UNIVERSITY OF NEVADA, RENO

Three Essays on Labor Supply Focusing on Entrepreneurship and Health Insurance

A dissertation submitted in partial fulfillment of
the requirements for the degree of Doctor of Philosophy in
Economics

by

Md. Mobarak Hossain

Sankar Mukhopadhyay / Dissertation Advisor

May, 2021

THE GRADUATE SCHOOL

We recommend that the dissertation

prepared under our supervision by

MD. MOBARAK HOSSAIN

entitled

Three Essays on Labor Supply Focusing on Entrepreneurship and Health Insurance

be accepted in partial fulfillment of the

requirements for the degree of

DOCTOR OF PHILOSOPHY

Sankar Mukhopadhyay, Ph.D., Advisor

Frank M. Fossen, Ph.D., Committee Member

Todd Sørensen, Ph.D., Committee Member

Mark D. Packard, Ph.D., Committee Member

Tomasz J. Kozubowski, Ph.D., Graduate School Representative

David Zeh, Ph.D., Dean, Graduate School

May, 2021

Abstract

This dissertation consists of three essays analyzing policies that affect different aspects of labor supply. Economists are interested in many different aspects of labor supply decisions, such as whether to participate in the labor market, amount of work conditional on participation and career choice, among others. At the same time, a wide range of policies can impact labor supply such as: tax, immigration, healthcare, and unionization policies, among others. Analyzing policies using appropriate methodology may assist policymakers by presenting alternative solutions, trade-offs, and intended and unintended consequences of economic policies. While some of these issues have been explored extensively, others remain relatively unexplored. In this dissertation, I explore three such unexplored problems. First, I analyze the dynamic decision processes of entrepreneurs. Second, I explore the impact of the cost of health insurance on entrepreneurial activities. Third, I evaluate the effect of the availability of health insurance on workplace absenteeism.

The first essay explores what induces individuals to become entrepreneurs creating jobs. Extant structural labor supply models used for ex-ante policy evaluations mostly exclude entrepreneurs. The first essay develops and estimates the

first dynamic structural micro-econometric model explicitly accounting for the employer and non-employer entrepreneurs. In the model, individuals in each period choose to work as an employee, as one of the two entrepreneur types, or be non-employed. Different types of work experiences may affect earnings in the three careers in different ways. The model, estimated using German survey data, replicates key data patterns. This essay simulates how policy scenarios would affect individuals' choices to become employers and non-employers.

The second essay explores whether the cost of health insurance affects entry into entrepreneurship. Entrepreneurship literature argues that a lack of access to health insurance is a potential barrier to become an entrepreneur (entrepreneurship lock), especially for individuals with a chronic health condition. Several papers have explored whether the guaranteed availability of health insurance brought about by the Affordable Care Act (ACA) of 2010 has increased the level of entrepreneurship, with conflicting results. However, the current literature focuses only on the availability of health insurance but not the cost of obtaining insurance. This essay explores whether the cost of health insurance rather than availability is a barrier to entrepreneurship. The results suggest that the probability of entry into self-employment is not sensitive to health insurance premiums.

The third essay examines whether the expansion of health insurance coverage brought on by the Patient Protection and Affordable Care Act of 2010 (ACA), led to a decline in absenteeism among overweight and obese individuals. This essay uses data from the National Health Interview Survey (NHIS) to compare absenteeism among overweight and obese workers to absenteeism among

normal-weight workers before and after the ACA. Results suggest that in the post-ACA period, the probability of being absent declined by about 1.3 (1.5) percentage points among obese (overweight) individuals. Disaggregated regressions suggest that the effect is significant among women but not among men. Furthermore, estimates (using a Tobit model) indicate that the obese (overweight) workers missed 0.32 (0.48) fewer days after the ACA. Again, the effect is concentrated among women. Results show that improved health outcomes led to reduced absenteeism. Results also show that there is no decline in absenteeism among elderly (age>=65) adults (who did not experience any increase in health insurance coverage as a result of the ACA), suggesting that the decline in absenteeism is indeed due to the expansion of health insurance coverage due to the ACA. Estimates of this essay imply that the ACA reduced the cost associated with absenteeism by about $350 million per year.

This dissertation is dedicated to my parents and children.

Acknowledgements

First and foremost, I would like to appreciate my advisor, Dr. Sankar Mukhopadhyay, who prepared me to complete this incredible journey. I could never imagine having a better advisor and mentor than Dr. Mukhopadhyay. He helped, encouraged, supported, and guided me with great patience which accelerated me to complete my graduate study in the department of Economics at the University of Nevada, Reno (UNR). I am very grateful to him. I am also grateful to the graduate school representative, Dr. Tomasz J. Kozubowski, for his great support in this journey. I also express my gratitude to my dissertation committee member, Dr. Frank M. Fossen, who not only gave me the best suggestions but also carefully helped me to complete my dissertation. I would not be where I am today without his effort. A special thanks also goes to my dissertation committee member and placement director, Dr. Todd Sorensen, who supported me academically and spiritually to complete my graduate study at UNR. I would like to thank Dr. Mark D. Packard for managing his valuable time to serve on my dissertation committee. I also would like to acknowledge my family's continuous support and encouragement. Moreover, I am very much thankful to all the faculty, graduate students, and staff in the department of Economics at UNR.

Contents

List of Tables

List of Figures

Chapter 1

Employer and Non-employer Entrepreneurs: A Dynamic Structural Microeconometric Model

(with Frank M. Fossen & Sankar Mukhopadhyay)

1.1 Introduction

The interplay between the accumulation of work experience, returns in form of labor earnings, and the corresponding incentives for labor supply over the life cycle has been studied intensively for wage and salary workers. However, we know little in this context about the roughly ten percent of the labor force who are entrepreneurs, although entrepreneurial activity is crucial for job creation, innovation, and growth. In particular, policymakers are eager to understand the individual decision to hire workers and to become an employer-entrepreneur: individual who starts business and hires employee(s). Employer-entrepreneurs exhibit higher growth ambitions, are more likely to be innovators, and therefore have a stronger impact on the economy than non-employer-entrepreneurs: individuals who start business but do not hire employee(s).

The values of work experience gained from being an employee, an employer, or a non-employer, may be different in each of these employment states. It is important to understand how experience from the different employment states is valued in the same and other states. The structure of returns to experience within and across labor market activities will determine the dynamics of individual decisions to supply labor as employees, employers and non-employers over the life cycle. Understanding how incentives drive these dynamics will not only advance the literatures on labor supply and entrepreneurship, but also help policymakers to design policies that influence entrepreneurship and job creation, e.g. through tax policy, social insurance or subsidies.

In this chapter, we develop a dynamic structural microeconometric model of labor supply over the life cycle with the choice options of being a non-employer-entrepreneur, an employer-entrepreneur, an employee, or non-employed. Earnings in the different alternatives depend on types of experience gained from the different employment forms. In our model individuals' preferences may be different in both observable and unobservable ways. This approach allows us to take into account selection into different careers due to observable and unobservable characteristics.

To estimate our dynamic structural model, we use the world's largest household panel survey that includes annual information over a sufficiently long time period and distinguishes between employer and non-employer entrepreneurs, the German Socio-Economic Panel Survey (SOEP). The estimated dynamic structural model fits key properties of the representative data well, such as the age-participation and age-earnings profiles of employees, non-employers, and employers. We use the estimated model to simulate the effects of hypothetical tax or subsidy policies on entrepreneurial activity.

Our main contribution to the literature is that we provide the first dynamic structural microeconometric model of labor supply that takes into account non-employer entrepreneurship and employer-entrepreneurship. Our results suggest that transferability of experience between sectors is one way. In particular, returns to work experience accumulated as an entrepreneur are negligible or even negative for employees. However, experience accumulated as an employee substantially increases the earnings of an entrepreneur. Our policy simulations are examples of the wide range of applied research questions that our estimated model is suitable to address.

The literature on dynamic structural labor supply models (e.g., Keane and Wolpin, 1997; Haan and Prowse, 2014, using the SOEP) has tremendously improved our understanding of labor market dynamics and lays the methodological foundation for this chapter. However, this literature mostly ignores entrepreneurship and usually excludes the entrepreneurs from the estimation samples. It does not address the possibility of selection bias arising from dropping roughly ten percent of the working population who make this choice at each point in time.

Parallel work by Hincapié (2020) and working papers by Dillon and Stanton (2017) and Humphries (2018) are notable exceptions. These authors include self-employment in dynamic structural models using different data for different countries, but do not distinguish between non-employers and employers. Given the importance of job creation and growth orientation in the policy debate on entrepreneurship, this distinction is an important contribution we make. Hurst and Pugsley (2011) report that most self-employed individuals without employees have no intention to grow or to hire workers. Therefore, modeling the self-employed with or without employees the same way can result in misleading policy conclusions for policymakers intending to promote entrepreneurship in order to stimulate economic growth and jobs.[1] In fact, in his conclusion, Hincapié (2020) calls for further research investigating which entrepreneurs hire workers to shed light on job creation. In this chapter, we document that employers and non-employers are very different.

An emerging literature investigates the determinants of becoming an employer, without estimating dynamic structural models. Haltiwanger et al. (2013) identify firm age as a crucial determinant of the decision to hire workers. Consistent with this, Fairlie and Miranda (2017) document that many entrepreneurs start solo and decide to hire their first employee within the first three years after start-up. These findings underline the importance of dynamic choices in this context, which is an important motivation for our chapter. Caliendo et al. (2019) test the effects of various individual characteristics on the choices to switch between non-employer and employer entrepreneurship, paid employment, and non-employment, in a reduced form approach using the SOEP. The results from

[1] The extent papers partially distinguish between entrepreneurs with incorporated and unincorporated businesses. This approach of capturing heterogeneity among entrepreneurs has important limitations, as the decision to incorporate depends on the current legal context around liability issues and specific tax rules that differ across countries and US states and often change over time.

this paper guide us to identify relevant variables for individual heterogeneity in our structural model. Åstebro and Tåg (2017) report that high-ability individuals are more likely to create jobs than low-ability individuals. This is in line with Caliendo et al. (2019), who find that higher education levels increase the probability of becoming an employer-entrepreneur, and consistent with our structural estimation results in this chapter. In contrast to our structural model, the results from these reduced form estimations cannot be used for ex-ante simulations of the effects of hypothetical policies on entrepreneurship.

A small literature proposes static structural or semi-structural models of entrepreneurship (e.g., Rees and Shah, 1986; Fossen, 2009; Wen and Gordon, 2014). Some of these models include lifetime earnings as an input into a utility function, but the decision to be an entrepreneur is assumed to be a static decision in these papers. We argue that it is important to model entrepreneurship in a dynamic programming framework because we observe in the data that most entrepreneurs start working as paid employees and switch to entrepreneurship many years later. We also frequently observe transitions back to paid employment after working as an entrepreneur. The existing static models cannot capture any of these dynamics, whereas our model fits the observational patterns well and provides a rationale for them. This makes policy simulations based on our estimated dynamic model more plausible and reliable. None of the structural models provided in the literature distinguish between non-employer and employer entrepreneurs, whereas we include the dynamics of the important decisions to hire and to keep workers.

This chapter proceeds as follows. Section 1.2 lays out the dynamic structural microeconometric model and Section 1.3 introduces the data. Section 1.4 presents the estimation results and Section 1.5 policy simulations. Section 1.6 concludes the analysis.

1.2 Dynamic Structural Model of Entrepreneurship

Our model focuses on four mutually exclusive and exhaustive choices for individuals. The four options to choose from as career are ($k = 1$) non-employer-entrepreneurship, ($k = 2$) employer-entrepreneurship, ($k = 3$) employed, and ($k = 4$) non-employed. Each individual's life span is finite, i.e. the individual optimization starts at age 18 and ends at age $t = T$. Individual i can choose any of the four options at any age t, which will maximize the sum of their current and expected future utility until the end of the life span. Suppose $d_{itk} = 1$ if alternative k is chosen by individual i at current age t, where $k = 1, 2, 3, 4$. We take into account unobserved heterogeneity by modeling three discrete unobserved types; $j(i) = 1, 2, 3$ indicates the type of individual i. The current period alternative-specific utility function for each individual is given by

$$\mathrm{U}_{it}^{k}(\cdot) = \begin{cases} \dfrac{c_{itk}^{1-\rho_{j(i)}}}{1-\rho_{j(i)}}; & \text{if } k = 1, 2, 3;\ \rho > 0 \\ \beta_{4,j(i)} + \gamma q_i; & \text{if } k = 4, \end{cases} \tag{1.1}$$

where consumption $c_{itk} = \max\left(w_{itk} - \tau_{itk},\, 0\right)$, and w_{itk} and τ_{itk} capture the annual earnings and the income tax liability of an individual, respectively. The coefficient of constant relative risk aversion, ρ, depends on unobserved types.

In Germany, unemployed individuals receive unemployment benefits for usually the first year of an unemployment spell, but only lower unemployment assistance thereafter. Therefore, utility for the non-employed individuals depends on two components, fixed utility from non-employment depending on type, $\beta_{4,j(i)}$, which includes utility from leisure and from unemployment assistance, and one-time unemployment benefit, γ. The indicator q_i equals one at the beginning of an individual's unemployment spell and zero thereafter.

Individuals' earnings from labor ($k = 1, 2, 3$) are stochastic and given by Mincer-type earnings functions as follows:

$$\begin{aligned}\ln(w_{itk}) =& \beta_{k,j(i)} + \alpha_{1k}e_{1,i,t-1} + \alpha_{2k}e^2_{1,i,t-1} + \alpha_{3k}e_{2,i,t-1} + \alpha_{4k}e^2_{2,i,t-1} \\ & + \alpha_{5k}e_{3,i,t-1} + \alpha_{6k}e^2_{3,i,t-1} + \xi_k \mathbb{1}(k_{it} \neq k_{i,t-1}) \\ & + \alpha_{7k}(\text{University degree}_{itk}) + \alpha_{8k}(\text{Academic track}_{itk}) + \epsilon_{itk},\end{aligned} \tag{1.2}$$

where ϵ_{itk} is a vector of serially uncorrelated jointly normally distributed shocks with mean 0 and variance-covariance matrix Σ. The indicator function, $\mathbb{1}$, takes value 1 if the choice in the previous period and the current choice are not the same. The coefficient ξ_k is the one period search and transition cost incurred by an individual. The two education dummies are one if an individual's highest educational attainment is a university degree or a high school leaving certificate that qualifies for university entry, respectively. A lower degree (vocational or no degree) is the omitted base category. The parameters $\alpha_{\nu k}$ represent the return to career-specific experience and educational attainment.

An individual's experience vector, e_{kit}, evolves according to

$$e_{kit} = e_{k,i,t-1} + d_{itk}. \tag{1.3}$$

We approximate the German progressive personal income tax schedule and social security contributions by estimating a regression of individual tax liabilities[2] as a non-linear function of before-tax annual earnings:

$$\begin{aligned}\tau_{itk} = & \pi_{0k} + \pi_{1k} w_{itk} + \pi_{2k} w_{itk}^2 + \pi_{3k} w_{itk}^3 + \pi_{4k} age + \pi_{5k} w_{itk} age \\ & + \pi_{6k} w_{itk}^2 age + \pi_{7k} w_{itk}^3 age + \epsilon_{ik}.\end{aligned} \tag{1.4}$$

We estimate this regression outside of the model; the estimated coefficients can be interpreted as tax parameters that capture the progressive tax schedule. The German personal income tax rules are mostly the same for income from different sources, in particular, from paid employment and self-employment, but social security contributions differ (the self-employed are exempt in most cases). We estimate separate tax regressions by employment state (employee, employer, non-employer) to allow for differences in effective taxation. This includes potentially better tax avoidance and evasion possibilities for the self-employed in comparison to employees (Kleven et al., 2011; Fossen et al., 2020). Although income taxes do not directly depend on age, the interaction terms with age capture individual circumstances that change with age on average,

[2] The SOEP survey asks respondents for before-tax and after-tax labor earnings in the month before the interview. We multiply by twelve to approximate annual labor earnings and calculate the tax liability as the difference.

such as marital status, number of children, and health expenses. These circumstances affect taxes, but are not modeled explicitly here.

In our model, an individual maximizes the present value of remaining lifetime utility at any age. Suppose $\mathrm{V}_{it}\left(\cdot|\Omega_{it}\right)$ is the value function of an individual with discount factor δ, set at 0.99. Ω_{it} represents the relevant components of the state space. The value function is given by

$$\mathrm{V}_{it}\left(\cdot|\Omega_{it}\right) = \max_{d_{itk}} \ \mathrm{E}\left[\sum_{s=t_0}^{T} \delta^{s-t_0} \sum_{k=1}^{4} \mathrm{U}_{is}^{k}\left(\cdot\right) d_{isk}|\Omega_{it}\right]. \tag{1.5}$$

Then the Bellman equation is given by

$$\begin{cases} \mathrm{V}_{i,t-1}^{k}\left(\cdot|\Omega_{i,t-1}\right) = \mathrm{U}_{i,t-1}^{k}\left(\cdot\right) + \delta \mathrm{E}\ \mathrm{V}_{it}\left(\cdot|\Omega_{it}\right); & \text{if } t < T \\ \mathrm{V}_{iT}^{k}\left(\cdot|\Omega_{it}\right) = \mathrm{U}_{iT}^{k}\left(\cdot\right); & \text{if } t = T. \end{cases} \tag{1.6}$$

We solve the Bellman equation by using backward recursion, beginning with the last period T. We use Monte Carlo integration to compute the multi-dimensional integrations necessary to calculate the expected value of the maximum of the alternative-specific value functions. We evaluate the value of the Emax function at every possible state point. The model is estimated by simulated maximum likelihood. [3]

Let O_{it} represent the outcomes (choices and earnings if individuals choose to work) of individual i at age t. Also, let I_i denote the set of initial conditions for that individual. Let $Pr(j(i) = 1|I_i)$ denote the type probability, which depends

[3] We only provide an outline of the solution and estimation methods since they have been described in detail elsewhere (Keane et al., 2011; Imai and Keane, 2004).

on initial conditions. The unobserved type is assumed to be known to the individual but not to the econometrician. The likelihood for individual i can be written as the product over the age-specific choice probabilities and the probability of observing the corresponding wage (if applicable), integrating over the unobserved type. Thus, the contribution to the likelihood of individual i is given by

$$L_i = \sum_{j=1}^{3} Pr\left(O_{it}|O_{i(t-1)}, ..., O_{it_0}; j(i) = j, I_i\right) Pr\left(j(i) = j|I_i\right). \tag{1.7}$$

To illustrate the calculation of the likelihood, suppose that the kth alternative is chosen by individual i, and we observe a wage at age t. The probability of observing that choice and wage combination outcome conditional on the state space (which includes $O_{i(t-1)}, ..., O_{it_0}, I_i$, and type) is

$$\begin{aligned} Pr\left(O_{it}|O_{i(t-1)}, ..., O_{it_0}; j(i) = j, I_i\right) & Pr\left(j(i) = j|I_i\right) \\ &= Pr\left(d_{itk} = 1, w_{it}|\Omega_{it}, I_i, j(i) = j\right) \\ &= Pr\left(d_{itk}, w_{it}|\Omega_{it}, I_i\right) f\left(w_{it}|\Omega_{it}, I_i, j(i) = j\right), \end{aligned} \tag{1.8}$$

where $f\left(w_{it}|\Omega_{it}, I_i, j(i) = j\right)$ is the wage density. The overall likelihood for $i = 1, ..., N$ individuals is the product over the individual likelihoods:

$$L = \prod_{i=1}^{N} L_i. \tag{1.9}$$

In our numerical implementation, we assume that there are three unobserved types, and that the type probabilities are multinomial logistic. In particular,

$$Pr(j(i) = 1|I_i) = \frac{e^{\eta I_i}}{1 + e^{\eta I_i} + e^{\zeta I_i}}, \tag{1.10}$$

$$Pr(j(i) = 2|I_i) = \frac{e^{\zeta I_i}}{1 + e^{\eta I_i} + e^{\zeta I_i}}, \tag{1.11}$$

$$Pr(j(i) = 3|I_i) = 1 - Pr(j(i) = 1|I_i) - Pr(j(i) = 2|I_i). \tag{1.12}$$

The vector of initial conditions I_i consists of a constant, a dummy variable indicating whether the respondent's father was self-employed when the respondent was 15 years old, the respondent's general willingness to take risk, locus of control, a dummy indicating whether the respondent lives in eastern Germany, and migration background. The corresponding coefficients are $\eta = \{\eta_0, \eta_1, \eta_2, \eta_3, \eta_4, \eta_5\}$ and $\zeta = \{\zeta_0, \zeta_1, \zeta_2, \zeta_3, \zeta_4, \zeta_5\}$. We allow risk preference (ρ), the constants in the choice-specific wage functions (β_k), and the value of leisure to vary across the unobserved types.

