivity measure of two is earning twice the minimum wage. Productivity in this figure is normalized because both piece rate wages and the hourly minimum wage vary over the sample period.

Shirking, if it occurs, could bias my results. In particular, if high temperatures or low wages lead to more pickers earning the minimum wage, and these pickers subsequently shirk, my econometric estimates will be biased upward. I address this concern in section 1.6 by re-estimating my primary results using only those picker-days where employees out-earn the minimum wage. My findings do not change when I eliminate these observations, suggesting that the threat to a picker of being fired if they consistently slack off is a sufficient incentive to keep them from shirking.

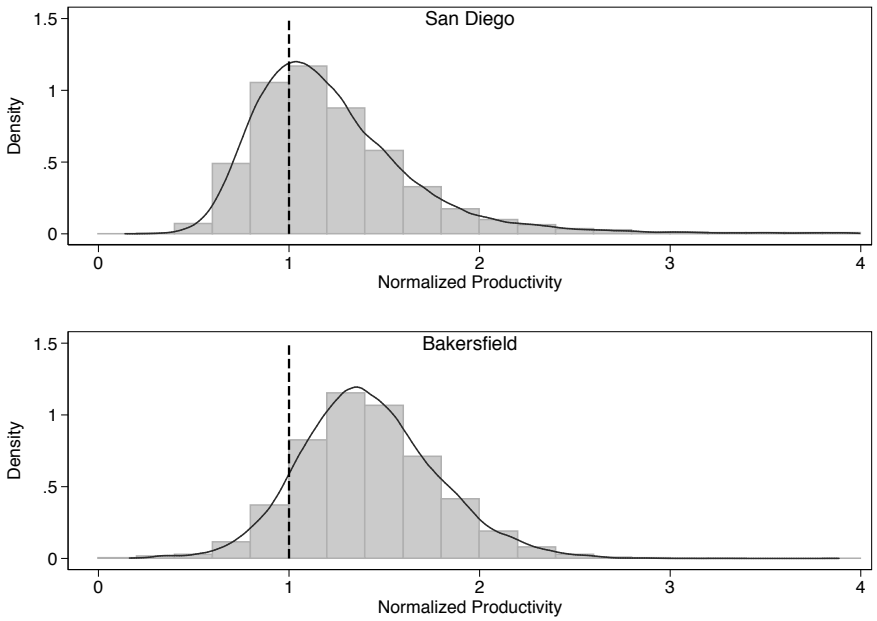
[22] Paarsch and Shearer (1999) explain clearly why a simple covariance between piece rate wages and worker productivity would suggest a negative elasticity of effort with respect to piece rates: a firm sets piece rates endogenously in response to the difficulty of the work. In the terms of the model from section 1.2.1, $dr/d\theta < 0$.

Figure 1.12: Average Productivity Across Observable Variables



This figure plots average worker productivity – pounds of blueberries picked per hour – for two California blueberry farms over three growing seasons and across four observable variables: the piece rate wage, hour of day, temperature, and worker tenure. The San Diego farm (solid lines) grows organic blueberries while the Bakersfield farm (dashed lines) grows conventional blueberries. For the temperature plot, observations are grouped into five-degree bins beginning at 50 degrees Fahrenheit. None of these plots is adjusted for seasonality. Sources: proprietary payroll and production data, MesoWest (temperature).

Figure 1.13: How Frequently Does the Minimum Wage Bind?



This figure plots the distribution of pickers' daily productivities over three growing seasons, normalized by the productivity necessary to exceed the hourly minimum wage rate. Normalization is necessary because both piece rate wages and the hourly minimum wage vary over the sample period. Pickers with a normalized productivity measure greater than one will earn more per hour than the minimum wage, while pickers with a normalized productivity measure less than one will be paid the hourly minimum wage. Pickers who consistently fall below this threshold receive additional training and are in some cases fired. The San Diego farm (top panel) grows organic blueberries while the Bakersfield farm (bottom panel) grows conventional blueberries. The kernel density estimates use an Epanechnikov kernel. Source: proprietary payroll data.

1.5 Empirical Strategy

The model presented in section 1.2.1 motivates my empirical strategy. In particular, my goal is to estimate the relationship between piece rate wages and labor productivity (X_r). The primary challenges to this undertaking are twofold. First, many observable and unobservable factors contribute to worker productivity which – if unaccounted for – could lead to omitted variable bias in my estimates of temperature and wage effects. Second, piece rate wages are endogenous to labor productivity.

To address factors other than the piece rate wage that could drive labor productivity, I exploit the richness of my data and include (1) flexible controls for temperature, and (2) a host of fixed effects. Most importantly, I include time fixed effects to capture seasonality (week-of-year), work patterns (day-of-week, hour-of-day), and season-specific shocks (year). I also include field-level fixed effects to capture variation in the productivity of different varieties and plantings of blueberry bushes. The combination of time- and field-level fixed effects gives me a credible control for the average density of blueberries available for harvest at a given time in a given field. In other words, these fixed effects allow me to control for resource abundance (θ). Further, I include worker-specific fixed effects to capture heterogeneity in picker ability. Lastly, I include a quadratic of worker tenure to allow for learning-by-doing. When estimating the effect of temperature on productivity, my identifying assumption is that individual realizations of temperature are as good as random after including the controls described here and the piece rate wage.

To address the endogeneity of piece rate wages to labor productivity, I instrument for these wages using California market prices for blueberries. In order for these prices (described in section 1.4.3) to be a valid instrument for wages, they must be correlated with farms' piece rates, but not affect labor productivity through any other channel. Figure 1.10 plots piece rate wages and market prices over time and suggests a strong correlation between the two variables. I provide formal evidence of this relationship in table 1.4, which I describe in detail in the following section. As evidence that the exclusion restriction holds – that market prices do not affect labor productivity except through wages – I rely on the size and heterogeneity of the California blueberry industry. Statewide market prices capture supply shocks from growing regions around the globe, each with different weather, growing conditions, and labor markets. To the extent that environmental conditions agronomically drive blueberry production, they do so differentially across different growing regions of California. Therefore, any one farm's temperature shocks in a given growing season do not determine aggregate blueberry supply.[23] Additionally, both of the farms I study are quite small in comparison to the statewide market: they are price-takers and cannot independently affect average prices. As a result, market prices capture exogenous variation in aggregate supply shocks and serve as an effective instrument for piece rate wages.

[23]The farms I study are located in San Diego and Kern Counties, which contain 0.1% and 14%, respectively, of all California blueberry acreage (California Blueberry Commission, 2015, p. 9).

Specifically, I estimate the following equation by two-stage least squares:

$$y_{itl} = \beta \hat{r}_{itl} + f(T_{tl}) + \boldsymbol{X}'_{it}\boldsymbol{\gamma} + \alpha_i + \boldsymbol{\delta}_t + \mu_l + \varepsilon_{itl} \tag{1.12}$$

where $\hat{r}_{itl}$ is estimated by the following first stage:

$$r_{itl} = bp_t + f(T_{tl}) + \boldsymbol{X}'_{it}\boldsymbol{g} + a_i + \boldsymbol{d}_t + m_l + \xi_{itl}. \tag{1.13}$$

In equations (1.12) and (1.13), y_{itl} is laborer i's production per hour at time t and location l. The piece rate wage is given by r, $f(T)$ is a flexible function of temperature, $\boldsymbol{X}$ is a vector of time-varying employee characteristics including days worked and days worked squared, α (a) is a worker fixed effect, $\boldsymbol{\delta}$ ($\boldsymbol{d}$) is a vector of time fixed effects (hour of day, day of week, week of year, year), μ (m) is a field fixed effect, and ε (ξ) is an error term.

In my baseline specification, I control for temperature with a series of five-degree temperature bins. In particular,

$$f(T_{tl}) = \sum_{i=0}^{10} \tau_{50+5i} \mathbb{1}(50 + 5i \leq T_{tl} < 55 + 5i) \tag{1.14}$$

where one of the eleven bins is omitted. I choose to omit the 80–85 degree bin. This approach allows me to remain agnostic to any particular functional parameterization of the temperature response function and capture any important non-linearities. It also is in line with previous research on the economic effects of temperature, and more precisely specified than some (very good) recent work.[24]

After estimating by baseline specification, I also estimate the effect of temperature on productivity using a piecewise-linear spline function with a single node at 88.5 degrees Fahrenheit:

$$f(T_{tl}) = \begin{cases} \tau_0 + \tau_1 T_{tl} & \text{if } \; T_{tl} \leq 88.5 \\ \tau_0 + \tau_1 88.5 + \tau_2 \left(T_{lt} - 88.5\right) & \text{if } \; T_{tl} \geq 88.5 \end{cases} \tag{1.15}$$

where the choice of 88.5 degrees is motivated by the results of my baseline specification.

In order to estimate how temperature affects productivity's responsiveness to piece rate wages, I estimate a variation of specification (1.12) without temperature controls on the observations in each five-degree temperature bin separately. I have too few observations to do this for the 50–55, 95–100, and 100–105 temperature bins, so I pool observations cooler than sixty degrees and hotter than ninety degrees, leaving me with eight separate estimates of how piece rate wages affect labor productivity across temperature.

In all my regressions, I two-way cluster my standard errors by day and worker to account for correlated error terms on the same day across different workers and for a single picker over time (Cameron et al., 2011).

[24] Baylis (2016), as one example, uses temperature bins of five degrees Celsius. These are equivalent to bins of nine degrees Fahrenheit.

1.6 Results

Table 1.2 presents the results of estimating my primary specification, equation (1.12), with different sets of controls. In column (1), I include only the instrumented piece rate wage and five-degree temperature bins. As expected, without controlling for seasonality or harvest field, I find a statistically significant negative effect of wages on productivity. I also find large and negative effects of cool (50–60 degree) temperatures on productivity. In each subsequent column, I add more controls: farm fixed effects, field fixed effects, worker tenure controls, time fixed effects (year, week-of-year, day-of-week, hour-of-day), and worker fixed effects. Including time fixed effects (moving from column (4) to column (5)) makes the largest difference to the sign and significance of my results. This makes sense, since seasonality and time-of-day dynamics are particularly relevant in the California blueberry context.

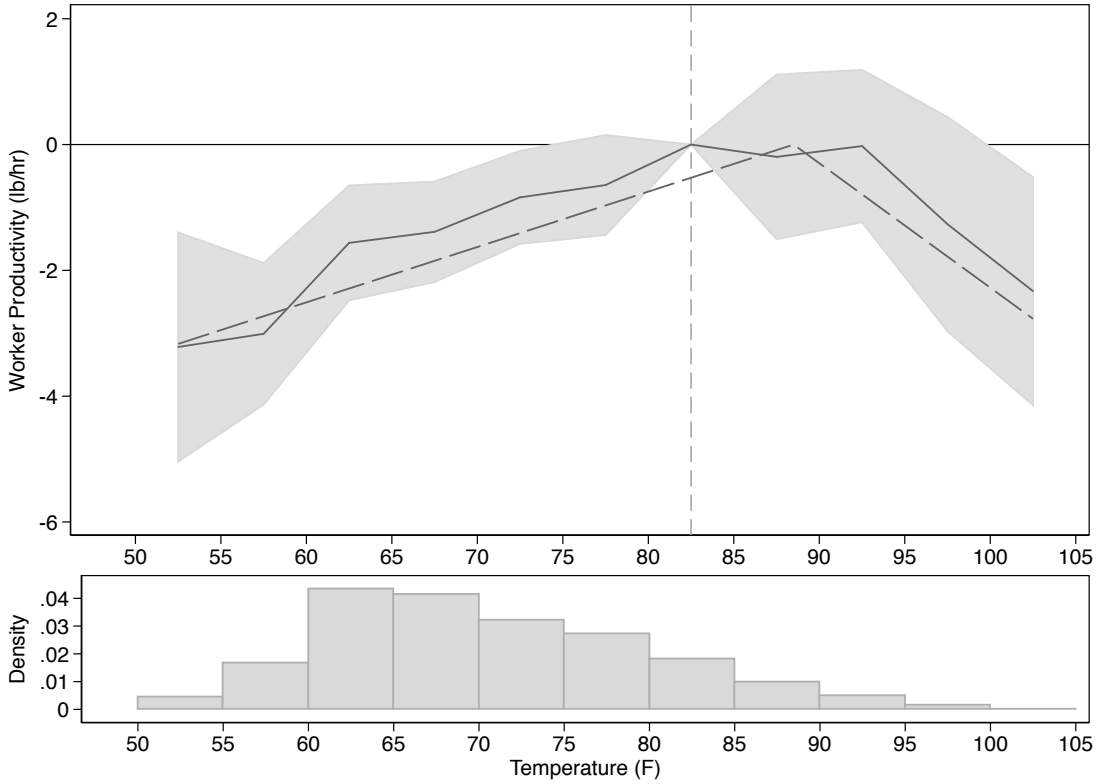
Column (6) of table 1.2 contains the results of my preferred specification using the temperature bins described in equation (1.14). By controlling for field and time fixed effects, (i.e. by controlling for resource abundance θ to the extent possible), the point-estimate for piece rate wages' effect on worker productivity switches from negative and statistically significant to positive but statistically indistinguishable from zero. The standard error on this effect is qualitatively small, meaning that I can reject even modest effects of wage on productivity.

I also find statistically significant negative effects of both cool temperatures (50–75 degrees Fahrenheit) and very hot temperatures (100+ degrees Fahrenheit) on picker productivity. The solid line in figure 1.14 plots this temperature-response function with a 95%-confidence interval. The relevant temperature point estimates (the τ_{50+5i} terms from equation (1.14)) represent the change in conditional average picker productivity (measured in pounds per hour) expected by replacing a picking period with a time-weighted average temperature between 80–85°F (the omitted temperature bin) with a picking period having a time-weighted average temperature within the corresponding temperature bin. I find that temperatures between 50 and 55 degrees lower productivity by 3.22 pounds per hour – a nearly 17% decrease, while temperatures over 100 degrees lower productivity by 2.33 pounds per hour – just over a 12% decrease.

Table 1.3 re-estimates my preferred specification using the piecewise-linear spline described in equation (1.15). I find that at temperatures below 88.5 degrees Fahrenheit, an additional degree of heat increases productivity by 0.088 pounds per hour, on average. At temperatures above 88.5 degrees, however, an additional degree of heat lowers productivity by 0.20 pounds per hour. The dashed line in figure 1.14 plots these effects, which are significant at the 0.001 and 0.05 levels, respectively.

In table 1.4, I provide evidence that blueberry market prices are an effective instrument for piece rate wages. Column (1) reports the results of estimating equation (1.12) by ordinary least squares without instrumenting for wages. While the estimated effect of wages on productivity in this specification is statistically insignificant, the point estimate is negative. Column (2) presents the results of regressing market prices, temperature, and other controls on the piece rate wage: my first stage. There is a large, positive, and statistically significant effect of prices on wages, while temperature has no meaningful effects on piece rates below

Figure 1.14: Effect of Temperature on Worker Productivity



This figure plots the relationship between temperature and worker productivity (pounds of fruit picked per hour) while controlling for instrumented piece rate wages, worker tenure, worker tenure squared, field fixed effects, time fixed effects (year, week-of-year, day-of-week, hour-of-day), and worker fixed effects. The solid line in the top panel displays the point-estimates for the five-degree temperature bins in specification (6) of table 1.2 with an omitted temperature bin of 80–85 degrees. The light gray region surrounding this line signifies a 95% confidence interval, two-way clustered on date and worker. The dashed line in the top panel displays a piecewise linear specification of the relationship between temperature and worker productivity with a single node at 88.5 degrees Fahrenheit as in table 1.3. Below $88.5°F$, an additional degree in temperature elicits an 0.088 lb/hr increase in worker productivity. Above $88.5°F$, however, an additional degree in temperature reduces worker productivity by 0.20 lb/hr. These effects are statistically significant at the 0.001 and 0.05 levels, respectively, again two-way clustering on date and worker. Finally, the bottom panel displays a histogram of temperature observations in my data.

95 degrees Fahrenheit. Column (3) gives results from a reduced form specification regressing market prices and controls on worker productivity directly, and column (4) provides the results of my preferred two-stage least squares specification instrumenting for wages with market prices. When I instrument for wages, their effect on worker productivity remains statistically insignificant, but the relevant point estimate becomes barely positive. The

temperature response function (as estimated by the point-estimates for each temperature bin) is quite stable across columns (1), (2), and (4), lending support to the conclusion that I accurately recover a true relationship.

While the richness of my data allows me to exploit intra-day variation in temperature, I can also collapse my data to the day-level and investigate how daily temperature affects daily worker productivity. Figure 1.15 reports the results of three different day-level temperature specifications. The first uses time-weighted average daily temperature experienced by each picker, the second uses daily maximum temperature, and the third uses daily minimum temperature. Overall, the results from these specifications support the qualitative results of my primary specification: extreme temperatures lower picker productivity, and cool temperatures are more damaging than very hot temperatures.

One threat to the credibility of my findings in tables 1.2 and 1.3 is that temperature and wages may affect workers' labor supply, both on the intensive and extensive margins. That is, workers may decide to work fewer hours on a particularly hot day, or choose not to come to work at all if the piece rate wage is particularly low.[25] Such behavior would bias my estimates of how temperature and wages affect productivity by introducing unobserved systematic selection into or out of my sample. I investigate this possibility in table 1.5 by regressing temperature, wages, and controls on both hours worked and the probability of working. In column (1), the dependent variable is the number of hours worked by a picker in a single day, and temperature is measured as a time-weighted average experienced by the picker during that (entire) day. Here, I control for a picker's start-time rather than their picking "midpoint." In column (3), the dependent variable is an indicator for whether a picker worked at all in a given day, and temperature is measured as a daily midpoint temperature: (Daily Max + Daily Min)/2. I use daily midpoint temperature in column (3) in order to provide a consistent comparison between employees who show up to work and employees who do not, since I do not know when or for how long these absent employees would have worked had they come to work. Figure 1.16 displays the relevant temperature results from columns (1) and (3) of table 1.5. Overall, table 1.5 reports that neither wages nor temperatures affect labor supply in a statistically significant way. Similar to Graff Zivin and Neidell (2012), I find the labor supply of agricultural workers to be highly inelastic in the short run. This also matches the findings of Sudarshan et al. (2015) for weaving workers in India. This evidence gives me confidence in the validity of my baseline results.[26]

I now turn to how temperature affects berry pickers' wage responsiveness. Table 1.6 reports the results of estimating a variant of equation (1.12) separately across eight temperature bins.[27] I find that wages have no meaningful effect on productivity at most temperatures,

[25] See Dickinson (1999) for an experimental treatment of these issues in a lab setting with both varying piece rates and labor supply choices.

[26] As an aside, the point-estimates in columns (1) and (2) of table 1.5, while not statistically significant, could be compared to the findings in Graff Zivin and Neidell (2014). These authors find that individuals reduce labor supply at hot temperatures and reduce time outside at cold temperatures. These same broad patterns, without statistical significance, appear at smaller magnitudes in my table 1.5.

[27] As discussed in the previous section, I pool cool and very hot observations to achieve the sample size

but have a statistically significant and positive effect on productivity at cool temperatures: those between 50 and 60 degrees. In particular, my estimate suggests an increase in the piece rate wage of one cent per pound at temperatures below 60 degrees increases average productivity by 0.28 pounds per hour. This reflects an elasticity of productivity with respect to the wage of roughly 1.6 at cool temperatures,[28] and an elasticity statistically indistinguishable from zero at other temperatures. This "productivity elasticity" is considerably smaller than the 2.14 number estimated by Paarsch and Shearer (1999).

Table 1.7, which repeats the analysis from table 1.6 using ordinary least squares (OLS), highlights the importance of instrumenting for piece rate wages. This table highlights two important things. First, the effects of wages on productivity at low temperatures do not show up in a statistically significant way without correctly instrumenting for wages with market prices. Second, I am able to rule out any dramatically large effect of wages on productivity at most temperatures.

Another threat to my findings is that workers who do not out-earn the hourly minimum wage in a given day may shirk ("slack off") when they know that additional productivity will not increase their take-home pay. Figure 1.13 reports the frequency with which workers fall below this minimum wage threshold. I face an econometric problem if the effects of temperature reduce workers' productivity, increase the probability that workers earn the minimum wage, and hence encourage shirking. To ensure my findings are not meaningfully altered by this phenomenon, I re-estimate my main results using only picker observations where the picker out-earns the minimum wage for the day. This procedure drops my number of picking period observations from 305,980 to 257,689: a decrease of 15.8%. Figure 1.17 and table 1.8 present the results of my main temperature and piece rate wage specifications using this subsample. My findings remain qualitatively stable and statistically significant.[29]

Finally, even if temperature and wages do not affect labor supply directly in a statistically significant manner, and even though worker-specific fixed effects capture individual workers' average productivity levels, I still face a potential adverse selection problem. Specifically, if variation in temperature and wages affects *which sorts* of workers choose to show up for work, my results may capture workforce compositional effects rather than individual productivity

necessary to estimate effects. The eight temperature bins, measured in degrees Fahrenheit, are: [50, 60), [60, 65), [65, 70), [70, 75), [75, 80), [80, 85), [85, 90), and [90, 105).

[28] An increase in productivity of 0.28 pounds per hour reflects an approximately 1.6% increase in productivity from the average 18.0 pounds per hour at temperatures between 50 and 60 degrees (column (1) of table 1.6). An increase in the piece rate wage of one cent reflects an approximately 1.0% increase in wages from the average 97 cents per pound across all observations (see panel C in table 1.1).

[29] Graff Zivin and Neidell (2012) face a similar problem on a larger scale: over 60% of their observations face a binding minimum wage policy. In that case, as a robustness exercise, the authors apply a Tobit model to a normalized measure of productivity across different crops and confirm their full-sample findings. While Tobit models rely on strong distributional assumptions, they attempt to recover the true relationship between dependent and independent variables, even for observations below some cutoff (in this case, observations below the minimum wage cutoff). My decision to drop observations below the minimum wage threshold is a more conservative approach that will understate the true relationship (if one exists) between independent and dependent variables. Thus, I remain confident in the results of my primary specifications.

effects. To address this concern, I re-estimate my results only using observations from those workers who work more than thirty days in the relevant season. The intention here is to focus on workers who are likely to have the least elastic extensive labor supply. The results of this robustness exercise are presented in figure 1.18 and table 1.9. Taken together with the other available evidence, these results largely support my baseline findings.

Table 1.2: Effects of Wage and Temperature on Worker Productivity

	(1)	(2)	(3)	(4)	(5)	(6)
Piece Rate Wage (¢/lb)	-0.20***	-0.14***	-0.16***	-0.074***	0.019	0.020
	(0.013)	(0.016)	(0.021)	(0.017)	(0.049)	(0.060)
Temperature $\in [50, 55)$	-8.68***	-9.36***	-9.46***	-9.62***	-2.93***	-3.22***
	(0.87)	(0.86)	(0.81)	(0.74)	(0.94)	(0.93)
Temperature $\in [55, 60)$	-2.78***	-3.09***	-3.37***	-3.34***	-2.78***	-3.01***
	(0.93)	(0.90)	(0.84)	(0.77)	(0.60)	(0.57)
Temperature $\in [60, 65)$	-1.34*	-1.23	-1.37*	-1.25*	-1.39***	-1.56***
	(0.76)	(0.75)	(0.71)	(0.67)	(0.49)	(0.47)
Temperature $\in [65, 70)$	-1.18	-1.09	-1.30*	-1.15*	-1.25***	-1.39***
	(0.73)	(0.72)	(0.67)	(0.65)	(0.43)	(0.41)
Temperature $\in [70, 75)$	-0.23	-0.37	-0.53	-0.54	-0.73*	-0.84**
	(0.58)	(0.54)	(0.52)	(0.48)	(0.38)	(0.38)
Temperature $\in [75, 80)$	-0.14	-0.056	-0.23	-0.22	-0.52	-0.64
	(0.67)	(0.65)	(0.61)	(0.59)	(0.43)	(0.41)
Temperature $\in [85, 90)$	-0.53	-0.69	-0.67	-0.70	-0.26	-0.20
	(0.90)	(0.90)	(0.87)	(0.80)	(0.68)	(0.67)
Temperature $\in [90, 95)$	-0.49	-0.38	-0.30	-0.59	-0.14	-0.024
	(0.93)	(0.83)	(0.77)	(0.70)	(0.65)	(0.62)
Temperature $\in [95, 100)$	-2.05**	-1.37*	-1.38	-2.15***	-1.06	-1.27
	(1.01)	(0.78)	(0.84)	(0.78)	(0.86)	(0.87)
Temperature $\in [100, 105)$	-1.40**	0.34	-0.64	-0.82	-1.62*	-2.33**
	(0.62)	(0.61)	(0.94)	(0.98)	(0.91)	(0.93)
Worker Tenure				0.36***	0.35***	0.36***
				(0.034)	(0.027)	(0.047)
Worker Tenure^2				-0.0051***	-0.0045***	-0.0052***
				(0.00068)	(0.00058)	(0.00071)
Number of Observations	305980	305980	305980	305980	305980	305980
Mean of Dependent Variable	19.1	19.1	19.1	19.1	19.1	19.1
Farm FE		X				
Field FE			X	X	X	X
Time FE					X	X
Worker FE						X

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Note: In all specifications, the dependent variable is worker productivity measured in pounds of fruit picked per hour (lb/hr). Results are estimated by two-stage least squares (2SLS) instrumenting for the piece rate wage with the market price for blueberries. Temperature is measured in degrees Fahrenheit, and the omitted temperature bin is 80–85 degrees. Worker tenure is the number of days an employee has worked in the current season at the time of weigh-in. Time fixed effects include year, week-of-year, day-of-week, and hour-of-day fixed effects. Standard errors are two-way clustered on date and worker.

Table 1.3: Piecewise Linear Effect of Temperature on Worker Productivity

	(1)
Piece Rate Wage (¢/lb)	0.026
	(0.060)
Temperature, $< 88.5°F$	0.088***
	(0.019)
Temperature, $> 88.5°F$	-0.20**
	(0.100)
Worker Tenure	0.36***
	(0.047)
Worker Tenure^2	-0.0052***
	(0.00070)
Number of Observations	305980
Mean of Dependent Variable	19.1
Field FE	X
Time FE	X
Worker FE	X

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Note: This table presents results of a piecewise linear relationship between temperature and worker productivity with a single node at 88.5°F. The dependent variable is worker productivity measured in pounds of fruit picked per hour (lb/hr). Results are estimated by two-stage least squares (2SLS) instrumenting for the piece rate wage with the market price for blueberries. Worker tenure is the number of days an employee has worked in the current season, measured on the day of an observation. Time fixed effects include year, week-of-year, day-of-week, and hour-of-day fixed effects. Standard errors are two-way clustered on date and worker.

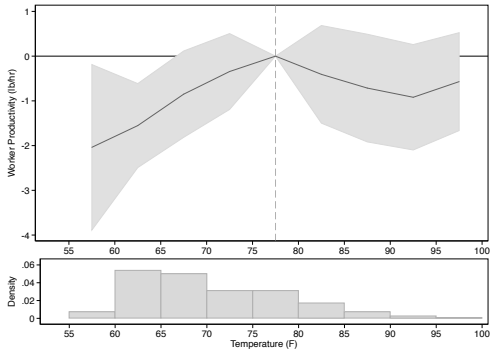
Table 1.4: Testing Market Price as an Instrument for Piece Rate Wage

	(1) OLS	(2) First Stage	(3) Reduced Form	(4) 2SLS
Piece Rate Wage (¢/lb)	-0.015			0.020
	(0.023)			(0.060)
Market Price ($/lb)		4.97***	0.100	
		(0.64)	(0.30)	
Temperature $\in [50, 55)$	-3.26***	-1.21	-3.24***	-3.22***
	(0.93)	(1.39)	(0.93)	(0.93)
Temperature $\in [55, 60)$	-3.03***	-1.22	-3.03***	-3.01***
	(0.57)	(1.13)	(0.57)	(0.57)
Temperature $\in [60, 65)$	-1.60***	-1.46*	-1.59***	-1.56***
	(0.45)	(0.85)	(0.45)	(0.47)
Temperature $\in [65, 70)$	-1.41***	-1.02	-1.41***	-1.39***
	(0.40)	(0.72)	(0.40)	(0.41)
Temperature $\in [70, 75)$	-0.85**	-0.52	-0.85**	-0.84**
	(0.38)	(0.58)	(0.38)	(0.38)
Temperature $\in [75, 80)$	-0.65	-0.19	-0.65	-0.64
	(0.41)	(0.43)	(0.40)	(0.41)
Temperature $\in [85, 90)$	-0.21	0.66	-0.18	-0.20
	(0.66)	(0.73)	(0.68)	(0.67)
Temperature $\in [90, 95)$	0.0038	1.39	0.0044	-0.024
	(0.62)	(0.88)	(0.63)	(0.62)
Temperature $\in [95, 100)$	-1.20	2.36**	-1.22	-1.27
	(0.87)	(0.96)	(0.88)	(0.87)
Temperature $\in [100, 105)$	-2.22**	6.25***	-2.20**	-2.33**
	(0.91)	(1.70)	(0.95)	(0.93)
Number of Observations	305980	305980	305980	305980
Mean of Dependent Variable	19.1	80.5	19.1	19.1
Controls for Worker Tenure	X	X	X	X
Field FE	X	X	X	X
Time FE	X	X	X	X
Worker FE	X	X	X	X

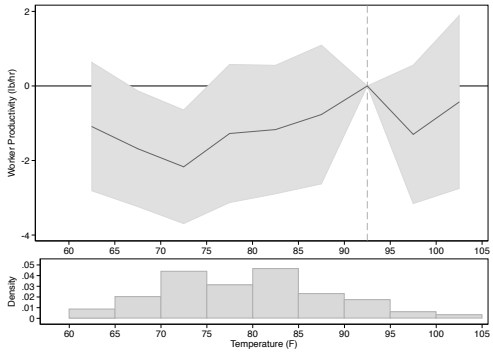
Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Note: In specifications (1), (3), and (4), the dependent variable is worker productivity measured in pounds of fruit picked per hour (lb/hr). In specification (2), the dependent variable is piece rate wage in cents per pound (¢/lb). All specifications include my preferred set of controls as in specification (6) from table 1.2. Temperature is measured in degrees Fahrenheit, and the omitted temperature bin is 80–85 degrees. Standard errors are two-way clustered on date and worker.

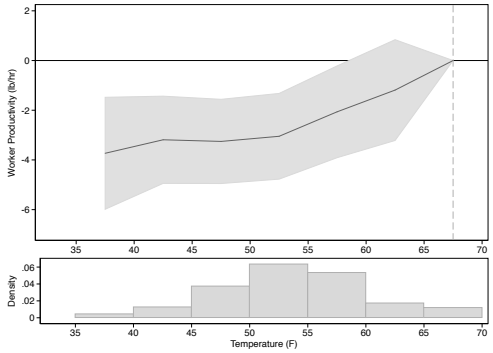
Figure 1.15: Effects of Daily Temperature on Daily Productivity



(a) Time-Weighted Daily Average Temperature



(b) Daily Maximum Temperature



(c) Daily Minimum Temperature

These figures repeat the analysis from figure 1.14, using different measures of daily temperature. Panel (a) reports the effect of time-weighted average daily exposed temperature on daily picker productivity, panel (b) reports the effect of daily maximum temperature on daily picker productivity, and panel (c) reports the effect of daily minimum temperature on daily picker productivity. These results are qualitatively similar to those in figure 1.14.

