Essays in Applied Economics

Norihiro Komura

A dissertation

submitted to the Faculty of

the department of Economics

in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

Boston College
Morrissey College of Arts and Sciences
Graduate School

August 2023

Essays in Applied Economics

Norihiro Komura

Advisor: Theodore Papageorgiou, Ph.D.

Chapter 1 examines whether moral hazard in unemployment insurance (UI) varies with the job-finding rate (JFR) and the job-losing rate (JLR). A search model allowing exogenous separation predicts that the JLR and the JFR depending on the tightness of the labor market could affect moral hazard differently. Consistent with the predictions, this chapter finds empirical evidence that moral hazard becomes smaller if JLR is higher while not varying with tightness. Policy makers should focus on JLR, which is a main driver of local unemployment rates, to adjust UI benefits according to labor market conditions. Chapter 2 estimates the impact of patient cost sharing on utilization over time by difference-in-discontinuities with population data. This chapter exploits the difference in coinsurance by cohort, 10\% vs 20\%, during ages 70-74. This chapter finds that the impact after five years on total spending is similar to the immediate impact, while heterogeneity by type of care exists. The reduction of discretionary care gets larger over time and remains persistent after age 75, possibly due to habit formation. In contrast, the impact on less discretionary care remains unchanged over time, diluting the strengthening impact of discretionary care and stabilizing the response dynamics of total spending as observed. Chapter 3 develops a theoretical framework to analyze the impact of social norms on the marriage-market outcome. Social norms have a husband work outside the home and a wife take care of children, irrespective of their comparative advantages. Social norms generate larger inefficiency if a high-skill wife has comparative advantages in working than a low-skill husband, making marriages less beneficial for high-skill women, hindering their match, and inducing them to leave the marriage market when outsourcing childcare is difficult. This effect decreases the aggregated surplus in an economy. At the same time, the decrease in the competitiveness of high-skill women could improve others in the marriage market, high-skill men or low-skill women.

Contents

List of Figures

Chapter 1

Losing or Finding? Why Should UI benefits Vary with Unemployment Rates

1.1 Introduction

Unemployment insurance (UI) provides key insurance for jobless workers. The optimal level of UI benefits depends on the marginal gain and the marginal cost (Baily (1978); Chetty (2006)). On the gain side, higher UI benefits help workers smooth their consumption upon the loss of a job. On the cost side, higher UI benefits could cause longer unemployment duration by discouraging search effort. This is known as moral hazard in UI, and numerous studies find that moral hazard does exist (Krueger and Meyer (2002); Schmieder and von Wachter (2016)).

Recently, literature examines whether moral hazard varies with labor market conditions in terms of the *level* of UI benefits (Sánchez (2008); Andersen and Svarer (2011); Kroft and Notowidigdo (2016)).[1] Kroft and Notowidigdo (2016) examine this issue both theoretically and empirically. In their model, labor market conditions, approximated by unemployment

[1]In terms of the *potential duration* of UI benefits, for example, Schmieder et al. (2012) find that the impact of longer potential duration does not vary with unemployment rates. Schmieder et al. (2012) also examine the *gain* side. Specifically, they find that the exhaustion rates of UI benefits increase substantially if unemployment rates are higher, which suggests that the *gain* of longer potential duration of UI benefits is higher if unemployment rates are higher. Together with their finding on the *cost* side that the impact of longer potential duration of UI benefits on unemployment duration does not vary with unemployment rates, they conclude that the potential duration of UI benefits should be longer if unemployment rates are higher.

rates, affect the size of moral hazard through job-finding rates (JFR). They assume that search effort and labor market conditions are compliment in the sense that search effort is more effective when a labor market condition is better. This assumption leads to the prediction that worse labor market conditions, higher unemployment rates, make smaller moral hazard smaller. Their model also assumes that once an agent finds a job, the job lasts forever; that is, they assume that job-losing rates (JLR) do not play a role in determining the magnitude of moral hazard. In their empirical analysis, they find that higher unemployment rates cause smaller moral hazard by exploiting state-year variation of UI benefits and unemployment rates with the Survey of Income and Program Participation (SIPP) spanning 1985-2000. Together with their finding on the marginal gain,[2] they conclude that the level of UI benefits should vary over the business cycle in a counter-cyclical manner.

I investigate how moral hazard in UI varies with unemployment rates by distinguishing the role of JLR from JFR that are the function of the tightness of the labor market, both theoretically and empirically. The distinction is crucial because literature finds that JLR are the driving force of differences in local unemployment rates (Bilal (2021)), which is also confirmed in my data, while JFR are the dominant force in driving unemployment rates over the business cycle (Shimer (2005); Hall (2005); Fujita and Ramey (2009): Krusell et al. (2017)). Thus, if the finding by Kroft and Notowidigdo (2016) is driven by JLR, it could suggest that UI benefits should vary across states, but not over the business cycle.

In theory part, I develop a simple partial equilibrium dynamic search model with two generalizations from Kroft and Notowidigdo (2016). First, I impose no assumption on the relationship between search effort and the tightness of the labor market; it can be compliment (Pissarides (2000); Sánchez (2008); Kroft and Notowidigdo (2016)) or substitute (Shimer (2004);Mukoyama et al. (2018)). Second, I allow exogenous job destruction after an unemployed agent finds a job to examine the role of JLR.

[2]Specifically, Kroft and Notowidigdo (2016) estimate how the consumption drop upon unemployment varies with unemployment rates. In the simulation, their calibration indicates that the consumption drop upon unemployment is larger if unemployment rates are higher. In reduced-form estimation, they do not find evidence that the consumption drop upon unemployment varies with unemployment rates.

The model indicates that JLR and the tightness of the labor market could affect the size of moral hazard differently. On JLR, the model gives a clear prediction; if JLR are higher, moral hazard becomes smaller. On JFR, the model provides an ambiguous prediction; the effect depends on the relationship between search effort and the tightness of the labor market, which I do not impose any assumption. The effect of the tightness of the labor market is an empirical question.

In the empirical part, I start from the same regression as Kroft and Notowidigdo (2016) by using SIPP spanning 1985-2000, periods used in their study, and 2001-2013. I find that the effect of unemployment rates on moral hazard, which Kroft and Notowidigdo (2016) find in 1985-2000, vanishes for 2001-2013; the sign of the estimated coefficient is flipped and it becomes statistically insignificant. Then, I examine the effect of JLR and the tightness of the labor market. I include in a regression both the interaction of UI benefits and JLR and the interaction of UI benefits and the tightness of the labor market. I find that moral hazard becomes smaller as JLR gets higher as the theory suggests, while not finding evidence that moral hazard varies with the tightness of the labor market. The results are stable for both periods, 1985-2000 and 2001-2013. The reason that the effect of unemployment rates is significant for 1985-2000 but not for 2001-2013 is that a correlation between the unemployment rate and JLR in state and year is large and positive for 1985-2000, 0.41, but substantially small for 2001-2013, 0.06, so that the unemployment rate does not capture the effect of JLR for the latter periods.

The magnitude of the effect of JLR is economically large: in my baseline specification, an increase in JLR by one standard deviation (an increase of 2.2 percentage points from a base, 3.8%) reduces the elasticity of unemployment duration with respect to UI benefits from 0.87 to 0.38 (a decline in the magnitude of roughly 57%). Although the precision of estimates is somewhat limited, the baseline result is fairly robust to potential concerns for identification: unobserved labor market conditions and composition bias.

The rest of the chapter proceeds as follows. Section 2 shows a simple partial equilibrium dynamic search model with exogenous job destruction and examines how JLR and the tightness of the labor market could play a different role in determining moral hazard. Section 3 describes data and empirical strategy and Section 4 shows and discusses the results.

1.2 Theory

1.2.1 Agent's Problem and Search Effort

Setup. Consider a partial equilibrium dynamic search model. An agent lives two periods, $t \in \{1, 2\}$. I assume that the interest rate and the discount rate are zero. Suppose that an agent enters a labor market without a job at period $t = 1$. At the beginning of period $t = 1$, an agent chooses search effort s given the average search effort in her labor market $\bar{s}$ and the tightness of labor market θ. Let $\pi = \pi(s, \bar{s}, \theta)$ is the probability of finding a job where π is concave in s. Also, let $\psi(s)$ denote the cost of search effort where $\psi(s)$ is strictly convex. If the search is successful, an agent starts working immediately in period $t = 1$. At the end of period $t = 1$, a job is subject to exogenous destruction at the rate of λ. If a job survives, an agent continues to work in period $t = 2$. Instead, if a job is destructed, an agent becomes unemployed in period $t = 2$. I assume that an agent does not search a job at period $t = 2$.

An agent earns $w_t = w$ while employed. While unemployed, an agent obtains UI benefits, $b_t = b < w$, that is financed by an actuarially fair tax τ. I assume that an agent cannot save. Let $c^e_t = c^e = w - \tau$ denote the agent's consumption while employed, and $c^u_t = c^u = b$ denote consumption while unemployed. Let $v(c^e) = v(w-\tau)$ and $v(c^u) = v(b)$ be the agent's flow utility as a function of consumption while employed and unemployed, respectively. I assume that v is strictly concave.

Agent's Problem. The value function of an agent who finds a job at the beginning of period $t = 1$ is

$$V_{t=1} = v(c^e) + \Big[(1-\lambda)v(c^e) + \lambda v(c^u)\Big] = (2-\lambda)v(c^e) + \lambda v(c^u). \tag{1.1}$$

Equation (1.1) indicates that the agent gets the flow utility while employed $v(c^e)$ in $t = 1$. In $t = 2$, with the probability of $1 - \lambda$, she keeps her job and get the flow utility while employed $v(c^e)$, and with the probability of λ, she loses her job and get the flow utility while unemployed $v(c^u)$.

The value function for an agent who fails to find a job at the beginning of $t = 1$ and remains unemployed is

$$U_{t=1} = v(c^u) + v(c^u) = 2v(c^u). \tag{1.2}$$

Equation (1.2) indicates that the agent gets the flow utility while unemployed $v(c^u)$ in $t = 1$ and also in $t = 2$.

The value of entering period $t = 1$ without a job is

$$J_{t=1} = \max_{s} \ \pi(s, \bar{s}, \theta) \cdot V_{t=1} + (1 - \pi(s, \bar{s}, \theta)) \cdot U_{t=1} - \psi(s). \tag{1.3}$$

An agent chooses search effort s to maximize her expected utility at the beginning of period $t = 1$, which is given by equation (1.3). Optimal search effort $s^* = s^*(w, \tau, b, \bar{s}, \theta, \lambda)$ is determined by the first-order condition:

$$\psi^{'}(s) = \pi_s(s, \bar{s}, \theta)(2 - \lambda)\big(v(c^e) - v(c^u)\big). \tag{1.4}$$

that equalizes the marginal cost of search effort (left-hand side) and the marginal expected return to search effort (right-hand side). Notice that the marginal expected return to search effort comprise of three elements: (1) the increased probability of finding a job by additional search effort $(\pi_s s, \bar{s}, \theta)$, (2) the probability of keeping a job over two periods $(2 - \lambda)$, and (3) the utility gain from being employed relative to being unemployed $\big(v(c^e) - v(c^u)\big)$. The higher marginal expected return to search effort motivates an agent to search for a job harder.

The first-order condition is more general than previous studies in two ways. First, in contrast to Chetty (2008) who assumes that $\pi = s$, I allow the probability of finding a job π to depend on the tightness of labor market θ. Furthermore, unlike previous studies (Pissarides (2000); Chetty (2008); Sánchez (2008); Kroft and Notowidigdo (2016)), I do not impose a restriction on the sign of $\pi_{s\theta}$ that represents whether search effort and the tightness of labor market are compliment or substitute. Second, in contrast to Chetty (2008) and Kroft and Notowidigdo (2016), I allow exogenous job destruction in my model; job-losing rates λ show up in the first-order condition. I will discuss the implication of these generalizations for the behavioral response to UI benefits in the next subsection.

1.2.2 Behavioral Response to UI benefits

In this section, I show how moral hazard in UI could vary with JLR as well as the tightness of the labor market. To keep my argument simple, I consider a symmetric Nash equilibrium where all workers choose the same search intensity so that $s = \bar{s}$. Then, the first-order condition becomes

$$\psi^{'}(s) = \pi_s(s, \theta)(2 - \lambda)\big(v(c^e) - v(c^u)\big). \tag{1.5}$$

Differentiating optimal search effort $s^* = s^*(w, \tau, b, \theta, \lambda)$ with respect to UI benefits b gives

$$\frac{\partial s^*}{\partial b} = \frac{\pi_s(s, \theta)(2 - \lambda)}{\psi''(s) - \pi_{ss}(s, \theta)(2 - \lambda)\big(v(c^e) - v(c^u)\big)} \cdot \frac{\partial\big(v(c^e) - v(c^u)\big)}{\partial b}. \tag{1.6}$$

This equation (1.6) captures the determinants of moral hazard. Under the assumptions mentioned above, the *sign* of $\partial s^*/\partial b$ is negative, meaning that more generous UI benefits decrease search effort, leading to longer unemployment duration. Notice that the decrease in the utility gain from being employed relative to being unemployed by UI benefits, $\frac{\partial(v(c^e) - v(c^u))}{\partial b} = -v^{'}(c^u)$, is the ultimate source of the behavioral response to UI benefits.

The *magnitude* of $\partial s^*/\partial b$ depends on two probabilities in the numerator of equation

(1.6): the increased probability of finding a job π_s and the probability of keeping a found job over two periods $(2-\lambda)$. In other words, those probabilities amplify or dampen the impact of the decrease in the utility gain from being employed relative to being unemployed by more generous UI benefits, $\frac{\partial(v(c^e)-v(c^u))}{\partial b}$. Specifically, if those probabilities are higher, moral hazard becomes larger. At the same time, these effects by two probabilities are denominated by net adjustment cost of search effort that depends on the second derivative of $\psi(s)$ and $\pi(s,\theta)$.

The equation (1.6) produces the prediction about how moral hazard varies with JLR and the tightness of the labor market. First, by taking derivatives with respect to the tightness of labor market θ, I obtain

$$\frac{\partial^2 s^*}{\partial b \partial \theta} = \frac{\pi_{s\theta}(s,\theta)\cdot\Phi_1 + \pi_{ss\theta}(s,\theta)\cdot\Phi_2}{\left[\psi''(s) - \pi_{ss}(s,\theta)(2-\lambda)\big(v(c^e)-v(c^u)\big)\right]^2} \cdot \frac{\partial\big(v(c^e)-v(c^u)\big)}{\partial b}. \tag{1.7}$$

where both $\Phi_1(= (2-\lambda)[\psi''(s) - \pi_{ss}(s,\theta)(2-\lambda)\big(v(c^e)-v(c^u)\big)])$ and $\Phi_2(= (2-\lambda)^2\pi_s(s,\theta)(v(c^e)-v(c^u)))$ are strictly positive.

The theory proposed by the numerical simulation in Kroft and Notowidigdo (2016) can be described as a special case of equation (1.7). In particular, they assume that $\pi(s,\theta) = s\theta$ and the equation boils down to

$$\frac{\partial^2 s^*}{\partial b \partial \theta} = \frac{1}{\psi''(s)} \cdot \frac{\partial\big(v(c^e)-v(c^u)\big)}{\partial b} < 0. \tag{1.8}$$

Thus, their theory predicts that a worse labor market, lower θ, makes moral hazard smaller. If the increased probability of finding a job by additional search effort is lower by a worse labor market, the marginal expected return to search effort is lower. The return when an agent gets a job relative to staying unemployed is not very relevant for her because she is less likely to get the job. She therefore does not care about the loss of the return by higher UI benefits.

However, in general, the effect of θ is unclear for two reasons. First, the relationship between the tightness of the labor market and search effort could be compliment or substitute; namely, the sign of $\pi_{s\theta}$ could be positive or negative. To date, there seems no broad consensus whether they are compliment or substitute. On the one hand, many studies assume that they are complement (for example, Pissarides (2000); Sánchez (2008); Kroft and Notowidigdo (2016)). On the other hand, recent findings by Shimer (2004) and Mukoyama et al. (2018) find that search effort is counter-cyclical; Mukoyama et al. (2018) show that the counter-cyclicality of search effort requires the tightness of labor market and search effort are substitute. Second, $\pi_{s\theta}$ and $\pi_{ss\theta}$ tend to take different signs. For example, Mukoyama et al. (2018) show a generalized CES matching function that nests some important cases, including the standard DMP matching function in Cobb-Douglas form, the Cobb-Douglas special case of Pissarides (2000); and the case to describe their finding of the counter-cyclical search effort. In this function, $\pi_{s\theta}$ and $\pi_{ss\theta}$ take opposite signs. Therefore, the effect of the tightness of the labor market on the behavioral response to UI benefits is an empirical question.

Second, by taking derivatives with respect to JLR λ, I obtain

$$\frac{\partial^2 s^*}{\partial b \partial \lambda} = \frac{-\pi_s(s,\theta) \cdot \psi''(s)}{\left[\psi''(s) - \pi_{ss}(s,\theta)(2-\lambda)\big(v(c^e) - v(c^u)\big)\right]^2} \cdot \frac{\partial\big(v(c^e) - v(c^u)\big)}{\partial b} > 0. \tag{1.9}$$

Thus, in contrast to the tightness of the labor market, the effect of JLR is clear qualitatively; higher JLR make moral hazard smaller. If the probability of keeping a found job is low, the marginal expected return to search effort is lower. The return when an agent gets a job relative to staying unemployed is not very relevant for her because she is more likely to lose the job. She therefore does not care the loss of the return by higher UI benefits. Like Kroft and Notowidigdo (2016), a lower marginal expected return to search effort is a key, but it is driven by a higher probability of losing a found job, not by a lower probability of finding a job.

Overall, the theory suggests that JLR and the tightness of the labor market could play a different role in determining moral hazard. Notice that Kroft and Notowidigdo (2016) use unemployment rates to approximate the tightness of labor market θ without using a variable corresponding to job-losing rates λ. Specifically, by using time and state variations, they estimate a Cox proportional hazard model where its dependent variable is the log of hazard rates of exiting from unemployment and its main independent variable is the interaction between the log of average UI benefits and the log of unemployment rates. However, since unemployment rates are related to λ as well as θ, the coefficient of the interaction term would capture the mixed effect of λ, through the probability of keeping a job, and θ, through the increased probability of finding a job. Since those effects could be different as I explained above, I will include both λ and θ in my regressions in the following empirical section.

1.3 Empirical Analysis

1.3.1 Data

UI Benefits, Job-losing Rates, Tightness of Labor Market

Regarding the level of UI benefits, I use (1) the average UI benefit amount for each state-year pair and (2) the statutory maximum UI benefit amount for each state-year pair (Chetty (2008); Krueger and Mueller (2010); Kroft and Notowidigdo (2016)). Both data are obtained from the Department of Labor. For calculating job-losing rates, I follow the method proposed by Shimer (2005) and Shimer (2012) by using the Current Population Survey (CPS). Regarding the tightness of labor market, I use the vacancy-unemployment ratio. Data on the state-level unemployment is obtained from Bureau of Labour Statistics. Regarding data on the state-level vacancy, I use two sets of data depending on periods. Specifically, for 1985-2000, I use the Conference Board's Help-Wanted Index (HWI). HWI is the monthly index of job advertisements in printed newspapers for major metropolitan areas.

I impute the average of WHI for each state, weighted by population in each belonging metropolitan area. Since some states are not covered by the HWI, I impute the HWI for those locations by following the method of Kroft and Notowidigdo (2016). In particular, I regress the log of vacancy index on year-month fixed effects, state fixed effects, and the log of state-level unemployment rates, and assign the predicted vacancy index to those non-covered states. For 2001-2013, I use the state-level estimates of job openings from the Job Openings and Labor Turnover Survey (JOLTS) that are available after February 2001.[3] JOLTS covers all 50 states and the District of Columbia. Due to the lack of consistent data on the state-level vacancy, I implement empirical analysis separately for 1985-2000 and 2001-2013.