The model parameters enter the likelihood function through the choice probabilities that are computed from the solution of the dynamic programming problem. We calculate the derivatives of the log likelihood function numerically. To calculate numerical derivatives, we use a step size equal to 1% of parameter estimates. The maximization of the likelihood function iterates between solving the dynamic program and calculating the likelihood.[4] We use a subroutine called HOPSPACK (Plantenga, 2009), a hybrid optimization parallel search package developed by Sandia National Laboratories. This subroutine uses a

[4]We used 500 simulations for each individual to calculate the likelihood.

Generating Set Search (GSS) algorithm for optimization. To obtain the standard errors of the estimates we invert the average of the product of the score matrices. This is known as the BHHH estimator (Berndt et al., 1974).

1.3 Panel Data

To estimate our dynamic structural model, we use the German Socio-Economic Panel Survey (SOEP). This large household panel survey is representative for Germany (Goebel et al., 2019) and provides annual information on individuals, including non-employers and employers, over a sufficiently long time period. We use the waves from 2000 to 2016 (the survey was significantly enlarged in 2000). We focus on men at the prime working age between 18 and 57 years of age. This way, we can abstain from modeling non-participation of married mothers and early retirement decisions. Respondents who answer that their primary labor activity is self-employment are asked whether they have no employees (labeled as non-employer), or 1-9, or 10 or more (both labeled as employers). A non-employer could have one or more partners in a partnership business, who are not employees. We exclude individuals who have less than three consecutive observations. After excluding individuals with missing values in our variables, we work with an unbalanced panel (minimum length 3 and maximum length 17 years) with 50,190 observations from 6,313 individuals.

Table 1.1 presents the descriptive statistics for the four employment states. In our data, the average age varies from 39.12 for the non-employed to 43.75 years for employers. We group individuals into three categories according to their

Table 1.1: Summary Statistics

Variable	Non-employer	Employer	Employed	Non-employed
Highest educ. degree				
Vocational	0.50 (0.50)	0.48 (0.50)	0.65 (0.48)	0.75 (0.43)
Academic track	0.33 (0.47)	0.23 (0.42)	0.22 (0.41)	0.17 (0.38)
University degree	0.17 (0.38)	0.29 (0.46)	0.13 (0.34)	0.08 (0.27)
Experience (years)				
Non-employer	5.57 (4.23)	2.30 (2.47)	0.17 (0.78)	0.56 (1.69)
Employer	2.47 (3.32)	8.01 (5.37)	0.14 (0.94)	0.35 (1.46)
Employed	11.36 (6.71)	10.67 (5.49)	18.40 (9.13)	12.57 (9.27)
Earnings	40.15 (35.06)	67.38 (63.07)	38.29 (22.69)	
Age	42.94 (8.09)	43.75 (7.50)	40.75 (8.98)	39.12 (10.56)
Father entrepreneur	0.10 (0.30)	0.18 (0.39)	0.07 (0.25)	0.07 (0.26)
Willingness to take risk	5.83 (2.17)	5.96 (2.08)	5.11 (2.12)	5.24 (2.32)
Locus of control	29.38 (5.96)	31.43 (5.72)	28.82 (5.70)	25.93 (6.42)
East Germany	0.25 (0.43)	0.19 (0.40)	0.22 (0.42)	0.33 (0.47)
Migration background	0.13 (0.34)	0.14 (0.34)	0.17 (0.37)	0.24 (0.43)
N	2396	3327	39204	5263

Notes: The table shows sample means by career choice group. Earnings (≥ 0) are annual in real \$1000 in prices of 2005. Standard deviations in parentheses.

highest formal educational degree obtained: vocational school track (the higher secondary school degree “Abitur”, which qualifies for entry into university in Germany, was not obtained), academic school track (“Abitur” obtained), and university degree. The table shows that employers on average have the highest level of formal education and the non-employed the lowest. About 29 percent of employers have a university degree compared to only 8 percent of the non-employed.

As discussed above, we distinguish between different types of experience. Individuals who are currently working as employees have, on average, 18.40 years of experience as employees, but little experience from entrepreneurship. In contrast, individuals who are working as non-employers have, on average, 5.57 years of experience as non-employers and 2.47 years as employers, but only

11.36 years of experience as an employee. Similarly, employers have, on average, 2.30 years of experience as non-employers, 8.01 years as employers, and 10.67 years as an employee.

A number of other characteristics that may affect individual labor market decisions are used to determine the unobserved type probabilities. These variables include the self-reported general willingness to take risk on a Likert scale from 0 (completely unwilling) to 10 (completely willing)[5] and a measure of locus of control. Individuals have an internal locus of control if they believe that outcomes are the consequences of their own actions rather than of luck or fate (Rotter, 1966).[6] Furthermore, we include family and demographic background variables such as migration background, whether a respondent is living in the area of former East Germany, and whether his or her father was self-employed when the respondent was 15 years old.

1.4 Estimation Results

In this section, we report the estimated parameters, assess how the estimated model fits the data, and present simulation results. Table 1.2 (Panel A) shows the estimates of the parameters of the earnings functions for non-employers, employers, and employed individuals. The returns to the higher secondary school degree "Abitur" (academic track) are comparable for employer entrepreneurs

[5]Dohmen et al. (2011) show that this survey measure of risk attitudes is a good predictor of actual risk taking behavior.

[6]In the survey, respondents are asked to state how much they agree with ten statements about themselves on a Likert scale from 1 to 7. This short inventory is used to calculate a score, with a high score indicating an internal locus of control.

and employees, while the returns to a university degree are highest for employers. Transition cost is largest for non-employer entrepreneurs; the parameter is added to wages in Eq. 1.2 for individuals making a transition. Our estimate suggest that an individual who becomes a non-employer entrepreneur experiences a one-time cost amounting to 16 percent of the year's earnings. Panel B displays the estimated parameters in the utility function (Eq. 1.1). Estimates of the constant relative risk aversion parameters suggest that Type 1 individuals are different from Type 2 and Type 3 individuals in terms of their risk preferences.

Table 1.2: Estimates of the Parameters of the Earnings and Utility Functions

Panel A: Earnings Functions				
Parameter	Variable	Non-employers ($k=1$)	Employers ($k=2$)	Employed ($k=3$)
$\beta_{k,1}$	Constant type 1	0.7666 (2.9539)	0.5154 (4.6325)	0.8601 (0.3172)
$\beta_{k,2}$	Constant type 2	1.7466 (2.9988)	2.1483 (4.0983)	2.9217 (0.1787)
$\beta_{k,3}$	Constant type 3	1.3045 (2.9358)	1.1937 (4.6901)	2.3410 (0.2061)
α_{1k}	Exp. non-employer linear	0.2370 (0.1198)	0.0559 (0.1234)	-0.0711 (0.0231)
α_{2k}	Exp. non-employer squared	-0.0120 (0.0006)	-0.0065 (0.0009)	0.0035 (0.0003)
α_{3k}	Exp. employer linear	0.0814 (0.1437)	0.3240 (0.0837)	-0.0040 (0.0308)
α_{4k}	Exp. employer squared	-0.0037 (0.0011)	-0.0155 (0.0004)	0.0002 (0.0002)
α_{5k}	Exp. employee linear	0.0811 (0.1138)	0.0752 (0.1678)	0.0820 (0.0075)
α_{6k}	Exp. employee squared	-0.0042 (0.0003)	-0.0043 (0.0005)	-0.0019 (1.82e-5)
α_{7k}	Academic track	0.2571 (0.5725)	0.3323 (0.6371)	0.3495 (0.0214)
α_{8k}	University degree	0.4197 (0.8412)	0.7153 (0.6806)	0.6211 (0.0344)
ξ_k	Transition cost	-0.1623 (0.0792)	-0.0389 (0.0316)	-0.0455 (0.0037)
Panel B: Utility Functions				
Parameter	Variable	Type 1 ($j=1$)	Type 2 ($j=2$)	Type 3 ($j=3$)
ρ_j	Constant Relative Risk Aversion coefficient	0.7604 (0.0066)	1.2528 (0.0030)	1.2616 (0.0007)
$\beta_{4,j}$	Constant utility of non-employed individuals	6.8225 (1.7138)	-2.0007 (0.4038)	-2.0006 (0.0937)
γ	One-period unemployment benefit (for all types)		0.0028 (0.0005)	

Notes: Panel A shows the estimated parameters of the earnings functions corresponding to the different choice alternatives ($k = 1, 2, 3$ represent non-employers, employers, and employed individuals, respectively). $\beta_{k,j}$ is the career-specific constant in the Mincer-type earnings function, which depends on unobserved types ($j = 1, 2, 3$). The parameters $\alpha_{\nu k}$ represent the return to career-specific experience and education categories relative to vocational degree. The parameters ξ_k represent one period search and transition cost.
Panel B shows the estimates of the parameters of the utility functions The CRRA coefficients and the constant utility of non-employed individuals depend on the unobserved type, but the one-period unemployment benefit parameter does not.
Standard errors at 1% deviation are in parentheses.

Further parameter estimates appear in the Appendix. Table 1.A.1 presents the estimates of the parameters of the determinants of the type probabilities. Table 1.A.2 shows the estimates of the variance co-variance matrix. Table 1.A.3 displays the estimated regression coefficients of the individual tax liability function. The coefficients reflect the nonlinear schedules of personal income taxes and social security contributions in Germany, tax-relevant circumstances that on average change with age, and differences in effective taxation between employees and the self-employed. Evaluated at the average age, the differences between the tax functions for the different groups are small.[7]

To assess how well the estimated model fits the data, we use it to simulate various statistics and compare them with sample statistics. Columns 2-3 in Table 1.3 compare the actual with the simulated participation rates in the four different employment states, averaged over the whole life-cycle. To generate these columns, we simulate[8] choices for all the individuals in our sample, starting with their initial conditions. Overall, the simulated model is able to replicate the career choice pattern in the data, although the participation rate of non-employer entrepreneurs is smaller in our simulations than what we observe in the data, and the non-employment rate is somewhat larger.

Before discussing the model fit further, we point out the importance of modeling unobserved heterogeneity. Estimates in Columns 4-6 in Table 1.3 suggest that 7.6 percent of individuals in our sample are Type 1, 47.5 percent are Type 2, and 44.9 percent are Type 3. To compare and contrast the types, in this table we simulate the participation probabilities in the different career choices

[7] Plots are available from the authors on request.
[8] Each individual has been simulated 1000 times at each time point over the life span.

Table 1.3: Participation Rates: Actual, Simulated, and Type Specific Simulated

	Actual	Simulated	Type 1	Type 2	Type 3
Non-employer	4.77	2.44	10.96	1.41	2.69
Employer	6.63	5.57	5.07	7.34	3.58
Employed	78.11	78.32	23.78	84.59	79.75
Non-employed	10.49	13.67	60.19	6.66	13.99
Share of type	100%	100%	7.64%	47.46%	44.90%

Notes: Columns 2-3 compare the actual and simulated (considering unobserved heterogeneity) participation rates. Columns 4-6 show the simulated participation rates pretending that everybody was a specific type. The last row shows the estimated shares of the specific types in the sample. The total number of observations is 50,190.

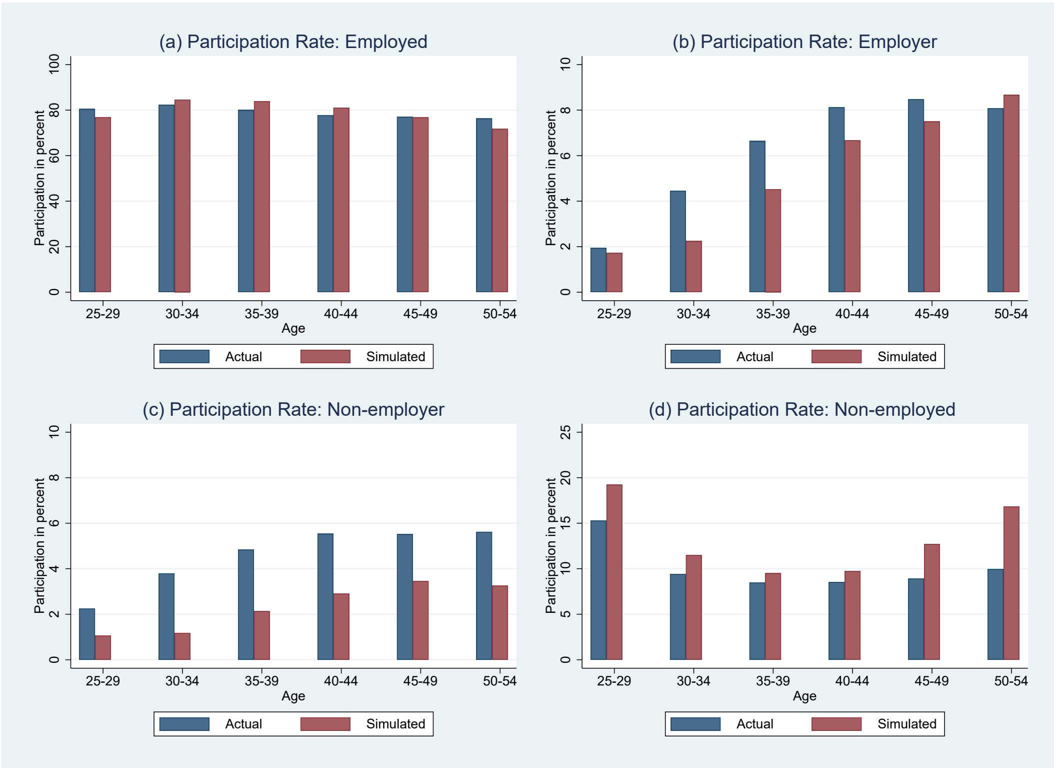
that would occur if everybody in the sample was of one particular type. For example, Type 1 individuals (7.6 percent of our sample) spend most of their time in non-employment (60.2 percent vs. 13.7 percent in the combined sample). They are also substantially more likely to be non-employer entrepreneurs compared to the other types. Type 2 individuals are less likely to be non-employed and more likely to be employer entrepreneurs or employees compared to other types.

Next, we compare the life cycle profiles of career choices. We show the model fit in five-year age groups.[9] In Germany, the share of employees declines slightly from more than 80 percent in the early thirties to below 80 percent above 40 years of age, as shown in panel (a) in Figure 1.1. The model simulation replicates this decreasing pattern. In addition, panel (b) (Figure 1.1) shows that the percentage of individuals who are employer-entrepreneurs increases with age from about 2 percent among 25-29 year old men to about eight percent by their early forties and then remains stable. The simulated data replicates both the

[9]We begin at age 25 since the number of entrepreneurs is small (less than one percent) in the early 20's and the focus of this chapter is on entrepreneurship.

increase with age and then the leveling off. However, in the early part of the life-cycle, the predicted rate of employer-entrepreneurship is lower in the simulations than in the observed data.

Figure 1.1: Participation Fit By Five Years Age Band



Our model is less successful in replicating the percentage of men who choose non-employer entrepreneurship. In the data, the percentage of individuals who are non-employers increases with age from about two percent among 25-29 year old men to almost six percent above 40 years, when the participation rate levels off (panel (c) in Figure 1.1). Our simulated data replicates this pattern, but the percentage of individuals in non-employer entrepreneurship in the simulated data increases from about one percent to little under four percent. Panel (d) (Figure 1.1) shows the participation rates of the non-employed individuals.

Figure 1.2: Earnings Fit By Five Years Age Band

Table 1.4: One Year Transition Matrix - Actual Versus Simulated

	Non-employers	Employers	Employed	Non-employed
Panel A: Actual				
Non-employers	73.73	13.49	5.41	7.37
Employers	8.99	84.22	4.78	2.02
Employed	0.39	0.51	95.35	3.74
Non-employed	4.28	1.54	28.13	66.06
Panel B: Simulated				
Non-employers	62.66	15.89	9.88	11.57
Employers	7.34	69.24	18.72	4.70
Employed	0.48	2.16	77.38	19.99
Non-employed	1.92	2.20	72.70	23.18

Note: The number of observations in each panel is 50,190.

Our simulated model is able to replicate the employment dynamics observed in the data fairly well. Table 1.4 presents a one-year transition matrix – Panel A shows the actual data and Panel B shows the simulated transition rates. The simulated rates replicate the overall pattern, although the persistence is lower than what we observe in the data, especially in non-employment.

Next, we assess the fit of individual labor earnings over the life-cycle. As shown in Figure 1.2, the simulated data from the estimated model replicates the growth in earnings in all three different careers. In the data, the age-earnings profile of employers is the steepest, followed by non-employers, and the age-earnings profile of employees is the flattest. Our model produces the same pattern.

The age-earnings profiles mask substantial heterogeneity in the own and cross returns to experience. Table 1.5 shows how the earnings in each of the careers change with own and cross experiences. In this table, we present the returns to own and cross experiences of an average individual in each of the three careers. For example, in our sample, an average non-employer entrepreneur has

Table 1.5: Return From Average Years of Experience From Each Career Choice

Variable	Non-employer	Employer	Employed
Return from exp. of non-employer	0.95	0.09	-0.01
Return from exp. of employer	0.18	1.60	0.00
Return from exp. of employed	0.37	0.31	0.87
Total	1.50	2.00	0.86

Note: Each number represents the return from average years of corresponding experience. Table 1.1 shows average years of experience from each career choice.

5.6 years of non-employer entrepreneur experience, 2.5 years of employer entrepreneur experience, and 11.4 years of experience as an employee (Table 1.1). Our estimates suggest that a non-employer who has 5.6 years of experience in the same career would earn 95 percent more than a non-employer who has no experience as a non-employer. Our estimates also indicate that a non-employer who has 2.5 years of experience as an employer would earn 18 percent more than a non-employer who has no experience as an employer. Finally, a non-employer who has 11.4 years of experience as an employee would earn 37 percent more than a non-employer who has no experience as an employee. The second column of Table 1.5 shows that for employer-entrepreneurs cross returns have a relatively small impact, but experience of an employer increases earnings substantially. The third column suggests that for employees, work experience accumulated as an employee increases their earnings but work experience accumulated as entrepreneurs has little or no effect on the earnings of employees. Taken together, these results suggest one-way transferability of skills across these three different types of careers: Experience as an employee is valuable in all career paths, but experience as a non-employer or employer entrepreneur does not increase earnings in paid employment.

Table 1.6: Return From Five Years of Experience From Each Career Choice

Variable	Non-employer	Employer	Employed
Return from exp. of non-employer	0.89	0.12	-0.27
Return from exp. of employer	0.32	1.23	-0.01
Return from exp. of employed	0.30	0.27	0.36

Note: Each number represents the return from five years of corresponding experience.

In interpreting the results of Table 1.5, one needs to take into account that the distribution of average years of own and cross experiences are different across careers. To provide a different view, we also present Table 1.6, which shows the returns to five years of own and cross experience for each of the careers. A few important differences across careers become apparent. The returns to same-career experience for entrepreneurs (both non-employers and employers) are substantially bigger than returns to same-career experience among employees: Five years of same-career experience increases earnings of non-employers by about 89 percent, employers by 123 percent, and employees by only 36 percent. However, these large percentage increases in earnings are coupled with relatively low starting wages for entrepreneurs. This suggests that entrepreneurs have to invest initially, and it takes some time to generate the returns from that investment.

The returns from other careers inform about transferability of skills across careers. The returns to cross-experience is highest for non-employers. In other words, experience from other careers increase earnings of non-employers most. For example, five years of experience as an employer (employee) increase the earnings of non-employers by about 32 (30) percent. However, experience as either a non-employer or an employer entrepreneur is not rewarded in terms of higher earnings as an employee. Somewhat surprisingly, experience as an employee is more valuable when working as an employer than experience gained as a non-employer.

1.5 Policy Simulations

In this section, we use the estimated structural model to simulate the effects of policies that are frequently discussed in the context of entrepreneurship, some of which are explicitly designed to promote entrepreneurship. A policy that is very relevant in practice is differential tax treatment of business income versus labor income from paid employment. Many countries have policies that reduce effective tax rates for entrepreneurs. For example, variants of the Dual Income Tax, as seen in some Scandinavian countries, effectively reduce tax rates for entrepreneurs by dividing business earnings into labor and capital income and applying a lower tax rate to the capital income portion. In many countries, earnings retained in the business are taxed at lower rates or taxation is deferred to the time when earnings are distributed, which results in lower effective tax rates. Moreover, it is often argued that it is easier for entrepreneurs to avoid or evade taxes because they self-report their earnings, whereas paid employees are subject to third-party reporting and face withholding taxes (Kleven et al., 2011; Fossen et al., 2020).