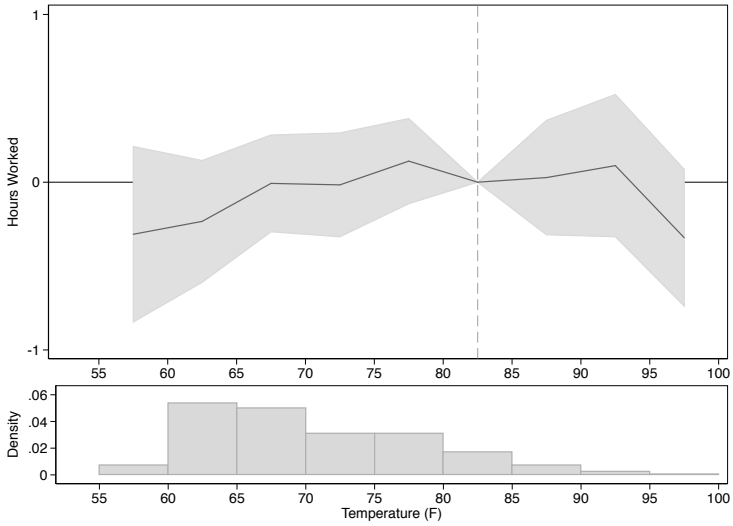
Table 1.5: Effects of Wage and Temperature on Labor Supply

	(1) Hours Worked	(2) Hours Worked	(3) Probability of Working
Piece Rate Wage (¢/lb)	0.0100		0.000078
	(0.014)		(0.0015)
Market Price ($/lb)		0.049	
		(0.070)	
Worked the Previous Day			0.72***
			(0.018)
Temperature $\in [50, 55)$			0.031
			(0.026)
Temperature $\in [55, 60)$	-0.31	-0.32	0.0030
	(0.27)	(0.27)	(0.021)
Temperature $\in [60, 65)$	-0.23	-0.26	0.0072
	(0.19)	(0.19)	(0.020)
Temperature $\in [65, 70)$	-0.0071	-0.036	-0.00059
	(0.15)	(0.15)	(0.018)
Temperature $\in [70, 75)$	-0.016	-0.026	-0.0019
	(0.16)	(0.16)	(0.021)
Temperature $\in [75, 80)$	0.13	0.11	0.024
	(0.13)	(0.13)	(0.044)
Temperature $\in [85, 90)$	0.028	0.044	
	(0.17)	(0.18)	
Temperature $\in [90, 95)$	0.099	0.11	
	(0.22)	(0.22)	
Temperature $\in [95, 100)$	-0.33	-0.31	
	(0.21)	(0.21)	
Number of Observations	47919	47919	201966
Mean of Dependent Variable	7.71	7.71	0.24
Controls for Worker Tenure	X	X	X
Start Hour FE	X	X	
Time FE	X	X	X
Worker FE	X	X	X

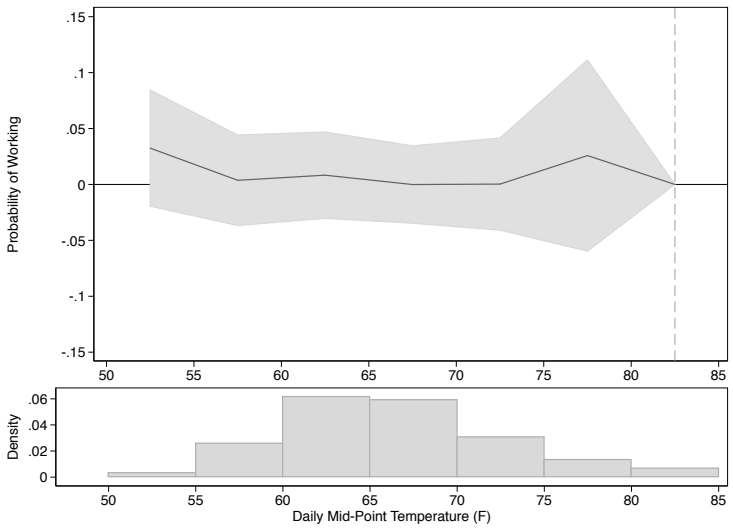
Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Note: This table reports the effects of temperature and wages on workers' labor supply, both on the intensive margin (hours worked) and on the extensive margin (probability of working). In specifications (1) and (2), the dependent variable is the number of hours worked by a picker in a single day, and temperature is measured as a time-weighted average experienced by the picker during that day. In specification (3), the dependent variable is an indicator for whether a picker worked at all in a given day, and temperature is measured as a daily midpoint temperature: (Daily Max + Daily Min)/2. Temperature is measured in degrees Fahrenheit, and the omitted temperature bin is 80–85 degrees. Specifications (1) and (3) are estimated by 2SLS, instrumenting for the piece rate wage with the market price for blueberries. Time fixed effects include year, week of year, and day of week. Standard errors are two-way clustered on date and worker.

Figure 1.16: Effects of Daily Temperature on Labor Supply



(a) Effect of Temperature on Hours Worked (Intensive Labor Supply)



(b) Effect of Temperature on Probability of Working (Extensive Labor Supply)

These figures display the results from table 1.5. Panel (a) reports the effect of temperature on pickers' intensive labor supply (the number of hours they work), and panel (b) reports the effect of temperature on pickers' extensive labor supply (the probability they show up to work on a given day). These results show that blueberry pickers have very inelastic labor supplies with respect to temperature.

Table 1.6: Effect of Piece Rate Wage on Worker Productivity by Temperature

	(1) [50, 60)	(2) [60, 65)	(3) [65, 70)	(4) [70, 75)	(5) [75, 80)	(6) [80, 85)	(7) [85, 90)	(8) [90, 105)
Piece Rate Wage (¢/lb)	0.28***	0.015	0.063	-0.029	-0.18	0.22	0.094	-0.29
	(0.097)	(0.091)	(0.083)	(0.073)	(0.13)	(0.15)	(0.091)	(0.19)
Number of Observations	32668	66265	63640	48915	41617	27912	15046	9917
Mean of Dependent Variable	18.0	18.9	19.1	20.1	19.7	19.1	17.9	17.7
Controls for Worker Tenure	X	X	X	X	X	X	X	X
Field FE	X	X	X	X	X	X	X	X
Time FE	X	X	X	X	X	X	X	X
Worker FE	X	X	X	X	X	X	X	X

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Note: In all regressions, the dependent variable is worker productivity measured in pounds of fruit picked per hour (lb/hr). Each column reports the effect of piece rate wages (instrumented by market prices via two-stage least squares) on worker productivity for a different range of temperature. Temperatures are pooled between 50 and 60 degrees and between 90 and 105 degrees to ensure a sufficient sample size to estimate effects in those ranges. Standard errors are two-way clustered on date and worker.

Table 1.7: Effect of Piece Rate Wage on Worker Productivity by Temperature: OLS

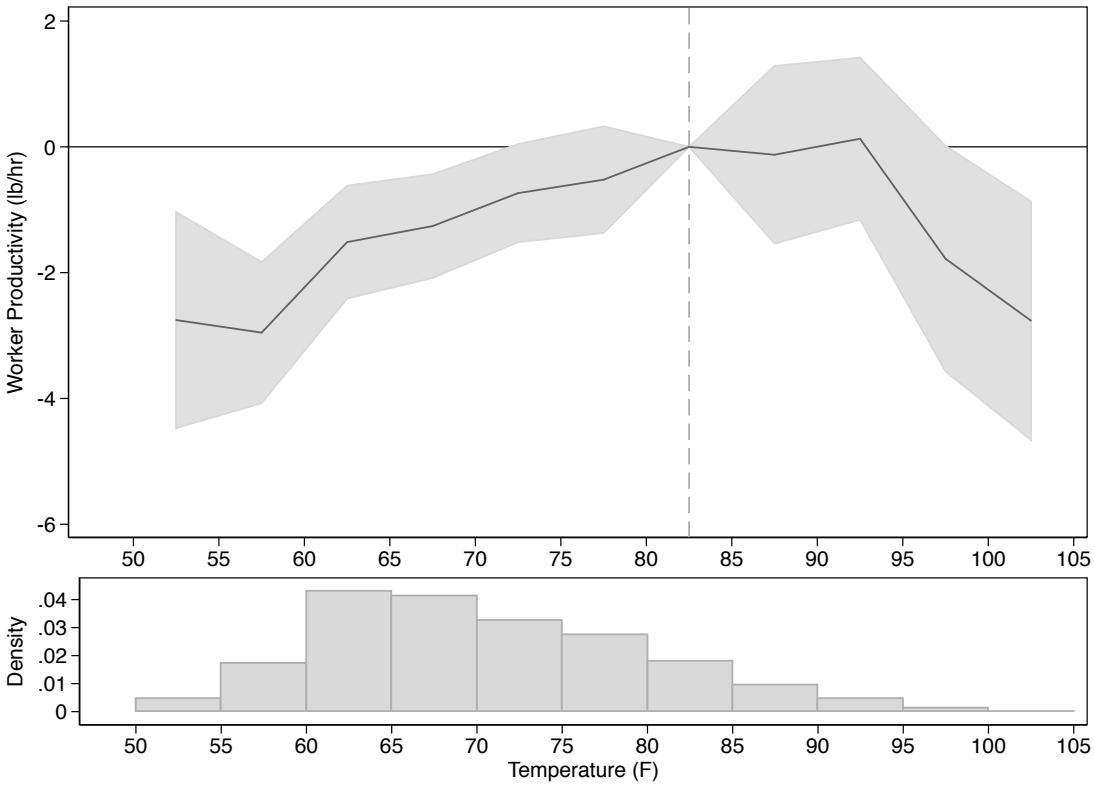
	(1) [50, 60)	(2) [60, 65)	(3) [65, 70)	(4) [70, 75)	(5) [75, 80)	(6) [80, 85)	(7) [85, 90)	(8) [90, 105)
Piece Rate Wage (¢/lb)	0.100*	-0.028	0.020	-0.032	-0.059	0.074	0.13	-0.19
	(0.050)	(0.048)	(0.041)	(0.031)	(0.054)	(0.063)	(0.078)	(0.15)
Number of Observations	32668	66265	63640	48915	41617	27912	15046	9917
Mean of Dependent Variable	18.0	18.9	19.1	20.1	19.7	19.1	17.9	17.7
Controls for Worker Tenure	X	X	X	X	X	X	X	X
Field FE	X	X	X	X	X	X	X	X
Time FE	X	X	X	X	X	X	X	X
Worker FE	X	X	X	X	X	X	X	X

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Note: This table replicates table 1.6 without instrumenting for piece rate wages. Instead, I run each regression using ordinary least squares (OLS). By comparing these results to those in table 1.6, it is clear that OLS biases my estimates toward zero. In all regressions, the dependent variable is worker productivity measured in pounds of fruit picked per hour (lb/hr). Each column reports the effect of piece rate wages on worker productivity for a different range of temperature. Temperatures are pooled between 50 and 60 degrees and between 90 and 105 degrees to ensure a sufficient sample size to estimate effects in those ranges. Standard errors are two-way clustered on date and worker.

Figure 1.17: Effect of Temperature on Worker Productivity Without Possible Shirkers



This figure plots the relationship between temperature and worker productivity (pounds of fruit picked per hour) while controlling for instrumented piece rate wages, worker tenure, worker tenure squared, field fixed effects, time fixed effects (year, week-of-year, day-of-week, hour-of-day), and worker fixed effects. This figure is estimated using only picker-day observations where pickers earn more than the hourly minimum wage. The top panel displays the point-estimates for the five-degree temperature bins as in specification (6) of table 1.2, while the bottom panel displays the support of temperature observations. The light gray region in the top panel signifies a 95% confidence interval, two-way clustered on date and worker. The omitted temperature bin is 80–85 degrees Fahrenheit.

Table 1.8: Piece Rate Wage Effect Without Possible Shirkers

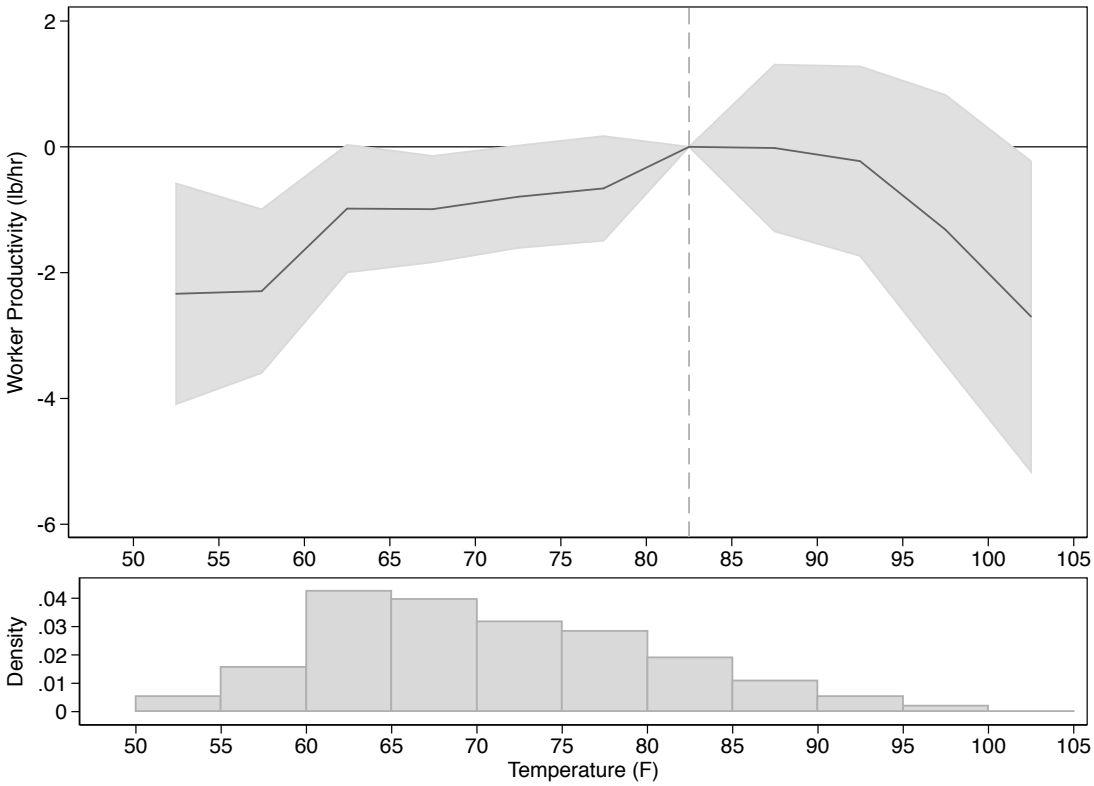
	(1) [50, 60)	(2) [60, 65)	(3) [65, 70)	(4) [70, 75)	(5) [75, 80)	(6) [80, 85)	(7) [85, 90)	(8) [90, 105)
Piece Rate Wage (¢/lb)	0.27***	0.013	0.047	-0.065	-0.20	0.20	0.014	-0.44
	(0.10)	(0.079)	(0.080)	(0.067)	(0.12)	(0.16)	(0.098)	(0.27)
Number of Observations	28494	55311	53630	41517	35300	23367	12192	7878
Mean of Dependent Variable	19.0	20.6	20.7	21.8	21.4	20.9	19.9	19.8
Controls for Worker Tenure	X	X	X	X	X	X	X	X
Field FE	X	X	X	X	X	X	X	X
Time FE	X	X	X	X	X	X	X	X
Worker FE	X	X	X	X	X	X	X	X

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Note: This table is analogous to table 1.6, with the distinction that here I include only those picking period observations for workers who earn more than the hourly minimum wage for the day. In all regressions, the dependent variable is worker productivity measured in pounds of fruit picked per hour (lb/hr). Each column reports the effect of piece rate wages (instrumented by market prices via two-stage least squares) on worker productivity for a different range of temperature. Temperatures are pooled between 50 and 60 degrees and between 90 and 105 degrees to ensure a sufficient sample size to estimate effects in those ranges. Standard errors are two-way clustered on date and worker.

Figure 1.18: Effect of Temperature on Worker Productivity Without Transient Workers



This figure plots the relationship between temperature and worker productivity (pounds of fruit picked per hour) while controlling for instrumented piece rate wages, worker tenure, worker tenure squared, field fixed effects, time fixed effects (year, week-of-year, day-of-week, hour-of-day), and worker fixed effects. This figure is estimated using only picker-day observations for workers who work more than thirty days in the relevant season. The top panel displays the point-estimates for the five-degree temperature bins as in specification (6) of table 1.2, while the bottom panel displays the support of temperature observations. The light gray region in the top panel signifies a 95% confidence interval, two-way clustered on date and worker. The omitted temperature bin is 80–85 degrees Fahrenheit.

Table 1.9: Piece Rate Wage Effect Without Transient Workers

	(1) [50, 60)	(2) [60, 65)	(3) [65, 70)	(4) [70, 75)	(5) [75, 80)	(6) [80, 85)	(7) [85, 90)	(8) [90, 105)
Piece Rate Wage (¢/lb)	0.45***	0.090	0.21**	0.13	0.25	0.39***	0.19***	-0.86***
	(0.11)	(0.090)	(0.099)	(0.10)	(0.21)	(0.13)	(0.052)	(0.18)
Number of Observations	16672	33872	31918	25130	22500	15136	8543	5973
Mean of Dependent Variable	17.8	17.3	17.8	18.3	18.0	17.0	15.6	15.8
Controls for Worker Tenure	X	X	X	X	X	X	X	X
Field FE	X	X	X	X	X	X	X	X
Time FE	X	X	X	X	X	X	X	X
Worker FE	X	X	X	X	X	X	X	X

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Note: This table is analogous to table 1.6, with the distinction that here I include only those picking period observations for workers who work more than thirty days in the relevant season. In all regressions, the dependent variable is worker productivity measured in pounds of fruit picked per hour (lb/hr). Each column reports the effect of piece rate wages (instrumented by market prices via two-stage least squares) on worker productivity for a different range of temperature. Temperatures are pooled between 50 and 60 degrees and between 90 and 105 degrees to ensure a sufficient sample size to estimate effects in those ranges. Standard errors are two-way clustered on date and worker.

1.7 Discussion

My results provide evidence of how blueberry pickers' productivity responds to piece rate wages and temperature, both independently and jointly. These findings have implications on several levels, each of which I address separately.

1.7.1 Wage effects

My primary finding is that labor productivity, on average, is very inelastic with respect to piece rate wages: I can reject with 95% confidence even modest positive elasticities of up to 0.7. This upper bound is considerably lower than the estimates derived by Paarsch and Shearer (1999) and Haley (2003). I show that, without controlling for seasonality, a regression of productivity on piece rate wages results in a negative and significant point estimate (see table 1.2). However, even once I control for seasonality, a naïve OLS regression of productivity on piece rate wage may be biased toward zero (see column (1) of table 1.4). By instrumenting for piece rate wages with the market price for blueberries, I can identify a precisely-estimated inelastic effect (see column (6) of table 1.2).

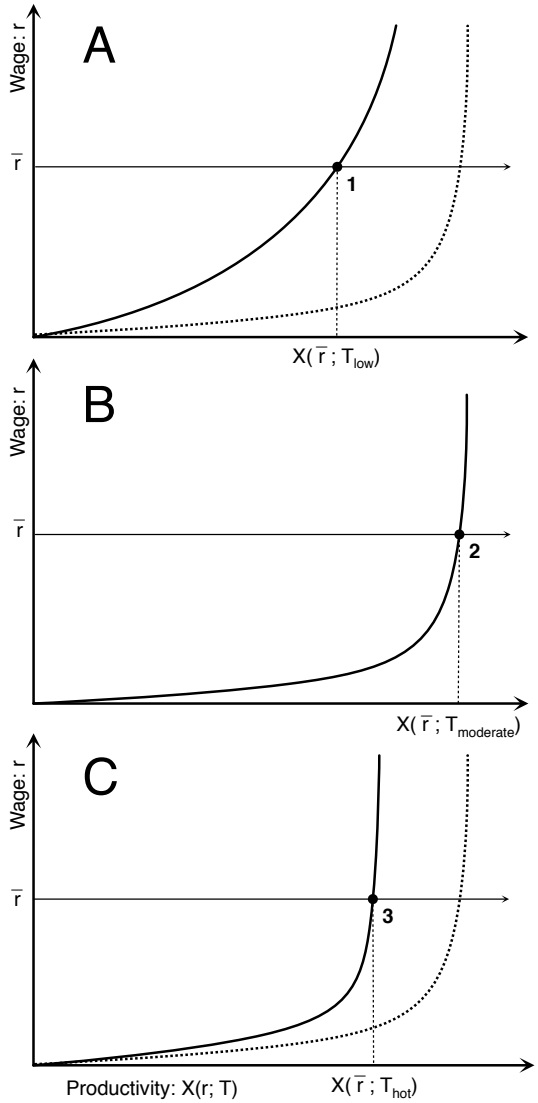
However, my primary specification makes the restrictive assumption that wages affect productivity linearly and in the same manner at all temperatures. Table 1.6 confirms that piece rates' effect on productivity is very much *non*-linear across different temperatures. Specifically, wages seem to spur productivity at cool temperatures (where workers' productivity is already depressed). At other temperatures, wages do not affect productivity in a statistically significant way. This empirical finding directly challenges one of the core assumptions of the model presented in section 1.2.1: that productivity always rises with the wage ($X_r > 0$). What is going on?

One possible explanation for my findings is that, at moderate to hot temperatures, workers' face some binding physiological constraint on effort that prevents them from responding to changes in their wage. Put bluntly, blueberry pickers in general may already be "giving all they've got" at the temperatures and wages I observe.[30] Figure 1.19 summarizes this possibility using the theoretical framework developed in section 1.2.1.

While the model in section 1.2.1 is straightforward and tractable, it is not the only way to conceptualize worker effort and productivity. In particular, rather than modeling effort as an unrestricted choice variable, one could assume each worker has a finite daily budget of effort that must be allocated across different activities throughout a day (for examples, see Becker (1965) and Becker (1977)). Such a model would allow X_r to be zero or even negative under certain conditions, implying a backwards-bending effort supply curve, somewhat analogous to the canonical backward-bending labor supply curve (Killingsworth, 1983). The downside

[30] I am not suggesting that blueberry pickers spend most of their working time on the brink of physiological exhaustion. Rather, I interpret my findings as evidence that most pickers spend most of their working time at or near a "maximum *sustainable* level of effort." Consider the analogy of a marathon runner: such a runner could certainly increase their speed for any arbitrary short distance if given a compelling incentive. However, maintaining that increased speed for the entire race would be physiologically impossible.

Figure 1.19: Graphical Summary of Findings



This figure uses the productivity curve derived in panel C of figure 1.1 to qualitatively describe my empirical findings. At moderate temperatures (panel B of this figure), I find that berry pickers are very inelastic to the piece rate wages they face. That is, the slope at point **2** is close to vertical. At very hot temperatures (over 100°F; panel C of this figure), productivity is similarly inelastic to wages (point **3**). However, these extreme temperatures have a direct negative effect on productivity (**3** is to the left of **2**). Finally, at cool temperatures (under 60°F; panel A of this figure), I find a positive elasticity of productivity with respect to the piece rate wage. That is, the slope at point **1** is less than vertical. However, cool temperatures also have a large direct negative effect on productivity (**1** is to the left of both **2** and **3**).

of such models is that they fail to provide comparative statics that can be tested with the data I observe in this setting.

1.7.2 Direct temperature effects

My econometric analyses allow me to estimate the direct effects of temperature on California berry-pickers' labor productivity. These findings contribute to a large literature studying the effects of temperature and other environmental conditions on a variety of economic outcomes.

1.7.2.1 Existing literature on temperature and other environmental conditions

A growing literature has rigorously documented the non-linear impact of temperature on everything from corn yields (Schlenker and Roberts, 2009) to cognitive performance (Graff Zivin et al., 2015), but has not focused specifically on how temperature affects agricultural workers.[31] Nevertheless, several recent papers in this literature seem particularly relevant to my findings. One strand of research has investigated how temperature affects labor productivity in a variety of different industries. Adhvaryu et al. (2016b) show that factory workers in India produce more output when heat-emitting conventional light bulbs are replaced LED lighting, especially on hot days. Sudarshan et al. (2015) find similar evidence that temperature reduces worker productivity in a variety of Indian manufacturing firms. Finally, Seppänen et al. (2006) show that temperature even has large effects on the productivity of office workers.[32]

Other researchers have asked broader questions about how temperature affects aggregate production or labor decisions at the county- or country-level. The growing consensus is that weather shocks – particularly exposures to extreme heat – reduce aggregate production in a wide variety of settings. For instance, Hsiang (2010) exploits natural variation in cyclones to find negative impacts of high temperatures in both agricultural and non-agricultural sectors at the country-level. Deryugina and Hsiang (2014) and Park (2016) find similar county-level effects of daily temperature in the United States, despite widespread adoption of air conditioning. Heal and Park (2014) document relevant findings throughout the economics literature and provide a useful theoretical link between heat's physiological effects and aggregate economic activity.[33] Extreme heat may reduce aggregate production through several

[31] For useful reviews of the economic literature on temperature, see Carleton and Hsiang (2016), Kahn (2016), and Heal and Park (2016).

[32] In a recent and extensive summary of the economic risks of climate change, Houser et al. (2015) highlight that relatively little is known about how temperature affects worker effort: "While we consider the effects of temperature on the number of hours worked, we do not assess the effects on the intensity of labor during working hours" (p. 167).

[33] There is a vast medical, physiological, and ergonomic literature documenting the ways in which temperature affects the human body. Hot temperatures consistently tax individuals' endurance, exacerbate fatigue, and diminish cognitive performance in a variety of experimental settings. For evidence on endurance and fatigue, see Nielsen et al. (1993), Galloway and Maughan (1997), and González-Alonso et al. (1999). For evidence on cognitive performance, see Epstein et al. (1980), Ramsey (1995), Pilcher et al. (2002), and

channels. The first possibility, discussed at length in the previous paragraph, is that employees are less productive while working at high temperatures. Another possibility is that employees may choose to work fewer hours when temperatures are particularly high. In other words, there may be a labor supply response to temperature on the extensive margin. Graff Zivin and Neidell (2014) provide support for this hypothesis by analyzing data from the American Time Use Survey. They find that at high temperatures, individuals reduce the time they spend working and increase the time they spend on indoor leisure. Finally, temperature can affect even broader aspects of the labor market like aggregate demand for agricultural labor in India (Colmer, 2016), or the composition of labor in urban vs. rural regions of Eastern Africa (Dou et al., 2016).

While this paper examines how a particularly salient environmental condition, temperature, affects labor productivity, previous research has shown that other environmental factors matter as well. Chang et al. (2016a), for instance, find that outdoor air pollution negatively affects the indoor productivity of pear packers. The same authors conduct a similar exercise using data from Chinese call-centers (Chang et al., 2016b) and find comparable results. Adhvaryu et al. (2016a) find a steep pollution-productivity gradient in the context of an Indian garment factory, and Graff Zivin and Neidell (2012) find large damages from ozone in an agricultural context somewhat similar to my own. In an older case study, Crocker and Horst, Jr. (1981) study seventeen citrus pickers in southern California and find negative effects of both high temperatures and air pollution. It is useful to think of temperature not as a single sufficient statistic to describe environmental quality, but rather as one condition among many that is relevant for understanding labor productivity.

This paper makes several important contributions to the literature discussed above. First, because I observe berry-pickers' productivity multiple times during a single day, the variation I observe in both productivity and temperature is much more temporally precise than in many previous studies. Additionally, since I use temperature observations that are taken hourly, and sometimes more frequently, I do not need to interpolate temperature over time. Second, I study a setting where both very hot and cool temperatures have negative effects on productivity, highlighting the particularities of different production processes when it comes to temperature impacts. Third, and most importantly, I look at how how environmental conditions and incentive schemes interact.[34]

Hancock et al. (2007).

[34] An early ergonomic laboratory experiment (Mackworth, 1947) found that incentivizing subjects to improve their performance on a simple physical task increased their physiological endurance at both normal and hot temperatures. However, incentives were not able to eliminate a stark performance drop-off as temperatures increased. More recently, Park (2016) noted that temperature's effect on labor productivity should be expected to vary with the incentive scheme: "...the optimal [temperature] response for someone who is paid a piece rate wage contract... will differ from someone who is paid on a fixed annual contract or simply by the hour" (p. 350).

1.7.2.2 Temperature effects on California blueberry pickers

Table 1.2 and figure 1.14 provide my estimates of the direct effects of temperature on labor productivity in the California blueberry industry. Whereas most previous studies have focused on the negative effects of extreme heat (Dell et al., 2012; Heal and Park, 2014; Deryugina and Hsiang, 2014), I find that cool temperatures (50–60 degrees Fahrenheit, in particular) have just as large negative effects as very hot temperatures, if not larger. While this may appear counterintuitive at first, two insights help to explain this result. First, blueberry farmers have already adapted to hot temperatures: pickers generally finish picking around 3:00 p.m. and avoid the hottest parts of the day. This means that I do not observe how workers would perform under temperatures above 100–105 degrees.[35] And looking at the temperature response function in figure 1.14, it is easy to imagine due to its overall inverse-parabolic shape that there would be even larger productivity losses at such high temperatures. Second, blueberry picking is a highly dextrous job requiring workers to use their bare hands to pick only ripe berries from the bush. At cooler temperatures, berry pickers lose finger dexterity and find it uncomfortable to maintain the same levels of productivity as at warmer temperatures.[36] Indeed, Enander and Hygge (1990) note that manual dexterity can start to be impaired at temperatures in the range of 12–15 degrees Celsius (53.6–59 degrees Fahrenheit).[37]

1.7.3 Implications for California blueberry growers

My empirical findings explain (1) how blueberry pickers respond to changes in their piece rate wage, (2) how temperature directly affects these pickers' labor productivity, and (3) how temperature affects these pickers' wage responsiveness. Combining these effects, I can predict how rising temperatures will affect labor market outcomes in the California blueberry picking sector. In appendix A, I develop a model of a firm choosing an optimal piece rate under some environmental condition. That model produces two particularly interesting comparative statics results.

First, condition (A.10) predicts that farmers should increase their piece rate wage in response to an increase in temperature if and only if $X_T < (p-r)X_{rT}$, where X is worker productivity, T is temperature, p is the market price of blueberries, and r is the piece rate wage. Using the results of my linear spline specification (table 1.3), I find that, at temperatures below $88.5°F$, $X_T = 0.088$ pounds per hour for a one-degree-Fahrenheit increase in temperature. Furthermore, for temperatures between 50 and 65 degrees, I estimate $X_{rT} \approx -0.035$

[35] In fact, all the temperature observations in my 100–105 degree bin are below 102 degrees Fahrenheit.