Unemployment Spells

On unemployment spells, I use the eleven SIPP panels spanning 1985-2013.[4] Relative to other data sets, such as CPS and the Panel of Study on Income Dynamics, the main advantage of the SIPP is the availability of data on UI benefits receipt, asset data, weekly data on employment status, and large sample size. As I mentioned above, I split the SIPP into two periods, 1985-2000 and 2001-2013. I focus on prime-age males who (a) report searching for a job, (2) are not on temporary layoff, (c) have at least 3 months of work history, and (d) claimed UI benefits. I also censor unemployment spells at 50 weeks. I drop observations from relatively small states.[5] These restrictions leave 4,307 unemployment spells for 1985-2000 and 2,886 unemployment spells for 2001-2013. On the measurement of unemployment spells, I follow Chetty (2008) and originally Cullen and Gruber (2000); see Appendix for details.

[3] I obtain the experimental estimates on August 15, 2020 from the Bureau of Labor Statistics (`https://www.bls.gov/jlt/jlt_statedata.htm`).

[4] I do not use the 1989 panel and the 1988 panel, like Cullen and Gruber (2000), Chetty (2008), and Kroft and Notowidigdo (2016), since they have only three waves.

[5] In particular, I drop observations from Maine, Vermont, Iowa, North Dakota, South Dakota, Alaska, Idaho, Montana, and Wyoming because the SIPP does not provide unique identifiers for individuals in these states (Chetty (2008)).

1.3.2 Regression Equation

I start my empirical analysis from the same regression as Kroft and Notowidigdo (2016) but for the different periods, 2001-2013, as well as for the same periods, 1985-2000. The regression equation is

$$\log h_{ist} = \gamma_1 \log(b_{st}) + \gamma_2 \log(b_{st}) \log(u_{st}) \\ + \gamma_3 \log(u_{st}) + \delta_t + \delta_s + \mathbf{X}_{ist}\Gamma + e_{ist} \tag{1.10}$$

where h_{ist} is the hazard rate of exiting from unemployment for individual i in state s at time t, b_{st} is the level of UI benefits in state s at time t, u_{st} is the unemployment rates in state s at time t. Like Kroft and Notowidigdo (2016), I assign the monthly state unemployment rate based on the month and state that individual i starts her unemployment spell. δ_t and δ_s are time and state fixed effects, and $\mathbf{X}_{ist}$ is a set of control variables.

Here, γ_1 represents the elasticity of the conditional probability of exiting from unemployment with respect to UI benefits, so that $-\gamma_1$ captures the elasticity of unemployment duration with respect to UI benefits. Thus, negative γ_1 means that higher UI benefits lead to longer unemployment duration, capturing moral hazard in UI. Then, $-\gamma_2$ represents the elasticity of unemployment duration with respect to the interaction of UI benefits and the unemployment rate. Kroft and Notowidigdo (2016) find that the estimate of γ_2 is positive and statistically significant for 1985-2000, meaning that higher unemployment rates lead to smaller moral hazard.

Then, to examine how the duration elasticity varies with JLR and the tightness of labor market, I replace the unemployment rate in equation (1.10) with JLR and the tightness of labor market. The baseline regression equation is the following:

$$\log h_{ist} = \beta_1 \log(b_{st}) + \beta_2 \log(b_{st}) \log(\lambda_{st}) + \beta_3 \log(b_{st}) \log(\theta_{st}) \\ + \beta_4 \log(\lambda_{st}) + \beta_5 \log(\theta_{st}) + \delta_t + \delta_s + \mathbf{X}_{ist}\Gamma + \varepsilon_{ist} \tag{1.11}$$

where λ_{st} is JLR in state s at time t, and θ_{st} is the tightness of labor market in state s at time t. Note that although λ and θ are negatively correlated, their relationship is not strong. In fact, their correlation is -0.10 for 2001-2013 when JOLTS is available. Collinearity does not therefore make estimates very unstable.

In the regression (1.11) (and regression (1.10)), I demean all variables. Thus, $-\beta_1$ represents the elasticity of unemployment duration with respect to the level of UI benefits at the average job-losing rates and the tightness of the labor market. The coefficient of interests are β_2 and β_3. In particular, $-\beta_2$ and $-\beta_3$ capture how the elasticity of unemployment duration varies with job-losing rates and with the tightness of the labor market, respectively. Recall that the theory predicts that β_2 should be positive, meaning that higher job-losing rates lead to a smaller behavioral response to UI benefits. In contrast, the sign of β_3 is ambiguous, depending on the relationship between search effort and the tightness of the labor market.

The regression equation (1.11) (and equation (1.10)) is a simplified representation in the sense that for estimation, the equation includes a (non-parametric) baseline hazard rate and separates baseline hazard rates for each quartile of net liquid wealth. I control for the third-order polynomial of age, marital status, years of education, the number of children, state fixed effects, year fixed effects, industry fixed effects, occupation fixed effects, a 10-knot linear spline in the log of the annual (pre-unemployment) wage income, and an indicator for being on the seam between interviews, and year fixed effects interacted with the log of the UI benefit amount. Standard errors are clustered by state.

As I mentioned in data section, for the level of UI benefits b_{st} I use (1) the average benefit amount for each state-year pair and (2) the statutory maximum benefit amount for each state-year pair. The statutory maximum benefit amount is the primary source of the average benefit amount. These measures have been exploited as exogenous policy variation in literature (Chetty (2008); Krueger and Mueller (2010); Kroft and Notowidigdo (2016)). In fact, Hsu et al. (2014) and Kroft and Notowidigdo (2016) show that the correlation between the level of UI benefits and various macroeconomic variables, including the unemployment rates, is relatively weak conditional on year fixed effects.

Still, there are two concerns about the identifying assumption: unobserved labor market conditions and composition bias. Regarding unobserved labor market conditions, while it is difficult to deal with the issue perfectly, the inclusion of both the log of job-losing rates and the log of the tightness of the labor market would help to mitigate the concern. In addition, as robustness checks, I control for the unemployment rates flexibly by following Kroft and Notowidigdo (2016). Composition bias is also a source of concern. In fact, several papers show that the composition of unemployed workers varies over the business cycle (Elsby et al. (2015); Mueller (2017)). Again, while it is difficult to deal with the issue perfectly, as robustness checks, I include interaction terms with UI benefits and individual characteristics, such as years of education and pre-unemployment wage, by following Kroft and Notowidigdo (2016). If composition bias is an important source of bias, these interaction variables are expected to significantly change the magnitude of the estimated coefficient of the interaction between UI benefits and job-losing rates as well as the tightness of the labor market.

1.4 Results

1.4.1 Main Results

Results are reported in Table 1.1. Part (A) uses the 1985-2000 SIPP and part (B) uses the 2001-2013 SIPP. Column KN in part (A) shows the result for regressing equation (1.10) for 1985-2000 so that this is a replication of Kroft and Notowidigdo (2016). The result is of course very similar to theirs. Specifically, the estimated coefficient for the interaction between UI benefits and unemployment rates (γ_2) is positive and statistically significant, meaning that behavioral responses to UI benefits are smaller if unemployment rates are higher. Column KN in part (B) reports the result of the same regression for 2001-2013. The result indicates that the coefficient of the interaction between UI benefits and unemployment rates (γ_2) becomes statistically insignificant, and moreover, its sign is flipped from positive to negative. The effect of the unemployment rate on behavioral responses vanishes for the latter period, 2001-2013.

Column Average and Max report the results for regressing equation (1.11), where I use the log of the average benefit amount paid in a given state-year and log of the statutory benefit amount paid in a given state-year, respectively, and interacts them with the log of job-losing rates and the log of the tightness of labor market.

In panel (A), in terms of job-losing rates (λ), the results show that the estimated coefficient of the interaction term (β_2) is positive and statistically significant at the 5% level (2.157 (SE 0.910) for column Average, for example). This result indicates that, as the theory predicts, the duration elasticity with respect to the level of UI benefits becomes smaller if the job-losing rate is higher, suggesting that higher job-losing rates lead to a smaller behavioral response to UI benefits. Furthermore, the magnitude of the effect of job-losing rates is economically large: in my baseline specification, an increase in job-losing rates by one standard deviation (an increase of 2.2 percentage points from a base, 3.8%) reduces the elasticity from 0.87 to 0.38 (a decline in the magnitude of roughly 57%). In contrast, in terms of the tightness of the labor market (θ), the estimated coefficient of the interaction term (β_3) is not statistically significant (-0.063 (SE 0.358) for column Average, for example). Thus, I do not find evidence that the behavioral response to UI benefits varies with the tightness of the labor market. In part (B), I repeat the same analysis by using the 2001-2013 SIPP. The results show a consistent pattern with Part (A). Specifically, the duration elasticity of job-finding with respect to the level of UI benefits gets smaller with job-losing rates (2.628 (SE 1.303) for column Average), while it does not vary with the tightness of the labor market (0.321 (SE 0.668) for column Average).

These results highlight an important difference between Kroft and Notowidigdo (2016) and this paper. Specifically, while they find that behavioral responses become smaller if the unemployment rate is higher for 1985-2000, I do not find such effects for 2001-2013. Furthermore, I find evidence that in both periods, 1985-2000 and 2001-2013, the behavioral response to UI benefits varies with job-losing rates but no evidence that it varies with the tightness of labor market. These results indicate that the result by Kroft and Notowidigdo (2016), which focuses on the former periods, 1985-2000, is driven by the effect of job-losing rates, while the unemployment rate does not capture the effect of job-losing rates for the latter periods, 2001-2013. In fact, the correlation between unemployment rates and job-losing rates in the former period is relatively high (0.41), while in contrast, the correlation becomes significantly weaker (0.06) in the latter period.[6] Accordingly, my findings indicate that "sufficient statistics" for the optimal level of UI benefits are the job-losing rate, instead of the unemployment rate itself or the tightness of labor market. Namely, policymakers should decide the level of UI benefits by looking at job-losing rates.

Note that my theoretical model assumes that JLR is exogenous. It is possible that JLR is endogenous and driven by UI benefits and other exogenous factors, like local productivity. As I show in Appendix, endogenizing JLR generates two effects: (1) the similar effect to my original model and (2) an additional effect depending on the cross derivative of JLR with respect to UI benefits and an exogenous factor whose sign and magnitude is ambiguous in literature. In regression, I should use those exogenous factors directly, instead of JLR. Meanwhile, given that literature does not reach a consensus about the driver of JLR, my current result should be informative because the interaction term between UI benefits and JLR in regression captures the effect of those exogenous factors determining JLR.

[6] It is beyond the scope of this paper to fully understand why the effect of unemployment rates becomes statistically insignificant and the sign of the associated coefficient is flipped in the latter period. However, I suspect that two factors contribute to the result: (1) relatively weak correlation between unemployment rates and job-losing rates in the latter period (0.06) as Shimer (2012) indicates and (2) the relationship between search effort and the tightness of the the labor market becomes more substitute as Shimer (2004) and Mukoyama et al. (2018) illustrate.

1.4.2 Robustness Check

In this section, I show the results of robustness tests by following Kroft and Notowidigdo (2016), focusing on unobserved labor market conditions and composition bias. Overall, the results show that the baseline results are robust to the tests.

Regarding unobserved labor market conditions, the results are reported in Table 1.2. I first control for the unemployment rate using flexible functional forms. In particular, in both parts (A) and (B), column (1) includes the log of the unemployment rate and column (2) includes the log of the unemployment rate and the square of the log of the unemployment rate. In column (3), I include state-specific linear trends. The results show that the baseline results are fairly robust. It should be noted, however, that in column (3) in part (B) the estimated coefficient of the interaction term with job-losing rates becomes statistically insignificant (p-value 0.121) while remaining positive (2.176 (SD 1.619)). The result might be due to the smaller sample size in the part (B) compared to part (A).

Regarding composition bias, the results by using the 2001-2013 SIPP are reported in Table 1.3. Specifically, I add interactions between UI benefits and the following individual characteristics: age in column (1), marital status in column (2), years of education in column (3), pre-unemployment wage in column (4), occupation fixed effects in column (5), and industry fixed effects in column (6). If the baseline results are primarily due to composition bias, one would expect that the inclusion of these interaction terms with individual characteristics changes the magnitude of the estimated coefficients of the interaction terms with job-losing rates and the tightness of the labor market. The results show that the baseline results are robust to including these interaction terms with individual characteristics. As Kroft and Notowidigdo (2016) mentioned, composition bias may be limited through the sample selection in this study as I limit the sample to prime-age men.

1.4.3 Discussion

Job-losing Rates versus Job-finding Rates

Literature finds that JFR are the dominant force in unemployment rates over the business cycle (Shimer (2005); Hall (2005); Fujita and Ramey (2009): Krusell et al. (2017)). In contrast, a recent paper by Bilal (2021) finds that JLR are a driving force of differences in local unemployment rates, which I do the same excise with his paper and confirm his finding in the United States by CPS (see Appendix like Kuhn et al. (2021). In this light, my finding suggests that behavioral cost of UI, moral hazard, should vary with local unemployment rates that are mainly driven by JLR, but not varying over the business cycle.

In fact, the graphical evidence of Kaplan-Meier survival curves depicted by Kroft and Notowidigdo (2016) essentially captures the difference in local unemployment rates. In particular, they split the sample into two sub-samples based on whether individuals began their unemployment spell in states with above- or below-median unemployment rates across states in a given year. Then, for each sub-sample, they further divide the sub-sample depending on whether the level of UI benefits at the start of unemployment spells is above or below the median level of UI benefits.

By this procedure, Kroft and Notowidigdo (2016) depict Figure 1.1 and 1.2 . In Figure 1.1, the survival curves are similar in both low-benefit and high-benefit groups where unemployment rates are high. Log-rank test for equality is not rejected at conventional levels ($p = 0.599$). In Figure 1.2, the curves diverge between low-benefit and high-benefit groups where unemployment rates are low, suggesting that lower unemployment rates lead to larger behavioral responses. Log-rank test for equality is rejected ($p = 0.004$). These are their findings.

Instead, if the effect of unemployment rates on moral hazard that Kroft and Notowidigdo (2016) finds reflects the effect of JLR, not the tightness of the labor market, one should find no effects of unemployment rates that capture variation over the business cycle. I test this possibility by changing how to split the sample. Specifically, I split the sample into two sub-samples based on whether individuals began their unemployment spell in the *year* with above or below median unemployment rates defined across *year*s in a given *state*.

The results are shown in Figure 1.3 and 1.4. The figures show that in both cases, high and low unemployment rates, the survival curves are similar in both low-benefit and high-benefit groups. Log-rank test for equality is not rejected ($p = 0.241$ for higher unemployment rates and $p = 0.476$ for lower unemployment rates).

Therefore, my results indicate that UI benefits would not need to vary over the business cycle. Rather, the results suggest that UI benefits should take different levels across locations depending on the difference in local unemployment rates that is mainly driven by JLR.

Time-Location Variation of Unemployment Rates

Beyond arguments on UI benefits, my study motivated the Bilal (2021)'s novel finding has an interesting implication for analyses using time-location variation in unemployment rates. In particular, many studies use time-location variation in unemployment rates "to exploit differences in the relative severity of recession" (for example, Aguiar et al. (2013); Haltiwanger et al. (2018); Mukoyama et al. (2018)). However, my study suggests that reserchers may need caution in interpreting results from analysis using time-location variation in unemployment rates. Specifically, if results are driven by JLR, rather than JFR, they might not capture "differences in the relative severity of recession" as intended, given the relatively limited role of JLR in determining aggregate fluctuations of unemployment rates along the business cycle.

1.5 Table

Table 1.1: The Elasticity of Job-finding Probability

	Part (A): 1985-2000			Part (B): 2001-2013		
	KN	Average	Max	KN	Average	Max
$\log(b)$	-0.668^{**}	-0.898^{**}	-0.440	-0.182	-0.412	-0.220
	(0.335)	(0.368)	(0.285)	(0.485)	(0.469)	(0.333)
$\log(b)\log(u)$	1.275^{**}			-0.770		
	(0.499)			(1.319)		
$\log(b)\log(\lambda)$		2.157^{**}	1.327^{**}		2.628^{**}	1.814^{**}
		(0.910)	(0.642)		(1.303)	(0.899)
$\log(b)\log(\theta)$		-0.063	-0.137		0.586	0.008
		(0.358)	(0.301)		(0.676)	(0.533)
$\log(u)$	0.063			-0.190		
	(0.119)			(0.248)		
$\log(\lambda)$		0.250	0.284		-0.365	-0.407
		(0.211)	(0.211)		(0.259)	(0.275)
$\log(\theta)$		0.016	0.040		0.054	0.050
		(0.086)	(0.086)		(0.144)	(0.145)
N of spells	4307	4307	4307	2886	2886	2886

Notes: All columns report the results from estimating the Cox proportional hazard model (**??**). Data are 1985–2000 SIPP for part (A) and 2001-2013 SIPP for part (B). In column KN and Average, the average benefit amount paid in a given state-year is used for b. In column Max, the statutory benefit amount paid in a given state-year is used for b. u is the unemployment rate, λ is job-losing rate, and θ is the tightness of labor market. All specifications include the third-order polynomial of age, marital status, years of education, the number of children, state fixed effects, year fixed effects, industry fixed effects, occupation fixed effects, a 10-knot linear spline in the log of the annual (pre-unemployment) wage income, and an indicator for being on the seam between interviews. All specifications also include year fixed effects interacted with the log of the UI benefit amount. All columns estimate non-parametric baseline hazards stratified by quartile of net liquid wealth. Standard errors are clustered by state and shown in parentheses.

$^{*}\, p < 0.10$, $^{**}\, p < 0.05$, $^{***}\, p < 0.01$

Table 1.2: Robustness Checks for Unobserved Labor Market Conditions

	Part (A): 1985-2000			Part (B): 2001-2013		
	(1)	(2)	(3)	(1)	(2)	(3)
$\log(b)$	-0.918^{**}	-0.874^{**}	-1.126^{**}	-0.413	-0.419	-0.355
	(0.386)	(0.439)	(0.563)	(0.465)	(0.483)	(0.745)
$\log(b)\log(\lambda)$	2.219^{**}	2.213^{**}	2.405^{**}	2.634^{**}	2.653^{**}	2.176
	(0.911)	(0.915)	(1.124)	(1.333)	(1.334)	(1.619)
$\log(b)\log(\theta)$	-0.039	-0.013	0.138	0.588	0.596	-0.750
	(0.360)	(0.376)	(0.464)	(0.668)	(0.80)	(0.693)
$\log(\lambda)$	0.196	0.199	0.396	-0.360	-0.356	-0.186
	(0.232)	(0.232)	(0.266)	(0.302)	(0.316)	(0.329)
$\log(\theta)$	0.106	0.104	-0.044	0.044	0.046	-0.010
	(0.192)	(0.194)	(0.115)	(0.248)	(0.248)	(0.141)
Linear in $\log(u)$	✓			✓		
Quadratic in $\log(u)$		✓			✓	
State-specific trends			✓			✓
N of spells	4307	4307	4307	2886	2886	2886

Notes: All columns report the results from estimating the Cox proportional hazard model (**??**). Data are 1985–2000 SIPP for part (A) and 2001-2013 SIPP for part (B). The average benefit amount paid in a given state-year is used for b. u is the unemployment rate, λ is job-losing rate, and θ is the tightness of labor market. All specifications include the third-order polynomial of age, marital status, years of education, the number of children, state fixed effects, year fixed effects, industry fixed effects, occupation fixed effects, a 10-knot linear spline in the log of the annual (pre-unemployment) wage income, and an indicator for being on the seam between interviews. All specifications also include year fixed effects interacted with the log of the UI benefit amount. All columns estimate non-parametric baseline hazards stratified by quartile of net liquid wealth. Standard errors are clustered by state and shown in parentheses.