To simulate this idea in our model, we introduce hypothetical earnings subsidies of one, five, and ten percent of before-tax earnings, in one case for all entrepreneurs (both non-employers and employers), and in another set of simulations only for employers. This is equivalent to a lower effective tax rate for entrepreneurs (or employers, respectively). The results from these policy simulations are shown in Table 1.7, with the baseline scenario in the top row and the two hypothetical policy scenarios below. The results suggest that a 5 percent subsidy for all entrepreneurs increases the non-employer-entrepreneurship rate by 0.6 percentage points, or by about 25 percent. A 10 percent subsidy for entrepreneurs increases the non-employer-entrepreneurship rate by 1.15 percentage points and the employer-entrepreneurship rate by 1.18 percentage points. These increases come at the expense of a 1.92 percentage points decrease in the share of paid employees. In contrast, the non-employment rate decreases by only 0.41 percentage points. Thus, only about 18 percent of the increase in entrepreneurship is due to individuals moving from non-employment to entrepreneurship. Most of the increase in entrepreneurship caused by the incentive comes from individuals changing career choice who would otherwise work as an employee.

Our simulations of the effects of a policy that increases only the earnings of employers, which appear in Panel B, show that a 10% subsidy of this type increases the share of employer-entrepreneurs from 5.57 percent to 7.29 percent, or by about 31 percent. Thus, the more targeted policy is more effective in encouraging individuals to become an employer.

Table 1.7: Partication Rates With Counterfactual Policy Scenarios

Subsidy	Non-employer	Employer	Employed	Non-employed
No subsidy	2.44	5.57	78.32	13.67
Panel A: Subsidy for all entrepreneurs				
1% earnings subsidy	2.57	5.67	78.13	13.63
5% earnings subsidy	3.04	6.12	77.37	13.47
10% earnings subsidy	3.59	6.75	76.40	13.26
Panel B: Subsidy for employers				
1% earnings subsidy	2.42	5.72	78.22	13.64
5% earnings subsidy	2.33	6.37	77.77	13.53
10% earnings subsidy	2.20	7.29	77.12	13.39

Notes: Using our estimated dynamic structural model, we simulate the effects of hypothetical policy scenarios on the participation rates in the different employment states. In panel A, we simulate the effects of a subsidy of 1%, 5% and 10% of an individual's earnings paid if this individual chooses to be an entrepreneur (non-employer or employer). In panel B, we simulate the effects of a subsidy of 1%, 5% and 10% of an individual's earnings paid if this individual chooses to be an employer. The first row repeats the baseline scenario without a subsidy for comparison. The number of observations in each simulation is 50,190.

1.6 Conclusion

In this chapter, we develop and estimate a dynamic structural model of career choice with a focus on entrepreneurial careers. Within entrepreneurs, we distinguish between non-employers and employers, who hire other individuals as employees. In our microeconometric model, we distinguish between different types of experience (experience as employee, as non-employer, and as employer). We solve the dynamic optimization problem by backward recursion and then estimate the model using the simulated maximum likelihood method.

Our results suggest that transferability of experience between sectors is one way. In particular, returns to work experience accumulated as an entrepreneur are negligible or even negative for employees. However, experience accumulated as an employee substantially increases the earnings of an entrepreneur.

Our model is able to explain the major patterns observed in the data, which suggests that this model will provide valuable information about how policy changes may affect entrepreneurship. To that effect, we use the estimated structural model to simulate hypothetical policy scenarios, in particular, increases in the earnings of entrepreneurs, which can be interpreted as tax breaks or subsidies. Variants of such policies have been discussed with the intention to promote entrepreneurship and to reduce unemployment. Our results suggest that effective subsidies to all entrepreneurs increase the share of individuals who would work as non-employer-entrepreneurs or employer-entrepreneurs, but most of these increases come from individuals who would otherwise work as employees, so the effectiveness of these policies with respect to reducing non-employment is limited.

One important avenue for future research would be to include capital accumulation in the model. This would greatly increase the state space and the required computerization power, but—when the computerization technology allows—the effort will eventually be worthwhile because the richer model would allow to simulate further relevant policies such as credit subsidies for entrepreneurs.

Appendix

1.A Tables

Table 1.A.1: Estimates of the Parameters of the Type Probability Determinants

Column A	Column B:	Type 1	Column C:	Type 2
Variable	Parameter	Estimate	Parameter	Estimate
Constant	η_0	-0.8108 (497.6509)	ζ_0	-1.5485 (204.2140)
Father self-employed	η_1	0.6826 (235.0465)	ζ_1	0.3914 (94.9959)
Willingness to take risk	η_2	0.0552 (3.4634)	ζ_2	0.0484 (1.2958)
Internal locus of control	η_3	-0.0565 (0.5376)	ζ_3	0.0627 (0.1973)
East Germany	η_4	0.3050 (92.7071)	ζ_4	-1.9192 (54.6780)
Migration background	η_5	0.3838 (104.1125)	ζ_5	-0.5991 (36.9636)

Notes: Column A represents the names of the variables that determine the unobserved type of an individual. Column B represents the names of the parameters and corresponding estimates that determine the probability of being Type 1, and Column C those that determine the probability of being Type 2. Standard errors at 1% deviation are in the parentheses.

Table 1.A.2: Estimates of the Variance-Covariance Matrix

Parameter	Variable	Estimate
σ_{11}	var(1,1)	0.8211 (0.4951)
σ_{21}	cov(2,1)	0.2275 (0.3896)
σ_{22}	var(2,2)	0.2872 (0.4829)
σ_{31}	cov(3,1)	0.6000 (1.46e-6)
σ_{32}	cov(3,2)	0.1651 (0.2061)
σ_{33}	var(3,3)	0.4491(0.1405)

Notes: The numbers 1, 2, and 3 represent the wage-shock vectors of non-employers, employers, and employees respectively. Standard errors at 1% deviation are in the parentheses.

Table 1.A.3: Estimated Coefficients of Tax Liability Functions by Employment State

Parameter	Variable	Non-employer ($k = 1$)	Employer ($k = 2$)	Employed ($k = 3$)
π_{0k}	Constant	-1.20 (1.76)	1.04 (4.25)	-4.93 (1.37)
π_{1k}	Earnings	0.27 (0.12)	0.26 (0.17)	0.58 (0.080)
π_{2k}	Earnings^2	0.0031 (0.0018)	0.0016 (0.0014)	-0.0015 (0.0011)
π_{3k}	Earnings^3	-1.00e-5 (4.2e-6)	-6.6e-7 (1.6e-6)	2.2e-6 (2.4e-6)
π_{4k}	Age	-0.0031 (0.044)	-0.091 (0.091)	0.051 (0.036)
π_{5k}	Earnings$\times$Age	0.0026 (0.0030)	0.0028 (0.0034)	-0.0036 (0.0021)
π_{6k}	Earnings$^2\times$Age	-6.6e-5 (4.1e-5)	-2.5e-5 (2.6e-5)	3.4e-5 (2.7e-5)
π_{7k}	Earnings$^3\times$Age	2.3e-7 (9.6e-8)	5.8e-9 (3.0e-8)	-5.0e-8 (6.4e-8)
N		1788	2406	35657
R^2		0.87	0.89	0.88

Notes: This table shows the estimated regression parameters (standard errors are in the parentheses) of the tax liability function of gross income corresponding to the different choice alternatives ($k = 1, 2, 3$). The dependent variable is the real annual tax liability. The tax liability and annual earnings are in real \$1000 in prices of 2005.

Chapter 2

The Cost of Health Insurance and Entry into Entrepreneurship

(with Frank M. Fossen & Sankar Mukhopadhyay)

2.1 Introduction

In the U.S., health insurance and employment are inextricably related. Most (about 61.2 percent) non-elderly adults in the U.S. are covered by employer-provided health insurance (EPHI). Then there are two government-provided insurance programs: Medicaid (14.5 percent; primarily low-income individuals) and Medicare (about 2 percent among age below 65; mostly disabled and/or blind). The rest either buy health insurance from the private non-group market (7.7 percent; primarily self-employed individuals) or remain uninsured (12.9 percent) (KFF - State Health Facts, 2019).

Individuals with pre-existing conditions faced many restrictions in buying health insurance from the private non-group market before the Affordable Care Act

(ACA) of 2010. Before the ACA, health insurance providers could deny health insurance to individuals who had pre-existing conditions or charge higher prices, a practice known as medical underwriting. These practices often made health insurance either unavailable or unaffordable for individuals with pre-existing conditions. Previous research has suggested that this unavailability or unaffordability of health insurance may prevent some individuals from starting entrepreneurial activities. This phenomenon, "entrepreneurship lock", refers to a situation where individuals are locked out of entrepreneurship.[1]

However, the ACA changed the private non-group market in dramatic ways. The ACA mandates health insurance providers to sell insurance to all individuals regardless of their pre-existing conditions. It also set up health insurance exchanges (HIX) for individuals to buy insurance in the non-group market, thus making health insurance available for everyone. [2] On the other hand, the ACA mandates that the insurance companies cannot charge different premiums based on the medical history of an individual. They can vary premiums based on age and smoking status only. The ACA also mandated a set of conditions that all insurance providers need to cover, known as the Essential Health Benefits (EHB). These provisions and the uncertainty associated with the new HIX put upward pressure on the cost of health insurance for the average consumer in the non-group insurance markets. While some individuals are shielded from

[1]This is a reminiscence to the job lock literature (Madrian, 1994; Gilleskie and Lutz, 2002), which is concerned with inhibited mobility of paid employees due to EPHI.

[2]Moreover, the ACA imposed a health insurance mandate for legal residents of the U.S., to reduce adverse selection and to keep the price of health insurance affordable. Non-compliant individuals were assessed with a tax penalty, but the Tax Cut and Jobs Act of 2017 reduced the individual mandate penalty to zero.

cost increases because of the ACA-associated subsidies, those who are not eligible for subsidies faced a substantial increase in health insurance cost in the first few years after the full implementation of the ACA in 2014.

If a non-elderly adult is not covered by EPHI or one of the government programs, then they need to buy health insurance from the private market. Both EPHI and Medicaid are heavily subsidized. Employers, on average, pay 67 percent of the premium (U. S. Bureau of Labor Statistics, 2020) , and Medicaid does not charge a premium in most states (Brooks et al., 2020) [3]. Those who buy health insurance from HIX may also receive a subsidy. The ACA introduced two types of subsidy: Annual Premium Tax Credit (APTC), which is available to individuals with income less than 400% of Federal Poverty Level (FPL), and Cost Sharing Reduction (CSR), which is available to individuals with income less than 250% of FPL. Individuals above 400% of the Federal Poverty Level (FPL) are not eligible for any subsidy. It is worth noting that among the individuals who remained uninsured in 2018, almost half had an income above 200% of FPL, and 16% was above 400% of FPL. Furthermore, about 45% cited the cost of health insurance as the primary reason (KFF - The Kaiser Commission on Medicaid and the Uninsured, 2015) behind their choice of not buying health insurance.

In this paper we use data on the cost of health insurance plans from the Robert Wood Johnson Foundation (RWJF) and individual-level data from the Annual Social and Economic Supplement of the Current Population Survey (CPS-ASEC)

[3]Only five states charge a small premium

to estimate the effect of the price of health insurance on the decision to become self-employed. Therefore, we extend the literature by exploring whether entrepreneurship lock exists due to health insurance costs. To the best of our knowledge, prior literature has not addressed this research question using the variation in cost prevalent in the HIX. Existing papers on the effect of the ACA on entrepreneurship instead focus on the potential effect of the availability of health insurance brought by ACA on entrepreneurship.

The remainder of this chapter is organized as follows: Section 2.2 provides a review of the literature on entrepreneurship lock. Section 2.3 discusses the relevant provisions of the ACA regarding our analysis. Section 2.4 introduces the data, and Section 2.5 explains the methodologies we use to answer our research question. Section 2.6 provides the results, and Section 2.7 concludes our analysis.

2.2 Literature Review

Extant literature mostly supports that the unavailability of health insurance discourages potential entrepreneurs from starting a business (Wellington, 2001; Fairlie et al., 2011), with some papers reporting this result for specific groups such as married women (Lombard, 2001) and older women (Jia, 2014). In particular, using CPS data, Fairlie et al. (2011) find that individuals without spousal coverage are significantly less likely to start a business than individuals who have spousal coverage. However, not all studies are conclusive, including the pioneering paper by Holtz-Eakin et al. (1996), which reports estimates that have large standard errors, and the paper by Zissimopoulos and Karoly (2007), whose authors focus on older individuals and conclude that the results from their study are only partially reconcilable with job lock.

Only a few papers investigate the effect of the cost of health insurance on entrepreneurship. Four studies analyze the effects of the 1986 Tax Reform Act and its amendments, which allowed self-employed individuals in the United States to deduct increasing portions of their health insurance premiums from their taxable income (Heim and Lurie, 2010; Gurley-Calvez, 2011; Velamuri, 2012; Gumus and Regan, 2015). These papers find that a lower after-tax price of health insurance in self-employment encourages self-employment, although Gumus and Regan (2015) report significant effects only for entry into self-employment for singles and married men whose wives lack employer-provided health insurance. Fossen and König (2017) analyze the health insurance system in Germany and find that higher health insurance costs in self-employment compared to the costs in paid employment decrease the probability of entry into self-employment. None of the existing papers quantitatively analyzing the effects of the cost of health insurance on entrepreneurship take into account the changes that the ACA introduced in the United States.

Several papers have used the increased availability of health insurance under various state-level insurance mandates to test whether "entrepreneurship lock" exists. For example, Li et al. (2017) found that state-level insurance mandates increased self-employment by 0.4 percentage points. They also found that most of this increase came from single individuals. Heim and Lurie (2014) found that increased availability of health insurance from the Massachusetts Health Reform Act also led to an increase in self-employment. More recently, several papers have used the variation generated by the ACA. While some papers find that the increased availability increased self-employment, others do not find any effect. For example, Blumberg et al. (2014) estimate that the number of self-employed individuals will increase by 1.5 million due to the availability of health insurance within the first two years of full implementation of the ACA. Bailey and Dave (2019) found that the ACA increased self-employment by 3-4% and full-time self-employment by 9%. On the other hand, Barber III and Kavoori (2015), Heim and Yang (2017) did not find any statistically significant effect of the ACA on self-employment. Bailey (2017) explored the effect of the dependent coverage mandate in the ACA and did not find any effect on the overall self-employment rate. However, they did report an increase of about 20 percent among disabled respondents. In a complementary paper, Barber III and Kavoori (2018) show that the ACA increased the likelihood of private purchase of non-group insurance among self-employed individuals.

Overall, the existing literature provides inconclusive results on the existence of the entrepreneurship lock and the ACA's role. A possible explanation of the inconclusiveness is the literature's focus on the extended availability of health insurance through ACA. We contribute to the literature by exploiting variation in the cost of health insurance introduced by the ACA.

2.3 The ACA Provision

The ACA requires each state to define one or more Rating Areas. States have used metropolitan areas (MSA), counties, or 3-digit zip codes to define Rating Areas. In the Health Insurance Exchanges (HIX), insurance providers have to

Table 2.3.1: Metal Tiers of Health Insurance Plans in Health Insurance Exchanges

Metal Tier	Actuarial Value	Number of Policies Sold	Market Share	% of Consumers with Subsidy
Catastrophic	-	76,920	1%	0%
Bronze	60%	1,872,457	21%	79%
Silver	70%	6,090,199	69%	94%
Gold	80%	573,641	6%	63%
Platinum	90%	225,074	3%	60%

Note: The numbers refer to 2015.
Source: ASPE Issue Brief (DeLeire and Marks, 2015).

charge the same premium, deductible, and maximum out of pocket (MOOP) to all persons within a Rating Area without medical underwriting, regardless of their health status. The premium may only depend on age and smoking status. Each plan has a fixed actuarial value. A plan can be a Bronze, Silver, Gold, Platinum, or a Catastrophic plan, where the metal tiers depend on the plan's actuarial value. A Bronze plan must pay for at least 60% of the expected value of healthcare expenditures for enrollees, a Silver plan 70%, Gold 80%, and Platinum 90% (see Table 2.3.1).

Legal residents and citizens who are not eligible for Medicaid and do not have EPHI are eligible for subsidies that reduce the cost of insurance. Individuals with income between 138%-400% of the federal poverty level (FPL) may be eligible to receive an Annual Premium Tax Credit (APTC) for an insurance plan purchased through the HIX in states that expanded Medicaid. In the states that did not expand Medicaid, individuals between 100% and 400% of the FPL are eligible for APTC.

The APTC caps the amount an individual would have to pay (as a percentage of his or her income). The premium for the Second Lowest Cost Silver Plan (SLCSP) in each Rating Area is used to compute the level of APTC. Table 2.3.2 shows the limits and how they vary with income. For example, an individual who has an income between the 100% and 133% of the FPL and purchases the SLCSP offered in her Rating Area would pay 2.08% of her income and the rest of the premium would be covered by the APTC. If this individual elected to purchase a different HIX plan with a higher premium, the magnitude of the APTC would remain unchanged and the individual would be responsible for the difference between the higher premium and the computed tax credit. This also implies that the individual may reduce the amount she pays to below 2.08% of her income if she buys a plan with a premium less than the SLCSP. In addition, individuals with income between 100% and 250% FPL may also receive cost sharing reduction (CSR) subsidies to reduce deductible and co-insurance payments if they purchase Silver plans through the HIX.

Table 2.3.2: Premium Caps by Income Level as % of the Federal Poverty Level (2019)

Income as % of FPL	Cap % (Lower End)	Cap % (Higher End)
Up to 133%	2.08%	2.08%
133%-150%	3.11%	4.15%
150%-200%	4.15%	6.54%
200%-250%	6.54%	8.36%
250%-300%	8.36%	9.86%
300%-400%	9.86%	9.86%

Note: The caps are expressed in percent of an individual's income.

Silver and Bronze plans accounted for 90% of HIX plans sold in 2015, with Silver plans accounting for the majority of these plans (see Table 2.3.1). Catastrophic plans have the smallest market share, because eligibility to purchase these plans is restricted to young adults who meet specific requirements.

2.4 Data

We use data from the Annual Social and Economic Supplement of the Current Population Survey (CPS-ASEC) from 2015 to 2019, which is collected annually in March. The ASEC is representative for the population in the United States and provides information on labor market outcomes including whether an individual is self-employed at the time of the survey, as well as extensive socio-economic information. However, the CPS does not include information on prices of health insurance plans available to the agents on the local HIX. Therefore, we use the HIX Compare data from the Robert Wood Johnson Foundation (RWJF), which provides plan-level information on HIX market plans offered during the years 2014-2018. The HIX Compare dataset includes information on each plan's metal level, the Rating Area, premiums, and cost-sharing provisions such as the deductible and MOOP. Therefore, by merging these two datasets, we obtain information on the premium, deductible, and MOOP of the insurances sold on the HIX in each Rating Area in every year.

As one would expect, premiums differ by metal level. There is also substantial variation within metal levels across Rating Areas and across years within each Rating Areas. In our analysis we use the cost of the SLCSP of a 50 year-old person in a Rating Area as the cost of insurance in that Rating Area. We use this cost measure for three reasons: First, the premium of SLCSP in each Rating Area is used to compute the level of APTC; second, as mentioned above, consumers have to buy Silver (though not necessarily SLCSP) to be eligible for CSR; and third, almost 70% of all insurances sold on the HIX are Silver.

As mentioned earlier, states use counties (40 states), MSAs (seven states), or 3-digit zip codes (three states) to define Rating Areas. California uses a combination of these geographical boundaries to define Rating Areas. As of 2020, there are 506 Rating Areas in the U.S., and county lines define a vast majority of them: 405 out of the 506 Rating areas. When a Rating Area is a county, we can simply merge CPS data with HIX Compare data, since both datasets identify county. If a Rating Area includes multiple counties then we can safely assume that the each of these counties has the same premium. If a county includes multiple Rating Areas then we assume that the county has the average premium prevailing in these Rating Areas as an approximation.

The CPS does not always identify the county of residence of a respondent to preserve the confidentiality of respondents; unidentified counties are those with a small number of inhabitants, mostly rural counties. According to the CPS, about 45% of households in recent years are located in a county that is identified. Therefore, the health insurance cost information is not available for the rest of the sample. We exclude these individuals from our analysis. We focus our attention to respondents between the ages of 26 to 64 years. Individuals below the age of 26 years may have health insurance coverage through their parents, and older individuals are usually covered by Medicare. We only include individuals in the sample who were paid employees and had EPHI coverage either as a policyholder or as a dependent in the calendar year before the interview. This sample restriction provides us clean treatment and control groups for our estimation of the probability of becoming self-employed between the previous year and the current interview (see below). We also exclude individuals earning between 250% and 400% of the FPL because they receive some, but not the full subsidies and therefore cannot be clearly classified into either the treatment or control groups. Individuals earning more than 2000 percent of the FPL are also excluded because they may have a different decision process about entry into entrepreneurship and may be less comparable to individuals with lower income.