[36] Ambient temperature is not the only consideration here. Berries themselves are cool to the touch each morning and compound the effect of air temperature on pickers' hand dexterity.

[37] The authors also write: "Tasks involving fine movements of the fingers and hands or manipulation of small objects are particularly sensitive to cold effects. Slow cooling, as found in many occupational settings, is more detrimental to manual performance than rapid cooling" (Enander and Hygge, 1990, p. 46).

pounds per hour for an increase of one degree.[38] Thus, $X_T/X_{rT} \approx -2.5 < 0 < p - r$, and a profit-maximizing farmer would choose to decrease the wage r for temperature increases at these temperatures.[39] Put more clearly, my results suggest that an optimizing farmer would pay pickers higher piece rate wages at particularly cool temperatures than at more moderate temperatures.

At warmer levels, between 90 and 105 degrees Fahrenheit, I am unable to repeat the analysis of the previous paragraph since the condition in expression (A.10) is only valid when $X_r > 0$. Therefore, I rely on more basic economic intuition when analyzing behavior at these temperatures: since there is no discernible benefit of higher wages at higher temperatures, an optimizing farmer would not raise piece rates at particularly hot temperatures.

The analyses of previous paragraphs relied on the important assumption that farmers are able to set piece rates that vary with temperature. In reality, a farmer only sets a single piece rate wage for the entire day. As a result, the entire distribution of expected temperature throughout a day is what matters to the farmer. Since most days during the blueberry season in California include a wide range of temperatures, farmers are likely already close to a second-best optimum given the constraint of a single daily wage.[40] Nonetheless, farmers may be able to boost profits by increasing their piece rates on particularly cool "cold snap" days, especially if the presence of such a premium could be tied to an external authority (i.e. *"I pay 2 ¢ per pound extra on days when the daily high temperature in the local newspaper is predicted to be below 75°F"*).

The second interesting prediction from the model in appendix A (see equation (A.13)) is that, at the margin, farm profits will move with X_T. Therefore, at low temperatures, a warming climate may increase California blueberry farms' profitability. However, if farm workers are then also exposed to higher, more extreme temperatures, profits may decrease at such temperatures.

How then, in sum, could climate change affect the California blueberry industry? My results suggest that average picker productivity would likely increase as workers face fewer "cold" hours in the fields. This effect will almost certainly dominate the marginal decreases in productivity at very hot temperatures, given the distribution of temperature to which pickers are currently exposed (see the lower panel in figure 1.14). Therefore, as far as labor considerations are concerned, moderate climate warming may increase profits for California blueberry growers. However, if pickers begin working in extreme heat conditions that they currently avoid, there may be significant negative impacts that are not captured by my analysis. From the farmer's perspective, an ideal response to climate concerns would be to

[38] I calculate this figure by subtracting the point estimate in column (2) of table 1.6 from the point estimate in column (1) of the same table and dividing by 7.5 (degrees). Specifically, $(0.28 - 0.015)/-7.5 \approx -0.035$.

[39] If $X_{rT} < 0$, as it is here, then the condition $dr/dT > 0 \iff X_T < (p-r)X_{rT}$ can be rewritten as $dr/dT > 0 \iff X_T/X_{rT} > (p-r)$.

[40] It is not even clear that an unconstrained variable piece rate wage regime would dominate the *status quo*, since a varying rate would introduce incentives for workers to manipulate that system, or horde fruit until times of the day when wages are highest. The behavioral losses incurred by such a system could easily outweigh its theoretical benefit.

hire more employees for fewer hours each day, focusing on times when temperatures are most amenable to high productivity. This is impractical, though, since many fruit pickers expect a full day's work and would likely look for other employers if a farmer failed to provide the opportunity for an eight-hour work day. As a final caveat, this paper explores only the effects of temperature on labor productivity. A changing climate is also likely to have important agronomic consequences (à la Schlenker and Roberts (2009)) that I do not address here.

1.8 Conclusion

In agriculture – as in many other industries – labor is a primary input, pay is tied to worker output, and firms cannot completely control important workplace environmental conditions like temperature. How do agricultural workers respond to changes in their piece rate wage? How does temperature affect this wage responsiveness? And what are the net effects of temperature on agricultural labor productivity? This paper addresses these questions in the context of California blueberry farmers and provides the following answers: (1) on average, blueberry pickers' productivity is very inelastic with respect to wages; (2) workers seem to face binding constraints on effort at moderate to hot temperatures, but display an elastic response to wages at cool temperatures; and (3) both very hot and cool temperatures have negative direct effects on berry pickers' productivity.

This paper makes a meaningful contribution to the empirical understanding of how wages affect worker productivity. While the basic theoretical prediction is straightforward (under piece rate wages, a higher wage should encourage more effort and higher output), previous studies have struggled to test this hypothesis directly. Doing so is difficult since, in settings where piece rates vary over time, their variation is endogenous to worker productivity. To isolate wages' effect on productivity, I instrument for blueberry pickers' piece rate wage using the market price for California blueberries. I find that on average, pickers' productivity is very inelastic with respect to piece rate wages, and I can reject even modest elasticities of up to 0.7. However, this finding hides important heterogeneity in the relationship across different temperatures. In particular, only at cool temperatures (50–60 degrees Fahrenheit) do higher wages have a statistically significant and positive effect on worker productivity. This result suggests that at most temperatures and wages, blueberry pickers face some sort of binding constraint on effort and cannot be incentivized to increase their productivity.

This research raises questions for future research both about firms' responses to changing temperatures and their choice of an optimal payment scheme. For instance, it would be helpful to analyze a different industry to see how temperature response functions differ across tasks. It would also be interesting to analyze, both theoretically and empirically, a varying wage scheme tied directly to exogenous factors such as market prices, resource abundance, and environmental conditions. With the advent of cheap, sophisticated monitoring technology, more and more industries are candidates for adopting piece rates, raising the importance for economists to deepen our understanding of the forces at work in such wage schemes.

Chapter 2

Reap What Your Friends Sow: Social Networks and Technology Adoption

joint with Fiona Burlig

2.1 Introduction

Technology adoption is an essential component of economic growth (Hall and Khan (2003); Foster and Rosenzweig (2010); Perla and Tonetti (2014)). In 2015 alone, the World Bank committed over eight billion dollars to projects encouraging people to adopt new technologies. Over the past decade, economists and policymakers have begun to recognize that social networks can facilitate technology adoption. In particular, information barriers hinder the take-up of new technologies; social networks can spread information and reduce these frictions. Understanding the ways in which these networks impact the take-up of new technologies is relevant for policymakers across the developed and developing world.

Economists face a fundamental challenge when trying to study social networks, since these networks are endogenously formed: people choose their own friends. Though there is a broad theoretical literature on social networks[1], endogenous network formation poses a significant challenge for empirical research (Manski (1993); Goldsmith-Pinkham and Imbens (2013); Jackson (2014); Choi et al. (2016)). In response to these difficulties, recent work in economics has relied on randomized experiments that act on or through existing social networks in field settings.[2] Other work uses detailed data on network structures to study how information moves within existing networks.[3]

[1] See Jackson (2008) for an overview of the theoretical work on networks.

[2] See, for example, Beaman and Magruder (2012), Banerjee et al. (2016), and Beaman et al. (2016).

[3] Prominent examples include Kremer and Miguel (2007), who examine peer effects in deworming in Kenya; Banerjee et al. (2013), who develop a model of information diffusion through networks using data from India; and Alatas et al. (2016), who model network-based information aggregation in Indonesian villages.

These papers represent a major development in our understanding of how information is transmitted through social networks. What they are unable to do, however, is analyze how naturally-arising changes in these networks affect economic activity. A small literature exists that attempts to address this issue by estimating the effects of plausibly exogenous shocks to existing social networks on economic outcomes. The majority of these papers in this focus on how social networks affect labor market outcomes (Munshi (2003), Edin et al. (2003), and Beaman (2012)).[4]

Though none of these papers studies technology adoption, there is a rich literature in economics studying the diffusion and take-up of new technologies, particularly in agricultural settings.[5] Our work is most closely related to several recent papers which study the role of social networks in agricultural technology adoption.[6] Foster and Rosenzweig (1995) and Munshi (2004) study the network determinants of technology adoption during India's Green Revolution. Conley and Udry (2001; 2010) study farmer learning about fertilizer use and pineapple in Ghana. Bandiera and Rasul (2006) find that family and religious communities matter for technology adoption in Mozambique.[7] Vasilaky (2013) and Vasilaky and Leonard (2014) randomly connect women with agricultural extension agents, and find that this dramatically improves productivity.

In this paper, we are able to directly estimate the causal effects of increases in network size and composition on technology adoption in agriculture. We take advantage of a unique natural experiment to isolate exogenous shocks to social networks. In particular, these shocks take the form of mergers between rural congregations of the American Lutheran Church between 1959 and 1964 in the Upper Midwest of the United States. These mergers were caused by national-level church mergers, church building fires, and pastoral employment constraints, all of which were beyond the control of individual congregations. Using county-level data from the American Census of Agriculture, we employ a difference-in-differences approach to study how these mergers affected farmers' adoption of inorganic nitrogen fertilizer – at the time, a relatively new yield-improving technology.

We demonstrate that congregational mergers had an economically meaningful effect on technology adoption among farmers. The number of farms using nitrogen fertilizer increased by over 7%, and the total fertilized acreage in these counties increased by over 13%, in counties with merging congregations, relative to those without. These increases were most pronounced on the region's major commercial crop: counties with mergers used 26% more fertilizer on corn. We perform a randomization inference test and a placebo exercise to demonstrate that our results are caused by congregational mergers and not other factors.

Our results are consistent with a model where information sharing is the primary mech-

[4]In a notable exception, Satyanath et al. (2017) show that social clubs were a determinant of Nazi party participation in prewar Germany.

[5]See Feder et al. (1985) and Evenson and Westphal (1994) for reviews. Griliches (1957) is the seminal paper in this area.

[6]See Foster and Rosenzweig (2010) for a review.

[7]Reviews of the economic impacts of religion can be found in Welch and Mueller (2001) and Jackson and Fleischer (2007)).

anism through which social networks facilitate technology adoption. Mergers only affected use of fertilizer, a new technology, and its complements. In contrast, congregational mergers did not lead to increases in the use of existing technologies. We find no effects of mergers on durable goods with high fixed costs, suggesting that mergers did not ease capital constraints.

The remainder of this paper is organized as follows: Section 2.2 describes the context in more detail. Section 2.3 presents a simple model of social networks and technology adoption. Section 2.4 details our data, and Section 2.5 describes our empirical strategy. Section 2.6 reports our results. Section 2.7 provides a discussion. Section 2.8 concludes.

2.2 Context

We study the effects of social networks on the adoption of a new technology in the Upper Midwest of the United States during the 1950s and 1960s: commercial fertilizer.[8] Between 1940 and 1970, the use of commercial fertilizer increased dramatically. Figure 2.1 displays the sharp increase in usage of chemical fertilizer for corn production in the United States.

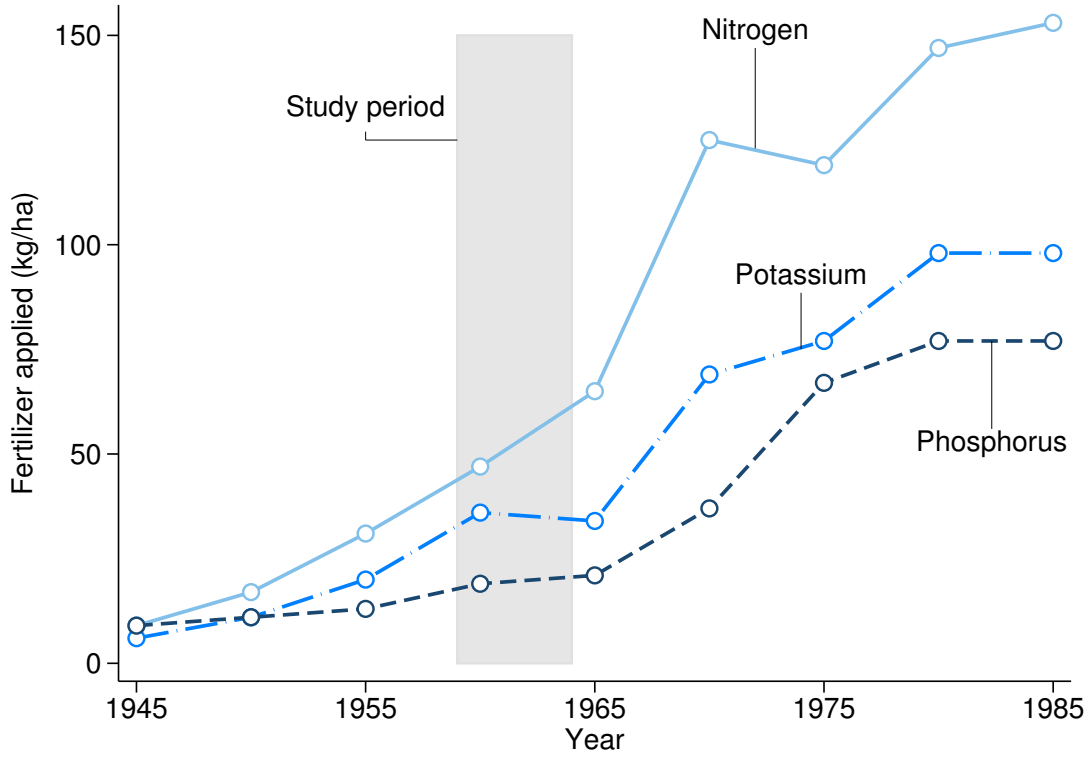
Between 1940 and 1949, average annual consumption of commercial fertilizer in the United States was 13.6 million tons; between 1950 and 1959, this number rose to 22.3 million tons; and between 1960 and 1969, use had increased further to 32.4 million tons (Campbell et al. (2004)).[9] This increase in usage had tangible results: between 1950 and 1975, agricultural productivity in the United States increased faster than ever before or since (Trautmann et al. (1998)). In 1950, the average American farmer supplied the materials to feed and clothe 14 people; by 1960, he was sustaining 26 (Rogers (1995)).

While today, over 95 percent of corn acres are fertilized (U.S. Department of Agriculture (2016)), and fertilizer is well-known to increase yields, during the 1950s and 1960s, farmers were far from being fully informed about optimal fertilizer usage and its benefits. Communication between farmers in different social circles was infrequent (Salamon (1992); Amato and Amato (2000); Cotter and Jackson (2001)), but information sharing within farmers' social networks was a major means of spreading professional knowledge. Religion was an important driver of farmers' social connections (Lazerwitz (1961); Azzi and Ehrenberg (1975); Swierenga (1997); Cotter and Jackson (2001)). The Upper Midwest had a high rate of religious adherence: according the Association of Religion Data Archives, in 1952, 64%, 62%, and 58% of the population of Minnesota, North Dakota, and South Dakota, respectively, were religious. We focus on these three states, because they contained large Lutheran populations: 51%, 48%, and 33% of religious Minnesotans, North Dakotans, and South Dakotans belonged to a Lutheran church. Figure 2.2 demonstrates the prevalence of religion in the United States in the 1950s, as well as the concentration of Lutheranism in Minnesota, North Dakota, and South Dakota.

[8] While farmers have been using organic fertilizer for thousands of years, commercial fertilizer as we know it today has its origins in the early twentieth century (Smil (2000)).

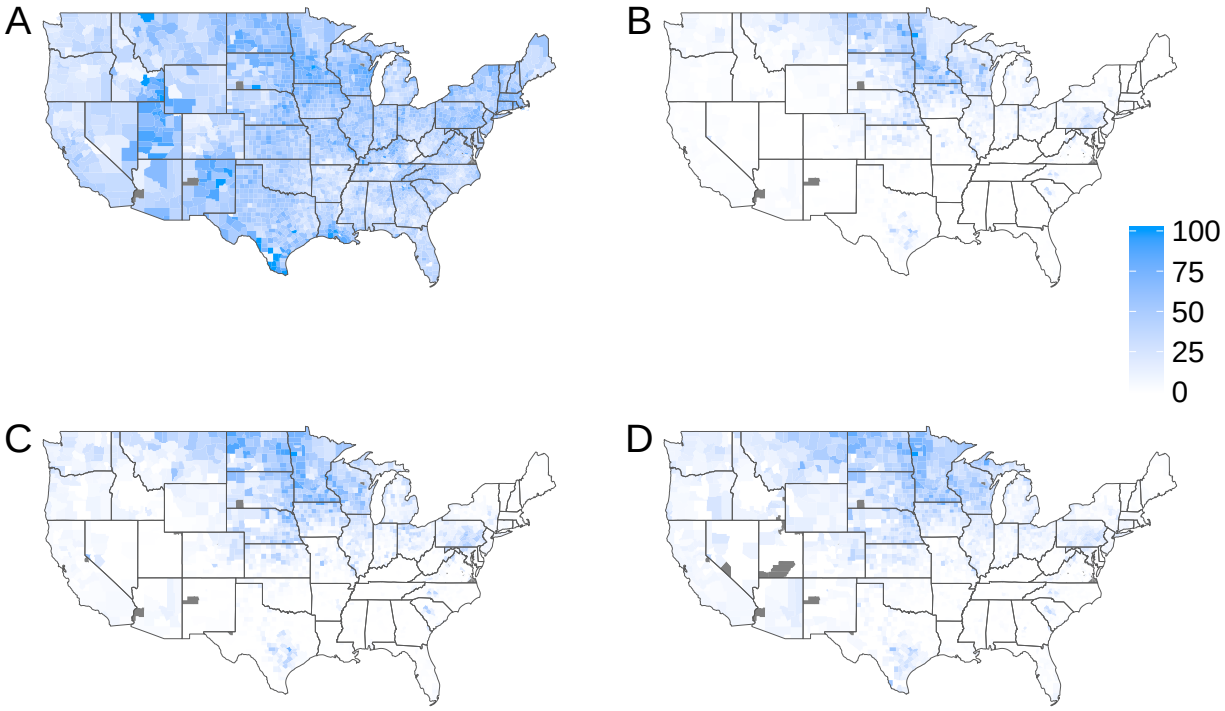
[9] Throughout this period, agricultural research centers were continually testing new types of ferilizer and best practices around their use (University of Minnesota Extension (1960)).

Figure 2.1: Fertilizer Applied in US Corn Production



Notes: This figure shows the growth in commercial fertilizer use in corn production in the United States between 1945 and 1985. None of the major fertilizer chemicals, nitrogen, potassium, or phosphorus, had reached a usage steady state by the time of our study, shaded in gray. The sharpest increase in fertilizer use occurred between 1965 and 1970. This figure was modified from Pimentel (1992).

Figure 2.2: Spatial Distribution of Lutherans



Notes: This figure presents data on religious adherence in the United States in 1952. Panel A displays the fraction of the population in each county that belonged to a religious organization. Panel B shows the fraction of each county's population that belonged to a Lutheran church branch. Panel C shows the fraction of religious members in each county that belonged to a Lutheran church branch. Panel D shows the fraction of churches or religious groups in each county that belonged to a Lutheran church branch. Minnesota, North Dakota, and South Dakota stand out as the most Lutheran states. Gray counties have missing data. County population data comes from the 1950 Census; religion data are from the Association of Religion Data Archives' 1952 county-level Churches and Church Membership in the United States data file (The Association of Religion Data Archives (2006)).

2.2.1 Church and congregational mergers

In the 1950s and 1960s, national Lutheran church bodies underwent significant institutional consolidation. At an April 1960 meeting in Minneapolis, Minnesota, three of the largest national Lutheran church bodies – the American Lutheran Church (ALC), the United Evangelical Lutheran Church (UELC), and the Evangelical Lutheran Church (ELC) – voted to merge and form The American Lutheran Church (TALC). This merger officially took effect on January 1, 1961.

A similar merger between the United Lutheran Church in America, the Finnish Evangelical Lutheran Church of America, the American Evangelical Lutheran Church, and the Augustana Evangelical Lutheran Church created the Lutheran Church in America (LCA) in 1962. In 1963, the Lutheran Free Church (LFC), composed largely of congregations that originally opted out of the 1960 TALC merger on theological grounds, decided to join TALC as well, extending the scope of this major Lutheran branch (Wolf (1966)).[10]

Figure 2.3 depicts the major mergers between Lutheran church bodies in the United States since the 1950s. For historical context, we focus primarily on TALC for two reasons. First, congregations of TALC were geographically clustered in the upper midwest whereas congregations of the LCA were more disperse throughout the country. Second, we have access to yearbooks from TALC detailing congregational-level statistics throughout the 1960s.

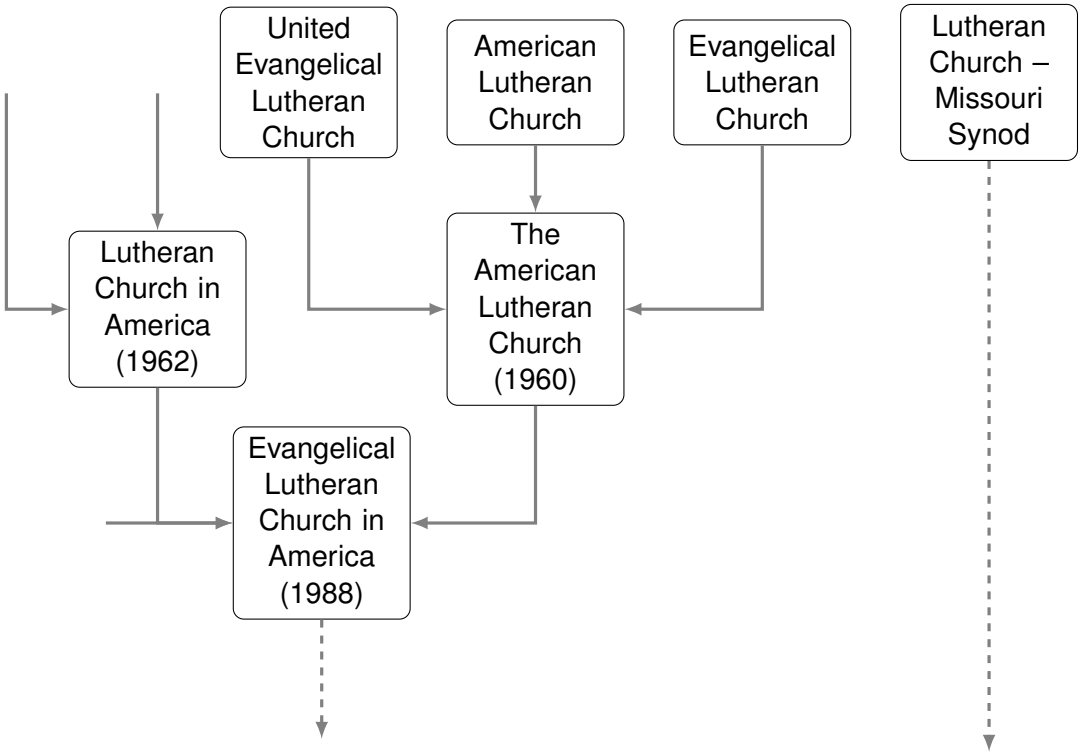
National-level mergers, arranged by the constituent churches' theological and institutional leadership, had far-reaching impacts. The TALC merger was reported in local newspapers across the Upper Midwest (Johnston (1960); Dugan (1960); Press (1960)). National mergers forced local congregations to adopt new constitutions, bringing them into alignment with the newly-formed national church (Nelson (1975)). Prior to the mergers, many towns had congregations from multiple church branches. As a result of the merger, these congregations suddenly found themselves in the same national denomination. This frequently led to mergers between local congregations that were previously impossible (Trinity Lutheran Church (2012); United Lutheran Church Laurel (2013)). These mergers brought previously socially disparate groups of people into contact with one another.

Each of the merging national-level church bodies (and their associated congregations) were linked to a different ethnic group: the ALC had German roots, the ELC had a Norwegian background, and the UELC was historically Danish (St. John Evangelical Lutheran Congregation (2014)). Especially in the early parts of the twentieth century, this often meant that congregations across the street from one another were holding services in different languages. Some congregations were even conducting multiple services, each in a different language (Bethel Lutheran Northfield Church (2014); Murray County (2014)).[11] Cross-branch mergers between local congregations were large shocks to churchgoers' social networks, since the congregants were not likely to have interacted frequently prior to the merger.

[10] TALC and the LCA later merged to found what is today the largest Lutheran church in the United States, the Evangelical Lutheran Church in America (ELCA).

[11] Over time, churches of all backgrounds conducted their services more and more frequently in English.

Figure 2.3: Lutheran Church Mergers



Notes: Each box represents a national-level Lutheran church branch. The mergers creating in The American Lutheran Church and the Lutheran Church in America in 1960 and 1962, respectively, prompted congregational-level mergers in the early 1960s.

In addition to the local mergers that were precipitated by national church changes, a number of congregational mergers resulted from other plausibly random events. Several congregations initiated mergers after natural disasters destroyed congregation buildings (Bethlehem Lutheran Church (2014); St. Mark's Lutheran Church (2014)). Other congregations merged due to difficulties hiring full-time clergy. Pastors, trained in centralized seminary programs, were a scarce and expensive resource, occasionally serving multiple congregations at once. Pastors in these roles frequently pressured their congregations to consolidate resources and merge into a single entity (Thoreson (2013); Grace Lutheran Church (2014)).[12] Importantly, these congregational mergers occurred for reasons that were unrelated to agricultural fertilizer use. Indeed, we show later in Table 2.1 that population trends are uncorrelated with congregational mergers.

[12]Congregations' steady, if begrudging, adoption of English as the primary or only language in church was also an impetus for congregational mergers (Lagerquist (1999)).

2.3 Theoretical framework

In this section, we develop a model of social networks and technology adoption to produce testable hypotheses and inform our subsequent empirical analysis.

Consider a network of n individuals indexed by i, each of whom is a potential adopter of some new technology. Every individual has a trait, ϕ_i, that determines whether or not it is optimal for them to adopt the new technology. For instance, if the new technology is fertilizer, then ϕ_i might describe land quality. Let $\bar{\phi}$ be a cutoff point: it is optimal for individuals with $\phi_i > \bar{\phi}$ to adopt the new technology, and optimal for individuals with $\phi_i < \bar{\phi}$ not to do so. Each individual knows their own ϕ_i, but does not know $\bar{\phi}$.

In addition, each individual has some trait θ_i which measures their apprehension around trying new technologies. Individuals with low levels of θ_i are relatively willing to experiment with new technologies, while individuals with high levels of θ_i are more cautious.

If an individual decides to "try" a new technology in some period t, he immediately observes whether $\phi_i < \bar{\phi}$ or $\phi_i > \bar{\phi}$. Let $A(t)$ be the number of individuals who have adopted the new technology by period t (i.e. who know in period t that $\phi_i > \bar{\phi}$).

In every period t, each individual in the network sends a signal $s_i(t)$ about whether they have previously tried the new technology. Individuals who have tried the technology in a previous period send a signal $s_i(t) = 1$, while individuals who have not tried the technology send a signal $s_i(t) = 0$. Define $S(t) = \sum_i s_i(t)$ to the total positive signal in the network at time t. Each individual i has a cut-off value $T_i(\theta_i)$ that determines if they will try the new technology in the subsequent period, such that $dT/d\theta_i > 0$ (i.e. more apprehensive individuals require more signals before trying a new technology). If $S(t) > T_i$, individual i will try the technology in period t.

We assume that the joint distribution of ϕ and θ is continuous. Additionally, we assume that the network is dense in θ-space. Specifically, let δ be the maximum distance between two adjacent individuals in θ-space: $\delta = \max_i\{\min_{j\neq i} |\theta_i - \theta_j|\}$. We assume that $T(\theta_i + \delta) \leq T(\theta_i) + 1$ for all i, such that every individual is "close" enough to enough other signals that they will try the new technology if all other nearby individuals send a signal. Additionally, for notational convenience, assume that $0 < \min_i\{T(\theta_i)\} < 1$. Finally, to initiate the model, we assume that in an initial period $t = 0$ there exists at least one individual i who has tried the new technology; that is, we assume $S(0) > 0$.

Proposition 1. *As long as there remain individuals who have not yet tried the technology ($S(t) < n$), the number of individuals adopting the new technology will weakly grow over time: $A(t+1) \geq A(t)$. Furthermore, in the long run, everyone will know whether or not it is optimal for them to adopt the technology.*

Proof. By induction. Consider the base case where $S(t) = 1$. By assumption, there exists at least one individual i s.t. $T(\theta_i) < 1 = S(t)$. Therefore, $s_i(t+1) = 1$ and $S(t+1) \geq 2$. If $\phi_i > \bar{\phi}$, then i adopts the technology and $A(t+1) > A(t)$. If $\phi_i < \bar{\phi}$, i does not adopt the technology, and $A(t+1) = A(t)$. Now consider the inductive step. Let $S(t) = l$. By the

definition of δ, there exists an individual j such that $l-1 < T(\theta_j) < l$. Therefore, $s_j(t) = 0$ and $s_j(t+1) = 1$, implying that $S(t+1) \geq l+1$. If $\phi_j > \bar{\phi}$, then j adopts the technology and $A(t+1) > A(t)$. If $\phi_j < \bar{\phi}$, j does not adopt the technology and it is possible that $A(t+1) = A(t)$. Finally consider the terminal case. If $S(t) = n$, all individuals have tried the technology and $A(t+1) = A(t)$. □

Proposition 2. *Consider two separate networks as described above, with populations n and m both greater than zero, respectively. These networks behave according to the assumptions above. Now consider a network merger, such that these two networks combine to form a larger network with population $n+m$. As long as not everyone has tried the new technology, adoption decisions will happen more quickly under the merger. That is: $A_n(t+1) - A_n(t) \leq A_{n+m}(t+1) - A_{n+m}(t)$.*

Proof. Let $S_n(t)$ denote $S(t)$ for the network of size n, $S_m(t)$ denote $S(t)$ for the network of size n, and $S_{n+m}(t)$ denote $S(t)$ for the merged networks. Suppose the two networks merge in period t. By construction, $S_{n+m}(t) > S_n(t)$. By the definition of δ, there exists some individual k such that $S_n(t) < T(\theta_k) < S_{n+m}(t)$, since $S_{n+m}(t) - S_n(t) \geq 1$. If $\phi_k > \bar{\phi}$, then k will adopt in $t+1$ if a merger occurs in time t, but will not adopt in the absence of a merger, because he has not yet observed enough signals. □

Proposition 3. *Network mergers will not change the adoption of technologies that all individuals have already tried, in other words, about which they are fully informed.*

Proof. If all individuals are fully informed, $S_n(t) = n$ and $S_m(t) = m$ in networks of size n and m. Therefore, there remain no individuals with $s_i = 0$, and no individuals with $\phi_i > \bar{\phi}$ who have not yet adopted. This implies $A_{n+m}(t+1) = A_n(t+1) + A_m(t+1)$. □

2.4 Data

We use detailed data on churches and agriculture to estimate the effects of social network expansions on technology adoption. In particular, we use data from 197 counties in Minnesota, North Dakota, and South Dakota. We exclude the 9 counties that included a Standard Metropolitan Statistical Area, according to the 1960 Census, in order to restrict our sample to agricultural regions.