$^{*}\,p < 0.10$, $^{**}\,p < 0.05$, $^{***}\,p < 0.01$

Table 1.3: Robustness Checks for Composition Bias

	2001-2013					
	(1)	(2)	(3)	(4)	(5)	(6)
$\log(b)$	-0.412	-0.417	-0.404	-0.373	-0.399	-0.381
	(0.469)	(0.461)	(0.475)	(0.480)	(0.481)	(0.466)
$\log(b)\log(\lambda)$	2.624^{**}	2.618^{**}	2.559^{**}	2.592^{**}	2.630^{**}	2.588^{**}
	(1.305)	(1.313)	(1.320)	(1.302)	(1.319)	(1.278)
$\log(b)\log(\theta)$	0.586	0.587	0.662	0.707	0.568	0.437
	(0.677)	(0.674)	(0.682)	(0.696)	(0.713)	(0.677)
$\log(b)$ Age	0.002					
	(0.011)					
$\log(b)$ Married		0.043				
		(0.243)				
$\log(b)$ Education			-0.056			
			(0.073)			
$\log(b)\log(\text{wage})$				-0.402^{*}		
				(0.206)		
$\log(b)$ Occupation FE					✓	
$\log(b)$ Industry FE						✓
N of spells	2886	2886	2886	2886	2886	2886

Notes: All columns report the results from estimating the Cox proportional hazard model (**??**). Data are 2001-2013 SIPP for part (B). The average benefit amount paid in a given state-year is used for b. λ is job-losing rate, and θ is the tightness of labor market. All specifications include the third-order polynomial of age, marital status, years of education, the number of children, state fixed effects, year fixed effects, industry fixed effects, occupation fixed effects, a 10-knot linear spline in the log of the annual (pre-unemployment) wage income, and an indicator for being on the seam between interviews. All specifications also include year fixed effects interacted with the log of the UI benefit amount. All columns estimate non-parametric baseline hazards stratified by quartile of net liquid wealth. Standard errors are clustered by state and shown in parentheses.

$^{*}\ p < 0.10$, $^{**}\ p < 0.05$, $^{***}\ p < 0.01$

1.6 Figure

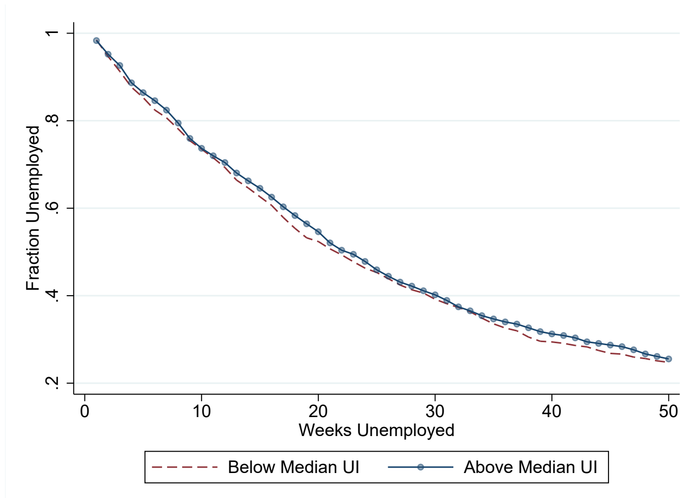


Figure 1.1: High Unemployment Rates (State Variation)

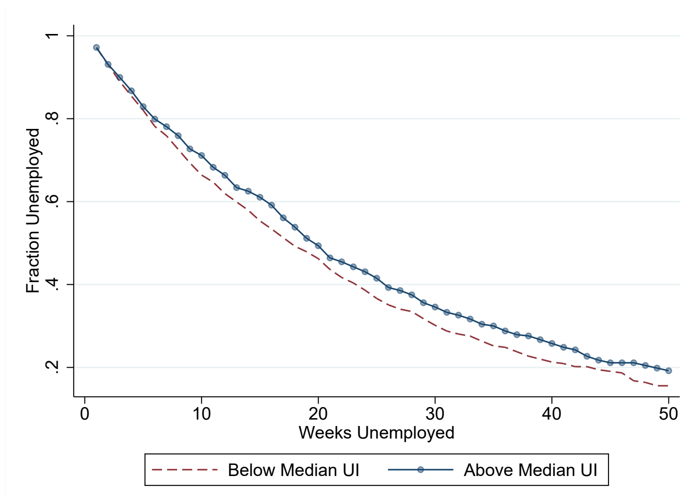


Figure 1.2: Low Unemployment Rates (State Variation)

Notes: I use the 1985-2000 SIPP. In Figure 1.1, the sample includes unemployment spells in states with unemployment rates above the median across states in a given year. Log-rank test indicate that $p = 0.599$. In Figure 1.2, the sample includes unemployment spells in states with unemployment rates below the median across states in a given year. Log-rank test indicate that $p = 0.004$. Each figure plots Kaplan-Meier survival curves for two groups of individuals based on whether UI benefits in an individual's state is above (solid navy) or below (dash maroon) the overall sample median.

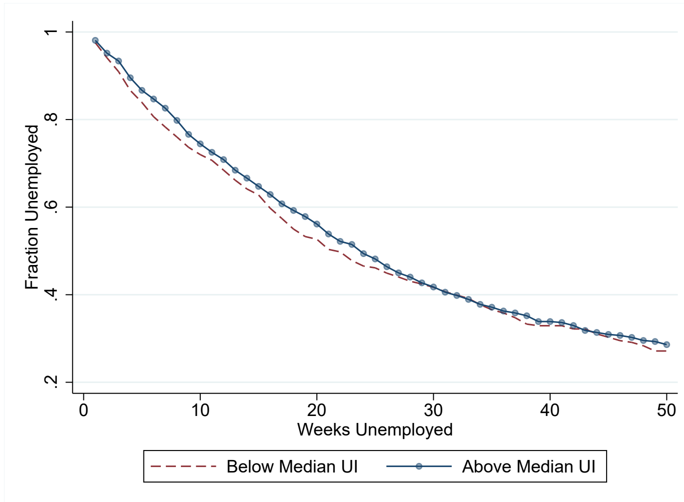


Figure 1.3: High Unemployment Rates (Time Variation)

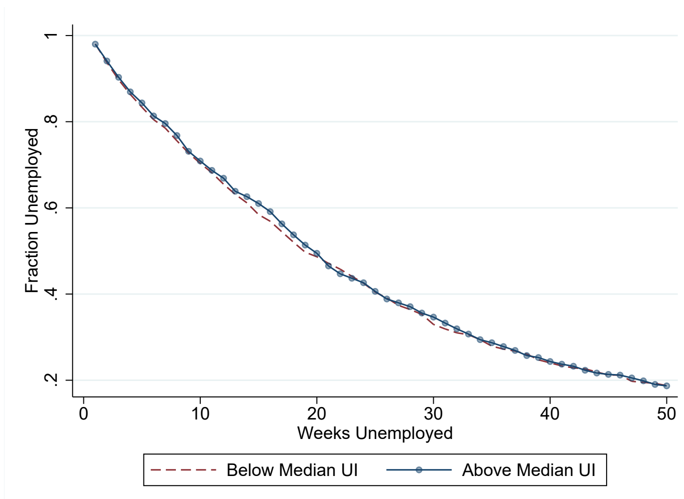


Figure 1.4: Low Unemployment Rates (Time Variation)

Notes: I use the 1985-2000 SIPP. In Figure 1.3, the sample includes unemployment spells in years with unemployment rates above the median across years in a given state. Log-rank test indicate that $p = 0.241$. In Figure 1.4, the sample includes unemployment spells in years with unemployment rates below the median across years in a given state. Log-rank test indicate that $p = 0.476$. Each figure plots Kaplan-Meier survival curves for two groups of individuals based on whether UI benefits in an individual's year is above (solid navy) or below (dash maroon) the overall sample median.

1.7 Appendix

Data

On the measurement of unemployment spells, I follow Chetty (2008) and originally Cullen and Gruber (2000). First, in SIPP, weekly employment status can take the following values:

ES1. with a job this week - working

ES2. with a job, not on layoff, but absent without pay

ES3. with a job, but on layoff

ES4. no job - looking for work

ES5. no job - not looking for work

I compute the unemployment spells by summing the number of consecutive weeks that the value of the weekly employment status is ES3, ES4, or ES5, starting at the date of changing the status from ES1 or ES2 to ES3, ES4, or ES5 and stopping at the date of reporting ES1 or ES2 in consecutive four weeks. Individuals are defined as being on searching a job if they report ES4 at any point of their unemployment spells. Individuals are defined as being on temporary layoff if they report ES3 at any point of their unemployment spells.

Job-losing Rates and Job-finding Rates (Bilal, 2021)

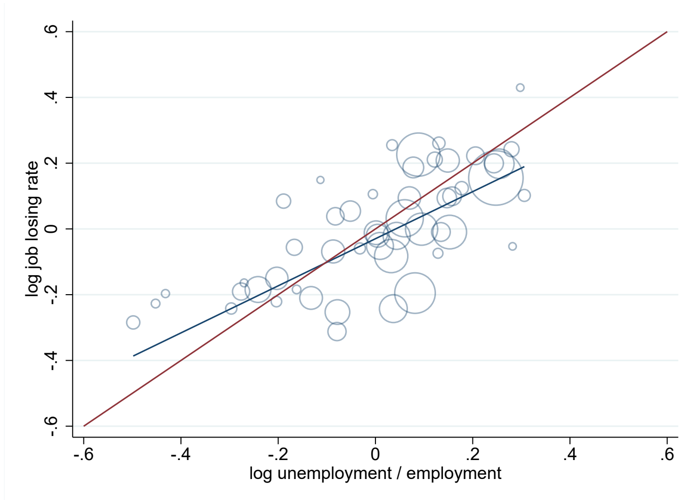


Figure 1.5: Job-losing Rates (Bilal, 2021)

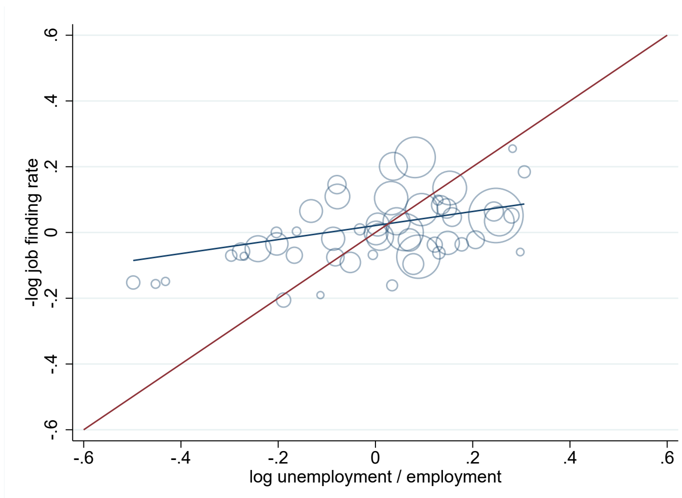


Figure 1.6: Job-finding Rates (Bilal, 2021)

Notes: I use the 1980-2013 CPS in this figure. Figure A1 is for Job-losing rates. Figure A2 is for Job-finding rates. The size of circles represents the relative size of state population.

Endogenous Job-losing Rates

What occurs in this model if $\lambda(b,\xi)$ where $\partial\lambda(b,\xi)/\partial b > 0$, and ξ is the exogenous component other than UI benefits, such as (the minus of) local productivity, and $\partial\lambda(b,\xi)/\partial\xi > 0$? Take two derivatives:

$$\frac{\partial s}{\partial b} = \Xi^{-1}[-\pi_s(s,\theta)(2-\lambda(b,\xi))\frac{\partial v(c^u(b))}{\partial b} - \pi_s(s,\theta)\frac{\partial\lambda(b,\xi)}{\partial b}\{v(c^e) - v(c^u(b))\}] < 0$$

where $\Xi = \psi''(s) - \pi_{ss}(s,\theta)(2-\lambda)\{v(c^e) - v(c^u(b))\}$ is assumed to be positive.

$$\frac{\partial^2 s}{\partial b\partial\xi} = \Xi^{-1}\left[\pi_s(s,\theta)\frac{\partial\lambda(b,\xi)}{\partial\xi}\frac{\partial v(c^u(b))}{\partial b} - \pi_s(s,\theta)\frac{\partial^2\lambda(b,\xi)}{\partial b\partial\xi}\{v(c^e) - v(c^u(b))\}\right]$$

In both derivatives, the second term (in red) shows up when JLR depends on UI benefits. The first derivative, $\partial s/\partial b$, shows that since the second term is negative, moral hazard becomes larger compared to the case that JLR is exogenous. Higher UI benefits raise JLR and the increased JLR lower the return to search effort. As to the second derivative, $\partial^2 s/\partial b\partial\xi$, the first term is positive, like the case of exogenous JLR, but the new second term depends on $\partial^2\lambda(b,\xi)/\partial b\partial\xi$. The prediction is modified; although the overall effect is ambiguous, $\partial^2 s/\partial b\partial\xi > 0$ holds, as long as $\partial^2\lambda(b,\xi)/\partial b\partial\xi$ is negative or positive but sufficiently small.

Chapter 2

The Dynamics of Consumer Response to Medical Price

Coauthored with: Shun-ichiro Bessho (Waseda University).

2.1 Introduction

Consumer responsiveness to medical care prices is a key ingredient for designing optimal health insurance.[1] Numerous studies have explored the responsiveness. Although there is a view that the demand of healthcare is completely inelastic, existing evidence compellingly shows that an increase in prices reduces healthcare spending immediately.[2]

It might not be the end of the story, however. Immediate responses to price changes, short-run responses, might differ from subsequent response dynamics, longer-run responses. If they are different, the estimate of immediate responses results in a wrong ingredient.

Theories say that they could be different and that the impact of price changes could get either smaller or larger over time. Consider an increase in cost sharing. The decreased uti-

[1]The impact on financial strain is also a key. For example, see Finkelstein and McKnight (2008); Engelhardt and Gruber (2011); Gross and Notowidigdo (2011); Finkelstein et al. (2012); Baicker et al. (2013); Shigeoka (2014); Barcellos and Jacobson (2015); Mazumder and Miller (2016); Hu et al. (2018).

[2]For a recent review, see Baicker and Goldman (2011), McGuire (2011), and Finkelstein et al. (2018).

lization by higher cost sharing could deteriorate health, boosting healthcare demand in the long run (Chandra et al., 2010). This possibility is often emphasized by healthcare professionals and researchers. *Behavioral hazard* (Baicker et al. (2015); Chandra et al. (2019)), meaning that in facing higher prices consumers reduce essential care by mistake, could exacerbate the damage on health. Several papers find evidence consistent with behavioral hazard, such as adverse effects on health and similar elasticities between high-value and low-value care.[3]

In contrast, the reduction of utilization could be larger in the long run, to which empirical literature pays little attention. One possibility is *ex-ante moral hazard* (Ehrlich and Becker, 1972). Higher cost sharing encourages health-improving behavior, resulting in subsequent reductions of utilization. Another possibility is *habit formation* (Stigler and Becker (1977); Becker and Murphy (1988)). Higher prices not only reduce utilization immediately, but also lower a habit stock. The lower stock gradually reduces utilization toward a new steady state by decreasing the marginal utility of utilization. While habit formation is paid limited attention in analyzing healthcare demand, there is ample evidence that it characterizes consumer behavior in other areas, like health behavior and energy consumption.[4]

Identifying the response dynamics to medical prices is challenging.[5] The reason is a lack of plausibly exogenous variation of prices to examine utilization, health, and health behavior over time; only randomized experiments, like the RAND Health Insurance Experiment, have been considered to be the best hope (Finkelstein et al., 2018).

This paper estimates the dynamics of consumer responses to medical care prices by using a quasi-experimental variation. We focus on the policy change of cost sharing in

[3]For example, see Chandra et al. (2010), Choudhry et al. (2011), Baicker et al. (2015), Brot-Goldberg et al. (2017), and Chandra et al. (2021)

[4]For example, see Allcott and Rogers (2014): Ito et al. (2018); Jessoe and Rapson (2014); Allcott and Kessler (2019); Charness and Gneezy (2009); Acland and Levy (2015); Royer et al. (2015); Loewenstein et al. (2016); Hussam et al. (2022); Just and Price (2013); List and Samek (2015); Schofield et al. (2015); Volpp et al. (2008); Volpp et al. (2009).

[5]A few studies focus on the provision of health insurance to children and find that the longer-term impact differs from the short-term impact, clarifying fiscal externalises of health insurance are smaller in the longer term (Wherry et al. (2018); Goodman-Bacon (2021)).

Japan's public insurance: the increase in coinsurance, from 10% to 20%, for those between ages 70-74, born after April 1944. This policy change produces a variation of coinsurance by cohort. Coinsurance for those born after April 1944 were 20%, while 10% for those born before April 1944, for the 5 years during ages 70-74. The five-year difference in coinsurance between those born before and after April 1944 enables us to estimate the causal impact of medical prices on utilization, health outcome, and health behavior over time.[6]

We use the Cohort Difference-in-discontinuities (Cohort Diff-in-disc) with administrative data.[7] This method is best understood by comparing with relevant methods. As we mentioned above, there is rich evidence on immediate responses, and the "Age Regression Discontinuity Design (Age RDD)" plays a big role there.[8] The Age RDD is the RDD using a specific age as the cutoff. It enables researchers to estimate short-term responses cleanly and precisely, because pooling multiple cohorts for each age helps control cohort fixed effects and increases sample size. Since the Age RDD does not allow to identify longer-term responses, we should take a different approach. The Cohort RDD, the RDD using a specific cohort as the cutoff, is a promising option (Wherry et al., 2018). Nonetheless, because it compares outcomes between two consecutive cohorts, cohort fixed effects and small sample size tend to be issues. The advantage and disadvantage are opposite between the Age RDD and the Cohort RDD. Our Cohort Diff-in-disc with population data gives estimates of longer-term responses while compensating the weaknesses of the Cohort RDD. We control cohort fixed effects by taking a difference between a jump at the threshold, April 1944, below the age of 70 when the coinsurance rate is the same for everyone at 30%, and a jump after age 70. Population data increases sample size.

We obtain data on utilization aggregated by month and year of birth from the National

[6]The Ministry of Health, Labour, and Welfare estimates the average life expectancy for each age in each year. In 2014, men who were 70 years old were expected to live 15.5 more years, while women were expected to live 19.8 more years on average. Thus, our study covers a reasonably long period of time for the elderly, one-third, or one-fourth of their remaining lives on average.

[7]Grembi et al. (2016) use Diff-in-disc to estimate the impact of fiscal rules on a fiscal discipline.

[8]See Card et al. (2008), Card et al. (2009), Anderson et al. (2012), Shigeoka (2014), Fukushima et al. (2016), Nilsson and Paul (2018), and Han et al. (2020), for example.

Database of Health Insurance Claims (NDB) covering almost all insurance claims. We are allowed to use total medical spending and the spending for some specific services, including outpatient and inpatient services. For health outcome, we examine mortality by the universal death records. We also examine self-reported health outcome and health behavior by using a government survey.

Our first finding is about the response dynamics of total spending. We find that total spending drops by 2.8% in 0.5 years later and by 3.5% in 3.5-4.5 years later. Namely, the impact after five years on total spending is similar to, or slightly larger than, the immediate impact. This finding suggests that an immediate response, which previous studies typically estimate, seems a good measure for consumer responsiveness to medical care prices.

We then move to the analysis by medical service. In the short run, like previous studies, consumer responsiveness is similar across services, while modestly larger for relatively discretionary care. For example, outpatient spending drops by 3.0% and laboratory tests decrease by 3.6%, while consumers reduce inpatient spending by 2.6%.