The CPS interviews respondents for four consecutive months and then again for four consecutive months after an eight-month gap. Therefore, if a respondent is included in the ASEC in March of year t, then the individual will be included in the ASEC in the March of year t+1 a second time, except for unplanned attrition. In each interview, respondents are asked whether they are currently self-employed. Respondents are also asked whether they were a paid employee or self-employed in the previous calendar year, about their health insurance status in the previous calendar year, and their annual income in the previous calendar year, among other things. Using the information contained in the retrospective question, we create a binary variable indicating entry into self-employment, denoted entry, which takes a value of one if the individual was a paid employee in the previous calendar year but is currently self-employed; otherwise, it takes a value of zero.

Table 2.4.1 represents the summary statistics of our analysis sample. The first column shows the full sample, and the next two columns split the sample based

Table 2.4.1: Summary Statistics

	Full sample	EPHI Policyholder	EPHI Dependent	Inc.$<250\%$	400<Inc. $<2000\%$
Entry	0.013	0.013	0.014	0.010	0.014
	(0.114)	(0.113)	(0.119)	(0.099)	(0.117)
Premium	3.989	4.038	3.710	4.030	3.980
	(1.057)	(1.092)	(0.772)	(1.153)	(1.035)
Deductible	3.200	3.224	3.065	3.333	3.171
	(1.489)	(1.501)	(1.414)	(1.494)	(1.486)
Moop	5.950	5.967	5.853	5.988	5.942
	(0.717)	(0.710)	(0.750)	(0.687)	(0.723)
Income	73.252	76.478	55.096	27.505	82.978
	(52.241)	(52.761)	(45.094)	(13.038)	(52.284)
Age	44.658	44.593	45.028	41.954	45.233
	(10.891)	(10.980)	(10.366)	(10.710)	(10.843)
# of Children	0.896	0.843	1.192	1.312	0.807
	(1.095)	(1.087)	(1.094)	(1.366)	(1.006)
Married	0.652	0.605	0.919	0.442	0.697
	(0.476)	(0.489)	(0.273)	(0.497)	(0.460)
Female	0.487	0.461	0.633	0.531	0.478
	(0.500)	(0.499)	(0.482)	(0.499)	(0.500)
White	0.788	0.781	0.824	0.701	0.806
	(0.409)	(0.413)	(0.381)	(0.458)	(0.395)
Black	0.103	0.109	0.071	0.197	0.083
	(0.304)	(0.312)	(0.257)	(0.398)	(0.276)
Other Race	0.109	0.110	0.105	0.102	0.110
	(0.311)	(0.312)	(0.306)	(0.303)	(0.313)
Less Than HS	0.035	0.035	0.037	0.119	0.017
	(0.184)	(0.184)	(0.188)	(0.324)	(0.131)
High School	0.198	0.198	0.200	0.348	0.166
	(0.399)	(0.399)	(0.400)	(0.476)	(0.373)
Some College	0.240	0.236	0.260	0.319	0.223
	(0.427)	(0.425)	(0.439)	(0.466)	(0.416)
College	0.527	0.531	0.503	0.214	0.593
	(0.499)	(0.499)	(0.500)	(0.410)	(0.491)
Poor Health	0.046	0.046	0.044	0.073	0.040
	(0.210)	(0.210)	(0.205)	(0.261)	(0.197)
N	42219	35850	6369	7402	34817

Notes: The table shows means and standard deviations in parentheses below.
Income, deductible and MOOP are in \$1000 per year; the premium is in \$100 per month; deflated to 2014 dollars.

on whether respondents were a primary policyholder (PH) of employer provided health insurance (EPHI) or covered by EPHI as a dependent in the previous calendar year. The fourth column includes respondents who had an income below 250% and the fifth column those who had an income between 400% and 2000% of the FPL in the previous calendar year. Individuals with income below 250% of FPL are eligible for both APTC and CSR. Individual with income above 400% of FPL are not eligible for either type of subsidies. We exclude individuals between 250% and 400% of FPL because they are eligible for APTC but not CSR. The average annual entry rate into self-employment is 1.3 percent for the full sample (column 1). The average entry rates of policyholders and dependents are 1.3 percent and 1.4 percent respectively. The average entry rates of low income (income<250%) and higher income (400%income<2000%) individuals are 1 percent and 1.4 percent respectively. Among EPHI dependents, the shares of women and the number of children are larger than among EPHI policyholders. Individuals who are white, married, and have a college degree have higher shares in the higher-income subsample than in the low-income subsample.

2.5 Identification and Methodology

2.5.1 Treatment and Control Groups and Graphical Evidence

We aim to estimate how a change in the SLCSP premium in the local HIX will affect an individual's decision to enter into self-employment. The primary treatment variable is the health insurance premium in the local HIX in a given year. It is a continuous treatment variable, with higher prices indicating a stronger treatment. However, certain groups of individuals are not affected by the HIX premium. This depends on two individual circumstances:

(i). Whether an individual needs to purchase health insurance through the HIX.

If an individual has an employed spouse and can get insurance through the spouse's employer then changes in the HIX insurance premiums will not affect the individual's decision. On the other hand if they have to purchase insurance thorough HIX then they can be affected by HIX price changes.

(ii). Even among those who purchase insurance through HIX, the exposure to the cost of insurance depends on their income levels. Individuals and families with income below 250% of the FPL are eligible for both APTC and CSR. Therefore, it is arguable that those with income below 250% of the FPL are largely shielded by subsidies. On the other hand, those above 400% of the FPL are not eligible for either and therefore bear the whole cost of insurance. Individuals with income between 250% and 400% of the FPL are eligible for reduced levels of APTC but no CSR, so they are partially affected by price changes on the HIX. We exclude these individual from the sample in order to have clean treatment and control groups.

Therefore, one can divide the population into four groups as shown in Table 2.5.1. Group C is the treatment group since individuals in this group have to buy insurance through the HIX if they become self-employed, and they are not shielded by the subsidy. Groups A, B and D are control groups. The subsidy provisions of the ACA imply that group A is (at least partly) shielded by the

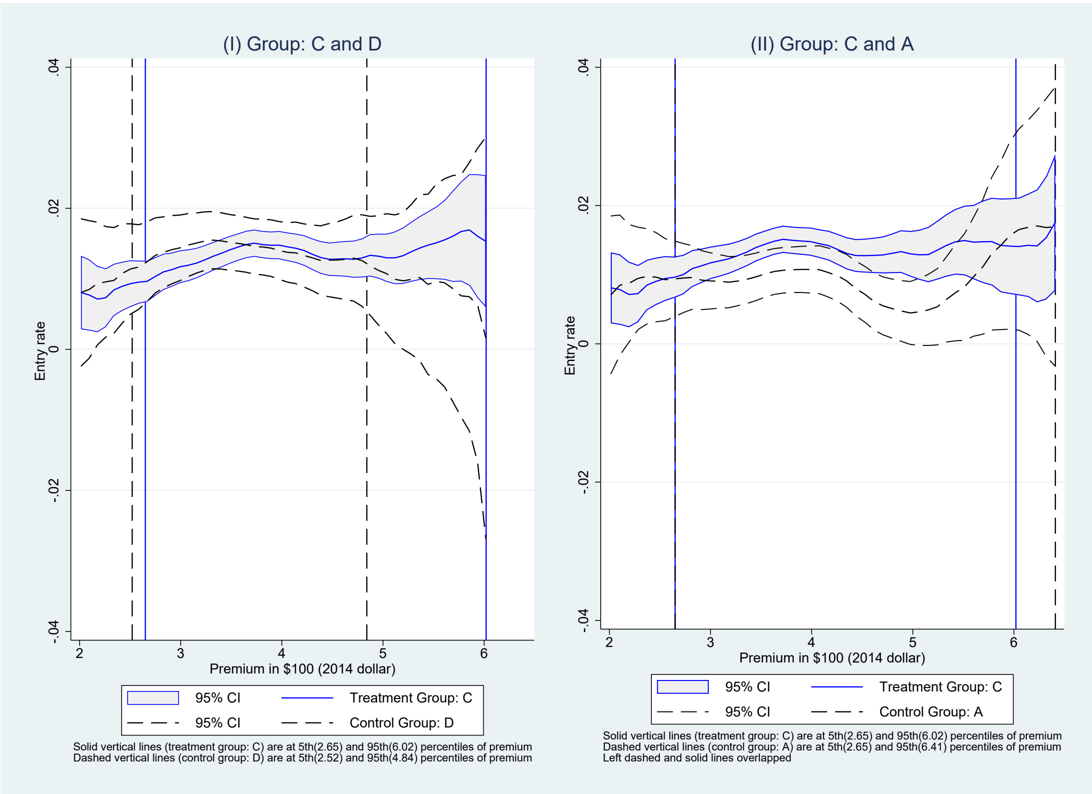
Table 2.5.1: Treatment and Control Groups

	EPHI: primary policyholder	EPHI: as dependent
<250% of FPL	A	B
400%-2000% of FPL	C	D

Note: Group C is the treatment group, the other groups are the control groups to identify the effect of HIX health insurance premiums on entry into self-employment. EPHI = employer-provided health insurance, FPL = federal poverty level.

APTC and CSR. We exploit the fact that the health insurance premiums only affect one of these groups in order to identify whether there is a causal effect of the health insurance premium on the decision to become self-employed.

Figure 2.5.1: Entry Rates Into Self-employment by HIX Premium in the Treatment and Control Groups.



With that in mind, Figure 2.5.1 shows how the entry rate into self-employment varies with the health insurance premium in the local HIX. Panel I shows how the entry rate varies with the premium for respondents in the higher income (groups C and D) by insurance status, and panel II the same for the EPHI policyholders (groups C and A) by income status. If higher health insurance costs in the HIX are a barrier to entry into self-employment, one would expect the entry rate for the treatment group to decrease relative to the entry rate of the control groups when the premium increases. However, in both panels, the relationship between the premium and the entry rates does not seem to differ between the treatment and the control groups, taking into account the 95% confidence intervals; if anything, the entry rate seems to increase in the treatment group relative to the control groups, contrary to expectation. Hence, the graphical evidence suggests that the HIX premium does not affect the probability of entry into self-employment.

2.5.2 Difference-in-differences and Triple Difference Models

As outlined above, the primary continuous treatment variable is the local HIX health insurance premium in a given year. It varies both across space and time. Importantly, the premium only affects individuals in group C in Table 2.5.1, the treatment group. This setting provides a quasi experiment that allows us to use difference-in-differences (DD) and triple-differences (DDD) methods to estimate the effect of health insurance premiums in the HIX on the probability of entry into self-employment.

We estimate two DD models. The first approach contrasts the treatment group C against the control group A based on the sample comprising these two groups C and A, and the second contrasts the same treatment group C against the control group D based on the sample comprising groups C and D. Finally, the DDD approach is based on the full sample comprising groups A, B, C, and D, and uses both dimensions to separate the treated from the controls, i.e., distinguishing between EPHI policyholders versus dependents and by income level. In the main estimations, both of the DD and DDD models, we control for county-level fixed effects, so we only use within-county changes in HIX premiums over time for identification. Any time-invariant unobserved differences between counties are controlled in these estimations.[4]

The estimation equations for the two DD models are

$$\begin{aligned}\text{Prob}\left(Entry_{it}=1|D_{1i,t-1},Premium_{j(i),t-1},\mu_{1j(i)},X_{it}^{c}\right)=\alpha_1 D_{1i,t-1}\\ +\alpha_2 Premium_{j(i),t-1}+\alpha_3(D_{1i,t-1}\times Premium_{j(i),t-1})+X_{it}^{c}\alpha_4^{c}+\mu_{1j(i)}+\epsilon_{1it}\end{aligned} \tag{2.1}$$

and

$$\begin{aligned}\text{Prob}\left(Entry_{it}=1|D_{2i,t-1},Premium_{j(i),t-1},\mu_{2j(i)},X_{it}^{c}\right)=\beta_1 D_{2i,t-1}\\ +\beta_2 Premium_{j(i),t-1}+\beta_3(D_{2i,t-1}\times Premium_{j(i),t-1})+X_{it}^{c}\beta_4^{c}+\mu_{2j(i)}+\epsilon_{2it},\end{aligned} \tag{2.2}$$

[4]We also estimate models without county fixed effects, additionally exploiting cross-sectional variation in HIX premiums. This increases efficiency of the estimation, but requires the stronger identifying assumption that unobserved differences between counties are uncorrelated with the local HIX premiums. In an additional robustness check, we control for individual-level fixed effects instead of county-level fixed effects.

and for the DDD model the equation is

$$\begin{aligned}\text{Prob}\left(Entry_{it} = 1|D_{1i,t-1}, D_{2i,t-1}, Premium_{j(i),t-1}, \mu_{3j(i)}, X^c_{it}\right) = \gamma_1 D_{1i,t-1}\\ +\gamma_2 D_{2i,t-1} + \gamma_3 Premium_{j(i),t-1} + \gamma_4(D_{1i,t-1} \times Premium_{j(i),t-1})\\ +\gamma_5(D_{2i,t-1} \times Premium_{j(i),t-1}) + \gamma_6(D_{1i,t-1} \times D_{2i,t-1})\\ +\gamma_7(D_{1i,t-1} \times D_{2i,t-1} \times Premium_{j(i),t-1}) + X^c_{it}\gamma^c_8 + \mu_{3j(i)} + \epsilon_{3it},\end{aligned} \tag{2.3}$$

where i denotes an individual and t the year the information in a variables pertains to, and $j(i)$ denotes the county individual i lives in. The outcome variable $Entry_{it}$ is a dummy indicating entry into self-employment between $t-1$ and t, and Prob(.) is the response probability. The continuous treatment variable $Premium_{j(i),t-1}$ is the premium in the local HIX in a given year. There are two treatment dummy variables, D_1 and D_2. $D_{1i,t-1} = 1$ if an individual has EPHI as the policyholder, and $D_{1i,t-1} = 0$ if the individual has EPHI as a dependent. $D_{2i,t-1} = 1$ if an individual has higher income (400%$<$ income $<$2000% FPL), and $D_{2i,t-1} = 0$ if the individual has low income (income $<$250% FPL). The three treatment variables are measured in $t-1$, before potential entry into self-employment. X^c is a row vector of control variables, μ are unobserved county fixed effects, and ϵ are the error terms. We estimate the linear probability models by OLS. Due to the binary nature of the dependent variable and because the health insurance premium varies across counties and time, we report standard errors robust to heteroscedasticity and clustered at the county level in all regressions.

In the DD models, the coefficients of the interaction terms, α_3 and β_3, respectively, capture the treatment effect on the treated. In the DDD model, this effect is given by the coefficient of the triple interaction, γ_7. If the local HIX premium is a barrier to entrepreneurship, we expect these coefficients to be negative.

The DD estimator relies on the assumption, in our context, that the entry rate into self-employment in the treatment and the control group would have trended the same over time if HIX premiums had not changed. Level differences in the entry rates do not distort the estimation of the treatment effect on the treated. Furthermore, any unobserved shocks that may be correlated with HIX premium changes, but affect the treatment and control groups in the same way, are controlled. The fact that we use two different DD models with different control groups serves as a safeguard in case the identifying assumption is not fully valid when using one of the control groups.

The DDD model (Gruber, 1994; Acemoglu et al., 2004; Fossen and Steiner, 2009; Olden and Møen, 2020) requires even weaker identifying assumptions. In our context, we only need to assume that the difference in the probability of entry between the groups C and D would have trended the same as the difference between the groups A and B in the absence of a change in the local HIX premium.[5] To see that the identifying assumption is weaker than in the DD estimations, consider this: In the DDD estimation, even if unobserved shocks are correlated with HIX premium changes and affect higher-income EPHI policyholders differently than low-income EPHI policyholders, these shocks are controlled if their differential effect is the same when comparing higher-income EPHI dependents to low-income EPHI dependents.

[5]Or, put differently, the assumption is that the difference in the entry probability between groups C and A would have trended the same as the difference between groups D and B.

To control for potential changes in the compositions of the treatment and control groups and to increase efficiency of the estimations, we include control variables in all our estimations. At the individual level, we include educational attainments, age, age squared, gender, number of children in the household, race, marital status, and health status (all observed in t); at the level of local HIX health insurance plans, we control for deductible and MOOP (in $t-1$). In a robustness check, we additionally control for total income in the calendar year $t-1$. We also include a full set of year dummies to control for potential general trends in the entry rate into self-employment or effects of the business cycle.

2.6 Empirical Results

Table 2.6.1 displays the results from estimating the main DD and DDD models. Columns 1 shows the DD estimations comparing EPHI policyholders (treatment group) to EPHI dependents (control group) based on the sample of higher income individuals who do not receive a subsidy for compensation. The DD estimate is the coefficient on the interaction term of the treatment dummy (abbreviated as D1) with the health insurance premium in the local HIX, which is the continuous treatment variable. Columns 2 provides the DD estimation comparing higher income individuals (treatment group, abbreviated as D2) to low income individuals (control group, because they are compensated for most of the health insurance costs in the HIX by the subsidies) based on the sample of EPHI policyholders. Both DD coefficients are insignificant, and the point estimates have positive signs, which is not consistent with entrepreneurship lock due to health insurance costs.

Table 2.6.1: Probability of Entry Into Self-employment

Model	(1)	(2)	(3)
Sample	C+D	A+C	A+B+C+D
D1 (EPHI PH)	-0.0092		-0.019
	(0.0068)		(0.023)
D2 (High Inc.)		-0.0011	-0.011
		(0.0051)	(0.026)
Premium	-0.0027	-0.00018	-0.0030
	(0.0021)	(0.0013)	(0.0059)
D1×Premium	0.0023		0.0023
	(0.0018)		(0.0059)
D2×Premium		0.00082	0.00093
		(0.0012)	(0.0064)
D1×D2			0.0099
			(0.026)
D1×D2×Premium			-0.000093
			(0.0065)
Age	0.0010**	0.00100**	0.00080**
	(0.00047)	(0.00040)	(0.00038)
Age2	-0.0000092*	-0.0000092**	-0.0000070
	(0.0000053)	(0.0000046)	(0.0000043)
# of Children	0.0020**	0.0011	0.0015**
	(0.00090)	(0.00076)	(0.00072)
High School	0.0017	0.0054*	0.0037
	(0.0042)	(0.0030)	(0.0030)
Some College	0.0030	0.0070**	0.0058**
	(0.0042)	(0.0028)	(0.0029)
College	0.0041	0.0067**	0.0061**
	(0.0039)	(0.0030)	(0.0030)
Married	0.0013	0.0016	0.0015
	(0.0016)	(0.0014)	(0.0014)
Female	-0.0067***	-0.0062***	-0.0062***
	(0.0013)	(0.0013)	(0.0012)
Race Black	-0.0035	-0.0051***	-0.0045**
	(0.0025)	(0.0019)	(0.0020)
Race Other	-0.0020	-0.0018	-0.0026
	(0.0022)	(0.0022)	(0.0019)
Poor Health	0.0034	0.0018	0.0021
	(0.0036)	(0.0031)	(0.0028)
Year FE	Yes	Yes	Yes
County FE	Yes	Yes	Yes
Mean Prob.	0.014	0.013	0.013
Rel. Effect Zize	16.8	6.27	-0.70
R^2	0.014	0.016	0.014
N	34817	35850	42219

[1] Linear probability models of entry into self-employment. DD and DDD estimates.
[2] Standard errors clustered at the county level in parentheses; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.
[3] D1=1 if the respondent has EPHI coverage as a policyholder (PH) and D1=0 if the respondent has EPHI coverage as a dependent.
[4] D2=1 if 400%<income<2000% of FPL and D2=0 if income<250% of FPL.