2.4.1 Church data

We acquired a unique dataset on mergers between Lutheran congregations based on the archives of the ELCA. Church archivists compiled a database of mergers between Lutheran congregations, the earliest of which occurred in 1810.[13] These data are derived from two main sources: annual national church yearbooks, which in turn were compiled from reports

[13]The most recent mergers in this dataset occurred in 2012, the year we got access to the data.

congregations made to their governing bodies; and The American Lutheran Church (1960), a record of all of the details surrounding the TALC merger. This dataset includes all mergers between congregations of the UELC, ALC, and ELC, which merged to form TALC in 1960, including mergers between congregations which no longer exist. For each congregation involved in a merger, the dataset records the state, county, local post office, location[14], synod[15], congregation name, founding date, merger date, details on which other congregations were involved in the merger, and additional historical notes.

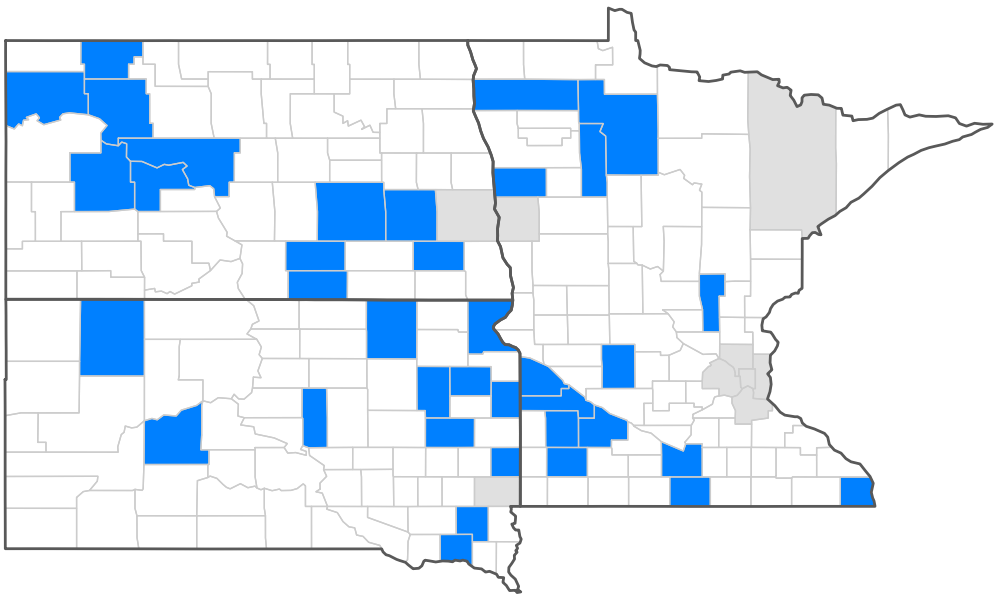
In our main analysis, we use the mergers that took place between 1959 and 1964.[16] A total of 41 mergers occurred during this time period in Minnesota, North Dakota, and South Dakota. In our main study period, between 1959 and 1964, 34 counties experienced one merger, two counties experienced two mergers, and one county experienced three mergers. We consider counties that experienced at least one merger between 1959 and 1964 to be treated, and use counties that did not experience a merger during this time as controls. Figure 2.4 displays the spatial distribution of treated and control counties. The counties in blue experienced at least one merger between 1959 and 1964; the counties in white did not; and the counties in gray included a major urban area and were excluded from the sample.

[14]Entries in the "location" field range from addresses to P.O. boxes to information such as "12SE," meaning 12 miles southeast of town. We match these data to the Census of Agriculture at the county level. Where possible, we cross-check the location information with the county. We find no major discrepancies.

[15]A synod is a Lutheran administrative region headed by a bishop. Lutheran synods are somewhat analogous to Catholic dioceses.

[16]We also use the 72 mergers that occurred between 1964 and 1969, after our main analysis period, to perform a placebo test.

Figure 2.4: Treated, control, and excluded counties



Notes: This figure displays our sample of counties within the states of Minnesota, North Dakota, and South Dakota. The 37 counties in blue experienced a congregational merger between 1959 and 1964. The 161 counties in white did not. The 9 counties in gray contained a major urban area, defined according to the 1960 Census, and were excluded from our analysis.

2.4.2 Agriculture data

We combine our data on merging congregations with data from the United States Department of Agriculture (USDA)'s Census of Agriculture. In the 1950s and 1960s, the Census was designed to have full coverage of every farm in the United States (U.S. Census Bureau (1967)). Censuses were taken every five years, and data gathering for these Censuses took place in the fall. Enumerators visited every dwelling, and administered the Census to any household engaged in agriculture. After collection, the Census underwent a multi-stage quality control process. The final dataset is available at the county level. Wherever possible, we use a digitized version of the dataset made available by the University of Michigan's Inter-university Consortium for Political and Social Research (Haines et al. (2015)). Several variables were unavailable in the digitized data; we hand-coded these from PDFs made available from the USDA's own archive.[17]

We use the 1954, 1959, and 1964 waves of the Census. Using earlier waves is impossible: the two earlier Censuses, taken in 1950 and 1944, did not include county-level information on fertilizer use. We are also unable to use later waves of data: after the 1964 wave, data

[17]These PDFs can be found here: `http://agcensus.mannlib.cornell.edu/AgCensus/`.

was only collected for farms selling over $2,500 worth of goods per year, and there is no way to reconcile the two sampling frames. We use the 1954 Census to test for differential trends among counties with and without congregational mergers. We perform our main analysis using the 1959 and 1964 Censues. We combine the Census of Agriculture data from these years with our congregational data to create a balanced panel of 197 counties.

The Census of Agriculture contains data on our main outcomes of interest: the number of farms using fertilizer, acres fertilized, tons of commercial fertilizer used, and corn acres fertilized, tons of dry and liquid fertilizer used on corn. It also contains information on the use of agricultural lime, a complement to nitrogen fertilizer. The Census also includes data on other agricultural practices, such as strip cropping and irrigation; other types of land use, such as orchards; and capital-intensive farm durables, including vehicles.

Table 2.1 presents summary statistics from the 1954 Census for counties that did and did not experience congregational mergers between 1959 and 1964. Treatment and control counties are statistically indistinguishable on most observable characteristics prior to the mergers.[18] The major exception is in harvested acreage: treatment counties harvested approximately 81,650 more acres than control counties in 1954. This difference is statistically significant at the one percent level. Treatment counties also harvested 14,200 more acres of corn relative to the control group, a difference which is statistically significant at the ten percent level.

Overall, these summary statistics reveal that treatment and control counties are relatively similar to one another prior to the congregational mergers that took place between 1959 and 1964. These statistics support the notion that mergers were not driven by the agricultural sector or other potentially endogenous factors.

[18]Not every variable that we use in our analysis is available in 1954.

Table 2.1: Pre-treatment Summary Statistics

Variable	Control	Treatment	Difference
Farms (nr.)	1,329.42	1,565.30	235.87
	(832.96)	(589.31)	[144.73]
Acres in county	679,127.45	756,808.65	77,681.20
	(389,333.41)	(410,248.03)	[71,696.29]
Acres in farms	565,570.96	661,454.41	95,883.44
	(334,078.67)	(363,495.21)	[61,926.92]
Acres harvested	268,707.05	350,355.78	81,648.73***
	(149,313.79)	(160,264.55)	[27,599.45]
Farms using fertilizer (nr.)	437.89	474.95	37.05
	(491.43)	(472.30)	[88.96]
Acres fertilized	29,932.11	30,689.14	757.02
	(46,217.63)	(41,836.45)	[8,285.16]
Fertilizer used (tons)	1,850.37	2,120.19	269.82
	(2,647.48)	(2,725.54)	[485.32]
Acres limed	957.80	503.54	−454.26
	(3,085.38)	(1,811.67)	[527.57]
Lime used (tons)	1,849.99	917.97	−932.01
	(6,515.08)	(3,502.70)	[1,107.52]
Acres harvested (corn)	48,872.45	63,074.46	14,202.01*
	(42,637.63)	(53,016.42)	[8,153.95]
% change in population (1950-1960)	−0.09	−3.95	−3.86
	(16.47)	(11.24)	[2.85]
Number of counties	161	37	

Notes: This table shows summary statistics in 1954 for counties that did not experience a congregational merger between 1959 and 1964 (control) and counties that did experience a merger during this time period (treatment). Standard deviations in parentheses; standard errors on the t-tests between control and treatment in brackets. Significance: *** $p < 0.01$, ** $p < 0,05$, * $p < 0.10$.

2.5 Empirical strategy

We employ a differences-in-differences approach to take the predictions of our model to data. We use data from 1959 and 1964 to estimate the following specification:

$$y_{cst} = \beta M_{cst} + \delta_t + \mu_{cs} + \gamma_{st} + \varepsilon_{cst} \tag{2.1}$$

where y_{cst} is an outcome of interest in county c belonging to state s in year t, M_{cst} is equal to one if county c experienced a merger between 1959 and 1964, δ_t is a time fixed effect, μ_{cs} are county fixed effects, γ_{st} are state-by-year fixed effects, and ε_{cst} is an idiosyncratic error term. The coefficient of interest is β, which captures the effect of congregational mergers on outcome y. We cluster our standard errors at the county level in order to account for arbitrary error dependence between observations in the same county.

Proposition 2 implies that congregational mergers should increase the adoption rate of new technologies. In our empirical context, this suggests that mergers should lead to increased adoption of commercial fertilizer. We test for adoption along the extensive margin by estimating Equation 2.1 with the number of farms using fertilizer as the outcome of interest. We also test for effects of mergers on the number of acres fertilized and tons of fertilizer applied, which capture both intensive and extensive margin effects.[19]

Though not directly predicted by our model, we expect the use of agricultural lime to increase with congregational mergers as well. Nitrogen, the primary component of commercial fertilizer, adds acidity to soil, which can impede crop growth. Agricultural lime helps to reverse this process, making it a natural complement to fertilizer use. We test for these effects by by estimating Equation 2.1 using the number of farms using lime, the number of acres limed, and tons of lime used as outcomes.

We also test for the effects of congregational mergers on fertilizer use on corn, which benefits greatly from the use of fertilizer (Barber and Stivers (1962)), and is one of the region's major commercial crops.[20] We expect congregational mergers to increase total fertilizer use on corn. The Census of Agriculture distinguishes between dry and liquid fertilizer used on corn. We expect to find stronger effects on dry fertilizer, since the major technological advances of the time period occurred in dry, rather than liquid, fertilizers (Russel and Williams (1977); Young and Hargett (1984)).

Proposition 3 implies that congregational mergers will not affect adoption of technologies that all farmers are already informed about. In the Upper Midwest in the 1950s, we can test this theory using three well-established technologies: strip cropping, irrigation, and orchards. Strip cropping, in which farmers alternate crop types in tight rows to prevent soil erosion, is an established practice in the United States. It was introduced to Minnesota in the early 1930s (Helms et al. (1996)), and was in widespread use in the region by 1940 (Granger and Kelly (2005)). Irrigation was another well-known technology: the most common irrigation

[19] Note that the model described in Section 2.3 could also apply to the intensive margin of technology adoption; simply consider the "technology" to be a level of fertilizer use.

[20] Nitrogen is the commercial fertilizer most heavily used in corn production (Pimentel (1992)).

Table 2.2: Test of Differential Trends

Variables	(1) Fertilizer: Farms	(2) Fertilizer: Acres	(3) Lime: Farms	(4) Lime: Acres
Year = 1959 × merger	20.59	8,357.89	1.71	−116.01
	(26.13)	(5,590.86)	(9.62)	(84.34)
Mean of dependent variable	488.30	43,790.71	52.57	1,013.40
R^2	0.55	0.64	0.13	0.06
Observations	396	396	396	396
Number of counties	198	198	198	198
County FE	YES	YES	YES	YES
State-by-year FE	YES	YES	YES	YES

Notes: This table shows results from estimating Equation (2.1), using data from 1954 and 1959. Year = 1959 × merger is equal to one if the year is 1959, and the county experienced a congregational merger between 1959 and 1964. The dependent variables are, in column (1), the number of farms reporting fertilizer use; in column (2) the number of acres fertilized; in column (3) the number of farms reporting lime use; and in column (4) the number of acres limed. Standard errors are clustered at the county level. Significance: *** $p < 0.01$, ** $p < 0,05$, * $p < 0.10$.

system in use in this area was the center pivot system, which had spread to farmers by the late 1950s (Kenney (1995); Granger and Kelly (2005)).

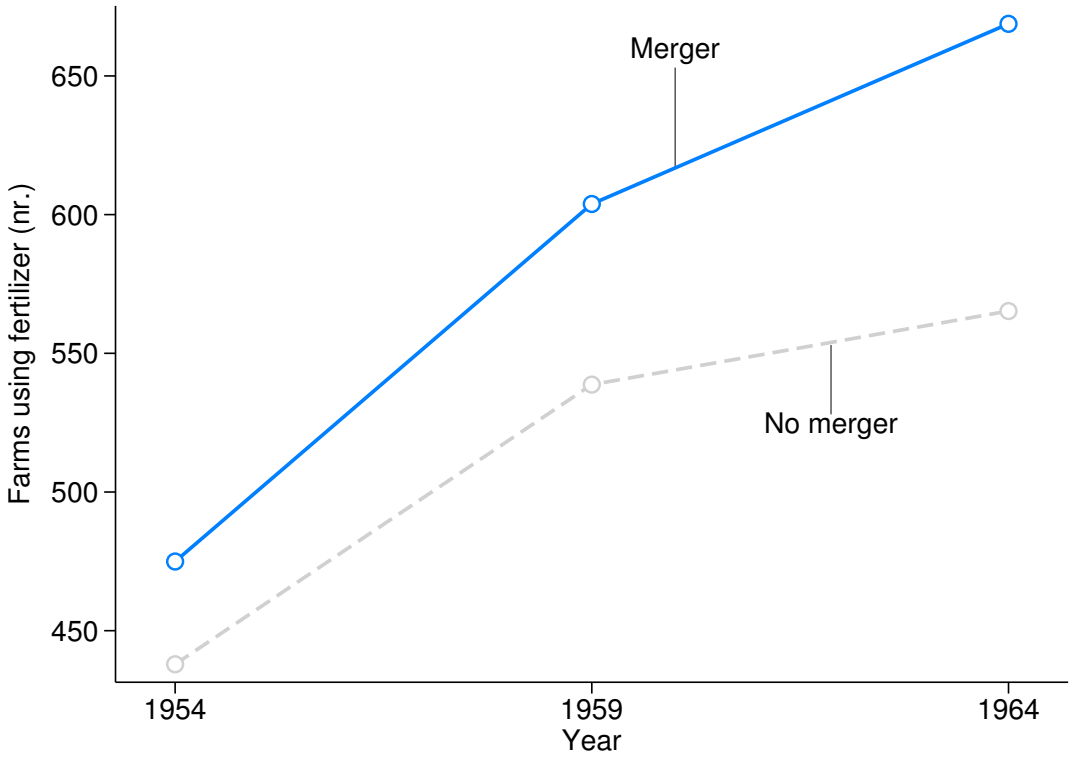
Finally, using land for orchards, vineyards, groves, and nut trees was a well-established practice in the Midwest by the 1950s (Gordon (1997); Burrows (2010); Smith (2011). Our model predicts that, as farmers were informed about them prior to our study period, strip cropping, irrigation, and orchard lands should not respond to congregational mergers.

Our difference-in-differences approach relies on the identifying assumption that $\mathbb{E}[\varepsilon_{cst}|M_{cst}, \mu_{cs}, \delta_t, \gamma_{st}) = 0$, or that there are no time-varying unobserved factors that are different between counties with and without mergers. We believe that this assumption is reasonable: as discussed in Section 2.2, exogenous factors, including national-level church branch mergers and building fires, caused the congregational mergers we study. While we fundamentally cannot empirically test our identifying assumption, we provide evidence in support of it in two ways. First, in Table 2.2, we estimate Equation 2.1 using data from 1954 and 1959, for four of our main outcomes of interest: the number of farms using fertilizer, acres fertilized, the number of farms using lime, and acres limed. Our "post" period, 1959, is before our congregational mergers, so we should expect to find no statistically significant effects of mergers on our outcomes of interest.

In all cases, we fail to reject the null hypothesis that counties with mergers are trending similarly to counties without mergers, prior to treatment. Figure 2.5 demonstrates this graphically, showing that, from 1959 to 1959, counties that experienced mergers were on a similar path to counties that did not. It was only after 1959, when our mergers occurred,

that the groups of counties began to diverge.

Figure 2.5: Parallel Pre-Trends in Farms Using Fertilizer



Notes: This figure shows the mean number of farms using fertilizer in 1954, 1959, and 1964 in the 37 counties that experienced congregational mergers between 1959 and 1964 and the 161 that did not. These two types of counties were trending similarly prior to these mergers. In Table 2.2, we estimate the difference in slopes in the pre-period. This difference is not statistically significant at conventional levels. Between 1959 and 1964, however, counties that experienced mergers began to diverge sharply from those that did not.

This evidence suggests that our empirical approach is valid. There may also be concerns about the Stable Unit Treatment Value Assumption, that is, that the treatment status of county c will not affect the outcome in any other county. We believe that this assumption is plausible in our empirical context, since none of the mergers in our data cross county boundaries. Furthermore, any spillovers from merging to non-merging counties are likely to be positive, which would attenuate our treatment effects. Our estimates are therefore lower bounds on the true effects of networks on technology adoption.

2.6 Results

We begin by testing Proposition 2. We first estimate the effects of congregational mergers on fertilizer use on the extensive margin, using the number of farms using fertilizer as the dependent variable. We estimate five specifications, each with a different set of controls. Table 2.3 reports the results. Column (1) is the most parsimonious specification, including only the interaction term of interest (Year = 1964 $\times$ merger), a 1964 dummy, and a "merger county" dummy. In column (2), we replace the "merger county" dummy with county fixed effects. Column (3) adds four weather controls: temperature, precipitation, heating degree days, and cooling degree days.[21] In order to control for time-varying unobservables, we also include state-by-year fixed effects in column (4). This is our preferred specification. In column (5), we include both state-by-year fixed effects and weather controls, though given our small sample, we expect this specification to be underpowered.

The results in Table 2.3 are consistent with Proposition 2: as expected, counties that experienced congregational mergers see higher rates of fertilizer adoption than those that did not. These effects are economically meaningful, and appear to be relatively consistent across specifications.[22]

Using our preferred specification, displayed in Column (4), we find that congregational mergers caused 40.07 additional farms per county to begin using fertilizer, a large increase of 7.3 percent over the mean in the control group. Our results are statistically significant at the 10 percent level, which, given our relatively small sample size, is encouraging.

We present the results from our main specification graphically in Figure 2.6. The dashed grey line is the kernel density of the change in the number of farms using fertilizer between 1959 and 1964 for counties that did not experience congregational mergers. The solid blue line is the same change for counties that did experience mergers.

The changes in the control distribution are centered around zero. In contrast, the treated distribution lies markedly to the right of the control distribution. This shift appears to be present throughout the distribution.

In order to ensure that these effects are not spurious, and instead result from congregational mergers, we implement a randomization inference procedure. We randomly reassign exactly 37 counties to treatment 10,000 times. For each run, we estimate every specification in Table 2.3, and store the estimated $\hat{\beta}$. We display the results of this procedure in Figure 2.7. The gray histograms show coefficients from these 10,000 random draws, and the blue lines denote the treatment effect using the real assignment vector. In each case, the real effect lies in the far right portion of the distribution - and in our preferred specification,

[21] We report estimates including weather controls for completeness, but weather controls and state-by-year fixed effects are removing similar variation in this context. Specifications including weather data use average annual precipitation, average annual temperature, heating degree days, and cooling degree days, aggregated to the climatic division level, from NOAA (NOAA (2007)).

[22] Column (5) is the exception, though we cannot reject that the estimates in Columns (1) through (4) are statistically different from this result.

Table 2.3: Impact of Congregational Mergers on Fertilizer Use – Farms

Variables	(1)	(2)	(3)	(4)	(5)
Year = 1964 × merger	38.42*	38.42*	39.66*	40.07*	33.00
	(22.83)	(22.80)	(20.48)	(21.35)	(21.15)
Year = 1964	26.53	26.53			
	(9.60)	(9.59)			
Merger (1/0)	65.12				
	(87.48)				
Mean of dependent variable	551.98	551.98	551.98	551.98	551.98
R^2	0.01	0.08	0.26	0.21	0.28
Observations	396	396	396	396	396
Number of counties	198	198	198	198	198
County FE	NO	YES	YES	YES	YES
Year FE	NO	NO	YES	YES	YES
State-by-year FE	NO	NO	NO	YES	YES
Weather controls	NO	NO	YES	NO	YES

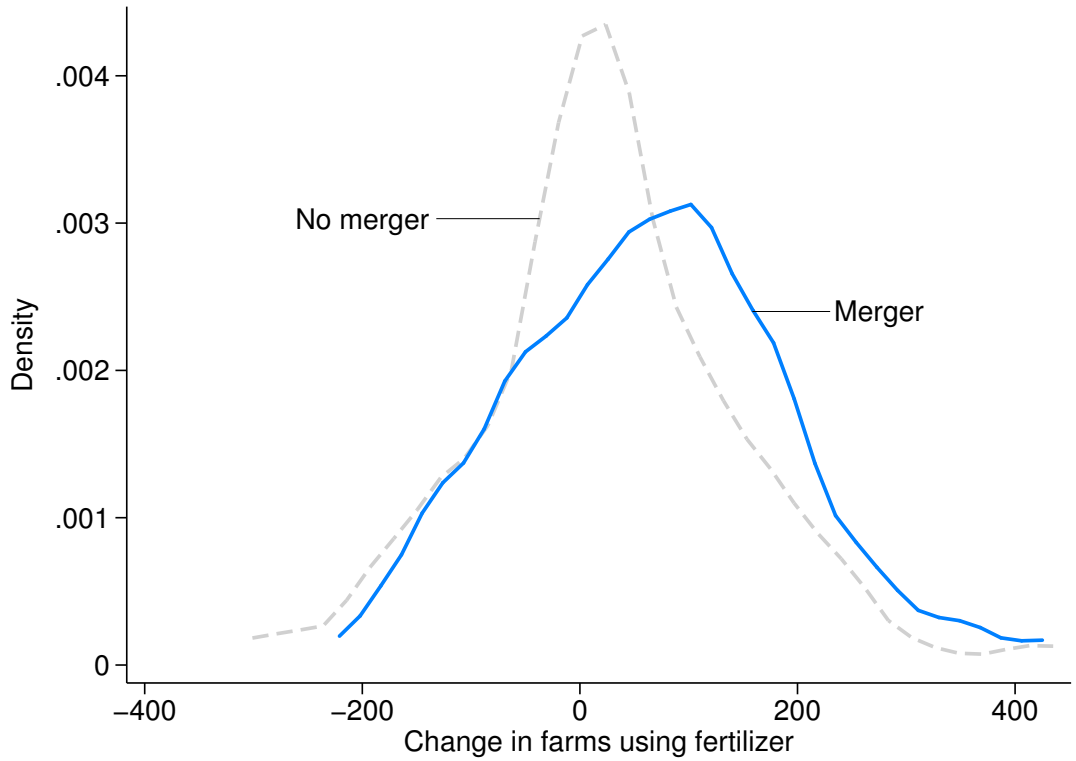
Notes: This table shows results from estimating Equation (2.1). The dependent variable is the number of farms in a county reporting fertilizer use. Year = 1964 × merger is equal to one if the year is 1964, and the county experienced a congregational merger between 1959 and 1964. Note that state-by-year fixed effects nest year fixed effects. Weather controls include temperature (° F), precipitation (in), heating degree days, and cooling degree days. Standard errors are clustered at the county level. Significance: *** $p < 0.01$, ** $p < 0,05$, * $p < 0.10$.

lies above the 96th percentile - which suggests that our results are not an artifact of random chance.

We next estimate Equation 2.1, our preferred specification, using acres fertilized and tonnage of fertilizer applied as dependent variables. We also test for effects on the number of farms using agricultural lime, acres limed, and the tons of lime applied. Since lime is a complement to fertilizer, we expect to find positive effects of congregational mergers on lime use. Table 2.4 presents these results.

Table 2.4 shows that, as expected, congregational mergers increased acres fertilized. Counties with mergers fertilized, on average, 8,370.5 acres more than counties without mergers, a 13.6 percent increase over the control group mean, and statistically significant at the 5 percent level. We do not find a corresponding increase in the tonnage of fertilizer applied on all crops, though this is likely to be driven in part by noise involved in measuring tonnage of fertilizer used. We do find the expected positive effects of congregational mergers on the number of farms using lime: 9.2 additional farms use lime in the treatment group relative to the control group, a large increase of 21.1 percent, statistically significant at the 5 percent level. We find corresponding increases in the number of acres limed and the tons of lime

Figure 2.6: Change in farms using fertilizer, 1959 to 1964



Notes: This figure plots kernel densities of the difference in the number of farms using fertilized between 1959 and 1964 for counties that experienced a congregational merger and those that did not. These densities were generated using an Epanechnikov kernel. The mean (SD) of the no-merger distribution is 26.53 (121.58), and the mean (SD) of the merger distribution is 64.95 (126.90).

used, with acreage limed increasing by nearly 24 percent; and tons of lime used increasing by close to 22 percent. These effects are statistically significant at the 5 and 10 percent level, respectively. Taken together, these results suggest that congregational mergers led to an economically meaningful and statistically significant increase in fertilizer and lime use, as predicted by our model.

Next, we look at corn. In the Census of Agriculture data, there is information about tonnage of both wet and dry commercial fertilizer applied for corn, so we will be able to separate the impact on the different types of inputs. In addition, there is information on fertilized acreage. Table 2.5 displays the results. In column (1), the dependent variable is corn acres fertilized; in column (2), the dependent variable is tons of dry commercial fertilizer used; column (3) looks at tons of liquid commercial fertilizer used, and column (4) looks at the total tonnage of commercial fertilizer used. We expect to see most of the positive effect on dry, rather than wet, tons.

Table 2.4: Impact of Congregational Mergers on Fertilizer Use – Fertilizer and Lime

Variables	(1) Fertilizer: Acres	(2) Fertilizer: Tons	(3) Lime: Farms	(4) Lime: Acres	(5) Lime: Tons
Year = 1964 × merger	8,370.50**	−558.35	9.20**	238.53**	487.72*
	(3,924.61)	(566.55)	(4.23)	(120.60)	(269.97)
Mean of dependent variable	61,656.96	1,779.68	43.60	1,005.12	2,323.20
R^2	0.24	0.54	0.07	0.04	0.03
Observations	396	396	396	396	396
Number of counties	198	198	198	198	198
County FE	YES	YES	YES	YES	YES
State-by-year FE	YES	YES	YES	YES	YES

Notes: This table shows results from estimating Equation (2.1). Year = 1964 × merger is equal to one if the year is 1964, and the county experienced a congregational merger between 1959 and 1964. The dependent variables are, in column (1), acres fertilized; in column (2), tons of commercial fertilizer used; in column (3), the number of farms reporting the use of lime; in column (4), acres limed; and in column (5), tons of lime used. Standard errors are clustered at the county level. Significance: *** $p < 0.01$, ** $p < 0,05$, * $p < 0.10$.

Table 2.5: Impact of Congregational Mergers on Fertilizer Use – Corn

Variables	(1) Acres	(2) Tons: Dry	(3) Tons: Liquid	(4) Tons: Total
Year = 1964 × merger	5,300.54***	391.41**	93.04	484.45***
	(1,962.03)	(163.65)	(72.56)	(183.72)
Mean of dependent variable	21,909.52	1,584.09	240.40	1,824.49
R^2	0.17	0.24	0.19	0.30
Observations	396	396	396	396
Number of counties	198	198	198	198
County FE	YES	YES	YES	YES
State-by-year FE	YES	YES	YES	YES

Notes: This table shows results from estimating Equation (2.1). Year = 1964 × merger is equal to one if the year is 1964, and the county experienced a congregational merger between 1959 and 1964. The dependent variables are, in column (1), corn acres fertilized; in column (2), tons of dry commercial fertilizer used on corn; in column (3), tons of liquid commercial fertilizer used on corn; and in column (4), the total tonnage of commercial fertilizer used on corn. Standard errors are clustered at the county level. Significance: *** $p < 0.01$, ** $p < 0,05$, * $p < 0.10$.

Table 2.6: Impact of Congregational Mergers on Land Use and Irrigation

Variables	(1) Strip: Farms	(2) Strip: Acres	(3) Irrigation: Farms	(4) Irrigation: Acres	(5) Orchard: Acres
Year = 1964 × merger	6.97	4,360.83	−1.48	−100.63	3.57
	(6.32)	(3,467.61)	(0.91)	(95.39)	(5.33)
Mean of dependent variable	100.43	18,991.61	8.89	974.45	24.11
R^2	0.14	0.18	0.03	0.06	0.12
Observations	396	396	396	396	396
Number of counties	198	198	198	198	198
County FE	YES	YES	YES	YES	YES
State-by-year FE	YES	YES	YES	YES	YES

Notes: This table shows results from estimating Equation (2.1). Year = 1964 × merger is equal to one if the year is 1964, and the county experienced a congregational merger between 1959 and 1964. The dependent variables are, in column (1), farms reporting the use of strip cropping; in column (2), acres strip cropped; in column (3), farms reporting the use of irrigation; in column (4), acres irrigated; and in column (5), acres in fruit orchards, groves, vineyards, and nut trees. Standard errors are clustered at the county level. Significance: *** $p < 0.01$, ** $p < 0,05$, * $p < 0.10$.