In the longer run, we find that responses are more diverged across services. In particular, discretionary services further decrease over time, while consumer responsiveness to less discretionary care does not change significantly. Outpatient spending drops by 4.4% and laboratory tests decrease by 5.0%, while consumers reduce inpatient spending by 3.0% in 3.5-4.5 years later.

To obtain the insight on what happens to discretionary services (and less discretionary services), we also examine the impact of higher coinsurance on health and health behavior. Ex-ante moral hazard predicts positive impacts on health and health behavior, while habit formation does not necessarily involve effects on them. We do not find discernible impacts on mortality, neither in the short nor longer run. Large and statistically significant impacts on self-reported health and health behavior are not observed as well. Given these results, our preferred explanation is that habit formation characterizes the response dynamics of discretionary care.

These findings allow for better understanding of our finding, the stable dynamics of total spending. While responses on discretionary services get larger over time, possibly owing to habit formation, stable responsiveness on less discretionary services dilutes those expanding impacts on total spending. Our study reveals that these contrasting forces exist and make the response dynamics of total spending relatively stable. We believe that this takeaway helps judge whether the estimate of immediate responses on each hand is a valid indicator for designing health insurance.

The remainder of this chapter is organized as follows. Section 2.2 describes the Japanese healthcare system and the policy change of coinsurance in 2014 in Japan. Section 2.3 explains our identification strategy: data and Cohort Diff-in-disc. Section 2.4 shows the results of Cohort Diff-in-disc. Section 2.5 discusses the external validity of our findings. Section 2.6 concludes this paper.

2.2 Background

In this section, we provide an overview of the Japanese healthcare system[9] and explain the change of coinsurance in 2014.

2.2.1 Japanese Healthcare System

The Japan's public health insurance provides universal coverage. Under the insurance, all citizens have the same benefit package of medical services. The package is comprehensive; it includes outpatient services, inpatient services, and prescription drugs. This setting helps circumvent an notable challenge in estimating consumer responsiveness to medical prices: an endogeneity problem stemming from the choice of a health insurance plan (Baicker and Goldman (2011); Shigeoka (2014); Fukushima et al. (2016); Ellis et al. (2017)).

[9] For details about Japanese healthcare system, see Ikegami et al. (2011), Kondo and Shigeoka (2013), Shigeoka (2014), and Fukushima et al. (2016).

Patients have access to any medical provider without going through a gatekeeper or without having a referral letter. For example, patients can visit large hospitals, rather than clinics, for outpatient care with less serious conditions.[10]

All fees for medical services are determined by the unique national fee schedule set by the Japanese government. As long as the same services are used, the same fees are applied to all patients and all medical providers. Therefore, from the view point of physicians, there are few incentives to influence healthcare demand of patients (Shigeoka (2014): Fukushima et al. (2016)). When patients use medical services, patients pay cost sharing of the fees at medical institutions, and medical providers are reimbursed from insurers.

2.2.2 Patient Cost Sharing and Change of Coinsurance in 2014

Cost sharing is characterized by coinsurance with out-of-pocket maximums (caps) imposed on a monthly basis; there is no deductible. Cost sharing is mainly determined by age and summarized in Table 2.1.[11] Until April 2014, the coinsurance rate is 30% for those ages 6-70. From the next month of reaching the age of 70, individuals become eligible to the lower coinsurance rate, 10%.

In December 2013, the Japanese government announced the policy change to raise the coinsurance rate from 10% to 20% for those between ages 70-74 from April 2014. Importantly, the new coinsurance rate, 20%, were applied to those born after April 2nd, 1944 only, while the lower coinsurance rate, 10%, remained assigned to those born before April 1st, 1944.

Figure 2.1 describes the schedules of coinsurance rates for those born before and after April 1944 by their age. For those born before April 1944 (dashed line), their coinsurance rates are 30% until the age of 70 and drop to 10% afterwards. For those born after April

[10] Outpatient visits to very large hospitals, like university hospitals, without a referral letter may require additional copayments.

[11] The coinsurance rate for high-income earners is 30%, irrespective of their age. As noted in Shigeoka (2014), only a limited fraction of patients above the age of 70, 7%, is classified as high-income earners (Ikegami et al., 2011).

1944 (solid line), their coinsurance rates are 30% until the age of 70 but become 20% for ages 70-74. The different coinsurance rates (20% vs 10%) are assigned based on the timing of birth (after April 1944 vs before April 1944) for 5 years between ages 70-74. By using this variation, we identify how consumers respond to the change of coinsurance from 10% to 20% over time.

One complication is non-linearlity imposed by cap where prices fall to zero (Keeler et al. (1977); Ellis (1986); Aron-Dine et al. (2013)). Caps decrease true shadow prices reflecting the zero price. If consumers respond to them, our study basically compares outcomes between two prices, different from 10% and 20%. Shigeoka (2014) argues that this complication is mitigated in Japan for two reasons. First, the cap is set monthly. The time for taking advantage of the zero price is limited. Even if patients immediately exceed the cap at the beginning of month, time with zero price is just one month, not one year. Second, the cap is set at the high level. Shigeoka (2014) shows that the probability of reaching the cap among claims for those who are above the age of 70 is 0.6% for outpatient visits and 0.0% for inpatient admissions, when coinsurance rates are 10%.

2.3 Data and Identification Strategy

2.3.1 Data

Utilization

Our data source for utilization is the National Database of Health Insurance Claims (NDB). The NDB covers almost all health insurance claims in Japan. We are allowed to use semi-aggregated data summed up three months from September to November in each year, from 2012 to 2019, by month and year of birth (cohort).[12] We therefore cannot distinguish those born on April 1st who were under the old policy from those born on April 2nd who were assigned higher coinsurance 20% during ages 70-74; we decide not to use information of

[12] We do not have access to the micro data of the NDB.

those born in April 1944.

Our data include the total medical spending, the number of health insurance claims, medical spending for outpatient visits, medical spending for inpatient admissions, and medical spending for some specific medical services. The specific services are laboratory tests for outpatients, prescription drugs for outpatients, outpatient visits with diagnosis classified as Ambulatory Care Sensitive Conditions (ACSCs) that the Agency for Healthcare Research and Quality develops for studying preventive care (Shigeoka, 2014), and inpatients with surgery.

Health and Health Behavior

We use health outcome similar to Finkelstein et al. (2012). The first is the monthly mortality rate. We use vital statistics covering almost all deaths in Japan from 2010 to 2021. The vital statistics report the date of birth, the date of death, and the cause of death by the ICD-10.[13] We calculate the number of deaths in each cell defined by cohort and the month and year of death, and divide by population to obtain monthly mortality rates for each cohort.[14]

We also use the Comprehensive Survey of Living Conditions (CSLC) by the Ministry of Health, Labour, and Welfare to examine other measures of health and health behavior. The timing of birth is reported at the month and year level; we again do not use information of those born in April 1944. The health-related survey of the CSLC is conducted in June every three years, and we use the 2013 wave, the 2016 wave, and the 2019 wave. Namely, we have data around 1 year before, 2 years after, and 5 years after the policy change.

The CSLC asks several questions that previous studies use to measure health status. For example, the CSLC includes self-reported health by five categories (very poor, poor, fair, good, and very good). The CSLC also contains information about health behavior. It

[13]We observe the surge of deaths in 2011 because of the Great East Japan Earthquake. We exclude deaths by T75.1 (drowning) and T14.9 (injury, unspecified) in ICD-10 from our sample.

[14]To estimate population by cohort and calendar month and year, we use the 2015 Census and the vital statistics. From the Census we can estimate population by cohort on October 1st in 2015. Then, by adding and subtracting the number of deaths calculated from the vital statistics, we estimate population by cohort at the beginning of each month in each year.

asks whether respondents take health-improving behaviors, including eating regular meals, doing exercises, and sleeping enough.

2.3.2 Identification Strategy

By focusing on the variation of coinsurance by cohort described in Figure 2.1, this study uses the "Cohort Difference-in-discontinuities Design (Cohort Diff-in-disc)" with population data to identify the impact of coinsurance over time. The design is best understood by comparing it with the "Age Regression Discontinuity Design (Age RDD)" and the "Cohort Regression Discontinuity Design (Cohort RDD)." Pros and cons of the Age RDD and the Cohort RDD are summarized in Table 2.2; they are opposite.

The Age RDD is often used to estimate consumer responsiveness to health care prices (Card et al. (2008); Card et al. (2009); Anderson et al. (2012); Shigeoka (2014); Fukushima et al. (2016); Nilsson and Paul (2018); Han et al. (2020)). The Age RDD is the RDD using a specific age as the cutoff. Card et al. (2008) and Card et al. (2009) use the age of Medicare eligibility to estimate its impact on utilization and health. Shigeoka (2014) and Fukushima et al. (2016) use the age of 70 when coinsurance sharply drops in Japan. Han et al. (2020) focuses on the cost-sharing subsidy for children under the age of 3 in Taiwan. These studies give clean and precise estimates of the impact of cost sharing. Because many cohorts can be pooled in each age, the Age RDD can control cohort fixed effects and increase sample size that is crucial for the RDD. These are the advantages of the Age RDD. In contrast, its disadvantage is that the Age RDD allows to estimate short-term impacts only. We should therefore deviate from the Age RDD.

The Cohort RDD, i.e. the RDD using a specific cohort as the cutoff, is an important option to identify longer-term impacts, because cohorts do not change over time. Wherry et al. (2018) exploits a discontinuity in childhood Medicaid eligibility based on cohorts to estimate its impact on later-life utilization of medical care. The Cohort RDD however comes with costs. Since pooling multiple cohorts is not possible, the Cohort RDD tends to have the problem of controlling cohort fixed effects and be bothered by a lack of power. Thus, the advantages and disadvantages between the Age RDD and Cohort RDD are exactly opposite.

Our Cohort Diff-in-disc with population data allows to estimate longer-term impacts while compensating the disadvantages of the Cohort RDD: cohort fixed effects and small sample size. The difference-in-discontinuities design takes the difference in two discontinuities; one is under the non-treatment period and another is under the treatment period (Grembi et al., 2016). The former discontinuity captures the original heterogeneity between the control group and the treatment group, while the latter captures treatment effects as well as the original heterogeneity. Our Cohort Diff-in-disc is the difference-in-discontinuities using a specific cohort as the cutoff. Specifically, we take the difference in two discontinuities between one below the age of 70, capturing cohort fixed effects, and another above 70, capturing cohort fixed effects and the causal impact of higher coinsurance. Population data increases the accuracy of estimates.

In particular, the regression equation is

$$\begin{aligned} Y_{ct} &= \alpha_0 + \beta_0 \cdot Post_c + f(c) + f(c) \cdot Post_c \\ &\quad + D_t \cdot \Big[\alpha_1 + \beta_1 \cdot Post_c + g(c) + g(c) \cdot Post_c\Big] + \varepsilon_{ct} \end{aligned} \tag{2.1}$$

where Y_{ct} is the outcome for cohort c at time t, $Post_c$ denotes the dummy variable for those born after April 1944, D_t are dummies omitting one specific time/age below the age of 70 as a base time/age, $f(c)$ and $g(c)$ are polynomial functions of cohorts without constants, and ε_{ct} is the unobserved error component.

β_0 captures cohort fixed effects identified by the information of a base time/age. The Cohort Diff-in-disc estimator is β_t. It represents the percent change of outcome by increasing coinsurance from 10% to 20% at time t. By using data at time t before or after April

2014, β_t captures the impact t years before or after the policy change. For example, if data in October 2018 is used, β_t represents the impact 4.5 years after the policy change.

In the utilization data and the mortality data, we observe individuals who used medical services or died only. We assume that the probability of utilization or death for underlying population smoothly changes with cohorts in a given time.[15] We thus use the log of outcome divided by population $log(y/pop)_{ct}$ for Y_{ct}. Linear functions are used for $f(c)$ and $g(c)$. We use 30 months both sides of the cutoff. We add a dummy for those below the age of 70, if necessary, whose coinsurance is 30%.

The health and health-behavior data is not population data. Moreover, linear functions seem to be restrictive. We therefore use 60 months both sides of the cutoff and quadratic functions for $f(c)$ and $g(c)$. As a robustness check, we use 30-months both sides of the cutoff and linear functions. We add a dummy for those below the age of 70, if necessary, whose coinsurance is 30%.

A challenge to infer the dynamics of consumer responses from estimating (2.1) is anticipatory spending (Einav et al. (2015); Brot-Goldberg et al. (2017)). Since individuals know the schedule of coinsurance by age, they could manipulate utilization around the time of changing coinsurance. If anticipatory spending exists and differs between those born before and after April 1944, estimates by (2.1) at around April 2014 (the age of 70) and April 2019 (the age of 75) are biased. For example, if anticipatory spending becomes larger when the experienced change of coinsurance is larger, we overestimate the reduction of utilization just after April 2014 and just before April 2019, hampering to understand the dynamics of consumer responses.

While the concern is legitimate, we argue that it is mitigated in our case. First, Shigeoka (2014) and Fukushima et al. (2016) estimate the change of utilization at the age of 70 in Japan and find that bias by anticipation is not large except for spending for inpatients with

[15] Card et al. (2008), Card et al. (2009), and Shigeoka (2014) assume that the probability of utilization for underlying population smoothly changes with ages, rather than cohorts. See Card et al. (2004) for discussions about this assumption.

surgery by comparing estimates based on the simple RDD and the donut-hole RDD. We therefore exclude spending for inpatients with surgery from our analysis. Second, those papers and other previous studies outside Japan tend to find that anticipatory spending is short-lived (Shigeoka (2014): Fukushima et al. (2016); Brot-Goldberg et al. (2017)). For example, Shigeoka (2014) exclude 6 months at maximum for his donut-hole RDD. Brot-Goldberg et al. (2017) find clear evidence that anticipatory spending with switching a free care plan to a high-deductible plan occurs within just 3 months. Thus, estimates of response dynamics excluding only few months around 70 and 75 should give valid inference. Finally, as we will show later, such anticipatory spending is not visible.

Consumer responsiveness to medical prices could vary with ages. This might also be a potential source of bias. Previous studies are suggestive that the elasticity of healthcare demand is smaller for older people because sicker individuals are less responsive to medical prices (Fukushima et al. (2016); Ellis et al. (2017)). Thus, this age effect could make our estimates getting smaller over time. We discuss this issue in the section of discussing results.

2.4 Results

2.4.1 Total Medical Spending

Age Profiles

Before estimating equation (2.1), we show the age profiles of total spending both for those born before April 1944 and for those born after April 1944. In Figure 2.2, by using multiple years of data, we plot the log of total spending per capita, controlling year fixed effects and birth-month fixed effects[16], against age for two groups: (i) the control group including those born during one year before April 1944, i.e. April 1943 to March 1944 (dashed line), and

[16] Specifically, we regress the log of total spending per capita on age, age squared, a dummy for age 70 or older, a dummy for age 75 or older, year fixed effects, and birth-month fixed effects by using the control group. We obtain the residual of the log of total spending per capita by subtracting those estimated fixed effects from the log of total spending per capita for each cell. Finally, we take the average of the residuals by age for each group.

(ii) the treatment group including those born during one year after April 1944, i.e. April 1944 to March 1945 (solid line). The coinsurance rate during ages 70-74 for the control group is 10% while that for the treatment group is 20%.

In this figure, four points are worth mentioning. First, below the age of 70 where both groups face 30% coinsurance, the levels and trends of their total spending are similar, although total spending of the treatment group seems slightly higher. This point indicates that both groups are comparable, while the Cohort Diff-in-disc, rather than Cohort RDD, is preferable. Second, at the age of 70 where coinsurance rates for both groups drop sharply, their total spending jump up discontinuously (Shigeoka (2014); Fukushima et al. (2016)). Furthermore, the control group that experiences the larger drop of coinsurance, from 30% to 10%, increases its spending more, by 5.1% (SE 0.5% and elasticity 0.047), while the treatment group experiencing the smaller drop of coinsurance, from 30% to 20%, raises their spending by 1.7% (SE 0.5% and elasticity 0.042).[17] These short-term responses are consistent with the consensus in previous studies that lower cost sharing increases utilization immediately. Furthermore, the similarity of the estimated elasticities suggests that a demand curve for medical care is close to log-linear, at around 10% to 30% coinsurance that are often found in health insurance contracts.[18] This finding indicates that while many studies, including ours, are limited to compare utilization under two prices around those ranges of coinsurance rates, estimated responses might be widely applicable.[19]

The last two points are directly related to our research question. Third and most impor-

[17]To estimate the increases in total spending at the age of 70 reported here, we use the Age RDD with the age of 70 as the cutoff (Shigeoka (2014); Fukushima et al. (2016)). Specifically, we regress the log of total spending per capita controlling year fixed effects and birth-month fixed effects on constant, age, age-squared, and a dummy for age 70 or older, fully interacted with a dummy for those born after April 1944. We use those aged 68-72 as the sample. We report elasticity calculated by using the log difference in coinsurance rates as denominator.

[18]To check the robustness of our finding on log-linearity, we do the same Age RDD by pooling multiple cohorts to increase its power. The result shows that elasticities for the change from 30% to 10% and for the one from 30% to 20% are almost identical, 0.049 vs 0.049, further supporting the log-linearity of the demand.

[19]Note that several states mention that the relationship between price and utilization is highly nonlinear, particularly for the outside of this range of coinsurance. Anderson et al. (2012) mentions that variations in the coinsurance rate from 25% to 95% do not affect inpatient admissions in the RAND experiment. Iizuka and Shigeoka (2022) shows that the zero price disproportionately boosts demand of child healthcare.

tantly, between ages 70-74 where the treatment and control group face different coinsurance rates, 20% vs 10%, their age profiles of total spending are almost parallel. The difference in their total spending does not change over time. The simple plot of age profiles provides an indicative answer to our research question: the longer-term response is similar to the short-term response. Lastly, above the age of 75 where both groups face the same 10% coinsurance again, the levels of their spending are back to the same level. Given that the treatment group uses medical care slightly more than the control group below the age of 70, there seems to exist small persistence of effects of coinsurance.

Cohort RDD

We next show Cohort RDD figures to describe how the Cohort Diff-in-disc works. Figure 2.3 plots the log of total spending per capita against cohort in 2012-2013 (Panel A), 2014-2015 (Panel B), 2016-2018 (Panel C), and 2019 (Panel D). We absorb birth-month fixed effects and jumps at the age of 70 and 75, if necessary. In the horizontal axis, we normalize the cutoff, April 1944, to zero. On the left (right) hand side of the cutoff with negative (positive) values, the outcome of those born after (before) April 1944 are plotted by solid (dashed) lines. Numbers in percentages from 10% to 30% represent coinsurance rates around the cutoff.

Panel B and Panel C exhibit that there are clear discontinuities at the cutoff during ages 70-74, after the policy change of coinsurance in 2014, and the magnitudes of the jumps are similar, 2.8% (SE 0.5%) in Panel B and 2.8% (SE 0.4%) in Panel C. Panel D shows a less visible discontinuity after individuals born around the cutoff reach the age of 75 and face the same 10% coinsurance. The jump corresponds to 1.2% (SE 0.6%). These discontinuities might not represent causal effects of coinsurance, however. In fact, Panel A shows that individuals born after April 1944 use more medical care originally by 1.6% (SE 0.5%). We assume that this original gap captures cohort fixed effects. Then, taking the difference of discontinuities between Panel B-D and Panel A gives us unbiased estimates of the impact of coinsurance. This is what equation (2.1) does. The reduction caused by higher coinsurance should be larger than the observed jumps showed in Panel B-D, presumably by about 1.6%.[20]

Cohort Difference-in-Discontinuities

We now show the results obtained by Cohort Diff-in-disc. We estimate equation (2.1) by using information in October 2013 as a base period; i.e., t = October 2013 is omitted from time dummies D_t. We use 25 months both sides of the cutoff. We add a dummy for 30% coinsurance for those below the age of 70, if necessary. Figure 2.4 shows the Cohort Diff-in-disc estimates β_t with 95 percent confidence intervals from t = October 2012 to October 2019, namely from 1.5 years before to 5.5 years after the policy change in April 2014 (from -1.5 to 5.5 in the figure). Note that a square marker at 0 represents an estimate from regressing the log of outcome on age and a dummy for age 70, interacted with a dummy for those born after April 1944, controlling year fixed effects and birth-month fixed effects, instead of using equation (2.1). All coefficients and standard errors are multiplied by 100 so that they are interpreted as percent changes.