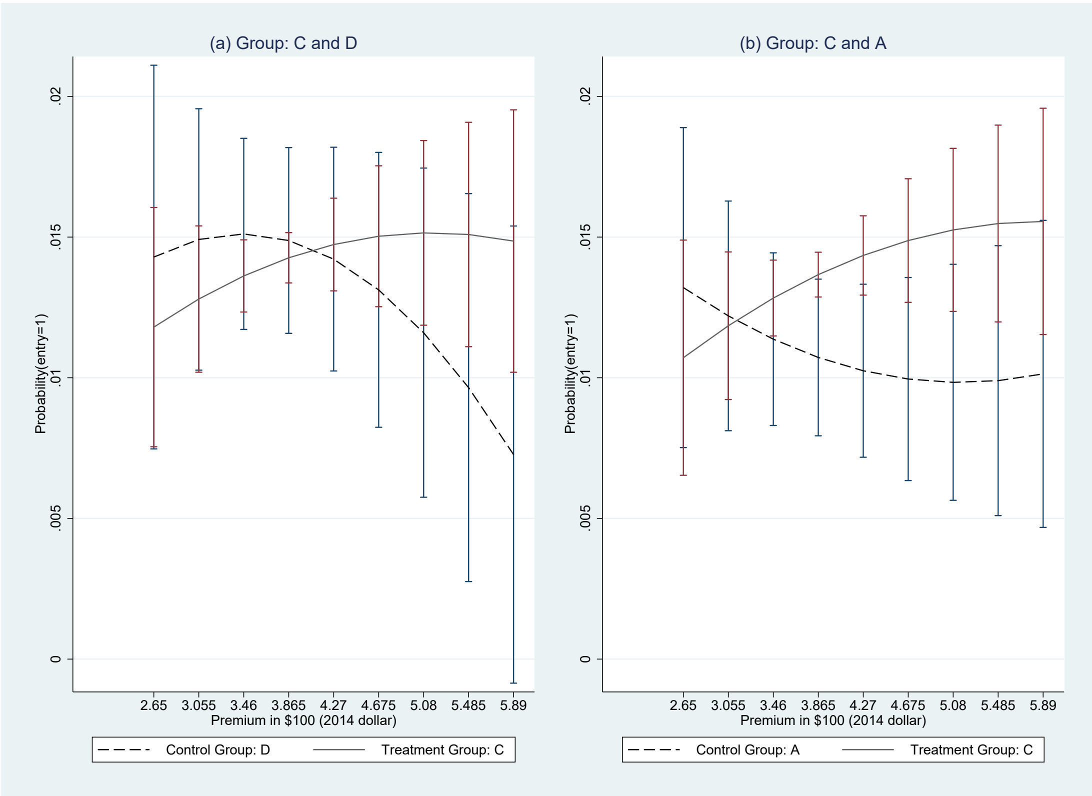
Column 3 shows the results from estimating the DDD models based on the full sample and including both the EPHI policyholder dummy, D1, and the higher income dummy, D2. The coefficient of the triple interaction is the DDD estimate of the treatment effect on the treated. This estimate is insignificant as well.

Thus, the estimated treatment effects on the treated are statistically insignificant in the three models. Are the point estimates economically significant? In the entry model in column 2, for example, if one took the point estimate of the interaction term at face value despite its insignificance, it would mean that increasing the premium by $100 per month would increase the annual probability of entry by 0.082 percentage points; relative to the baseline annual probability of entry of 1.3% this would correspond to an increase in the entry rate by 6.3% (as indicated at the bottom of the table). The mean monthly premium among EPHI policyholders is about $400 in the sample (Table 2.4.1), so the elasticity of the probability of entry with respect to the premium would be small (6.3% / 25% = 0.25). In column 1, the point estimate is larger, but in the DDD model (column 3), which requires the weakest assumptions, the point estimate is very close to zero. Hence, not only is the effect insignificant, but the point estimate of the economic effect size is also small at least in the most robust model, the DDD model.

The coefficients of the control variables confirm results from prior literature (Parker, 2018). Women are less likely to become self-employed than men, Black Americans less likely than whites, and a college degree increases these probabilities (the latter two effects are insignificant when only using the sample of EPHI policyholders in column 1).

We also estimate models including a quadratic term of the HIX premium to allow for potential non-linear effects. In the DD and DDD models, this means interacting not only the premium variable, but also the squared term with the treatment dummies, thus, doubling the number of double and triple interaction terms. We include the same control variables as in the linear models. Figure 2.6.1 shows the predicted probabilities of entry into self-employment using the estimated coefficients from the quadratic DD models. Panel (a) compares EPHI policyholders (the treatment group) to EPHI dependents, based on the sample of higher-income individuals, like column 1 in Table 2.6.1, and Panel (b) compares higher income individuals (the treatment group) to low-income individuals, based on the sample of EPHI policyholders, like column 2 in the table. The predicted relationships between the premium and entry for the treatment group are of course the same in both panels because the treatment group is the same. Importantly, the curves do not differ significantly between the treatment group and the control groups in both panels (the confidence intervals overlap). If anything, the probability of entry increases for the treatment group relative to the control group with higher HIX premiums, which confirms the graphical evidence presented above, but is not consistent with entrepreneurship lock. Based on the estimated quadratic DDD model, we test whether the coefficients of interest (the two coefficients of the triple interactions) are jointly significant, and we obtain a p-value of 0.95. Thus, our estimates do not support any differential impact of the price of health insurance on entry into self-employment for the treated individuals.

Figure 2.6.1: Predicted Probabilities of Quadratic Models With 95% Confidence Intervals Using County Fixed Effect Where Prediction Shown From 5 to 95 Percentiles of the Independent Variable Premium



We assess the robustness of our results with respect to various further specifications. In all these robustness checks, we include the same control variables as in the main Table 2.6.1. First, to account for possible non-linearities in the effect of the health insurance premium on entrepreneurial choices not captured well by the quadratic specification described above, we use three splines of the premium (using the 33 and 66 percentiles) and interact the slope coefficients in each interval with the treatment dummies. This allows us to estimate different treatment effects for each interval of the premium (Table 2.A.1). Second, going back to the main linear model, we include income in the previous calendar year as an additional control variable (Table 2.A.2). We do not include this variable in the main estimations because of its correlation with the higher income dummy variable, which may decrease efficiency of the estimation. Third, we use the HIX premium at time t instead of the premium at time $t-1$ (Table 2.A.3). The rationale is that individuals may anticipate next year's health insurance price when deciding to become self-employed. Fourth, in Table 2.A.4 in the Appendix (not including income), we include individual-level fixed effects instead of county-level fixed effects.[6] Finally, we remove any fixed effects (except for the time dummies) from the regression (Table 2.A.5). This way, we use cross-sectional variation in health insurance premiums for identification in addition to the variation over time used in the main analysis. This increases the statistical power, but comes at the cost of stronger identifying assumption, as mentioned in Section 2.5. In all robustness checks, none of the estimated treatment effects on the treated is significantly different from zero. We conclude that our findings are robust: There are no detectable effects of the health insurance premium in the HIX on the probability of entry into self-employment.

[6] This is not our preferred specification because the main independent variable of interest, the health insurance premium, varies across counties and time, so county-level unobservables are a larger concern than individual-level unobservables, and individual-level fixed effects may not fully capture county-level unobservables to the extent that individuals move across counties. On the other hand, there are not enough movers to separately identify both, county and individual-level fixed effects.

2.7 Conclusion

In this chapter we examine whether "entrepreneurship lock" exists due to high prices of health insurance. To examine the impact of the price of health insurance on entry into self-employment we exploit changes of health insurance premiums in local Health Insurance Exchanges after the introduction of the Affordable Care Act as natural experiments. We estimate difference-in-differences and triple difference models using individual-level data from the CPS-ASEC merged with panel data on health insurance plans. Individuals who were covered by EPHI as policyholders and had higher income before potential entry into self-employment are the treatment group because they lose their EPHI coverage when becoming self-employed and have to bear the cost of health insurance bought at the HIX. In contrast, EPHI dependents and low-income individuals are control groups because they are not affected by HIX premiums due to health insurance coverage through a spouse or subsidies, respectively. Our DD and DDD estimates indicate that the price of health insurance does not have any differential impact on the groups concerning their probabilities of entry into self-employment. Hence, our findings show that higher prices of health insurance do not deter individuals from becoming self-employed. This result does not invalidate prior literature reporting entrepreneurship lock due to the unavailability of health insurance, because the ACA made health insurance universally available; however, the novel insight we provide is that the price of health insurance in the HIX does not play a significant role for the decision to become self-employed. For future research, larger data samples are needed to increase precision of estimates in order to be able to detect potentially small effect sizes.

Appendix

2.A Tables

Table 2.A.1: Probability of Entry Into Self-employment Using Splines

Model	(1)	(2)	(3)
Sample	C+D	A+C	A+B+C+D
D1 (EPHI PH)	-0.00061		-0.080
	(0.017)		(0.082)
D2 (High Inc.)		-0.022	-0.10
		(0.020)	(0.087)
Premium1	0.0057	-0.0026	-0.025
	(0.0059)	(0.0054)	(0.026)
Premium2	-0.0073	-0.0019	0.027
	(0.0076)	(0.0054)	(0.022)
Premium3	-0.0045	0.00096	-0.0086
	(0.0029)	(0.0018)	(0.0079)
D1×Premium1	-0.00060		0.023
	(0.0057)		(0.025)
D1×Premium2	0.0057		-0.031
	(0.0076)		(0.021)
D1×Premium3	0.0034		0.0089
	(0.0025)		(0.0081)
D2×Premium1		0.0071	0.031
		(0.0061)	(0.027)
D2×Premium2		0.0038	-0.034
		(0.0052)	(0.023)
D2×Premium3		-0.0013	0.0047
		(0.0019)	(0.0079)
D1×D2			0.080
			(0.086)
D1×D2×Premium1			-0.024
			(0.026)

D1×D2×Premium2			0.037
			(0.023)
D1×D2×Premium3			-0.0057
			(0.0083)
Control Var.	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
County FE	Yes	Yes	Yes
Mean Prob.	0.014	0.013	0.013
R^2	0.014	0.016	0.014
N	34817	35850	42219

[1] Linear probability models of entry into self-employment. DD and DDD estimates.

[2] Standard errors clustered at the county level in parentheses; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

[3] D1=1 if the respondent has EPHI coverage as a policyholder (PH) and D1=0 if the respondent has EPHI coverage as a dependent.

[4] D2=1 if 400%<income<2000% of FPL and D2=0 if income<250% of FPL.

Table 2.A.2: Probability of Entry Into Self-employment With Income as Additional Control Variable

Model	(1)	(2)	(3)
Sample	C+D	A+C	A+B+C+D
D1 (EPHI PH)	-0.0095		-0.019
	(0.0069)		(0.023)
D2 (High Inc.)		-0.0015	-0.011
		(0.0051)	(0.026)
Premium	-0.0027	-0.00018	-0.0030
	(0.0021)	(0.0013)	(0.0059)
D1×Premium	0.0023		0.0023
	(0.0017)		(0.0059)
D2×Premium		0.00081	0.00093
		(0.0012)	(0.0064)
D1×D2			0.0097
			(0.026)
D1×D2×Premium			-0.000097
			(0.0065)
Income	0.000012	0.0000096	0.000014
	(0.000018)	(0.000016)	(0.000017)
Control Var.	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
County FE	Yes	Yes	Yes
Control for Inc.	Yes	Yes	Yes
Mean Prob.	0.014	0.013	0.013
Rel. Effect Size	16.8	6.26	-0.73
R^2	0.014	0.016	0.014
N	34817	35850	42219

[1] Linear probability models of entry into self-employment. DD and DDD estimates.

[2] Standard errors clustered at the county level in parentheses; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

[3] D1=1 if the respondent has EPHI coverage as a policyholder (PH) and D1=0 if the respondent has EPHI coverage as a dependent.

[4] D2=1 if 400%<income<2000% of FPL and D2=0 if income<250% of FPL.

Table 2.A.3: Probability of Entry Into Self-employment Using Premium at Time t

Model	(1)	(2)	(3)
Sample	C+D	A+C	A+B+C+D
D1 (EPHI PH)	-0.0073		0.0032
	(0.0050)		(0.020)
D2 (High Inc.)		-0.0022	0.0084
		(0.0048)	(0.021)
Premium	-0.0032***	-0.0019	0.00032
	(0.0012)	(0.0011)	(0.0051)
D1×Premium	0.0018		-0.0026
	(0.0011)		(0.0050)
D2×Premium		0.00092	-0.0034
		(0.0011)	(0.0051)
D1×D2			-0.011
			(0.021)
D1×D2×Premium			0.0044
			(0.0053)
Control Var.	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
County FE	Yes	Yes	Yes
Mean Prob.	0.013	0.012	0.013
Rel. Effect Size	13.2	7.50	35.3
R^2	0.013	0.014	0.012
N	43565	44796	53156

[1] Linear probability models of entry into self-employment. DD and DDD estimates.

[2] Standard errors clustered at the county level in parentheses; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

[3] D1=1 if the respondent has EPHI coverage as a policyholder (PH) and D1=0 if the respondent has EPHI coverage as a dependent.

[4] D2=1 if 400%<income<2000% of FPL and D2=0 if income<250% of FPL.

Table 2.A.4: Probability of Entry Into Self-employment With Individual Fixed Effects

Model	(1)	(2)	(3)
Sample	C+D	A+C	A+B+C+D
D1 (EPHI PH)	0.0074		-0.00016
	(0.026)		(0.017)
D2 (High Inc.)		-0.0020	-0.011
		(0.031)	(0.022)
Premium	0.00088	-0.0011	-0.00097
	(0.0074)	(0.0030)	(0.0025)
D1×Premium	-0.00049		0.00043
	(0.0073)		(0.0036)
D2×Premium		0.00048	0.0021
		(0.0056)	(0.0061)
D1×D2			0.0079
			(0.033)
D1×D2×Premium			-0.0011
			(0.0080)
Control Var.	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
Individual FE	Yes	Yes	Yes
Mean Prob.	0.014	0.013	0.013
Rel. Effect Zize	-3.48	3.69	-8.17
R^2	0.95	0.96	0.95
N	34817	35850	42219

[1] Linear probability models of entry into self-employment. DD and DDD estimates.

[2] Standard errors clustered at the county level in parentheses; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

[3] D1=1 if the respondent has EPHI coverage as a policyholder (PH) and D1=0 if the respondent has EPHI coverage as a dependent.

[4] D2=1 if 400%<income<2000% of FPL and D2=0 if income<250% of FPL.

Table 2.A.5: Probability of Entry Into Self-employment Without Fixed Effects

Model	(1)	(2)	(3)
Sample	C+D	A+C	A+B+C+D
D1 (EPHI PH)	-0.0074		-0.014
	(0.0067)		(0.023)
D2 (High Inc.)		0.00017	-0.0055
		(0.0049)	(0.025)
Premium	-0.0014	0.00011	-0.0010
	(0.0016)	(0.0011)	(0.0058)
D1×Premium	0.0019		0.0011
	(0.0017)		(0.0058)
D2×Premium		0.00060	-0.00039
		(0.0012)	(0.0063)
D1×D2			0.0058
			(0.025)
D1×D2×Premium			0.00098
			(0.0064)
Control Var.	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
County/Ind. FE	No	No	No
Mean Prob.	0.014	0.013	0.013
Rel. Effect Size	13.9	4.62	7.38
R^2	0.0021	0.0023	0.0023
N	34817	35850	42219

[1] Linear probability models of entry into self-employment. DD and DDD estimates.

[2] Standard errors clustered at the county level in parentheses; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

[3] D1=1 if the respondent has EPHI coverage as a policyholder (PH) and D1=0 if the respondent has EPHI coverage as a dependent.

[4] D2=1 if 400%<income<2000% of FPL and D2=0 if income<250% of FPL.

Chapter 3

The Effects of the Affordable Care Act on Workplace Absenteeism of Overweight and Obese Workers

(with Ethan Grumstrup; Sankar Mukhopadhyay & Olga Shapoval)

3.1 Introduction

Obesity is a growing problem in the U.S. According to the Centers for Disease Control and Prevention (CDC), 42.4 percent of adults in the U.S. were obese during 2017 to 2018 (Hales et al., 2020). The prevalence of obesity was 40 percent among younger adults, 44.8 percent among middle-aged adults, and 42.8 percent among older adults. Obesity is associated with several health conditions, such as cardiovascular disorders, musculoskeletal disorders, and sleep apnea (World Health Organisation, 2018), which may not necessarily lead to disability but may affect workplace absenteeism. Two systematic reviews (Neovius et al., 2009; Schmier et al., 2006) found that obese individuals are more likely to miss work compared to their normal-weight counterparts. Cawley et al. (2007) estimated that the total cost of obesity-related absenteeism is about $4.3 billion in the U.S. In addition, Finkelstein et al. (2010) estimate that the total cost associated with all obese full-time employees is about $73.1 billion and about 18 percent of this cost (or about $13 billion) is due to increased absenteeism. Access to health insurance is often suggested as a potential solution to this problem (Trogdon et al., 2008; Cawley et al., 2007).

In this chapter, we examine whether the expansion of health insurance coverage brought by the Patient Protection and Affordable Care Act of 2010 (ACA), led to a decline in absenteeism among overweight and obese individuals. While several papers have examined the effects of different provisions of the ACA on labor market outcomes, we are not aware of any previous paper that has attempted to explore the effects of the ACA on absenteeism.

On January 1, 2014, the individual mandate of the ACA, which required most citizens and legal residents to have health insurance (or pay a fine), went into effect [1]. At the same time, 25 states (including DC) expanded Medicaid to cover individuals with earnings up to 138 percent of the federal poverty level. Later, eight more states adopted the Medicaid expansion. These two changes reduced the percentage of non-elderly adults (ages 19-64) without health insurance from 20.3 percent in 2013 to 12.2 percent in 2016. We use individual-level data from the National Health Interview Survey (NHIS) to estimate the effect of the ACA

[1] The individual mandate was repealed in late 2017. Early evidence suggests that it did not affect the percentage of individuals with health insurance. Our study does not include the period after repeal.

on absenteeism among overweight ($25 \leq \text{BMI} < 30$) and obese ($\text{BMI} \geq 30$) individuals. We compare absenteeism among overweight and obese individuals, to absenteeism among normal-weight ($18.5 \leq \text{BMI} < 25$) individuals before and after the ACA, to estimate the effect of the ACA on absenteeism.

The ACA increased health insurance coverage rates for all weight groups (underweight, normal-weight, overweight, and obese). Therefore, we do not have an untreated (control) group per se. Therefore, our structure is somewhat different from a canonical difference-in-difference (DD) model. We hypothesize that health insurance coverage may have a different effect on absenteeism of overweight and obese workers compared to normal-weight workers. Given that, we define the "treatment effect" as the effect of the insurance expansion on absenteeism of overweight/obese workers compared to normal-weight workers. Once we define our "treatment effect" this way, we have a structure analogous to a DD model.

Our results suggest that in the post-ACA period, the probability of being absent declined by about 1.3 (1.5) percentage points among obese (overweight) individuals. The effect on women is more significant than on men. The probability of being absent declined by about 2.3 (2.6) percentage points in the post-ACA period among obese (overweight) women. The effect on men was comparatively smaller and statistically insignificant. Furthermore, our estimates (using a Tobit model) indicate that the obese (overweight) workers missed 0.32 (0.48) fewer days after the ACA. Again, the effect was stronger among overweight women (a statistically significant decline of 0.81 days) compared to overweight men (0.27 days and insignificant). The same holds for obese workers (a statistically significant decline of 0.68 days for obese women compared to an insignificant 0.05 days for obese men). In the pre-ACA period, obese (overweight) women missed 3.9 (2.9) days of work per year; therefore, this change translates to about a 17 (28) percent decline in days of missed work among obese (overweight) women.

While the ACA expanded health insurance coverage among non-elderly adults, the health insurance coverage rate of the elderly (age 65 and above) did not change during this time. In this group, the percentage of uninsured individuals declined only by 0.1 percentage points (from 1.2 percent in 2005 (DeNavas-Walt et al., 2013) to 1.1 percent in 2016 (Barnett and Berchick, 2017)). This group, therefore, provides us with an opportunity (albeit imperfect) to test whether the effect reported above is due to unrelated time effects. If the decline in absenteeism were due to some unrelated time trend, we would expect to see a similar effect among elderly obese individuals. Our results show that there is no decline in absenteeism among elderly adults, suggesting that the decline in absenteeism is indeed due to the expansion of health insurance coverage due to the ACA.

The rest of the chapter is structured in the following way. Section 3.2 provides a background, Section 3.3 discusses data, Section 3.4 presents the results, and Section 3.5 concludes.

3.2 Background

Two previous reviews (Neovius et al., 2009; Schmier et al., 2006) have concluded that overweight and obese workers are more likely to be absent from work compared to normal-weight workers. Related literature shows that access to health insurance reduces both the probability of missing work and the number of days missed (Gilleskie, 1998; Lofland and Frick, 2006). In particular, Dizioli and Pinheiro (2016) control for endogeneity of health insurance status, and find that workers with health insurance missed 76.54 percent (or 5.5 days) fewer workdays over two years compared to workers without health insurance. It is plausible that access to health insurance may allow obese individuals to address some of the chronic conditions associated with obesity. This has led researchers and policymakers to suggest that expanding health insurance coverage may lead to reduced absenteeism among obese workers. For example, Cawley et al. (2007) conclude that providing health insurance may be a solution. However, an individual who is insured against health risks may have less incentive to invest in dieting and exercise (Bhattacharya and Sood, 2006)[2]. Some studies report that access to health insurance is not associated with absenteeism (Xu and Jensen, 2012) or even associated with an increase in absenteeism (Vistnes, 1997). Therefore, it is an empirical question whether the expansion of health insurance under the ACA reduced absenteeism of overweight and obese workers.