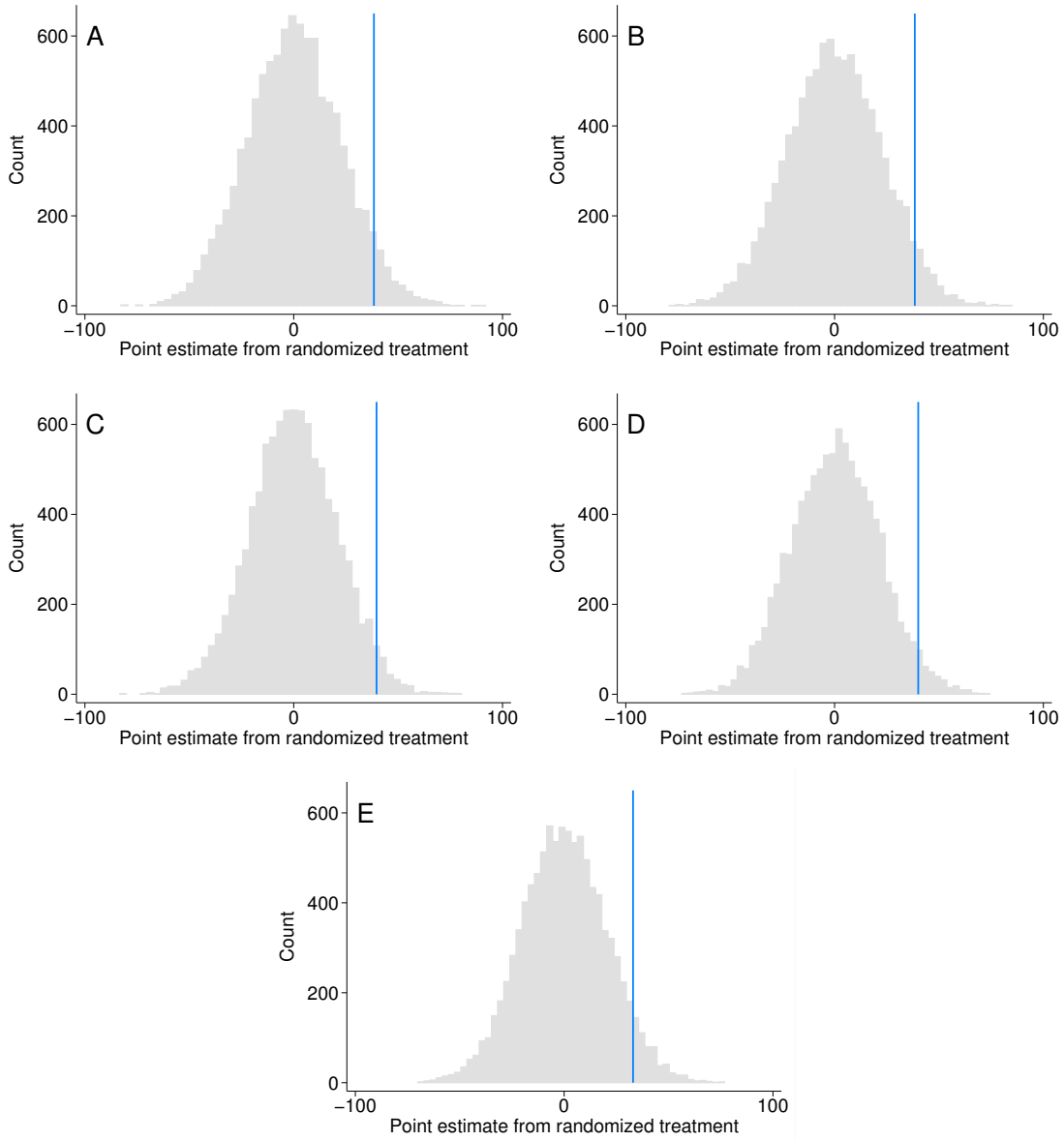
The results from Table 2.5 are in line with Proposition 2, suggesting that fertilizer used on corn increases as a result of congregational mergers. Acreage fertilized increases by 5,300.54, a change of 24.2 percent. This is statistically significant at the 1 percent level. Columns (2), (3), and (4) demonstrate that there is an increase in tonnage of fertilizer used on corn, and that this increase is driven by dry fertilizer use. We find an increase of 391.41 tons of dry fertilizer, statistically significant at the 5 percent level, which represents a 24.7 percent increase over the mean; and no statistically significant increase in the tonnage of wet fertilizer applied. The total tonnage of fertilizer applied to corn increases by 484.45 tons, an 26.6 percent increase, statistically significant at the 1 percent level. The impact congregational mergers have on corn fertilizer use is not only statistically significant, but also represents an economically meaningful change. In sum, the information in Table 2.3, Table 2.4, and Table 2.5, demonstrates that congregational mergers increase fertilizer (and lime) use, in keeping with Proposition 2.

Next, we test Proposition 3: that congregational mergers will not affect technology adoption when all potential adopters are fully informed. Strip cropping, irrigation, and orchards are technologies for which we expect no effect. Table 2.6 demonstrates the impact congregational mergers have on the number of farms using strip cropping and the acres under strip cropping; the number of farms reporting irrigation use and the total number of irrigated acres; and the total number of acres in fruit orchards, groves, vineyards, and nut trees.

Table 2.6 confirms our hypothesis: we see no statistically significant impacts of congrega-

tional mergers on strip cropping, irrigation, or orchard acreage. Furthermore, the magnitudes of the coefficients we do see are small, and in the case of irrigation, have negative signs. The absence of results in this table further supports the fact that congregational mergers are driving the changes in input uses observed above, because mergers are affecting farmers on the dimensions we would expect, but we are *not* seeing impacts on agriculture that should not be affected by mergers.

Figure 2.7: Randomization Inference – Number of Farms Using Fertilizer



Notes: This figure displays the results of a randomization inference procedure the five regression specifications used to estimate the effects of congregational mergers on the number of farms using fertilizer in Table 2.3. Panel A includes the difference-in-difference treatment variable, a post-treatment dummy, and a merging county dummy on the right-hand side. Panel B includes the treatment variable, a post-treatment dummy, and county fixed effects. Panel C includes the treatment variable, county fixed effects, and state-by-year fixed effects. Panel D includes the treatment variable, a post-treatment dummy, county fixed effects, and weather controls. Panel E includes the treatment variable, county fixed effects, state-by-year fixed effects, and weather controls. These panels correspond to Columns (1), (2), (3), (4), and (5) of Table 2.3, respectively. For each specification, we randomly reassigned mergers to exactly 37 counties 10,000 times, re-estimated the appropriate specification, and saved the coefficient of interest. These coefficients are plotted in gray. The blue lines denote the $\hat{\beta}$ values estimated using the real merger assignment vector. These $\hat{\beta}$s are above the 95th, 95th, 97th, 96th, and 94th percentiles of the distributions, respectively, suggesting that it is highly unlikely that our results are spurious.

2.7 Discussion

With any regression analysis, it is important to ensure that the results are robust. We use three pieces of evidence to demonstrate that the results we present in this paper hold up to robustness checks. First, we saw impacts of congregational mergers where we expected them: on fertilizer and lime use and on corn. We did not see impacts where we expected them not to be: on strip cropping, irrigation, and orchards, as presented above. In addition, our results are robust to a variety of different land use variables provided in the census of agriculture data: the outcomes measured in number of farms do not change significantly if instead of total farms, we use commercial farms or cash-grain farms; the outcomes measured in acreage and tonnage do not change if we use acres in the county or acres harvested rather than acres in farms.[23] In this section of the paper, we perform a placebo test. We also investigate channels other than information through which congregational mergers might be driving fertilizer adoption, and provide evidence against these other possible explanations.

2.7.1 Placebo test

Table 2.7 has the placebo analysis. We run this check in order to see whether mergers that occurred from 1964 to 1967 impact our outcomes in 1964, before these mergers actually occurred. We use the time period 1964 to 1967 because it includes the same number of years as our actual treatment period. Column (1) looks at farms using fertilizer; (2) at acres fertilized; (3) at tons of fertilizer used; (4) at corn acres fertilized; (5) at farms using lime; and (6) at acres limed. We expect to see no statistically significant impacts of "future" congregational mergers on these outcomes.

Indeed, Table 2.7 shows that, with the exception of the number of farms using lime, there is no statistically significant effect of future congregational mergers on 1964 input outcomes. In addition, comparing these effects to those in Tables 2.3, 2.4, and 2.5, the magnitudes of the coefficients are quite small. This helps confirm that the effects we are observing above are real and driven by the congregational mergers we observe, rather than by something unobserved. To assuage concerns about small-sample inference, we also run a permutation test, in which we randomly assign 39 counties to treatment 10,000 times. Using each of these treatment assignment vectors as a placebo treatment, we find our actual effect on farms using fertilizer is larger than all but 4 percent of the randomly drawn treatment vectors. Figure 2.7 displays the results of this procedure.

[23] Regression results are available from the authors upon request.

Table 2.7: Impact of Congregational Mergers on Inputs – Placebo Tests

Variables	(1) Fertilizer: Farms	(2) Fertilizer: Acres	(3) Fertilizer: Tons	(4) Fertilizer: Acres (Corn)	(5) Lime: Farms	(6) Lime: Acres
Year = 1964× future merger	−4.40	4,686.36	−467.65	1,412.22	7.23*	153.86
	(19.74)	(3,440.47)	(756.61)	(1,954.61)	(3.93)	(130.37)
Mean of dependent variable	551.98	61,656.96	1,779.68	21,909.52	43.60	1,005.12
R^2	0.20	0.22	0.54	0.13	0.07	0.04
Observations	396	396	396	396	396	396
Number of counties	198	198	198	198	198	198
County FE	YES	YES	YES	YES	YES	YES
State-by-year FE	YES	YES	YES	YES	YES	YES

Notes: This table shows results from estimating Equation (2.1). Year = 1964 × future merger is equal to one if the year is 1964, and the county experienced a congregational merger between 1964 and 1969. Counties with mergers between 1959 and 1964 and between 1964 and 1969 are treated as controls. The dependent variables are, in column (1), farms reporting fertilizer use; in column (2), acres fertilized; in column (3), tons of commercial fertilizer used; in column (4), acres of corn fertilized; in column (5), farms reporting lime use; and in column (6), acres limed. Standard errors are clustered at the county level. Significance: *** $p < 0.01$, ** $p < 0,05$, * $p < 0.10$.

2.7.2 Alternative explanations

It is still possible that our results are being driven by something other than a congregational-merger driven information effect. Here, we explore two other possible explanations for our results. The first is the presence of agricultural extension. Agricultural extension, formally introduced in the United States by the Smith-Lever Act of 1914, plays a major role in information dissemination in agriculture. There is a large literature on the effect of agricultural extension, both in the United States and elsewhere, on agricultural productivity and technology adoption (Huffman (1974); Huffman (1977); Birkhaeuser et al. (1991); Dercon et al. (2009)). Despite the importance of extension, we argue that it is in fact congregational mergers and not extension services that generate the results we find in this paper: because of the fixed effects strategy, in order for agricultural extension to be driving these results, we would need to see agricultural extension services changing differently over time in treatment counties than in control counties, having removed the state time trend, only over the 1959 to 1964 time period. This is potentially plausible, but seems unlikely, especially because extension funding and the number of extension agents allowed is governed by state laws, which do not change often. For example, the Minnesota statutes outlining extension were first passed in 1923, updated in 1953, and were not revised again until 1969 (Minnesota Legislature (2013)). The law allows for "the formation of one county corporation in each county in [Minnesota]" to act as an extension agency, with in most cases one extension agent and a specified budget, based on the number of townships in the county (Minnesota Legislature (1953)). While county extension offices documented their activities for mandatory state reports, these reports were inconsistent across different counties and years. Also, many of

Table 2.8: Impact of Congregational Mergers on Capital

Variables	(1) Cars	(2) Trucks	(3) Tractors	(4) Bailers	(5) Freezers
Year = 1964 × merger	0.16	−15.09	27.48	5.29	2.92
	(15.93)	(16.74)	(51.84)	(10.49)	(20.64)
Mean of dependent variable	1,034.37	776.06	990.63	364.38	775.42
R^2	0.65	0.50	0.19	0.60	0.66
Observations	396	395	395	394	395
Number of counties	198	198	198	198	198
County FE	YES	YES	YES	YES	YES
State-by-year FE	YES	YES	YES	YES	YES

Notes: This table shows results from estimating Equation (2.1). Year = 1964 × merger is equal to one if the year is 1964, and the county experienced a congregational merger between 1959 and 1964. The dependent variables are, in column (1), farms reporting car ownership; in column (2), farms reporting truck ownership; in column (3), farms reporting tractor ownership; in column (4), farms reporting bailer ownership; and in column (5), farms reporting freezer ownership. Standard errors are clustered at the county level. Significance: *** $p < 0.01$, ** $p < 0,05$, * $p < 0.10$.

the variables measured were endogenous, such as the number of phone calls received or the number of attendees at extension events. As a result, it is impossible to credibly measure the intensity and efficacy of extension efforts over our sample period.[24]

We argue in this paper that congregational mergers impact fertilizer use through information. Another plausible explanation would be that the mergers also facilitated increased access to capital. In order to provide evidence against this possibility, we estimate Equation (2.1) again, this time with the number of farms with each of a variety of capital-intensive technologies as outcome variables. Table 2.8 shows the impact congregational mergers have on the number of farms with cars, trucks, tractors, bailers, and freezers.

As expected, we find no statistically or economically significant effect of congregational mergers on capital-intensive inputs: the standard errors are quite wide, and the effect sizes small: the coefficient on cars, for example, is only a 0.01 percent increase relative to the control group mean, and the standard error is almost one hundred times the size of the coefficient. This suggests that congregational mergers did not substantially increase access to capital, and provides additional evidence that information is the main channel through which congregational mergers impacted technology adoption.

Finally, one might worry that by only using TALC congregational mergers in our analysis, we are understating the true treatment effect. We argue above that the TALC mergers

[24]After contacting state archives in North Dakota and South Dakota and doing significant research at the University of Minnesota Archives, we are convinced that no adequate measure of extension outreach or capacity exists that would allow us to test extension as an alternative explanation to our findings.

are exogenous, and, due to the heavily Lutheran populations in these regions, the mergers where we would expect to see an effect. Indeed, the congregations that are merging in these data have, on average, 492 baptized members, so seeing an additional 35 farms begin to use fertilizer is an entirely reasonable effect size. There is another major Lutheran church branch, the Lutheran Church - Missouri Synod (LCMS), that was not directly involved in the TALC merger, but whose mergers could be attributed to increased discussion about merger surrounding TALC. We collected data from Concordia Historical Institute, the LCMS seminary, on congregational mergers between LCMS churches during the sample period. There is only one merger that occurs in a non-metropolitan county during this time period, and the inclusion of said merger does not produce a statistically distinguishable result from using only the TALC mergers.

Ultimately, given the range of tests that we perform, we have confidence that our results are robust and that we are correctly attributing them to the information effect of congregational mergers.

2.8 Conclusion

In this paper, we study the impacts of social networks on technology adoption. We combine data on American agriculture in the 1960s Midwest with unique information on mergers between Lutheran congregations. Using these mergers as exogenous shocks to social networks, we demonstrate that counties that experienced congregational mergers have over 7% more farms using fertilizer, and fertilize over 13% more acres than do other counties without mergers. Counties with mergers also use 21% more agricultural lime, a complement to fertilizer, and over 26% more fertilizer on corn. These effects are economically and statistically significant, and in line with theoretical predictions. We also provide evidence for a mechanism: our results are consistent with increased information spreading through larger social networks. Our results are robust to a variety of specifications, as well as to a placebo test.

Despite the fact that this is an isolated social networks experiment under specific conditions, these results have policy implications: they suggest that contact with innovators can be a meaningful way to increase technology adoption in agriculture. Furthermore, our results shed light on the importance of networks that are not explicitly arranged around economic activity. In particular, we find that religious social networks are an important factor in economic decision-making, as churches represent a key focal point for information diffusion among midwestern farmers. Future research should seek out other naturally-occurring shocks to social networks to further understand the within-network interactions that drive information diffusion.

Chapter 3

Fueling Local Water Pollution: Ethanol Refineries, Land Use, and Nitrate Runoff.

3.1 Introduction

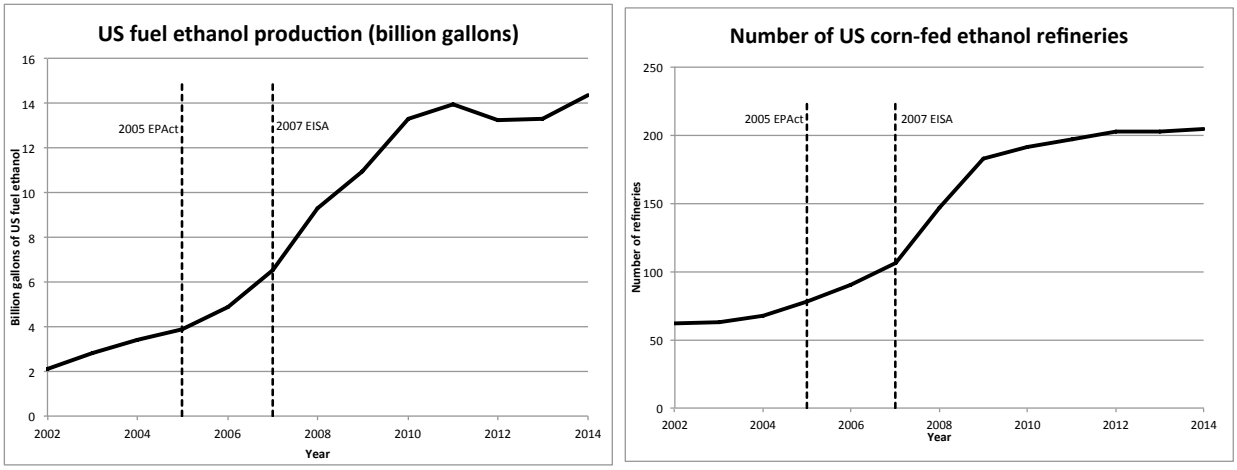
Since the early 2000s, US ethanol production has exploded in response to federal policies incentivizing the production of renewable fuels. In 2005, Congress passed the Energy Policy Act (EPAct) introducing a Renewable Fuel Standard (RFS) mandating that 2.78% of gasoline sold in the US be from renewable sources. In 2007, Congress passed the Energy Independence and Security Act (EISA) setting annual renewable fuel mandates for US production with an ultimate goal of 36 billion gallons by 2022. Of these 36 billion gallons, 15 billion are to be conventional biofuels – corn-based ethanol in particular.

The US ethanol industry has clearly responded to the Renewable Fuel Standards established in the EPAct and EISA. Between 2002 and 2014, US ethanol production has increased from just over 2 billion gallons per year to over 14 billion gallons per year (Figure 3.1a). In order to produce such quantities of ethanol, the number of corn ethanol refineries in the US has increased from 62 in 2002 to 204 in 2014 (Figure 3.1b).

The striking increase in US corn ethanol production has raised several important questions about its unintended consequences. One strand of research has explored how increased demand for ethanol has affected land use in the US corn belt as aggregate demand for corn increases (Fatal and Thurman, 2014; Miao, 2013; Feng and Babcock, 2010). Another strand of research has been more concerned about the environmental externalities of changing agricultural patterns, particularly focused on nitrate runoff and water pollution (Donner and Kucharik, 2008; Thomas et al., 2009).

In this chapter, I explore both the land use change effects and environmental effects of expanding ethanol production. In particular, I study the geospatial effect of ethanol refineries' placement on nearby land use change and use my results to estimate environmental

Figure 3.1: Growth of US Ethanol Production and Refineries, 2002–2014.



(a) US Ethanol Production

(b) US Ethanol Refineries

Source: Renewable Fuels Association.

consequences. I am specifically interested in how the location of ethanol refineries spatially affects agricultural land, and I do not attempt to identify the full general equilibrium effect of the 14 billion gallon US corn ethanol industry. Put another way, I study how the distribution of ethanol refineries differentially affects different agricultural areas net of the ethanol industry's aggregate effect on corn prices.

I find that within a population of almost 114 million acres of agricultural land in Illinois, Indiana, Iowa, and Nebraska, nearly 300,000 more acres of corn were grown in 2014 than in 2002 due merely to ethanol refinery location effects. This represents approximately 21,000 tons of nitrogen applied as fertilizer. Almost all the 300,000 acres of increased corn acreage exist within 30 miles of an ethanol refinery, suggesting that these refineries have strong local effects on land use change and nitrogen use.

There is clear economic intuition for why ethanol refineries would differentially affect nearby and faraway agricultural land. When a corn-fed ethanol refinery is built, it represents a new terminal market for corn. Since refineries operate continuously, they have an inelastic demand for this input. And since transportation costs are significant for grains, one would expect an ethanol refinery to source its corn from the nearest producers. Thus, by reducing transportation costs for nearby producers (reducing basis), ethanol refineries essentially subsidize corn production for nearby farmers. On the margin, this subsidy incentivizes farmers to grow more corn – or grow corn more often – than they otherwise would. As corn production increases, so will nitrogen fertilizer use. Corn requires higher levels of nitrogen fertilizer than other Corn Belt crops, and particularly high levels of fertilizer when grown successively corn-after-corn. Thus, economic intuition suggests ethanol refineries would have

a localized effect increasing corn production and nitrogen fertilizer use. Consequently, these refineries would also have an effect on localized nitrate runoff due to the increased nitrogen fertilizer use.

Researchers have previously addressed different components of the ethanol industry's effects on land use change and nitrate runoff. One line of research has explored whether the hypothesized local corn subsidy provided by nearby ethanol refineries actually exists. In a frequently cited paper, McNew and Griffith (2005) find that corn prices at an ethanol refinery are 12.5¢ higher than average, that the effect is slightly stronger for "upstream" refineries than for "downstream" refineries, and that price effects can be detected up to 68 miles from a refinery. However, Katchova (2009) and O'Brien (2009) both fail to find such a subsidy. Gallagher et al. (2005) highlight that locally-owned and non-locally-owned refineries have different effects on corn prices: the authors find that corn prices are increased by proximity to a non-locally-owned refinery, but not by proximity to a locally-owned refinery. Finally, Lewis (2010) finds different results in different states: ethanol refineries in Michigan and Kansas affect local corn prices, but refineries in Iowa and Indiana do not.

Other authors have explored whether ethanol refineries have an effect on land use. Fatal and Thurman (2014) use county-level data to estimate the corn acreage effect of ethanol refineries. They find that a typical ethanol refinery increases corn acreage in its home county by over 500 acres and has effects that can persist for up to 300 miles. Miao (2013) also uses county-level data and finds a significant effect of ethanol refineries on corn acreage, as well as a differential effect between locally-owned and non-locally-owned refineries. Turnquist et al. (2008), in contrast to more recent studies, fail to find any significant agricultural land conversion in areas near Wisconsin ethanol refineries. Finally, Feng and Babcock (2010) explore the full general equilibrium effect of increased ethanol production and find an unambiguous increase in corn acreage.

Several researchers have focused on how ethanol production affects water quality and nitrate runoff. Donner and Kucharik (2008) highlight how the aggregate impact of the EISA will likely make achieving nitrate level goals in the Mississippi impossible. Thomas et al. (2009) use hydrologic models to estimate the water quality impacts of corn production caused by increased demand due to biofuel mandates. They find significant negative results.

While it is likely true that "refineries cause corn," it is also likely true that "corn causes refineries." Ethanol refineries are not located at random, and several researchers have explored the topic of ethanol refinery placement. A series of papers have shown, unsurprisingly, that ethanol refineries are more likely to locate near areas with large corn production, near transportation infrastructure, and not near existing ethanol refineries (Sarmiento et al., 2012; Haddad et al., 2010; Lambert et al., 2008). This finding is important because it highlights that ethanol refinery placement cannot be treated as truly random in econometric analyses without accounting for the underlying drivers of this placement.

In my analysis, I argue that field-level fixed effects appropriately account for the major determinants of refinery placement. In particular, I study how distance-to-nearest-refinery affects the probability of a field being planted to corn. Whenever a new refinery is built, its presence differentially affects fields close to it relative to fields slightly farther away.

However, due to the spatial characteristics of soil quality and topology, "more-treated" and "less-treated" fields are qualitatively comparable.

My project improves upon previous work by leveraging new sources of field-level land use data and exploiting a finer-scaled panel of observations than previous authors. I exploit both the Cropland Data Layer (CDL) and Common Land Unit (CLU) to create annual observations of field-level land use. These agricultural micro-data allow for much more nuanced econometric estimation than in previous studies. Other authors have exploited similar micro-data in agricultural research to great effect (Livingston et al., 2015; Hendricks et al., 2014; Wright and Wimberly, 2013). I also highlight the locality effect of ethanol refineries rather than the general equilibrium effect, focusing on small-scale heterogeneous effects that have not been well identified in previous work.

The remainder of this paper is divided into a theoretical framework (model), a summary of my data, an overview of my econometric methods, a discussion of my results, and a conclusion.

3.2 Model of Optimal Crop Choice

Consider a farmer maximizing expected profits from an agricultural field. I assume the farmer is not forward-looking, and maximizes only the current year's expected profits.[1] The farmer observes input prices, expected output prices, the locations of terminal markets for possible crops, and the field's planting history. Then, the farmer chooses to plant either corn (C), soy (S), or another crop (O) to maximize:

$$\begin{aligned} \mathbb{E}\left[\Pi_i\right] &= \max\left\{\mathbb{E}\left[\Pi_i(C)\right], \mathbb{E}\left[\Pi_i(S)\right], \mathbb{E}\left[\Pi_i(O)\right]\right\} \\ &= \max_{x \in \{C,S,O\}} \left\{\mathbb{E}\left[\left(p_x - b_{i,x}\right) f_x\left(\mathbf{v}_x | x_{i,-1}\right) - \mathbf{z}_x \cdot \mathbf{v}_x\right]\right\} \end{aligned} \tag{3.1}$$

where Π_i are profits for field i, p_x is the output price for crop x, $b_{i,x}$ is the basis for crop x on field i, f_x is the production function for crop x, $\mathbf{v}_x$ is a vector of inputs to produce crop x, $x_{i,-1}$ is the crop planted on field i in the previous period, and $\mathbf{z}_x$ is a vector of input prices for inputs $\mathbf{v}_x$. Basis $b_{i,x}$ is assumed to be a linear function of distance from field i to the nearest terminal market for crop x, and input quantities $\mathbf{v}_x$ are determined my maximizing Π_i conditional on x and $x_{i,-1}$.

Given the problem outlined above, a farmer's optimal decision is deterministic. However, for an econometrician who does not observe all relevant data and production functions, the above problem gives rise to a probability that field i will be planted crop x given previous

[1] Of course, we expect farmers to be forward-looking and dynamically optimizing their cropping decisions. Livingston et al. (2015) provide an excellent treatment of how these dynamics affect farmers' optimal choices. These authors find relatively little difference in the optimal behavior of a myopic farmer compared to that of an infinitely-forward-looking farmer, suggesting my own model is an acceptable approximation of reality. However, one could easily incorporate the dynamics from Livingston et al. (2015) into a more complicated version of the model I present here.

planting decisions: $P_i(x|x_{i,-1})$. Summing across different possible planting histories, this gives rise to an unconditional probability that field i will be planted to crop x: $P_i(x)$. In this project, I am interested in how ethanol refineries affect the probability a field will be planted to corn: $P_i(C)$.

Now consider the introduction of a new corn-fed ethanol refinery. There are two interesting cases. First, suppose the new refinery is closer to field i than any existing refineries, but further away than field i's current terminal market for corn. In this case, field i's distance-to-nearest-refinery changes, but its distance-to-nearest-corn-market remains unchanged and its basis for corn, $b_{i,C}$ remains the same as before. In the second case, suppose the new refinery is closer to field i than field i's current terminal market for corn. In this case, field i's basis for corn, $b_{i,C}$ gets smaller while its bases for soy and other crops, $b_{i,S}$ and$b_{i,O}$, remain unchanged.[2]

Within the second case outlined above, the specific placement of the ethanol refinery matters. Since I assume basis is linear in distance to terminal market, $b_{i,C}$ will increase with distance to the new ethanol refinery as long as the refinery is closer to field i than the next-nearest terminal market.

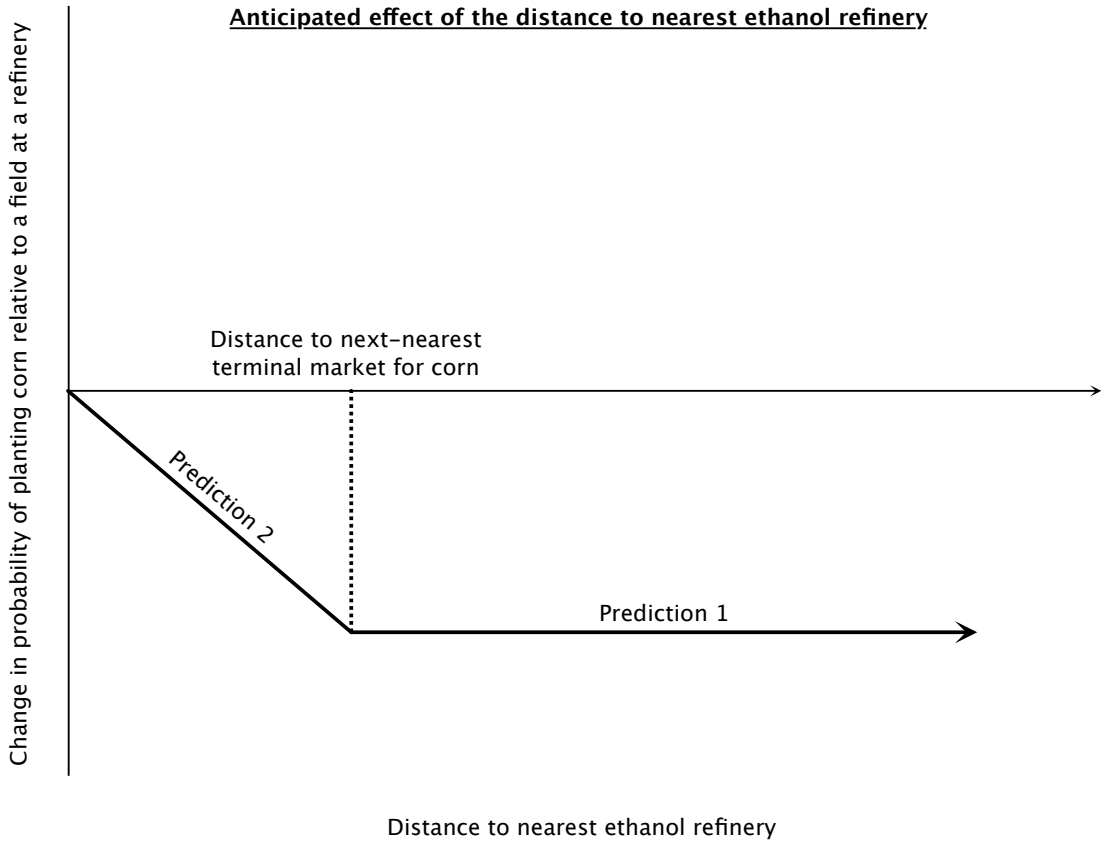
These assumptions give rise to two predictions about the effect of corn-fed ethanol refineries on the probability corn is planted on any field i. First, regardless of planting history, if the new refinery is further away than the current terminal market for corn, there will be no impact on field i's probability of being planted to corn $P_i(C)$. Second, regardless of planting history, if the new refinery is closer than the current terminal market for corn, the new refinery will increase $P_i(C)$ linearly in distance. Figure 3.2 gives a graphical representation of these two predictions.

Unfortunately, in reality, I only observe a field's distance to its nearest ethanol refinery. I do not observe its distance to all nearest terminal markets for corn or any other crop. If I were able to observe locations for all grain elevators or other terminal markets, I might be able to more explicitly test for both predictions summarized in Figure 3.2. As it is, I can only hope to observe an effect that has the general shape outlined in Figure 3.2.

An important point to note is that the effect of distance to the nearest ethanol refinery on the probability of planting corn is not strictly linear; it is piecewise-linear. This suggests that running any regression with simply a linear "distance to nearest ethanol refinery" covariate on the right-hand-side will obfuscate the true underlying relationship. A non-linear specification is required.