Before the policy change, the estimate in 1.5 years before the policy change (-1.5) is almost zero and statistically insignificant at the conventional level (Coeff 0.2% and SE 0.9%), reassuring the validity of our identification strategy.

After the policy change, estimates by Cohort Diff-in-disc turn to negative statistically significantly. The estimates show that the change of coinsurance from 10% to 20% reduce total spending by around 3.5% to 4.0% (elasticity 0.05 to 0.06), and the reduction is almost unchanged, or becomes slightly larger in absolute value, over time from 0 to 4.5 years.[21]

[20] Note that these figures plot the outcome controlling for birth-month fixed effects. Without the fixed effects, the observed patterns in these figures are more emphasized. Thus, although including birth-month fixed effects alleviate bias owing to cohort specific effects, the Cohort Diff-in-disc is needed.

[21] Aron-Dine et al. (2013) argues that summarizing the price responsiveness of non-linear insurance contract by single elasticity needs considerable caution. With this caution in mind, a comparison of elasticity with recent studies targeting adults or the elderly tells us that our estimated elasticity is close to theirs that range from around one-quarter to one-half of the well-mentioned RAND estimates by Keeler and Rolph (1988) (Chandra et al. (2010); Shigeoka (2014); Fukushima et al. (2016); Brot-Goldberg et al. (2017)) once we use

In particular, the reduction is 3.4% (SE 0.6% and elasticity 0.049) at the time of change and 3.4% (SE 0.8%) in 0.5 years while 4.0% (SE 0.8%) in 3.5 years and 4.2% (SE 0.8% and elasticity 0.061) in 4.5 years. Thus, the reduction becomes larger by around 23.5% (0.8 p.p.) for 4.5 years. This suggests that the longer-term impact of coinsurance on total spending is similar to, or slightly larger than, the short-term impact.

Furthermore and interestingly, in 5.5 years after the policy change when individuals around the cutoff reach the age of 75, those born after April 1944 who were exposed to higher coinsurance during ages 70-74 use medical services less by 1.8% (SE 0.8%). This implies that even though those both sides of the cutoff face the same coinsurance rate, only 57.1% ($= (1.8\% - 4.2\%)/4.2\%$) of the gap between them at 4.5 years after the policy change is narrowed; that is, 42.9% of the effect of higher coinsurance remains existing. This means that the effect of coinsurance is persistent. In summary, the similarity of short-term and longer-term impact and the persist impact of coinsurance are our main findings on total spending.

2.4.2 Utilization by Type of Medical Services

The finding that the impact of coinsurance on total spending does not change significantly over time could mask important heterogeneity across type of medical services. We thus repeat the same analyses as total spending for three subcategories: the number of health insurance claims, medical spending for outpatients, and medical spending for inpatients. Note that while total spending depends on the amount of medical resources used, the total number of claims reflects whether individuals use medical resources within a specific length of time that health insurance claims are issued.[22] Comparing these outcomes tells us of which services consumers change their utilization, relatively cheap services or expensive services. For example, if consumers reduce the number of claims but do not change total

the same measure of elasticity, including arc elasticity. See Appendix for details.

[22] In Japan, if a patient goes to a medical provider, one health insurance claim is issued per month.

spending significantly, they should reduce relatively cheap services.

Age Profiles

We again begin with showing the age profiles of three subcategories of medical services. We first discuss the total number of claims and spending for outpatients because their results are qualitatively similar. Figure 2.5 plots the age profile for total number of claims and Figure 2.6 for spending for outpatients. The first two points for total spending mentioned in previous Section are also applied; the control and treatment group are comparable, and there are clear jumps at the age of 70 and the jump for the control group is larger. More interestingly, the third and fourth points are different, at least quantitatively, from the ones for total spending. Specifically, the difference between two groups becomes larger over time, and the utilization of these services by the treatment group are clearly lower even after the age of 75.

Compare to those figures, Figure 2.7 that depicts the age profile of spending for inpatient is contrasting. In particular, the profiles for the control and treatment group are almost parallel during ages 70-74, indicating that the reduction by higher coinsurance does not change over time. In addition, given that the treatment groups use more inpatient services below 70, there seems not strong persistent effects of coinsurance.

These three profiles suggest an interesting pattern of consumer responses. By assuming that cheaper and outpatient services are relatively discretionary while inpatient services are involved with more serious conditions, the profiles indicate that the impact of coinsurance becomes larger and persistent for more discretionary services. In the following, we check whether this pattern exists formally by Cohort Diff-in-disc.

Cohort Difference-in-Discontinuities

The results by Cohort Diff-in-disc are represented in Figure 2.8 for the total number of claims, in Figure 2.9 for outpatient spending, and in Figure 2.10 for inpatient spending. We use the same bandwidth and specification as for total spending.

Again, the results for the first two services are qualitatively similar, while estimates for outpatient spending are not as precise as the number of claims. The impacts of higher coinsurance become larger over time and are strongly persistent. More specifically, regarding the total number of claims, the reduction is 3.4% (SE 0.3% and elasticity 0.049) at the time of change and 4.4% (SE 0.4%) in 0.5 years while 5.5% (SE 0.3%) in 3.5 years and 5.8% (SE 0.3% and elasticity 0.084) in 4.5 years. Thus, the reduction becomes larger by around 70.6% (2.4 p.p.) for 4.5 years. Furthermore, the reduction is 3.7% (SE 0.3%) in 5.5 years. This implies that even though those both sides of the cutoff face the same coinsurance rate, only 36.2% ($= (3.7\% - 5.8\%)/5.8\%$) of the gap between them at 4.5 years after the policy change is narrowed; that is, 63.8% of the effect of higher coinsurance remains existing.

As to outpatient spending, the reduction is 3.5% (SE 0.6% and elasticity 0.050) at the time of change and 4.3% (SE 0.9%) in 0.5 years while 5.5% (SE 0.8%) in 3.5 years and 5.0% (SE 0.9% and elasticity 0.072) in 4.5 years. The reduction becomes larger by around 42.9% (1.5 p.p). In addition, the reduction is 2.8% (SE 0.8%) in 5.5 years, so 56.0% of the effect of higher coinsurance remains.

In contrast, we do not find clear evidence that the longer-term impact is different from the short-term impact and the impact is persistent. In particular, the reduction is 3.4% (SE 0.8% and elasticity 0.049) at the time of change and 2.7% (SE 0.9%) in 0.5 years while 2.9% (SE 1.0%) in 3.5 years and 3.7% (SE 0.9% and elasticity 0.053) in 4.5 years. The reduction is 1.2% (SE 1.0%) in 5.5 years, which is not statistically significant at the conventional level.

These results for three subcategories increase a resolution for our findings on total spending. The difference between the short-term impact and the longer-term impact of total spending is mostly driven by cheaper and outpatient services, which tend to be discretionary, while the impact on inpatient services does not vary over time. As a result of the combination of these effects, we find that the longer-term impact on total spending is similar, or slightly larger than, the short-term impact. Specifically, the reduction of total spending in 4.5 years is larger than the one at 0 by 0.8 percentage points. A back of envelop calculation suggests that the contribution of outpatient spending is 0.6 percentage points ($= 1.5$ p.p.$*42.3\%$ (the share of outpatient spending)), explaining most of the reduction of total spending. We also find that the reduction is 1.8% in 5.5 years. The contribution of outpatient spending is 1.2% ($= 2.8\% * 42.3\%$).

More generally, this paper asks whether short-term estimates are valid for measuring ex-post moral hazard of lower patient cost sharing. Our results indicate that the answer to the question depends on type of services that a health insurance contract of interest covers. If a contract covers more discretionary services, like cheaper and outpatient services, the associated short-term estimate may not capture ex-post moral hazard fully. Instead, if an contract is skewed to services for serious conditions, short-term estimates are a good measure. For instance, in our case that focuses on Japanese public health insurance covering both outpatient and inpatient care, short-term estimates, like Shigeoka (2014) and Iizuka and Shigeoka (2021), should not be far from true costs.

2.4.3 Health and Health Behavior

In previous section, we find that consumers gradually decrease the utilization of discretionary services over time. This is consistent either with ex-ante moral hazard or habit formation. A key to distinguish them is whether health or health behavior changes. If ex-ante moral hazard drives our finding, we should observe positive impacts on health and health behavior. Instead, if habit formation is a driver, such changes are not expected to occur.

Mortality

Age Profiles

Figure 2.11 shows the age profiles of mortality rates for the control group and for the treatment group. The control group includes those born between April 1943 and March 1944, while the treatment group includes those born between April 1944 and March 1945. For each group, we plot the average of the log of monthly mortality rates by age, excluding death-year fixed effects, death-month fixed effects, and birth-month fixed effects.[23]

Unlike utilization, there is no visible difference between two groups. Below the age of 70, the levels and trends of mortality rates for those groups are similar. At the age of 70, we do not find large and statistically significant discontinuities (Coef. -0.3% and SE 1.3% for the control group and Coef. 0.3% and SE 1.1% for the treatment group).[24] Above 70 mortality rates for those groups move in the same way. These observations indicate that the change of coinsurance from 10% to 20% does not affect mortality rates significantly.

Cohort Difference-in-Discontinuities

Unlike utilization, our mortality data includes the number of deaths in every month. Although we basically do the same estimation as utilization, we make two modifications. First, we pool 12 months of data to estimate an impact on mortality at each point of time, from 3.5 years before and 6.5 years after the change of coinsurance. For example, we use data from April 2018 to March 2019 to identify the impact on mortality 4.5 years later. Second, we use the age of 69 as a base time taking $D_t = 0$ in equation (2.1). The age is desirable as a base time since everyone faces the same coinsurance, 30%.

[23]Specifically, we first collapse data into cells defined by the month and year of birth (cohort) and the month and year of death, and we divide them by population to obtain monthly mortality rates for each cohort. We then regress the log of mortality rates on age, age squared, death-year fixed effects, death-month fixed effects, and birth-month fixed effects by using the control group. We obtain the residual of the log of mortality rates by subtracting those estimated fixed effects from the log of mortality rates for each cell. Finally, we take the average of the residuals by age for each group.

[24]We regress the log of mortality rates on age, a dummy for age 70 or older, and the interaction term between age and the dummy, for each group, aged 67.5-74.5.

The estimates by Cohort Diff-in-disc are reported in Figure 2.12.[25] Overall, the estimates are not very large and statistically insignificant at any point of time. Most of the point estimates fall within the range of -2% to 2%. The monthly mortality rate for those born in March 1944 between ages 70-74, for example, is around 110 deaths per 100,000 people. Thus, the estimates indicate that the increase in coinsurance rates might cause ± 2 deaths per 100,000 people and are not statistically significant. In summary, we do not find any large impact on mortality over time.

Other Health Outcome and Health Behavior

We also investigate other measures of health and health behavior by Cohort Diff-in-disc.[26] The results of health are summarized in Table 2.4. Panel A reports the results of using the 2016 wave of the CSLC and Panel B uses its 2019 wave. Column (1) examines self-reported health fair, good, or very good. Column (2) examines whether physical or mental health impaired usual activities in the past month, while Column (3) uses the number of days impaired by physical or mental health in the past month. Column (4) uses a question about whether respondents were sick or injured within a few days. Column (5) focuses on whether respondents feel stress about their health now. Column (6) examines whether health problems affect their daily life now. For any measure, the impacts of higher coinsurance, represented by the coefficients of the interaction term between $Post_c$ and D_t, do not seem very large and can not distinguish from zero. We do not find discernible impacts on health outcome neither two years later nor five years later.

Table 2.5 shows the results of health behavior. We construct binary measures from questions asking whether respondents do the followings for maintaining health: (1) eating regular meals, (2) doing exercise, (3) sleeping enough, (4) not smoking, (5) not drinking, and (6) nothing. Like health outcome, we do not find large and statistically significant changes of health behavior in both waves.

[25] We thank Toshiaki Iizuka for his suggestion to examine the impact over time.
[26] RD figures are in Appendix.

In summary, we do not find discernible impacts on health and health behavior neither in the short run nor in the longer run. While the reduction of discretionary services gets larger over time, this is not accompanied with the change of health and health behavior. This evidence is suggestive that habit formation drives the dynamics of consumer responses for discretionary services, but it does not strongly influence the consumption of less discretionary services. While we do not take ex-ante moral hazard as the driver of response dynamics of discretionary services, this conclusion is consistent with literature in which ex-ante moral hazard in health insurance is considered to be irrelevant (Einav and Finkelstein, 2017).

As we discuss above, the effect of getting older could contaminate our estimates, possibly making our estimates getting smaller over time. We find that consumer responsiveness does not change over time for less discretionary care and gets larger for discretionary care, possibly due to habit formation. Thus, if the age effect is important, the true impact of coinsurance should be larger than our estimates in the longer run.

2.5 Discussion - External Validity -

We find that the longer-term impact of coinsurance on total spending is similar to, or slightly larger than, the immediate impact, as a result of the combination of the contrasting forces between discretionary care, possibly characterized by habit formation, and less discretionary care. While we believe that this finding is helpful to judge, at least qualitatively, whether immediate responses can be used for other cases than Japan's public health insurance, the quantitative relationship between the short-term and longer-term impacts might depend on other factors as well.

In particular, we do not find that the reduction of utilization deteriorates health and boosts healthcare demand in the long run, in net. While we cannot identify its reason here, one possibility is that access to healthcare is relatively good in Japan. Access to medical services not only depend on cost sharing but also other factors, including full price of medical services, a gatekeeping, and the number of medical providers. Some statistics indicate that the goodness of access in Japan. For example, the number of beds per 1,000 is 13.0 while 2.9 for the United States, and the annual average number of outpatient visits is around 13.0 in Japan while 4.0 for the United States (OECD Health Statistics 2020).

Good access could make cost sharing less blunt because better access leads to higher utilization, making the marginal value of care lower. The reduction of marginal services has smaller impacts on health. Better access could also mitigate behavioral hazard. Intuitively, individuals are relatively easy to know what services are unnecessary, if they visit a physician's office with very minor deceases.

We recognize existing evidence on behavioral hazard (Baicker et al. (2015); Chandra et al. (2010); Chandra et al. (2021)). In cases where behavioral hazard poses serious effects, the dynamics of total spending might tilt toward the smaller reduction in the long run. Thus, while our study is useful to pop up the response dynamics of discretionary services possibly by habit formation, we believe that this paper also highlights the importance of research on behavioral hazard to understand what makes it more serious.

2.6 Conclusion

This paper estimates the dynamics of consumer responses to medical care prices. The estimation is considered to be challenging because of a lack of plausible variation (Finkelstein et al., 2018). We overcome this challenge by exploiting a quasi-experimental variation. Specifically, we use the increase in coinsurance rates, from 10% to 20%, for those between ages 70-74, born after April 1944. This variation allows us to estimate the impact of coinsurance on utilization over time by controlling cohort fixed effects by Cohort Diff-in-disc. Our population data helps increase the sample size.

We find that the impact of coinsurance on total medical spending is stable, or slightly gets larger, over time. This suggests that immediate responses, typically estimated by previous studies, are valid for measuring consumer responsiveness to medical care prices. Furthermore, our study reveals the determinants of the response dynamics of total spending. We find contrasting forces. First, the change of discretionary services becomes larger over time. Given no discernible impacts on health and health behavior, our preferred explanation is that habit formation drives the dynamics of these services. As far as we know, such increasing impacts on discretionary services by habit formation are not highlighted in literature. Second, consumer responses to less discretionary services remain unchanged over time. Their response dynamics is not strongly affected by habit formation. The combination of these forces determines the dynamics of total spending, which tends to make it relatively stable.

2.7 Table

Table 2.1: Patient Cost Sharing in Japan

	Below the Age of 70		Between Ages 70 - 74: Born Before April 1944		Between Ages 70 - 74: Born After April 1944		Above the Age of 75	
	CO (%)	Cap (¥K)	**CO (%)**	Cap (¥K)	**CO (%)**	Cap (¥K)	CO (%)	Cap (¥K)
Apr 2008	30	$80.1 + (TC - 267) \times 1\%$	**10**	44.4	**-**		10	44.4
Apr 2014	30	$80.1 + (TC - 267) \times 1\%$	**10**	44.4	**20**	44.4	10	44.4
Aug 2017	30	$80.1 + (TC - 267) \times 1\%$	**10**	57.6	**20**	57.6	10	57.6
Aug 2018	30	$80.1 + (TC - 267) \times 1\%$	**10**	57.6	**20**	57.6	10	57.6

Notes: CO stands for coinsurance. Caps are expressed in thousand JPY (¥K) and TC means household's total medical expenditure per month. Caps are imposed monthly at the household level. From August 2017, if a household exceeds its cap three times within a single year, a cap for the remaining months drops to 44.4 thousand JPY. In addition to the caps in the table, there are caps for outpatient visits for individuals above the age of 70: 12.0 thousand yen until July 2017, 14.0 thousand yen from August 2017 to July 2018, and 18.0 thousand yen from August 2018.

Table 2.2: Comparison between Age RDD and Cohort RDD

	Pros	Cons
Age RDD	Cohort fixed effects & Sample size	Short-term impacts
Cohort RDD	Longer-term impacts	Cohort fixed effects & Sample size

Table 2.3: Cohort Diff-in-disc Estimates by Type of Services

	0.5 yrs after		3.5 yrs after		4.5 yrs after	
	Coeff	Elast	Coeff	Elast	Coeff	Elast
Total						
Spending	-2.79	0.040	-3.42	0.049	-3.64	0.053
	(0.90)		(0.78)		(0.79)	
No. of Claims	-3.99	0.058	-5.15	0.074	-5.41	0.078
	(0.35)		(0.32)		(0.31)	
Outpatient						
All	-3.02	0.044	-4.63	0.067	-4.13	0.060
	(0.96)		(0.77)		(0.92)	
Lab Tests	-3.61	0.052	-4.86	0.070	-5.18	0.075
	(0.51)		(0.41)		(0.47)	
Drugs	-4.04	0.058	-5.02	0.072	-5.40	0.078
	(0.37)		(0.32)		(0.33)	
ACSCs	-1.39	0.020	-0.05	0.001	1.04	-0.015
	(2.25)		(1.73)		(1.88)	
Inpatient						
Without Surgery	-2.62	0.038	-2.61	0.038	-3.30	0.048
	(0.99)		(0.98)		(0.98)	

Notes: Each cell in the column ”Coefficient” reports the estimate of coefficient for post dummy that takes the vale of one for those born after April 1944 in equation (**??**). Robust standard errors are in parentheses. All estimates of coefficients and standard errors are multiplied by 100 so that they can be interpreted as percent changes. In the column ”Elasticity,” we report the implied elasticity that is obtained by dividing the estimate of coefficient by $(ln(0.2) - ln(0.1))$. We use data in 2015-2018, and collapse them by month and year of birth. We use 15 months in both sides from the cutoff, April 1944; sample size is 30. The specification is linear in month and year of birth, interacted with post dummy, and fixed effects for month of birth. See RD figures in appendix **??**

Table 2.4: Cohort Diff-in-disc Estimates: Health Outcome

	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: June 2016						
Post	-0.03	-0.34	-0.06	-0.72	1.42	1.80
	(1.33)	(0.95)	(0.01)	(1.59)	(0.95)	(1.15)
Post * D	1.33	-0.17	-0.02	-1.07	0.31	-1.39
	(1.64)	(1.05)	(0.01)	(1.96)	(1.47)	(1.68)
Observations	160,343	158,902	158,463	161,337	159,863	157,512
Panel B: June 2019						
Post	-0.04	-0.48	-0.06	-0.82	1.46	1.78
	(1.33)	(0.93)	(0.01)	(1.56)	(0.97)	(1.17)
Post * D	-1.24	0.60	0.02	1.03	-1.48	0.83
	(2.16)	(1.23)	(0.02)	(1.94)	(1.63)	(2.11)
Observations	152,427	150,819	150,605	153,252	151,746	149,260

Notes: This table reports coefficients of estimating equation (2.1). Robust standard errors are in parentheses. All coefficients and standard errors, except for Column (3), are multiplied by 100 so that they can be interpreted as percent changes. Quadratic functions are used for polynomial functions of cohorts. We use 60 months both sides of the cutoff. Panel A uses the 2016 wave and Panel B uses the 2019 wave. Column (1)-(6) use different measures of health outcome. Column (1) uses a binary measure about whether self-reported health is fair, good, or very good. Column (2) uses a binary measure about whether physical or mental health impaired usual activities in the past month. Column (3) uses the number of days impaired by physical or mental health in the past month. Column (4) uses a binary measure about whether respondents were sick or injured within a few days. Column (5) uses a binary measure about whether respondents feel stress about their health now. Column (6) uses a binary measure about whether health problems affect their daily life now.