Several studies examine the effect of the ACA on labor market outcomes. They find limited or no impact on young adults (Dahlen, 2015; Heim et al., 2015;

[2]Simon et al. (2017) do not find any evidence that increased health insurance coverage under the ACA led to increased obesity.

Slusky, 2017). Studies that use the Medicaid expansion and employer coverage mandate, also find no effect on labor force participation (Gooptu et al., 2016; Kaestner et al., 2017; Leung and Mas, 2016; Moriya et al., 2016). Dong et al. (2017) found that newer retirees are expecting to work longer than previous generations. However, other studies find that the ACA had little to no significant impact on those retirement decisions (Ayyagari, 2019; French et al., 2016; Gustman et al., 2016; Kaestner et al., 2017; Levy et al., 2018). We are not aware of any previous paper that has attempted to explore the effects of the ACA on absenteeism among different weight groups.

3.3 Data

We use data from the National Health Interview Survey (NHIS) (Blewett et al., 2016) for the years, including 2005 to 2018. Our baseline sample is limited to individuals who are older than 26 and younger than 65. Individuals below the age of 26 were affected by the dependent care mandate of the ACA, which went into effect on September 23, 2010. Since it expanded health insurance coverage in this group after 2011, we exclude them from our analysis. Individuals 65 and above are eligible for Medicare, and therefore were not directly affected by Medicaid expansion and the individual mandate. Since the ACA did not make any significant changes in Medicare, we use this group for a falsification exercise.

In the NHIS, respondents were asked how many days of work they missed "*because of an illness or injury in the last 12 months*" (Blewett et al., 2016). We use answers to this question to construct our dependent variables. Since interviews are conducted throughout a given year, and this question refers to absence from work during the 12 months preceding an interview, the relevant periods do not match calendar years. For example, an individual interviewed in March of 2014, would report the number of days he/she missed work between March of 2013 and February of 2014. This is an important point because the ACA provisions that pertain to our study (individual mandate and Medicaid expansion) were implemented on January 1, 2014. In our baseline analysis, we treat the survey year 2014 as a pre-treatment year (since respondents reported their experiences of 2013, at least in part). The post-ACA binary variable takes the value one if the survey year is 2015, 2016, 2017, or 2018 (representing the data year 2014, 2015, 2016, or 2017 respectively) and zero otherwise. Therefore, even though we use data from survey years 2005-2018, the data pertains to years 2004-2017. In the rest of the chapter, when we refer to the year variable, we refer to the year of the data, as opposed to the interview year. As a part of our robustness checks, we excluded 2014 from our analysis, but the results are not sensitive to this change (see Section 3.4.2 for details).

We impose some sample restrictions to reduce the effect of outliers. We exclude individuals with more than 75 days of missed work per year (99th percentile of days absent distribution). We also exclude individuals with Body Mass Index (BMI) below 10 and above 60. In addition, we restrict our attention to individuals who worked for the full year since the question about absenteeism in the NHIS asks about the number of days of missed work in the previous 12 months. Individuals who only worked part of the year may miss fewer days simply because they did not work for the whole year. Our qualitative results are not sensitive to any of these sample restrictions (see Section 3.4.2). We also exclude women that were pregnant at the time of interview.

We present the summary statistics in Table 3.6.1. We organized the summary statistics for the full sample[3] separated into four weight categories: underweight (BMI $<$ 18.5), normal-weight (18.5 $\leq$ BMI $<$ 25), overweight (25 BMI $<$ 30), and obese (BMI $\geq$ 30). About 59.3 percent of normal-weight respondents are female, and 48.2 percent of obese respondents are female. The average age among normal-weight respondents is 43.4, and among obese respondents, it is 45.1. The percentage of married respondents varies between 50.0 percent and 55.9 percent among the groups. The average number of children among normal-weight individuals is 0.81, compared to 0.88 among obese individuals. About 66.0 percent (60.9 percent) of normal-weight (obese) respondents are non-Hispanic, White. Normal-weight respondents are more likely to have college degrees (46.1 percent) compared to obese respondents (28.7 percent). In addition, they have better self-reported health status; 42.5 of normal-weight respondents have excellent health compared to 18.6 percent of obese respondents. The summary statistics show that overweight (obese) individuals are different from normal-weight individuals in several observable dimensions, which imposes a challenge for our analysis. We discuss this issue in more details in the Results section.

3.4 Results

In this section, we present the empirical results. We compare the outcomes for our three groups of interest (underweight, overweight, and obese respondents)

[3]The descriptive statistics separated by gender and then by weight categories are presented in appendix Tables 3.A.1 and 3.A.2

to normal-weight respondents. Based on the descriptive statistics (Table 3.6.1), normal-weight group differs from our groups of interest in baseline covariates. Therefore, we refer to the normal-weight individuals as a comparison group (as opposed to a control group). Since the objective is to difference out the trends in absenteeism that would have affected all workers, a comparison group may be sufficient (Kossoudji and Cobb-Clark, 2002). However, we still need the parallel trend assumption to hold. We test for it and present evidence to that effect in Section 3.4.1.2.1.

First, we present the results using mean DD and then DD regressions (Section 3.4.1). In Section 3.4.1.2.1, we show that the parallel trend assumption holds and therefore DD estimates may represent the causal effect of the ACA on absenteeism of overweight and obese workers. In Section 3.4.2, we show a number of robustness checks. Section 3.4.3 presents a falsification exercise using individuals 65 and older. In section 3.4.4, we explore how the ACA affected absenteeism.

3.4.1 Baseline Results

3.4.1.1 Mean DD Results

Table 3.6.2 shows the mean DD estimates for our two outcome variables. Panel A shows the estimates for the rate of absenteeism (i.e., probability of missing at least one day of work in the 12 months preceding an interview) and panel B shows the estimates for the number of days absent. Panel A (column 1) shows that 43.8 percent of normal-weight individuals missed at least one day of work before the ACA, and it declined to 42.8 percent in the post-ACA period; a decline of 1.0 percentage points.

On the other hand, 50.3 percent of obese individuals missed at least one day of work before the ACA, but 48.2 percent after the ACA; a decline of 2.1 percentage points. Thus, the mean DD estimate suggests that the ACA reduced the rate of absenteeism among obese people by 1.1 percentage point. However, this estimate is not statistically significant at conventional levels. Similarly, the mean DD estimate suggests that the ACA reduced the rate of absenteeism among overweight individuals by 1.2 percentage points (not significant).

Panel B in Table 3.6.2 shows that individuals in the normal-weight category missed 2.20 days of work per year before the ACA and declined by 0.09 days to 2.11 days of missed work in the post-ACA period. On the other hand, individuals in the obese category missed 3.25 days of work per year before the ACA, and it declined by 0.21 days to 3.04 days per year. Thus, the mean DD estimate suggests that the ACA reduced absenteeism among obese individuals by 0.12 days per year (not significant). Similarly, the mean DD estimate suggests that the ACA reduced the absenteeism among overweight individuals by 0.18 days per year (significant at 5 percent level).

Tables 3.A.3 and 3.A.4 in the appendix show the mean DD estimates for women and men respectively. Estimates suggest that the effect of the ACA is concentrated among women. For example, the rate of absenteeism declined by a significant 2.1 percentage points among overweight women (Panel A of Table 3.A.3) vs. only 0.8 percentage points (not significant) among men (Panel A of Table 3.A.4). A similar result holds when the number of days absent is the outcome variable.

3.4.1.2 Regression Results

The mean difference-in-difference estimates are consistent if the treatment and control groups are similar, except for the treatment status. However, we do not have a control group; our identification relies on a comparison group. This also implies that we need to control for observable differences across groups using a regression framework. Controlling for observable characteristics may also reduce standard errors and may lead to more precise estimates.

In our regressions, we include controls for age (quadratic), educational categories (less than High School, High School, Some College, College Degree), marital status (married or not), number of children, whether the youngest child is less than six years old, health status (very good, good, fair, and poor; excellent health is the omitted category), race, gender, region of residence, and survey year. We analyze the effects of the ACA on two outcome variables: the probability of absenteeism and number of days absent from work.

Before we discuss regression results we have to discuss parallel trend assumption. The DD estimates may be interpreted as causal effects of the expansion of insurance coverage brought on by the ACA if the parallel trends assumption holds. This assumption requires that the trends in absenteeism among treatment groups (obese, overweight, underweight individuals) would have been the same as the trend in the comparison group (normal-weight individuals) in the absence of the ACA. This assumption is not testable. As an alternative, econometric studies check whether the pre-intervention trends were similar across groups. Since we have multiple periods of pre-intervention data, we check whether the trend in absenteeism was the same in the pre-intervention period, i.e., between the years of 2004 and 2013.

3.4.1.2.1 Parallel Trends We use a regression-based approach to test for parallel trends. We estimate regressions that interact with the three treatment group (obese, overweight, and underweight) indicators with year indicator variables for all years except 2004, which is our base year. If the trend in absenteeism among obese individuals was the same as the trend in absenteeism among normal-weight individuals then all the coefficients of the interaction terms (i.e., obese interacted with the years) for the years, 2005-2013 should be jointly insignificant. To formally test the parallel trends assumption, we test the null hypothesis that the coefficients of all pre-2014 (2005-2013) interaction terms (i.e., obese * year dummy) jointly equal zero. Table 3.6.3 presents the p-values for the test of joint significance for pre-ACA (2005-2013) interaction terms for all six samples used in Table 3.6.4. Appendix Figures 3.B.1, 3.B.3, and 3.B.3 present the (year*treatment group) coefficients (when number of days absent is the outcome variable) and marginal effects (when the outcome variable is absent or not) for all respondents, women, and men, respectively. Figures show that all the coefficients in the pre-ACA period are statistically insignificant. Results of F-tests (presented in Table 3.6.3) suggest that in all six samples, the interaction terms are also jointly insignificant, suggesting that the pre-intervention trend was the same across groups and therefore the use of a DD structure is appropriate.

3.4.1.2.2 Regression Estimates Table 3.6.4 presents the regression estimates. In the first three columns, the outcome variable is whether the respondent missed any work or not. Since the outcome variable is binary, we use Probit regressions.

The estimates presented in columns 1-3 of Table 3.6.4 are average marginal effects. In the last three columns, the outcome variable is number of days of missed work. Since our second outcome variable (number of workdays missed in a year) is zero for about 56 percent of the respondents, we use a Tobit model to account for censoring.

We only report the coefficients of primary interest in the text. The complete table with estimates for control variables is in the Appendix Table 3.A.5. Our estimates suggest that among obese individuals, the probability of being absent declined by 1.3 percentage points in the post-ACA period (not significant). Our results also suggest that the probability of being absent declined by 1.5 percentage points (significant at 5 percent) among overweight workers. The ACA did not affect the probability of absenteeism among underweight workers.

Columns 2 and 3 of Table 3.6.4 present the results for women and men, respectively. Results suggest that the effect of the ACA is more extensive (in absolute values) among women compared to men. In the post-ACA period, the probability of being absent declined by 2.3 percentage points among obese women (significant at 5 percent) compared to 0.4 percentage points (not significant) among obese men. It is worth noting that women were about eight percentage points more likely to miss work compared to men in the pre-ACA period (46.7 percent for women vs. 38.7 percent for men). In other words, they had more room for improvement. In the post-ACA period, the probability of being absent declined by 2.6 percentage points among overweight women (significant at 5 percent) compared to 0.9 percentage points among overweight men (not significant).

Columns 4-6 present the results from Tobit regressions with the number of days of missed work as the outcome variable. The fourth column shows the results for all workers, and the next two columns split the sample by gender. Estimates in column 4 suggest that the number of days absent from work declined by 0.32 days for obese workers and 0.48 days for overweight workers. There was no statistically significant change in absenteeism for underweight individuals. Given the pre-ACA obese (overweight) workers missed 3.2 (2.3) days of work on average, these estimates suggest a 10.3 percent (20.0 percent) reduction in the number of days absent among the obese (overweight) workers. Estimates in columns 5 (women) and 6 (men) suggest that the effect of the ACA is stronger among women compared to men. This is consistent with the results in columns 2 and 3. In the post-ACA period, the number of days absent declined by 0.68 days among obese women (significant at 5 percent) compared to 0.05 days among obese men (not significant). Again, we should note that women missed more days of work compared to men in the pre-ACA period (3.05 days for women vs. 2.20 days for men). A similar result holds for overweight individuals. In the post-ACA period, the number of missed workdays declined by 0.81 days among overweight women (significant at 1 percent) compared to 0.27 days among overweight men (not significant). Therefore, our results suggest that there was a substantial reduction in absenteeism among overweight and obese women in the post-ACA period.

3.4.2 Robustness Checks

3.4.2.1 Excluding the Survey Year 2014 from Data

As we discussed earlier, the data from the survey year 2014 covers part of 2013 and part of 2014. In our baseline analysis (Table 3.6.4), we assumed that the data from the survey year 2014 are part of the pre-ACA period. In this section, we exclude the data from the survey year 2014 since technically it covers part of the pre-ACA and post-ACA period. The results are presented in Panel A of Table 3.6.5. The results for both men and women samples are similar to the results presented in Table 3.6.4 in both Probit and Tobit regressions. The estimates for overweight respondents in the full sample (men and women combined) are also similar to Table 3.6.4, but the coefficients for obese are not statistically significant anymore. Therefore, overall, these results suggest that the baseline results are robust to this change.

3.4.2.2 Robustness to Sample Selection Criteria

In our baseline analysis (Table 3.6.4), we excluded individuals who worked for less than 12 months in the year before an interview. In this section, we include those individuals. The sample size increases from 156,623 to 194,651. The results for this sample are presented in Panel B of Table 3.6.5. All the results are similar to Table 3.6.4, suggesting that results are robust to this assumption as well.

3.4.2.3 Controlling for Occupation

In Panel C, we add occupation-specific fixed effects (one digit level) as additional controls. Since different occupations require different amounts of physical intensity, identification from within occupation variation may provide better information. However, occupation selection may be endogenous. Nonetheless, we control for occupation as a robustness check. The results (presented in Panel C of Table 3.6.5) are similar to our baseline results in Table 3.6.4. However, the coefficient for obese individuals loses statistical significance in Probit regressions.

3.4.3 Counterfactual Using Respondents 65 and Older

While the ACA increased the percentage of non-elderly insured individuals (age<65), it did not change the percentage of elderly (age>=65) who have health insurance. For example, 79.1 percent of 27-64-year-olds had had health insurance in 2005 (DeNavas-Walt et al., 2013), which increased to 88.1 percent by 2016 (Barnett and Berchick, 2017) an increase of 9.0 percentage points. On the other hand, among the elderly (age>=65) the percentage of individuals with health insurance increased only slightly from 98.8 percent in 2005 (DeNavas-Walt et al., 2013) to 98.9 percent in 2016 (Barnett and Berchick, 2017). Therefore, if the expansion of health insurance drives the reduction in absenteeism in the post-ACA period then we would not expect to see an effect in individuals 65 and over. This group (age>=65), therefore, provides us with an opportunity to test whether the effect reported above is due to unrelated time effects. Table

3.6.6 presents the results. Estimates suggest that neither the probability of missing work nor the number of days of missed work changed after the ACA among the elderly (age >=65) obese and overweight workers. This is true in the full sample, among women, and men. This suggests that the decline in absenteeism is indeed due to health insurance coverage expansion due to the ACA.

3.4.4 Why Did the Absenteeism Decline?

In this sub-section, we explore potential mechanisms through which the ACA may have affected absenteeism among workers who are overweight or obese. In the Introduction section, we noted that several researchers have argued that access to health insurance may allow them to better manage their health conditions leading to lower absenteeism. Table 3.6.7 shows the effect of the ACA on the number of days an individual was forced to spend in bed due to an illness or injury[4]. Estimates from Tobit regressions suggest that compared to normal-weight women, the number of disability days went down by 0.80 days among over-weight women (statistically significant) and 0.29 days among obese women (not significant). In Table 3.6.4, we reported that the number of days absent went down by 0.81 days among overweight women and 0.67 days among obese women. These numbers are consistent with that result.

[4] The exact question was “During the PAST 12 MONTHS, how many days did illness or injury keep you in bed more than half of the day(include days while an overnight patient in a hospital)?"

3.5 Conclusion

In this chapter, we explore whether the health insurance expansion brought on by the ACA reduced absenteeism among overweight and obese workers. To estimate the effect of the ACA, we use a difference-in-differences structure with normal-weight individuals as our comparison group. Our results suggest that the effects on women are both economically and statistically significant. The probability of being absent declined by about 2.3 (2.6) percentage points in the post-ACA period among obese (overweight) women. Furthermore, our estimates (using a Tobit model) indicate that obese (overweight) women missed 0.675 (0.810) fewer days after the ACA, compared to normal-weight women. On the other hand, the effects on men are smaller and often statistically insignificant. Our results also suggest that improved physical health outcomes led to this reduced absenteeism.

These results are consistent with the hypothesis that access to healthcare may allow overweight and obese workers to address the chronic conditions which may limit their ability to consistently be present at work. Since normal-weight and underweight workers are less likely to have health problems, access to health insurance may not be as important. Apart from being statistically significant, the reduction in absenteeism is economically significant. Women who are obese, on average, missed about 3.9 days of work. Our estimates suggest that in the post-ACA period, absenteeism in this group was reduced by about 0.68 days, which is a 17 percent reduction in absenteeism. Cawley et al. (2007) estimated that the cost associated with absenteeism among all (men and women) obese workers is about $4.3 billion per year. Therefore, a 17 percent reduction in absenteeism among obese women may save about $350 million per year.

3.6 Tables

Table 3.6.1: Summary Statistics

	Normal Weight	Underweight	Overweight	Obese
Work loss days	2.184	2.482	2.325	3.201
	(5.746)	(6.788)	(6.249)	(7.552)
Lost any work days	0.436	0.448	0.433	0.498
	(0.496)	(0.497)	(0.495)	(0.500)
Body mass index	22.548	17.608	27.287	35.173
	(1.633)	(0.864)	(1.444)	(4.872)
Female	0.593	0.816	0.374	0.482
	(0.491)	(0.388)	(0.484)	(0.500)
Age	43.357	42.115	44.816	45.110
	(10.606)	(10.860)	(10.377)	(10.209)
Married	0.519	0.501	0.559	0.520
	(0.500)	(0.500)	(0.496)	(0.500)
Number of own children	0.806	0.765	0.874	0.884
	(1.091)	(1.069)	(1.138)	(1.158)
Youngest child less than 6	0.151	0.153	0.155	0.137
	(0.358)	(0.360)	(0.362)	(0.343)
Race				
White, Non-hispanic	0.660	0.659	0.626	0.609
	(0.474)	(0.474)	(0.484)	(0.488)
Black, Non-hispanic	0.091	0.071	0.125	0.180
	(0.287)	(0.257)	(0.331)	(0.384)
Hispanic	0.132	0.078	0.181	0.170
	(0.339)	(0.268)	(0.385)	(0.376)
Other	0.117	0.192	0.068	0.041
	(0.322)	(0.394)	(0.252)	(0.199)
Education				
Less than HS	0.073	0.066	0.102	0.101
	(0.260)	(0.249)	(0.303)	(0.301)
High School	0.198	0.224	0.231	0.263
	(0.399)	(0.417)	(0.422)	(0.440)
Some College	0.268	0.250	0.297	0.349
	(0.443)	(0.433)	(0.457)	(0.477)
College Degree	0.461	0.461	0.369	0.287
	(0.498)	(0.499)	(0.483)	(0.453)
Health Status				
Excellent	0.425	0.385	0.332	0.186
	(0.494)	(0.487)	(0.471)	(0.389)
Very Good	0.357	0.334	0.388	0.361
	(0.479)	(0.472)	(0.487)	(0.480)
Good	0.181	0.212	0.233	0.353
	(0.385)	(0.409)	(0.423)	(0.478)
Fair	0.034	0.060	0.043	0.091
	(0.181)	(0.237)	(0.204)	(0.288)
Poor	0.003	0.009	0.004	0.009
	(0.059)	(0.094)	(0.062)	(0.096)
N	49,950	1,570	58,331	46,772

Means presented, standard deviations in parentheses. Work lost days are in last 12 months.