[2]This model makes the implicit assumption that a field's nearest terminal market for any crop x is able to absorb the product of all fields for whom it is the nearest terminal market. In the case of ethanol refineries, this may not be true. Rather, refinery production capacities may impose additional constraints or relax other constraints. I leave this issue to future work.

Figure 3.2: Model prediction of the effect of distance to nearest ethanol refinery on probability a field is planted to corn



3.3 Data

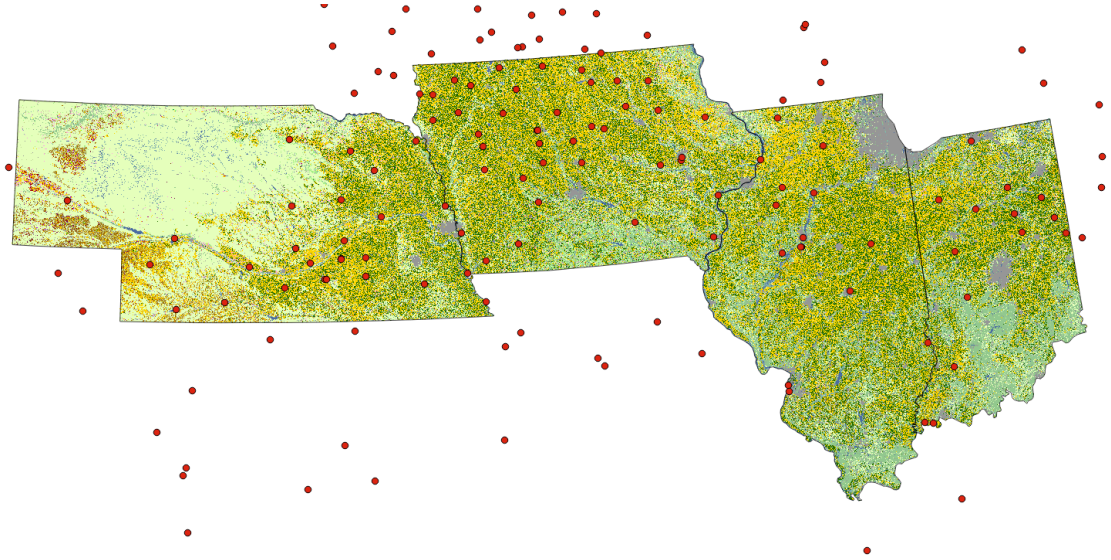
To conduct my analysis, I construct a balanced panel of annual crop choices for 2,145,848 agricultural fields in Illinois, Indiana, Iowa, and Nebraska over thirteen years from 2002 to 2014. In each year, I also calculate the distance from each field to the nearest ethanol refinery. I rely on three data sources to create my panel: the Cropland Data Layer, Common Land Unit, and ethanol refinery locations.

3.3.1 Cropland Data Layer

The Cropland Data Layer (CDL) is a raster dataset of landcover in the United States collected and maintained by the National Agricultural Statistics Service (NASS) of the USDA. A satellite records the electro-magnetic wavelengths of light reflected from different points on the earth's surface and uses a ground-tested algorithm to assign each pixel a single land-cover type for the year. Pixels measure 30 meters by 30 meters, except for years 2006-2009 when

pixels measured 56 meters by 56 meters.[3] The CDL provides remarkably high-resolution land cover data and is able to distinguish between many different types of vegetation. Figure 3.3 displays the CDL for Illinois, Indiana, Iowa, and Nebraska in 2014. Yellow pixels represent corn, dark green pixels represent soy, and light green pixels represent other grassland-like land covers. Red dots mark ethanol refineries operating in 2014.

Figure 3.3: 2014 Cropland Data Layer (CDL) with ethanol refinery locations.



Sources: NASS & Renewable Fuels Association.

The CDL identifies different crops with different accuracies. For major row crops, CDL accuracy is usually between 85% and 95% (Boryan et al., 2011). For corn and soy, pixel accuracy is particularly high. For example, according to the 2014 CDL metadata, accuracy for both corn and soy in Iowa was over 97%. However, the CDL is considerably less accurate at distinguishing between more similar land covers, such as between alfalfa, rangeland, and grassland. This fact has complicated research that explores extensive land use change in the Western Corn Belt (Wright and Wimberly, 2013), but is not a concern for research focused on corn and soy.

One problem with using raw CDL data is that a 30 meter by 30 meter pixel is likely not the appropriate unit of analysis. Rather, economists are more interested in observing field-level crop choices. Additionally, while CDL data are quite accurate for primary row crops, it is apparent that individual pixels are frequently mis-measured. For instance, upon

[3]Data collection for the CDL began in the late 1990s in only three states. By 2008, all 48 contiguous states had been included in the CDL. Changes in CDL technical specifications – such as different pixel sizes in different years – can be attributed to a growing data collection program with evolving hardware and software resources.

visual inspection of a CDL image, it is not uncommon to observe what is clearly a large field of more than 100 pixels planted to soybeans, with one or two pixels somewhere in the field reported as corn. If analysis is conducted at the pixel level rather than the field level, such mis-measurements become a large concern. To address this concern, I exploit Common Land Unit data to construct field-level crop cover observations.

3.3.2 Common Land Unit

According to the Farm Service Agency (FSA) of the USDA, a Common Land Unit (CLU) is "an individual contiguous farming parcel, which is the smallest unit of land that has a permanent, contiguous boundary, common land cover and land management, a common owner, and/or a common producer association" (Farm Service Agency, 2012). Practically, a CLU represents a single agricultural field. Polygon shapefiles of CLUs are maintained by the FSA, but are not currently publicly available.

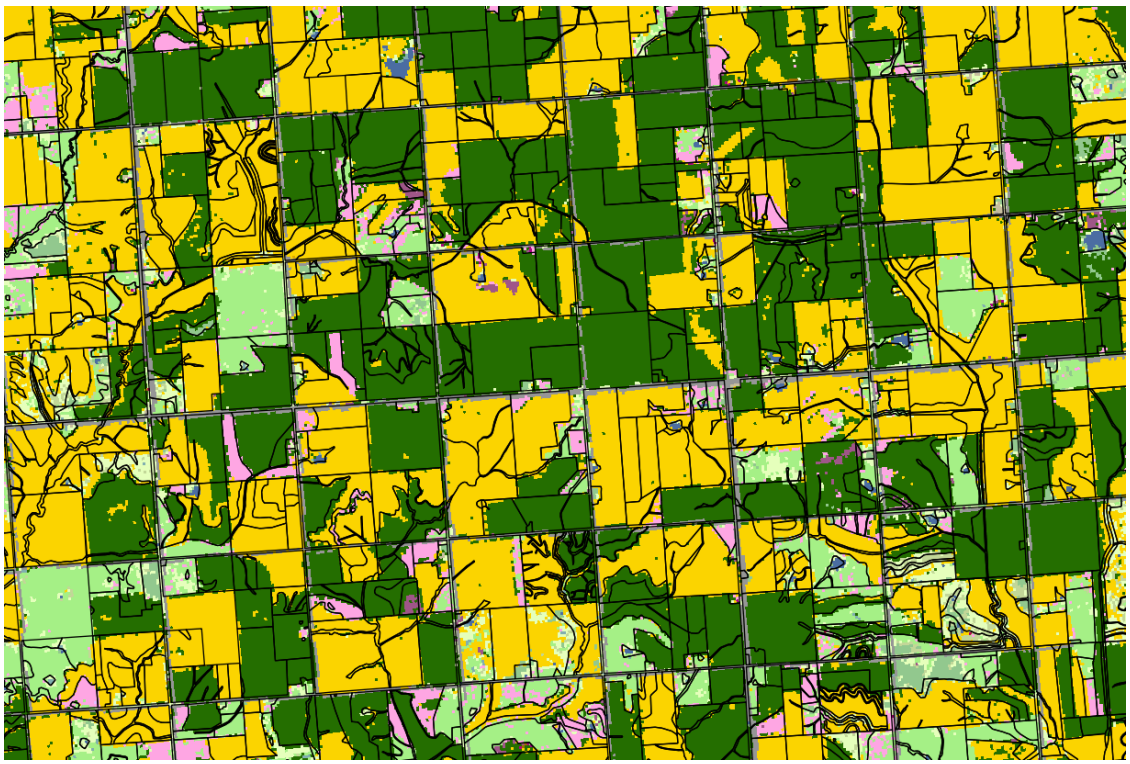
I obtain CLU data for Illinois, Indiana, Iowa, and Nebraska from the website GeoCommunity (`http://www.geocomm.com`). These data contain shapefiles from the mid 2000s, before CLU data were removed from the public domain. In this research, I implicitly assume that individual CLUs do not change over time: a reasonable assumption given the FSA definition. In reality, the FSA does adjust individual CLU definitions on a case-by-case basis if necessary, but I assume these adjustments to be negligible as in previous similar studies (Hendricks et al., 2014).

Using the geospatial software ArcGIS, I overlay the CDL raster data with CLU polygons as shown in Figure 3.4. Upon visual inspection, the fit is quite good: CLU boundaries line up with crop changes in the CDL, roads appear clearly in both datasets, and geographical features such as waterways and elevation changes are visible. One concern is that many CLUs are quite small and appear to outline geographical features such as gullies, rather than larger constituent fields. This is particularly pronounced in areas near urban sprawl. Therefore, to maintain confidence that the fields I study are actually "fields" in the way we think of them, I drop all CLUs from my dataset with areas of less than 10 acres. I also drop CLUs with areas of greater than 10,000 acres, based on an assumption that these CLUs are incorrect.[4]

To assign each CLU a single crop cover, I calculate the modal value of the raster pixels contained within each CLU polygon. I then assign that modal value to the entire CLU. This procedure enforces the assumption that each field (CLU) is planted to a single crop – an assumption strongly supported by a visual examination of the data. To my knowledge, this is the first instance of using modal statistics to interact the CDL and CLU datasets. Previous research (Hendricks et al., 2014) has used an off-center centroid to sample a single point of the underlying raster data. My procedure is preferable in that it reduces the chance of idiosyncratic mis-measurement of the field's true land cover.

[4]I drop 3,201,933 fields that have areas under 10 acres. The aggregate area dropped is 10,043,985 acres. I drop 108 fields that have areas over 10,000 acres. The aggregate area dropped is 1,965,324 acres. The total acreage dropped is less than 10% of the 125,987,632 acres in my initial dataset.

Figure 3.4: Detail of Cropland Data Layer (CDL) and Common Land Unit (CLU) data: Iowa, 2014



Yellow pixels represent corn, dark green pixels represent soy, light green pixels represent grassland, and pink pixels represent alfalfa. Black lines are CLU borders. Sources: NASS & FSA.

Finally, for each CLU polygon, I construct a centroid for the field. This centroid is constrained to exist within the boundaries of its parent CLU polygon. I then use these CLU centroids to calculate distances from each field to its nearest ethanol refinery in each year.

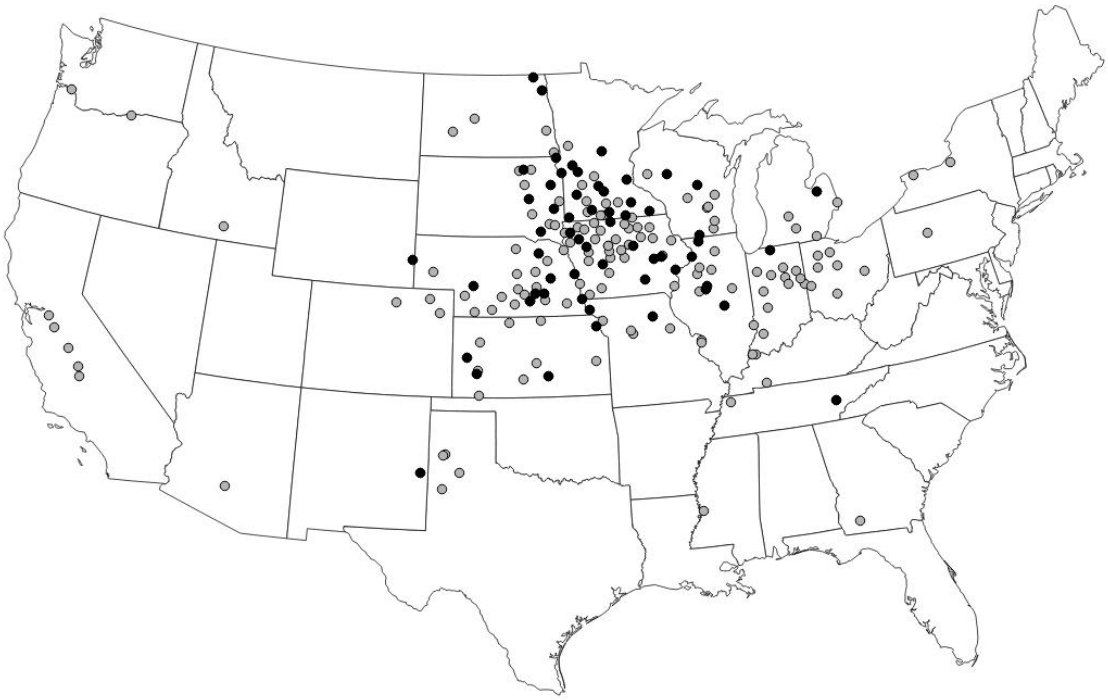
3.3.3 Ethanol Refineries

I obtain data on ethanol refinery location, capacity, and opening date from the Renewable Fuels Association (RFA). The RFA has comprehensive data on ethanol refineries each year starting in 2002.[5] Using these data, I geo-code the locations of over 200 ethanol refineries in the US. Since new ethanol refineries open each year, I create a separate dataset of operating ethanol refineries for each year from 2002 to 2014. I only include refineries that can use

[5]Data on ethanol refineries is unavailable for 2013. In my analysis, I assume all "new" ethanol refineries in 2014 opened in 2014, even though some of them may have opened in 2013.

corn as an input and omit refineries that only accept cellulosic biomass or non-corn inputs. Figure 3.5 displays the geographic expansion of corn-fed ethanol refineries between 2002 and 2014.

Figure 3.5: Ethanol refinery locations in 2002 (black) and 2014 (gray)

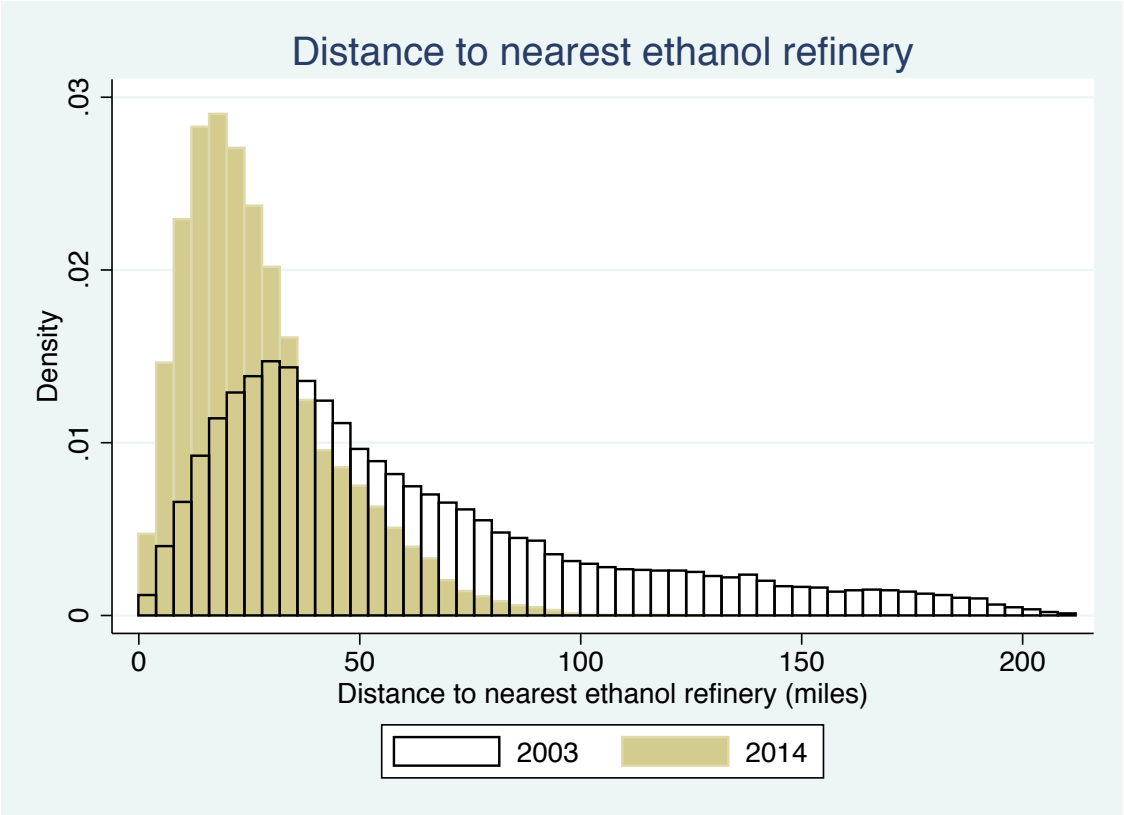


Source: Renewable Fuels Association.

Using ArcGIS, I calculate the distance from each CLU centroid to the nearest ethanol refinery for each year from 2002 to 2014. As new ethanol refineries are constructed, this distance will decrease for nearby fields. This change in distance-to-nearest-refinery is the variation I will use to identify my econometric analysis. Figure **??** displays the change in distribution of nearest-distances from 2002 to 2014. Distributions for years 2003-2013 are omitted for clarity, but the distribution skews more and more to the left in each year.

In the current project, I do not incorporate ethanol refinery production capacity into my analysis. Rather, I treat each refinery as identical. Thus, there is no analytic difference between a field 30 miles from a 10 million-gallons-per-year (mgy) refinery and a field 30 miles from a 100 mgy refinery. I plan to exploit production capacity information in future work.

Figure 3.6: Distributions of field-level distance-to-nearest-refinery in 2003 and 2014.



Each distribution represents 2,145,848 agricultural fields. Source: author's calculations.

3.3.4 Summary Statistics

The final dataset I use for my analysis consists of a balanced panel of 2,145,853 agricultural fields over 13 years. Table 3.1 presents summary statistics for the data. The average field is 53.12 acres in area, is 62.14 miles away from the nearest ethanol refinery in 2002, and 27.52 miles away from the nearest refinery in 2014.

In any particular year, an average of 32.85% of all fields are planted to corn representing 33.27% of all acreage. Similarly, 25.28% of fields are planted to soy representing 23.51% of all acreage. Remaining fields and acreage are left to other crops.

Table 3.1: Summary Statistics

VARIABLE	Mean	Median	Std. Dev.	Minimum	Maximum	Observations
Field area (acres)	53.12	30.15	123.60	10	9,934.94	2,145,853
Fields of corn	704,992	699,640	54,088	627,912	785,925	13
% of fields corn	32.85	32.60	2.52	29.26	36.63	13
Acreage of corn	37,915,466	38,217,232	2,819,787	33,887,520	41,720,688	13
% of acreage corn	33.27	33.53	2.47	29.73	36.60	13
Fields of soy	542,552	545,443	39,075	434,344	593,006	13
% of fields soy	25.28	25.42	1.82	20.24	27.63	13
Acreage of soy	26,792,631	27,205,006	1,809,211	21,987,108	29,220,192	13
% of acreage soy	23.51	23.87	1.59	19.29	25.64	13
Fields of other crops	898,309	901,122	58,856	816,503	981,631	13
% of fields other crops	41.86	41.99	2.74	38.05	45.75	13
Acreage of other crops	49,270,226	49,726,160	2,207,439	46,356,128	52,840,092	13
% of acreage other crops	43.23	43.63	1.94	40.67	46.36	13
Dist. to nearest refinery, 2002	62.14	48.95	44.05	0.01	211.84	2,145,853
Dist. to nearest refinery, 2003	61.60	47.96	44.19	0.01	211.84	2,145,853
Dist. to nearest refinery, 2004	61.58	47.96	44.19	0.01	211.84	2,145,853
Dist. to nearest refinery, 2005	49.38	41.66	30.99	0.01	157.32	2,145,853
Dist. to nearest refinery, 2006	48.55	40.89	31.40	0.01	157.32	2,145,853
Dist. to nearest refinery, 2007	46.77	38.44	31.19	0.01	157.32	2,145,853
Dist. to nearest refinery, 2008	34.15	28.36	22.86	0.01	131.05	2,145,853
Dist. to nearest refinery, 2009	30.07	25.57	19.54	0.01	128.41	2,145,853
Dist. to nearest refinery, 2010	29.44	24.81	19.48	0.01	128.41	2,145,853
Dist. to nearest refinery, 2011	28.91	24.59	18.87	0.01	126.30	2,145,853
Dist. to nearest refinery, 2012	27.52	23.75	17.07	0.01	126.30	2,145,853
Dist. to nearest refinery, 2013*	27.52	23.75	17.07	0.01	126.30	2,145,853
Dist. to nearest refinery, 2014**	27.52	23.75	17.07	0.01	126.30	2,145,853

Notes: Total number of fields: 2,145,853. Total acreage: 113,978,323 acres. Variables with 2,145,853 observations are measured at the field level. Variables with 13 observations are measured at the year level. All distances measured in miles. *The Renewable Fuels Association did not publish data on ethanol refineries in 2013, so there is no change in distance in my data between 2012 and 2013. **Although the mean and median distances for 2014 appear identical to those for 2013, this is only due to rounding. The values for 2014 are in fact smaller.

3.4 Econometric Methods

Using a balanced panel of 2,145,848 agricultural fields over 13 years, I estimate a linear probability model (LPM) of the probability that an individual field is planted to corn. I include as independent variables a field-level fixed effect to capture time-invariant unobserved characteristics of individual fields, state-by-year fixed effects to capture input and output prices, and distance-bin dummy variables measuring distance to the nearest ethanol refinery. These distance-bin variables allow me to observe a non-linear relationship between a field's distance to its nearest ethanol refinery and the probability that field is planted to corn. The

distance-bins are the covariates of interest. In particular, I estimate equation 3.2:

$$P_{it}(C) = \beta_0 + \sum_{k=1}^{21} \beta_k \text{bin}(10 \times k)_{it} + \gamma_{st} + \alpha_i + \varepsilon_{it} \tag{3.2}$$

where i indexes field, t indexes year, $P_{it}(C)$ is the probability of field i growing corn in year t, $\text{bin}10, ..., \text{bin}210$ are sequential distance bin dummies beginning at the 10 mile mark and each representing a range of 10 miles (the omitted bin in 0-10 miles), γ_{st} is a state-by-year fixed effect, α_i is a field-level fixed effect, and ε_{it} is an error term. I cluster standard errors at the field level to control for heteroskedasticity and correlation over time.

Equation 3.2 has several desirable qualities. First, and most importantly, an LPM allows me to control for field-level fixed effects by exploiting the within-transformation. This allows me to ignore any characteristics of a field that do not change over time, such as soil quality, average weather patterns, and field slope. Second, an LPM allows me to easily interpret the coefficients $\beta_1, ..., \beta_{21}$ as marginal effects. For instance, a field 25 miles away from its nearest ethanol plant is $100 \times \beta_2$ percent more likely to grow corn than a field 0-10 miles away from its nearest ethanol plant, *ceteris paribus*.

Linear probability models have drawbacks as well. LPMs can result in coefficients that will predict outcomes outside the [0,1] interval, as opposed to discrete-choice models such as logit or probit. The reason I use an LPM rather than a discrete-choice model is that fixed-effects are easy to incorporate in such a framework and the results are easy to interpret. Additionally, with large sample sizes, LPMs often perform quite similarly to discrete choice models.

In addition to equation 3.2, I estimate similar equations conditioning on a field's land use during the previous year. In particular, I estimate $P_{it}(C|C_{-1})$ (the probability of growing corn given that corn was grown last year), $P_{it}(C|S_{-1})$ (the probability of growing corn given that soy was grown last year), and $P_{it}(C|O_{-1})$ (the probability of growing corn given something other than corn or soy was grown last year). These specifications allow me to explore how ethanol refineries affect specific crop rotation dynamics rather than merely an aggregate effect.

To interpret the coefficients of interest in equation 3.2 ($\beta_1, ..., \beta_{21}$) as causal, I rely on the following identification assumption: that the placement of ethanol refineries is as good as random after accounting for field fixed effects and time fixed effects. More precisely, I assume that a field's distance to the nearest ethanol refinery is uncorrelated with the error term ε_{it} after accounting for the field fixed effect α_i and state year fixed effects γ_{st}. This is a reasonable assumption given that field fixed effects capture any observable or unobservable place-based characteristics that would drive refinery placement, and state year fixed effects capture aggregate ethanol production capacity over time. This argument requires me to treat ethanol placement to be an irreversible investment. In the period from 2002 to 2014, this is a reasonable assumption.

Additionally, I note that important determinants of crop choice (soil quality, topography, etc.) change in largely continuous ways across the Corn Belt landscape. Because I am

focused on the "treatment" of the distance of a refinery on a field's likelihood of growing corn, a refinery's presence effectively "treats" near fields more severely than far fields, even though the near and far fields are likely to share soil characteristics and are otherwise largely comparable. Field fixed effects will capture these common determinants of likely crop choice, allowing me to interpret my estimated effect as causal.

3.5 Results

I estimate equation 3.2 using the `reghdfe` command in `stata`. This command optimizes the estimation of high-dimensional fixed effects models and runs considerably faster than `xtreg`. However, the `reghdfe` command subsumes the constant term β_0 which must be reconstructed after the regression has been estimated. Thus, the constants reported in Table 3.2 do not include standard errors.

Table 3.2 presents my results for four different econometric specifications. In all cases, the dependent variable is the probability of a field being planted to corn. In specification (1), this probability in unconditional. In specifications (2), (3), and (4), the probability is conditional upon corn, soy, or neither (respectively) being grown on the field in the previous year.

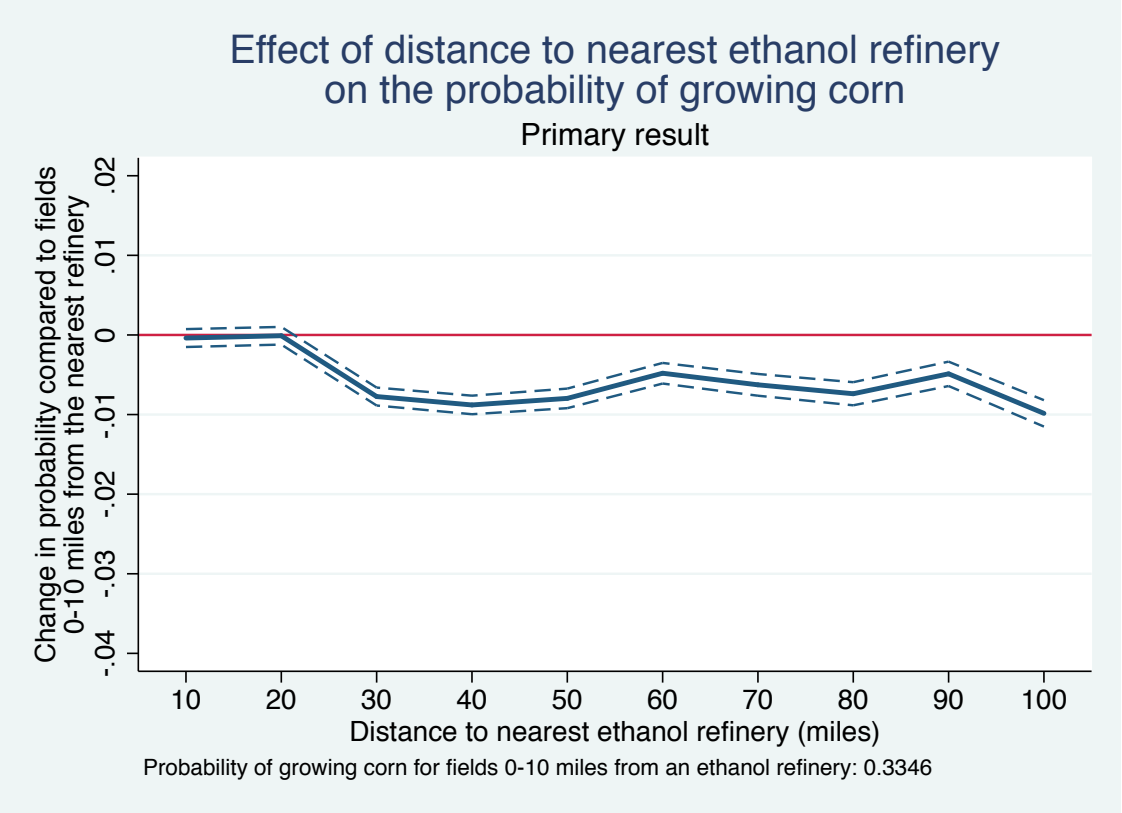
The unconditional probability of a field 0-10 miles from its nearest ethanol refinery being planted to corn is 33.46%. This is only slightly higher than the 32.49% probability of being planted to corn after being planted to corn last year. However, the same probability after being planted to soy is more than twice as large at 76.79%. Corn is relatively unlikely to be grown after crops other than corn or soy, with a probability of only 8.25%.

It is easiest to interpret the coefficients on each of the distance bins by plotting them on a graph. Figures 3.7, 3.8, 3.9, and 3.10 correspond to specifications (1), (2), (3), and (4), respectively. In each case, vertical distance on the graph measures changes in the probability that a field is planted to corn relative to fields 0-10 miles from their nearest ethanol refineries. Also, recall that by 2014, the mean distance to a field's nearest ethanol plant is 27.52 miles, the median distance is 23.75 miles, and the maximum distance is 126.30 miles. I highlight these facts to focus readers' attention on the areas of the following graphs most relevant to the underlying population of fields.

Consider Figure 3.7. The first result to note is an overall shape between the 20 and 100 mile markers that looks remarkably similar to the shape of the piecewise-linear curve in Figure 3.2. If we rely heavily on our theoretical framework, Figure 3.7 may suggest the average field is approximately 30-40 miles away from its nearest non-ethanol-refinery terminal market for corn. Unlike my model's prediction, however, Figure 3.7 displays no statistically significant effects of distance in the 10 and 20 mile bins. Broadly, specification (1) suggests that fields 0-10 miles from their nearest ethanol refineries are approximately 1% more likely to grow corn in any given year than fields 40-100 miles away from their nearest refinery.