Table 2.5: Cohort Diff-in-disc Estimates: Health Behavior

	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: June 2016						
Post	-0.86	-0.87	-1.79	0.34	0.08	0.25
	(1.03)	(1.33)	(1.35)	(1.24)	(1.30)	(0.58)
Post * D	1.27	1.87	2.49	1.88	-0.31	-0.89
	(1.49)	(2.24)	(2.23)	(2.10)	(1.93)	(0.85)
Observations	162,799	162,799	162,799	162,799	162,799	162,799
Panel B: June 2019						
Post	-0.88	-0.87	-1.97	0.39	0.10	0.28
	(1.02)	(1.35)	(1.35)	(1.26)	(1.33)	(0.58)
Post * D	0.15	0.32	3.06	0.13	-1.19	-0.40
	(1.35)	(1.94)	(2.02)	(1.88)	(1.95)	(0.73)
Observations	154,691	154,691	154,691	154,691	154,691	154,691

Notes: This table reports coefficients of estimating equation (2.1). Robust standard errors are in parentheses. All coefficients and standard errors are multiplied by 100 so that they can be interpreted as percent changes. Quadratic functions are used for polynomial functions of cohorts. We use 60 months both sides of the cutoff. Panel A uses the 2016 wave and Panel B uses the 2019 wave. Column (1)-(6) use different measures of health behavior. Column (1) uses a binary measure about whether respondents eat regular meals. Column (2) use a binary measure about whether respondents do exercises. Column (3) use a binary measure about whether respondents take enough sleep. Column (4) use a binary measure about whether respondents do not smoke. Column (5) use a binary measure about whether respondents do not drink. Column (6) use a binary measure about whether respondents do nothing.

2.8 Figure

Figure 2.1: The Schedule of Coinsurance Rates around Ages 70-74

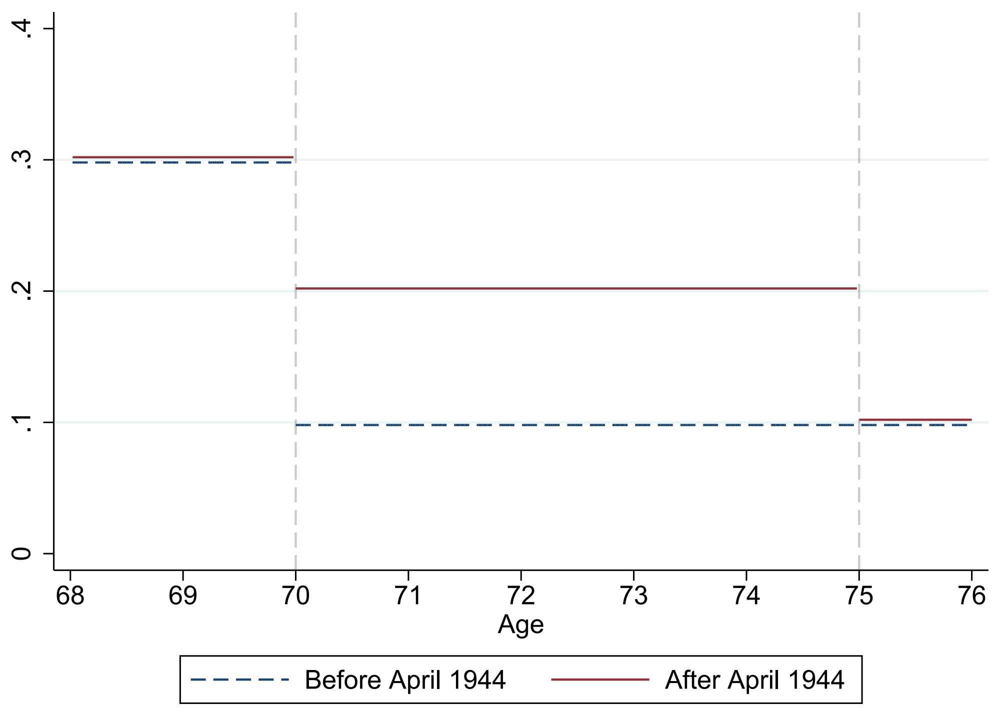
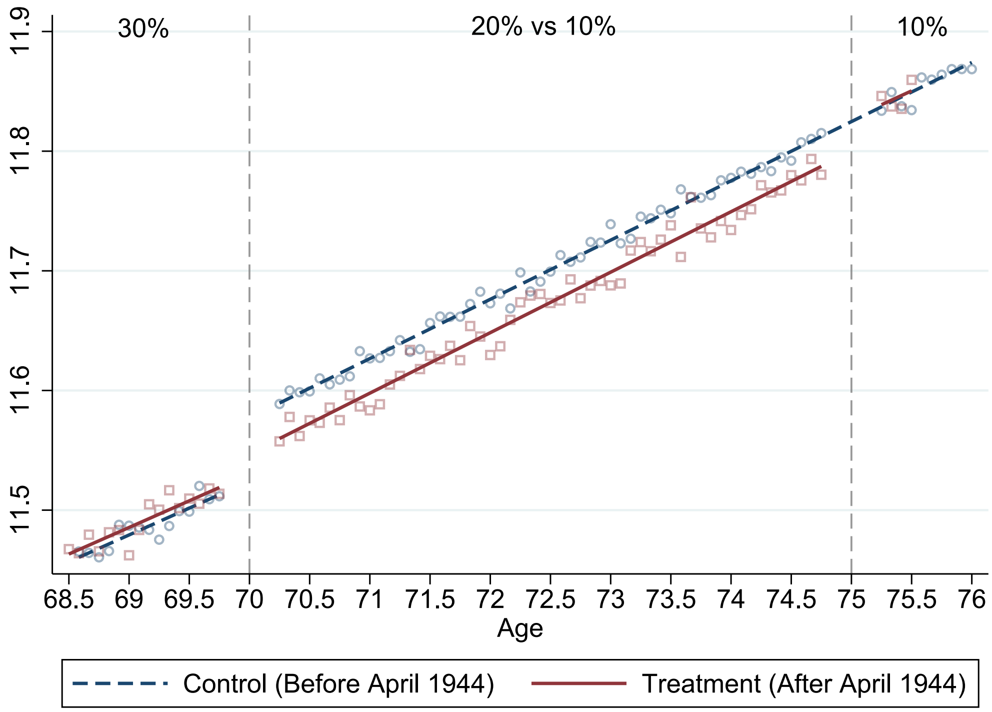
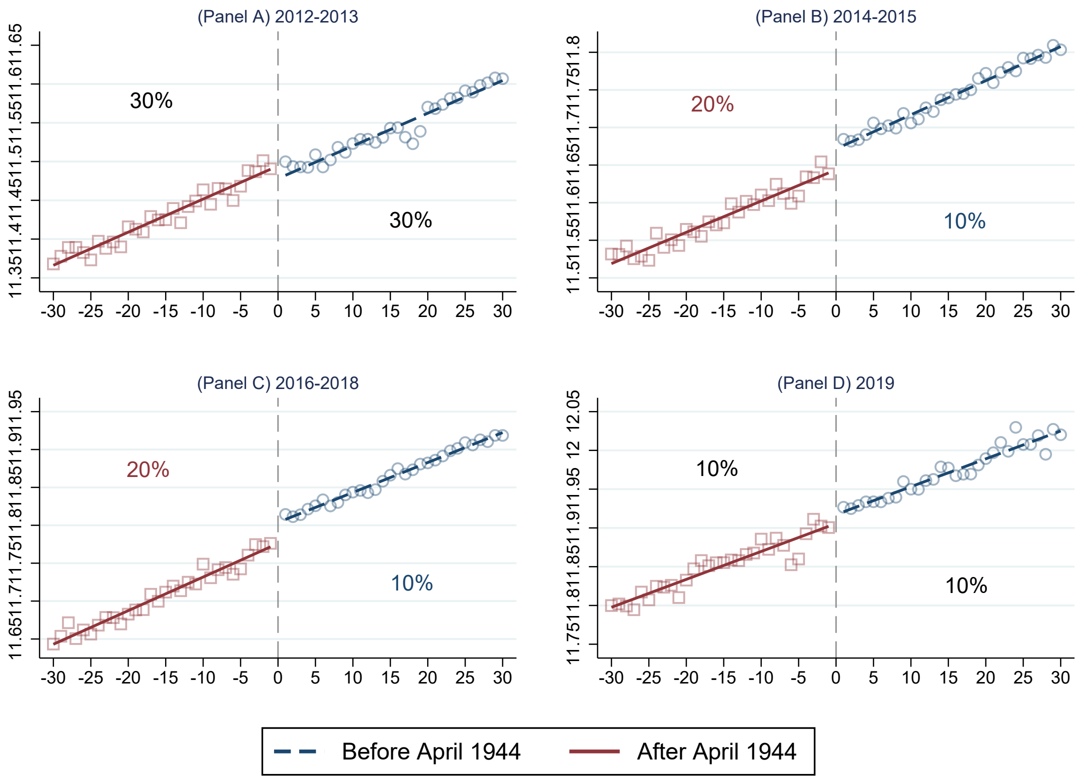


Figure 2.2: The Age Profiles of Total Medical Spending



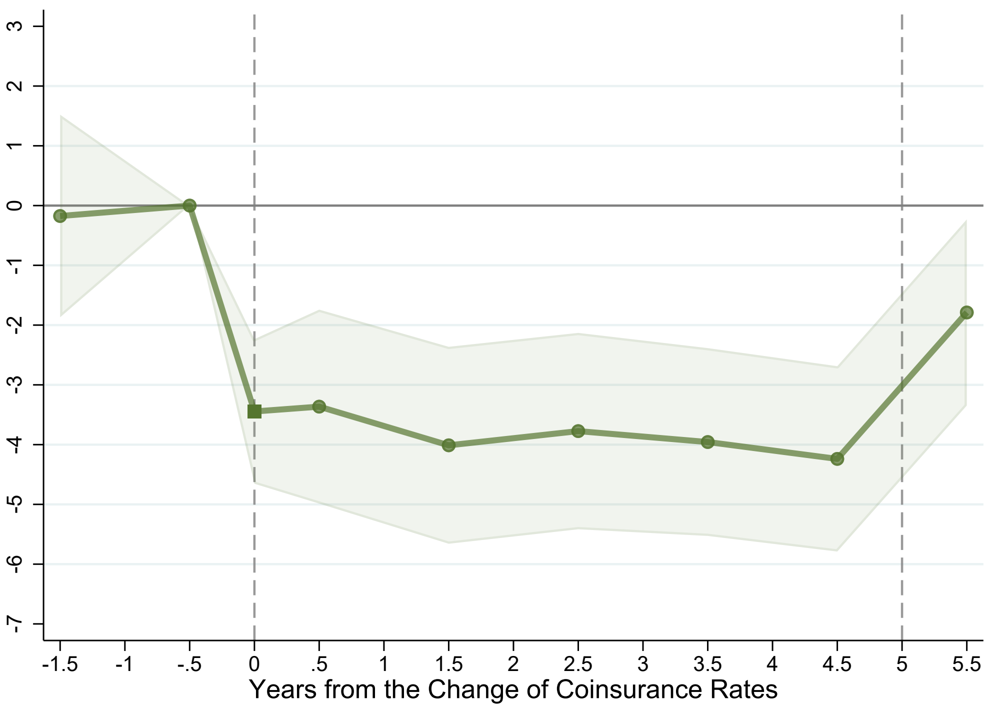
Notes: This figure plots the log of total medical expenditure per capita by age for treated group including those born after April 1944 (solid line) and for the control group including those born before April 1944 (dashed line). The treated group includes those who were born between April 1942 and March 1944. Their coinsurance rate between 70-74 is 20%. The control group includes those born between April 1944 and March 1946. Their coinsurance rate between 70-74 is 10%. Each marker represents the log of outcome excluding year fixed effects and birth-month fixed effects. These effects are identified by regressing the log of outcome on quadratic in age, year fixed effects, and birth-month fixed effects for the control group. Since the data is the sum of three months (September to November) of the total medical expenditure, age in months in this figure represents age in October. Markers around age 70 and 75 are excluded.

Figure 2.3: Cohort RDD: Total Medical Spending



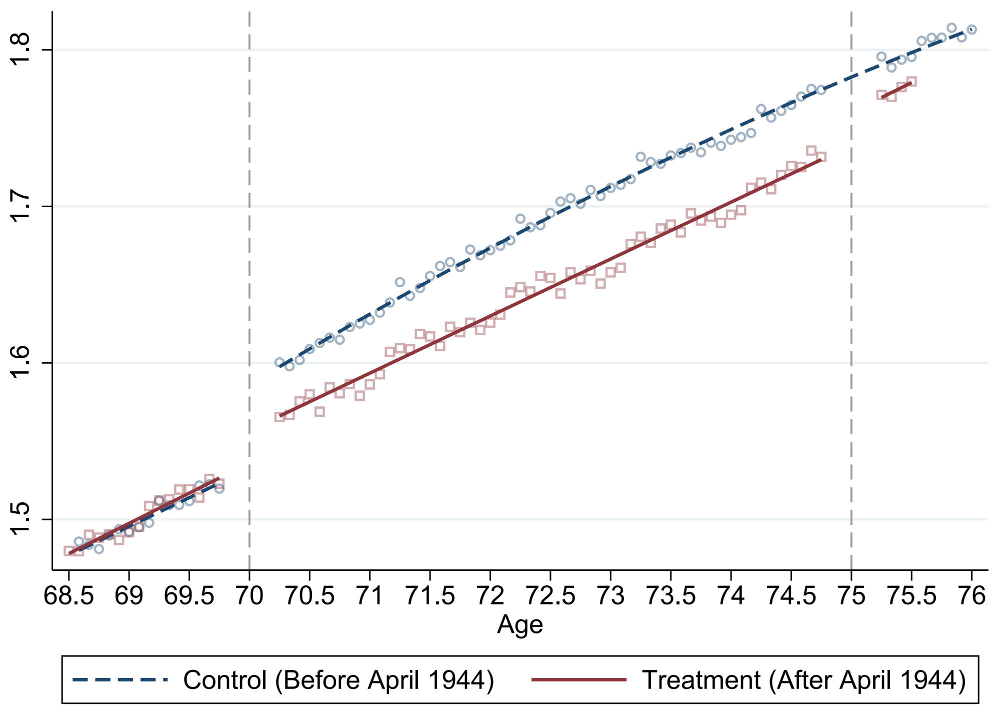
Notes: Panel A pools data in 2012 and 2013, Panel B in 2014 and 2015, and Panel C in 2016, 2017, and 2018. Panel D uses data in 2019. Numbers in percentages from 10% to 30% are coinsurance rates that individuals around the cutoff belongs to. Each marker represents the log of outcome excluding birth-month fixed effects and jumps at age 70 and 75. The lines represent fitted values regressing on linear in month and year of birth, interacted with a dummy that represents those born after April 1944. Horizontal axes represent distance from April 1944. Larger numbers represent older cohorts. For example, 12 indicates those born in April 1945.

Figure 2.4: Cohort Diff-in-disc Estimates by Year: Total Medical Spending



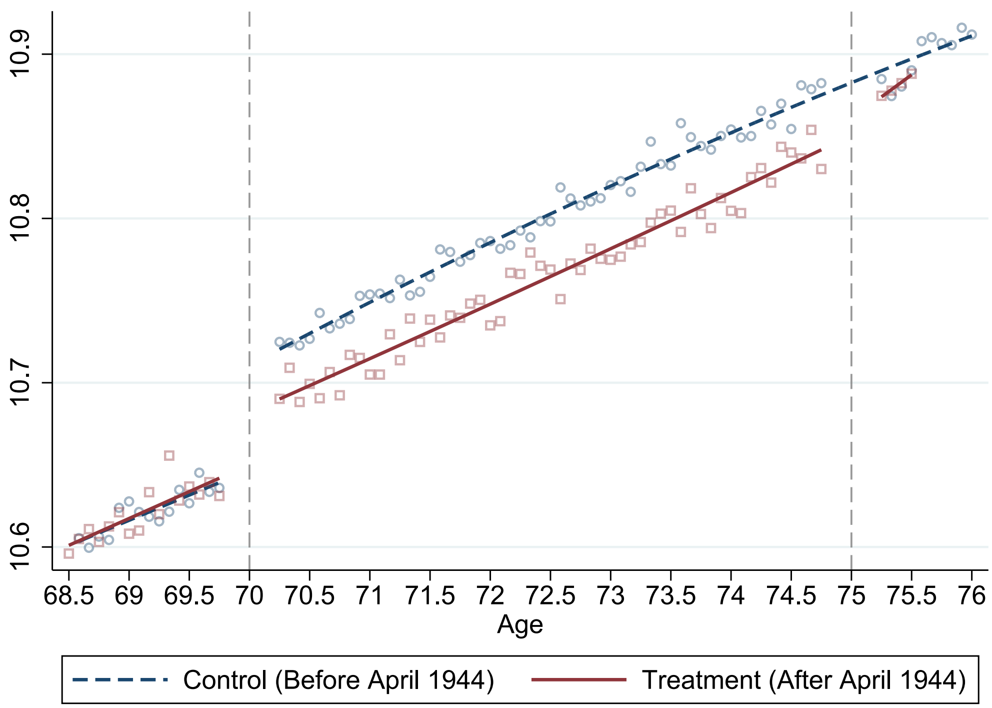
Notes: Each marker represents the RD estimate by year for post dummy in equation (**??**) with 95% confidence interval. For example, the marker for 3.5 years represents the RD estimate by using the data in September-November 2017. In each regression, we use 30 months both sides of the cutoff. We include a dummy that takes the value of one for those who are below the age of 70 for -1.5, -0.5, 0.5, and 1.5. The specification is linear in month and year of birth, interacted with post dummy, and fixed effects for month of birth. All RD estimates and standard errors are multiplied by 100 so that they can be interpreted as percent changes.