Table 3.6.2: Mean Difference-in-Differences

	Normal Weight	Underweight	Overweight	Obese
Panel A: Lost any work days				
Before	0.438***	0.456***	0.437***	0.503***
	[0.496]	[0.498]	[0.496]	[0.500]
After	0.428***	0.416***	0.415***	0.482***
	[0.495]	[0.494]	[0.493]	[0.500]
Difference	-0.010*	-0.039	-0.022***	-0.021***
	(0.006)	(0.032)	(0.005)	(0.006)
Diff-in-Diff	.	-0.029	-0.012	-0.011
	.	(0.032)	(0.008)	(0.008)
Panel B: Work days lost				
Before	2.201***	2.489***	2.379***	3.247***
	[5.774]	[6.787]	[6.350]	[7.579]
After	2.110***	2.451***	2.106***	3.035***
	[5.629]	[6.805]	[5.820]	[7.453]
Difference	-0.091	-0.039	-0.273***	-0.212**
	(0.064)	(0.440)	(0.062)	(0.084)
Diff-in-Diff	.	0.052	-0.182**	-0.121
	.	(0.445)	(0.089)	(0.106)
N	49,950	1,570	58,331	46,772

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$; Standard Errors in parentheses, Standard Deviations in square brackets.

Table 3.6.3: Parallel Trend Test

	Probit			Tobit		
	Full Sample	Women	Men	Full Sample	Women	Men
Underweight	0.582	0.219	0.224	0.417	0.366	0.740
Overweight	0.379	0.587	0.256	0.129	0.367	0.242
Obese	0.473	0.220	0.863	0.409	0.658	0.588
N	156,623	75,273	81,350	156,623	75,273	81,350

Controls included in the model are: Sex, marital status, age, age squared, race dummies, educational attainment dummies, number of children, and an indicator if the youngest child is less than 6 years old. P-values for test of joint significance of coefficients for years 2005 to 2013 are presented.

Table 3.6.4: Regressions Results Showing the Effect of the ACA

	Probit			Tobit		
	Full Sample	Women	Men	Full Sample	Women	Men
Underweight X Post-ACA	-0.022	-0.012	-0.055	-0.087	-0.163	0.454
	(0.032)	(0.036)	(0.075)	(0.862)	(0.931)	(2.277)
Overweight X Post-ACA	-0.015**	-0.026**	-0.009	-0.475**	-0.810***	-0.266
	(0.008)	(0.011)	(0.011)	(0.188)	(0.269)	(0.276)
Obese X Post-ACA	-0.013	-0.023**	-0.004	-0.319	-0.675**	0.048
	(0.008)	(0.011)	(0.011)	(0.200)	(0.269)	(0.301)
Region Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
N	156,623	75,273	81,350	156,623	75,273	81,350

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$; Standard errors in parentheses; Columns (1) to (3) display average marginal effects for discrete change of dummy variable from 0 to 1; all models include constants. Controls included in the model are: Sex, marital status, age, age2, race, educational attainment, number of children, and an indicator if the youngest child is less than 6 years old.

Table 3.6.5: Robustness Checks

	Probit			Tobit		
	Full Sample	Women	Men	Full Sample	Women	Men
Panel A: Excluding 2014						
Underweight X Post-ACA	-0.021	-0.011	-0.053	-0.175	-0.375	0.994
	(0.032)	(0.036)	(0.075)	(0.862)	(0.921)	(2.341)
Overweight X Post-ACA	-0.016*	-0.027*	-0.011	-0.511**	-0.861**	-0.317
	(0.007)	(0.011)	(0.011)	(0.189)	(0.270)	(0.278)
Obese X Post-ACA	-0.014	-0.022*	-0.007	-0.393	-0.736**	-0.056
	(0.008)	(0.011)	(0.011)	(0.201)	(0.271)	(0.303)
N	146,406	70,367	76,039	146,406	70,367	76,039
Panel B: Including workers with less than 12 months of work						
Underweight X Post-ACA	-0.028	-0.020	-0.060	-0.861	-0.960	-0.267
	(0.028)	(0.032)	(0.067)	(0.792)	(0.844)	(2.223)
Overweight X Post-ACA	-0.015*	-0.026**	-0.007	-0.535**	-0.870***	-0.281
	(0.007)	(0.010)	(0.010)	(0.186)	(0.261)	(0.277)
Obese X Post-ACA	-0.014*	-0.023*	-0.004	-0.432*	-0.712**	-0.099
	(0.007)	(0.009)	(0.010)	(0.195)	(0.261)	(0.299)
N	194,651	96,612	98,039	194,651	96,612	98,039
Panel C: Controlling for Occupation						
Underweight X Post-ACA	-0.017	-0.008	-0.052	-0.138	-0.387	1.067
	(0.032)	(0.036)	(0.075)	(0.870)	(0.930)	(2.356)
Overweight X Post-ACA	-0.015*	-0.024*	-0.010	-0.491*	-0.799**	-0.301
	(0.008)	(0.011)	(0.011)	(0.191)	(0.272)	(0.279)
Obese X Post-ACA	-0.012	-0.020	-0.005	-0.349	-0.667*	-0.035
	(0.008)	(0.011)	(0.011)	(0.202)	(0.272)	(0.304)
N	143,626	69,176	74,450	143,626	69,176	74,450

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$; Standard errors in parentheses; Columns (1) to (3) display average marginal effects for discrete change of dummy variable from 0 to 1; all models include controls and constants. Controls included in the model are: Sex, marital status, age, age squared , race dummies, educational attainment dummies, number of children, and an indicator if the youngest child is less than 6 years old.

Table 3.6.6: Results for Individuals 65 and Above

	Probit			Tobit		
	Full Sample	Women	Men	Full Sample	Women	Men
Underweight X Post-ACA	-0.047	-0.044	-0.072	-2.708	-2.544	-3.377
	(0.091)	(0.103)	(0.237)	(3.500)	(3.621)	(9.925)
Overweight X Post-ACA	-0.012	-0.020	-0.013	-0.889	-1.090	-1.110
	(0.022)	(0.032)	(0.030)	(0.950)	(1.232)	(1.497)
Obese X Post-ACA	-0.001	0.015	-0.022	-0.048	0.104	-0.610
	(0.024)	(0.034)	(0.034)	(1.064)	(1.324)	(1.716)
N	11,959	5,786	6,171	11,959	5,788	6,171

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$; Standard errors in parentheses; Columns (1) to (3) display average marginal effects for discrete change of dummy variable from 0 to 1; all models include controls and constants. Controls included in the model are: Sex, marital status, age, age squared , race dummies, educational attainment dummies, number of children, and an indicator if the youngest child is less than 6 years old.

Table 3.6.7: Tobit Regressions of Bed Disability Days, Past 12 Months

	Full Sample	Women	Men
Overweight X Post-ACA	-0.511***	-0.804**	-0.165
	(0.205)	(0.325)	(0.248)
Obese X Post-ACA	-0.264	-0.293	-0.127
	(0.224)	(0.339)	(0.273)
N	156,623	75,273	81,350

Standard errors in parentheses; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. All models include controls and constants. Controls included in the model are: Sex, marital status, age, age squared , race dummies, educational attainment dummies, number of children, and an indicator if the youngest child is less than 6 years old.

Appendix

3.A Tables

Table 3.A.1: Summary Statistics - Women Only

	Normal Weight	Underweight	Overweight	Obese
Work loss days	2.368	2.541	2.870	3.858
	(5.968)	(6.915)	(6.957)	(8.294)
Lost any work days	0.469	0.458	0.497	0.562
	(0.499)	(0.498)	(0.500)	(0.496)
Body mass index	22.267	17.637	27.308	35.928
	(1.696)	(0.780)	(1.428)	(5.333)
Female	1.000	1.000	1.000	1.000
	(0.000)	(0.000)	(0.000)	(0.000)
Age	43.782	42.129	45.411	45.265
	(10.503)	(10.798)	(10.380)	(10.288)
Married	0.515	0.504	0.495	0.425
	(0.500)	(0.500)	(0.500)	(0.494)
Number of own children	0.855	0.799	0.916	0.900
	(1.083)	(1.078)	(1.121)	(1.133)
Youngest child less than 6	0.142	0.148	0.130	0.120
	(0.349)	(0.356)	(0.336)	(0.325)
Race				
White, Non-hispanic	0.675	0.680	0.594	0.565
	(0.469)	(0.467)	(0.491)	(0.496)
Black, Non-hispanic	0.089	0.060	0.168	0.232
	(0.285)	(0.238)	(0.374)	(0.422)
Hispanic	0.131	0.070	0.178	0.163
	(0.338)	(0.256)	(0.382)	(0.369)
Other	0.105	0.190	0.061	0.040
	(0.307)	(0.392)	(0.239)	(0.196)
Education				
Less than HS	0.058	0.047	0.095	0.095
	(0.233)	(0.211)	(0.293)	(0.293)
High School	0.183	0.207	0.226	0.243
	(0.387)	(0.405)	(0.418)	(0.429)
Some College	0.284	0.259	0.331	0.373
	(0.451)	(0.438)	(0.470)	(0.484)
College Degree	0.475	0.487	0.349	0.289
	(0.499)	(0.500)	(0.477)	(0.453)
Health Status				
Excellent	0.428	0.414	0.291	0.163
	(0.495)	(0.493)	(0.454)	(0.370)
Very Good	0.363	0.337	0.391	0.352
	(0.481)	(0.473)	(0.488)	(0.478)
Good	0.176	0.191	0.260	0.373
	(0.381)	(0.393)	(0.439)	(0.483)
Fair	0.030	0.050	0.053	0.101
	(0.171)	(0.218)	(0.223)	(0.301)
Poor	0.004	0.008	0.004	0.011
	(0.059)	(0.088)	(0.067)	(0.103)
N	29,641	1,281	21,801	22,550

Means presented, standard deviations in parentheses. Work lost days are in last 12 months.

Table 3.A.2: Summary Statistics - Men Only

	Normal Weight	Underweight	Overweight	Obese
Work loss days	1.915	2.221	1.999	2.590
	(5.395)	(6.199)	(5.761)	(6.732)
Lost any work days	0.389	0.405	0.394	0.439
	(0.488)	(0.492)	(0.489)	(0.496)
Body mass index	22.959	17.480	27.275	34.470
	(1.440)	(1.160)	(1.454)	(4.283)
Female	0.000	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)
Age	42.736	42.052	44.461	44.966
	(10.724)	(11.146)	(10.358)	(10.133)
Married	0.526	0.484	0.598	0.609
	(0.499)	(0.501)	(0.490)	(0.488)
Number of own children	0.734	0.612	0.848	0.870
	(1.100)	(1.018)	(1.148)	(1.181)
Youngest child less than 6	0.164	0.173	0.170	0.152
	(0.371)	(0.379)	(0.376)	(0.359)
Race				
White, Non-hispanic	0.638	0.567	0.645	0.650
	(0.481)	(0.496)	(0.478)	(0.477)
Black, Non-hispanic	0.093	0.121	0.100	0.131
	(0.291)	(0.327)	(0.300)	(0.337)
Hispanic	0.134	0.111	0.182	0.177
	(0.341)	(0.314)	(0.386)	(0.381)
Other	0.135	0.201	0.073	0.043
	(0.341)	(0.401)	(0.259)	(0.202)
Education				
Less than HS	0.095	0.152	0.107	0.106
	(0.294)	(0.360)	(0.309)	(0.308)
High School	0.220	0.298	0.234	0.282
	(0.414)	(0.458)	(0.424)	(0.450)
Some College	0.245	0.208	0.278	0.326
	(0.430)	(0.406)	(0.448)	(0.469)
College Degree	0.440	0.343	0.381	0.286
	(0.496)	(0.475)	(0.486)	(0.452)
Health Status				
Excellent	0.421	0.260	0.356	0.207
	(0.494)	(0.439)	(0.479)	(0.405)
Very Good	0.347	0.318	0.386	0.369
	(0.476)	(0.467)	(0.487)	(0.482)
Good	0.189	0.304	0.217	0.334
	(0.391)	(0.461)	(0.412)	(0.472)
Fair	0.040	0.104	0.038	0.082
	(0.195)	(0.306)	(0.191)	(0.275)
Poor	0.003	0.014	0.003	0.008
	(0.058)	(0.117)	(0.059)	(0.089)
N	20,309	289	36,530	24,222

Means presented, standard deviations in parentheses. Work lost days are in last 12 months.

Table 3.A.3: Mean Difference-in-Differences, Women Only

	Normal Weight	Underweight	Overweight	Obese
Panel A: Lost any work days				
Before	0.471***	0.465***	0.504***	0.568***
	[0.499]	[0.499]	[0.500]	[0.495]
After	0.459***	0.429***	0.471***	0.540***
	[0.498]	[0.496]	[0.499]	[0.498]
Difference	-0.012	-0.036	-0.032***	-0.028***
	(0.007)	(0.035)	(0.008)	(0.008)
Diff-in-Diff	.	-0.024	-0.021*	-0.016
	.	(0.036)	(0.011)	(0.011)
Panel B: Work days lost				
Before	2.388***	2.580***	2.970***	3.956***
	[5.986]	[7.039]	[7.147]	[8.417]
After	2.285***	2.371***	2.466***	3.491***
	[6.359]	[6.112]	[7.807]	
Difference	-0.102	-0.209	-0.504***	-0.465***
	(0.087)	(0.464)	(0.108)	(0.129)
Diff-in-Diff	.	-0.107	-0.402***	-0.363**
	.	(0.472)	(0.139)	(0.156)
N	29,641	1,281	21,801	22,550

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$; Standard Errors in parentheses, Standard Deviations in square brackets.

Table 3.A.4: Mean Difference-in-Differences, Men Only

	Normal Weight	Underweight	Overweight	Obese
Panel A: Lost any work days				
Before	0.391***	0.415***	0.397***	0.441***
	[0.488]	[0.494]	[0.489]	[0.497]
After	0.383***	0.358***	0.381***	0.431***
	[0.486]	[0.484]	[0.486]	[0.495]
Difference	-0.008	-0.057	-0.016**	-0.011
	(0.009)	(0.073)	(0.006)	(0.008)
Diff-in-Diff	.	-0.049	-0.008	-0.003
	.	(0.074)	(0.011)	(0.012)
Panel B: Work days lost				
Before	1.929***	2.089***	2.026***	2.580***
	[5.438]	[5.534]	[5.794]	[6.627]
After	1.855***	2.811**	1.892***	2.626***
	[5.208]	[8.602]	[5.628]	[7.095]
Difference	-0.074	0.722	-0.134*	0.046
	(0.094)	(1.224)	(0.074)	(0.109)
Diff-in-Diff	.	0.796	-0.060	0.120
	.	(1.228)	(0.120)	(0.143)
N	20,309	289	36,530	24,222

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$; Standard Errors in parentheses, Standard Deviations in square brackets.

Table 3.A.5: Regressions With 4 Weight Categories

	Probit			Tobit		
	Full Sample	Women	Men	Full Sample	Women	Men
Underweight X Post-ACA	-0.022	-0.012	-0.055	-0.087	-0.163	0.454
	(0.032)	(0.036)	(0.075)	(0.862)	(0.931)	(2.277)
Overweight X Post-ACA	-0.015**	-0.026**	-0.009	-0.475**	-0.810***	-0.266
	(0.008)	(0.011)	(0.011)	(0.188)	(0.269)	(0.276)
Obese X Post-ACA	-0.013	-0.023**	-0.004	-0.319	-0.675**	0.048
	(0.008)	(0.011)	(0.011)	(0.200)	(0.269)	(0.301)
Post-ACA	0.020**	0.011	0.029**	0.174	0.089	0.276
	(0.008)	(0.011)	(0.012)	(0.201)	(0.270)	(0.304)
Underweight	-0.018	-0.020	0.013	-0.302	-0.187	-0.125
	(0.014)	(0.015)	(0.032)	(0.361)	(0.406)	(0.787)
Overweight	0.028***	0.040***	0.013***	0.673***	0.967***	0.275**
	(0.003)	(0.005)	(0.005)	(0.086)	(0.124)	(0.123)
Obese	0.063***	0.083***	0.040***	1.599***	2.066***	0.960***
	(0.004)	(0.005)	(0.005)	(0.095)	(0.130)	(0.139)
Female	0.096***			2.219***		
	(0.003)			(0.067)		
Married	-0.020***	-0.031***	-0.012***	-0.677***	-0.969***	-0.388***
	(0.003)	(0.004)	(0.004)	(0.070)	(0.096)	(0.106)
Age	-0.001	0.002	-0.003**	0.006	0.100**	-0.090**
	(0.001)	(0.002)	(0.002)	(0.030)	(0.042)	(0.042)
Age squared	-0.000***	-0.000***	-0.000	-0.001**	-0.002***	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Black, Nonhispanic	-0.076***	-0.080***	-0.074***	-1.272***	-1.306***	-1.305***
	(0.004)	(0.005)	(0.006)	(0.108)	(0.146)	(0.163)
Hispanic	-0.111***	-0.112***	-0.109***	-2.450***	-2.290***	-2.612***
	(0.004)	(0.006)	(0.005)	(0.103)	(0.148)	(0.144)
Other	-0.092***	-0.105***	-0.083***	-1.916***	-2.095***	-1.802***
	(0.005)	(0.007)	(0.006)	(0.122)	(0.177)	(0.169)
Less than HS	-0.055***	-0.061***	-0.052***	-1.194***	-1.250***	-1.168***
	(0.005)	(0.008)	(0.006)	(0.147)	(0.221)	(0.197)
Some College	0.057***	0.061***	0.053***	1.233***	1.374***	1.095***
	(0.003)	(0.005)	(0.005)	(0.093)	(0.134)	(0.131)
College Degree	0.084***	0.087***	0.082***	1.196***	1.325***	1.050***
	(0.003)	(0.005)	(0.005)	(0.088)	(0.128)	(0.120)
Very Good	0.082***	0.086***	0.078***	2.028***	2.100***	1.948***
	(0.003)	(0.004)	(0.004)	(0.074)	(0.106)	(0.104)
Good	0.124***	0.125***	0.124***	3.666***	3.774***	3.539***
	(0.003)	(0.005)	(0.005)	(0.092)	(0.131)	(0.130)
Fair	0.210***	0.206***	0.213***	7.288***	7.453***	7.057***
	(0.006)	(0.008)	(0.008)	(0.183)	(0.256)	(0.261)
Poor	0.318***	0.306***	0.329***	13.314***	13.259***	13.297***
	(0.015)	(0.020)	(0.023)	(0.619)	(0.859)	(0.888)
Number of own children	-0.005***	-0.012***	0.001	-0.161***	-0.313***	-0.044
	(0.001)	(0.002)	(0.002)	(0.035)	(0.049)	(0.050)
Youngest child less than 6	0.004	-0.022***	0.019***	0.213**	-0.165	0.445***
	(0.004)	(0.006)	(0.006)	(0.103)	(0.151)	(0.142)
Region Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
N	156,623	75,273	81,350	156,623	75,273	81,350

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$; Marginal effects; Standard errors in parentheses. Columns (1) to (3) display average marginal effects for discrete change of dummy variable from 0 to 1; all models include constants.