Figure 3.8 presents results conditioning on corn being planted in the previous year (specification (2)). In this case, we see a stronger and more pronounced effect of distance to

Figure 3.7: Unconditional probability of growing corn: specification (1)



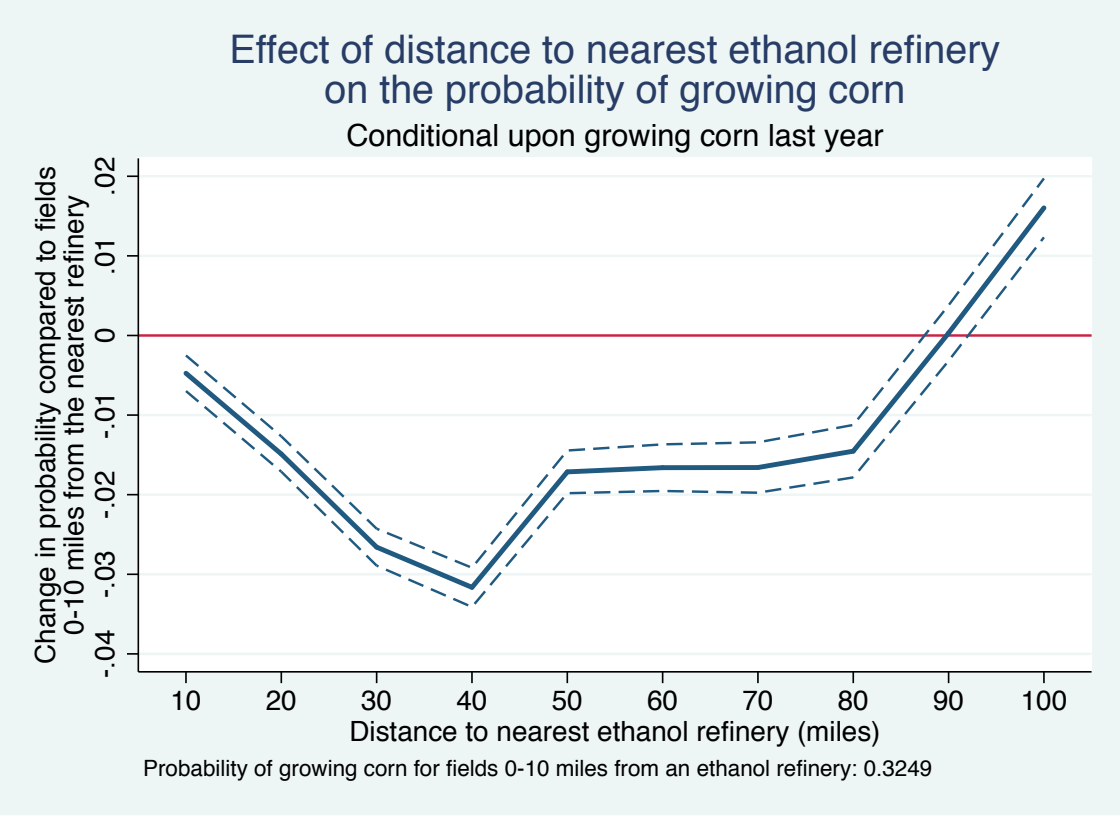
nearest ethanol refinery on the probability of planting corn in the region 10-40 miles away from a refinery. Here we more clearly see the sloped portion of the curve predicted in Figure 3.2. The interpretation of this regression is that as ethanol refineries are constructed, they strongly incentivize nearby fields to grow corn-after-corn relative to fields further away. This behavior particularly exacerbates any negative externalities of nitrogen fertilizer use.

Figure 3.9 presents results conditioning on soy being planted in the previous year. This figure displays a puzzling relationship: within the range of 10-70 miles, fields close to ethanol refineries are less likely to grow corn-after-soy than are fields further away. This result is contrary to the prediction developed in Figure 3.2, and has no straightforward explanation based on the model developed in this paper.

Figure 3.10 presents results conditioning on something other than corn or soy being planted in the previous year. It displays a downward trend between 20 and 100 miles from the nearest ethanol refinery, but does not clearly reflect the prediction outlined in Figure 3.2.

Taken together, and relying primarily on the unconditional results shown in Figure 3.7, I conclude that the distance to a field's nearest ethanol plant does affect that field's probability

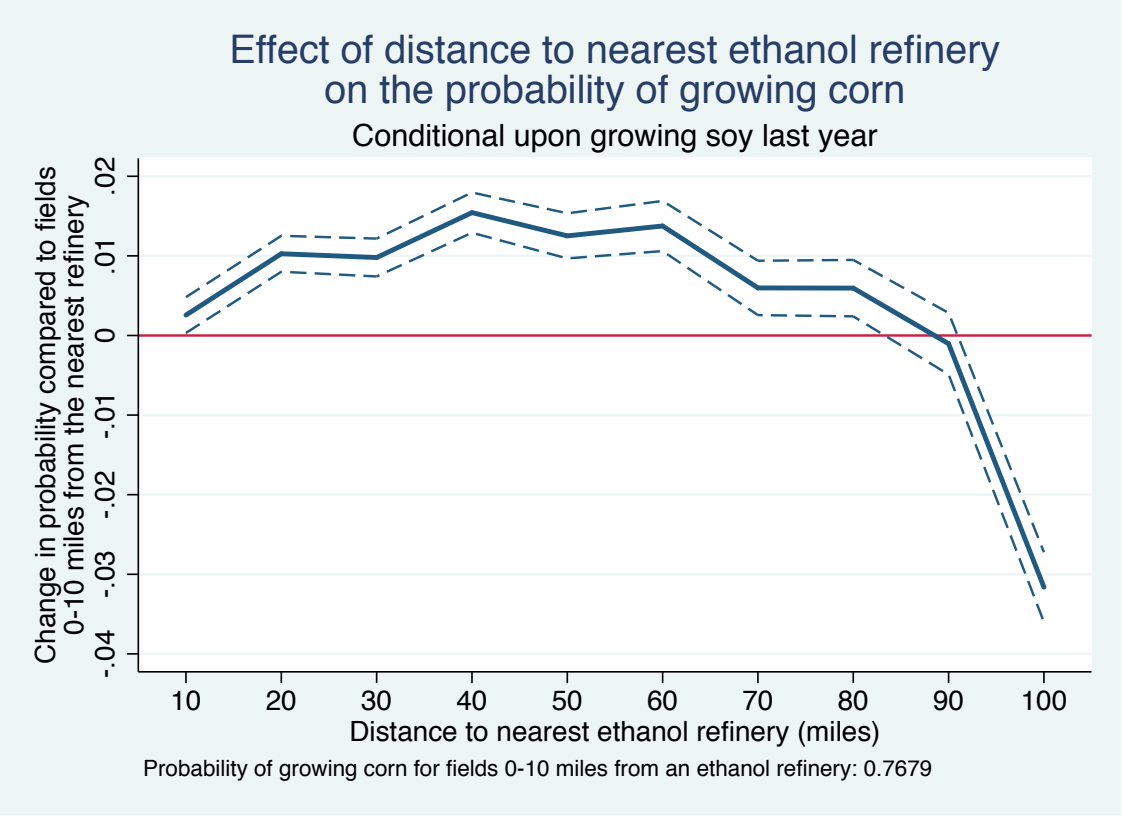
Figure 3.8: Probability of growing corn conditional on growing corn in the prior year: specification (2)



of being planted to corn in a non-linear way. If anything, results from Figure 3.8 suggest that farmers are realizing this effect in a way that maximizes nutrient strain on crop rotations and that exacerbates the negative externalities associated with nitrogen fertilizer use.

Next, I use the results from specification (1) to determine how the entry of ethanol refineries between 2002 and 2014 affected corn acreage in my population of fields. For each of the 2,145,853 fields in my dataset, I use their distance to the nearest ethanol refinery in 2002 and the relevant coefficient from specification (1) to approximate the unconditional probability of that field being planted to corn in 2002. I then repeat this process using each field's distance to nearest ethanol refinery in 2014. Subtracting the former probability from the latter, I construct the change in unconditional probability between 2002 and 2014 of each field growing corn in any particular year. Note that this change in probability is entirely attributable to the distance-to-nearest-refinery effect, and does not depend on level-shifts in corn demand between 2002 and 2014. I then multiply each field's change in probability by its acreage and sum across all fields to find the total net change in acreage between 2002 and 2014. I find a total increase of 298,718 acres in my population of 113,978,323 acres. Figure

Figure 3.9: Probability of growing corn conditional on growing soy in the prior year: specification (3)



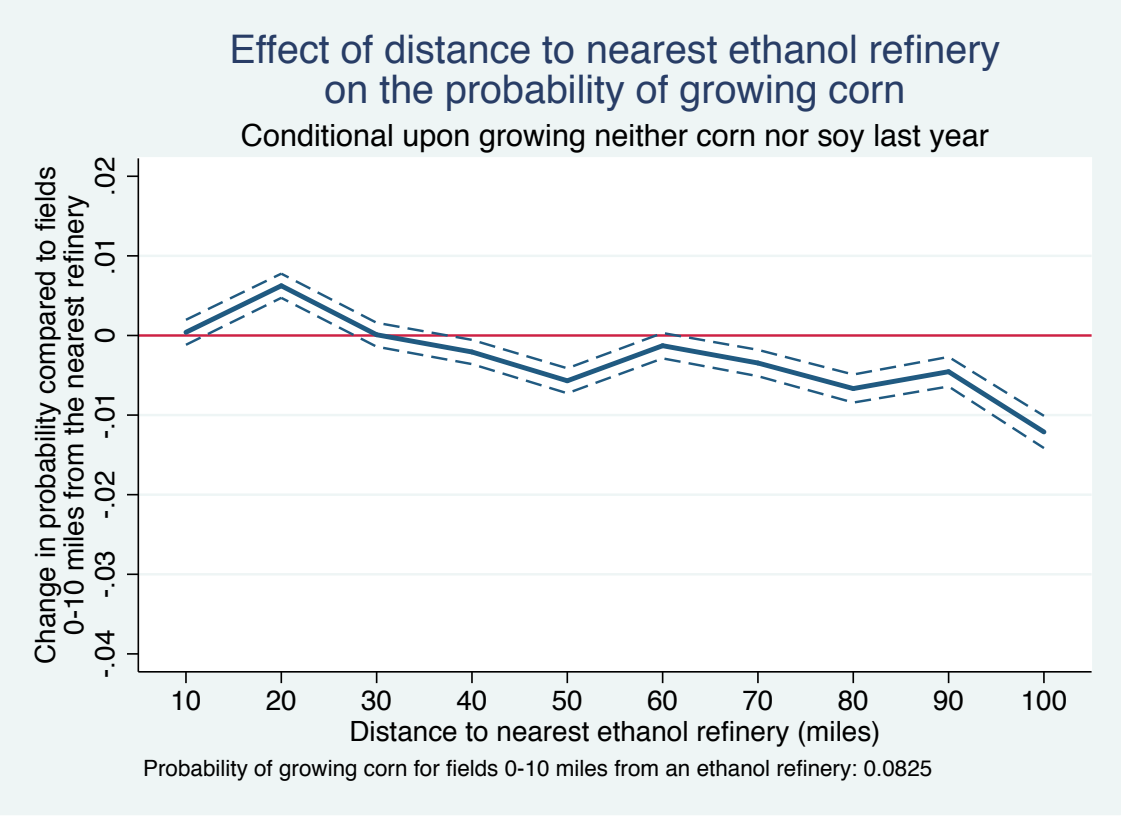
3.11 presents this change in acreage divided into distance-bins of 10 miles each.

The net increase in corn acreage of 298,718 that I find is only 0.26% of the 113,978,323 acres in my population, but it is 0.76% of all corn acreage in my population in 2014. This is a significant number given that it can be attributed to only the distance-to-nearest-refinery effect. In other words, the effect of new ethanol refineries since 2002 on lowering transportation costs (and *not* the general-equilibrium effect of ethanol increasing aggregate demand for corn) can explain almost 300,000 acres of the corn grown in a subset of the fields across Illinois, Indiana, Iowa, and Nebraska.

Figure 3.11 highlights that the entirety of this acreage effect is captured by fields less than 30 miles from the nearest ethanol refinery. This result matches incredibly well with the predictions outlined in Figure 3.2. It also demonstrates that any spatial externalities associated with increased corn cultivation due to ethanol refinery location occur entirely within 30 miles of ethanol refineries. This suggests highly localized effects.

What do the acreage increases highlighted in Figure 3.11 mean for nitrogen application? A 2007 Iowa State University Extension publication suggests that optimal nitrogen applica-

Figure 3.10: Probability of growing corn conditional on growing something other than corn or soy in the prior year: specification (4)



tion for corn-after-soy is 125 lb N/acre, and optimal application for corn-after-corn is 175 lb N/acre (Sawyer, 2007). Taking a middle value of 140 lb N/acre, (recall that most corn is grown after soy), the 298,718 acres of increased corn acreage estimated in Figure 3.11 represent 41,820,520 lbs, or almost 21,000 tons of extra nitrogen.

These 21,000 tons of additional nitrogen that are attributable to the distance effect of ethanol refinery placement are essentially all applied to areas within 30 miles of an ethanol refinery. While this number is relatively small relative to the total application of nitrogen in the US Corn Belt, there is cause for concern about localized geographic effects. Nitrate runoff into local water sources is harmful to water ecosystems, animals, and humans, and has been a growing problem in the US Corn Belt (Donner and Kucharik, 2008; Mueller and Helsel, 1996). Local water quality data from the USGS could be used in future research to look for an effect of ethanol refineries on nitrate levels directly.

Figure 3.11: Effect of distance to nearest ethanol refinery on corn acreage

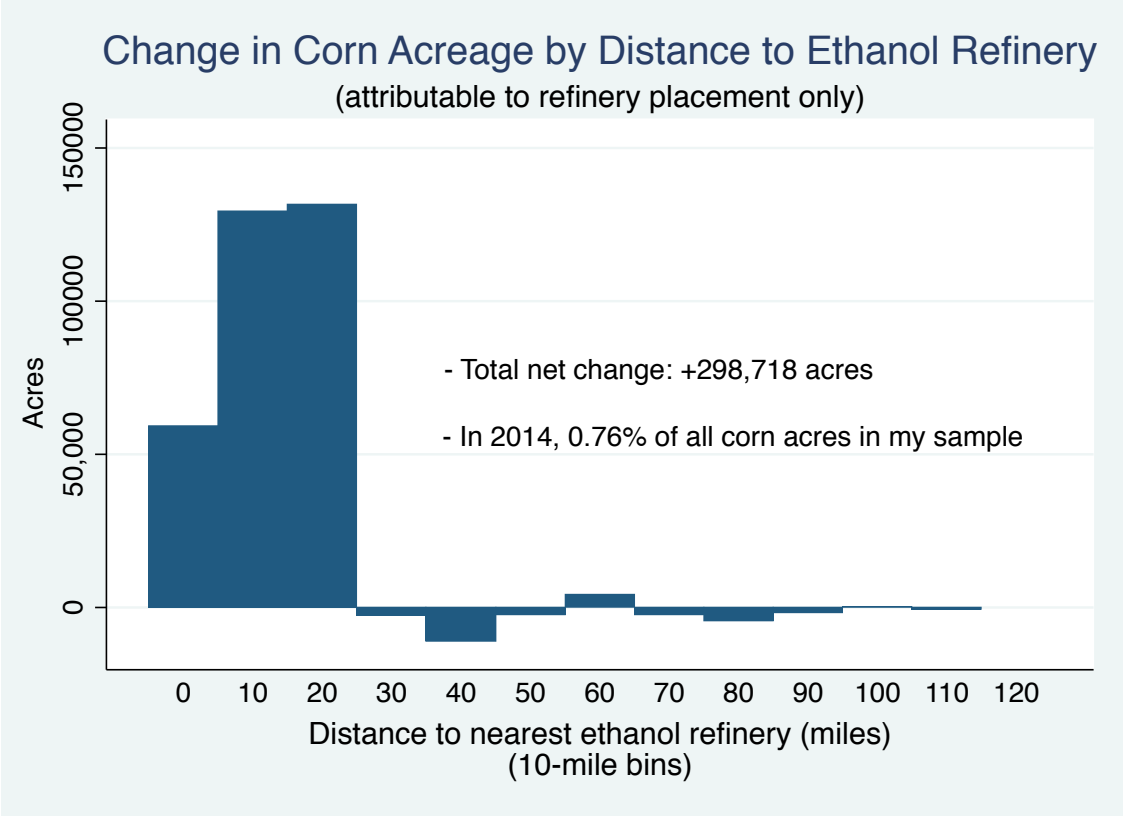


Table 3.2: Probability of a field being planted to corn

VARIABLES	(1) $P_{it}(C)$	(2) $P_{it}(C\|C_{-1})$	(3) $P_{it}(C\|S_{-1})$	(4) $P_{it}(C\|O_{-1})$
Constant (recovered)	0.3346	0.3249	0.7679	0.0825
Distance bin: 10-20 miles	-0.0004	-0.0047***	0.0026**	0.0004
	(0.0006)	(0.0011)	(0.0011)	(0.0008)
Distance bin: 20-30 miles	-0.0001	-0.0149***	0.0103***	0.0062***
	(0.0006)	(0.0011)	(0.0012)	(0.0008)
Distance bin: 30-40 miles	-0.0077***	-0.0266***	0.0098***	0.0001
	(0.0006)	(0.0012)	(0.0012)	(0.0008)
Distance bin: 40-50 miles	-0.0088***	-0.0316***	0.0154***	-0.0021***
	(0.0006)	(0.0013)	(0.0013)	(0.0008)
Distance bin: 50-60 miles	-0.0080***	-0.0171***	0.0125***	-0.0057***
	(0.0006)	(0.0014)	(0.0014)	(0.0008)
Distance bin: 60-70 miles	-0.0048***	-0.0166***	0.0138***	-0.0013
	(0.0007)	(0.0015)	(0.0016)	(0.0008)
Distance bin: 70-80 miles	-0.0062***	-0.0166***	0.0060***	-0.0035***
	(0.0007)	(0.0016)	(0.0017)	(0.0008)
Distance bin: 80-90 miles	-0.0074***	-0.0145***	0.0060***	-0.0067***
	(0.0007)	(0.0017)	(0.0018)	(0.0009)
Distance bin: 90-100 miles	-0.0049***	0.0003	-0.0010	-0.0046***
	(0.0008)	(0.0018)	(0.0020)	(0.0009)
Distance bin: 100-110 miles	-0.0098***	0.0160***	-0.031***	-0.0121***
	(0.0008)	(0.0019)	(0.0022)	(0.0010)
Distance bin: 110-120 miles	-0.0144***	0.0339***	-0.0415***	-0.0295***
	(0.0009)	(0.0021)	(0.0024)	(0.0012)
Distance bin: 120-130 miles	-0.0084***	0.0211***	-0.0254***	-0.0244***
	(0.0010)	(0.0022)	(0.0026)	(0.0013)
Distance bin: 130-140 miles	-0.0058***	0.0089***	-0.0454***	-0.0150***
	(0.0011)	(0.0026)	(0.0030)	(0.0013)
Distance bin: 140-150 miles	-0.0112***	0.0045	-0.0804***	-0.0193***
	(0.0012)	(0.0033)	(0.0036)	(0.0015)
Distance bin: 150-160 miles	-0.0113***	0.0141***	-0.1125***	-0.0031*
	(0.0014)	(0.0040)	(0.0044)	(0.0016)
Distance bin: 160-170 miles	-0.0108***	0.0246***	-0.0999***	-0.0081***
	(0.0015)	(0.0046)	(0.0049)	(0.0015)
Distance bin: 170-180 miles	-0.0077***	0.0198***	-0.1071***	-0.0129***
	(0.0015)	(0.0050)	(0.0052)	(0.0016)
Distance bin: 180-190 miles	0.0046***	0.0273***	-0.1079***	-0.0089***
	(0.0016)	(0.0053)	(0.0060)	(0.0017)
Distance bin: 190-200 miles	0.0051**	0.0368***	-0.0837***	-0.0192***
	(0.0020)	(0.0092)	(0.0073)	(0.0021)
Distance bin: 200-210 miles	0.0206***	0.0388**	-0.0630***	-0.0167***
	(0.0027)	(0.0155)	(0.0144)	(0.0028)
Distance bin: 210-220 miles	0.0193	0.0412	-0.3913***	-0.0090
	(0.0104)	(0.0896)	(0.0690)	(0.0090)
Field FE	YES	YES	YES	YES
State-by-Year FE	YES	YES	YES	YES
Observations	25,750,236	8,424,413	6,477,454	10,848,369
Number of fields	2,145,853	1,549,958	1,451,343	1,414,126
R-squared	0.3324	0.4834	0.5052	0.5021

Notes: In each specification, the dependent variable is the probability of a field being planted to corn. In specification (1), the probability is unconditional. In specifications (2), (3), and (4), the probability is conditional upon corn, soy, or neither (respectively) being grown on the field in the previous year. Distance bins are dummy variables for 10-mile ranges of distance to the nearest ethanol refinery. The omitted bin is 0-10 miles, so the constant term represents the probability of a field 0-10 miles from the nearest ethanol plant being planted to corn. Coefficients on the distance bins are interpreted as marginal effects relative to the constant term. Standard errors clustered at the field level are in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

3.6 Conclusion

In this paper, I have demonstrated that ethanol refineries exert a statistically significant effect on the land use of surrounding fields. Increases in corn acreage and nitrogen application occur within 30 miles of ethanol refineries, suggesting a highly localized effect. These findings are consistent with a model of ethanol refineries lowering corn basis for nearby farmers. Within a sample of almost 114 million acres, I find nearly 300,000 acres of the corn grown in 2014 can be attributed to ethanol placement effects accumulated over the years between 2002 and 2014.

This project makes several important contributions to the existing literature and improves upon previous research. Most importantly, I leverage field-level observations of land use to create a thirteen year panel of over two million observations. This allows me to estimate a highly nonlinear relationship between distance to a field's nearest ethanol refinery and that field's probability of growing corn. My panel also allows me to include field-level fixed effects that control for time-invariant characteristics of each field such as soil type.

In three econometric specifications that condition on the previous year's land use, I find interesting patterns. Ethanol refineries seem to strongly incentivize nearby fields to grow corn-after-corn, while the effect appears opposite for corn-after-soy. The result for corn-after-soy is puzzling and is not predicted by my model. Future work may attempt to better understand this result. Nonetheless, the net effect of these two individual effects is that farmers appear to be growing more corn near ethanol refineries in the way the most stresses crop rotations and most exacerbates the use of nitrate-producing fertilizer.

There is considerable room for further work on these questions. First, this analysis treats all refineries as identical. In reality, different refineries have different production capacities and ownership structures, and may have heterogeneous effects on surrounding land. Second, there is room to explore a wider range of econometric specifications beyond the linear probability model estimated in this paper. Finally, future work should explore data from the US Geological Survey to test whether water nitrate levels directly reflect the effect derived in the current project.

While the findings of this paper appear relatively small in the context of the entire US Corn Belt, they are strongly statistically significant and demonstrate a real and important localized effect of ethanol refinery placement. My results are useful for anyone interested in a fuller understanding of the spatial forces driving land use change and nitrogen application in agriculture.

Bibliography

Adhvaryu, A., N. Kala, and A. Nyshadham (2016a). Management and Shocks to Worker Productivity. Working Paper.

Adhvaryu, A., N. Kala, and A. Nyshadham (2016b). The Light and the Heat: Productivity Co-benefits of Energy-saving Technology. Working Paper.

Akerlof, G. A. and J. L. Yellen (Eds.) (1986). *Efficiency Wage Models of the Labor Market*. Cambridge, UK: Cambridge University Press.

Alatas, V., A. Banerjee, A. G. Chandrasekhar, R. Hanna, and B. A. Olken (2016). Network Structure and the Aggregation of Information: Theory and Evidence from Indonesia. *American Economic Review 106*(7), 1663–1704.

Allen, D. W. and D. Lueck (2002). *The Nature of the Farm: Contracts, Risk, and Organization in Agriculture*. Cambridge, MA: The MIT Press.

Amato, J. A. and A. Amato (2000). Minnesota, Real and Imagined: A View from the Countryside. *Daedalus 129*(3), 55–80.

Antle, J. M. and S. M. Capalbo (2010). Adaptation of Agricultural and Food Systems to Climate Change: An Economic and Policy Perspective. *Applied Economic Perspectives and Policy 32*(3), 386–416.

Azzi, C. and R. G. Ehrenberg (1975). Household Allocation of Time and Church Attendance. *Journal of Political Economy 83*(1), 27–56.

Bandiera, O., I. Barankay, and I. Rasul (2005). Social Preferences and the Response to Incentives: Evidence from Personnel Data. *The Quarterly Journal of Economics 120*(3), 917–962.

Bandiera, O. and I. Rasul (2006). Social Networks and Technology Adoption in Northern Mozambique. *The Economic Journal 116*(514), 869–902.

Banerjee, A., A. G. Chandrasekhar, E. Duflo, and M. O. Jackson (2013). The Diffusion of Microfinance. *Science 341*(6144).

Banerjee, A., A. G. Chandrasekhar, E. Duflo, and M. O. Jackson (2016). Gossip: Identifying Central Individuals in a Social Network.

Barber, S. A. and R. K. Stivers (1962). Fertilizing Farm Fields in Indiana. *Extension Circular* (474).

Baylis, P. (2016). Temperature and Temperament: Evidence from a billion tweets. Working Paper.

Beaman, L., A. BenYishay, J. Magruder, and A. M. Mobarak (2016). Can Network Theory-based Targeting Increase Technology Adoption? Unpublished working paper.

Beaman, L. and J. Magruder (2012). Who Gets the Job Referral? Evidence from a Social Networks Experiment. *American Economic Review 102*(7), 3574–3593.

Beaman, L. A. (2012). Social Networks and the Dynamics of Labour Market Outcomes: Evidence from Refugees Resettled in the U.S. *The Review of Economic Studies 79*(1), 128–161.

Becker, G. S. (1965). A Theory of the Allocation of Time. *The Economic Journal 75*(299), 493–517.

Becker, G. S. (1977). A Theory of the Production and Allocation of Effort. NBER Working Paper 184.

Bethel Lutheran Northfield Church (2014). Homepage. `http://bethelnorthfield.org/our-history/`.

Bethlehem Lutheran Church (2014). Homepage. `http://www.bethlehemff.org/welcome.html`.

Billikopf, G. (2008). Designing an Effective Piece Rate. `https://nature.berkeley.edu/ucce50/ag-labor/7research/7calag06.htm`. Accessed: September 3, 2016.

Billikopf, G. E. (1995). High piece-rate wages do not reduce hours worked. *California Agriculture 49*(1), 17–18.

Billikopf, G. E. and M. V. Norton (1992). Pay method affects vineyard pruner performance. *California Agriculture 46*(5), 12–13.

Birkhaeuser, D., R. E. Evenson, and G. Feder (1991). The Economic Impact of Agricultural Extension: A Review. *Economic Development and Cultural Change 39*(3), 607–650.

Boryan, C., Z. Yang, R. Mueller, and M. Craig (2011). Monitoring US agriculture: the US department of agriculture, national agricultural statistics service, cropland data layer program. *Geocarto International 26*(5), 341–358.

Brown, C. (1990). Firms' Choice of Method of Pay. *Industrial and Labor Relations Review 43*(3), 165S–182S.

Burrows, R. (2010). Starting a Commercial Grape Vineyard in South Dakota. `http://igrow.org/up/resources/06-2018-2012.pdf`.

California Blueberry Commission (2015). Annual Report 2014–2015. `https://www.cdfa.ca.gov/statistics/PDFs/2015Report.pdf`. Accessed: September 4, 2016.

Camerer, C., L. Babcock, G. Loewenstein, and R. Thaler (1997). Labor Supply of New York City Cabdrivers: One Day at a Time. *Quarterly Journal of Economics 112*(2), 407–441.

Cameron, A. C., J. B. Gelbach, and D. L. Miller (2011). Robust Inference with Multiway Clustering. *Journal of Business and Economic Statistics 29*, 238–249.

Campbell, R. R., M. Campbell, and C. Hughes (2004). *A Revolution in the Heartland: Changes in Rural Culture, Family and Communities: 1900-2000*. University of Missouri Press.

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

Chang, T., J. Graff Zivin, T. Gross, and M. Neidell (2016a). Particulate Pollution and the Productivity of Pear Pickers. *American Economic Journal: Economic Policy 8*(3), 141–169.

Chang, T., J. Graff Zivin, T. Gross, and M. Neidell (2016b). The Effect of Pollution on Worker Productivity: Evidence from Call-Center Workers in China. NBER Working Paper 22328.

Chang, T. and T. Gross (2014). How many pears would a pear packer pack if a pear packer could pack pears at quasi-exogenously varying piece rates? *Journal of Economic Behavior & Organization 99*, 1–17.

Chen, M. K. and M. Sheldon (2015). Dynamic Pricing in a Labor Market: Surge Pricing and Flexible Work on the Uber Platform. Working Paper.

Choi, S., E. Gallo, and S. Kariv (2016). Networks in the laboratory. *The Oxford Handbook of the Economics of Networks*. Eds: Yann Bramoulle, Andrea Galeotti, and Brian Rogers.

Colmer, J. (2016). Weather, Labor Reallocation, and Industrial Production: Evidence from India. Working Paper.

Conley, T. and C. Udry (2001). Social Learning through Networks: The Adoption of New Agricultural Technologies in Ghana. *American Journal of Agricultural Economics 83*(3), 668–673.

Conley, T. G. and C. R. Udry (2010). Learning about a New Technology: Pineapple in Ghana. *American Economic Review 100*(1), 35–69.

Cotter, M. and B. Jackson (2001). *Growing Up on a Minnesota Farm.* Arcadia Publishing.

Crawford, V. P. and J. Meng (2011). New York City Cab Drivers' Labor Supply Revisited: Reference-Dependent Preferences with Rational Expectations Targets for Hours and Income. *The American Economic Review 101*(5), 1912–1932.

Crocker, T. D. and R. L. Horst, Jr. (1981). Hours of Work, Labor Productivity, and Environmental Conditions: A Case Study. *The Review of Economics and Statistics 63*(3), 361–368.

Dell, M., B. F. Jones, and B. A. Olken (2012). Temperature Shocks and Economic Growth: Evidence from the Last Half Century. *American Economic Journal: Macroeconomics 4*(3), 66–95.

Dercon, S., D. O. Gilligan, J. Hoddinott, and T. Woldehanna (2009). The Impact of Agricultural Extension and Roads on Poverty and Consumption Growth in Fifteen Ethiopian Villages. *American Journal of Agricultural Economics 91*(4), 1007–1021.

Deryugina, T. and S. M. Hsiang (2014). Does the Environment Still Matter? Daily Temperature and Income in the United States. NBER Working Paper 20750.

Dickinson, D. L. (1999). An Experimental Examination of Labor Supply and Work Intensities. *Journal of Labor Economics 17*(4), 638–670.

Dillon, A., J. Friedman, and P. Serneels (2014). Health Information, Treatment, and Worker Productivity: Experimental Evidence from Malaria Testing and Treatment among Nigerian Sugarcane Cutters. World Bank Group Impact Evaluation Series, Policy Research Working Paper WPS 7120.

Donner, S. D. and C. J. Kucharik (2008). Corn-based ethanol production compromises goal of reducing nitrogen export by the Mississippi river. *Proceedings of the National Academy of Sciences 105*(11), 4513–4518.

Dou, X., C. Gray, V. Mueller, and G. Sheriff (2016). Labor Adaptation to Climate Variability in Eastern Africa. International Food Policy Research Institute Discussion Paper 01537.

Dugan, G. (1960). Lutheran Groups Plan 2 Mergers; Bodies in Denomination Will Be Cut From 16 to 11 as Wider Union Is Sought. *The New York Times*.