Figure 2.5: The Age Profiles of Total Number of Health Insurance Claims



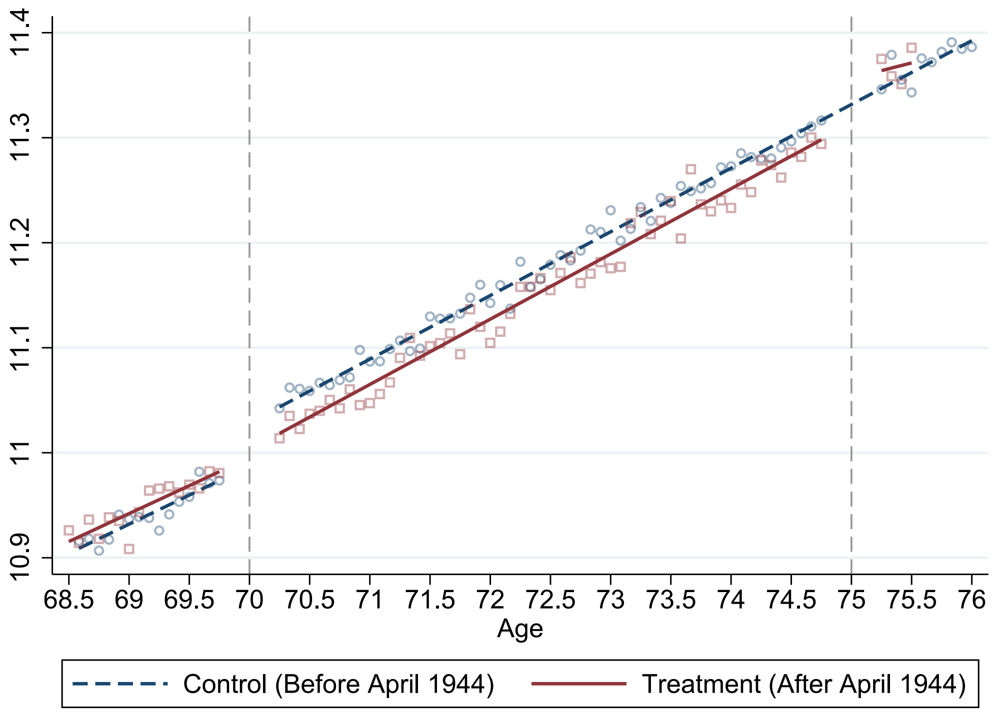
Notes: This figure plots the log of total number of health insurance claims per capita by age for treated group including those born after April 1944 (maroon solid curve) and non-treated group including those born before April 1944 (navy dash curve). The treated group includes those who were born between April 1942 and March 1944. Their coinsurance rate between 70-74 is 20 percent. The non-treated group includes those born between April 1944 and March 1946. Their coinsurance rate between 70-74 is 10 percent. Each marker represents the log of outcome excluding year specific effects and birth-month specific effects. These effects are identified by regressing the log of outcome on quadratic in age, year fixed effects, and birth-month fixed effects for the non-treated group. Since the data is the sum of three months (September to November) of the total number of claims, age in months in this figure represents age in October. Markers around age 70 and 75 are excluded.

Figure 2.6: The Age Profiles of Spending for Outpatients



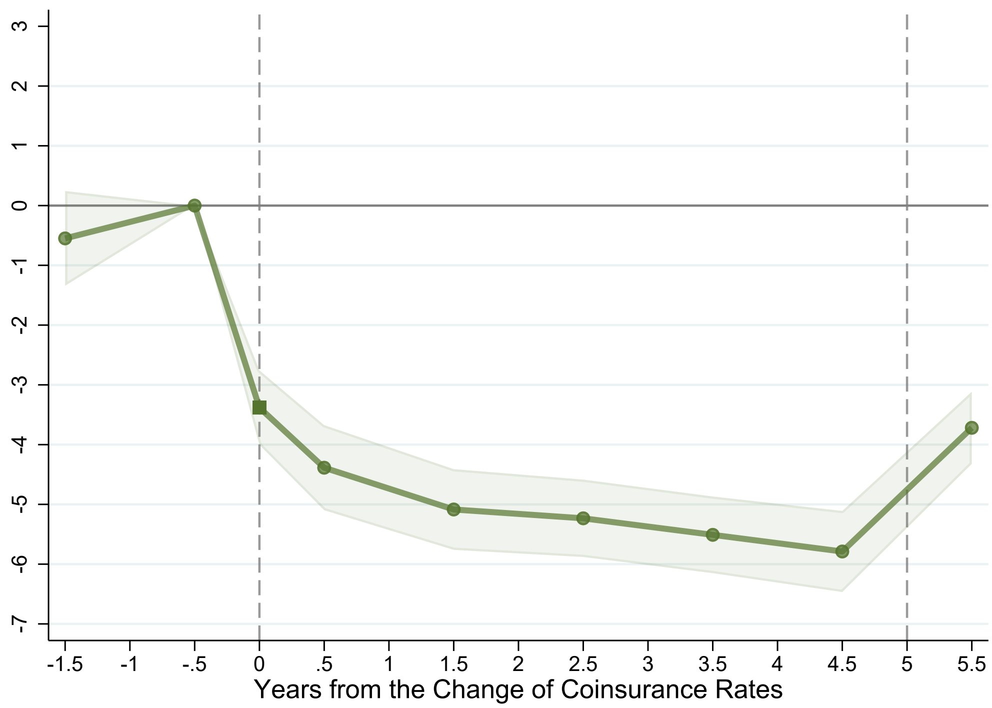
Notes: This figure plots the log of total number of health insurance claims per capita by age for treated group including those born after April 1944 (maroon solid curve) and non-treated group including those born before April 1944 (navy dash curve). The treated group includes those who were born between April 1942 and March 1944. Their coinsurance rate between 70-74 is 20 percent. The non-treated group includes those born between April 1944 and March 1946. Their coinsurance rate between 70-74 is 10 percent. Each marker represents the log of outcome excluding year specific effects and birth-month specific effects. These effects are identified by regressing the log of outcome on quadratic in age, year fixed effects, and birth-month fixed effects for the non-treated group. Since the data is the sum of three months (September to November) of the total number of claims, age in months in this figure represents age in October. Markers around age 70 and 75 are excluded.

Figure 2.7: The Age Profiles of Spending for Inpatients without Surgery



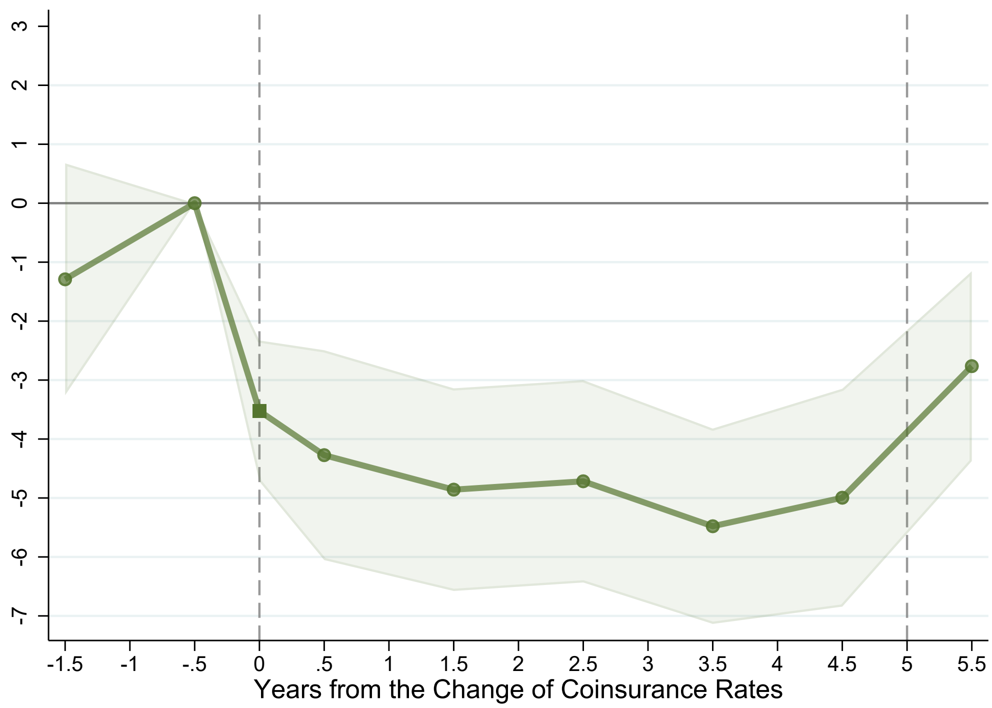
Notes: This figure plots the log of total number of health insurance claims per capita by age for treated group including those born after April 1944 (maroon solid curve) and non-treated group including those born before April 1944 (navy dash curve). The treated group includes those who were born between April 1942 and March 1944. Their coinsurance rate between 70-74 is 20 percent. The non-treated group includes those born between April 1944 and March 1946. Their coinsurance rate between 70-74 is 10 percent. Each marker represents the log of outcome excluding year specific effects and birth-month specific effects. These effects are identified by regressing the log of outcome on quadratic in age, year fixed effects, and birth-month fixed effects for the non-treated group. Since the data is the sum of three months (September to November) of the total number of claims, age in months in this figure represents age in October. Markers around age 70 and 75 are excluded.

Figure 2.8: Cohort Diff-in-disc Estimates by Year: Total Number of Claims



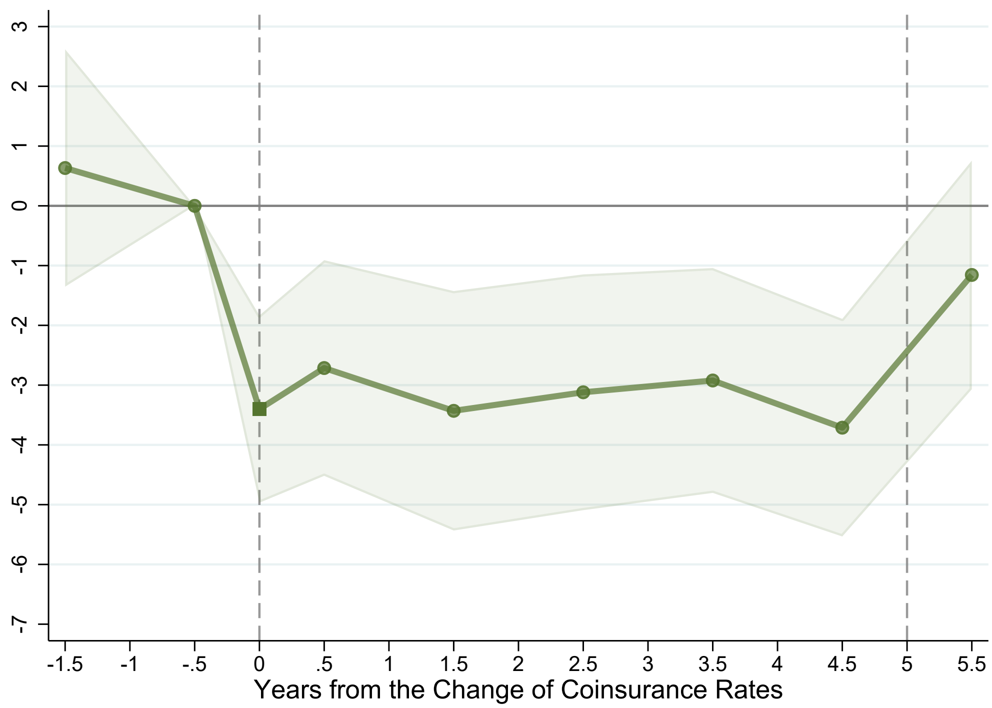
Notes: Each marker represents the RD estimate by year for post dummy in equation (**??**) with 95% confidence interval. For example, the marker for 3.5 years represents the RD estimate by using the data in September-November 2017. In each regression, we use 30 months both sides of the cutoff. We include a dummy that takes the value of one for those who are below the age of 70 for -1.5, -0.5, 0.5, and 1.5. The specification is linear in month and year of birth, interacted with post dummy, and fixed effects for month of birth. All RD estimates and standard errors are multiplied by 100 so that they can be interpreted as percent changes.

Figure 2.9: Cohort Diff-in-disc Estimates by Year: Spending for Outpatients



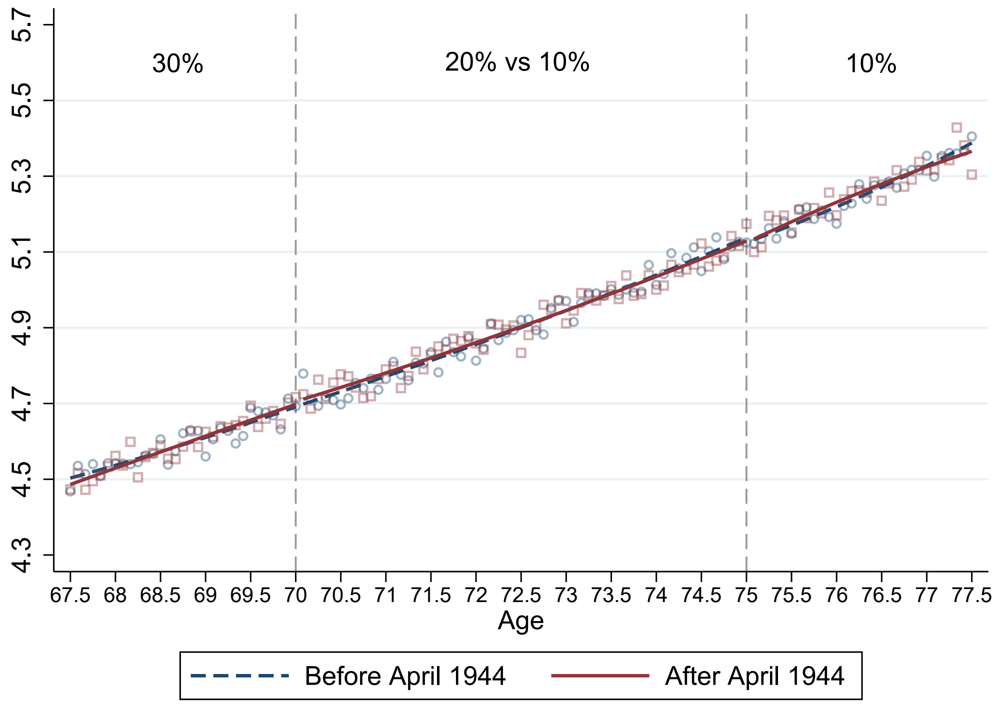
Notes: Each marker represents the RD estimate by year for post dummy in equation (**??**) with 95% confidence interval. For example, the marker for 3.5 years represents the RD estimate by using the data in September-November 2017. In each regression, we use 30 months both sides of the cutoff. We include a dummy that takes the value of one for those who are below the age of 70 for -1.5, -0.5, 0.5, and 1.5. The specification is linear in month and year of birth, interacted with post dummy, and fixed effects for month of birth. All RD estimates and standard errors are multiplied by 100 so that they can be interpreted as percent changes.

Figure 2.10: Cohort Diff-in-disc Estimates by Year: Spending for Inpatients without Surgery



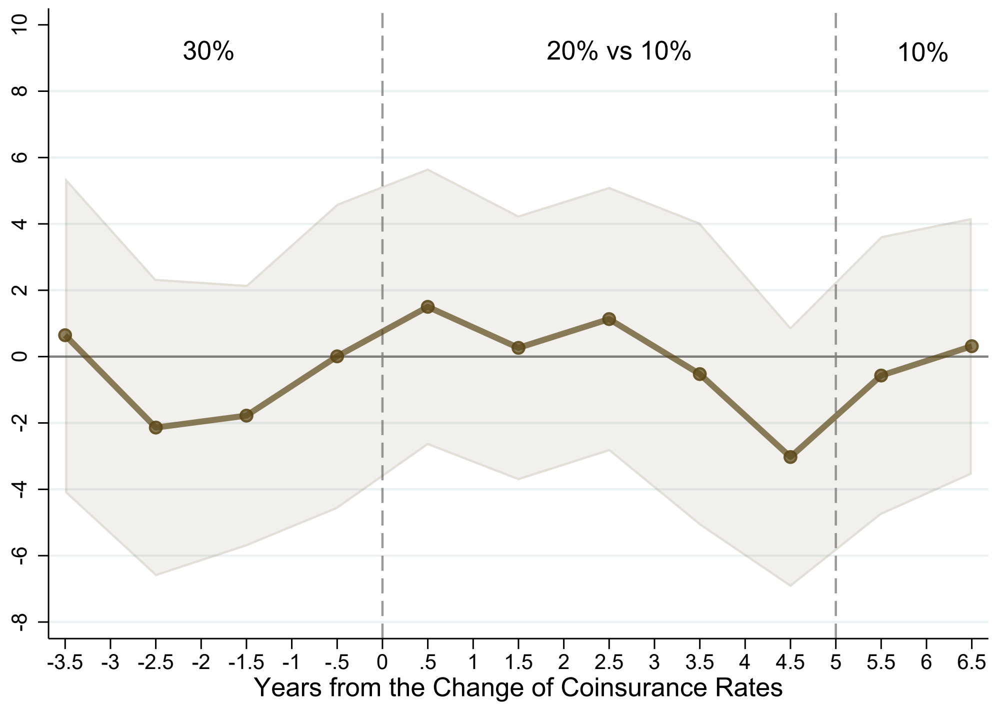
Notes: Each marker represents the RD estimate by year for post dummy in equation (**??**) with 95% confidence interval. For example, the marker for 3.5 years represents the RD estimate by using the data in September-November 2017. In each regression, we use 30 months both sides of the cutoff. We include a dummy that takes the value of one for those who are below the age of 70 for -1.5, -0.5, 0.5, and 1.5. The specification is linear in month and year of birth, interacted with post dummy, and fixed effects for month of birth. All RD estimates and standard errors are multiplied by 100 so that they can be interpreted as percent changes.

Figure 2.11: The Age Profiles of Mortality Rates



Notes: This figure plots the log of the total number of deaths per 100,000 people (outcome) by age for the treatment group (red and solid lines) and for the control group (blue and dashed lines) by using data from multiple time points. The treatment group includes those born between April 1944 and March 1945. Their coinsurance rate between ages 70-74 is 20%. The control group includes those born between April 1943 and March 1944. Their coinsurance rate between ages 70-74 is 10%. Each marker represents the residual of the outcome by excluding death-year fixed effects, death-month fixed effects, and birth-month fixed effects. These effects are identified by regressing the outcome on age, quadratic in age, death-year fixed effects, death-month fixed effects, and birth-month fixed effects for the control group. Lines represent quadratic fitted curves to the residuals for each group by three ranges of age depending on coinsurance, below 70, between 70 and 75, and above 75.

Figure 2.12: Cohort Diff-in-disc Estimates over Time: Mortality Rates



Notes: Each marker represents the point estimate by Cohort diff-in-disc at each point of time, from 3.5 years before to 6.5 years after the change of coinsurance, with 95% confidence interval. Specifically, we obtain each estimate by regressing equation (2.1) by using mortality data for 12 months, including before and after each point of time, and for the age of 69. For example, the marker at 3.5 years from the change of coinsurance represents the estimate by using data from April 2017 to March 2018. In each regression, we use 30 months both sides of the cutoff. We use a linear function of cohorts. All estimates and standard errors are multiplied by 100 so that they can be interpreted as percent changes.

Chapter 3

Welfare Implications of Social Norms - Marriage Market with Child Penalty -

Coauthored with: Kenzo Imamura (University of Tokyo).

3.1 Introduction

It is documented that there exists the gender asymmetry in roles within household. Men work outside the home, while women take care of children. Recent studies emphasize the importance of social norms as a factor contributing to this asymmetry. For example, research on the child penalty (CP) shows that while there is no significant difference in wage rates and labor supply between men and women until they have children, men's labor income does not change but women's labor income decreases significantly after they have children. Accumulating evidence suggests that this pattern cannot be explained by comparative advantages on labor and childbearing, but is due to social norms (Kleven et al. (2019a); Kleven et al. (2019b); Kleven et al. (2021); Andresen and Nix (2022); Kleven (2022)).

Becker (1973) argues that the marriage-market outcome should be analyzed in a matching model. Its key element is the joint surplus generated by each potential couple. From this view point, it is natural to suppose that social norms affect the marriage-market outcome by influencing the distribution of housework and childcare within a couple and thereby changing the joint surplus they generate.