3.B Figures

Figure 3.B.1: Testing for Parallel Trends: All Respondents

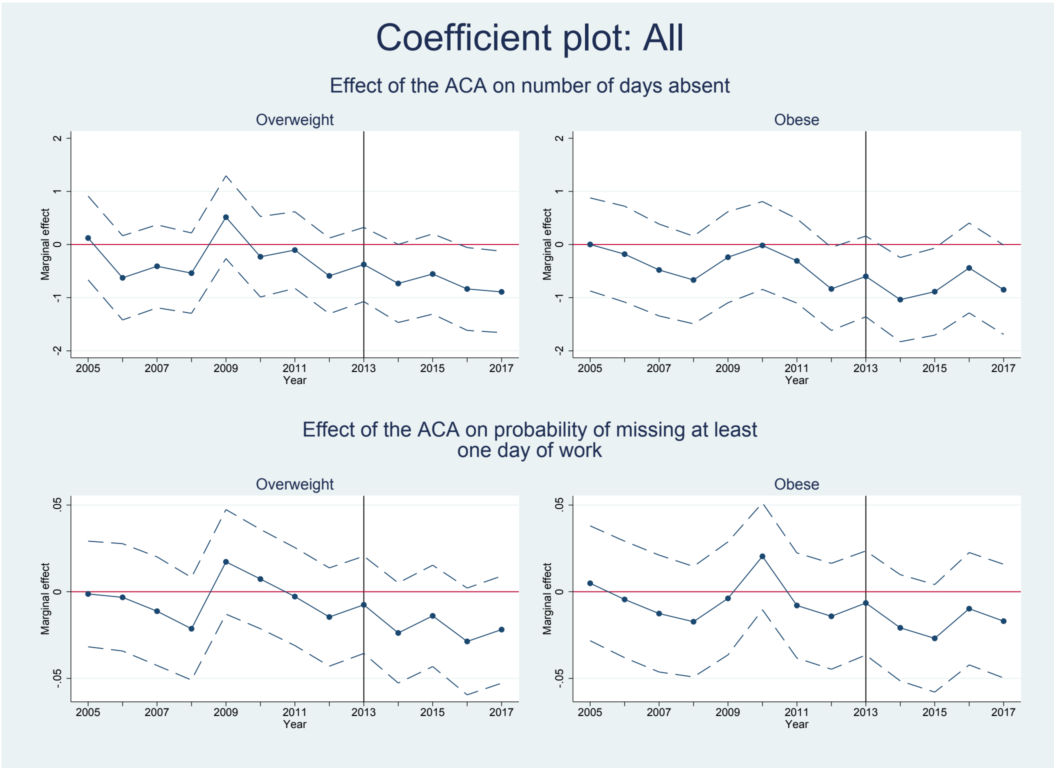


Figure 3.B.2: Testing for Parallel Trends: Women

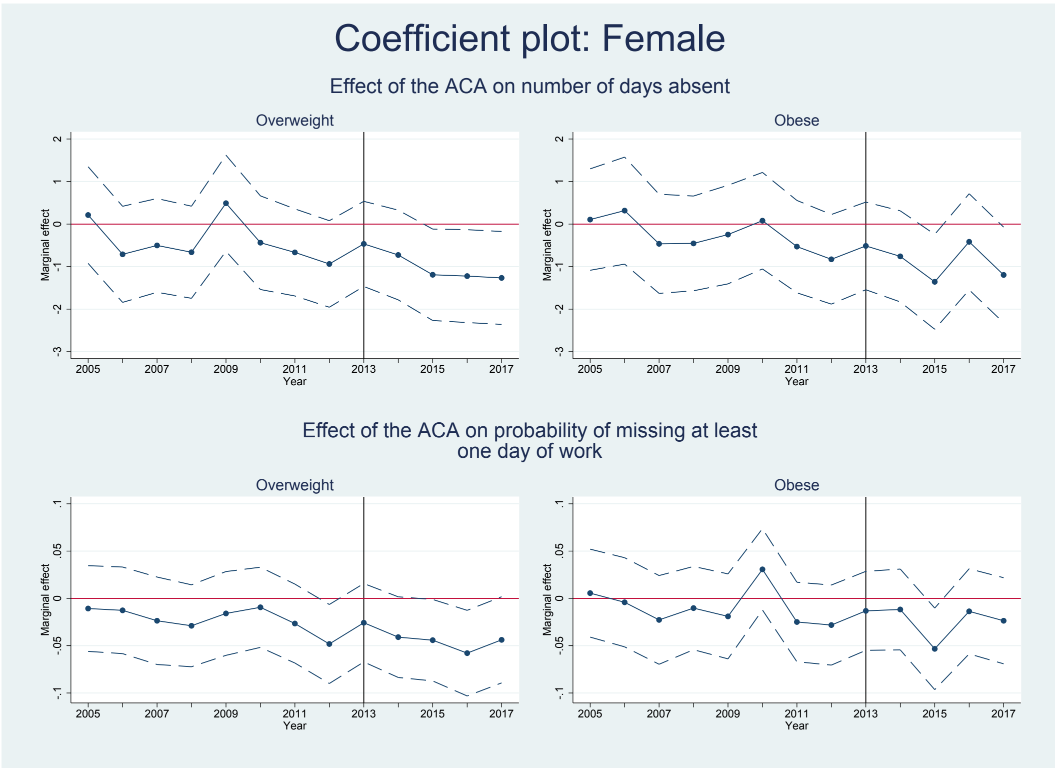
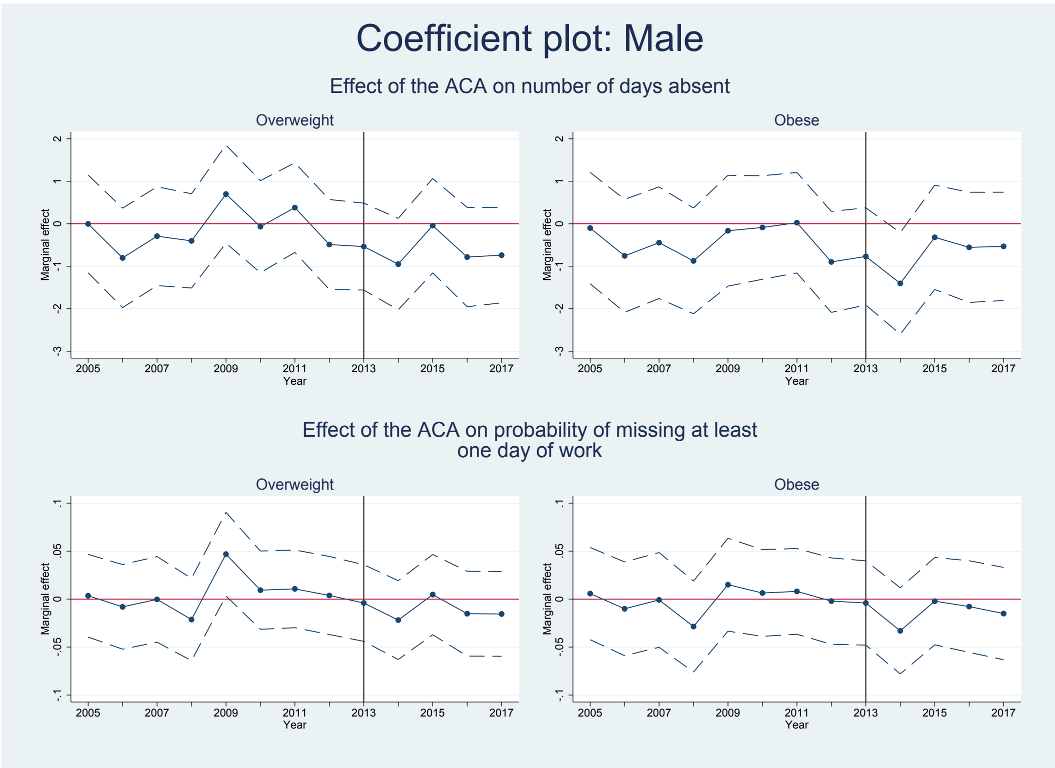


Figure 3.B.3: Testing for Parallel Trends: Men



Bibliography

Acemoglu, D., D. H. Autor, and D. Lyle (2004). Women, war, and wages: The effect of female labor supply on the wage structure at midcentury. *Journal of Political Economy 112*(3), 497–551.

Åstebro, T. and J. Tåg (2017). Gross, net, and new job creation by entrepreneurs. *Journal of Business Venturing Insights 8*, 64–70.

Ayyagari, P. (2019). Health insurance and early retirement plans: Evidence from the affordable care act. *American Journal of Health Economics 5*(4), 533–560.

Bailey, J. (2017). Health insurance and the supply of entrepreneurs: New evidence from the Affordable Care Act. *Small Business Economics 49*(3), 627–646.

Bailey, J. and D. Dave (2019). The effect of the Affordable Care Act on entrepreneurship among older adults. *Eastern Economic Journal 45*(1), 141–159.

Barber III, D. and T. Kavoori (2015). The effects of the Affordable Care Act on the self-employed. *Academy of Economics and Finance Journal 6*, 15–23.

Barber III, D. and T. Kavoori (2018). The Affordable Care Act, state exchanges and the self-employed in the USA. *Journal of Small Business & Entrepreneurship 30*(6), 499–517.

Barnett, J. C. and E. R. Berchick (2017). Health insurance coverage in the united states: 2016. united states census bureau. Technical report, Report.

Berndt, E. R., B. H. Hall, R. E. Hall, and J. A. Hausman (1974). Estimation and inference in nonlinear structural models. In *Annals of Economic and Social Measurement, Volume 3, number 4*, pp. 653–665. NBER.

Bhattacharya, J. and N. Sood (2006). *Health insurance and the obesity externality*. Emerald Group Publishing Limited.

Blewett, L. A., J. A. R. Drew, R. Griffin, M. L. King, and K. Williams (2016). Ipums health surveys: National health interview survey, version 6.2. *Minneapolis: University of Minnesota 10*, D070.

Blumberg, L. J., S. Corlette, and K. Lucia (2014). The Affordable Care Act: Improving incentives for entrepreneurship and self-employment. *Public Policy & Aging Report 24*(4), 162–167.

Brooks, T., L. Roygardner, S. Artiga, O. Pham, and R. Dolan (2020). Medicaid and chip eligibility, enrollment, and cost sharing policies as of january 2019: findings from a 50-state survey. *The Henry J. Kaiser Family Foundation*.

Caliendo, M., F. M. Fossen, and A. S. Kritikos (2019). What makes an employer-entrepreneur? DIW Berlin Discussion Paper 1829.

Cawley, J., J. A. Rizzo, and K. Haas (2007). Occupation-specific absenteeism costs associated with obesity and morbid obesity. *Journal of Occupational and Environmental Medicine 49*(12), 1317–1324.

Dahlen, H. M. (2015). “aging out" of dependent coverage and the effects on us labor market and health insurance choices. *American journal of public health 105*(S5), S640–S650.

DeLeire, T. and C. Marks (2015). Consumer decisions regarding health plan choices in the 2014 and 2015 marketplaces. *Off Assist Secr Plan Eval 21*.

DeNavas-Walt, C., B. D. Proctor, and J. C. Smith (2013). Income, poverty, and health insurance coverage in the united states: 2012. current population reports p60-245. *US Census Bureau*.

Dillon, E. W. and C. T. Stanton (2017). Self-employment dynamics and the returns to entrepreneurship. NBER Working Paper No. 23168, National Bureau of Economic Research.

Dizioli, A. and R. Pinheiro (2016). Health insurance as a productive factor. *Labour Economics 40*, 1–24.

Dohmen, T., A. Falk, D. Huffman, U. Sunde, J. Schupp, and G. G. Wagner (2011). Individual risk attitudes: Measurement, determinants, and behavioral consequences. *Journal of the European Economic Association 9*(3), 522–550.

Dong, X. S., X. Wang, K. Ringen, and R. Sokas (2017). Baby boomers in the united states: factors associated with working longer and delaying retirement. *American journal of industrial medicine 60*(4), 315–328.

Fairlie, R. W., K. Kapur, and S. Gates (2011). Is employer-based health insurance a barrier to entrepreneurship? *Journal of Health Economics 30*(1), 146–162.

Fairlie, R. W. and J. Miranda (2017). Taking the leap: The determinants of entrepreneurs hiring their first employee. *Journal of Economics & Management Strategy 26*(1), 3–34.

Finkelstein, E. A., M. daCosta DiBonaventura, S. M. Burgess, B. C. Hale, et al. (2010). The costs of obesity in the workplace. *Journal of Occupational and Environmental Medicine 52*(10), 971–976.

Fossen, F. M. (2009). Would a flat-rate tax stimulate entrepreneurship in Germany? A behavioural microsimulation analysis allowing for risk. *Fiscal Studies 30*(2), 179–218.

Fossen, F. M. and J. König (2017). Public health insurance, individual health, and entry into self-employment. *Small Business Economics 49*(3), 647–669.

Fossen, F. M., R. Rees, D. Rostam-Afschar, and V. Steiner (2020). The effects of income taxation on entrepreneurial investment: A puzzle? *International Tax and Public Finance 27*, 1321–1363.

Fossen, F. M. and V. Steiner (2009). Income taxes and entrepreneurial choice: Empirical evidence from two German natural experiments. *Empirical Economics 36*(3), 487–513.

French, E., H.-M. von Gaudecker, and J. B. Jones (2016). The effect of the affordable care act on the labor supply, savings, and social security of older americans. *Michigan Retirement Research Center Research Paper* (2016-354).

Gilleskie, D. B. (1998). A dynamic stochastic model of medical care use and work absence. *Econometrica*, 1–45.

Gilleskie, D. B. and B. F. Lutz (2002). The impact of employer-provided health insurance on dynamic employment transitions. *Journal of Human Resources 37*(1), 129–162.

Goebel, J., M. M. Grabka, S. Liebig, M. Kroh, D. Richter, C. Schröder, and J. Schupp (2019). The German Socio-economic Panel (SOEP). *Jahrbücher für Nationalökonomie und Statistik–Journal of Economics and Statistics 239*(2), 345–360.

Gooptu, A., A. S. Moriya, K. I. Simon, and B. D. Sommers (2016). Medicaid expansion did not result in significant employment changes or job reductions in 2014. *Health affairs 35*(1), 111–118.

Gruber, J. (1994). The incidence of mandated maternity benefits. *The American Economic Review 84*(3), 622–641.

Gumus, G. and T. L. Regan (2015). Self-employment and the role of health insurance in the us. *Journal of Business Venturing 30*(3), 357–374.

Gurley-Calvez, T. (2011). Will tax-based health insurance reforms help the self-employed stay in business? *Contemporary Economic Policy 29*(3), 441–460.

Gustman, A. L., T. L. Steinmeier, and N. Tabatabai (2016). The affordable care act as retiree health insurance: Implications for retirement and social security claiming. Technical report, National Bureau of Economic Research.

Haan, P. and V. Prowse (2014). Longevity, life-cycle behavior and pension reform. *Journal of Econometrics 178*, 582–601.

Hales, C., M. Carroll, C. Fryar, and C. Ogden (2020). Prevalence of obesity and severe obesity among adults: United States, 2017–2018. NCHS Data Brief, no 360. Available at https://www.cdc.gov/nchs/products/databriefs/db360.htm. Accessed April 21, 2021.

Haltiwanger, J., R. S. Jarmin, and J. Miranda (2013). Who creates jobs? Small versus large versus young. *Review of Economics and Statistics 95*(2), 347–361.

Heim, B., I. Lurie, and K. Simon (2015). The impact of the affordable care act young adult provision on labor market outcomes: Evidence from tax data. *Tax Policy and Economy 29*(1), 133–157.

Heim, B. T. and I. Z. Lurie (2010). The effect of self-employed health insurance subsidies on self-employment. *Journal of Public Economics 94*(11-12), 995–1007.

Heim, B. T. and I. Z. Lurie (2014). Does health reform affect self-employment? Evidence from Massachusetts. *Small Business Economics 43*(4), 917–930.

Heim, B. T. and L. K. Yang (2017). The impact of the Affordable Care Act on self-employment. *Health Economics 26*(12), e256–e273.

Hincapié, A. (2020). Entrepreneurship over the life cycle: Where are the young entrepreneurs? *International Economic Review 61*(2), 617–681.

Holtz-Eakin, D., J. R. Penrod, and H. S. Rosen (1996). Health insurance and the supply of entrepreneurs. *Journal of Public Economics 62*(1-2), 209–235.

Humphries, J. E. (2018). The causes and consequences of self-employment over the life cycle. Working paper, Yale University.

Hurst, E. and B. W. Pugsley (2011). What do small businesses do? *Brookings Papers on Economic Activity*, 73–143.

Imai, S. and M. P. Keane (2004). Intertemporal labor supply and human capital accumulation. *International Economic Review 45*(2), 601–641.

Jia, Y. G. (2014). Health insurance coverage and self-employment among young US adults. *Available at SSRN 2533648*.

Kaestner, R., B. Garrett, J. Chen, A. Gangopadhyaya, and C. Fleming (2017). Effects of aca medicaid expansions on health insurance coverage and labor supply. *Journal of Policy Analysis and Management 36*(3), 608–642.

Keane, M. P., P. E. Todd, and K. I. Wolpin (2011). The structural estimation of behavioral models: Discrete choice dynamic programming methods and applications. In *Handbook of labor economics*, Volume 4, pp. 331–461. Elsevier.

Keane, M. P. and K. I. Wolpin (1997). The career decisions of young men. *Journal of Political Economy 105*(3), 473–522.

KFF - State Health Facts (2019). Health Insurance Coverage of Adults 19-64. Available at `https://bit.ly/2RBMxCW`. Accessed April 18, 2021.

KFF - The Kaiser Commission on Medicaid and the Uninsured (2015). Key Facts about the Uninsured Population. Available at `https://bit.ly/3eja5EC`. Accessed April 18, 2021.

Kleven, H. J., M. B. Knudsen, C. T. Kreiner, S. Pedersen, and E. Saez (2011). Unwilling or unable to cheat? Evidence from a tax audit experiment in Denmark. *Econometrica 79*(3), 651–692.

Kossoudji, S. A. and D. A. Cobb-Clark (2002). Coming out of the shadows: Learning about legal status and wages from the legalized population. *Journal of Labor Economics 20*(3), 598–628.

Leung, P. and A. Mas (2016). Employment effects of the aca medicaid expansions. Technical report, National Bureau of Economic Research.

Levy, H., T. C. Buchmueller, and S. Nikpay (2018). Health reform and retirement. *The Journals of Gerontology: Series B 73*(4), 713–722.

Li, Y., M. A. Palma, and S. Towne (2017). Does health insurance provision improve self-employment and entrepreneurship? Evidence from state insurance mandates. Technical report, working paper, DOI: 10.22004/ag.econ.258399.

Lofland, J. H. and K. D. Frick (2006). Effect of health insurance on workplace absenteeism in the us workforce. *Journal of occupational and environmental medicine 48*(1), 13–21.

Lombard, K. V. (2001). Female self-employment and demand for flexible, nonstandard work schedules. *Economic inquiry 39*(2), 214–237.

Madrian, B. C. (1994). Employment-based health insurance and job mobility: Is there evidence of job-lock? *The Quarterly Journal of Economics 109*(1), 27–54.

Moriya, A. S., T. M. Selden, and K. I. Simon (2016). Little change seen in part-time employment as a result of the affordable care act. *Health affairs 35*(1), 119–123.

Neovius, K., K. Johansson, M. Kark, and M. Neovius (2009). Obesity status and sick leave: a systematic review. *Obesity reviews 10*(1), 17–27.

Olden, A. and J. Møen (2020). The triple difference estimator. *NHH Dept. of Business and Management Science Discussion Paper* (2020/1).

Parker, S. C. (2018). *The economics of entrepreneurship*. Cambridge University Press, DOI: 10.1017/9781316756706.

Plantenga, T. D. (2009). Hopspack 2.0 user manual. *Sandia National Laboratories Technical Report Sandia National Laboratories Technical Report SAND2009-6265*.

Rees, H. and A. Shah (1986). An empirical analysis of self-employment in the UK. *Journal of Applied Econometrics 1*(1), 95–108.

Rotter, J. B. (1966). Generalized expectancies for internal versus external control of reinforcement. *Psychological Monographs: General and Applied 80*(1), 1.

Schmier, J. K., M. L. Jones, and M. T. Halpern (2006). Cost of obesity in the workplace. *Scandinavian journal of work, environment & health*, 5–11.

Simon, K., A. Soni, and J. Cawley (2017). The impact of health insurance on preventive care and health behaviors: evidence from the first two years of the aca medicaid expansions. *Journal of Policy Analysis and Management 36*(2), 390–417.

Slusky, D. J. (2017). Significant placebo results in difference-in-differences analysis: The case of the acaâs parental mandate. *Eastern Economic Journal 43*(4), 580–603.

Trogdon, J., E. A. Finkelstein, T. Hylands, P. S. Dellea, and S. Kamal-Bahl (2008). Indirect costs of obesity: a review of the current literature. *Obesity Reviews 9*(5), 489–500.

U. S. Bureau of Labor Statistics (2020). Ted: The Economics Daily-Average employee medical premium $6,797 for family coverage in 2020. Available at https://bit.ly/2OXOgS6. Accessed April 18, 2021.

Velamuri, M. (2012). Taxes, health insurance, and women's self-employment. *Contemporary Economic Policy 30*(2), 162–177.

Vistnes, J. P. (1997). Gender differences in days lost from work due to illness. *ILR Review 50*(2), 304–323.

Wellington, A. J. (2001). Health insurance coverage and entrepreneurship. *Contemporary Economic Policy 19*(4), 465–478.

Wen, J.-F. and D. V. Gordon (2014). An empirical model of tax convexity and self-employment. *Review of Economics and Statistics 96*(3), 471–482.

World Health Organisation (2018). Available at https://bit.ly/3mWGxjM. Accessed April 18, 2021.

Xu, X. and G. A. Jensen (2012). Does health insurance reduce illness-related worker absenteeism? *Applied Economics 44*(35), 4591–4603.

Zissimopoulos, J. M. and L. A. Karoly (2007). Transitions to self-employment at older ages: The role of wealth, health, health insurance and other factors. *Labour Economics 14*(2), 269–295.