Edin, P.-A., P. Fredriksson, and O. ÃĔslund (2003). Ethnic Enclaves and the Economic Success of ImmigrantsâĂŤEvidence from a Natural Experiment. *The Quarterly Journal of Economics 118*(1), 329–357.

Enander, A. E. and S. Hygge (1990). Thermal stress and human performance. *Scandinavian Journal of Work, Environment & Health 16*(Supplement 1), 44–50.

Epstein, Y., G. Keren, J. Moisseiev, O. Gasko, and S. Yachin (1980). Psychomotor deterioration during exposure to heat. *Aviation, Space, and Environmental Medicine 51*(6), 607–610.

Eswaran, M. and A. Kotwal (1885). A Theory of Contractual Structure in Agriculture. *The American Economic Review 75*(3), 352–367.

Evenson, R. E. and L. E. Westphal (1994). Handbook of Development Economics. Volume 3. North Holland.

Fan, M., S. Gabbard, A. A. Pena, and J. M. Perloff (2015). Why Do Fewer Agricultural Workers Migrate Now? *American Journal of Agricultural Economics 97*(3), 665–679.

Farber, H. S. (2005). Is Tomorrow Another Day? The Labor Supply of New York City Cabdrivers. *Journal of Political Economy 113*(1), 46–82.

Farber, H. S. (2008). Reference-Dependent Preferences and Labor Supply: The Case of New York City Taxi Drivers. *The American Economic Review 98*(3), 1069–1082.

Farm Service Agency (2012). Common Land Unit (CLU): Information sheet. Technical report, USDA.

Fatal, Y. S. and W. N. Thurman (2014). The response of corn acreage to ethanol plant siting. *Journal of Agricultural and Applied Economics 46*(02).

Feder, G., R. E. Just, and D. Zilberman (1985). Adoption of Agricultural Innovations in Developing Countries: A Survey. *Economic Development and Cultural Change 33*(2), 225–298.

Fehr, E. and L. Götte (2007). Do Workers Work More if Wages Are High? Evidence from a Randomized Field Experiment. *The American Economic Review 97*(1), 298–317.

Feng, H. and B. A. Babcock (2010). Impacts of ethanol on planted acreage in market equilibrium. *American Journal of Agricultural Economics 92*(3), 789–802.

Fisher, D. U. and R. D. Knutson (2013). Uniqueness of Agricultural Labor Markets. *American Journal of Agricultural Economics 95*(2), 463–469.

Foster, A. D. and M. R. Rosenzweig (1994). A Test for Moral Hazard in the Labor Market: Contractual Arrangements, Effort, and Health. *The Review of Economics and Statistics 76*(2), 213–227.

Foster, A. D. and M. R. Rosenzweig (1995). Learning by Doing and Learning from Others: Human Capital and Technical Change in Agriculture. *Journal of Political Economy 103*(6), 1176–1209.

Foster, A. D. and M. R. Rosenzweig (2010). Microeconomics of Technology Adoption. *Economic Growth Center Discussion Paper 984*.

Gabbard, S. M. and J. M. Perloff (1997). The Effects of Pay and Work Conditions on Farmworker Retention. *Industrial Relations 36*(4), 474–488.

Gallagher, P., R. Wisner, and H. Brubacker (2005). Price relationships in processors' input market areas: Testing theories for corn prices near ethanol plants. *Canadian Journal of Agricultural Economics/Revue canadienne d'agroeconomie 53*(2-3), 117–139.

Gallardo, R. K. and D. Zilberman (2016). The Economic Feasibility of Adopting Mechanical Harvesters by the Highbush Blueberry Industry. *HortTechnology 26*(3), 299–308.

Galloway, S. D. and R. J. Maughan (1997). Effects of ambient temperature on the capacity to perform prolonged cycle exercise in man. *Medicine and Science in Sports and Exercise 29*(9), 1240–1249.

Giné, X., M. Martínez-Bravo, and M. Vidal-Fernández (2010). Intertemporal Substitution or Reference-Dependent Preferences? Evidence from Daily Labor Supply of South Indian Boat-owners. Working Paper.

Goldsmith-Pinkham, P. and G. W. Imbens (2013). Social Networks and the Identification of Peer Effects. *Journal of Business & Economic Statistics 31*(3), 253–264.

González-Alonso, J., C. Teller, S. L. Andersen, F. B. Jensen, T. Hyldig, and B. Nielsen (1999). Influence of body temperature on the development of fatigue during prolonged exercise in the heat. *Journal of Applied Physiology 86*(3), 1032–1039.

Gordon, D. (1997). *Growing Fruit in the Upper Midwest*. University of Minnesota Press.

Grace Lutheran Church (2014). Our History. `http://www.gracedawson.org/clients/our-history`.

Graff Zivin, J. and M. Neidell (2012). The Impact of Pollution on Worker Productivity. *The American Economic Review 102*(7), 3652–3673.

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.

Graff Zivin, J. S., S. M. Hsiang, and M. J. Neidell (2015). Temperature and human capital in the short- and long-run. NBER Working Paper 21157.

Granger, S. and S. Kelly (2005). Historic Context Study of Minnesota Farms, 1820-1960, Volume 2.

Griliches, Z. (1957). Hybrid Corn: An Exploration in the Economics of Technological Change. *Econometrica 25*(4), 501–522.

Grossman, S. J. and O. D. Hart (1983). An Analysis of the Principal-Agent Problem. *Econometrica 51*(1), 7–45.

Guiteras, R. P. and K. Jack (2017). Productivity in Piece-Rate Labor Markets: Evidence from Rural Malawi. Working Paper.

Guthman, J. (2017). Paradoxes of the Border: Labor Shortages and Farmworker Minor Agency in Reworking California's Strawberry Fields. *Economic Geography 93*(1), 24–43.

Haddad, M. A., G. Taylor, and F. Owusu (2010). Locational choices of the ethanol industry in the midwest corn belt. *Economic Development Quarterly 24*(1), 74–86.

Haines, M., P. Fishback, and P. Rhode (2015). United States Agriculture Data, 1840 - 2010. ICPSR35206-v2.

Haley, M. R. (2003). The Response of Worker Effort to Piece Rates: Evidence from the Midwest Logging Industry. *Journal of Human Resources 38*(4), 881–890.

Hall, B. H. and B. Khan (2003). Adoption of New Technology. Unpublished working paper.

Hancock, P. A., J. M. Ross, and J. L. Szalma (2007). A meta-analysis of performance response under thermal stressors. *Human Factors 49*(5), 851–877.

Hart, O. and B. Holmström (1987). The theory of contracts. In T. F. Bewley (Ed.), *Advances in Economic Theory Fifth World Congress*, Chapter 3, pp. 71–155. Cambridge, UK: Cambridge University Press.

Heal, G. and J. Park (2016). Temperature Stress and the Direct Impact of Climate Change: A Review of an Emerging Literature. *Review of Environmental Economics and Policy 10*(2), 347–362.

Heal, G. M. and J. Park (2014, January). Feeling the Heat: Temperature, Physiology & the Wealth of Nations. Discussion Paper 2014-51: Harvard Environmental Economics Program.

Helms, J. D., G. A. Pavelis, S. Argabright, R. G. Cronshey, and H. Raymond Sinclair Jr. (1996). National Soil Conservation Policies: A Historical Case Study of the Driftless Area. *Agricultural History 70*(2), 377–394.

Hendricks, N. P., A. Smith, and D. A. Sumner (2014). Crop supply dynamics and the illusion of partial adjustment. *American Journal of Agricultural Economics 96*(5), 1469–1491.

Houser, T., S. Hsiang, R. Kopp, and K. Larsen (2015). *Economic Risks of Climate Change: An American Prospectus*. New York: Columbia University Press.

Hsiang, Solomon, M. (2010). Temperatures and cyclones strongly associated with economic production in the Caribbean and Central America. *Proceedings of the National Academy of Sciences 107*(35), 15367–15372.

Huffman, W. E. (1974). Decision Making: The Role of Education. *American Journal of Agricultural Economics 56*(1), 85–97.

Huffman, W. E. (1977). Allocative Efficiency: The Role of Human Capital. *The Quarterly Journal of Economics 91*(1), 59–79.

Isé, S. J., J. M. Perloff, S. R. Sutter, and S. Vaupel (1996). Grower beliefs determine hiring practices. *California Agriculture 50*(2), 17–20.

Ito, K. (2014). Do Consumers Respond to Marginal or Average Price? Evidence from Nonlinear Electricity Pricing. *The American Economic Review 104*(2), 537–563.

Jackson, M. O. (2008). *Social and Economic Networks*. Princeton University Press.

Jackson, M. O. (2014). Networks in the Understanding of Economic Behaviors. *Journal of Economic Perspectives 28*(4), 3–22.

Jackson, P. and C. Fleischer (2007). Religion and Economics: A Literature Review. Religions and Development Working Paper: Univeristy of Birmingham.

Jimenez, M., K. M. Klonsky, and R. L. De Moura (2009). Sample costs to establish and produce fresh market blueberries. University of California Cooperative Extension.

Johnston, J. M. (1960). New Church Is Born With the New Year. *The Milwaukee Sentinel*.

Kahn, M. E. (2016). The Climate Change Adaptation Literature. *Review of Environmental Economics and Policy 10*(1), 166–178.

Katchova, A. L. (2009, July). The spatial effect of ethanol biorefinery locations on local corn prices. In *Proceedings of the Agricultural and Applied Economics Association AAEA and ACCI Joint Annual Meeting*, Milwaukee, WI.

Kenney, L. (1995). *The Past is Never Far Away: A History of the Red River Valley Potato Industry*. Red River Valley Potato Growers Association.

Kern County Department of Agriculture and Measurement Standards (2016). Kern County Crop Statistics. `http://www.kernag.com/dept/stats/crop-stats.asp`. Accessed: August 31, 2016.

Killingsworth, M. R. (1983). *Labor Supply*. Cambridge, UK: Cambridge University Press.

Kőszegi, B. and M. Rabin (2006). A model of reference-dependent preferences. *The Quarterly Journal of Economics 121*(4), 1133–1165.

Kremer, M. and E. Miguel (2007). The Illusion of Sustainability. *The Quarterly Journal of Economics 122*(3), 1007–1065.

Lagerquist, L. D. (1999). *The Lutherans*. Greenwood Press.

Lambert, D., M. Wilcox, A. English, and L. Stewart (2008, 4). Ethanol plant location determinants and county comparative advantage. *Journal of Agricultural and Applied Economics 40*, 117–135.

Lazear, E. P. (2000). Performance Pay and Productivity. *The American Economic Review 90*(5), 1346–1361.

Lazerwitz, B. (1961). Some Factors Associated with Variations in Church Attendance. *Social Forces 39*(4), 301–309.

Lewis, K. E. (2010). The impact of ethanol on corn market relationships and corn price basis levels. Masters thesis, Michigan State University.

Livingston, M., M. J. Roberts, and Y. Zhang (2015). Optimal sequential plantings of corn and soybeans under price uncertainty. *American Journal of Agricultural Economics 97*(3), 855–878.

Mackworth, N. H. (1947). High Incentives versus Hot and Humid Atmospheres in a Physical Effort Task. *British Journal of Psychology. General Section 38*(2), 90–102.

Manski, C. F. (1993). Identification of Endogenous Social Effects: The Reflection Problem. *The Review of Economic Studies 60*(3), 531–542.

McNew, K. and D. Griffith (2005). Measuring the impact of ethanol plants on local grain prices. *Applied Economic Perspectives and Policy 27*(2), 164–180.

Meese, G., R. Kok, M. Lewis, and D. Wyon (1984). A laboratory study of the effects of moderate thermal stress on the performance of factory workers. *Ergonomics 27*(1), 19–43.

Miao, R. (2013). Impact of ethanol plants on local land use change. *Agricultural and Resource Economics Review 42*(2).

Minnesota Legislature (1953). Session Laws of Minnesota for 1953: Chapter 202 – S.F. No. 1236.

Minnesota Legislature (2013). 2013 Minnesota Statutes - 38.37 County Extension Committee; Program, County Agents.

Moretti, E. and J. M. Perloff (2002). Efficiency Wages, Deferred Payments, and Direct Incentives in Agriculture. *American Journal of Agricultural Economics 84*(4), 1144–1155.

Mueller, D. K. and D. R. Helsel (1996). Nutrients in the nation's waters–too much of a good thing? Technical Report 1136, US Geological Survey.

Munshi, K. (2003). Networks in the Modern Economy: Mexican Migrants in the U. S. Labor Market. *The Quarterly Journal of Economics 118*(2), 549–599.

Munshi, K. (2004). Social learning in a heterogeneous population: technology diffusion in the Indian Green Revolution. *Journal of Development Economics 73*(1), 185–213.

Murray County (2014). Murray County, Minnesota.

National Agricultural Statistics Service (2015). 2012 Census of Agriculture: Specialty Crops. `https://www.agcensus.usda.gov/Publications/2012/Online_Resources/Specialty_Crops/SCROPS.pdf`. Accessed: September 2, 2016.

National Agricultural Statistics Service (2016). Noncitrus Fruits and Nuts 2015 Summary. `http://usda.mannlib.cornell.edu/usda/current/NoncFruiNu/NoncFruiNu-07-06-2016.pdf`. Accessed: September 2, 2016.

Nelson, E. C. (1975). *The Lutherans in North America*. Fortress Press.

Newbery, D. M. and J. E. Stiglitz (1987). Wage Rigidity, Implicit Contracts, Unemployment and Economic Efficiency. *The Economic Journal 97*(386), 416–430.

Nielsen, B., J. R. Hales, S. Strange, N. J. Christensen, J. Warberg, and B. Saltin (1993, January). Human circulatory and thermoregulatory adaptations with heat acclimation and exercise in a hot, dry environment. *Journal of Physiology 460*, 467–485.

NOAA (2007). Time Bias Corrected Divisional Temperature-Precipitation-Drought Index (TD-9640).

O'Brien, D. M. (2009). The effects of the micro-market structure for kansas grain elevators on spatial grain price differentials. In *Proceedings of the NCCC-134 Conference on Applied Commodity Price Analysis, Forecasting, and Market Risk Management*, St. Louis, MO.

Oettinger, G. S. (1999). An Empirical Analysis of the Daily Labor supply of Stadium Vendors. *Journal of Political Economy 107*(2), 360–392.

Paarsch, H. J. and B. S. Shearer (1999). The Response of Worker Effort to Piece Rates: Evidence from the British Columbia Tree-Planting Industry. *Journal of Human Resources 34*(4), 643–667.

Paarsch, H. J. and B. S. Shearer (2009). The response to incentives and contractual efficiency: Evidence from a field experiment. *European Economic Review 53*(5), 481–494.

Park, J. (2016). Will We Adapt? Temperature Shocks, Labor Productivity, and Adaptation to Climate Change in the United States (1986–2012). Discussion Paper 2016–81, Harvard Project on Climate Agreements.

Perla, J. and C. Tonetti (2014). Equilibrium Imitation and Growth. *Journal of Political Economy 122*(1), 52–76.

Pilcher, J. J., E. Nadler, and C. Busch (2002). Effects of hot and cold temperature exposure on performance: A meta-analytic review. *Ergonomics 45*(10), 682–698.

Pimentel, D. (1992). Energy Inputs in Production Agriculture. In R. C. Fluck (Ed.), *Energy in Farm Production*, Volume 6 of *Energy in World Agriculture*. Elsevier Science Publishing Company.

Prendergast, C. (1999). The Provision of Incentives in Firms. *Journal of Economic Literature 37*(1), 7–63.

Press, A. (1960). Lutherans Unite 3 Denominations; Rift Along Nationality Lines Is Closed in Merger of 2 1/4 Million Members. *The New York Times*.

Ramsey, J. D. (1995). Task performance in heat: A review. *Ergonomics 38*(1), 154–165.

Robbins, L. (1930). On the Elasticity of Demand for Income in Terms of Effort. *Economica* (29), 123–129.

Rogers, E. M. (1995). *Diffusion of Innovations* (4th ed.). The Free Press.

Rubin, D. K. and J. M. Perloff (1993). Who Works for Piece Rates and Why. *American Journal of Agricultural Economics 75*(4), 1036–1043.

Russel, D. A. and G. G. Williams (1977). History of Chemical Fertilizer Development. *Soil Science Society of America Journal 41*, 260–265.

Salamon, S. (1992). *Prairie Patrimony: Family, Farming, and Community in the Midwest*. The University of North Carolina Press.

San Diego Farm Bureau (2016). Harvest Calendar. `https://www.sdfarmbureau.org/BuyLocal/Harvest-Calendar.php`. Accessed: August 31, 2016.

Sarmiento, C., W. W. Wilson, and B. Dahl (2012). Spatial competition and ethanol plant location decisions. *Agribusiness 28*(3), 260–273.

Satyanath, S., N. Voigtlaender, and H.-J. Voth (2017). Bowling for Fascism: Social Capital and the Rise of the Nazi Party. *Journal of Political Economy 125*(2), 478–526.

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.

Seppänen, O., W. J. Fisk, and Q. H. Lei (2006). Effect of Temperature on Task Performance in Office Environment. Lawrence Berkeley National Laboratory Working Paper LBNL–60946.

Shapiro, C. and J. E. Stiglitz (1984). Equilibrium Unemployment as a Worker Discipline Device. *The American Economic Review 74*(3), 433–444.

Shearer, B. (2004). Piece Rates, Fixed Wages and Incentives: Evidence from a Field Experiment. *Review of Economic Studies 71*(2), 513–534.

Sheldon, M. (2016). Income Targeting and the Ridesharing Market. Working Paper.

Shi, L. (2010). Incentive Effect of Piece-Rate Contracts: Evidence from Two Small Field Experiments. *The B.E. Journal of Economic Analysis & Policy 10*(1). Article 61.

Smil, V. (2000). *Enriching the Earth: Fritz Haber, Carl Bosch, and the Transformation of World Food Production*. The MIT Press.

Smith, R. (2011). Starting an Apple Orchard In North Dakota. `http://www.ag.ndsu.edu/pubs/plantsci/hortcrop/h1547.pdf`.

St. John Evangelical Lutheran Congregation (2014). Homepage. `http://www.stjohnschwer.com/our-history`.

St. Mark's Lutheran Church (2014). Church History. `http://stmarkschurchsp.org/about-st-marks/church-history.html`.

Stiglitz, J. E. (1975). Incentives, Risk, and Information: Notes Towards a Theory of Hierarchy. *The Bell Journal of Economics 6*(2), 552–579.

Sudarshan, A., E. Somanathan, R. Somanathan, and M. Tewari (2015). The Impact of Temperature on Productivity and Labor Supply: Evidence from Indian Manufacturing. Delhi School of Economics Centre for Development Economics Working Paper 244.

Swierenga, R. P. (1997). The Little White Church: Religion in Rural America. *Agricultural History 71*(4), 415–441.

The American Lutheran Church (1960). *1960 Yearbook of the Uniting Churches: American Lutheran Church, the Evangelical Lutheran Church, United Evangelical Lutheran Church*. Augsburg Publishing House.

The Association of Religion Data Archives (2006). Churches and Church Membership in the United States, 1952 (Counties).

Thomas, M. A., B. A. Engel, and I. Chaubey (2009). Water quality impacts of corn production to meet biofuel demands. *Journal of Environmental Engineering 135*(11), 1123–1135.

Thoreson, J. (2013). Personal communication between Andrew Stevens and Joel Thoreson (ELCA). Email.

Trautmann, N. M., K. S. Porter, and R. J. Wagenet (1998). Modern Agriculture: Its Effects on the Environment. Fact Sheet.

Treble, J. G. (2003). Intertemporal Substitution of Effort: Some Empirical Evidence. *Economica 70*(280), 579–595.

Trinity Lutheran Church (2012). Trinity Lutheran Church: Church History. `http://www.newhamptoniowa.com/trinity/history.htm`.

Turnquist, A., T. R. Fortenbery, and J. Foltz (2008, July). Progress or devastation? The effects of ethanol plant location on local land use. In *Proceedings of the American Agricultural Economics Association Annual Meeting*, Orlando, FL.

United Lutheran Church Laurel (2013). About United Lutheran Church. `http://www.ulc-laurel.org/aboutus.htm`.

United States Department of Agriculture (2014). USDA Definition of Specialty Crop. `https://www.ams.usda.gov/sites/default/files/media/USDASpecialtyCropDefinition.pdf`. Accessed: September 3, 2016.

University of Minnesota Extension (1960). *Field Research in Soil Science*.

U.S. Census Bureau (1967). *Census of Agriculture, 1964: Data-Collection Forms and Procedures for Census and Related Surveys*, Volume 3. U.S. Government Printing Office.

U.S. Department of Agriculture (2016). Chemical use survey.

Vasilaky, K. (2013). Female Social Networks and Farmer Training: Can Randomized Information Exchange Improve Outcomes? *American Journal of Agricultural Economics 95*(2), 376–383.

Vasilaky, K. and K. L. Leonard (2014). As Good as the Networks They Keep?: Improving Farmers' Social Networks via Randomized Information Exchange in Rural Uganda. Unpublished working paper.

Welch, P. and J. J. Mueller (2001). The Relationships of Religion to Economics. *Review of Social Economy 59*(2), 185–202.

Wolf, R. C. (1966). *Documents of Lutheran Unity in America*. Fortress Press.

Wright, C. K. and M. C. Wimberly (2013). Recent land use change in the western corn belt threatens grasslands and wetlands. *Proceedings of the National Academy of Sciences 110*(10), 4134–4139.

Wyndham, C. H., N. B. Strydom, J. F. Morrison, C. G. Williams, G. A. G. Bredell, J. S. Maritz, and A. Munro (1965). Criteria for physiological limits for work in heat. *Journal of Applied Physiology 20*(1), 37–45.

Yellen, J. L. (1984). Efficiency Wage Models of Unemployment. *The American Economic Review 74*(2), 200–205.

Young, R. D. and N. L. Hargett (1984). Fluid Fertilizers: History, Growth, and Status. In R. Mikkelsen (Ed.), *TVA Fluid Fertilizers*. Tennessee Valley Authority.

Appendix A

A Model of Optimal Piece Rate Wages

In this appendix, I develop a simple model of a firm employing workers under a piece rate wage. This framework generates comparative statics on the firm's optimal piece rate wage with respect to output prices, resource abundance, and environmental conditions, which will allow me to predict how rising temperatures will affect the California blueberry industry.[1]

Consider a firm that produces output by employing labor to harvest some resource. In the short-run, the total amount of the resource, B, is fixed. Let θ be the abundance or density of the resource, such that $B = b\theta$ for some constant b. Because B is fixed in the short-run, the firm's only discretionary input is labor. The firm chooses both the number of workers to hire, n, and a piece rate wage r. The workers then each produce output X, which the firm sells at market price p. The firm is assumed to be a price-taker in the output market, but to have at least some monopsony power in the labor market to set r.[2] Also, if there exists a minimum hourly wage below which laborers cannot be paid, this constraint is assumed not to bind in equilibrium (i.e., all laborers working for the piece rate wage end up earning more than they would under the hourly minimum wage).

Employed laborers observe the piece rate wage r, the abundance of the resource θ, and an environmental condition T. They then endogenously select an unobserved level of effort that results in output $X(r, \theta, T)$ as described in section 1.2.1.[3] Laborers' supply of output, X, is increasing in r, since higher piece rate wages incentivize higher levels of effort holding θ and T constant. X is also increasing in θ, since a more abundant resource is easier to harvest.

[1] Heal and Park (2014) have previously considered the relationship between temperature and effective labor supply, deriving theoretical predictions based on worker-level micro-foundations. My model differs by directly accounting for wages that may also adjust to changing environmental conditions.

[2] In the California blueberry market, there is certainly competition between farms in the labor market. However, different farms provide different work environments and require workers to perform different physical tasks, meaning potential workers have *a priori* preferences over different possible employers. As long as a farmer offers a wage that is considered relatively comparable to what other farmers offer, conditional on the nature of the work, the farmer has some latitude to choose the particular wage level.

[3] Laborers may also differ in individual ability or skill ϕ. If skill simply affects optimized production multiplicatively, (i.e., $\phi X(r, \theta, T)$), then it will have no impact on the results derived in this model. My empirical strategy controls for individual fixed effects to address this concern.

The effect of T on laborers' supply of output is ambiguous *a priori*, and depends both on the environmental condition in question and its level.[4] Finally, the workers' productivity function is assumed to be concave in r. To summarize, $X_r > 0$, $X_\theta > 0$, and $X_{rr} < 0$. The firm is assumed to be sufficiently knowledgeable about its production process that it knows the function $X(r, \theta, T)$.[5]

I assume the firm must choose n before choosing r, and that the realization of n is somewhat stochastic. This closely matches reality in the California blueberry industry: a farmer considers how many berries will be ripe on the following day, and determines the ideal number of pickers he would like to hire. The farmer then calls workers directly or calls a labor contractor to coordinate the right number of pickers to show up the next day. However, the number of pickers who arrive the next morning may not exactly match the farmer's expectations. Once the farmer observes how many pickers show up, he *then* sets the daily piece rate wage. I further assume the farmer pays some constant per-worker cost h, to reflect various managerial costs that scale with the size of the farm workforce.

The firm faces the following "day-ahead" profit maximization problem:

$$\max_{n,r,\lambda} \quad (p - r)nX - hn + \lambda(b\theta - nX), \tag{A.1}$$

which implicitly gives an expected (non-binding) piece rate wage $\mathbb{E}[r^*]$ and the following optimal n:

$$n^* = \frac{b\theta}{\sqrt{hX_r}}. \tag{A.2}$$

Here, the labor participation constraint is $\mathbb{E}[r^*]X(\mathbb{E}[r^*]) > w_0$ where w_0 is a worker's daily reservation wage. Note that in equation (A.2), the optimal number of workers is mainly driven by resource abundance θ and labor costs h.

Once workers arrive for a day's work, n (which may differ somewhat from n^*) is fixed. The farmer then sets r to maximize daily profits:

$$\max_{r} \quad (p - r)nX - hn, \tag{A.3}$$

[4] If T is a measure of air pollution, one would expect $X_T < 0$ since pollution reduces both the productivity of laborers' effort as well as their willingness to provide it (Graff Zivin and Neidell, 2012). Similarly, if T is temperature, one would expect $X_T < 0$ for hot temperatures (Adhvaryu et al., 2016b; Sudarshan et al., 2015). However, at low temperatures, additional heat may be welcome and actually increase laborers' supply of output: $X_T > 0$ (Seppänen et al., 2006; Meese et al., 1984).

[5] In my empirical setting, this is close to true. Farmers understand particularly well the relationship between the density of ripe berries in the field and worker output: $X(\theta)$. Additionally, employers in other piece rate settings clearly acknowledge the negative effect of temperature on output: "Managers claimed that during the hottest months, daily wage workers preferred to go home to their villages... rather than work under the much more strenuous conditions at the factory. Some owners said they were actively considering the possibility of combating this preference for less taxing work by temporarily raising wages through a summer attendance bonus" (Sudarshan et al., 2015, p. 44).

which gives the following first-order condition:

$$(p - r)X_r - X = 0. \tag{A.4}$$

Equation (A.4) implicitly defines an optimal piece rate wage as a function of three exogenous parameters: output price, resource abundance, and the environmental characteristic: $r(p, \theta, T)$. Differentiating equation (A.4) once more by r gives the following second-order condition:

$$(p - r)X_{rr} - 2X_r < 0 \tag{A.5}$$

where the inequality follows from the earlier assumptions that $X_{rr} < 0$ and $X_r > 0$.

Now, I investigate the comparative statics of the optimal piece rate wage $r(p, \theta, T)$. I begin by analyzing the effects of output price p on r. Totally differentiating equation (A.4) with respect to p and rearranging gives:

$$\frac{dr}{dp} = \frac{-X_r}{(p - r)X_{rr} - 2X_r} > 0 \tag{A.6}$$

where the inequality comes from the facts that the denominator in equation (A.6) is simply the second-order condition from expression (A.5), and that X_r is assumed to be positive.

Next, I consider resource abundance θ. Totally differentiating equation (A.4) with respect to θ and rearranging gives:

$$\frac{dr}{d\theta} = \frac{X_\theta - (p - r)X_{r\theta}}{(p - r)X_{rr} - 2X_r}. \tag{A.7}$$

Again, note that the denominator in equation (A.7) is the second-order condition from expression (A.5), and therefore negative. Thus, the sign of $dr/d\theta$ depends on the numerator of equation (A.7). In particular,

$$\frac{dr}{d\theta} < 0 \iff X_{r\theta} < \frac{X_\theta}{p - r}. \tag{A.8}$$

This gives the sufficient condition that if $X_{r\theta} < 0$, then $dr/d\theta < 0$. The condition $X_{r\theta} < 0$ implies that the marginal increase in output induced by an increase in r is decreasing in the resource's abundance. This may or may not be a reasonable assumption.[6]

Finally, I consider the environmental condition T. Totally differentiating equation (A.4) with respect to T and rearranging gives:

$$\frac{dr}{dT} = \frac{X_T - (p - r)X_{rT}}{(p - r)X_{rr} - 2X_r}. \tag{A.9}$$

[6]In my empirical setting, I observe $dr/d\theta < 0$.

As before, the denominator in equation (A.9) is the second-order condition from expression (A.5) and thus negative. Consequently, the sign of dr/dT depends on the numerator of equation (A.9). In particular,

$$\frac{dr}{dT} > 0 \iff X_T < (p-r)X_{rT}. \tag{A.10}$$

This implies that, in general, the sign of dr/dT is ambiguous and therefore an empirical question.

In addition to the above exercises, I also analyze how changes in the three truly exogenous variables – p, θ, and T – affect farm profits Π:

$$\frac{d\Pi}{dp} = \frac{dr}{dp}\underbrace{((p-r)X_r - X)}_{=0 \text{ by envelope thm}} + 1 = 1 > 0 \tag{A.11}$$

$$\frac{d\Pi}{d\theta} = \frac{dr}{d\theta}\underbrace{((p-r)X_r - X)}_{=0 \text{ by envelope thm}} + (p-r)X_\theta = (p-r)X_\theta > 0 \tag{A.12}$$

$$\frac{d\Pi}{dT} = \frac{dr}{dT}\underbrace{((p-r)X_r - X)}_{=0 \text{ by envelope thm}} + (p-r)X_T = (p-r)X_T. \tag{A.13}$$

Equation (A.13) implies that the sign of $d\Pi/dT$ will match the sign of X_T. All three results match intuition.

The model outlined above produces two particularly interesting predictions. First, equation (A.10) suggests that the impact of environmental conditions (such as temperature) on the optimal piece rate wage r can be determined by estimating two effects – X_T and X_{rT} – over wide supports of r, θ, and T.[7] Second, equation (A.13) states that the effect of the environmental condition T on firm profits will have the same sign as X_T. Thus, by credibly estimating X_T, I am able to provide a prediction on how changing levels of T will affect firm profits.

[7]The model also highlights the need to control for r's endogeneity. In particular, equation (A.7) suggests that r will depend on resource abundance θ. However, equation (A.6) also suggests that r will be positively correlated with output price p, justifying my use of the market price p as an instrument in my empirical strategy.