This paper presents a plain-vanilla theoretical framework for analyzing the impact of social norms on the marriage-market outcome by using a matching model in which individuals have transferable utility (TU). Specifically, we make two modifications to the model of Chiappori et al. (2017), which analyzes how the burden of housework and childcare on the marriage-market outcome, based on empirical facts.

First, we introduce the role of social norms. In Chiappori et al. (2017), the size of the housework and childcare burden is determined by technology (e.g., household appliances), and its division between spouses is based on comparative advantages. There is no gender asymmetry in their model if wages are the same between men and women. We introduce social norms influencing the division of the housework and childcare burden. This is based on rich empirical evidence, including our own, that men have a dominant roles in work and women in housework and childcare, even when women have comparative advantages in work.

Second, while Chiappori et al. (2017) assume that the childcare burden is common for all individuals, we generalize and make it dependent on education. This is based on our empirical finding that the CP is larger for educated women.

By using the model, we conduct a comparative statics on how the existence of social norms affects matching patterns, the aggregated surplus in a whole economy, and the distribution of the surplus. The key points of the comparative statics results can be summarized as follows. Overall, social norms have a broader and more complex impact than can be summarized by the CP. Given the larger CP for high-skill women, social norms generate larger inefficiency when a high-skill woman matches with a low-skill man. If high-skill women have comparative advantages in work than low-skill men, it is efficient for the men to do childcare. Social norms however have women do so. This makes marriages less beneficial for high-skill women, preventing them from getting married, and get them leave from the marriage market. This decreases the aggregated surplus in an economy by vanishing marriages potentially generating the surplus. At the same time, the decreased competitiveness of high-skill women in the marriage market could benefit other participants, like high-skill men and low-skill women. Social norms cause the CP but the effect spreads out in an economy by the equilibrium effects of the marriage market.

The impact of social norms on the marriage-market outcome has not been analyzed extensively. An important exception is Bertrand et al. (2020). Their model assumes that the allocation of childcare time is determined by comparative advantages and also assumes that in any couple, a man has comparative advantages in labor and a woman has comparative advantage in childcare, so that wives are always responsible for childcare. Social norms are then modeled as a factor influencing the magnitude of the childcare burden. Namely, in their model, the size of childcare burden is determined by social norms, its allocation between spouses is determined by comparative advantages which men always have in labor. Their model demonstrates that strengthening social norms, given women's wage rates, has a negative impact on the marriage outcomes of educated women relative to less educated women, widening the skilled-unskilled marriage gap among women.

In contrast, our model lets social norms affect not only the size of the household childcare burden, but also the division of labor within household, rather than comparative advantages based on empirical facts. In this sense, our study is a generalization of Bertrand et al. (2020). Considering the effect of social norms on the division of labor deepens discussions how social norms affect matching patterns, the aggregated surplus, and its distribution across citizens. Furthermore, it helps broaden discussions about family policies. A prominent example is mandating and motivating low-skill men to take on housework and childcare. In Bertrand et al. (2020)'s model, since low-skill men have comparative advantages in labor against any women, the policy always exacerbates the aggregated surplus. In contrast, our model shows that under social norms, because the allocation of childcare time by couples in which women have comparative advantages in labor is inefficient, the policy increases the aggregated welfare and could be Pareto improving if appropriate re-distribution policies are taken.

The structure of this chapter is as follows. In Chapter 2, we present empirical facts on gender role asymmetries. Chapter 3 discusses a framework for analyzing the impact of social norms on marriage market outcomes through their impact on the burden of household chores, based on a matching model between individuals with TU. Chapter 4 presents a qualitative policy welfare analysis. Chapter 5 concludes this paper.

3.2 Empirical Facts

Main Data: The Japanese Panel Survey of Consumers

The data used primarily in this paper are from the Japanese panel survey of consumers (JPSC) from 1993 to 2019. It consists of five cohorts of women, with 1,500 women aged 24-34 as of 1993, 1,500 women aged 24-27 as of 1997 500 women aged 24-27 as of 1993, 836 women aged 24-28 as of 2003, 636 women aged 24-28 as of 2008, and 648 women aged 24-28 as of 2013. Males are observed only as spouses of females.

The Division of Labor within Household by Wife's Income

Previous studies indicate that social norms influence the division of labor between spouses within a household. First, there is abundant evidence that the role of a husband and wife is not determined solely by comparative advantage. Even when a wife obtains higher wages than her husband, she has a greater role in housework and childcare. Figure 3.1 plots the wife's share of income on the horizontal axis against the wife's share of the time for doing housework and childcare on the vertical axis. If comparative advantages are completely determined by income, the data should be observed around the 45-degree line. Although the fitted line is downward sloping, indicating higher share of income leads to lower share of time for doing housework and childcare, it is far above from the 45-degree line. Moreover, the slop is much shallow. This figure suggests that there exist some significant force to determine the division of labor within household, other than comparative advantages in earnings. Akerlof and Kranton (2000) show a similar figure in the U.S. and take it suggestive evidence of social norms.

Child Penalty

Recently, the clear differentiation of the roles of men and women in the household at the time of having children has attracted attention as a factor explaining the gender wage gap. The difference between genders in earnings after having children is the CP. Prior studies show that the CP is determined by social norms rather than comparative advantages. For example, the size of CP is associate with whether grand mother worked or not (Kleven et al. 2019a), with the measure of conservatives (Kleven et al. 2019b), with geographic culture (Kleven 2022), and with whether a partner is the same sex or not (Nix and Andersen 2021). One test for comparative advantages is to compare the CP between biological children and adopted children and Kleven et al. (2021) find no large difference. Other studies test whether the CP is related to the pre-marital difference in wages between spouses and cannot reject the null hypothesis that it does not influence the CP (Kleven et al. 2019a; Nix and Andersen 2021; Cortes and Pan 2023).

We also estimate the Japanese CP and then conduct a similar test for comparative advantages in wages. We use individuals who have their first child, who are observed both before and after having their first child, and who have responded at least five times in all. The basic estimating equation is as follows.

$$Y_{it}^{g} = \alpha^{\mathbf{g}}\mathbf{D}_{\mathbf{it}}^{\mathbf{Event}} + \beta^{\mathbf{g}}\mathbf{D}_{\mathbf{it}}^{\mathbf{Age}} + \gamma^{\mathbf{g}}\mathbf{D}_{\mathbf{it}}^{\mathbf{Year}} + \varepsilon_{it}^{g} \tag{3.1}$$

where Y_{it}^g is earnings of individual i whose gender is g at time t, $\mathbf{D_{it}^{Event}}$ is event time, $\mathbf{D_{it}^{Age}}$ are age dummies, $\mathbf{D_{it}^{Year}}$ are year dummies, and ε_{it}^g is error components.

The results are shown in Figure 3.2. Like other countries, while men's labor income does not change significantly before and after having their first child, women's labor income drops significantly and then recovers slowly. The CP after 10 years is about 65%, slightly larger than in Austria and Germany.

For the test of comparative advantages in wage, we interact the difference in pre-marital earnings with event dummies. We pool multiple years, from -6 to -4, from -3 to -1, from 0 to 2, from 3 to 5, from 6 to 8, from 9 to 11, and from 12 to 14. Similar to previous studies, the coefficients on comparative advantages in wages are largely insignificant both economically and statistically. Without any controls, our result suggests that an initial difference by $1,000,000$ JPY makes the CP smaller by around $30,000$ JPY to $80,000$ JPY, which are not statistically significant. With controls, the impact is much smaller. See Table 3.1 in details.

Child Penalty by Educational Attainment

Next, Figure 3.3 shows the CP by education level. College-graduated women are represented by solid lines and high school-graduated women by dashed lines. Immediately after having their first child, labor income falls by 60-70% for both types of women. The difference is however gradually widening, and 15 years after having their first child, the CP for high school-educated women is around 30-40%, while for college-educated women around 70%.

What are the reasons for the larger CP for high-skill women? Figures 3.4 and 3.5 examine the dynamics of the employment rate and the full-time rate before and after childbirth by education level. Because the data is somewhat noisy, we pool multiple years to compute the employment rate and the full-time rate. The figures show that their trends are almost parallel, suggesting that the difference in the CP between college- and high school-graduated women is determined by other factors.

Using the estimated CP by education level, Figure 3.6 compares the observed dynamics of labor income when having children with the counterfactual dynamics when there were no children by educational attainment. College-educated women are represented by solid lines and high school-educated women by dashed lines. This shows that while the labor income of the two groups spread out when they do not have children, there is almost no difference between the two groups when they have children. In other words, when they do not have children, the difference between the two groups gradually widens, but once they have children, the educated women become as if less educated and the difference disappears. In summary, while the impact of the employment rate and the full-time rate is similar, the cost of doing so is much higher for high-skill women.

Marriage Pattern

Table 3.2 shows the unmarried rate for 35-44 year olds by educational attainment using the 1990-2020 Census. The table shows that for women, the unmarriage rate is higher for those with higher education than for those with lower education, while for men, the unmarriage rate is higher for those with lower education than for those with higher education. Bertrand et al. (2021) focus on this skilled-unskilled marriage gap. They find that in 1995, the gap for women was negative in many countries, with less-educated women having lower marriage rates. By 2010, the gap had turned positive in North America and Northern Europe, while it remained negative in other countries, including East Asia. As discussed below, they present a theoretical model to explain this pattern.

3.3 Theoretical Framework

In this section, we develop a theoretical framework for analyzing the impact of social norms on marriage market outcome. Our model is closely related to Chiappori et al. (2017) who explore the impact of childcare on the marriage market outcome. Their model consists of two stages. In the first stage, individuals enter the marriage market. The marriage market determines who marries whom. In the second stage, households solve the utility maximization problem. A household allocates childcare time, which is common for all individuals and is exogenously given for a married couple in their budget constraint in the model, hence representing the opportunity cost of children, based on comparative advantages. Namely, a spouse with lower wage does childcare, and reduces hours worked.

We modify their model in two ways based on empirical facts. First, we introduce the role of social norms determining the allocation of childcare time between spouses. To examine welfare implications of social norms, we compare marriage-market outcome between the world where the allocation is determined by comparative advantages and social norms. Second, we allow heterogeniety of the opportunity cost of children.

Second Stage: Household Problem

When married, a household is endowed with a child. A household obtains utility from husband's private consumption q_m, wife's private consumption q_f, and public consumption, Q, interpreted as expenditure for a child. A necessity and sufficient condition to be TU is that utility function belongs to generalized quasi-linear form. Like Chiappori et al. (2017), we use Cobb-Douglas utility function, $(q_m + q_f)Q^{\alpha}$ where $0 < \alpha \leq 1$.

Wage rate is $e_g W > 1$ where $e_g = L = 1$ for low-skill workers and $e_g = H > 1$ for high-skill workers. That is, H represents skill premium. All individuals are endowed with total time 1. Having a child reduces earnings by $1 - \delta_g(e_g)$. Then, $-\delta_g(e_g)e_g W$ represents the cost of having a child, and a difference between men and women is CP. The ultimate source of this cost is childcare time. The cost includes not only a direct impact on earnings due to the decrease in hours worked, but also skill deprecation. Notice that CP depends on the level of skill e_g.

The utility maximization problem of married couples are

$$\max_{q_m+q_f,Q} \ (q_m+q_f)Q^{\alpha} \quad \text{s.t.} \ \ (q_m+q_f)+Q=\Big\{1-\delta_m(e_m)\Big\}e_mW+\Big\{1-\delta_f(e_f)\Big\}e_fW. \tag{3.2}$$

Here, solutions for (q_m+q_f) and Q are as follows.

$$(q_m+q_f)^* = \frac{1}{1+\alpha}\cdot\Big[\{1-\delta_m(e_m)\}e_mW+\{1-\delta_f(e_f)\}e_fW\Big] \tag{3.3}$$

$$Q^* = \frac{\alpha}{1+\alpha}\cdot\Big[\{1-\delta_m(e_m)\}e_mW+\{1-\delta_f(e_f)\}e_fW\Big] \tag{3.4}$$

By substituting these solutions into the household utility function, we obtain joint utility:

$$G(e_m,e_f)=K\Big[\{1-\delta_m(e_m)\}e_mW+\{1-\delta_f(e_f)\}e_fW\Big]^{1+\alpha} \tag{3.5}$$

where $K=\alpha^{\alpha}/(1+\alpha)^{1+\alpha}$.

If single, individuals simply enjoy their private consumption q_g. They use their earnings for q_g and allocate all of their time to labor. Thus, utility being single is the following.

$$G(e_g) = e_g W \tag{3.6}$$

By subtracting husband's utility being single and wife's utility being single from couple's joint utility, we obtain couple's joint surplus:

$$S(e_m, e_f) = G(e_m, e_f) - G(e_m) - G(e_f). \tag{3.7}$$

Comparative Advantage vs Social Norms

Social norms affect an allocation of the cost of having a child between spouses: $\delta_m(e_m)$ and $\delta_f(e_f)$. If there is no social norms, we assume that $\delta_m(e_m)$ and $\delta_f(e_f)$ are determined by comparative advantages. If a husband and wife have the same level of education, then $\delta_m(e_m) = \delta_f(e_f) = \delta(e)/2$. If $e_m = H > e_f = L$, it is optimal for a household to set $\delta_m = 0$ and $\delta_f = \delta(L)$. If $e_m = L < e_f = H$, then $\delta_m = \delta(L)$ and $\delta_f = 0$. In contrast, if a society is under social norms, a wife takes care of her child, irrespective of her own and husband's education. Namely, $\delta_m(e_m) = 0$ and $\delta_f(e_f) \neq 0$.

To focus on the relative size of childcare cost between skills, we assume that $\delta_m(L) = \delta_f(L) = 0$; there is no CP for low-skill workers. High-skill workers experience the drop of earnings by $\delta_m(H) = \delta_f(H) = \delta$ if they do all of the childcare. We also assume $0 < \delta < 1 - 1/H$. This assumption means that at the maximum of δ, high-skilled workers are not distinguishable form low-skill workers in terms of wages when married.

With the setting above, we explain joint surplus by the combination of husband's skill and wife's skill both under comparative advantages and social norms. First, joint surplus under comparative advantages are as follows.

$$S(H,H) = K\Big\{(2-\delta)HW\Big\}^{1+\alpha} - 2HW \tag{3.8}$$

$$S(H,L) = K\Big\{(1+H)W\Big\}^{1+\alpha} - (1+H)W \tag{3.9}$$

$$\overline{S(L,H)} = S(H,L) \tag{3.10}$$

$$S(L,L) = K\Big\{2W\Big\}^{1+\alpha} - 2W \tag{3.11}$$

Here, $S(H,H)$ is a decreasing function of δ. Others do not depend on δ and are always positive. Note that joint surplus generated by low-skill men and high-skill women uses a slightly different notation, $\overline{S(L,H)}$. We will explain the meaning later.

In a society under social norms, joint surplus generated by low-skill men and high-skill is different from the world under comparative advantages. Namely, $\underline{S(L,H)} \neq \overline{S(L,H)}$. In particular, women take care of their child even if they are high skilled, its surplus becomes

$$\underline{S(L,H)} = K\Big[\big\{1+(1-\delta)H\big\}W\Big]^{1+\alpha} - (1+H)W. \tag{3.12}$$

This is also decreasing in δ. Thus, if $\delta > 0$, $\underline{S(L,H)} < \overline{S(L,H)}$.

First Stage: Marriage Market

We next consider matching in the marriage market. There are three main cases depending on the value of δ as summerized in Table 3.3. First, if δ is sufficiently small, the joint surplus is super-modular regardless of the existence of social norms. Second, if δ gets larger, the joint surplus becomes sub-modular if the social norm does not exist, but remains super-modular if it does exist. While only the joint surplus of high-skilled men and high-skilled women will fall if social norms do not exist, the joint surplus of low-skilled men and high-skilled women will also fall if social norms exist. Social norms tend to make joint surplus super-modular compared to comparative advantages. Thirdly, when δ is sufficiently large, the joint surplus of marriages involving high-skilled women may become negative. Marriages are not attractive for high-skill women, and they do not join the marriage market.

Case 1: super-modular both under CA and under SN

If δ is sufficiently small, the joint surplus is super-modular regardless of the existence of social norms. In this case, positive assortative matching is realized; high-skilled men and high-skilled women marry, and low-skilled men and low-skilled women marry. The aggregated surplus is $S(H,H)+S(L,L)$, the same for both worlds, CA and SN.

Each individual's share of the joint surplus is as follows.

$$\begin{aligned}
&\tilde{u}_L \in \Big[0, \quad S(L,L)\Big], \\
&\tilde{u}_H \in \Big[S(H,L)-S(L,L), \quad S(H,H)-\big\{S(L,H)-S(L,L)\big\}\Big], \\
&\tilde{v}_L \in \Big[0, \quad S(L,L)\Big], \quad \text{and} \\
&\tilde{v}_H \in \Big[S(L,H)-S(L,L), \quad S(H,H)-\big\{S(H,L)-S(L,L)\big\}\Big]
\end{aligned}$$

where $\tilde{u}_L$ is the share for a low-skill man, $\tilde{u}_H$ is the share for a high-skill man, $\tilde{v}_L$ is the share for a low-skill woman, and $\tilde{v}_H$ is the share for a high-skill woman.

Each person's utility is this surplus share plus his or her utility when unmarried.

$$
\begin{aligned}
u_L &\in \Big[W, \quad S(L,L)+W\Big], \\
u_H &\in \Big[S(H,L)-S(L,L)+HW, \quad S(H,H)-\big\{S(L,H)-S(L,L)\big\}+HW\Big], \\
v_L &\in \Big[W, \quad S(L,L)+W\Big], \text{ and} \\
v_H &\in \Big[S(L,H)-S(L,L)+HW, \quad S(H,H)-\big\{S(H,L)-S(L,L)\big\}+HW\Big]
\end{aligned}
$$

where u_L is the share for a low-skill man, u_H is the share for a high-skill man, v_L is the share for a low-skill woman, and v_H is the share for a high-skill woman.

Importantly, only the surplus generated by a couple of a low-skilled man and high-skilled woman differs between comparative advantages and under social norms. Under comparative advantages, $S(H,L) = \overline{S(L,H)}$, while under social norms, $S(H,L) > \underline{S(L,H)}$. Now $S(L,H)$ is in the upper bound of $\tilde{u}_H$ and u_H and the lower bound of $\tilde{v}_H$ and v_H. Thus, compared to under comparative advantage, social norms make the allocation desirable for high-skilled men and undesirable for high-skilled women.

Case 2: sub-modular and increasing under CA, but super-modular under SN

Case 2-1: $S(H,H) > S(H,L) = S(L,H)$

As δ increases from Case 1, it is possible that the joint surplus is sub-modular if the social norm does not exist, but if it does, it remains super-modular as in Case 1. In the former case, matching becomes negative assortative. The aggregated surplus is $S(H,L) + S(L,H) = S(H,L) + \overline{S(L,H)}$. Note that they are no longer a function of δ. In the latter case, matching remains positive assortative and the aggregated surplus is $S(H,H) + S(L,L)$.

Each individual's share of the joint surplus is as follows.

$$\tilde{u}_L \in \Big[0, \quad S(L,H) - \big\{S(H,H) - S(H,L)\big\}\Big],$$
$$\tilde{u}_H \in \Big[S(H,H) - S(L,H), \quad S(H,L)\Big],$$
$$\tilde{v}_L \in \Big[0, \quad S(H,L) - \big\{S(H,H) - S(L,H)\big\}\Big], \text{ and}$$
$$\tilde{v}_H \in \Big[S(H,H) - S(H,L), \quad S(L,H)\Big]$$

Each person's utility is this surplus share plus his or her utility when unmarried.

$$u_L \in \Big[W, \quad S(L,H) - \big\{S(H,H) - S(H,L)\big\} + W\Big],$$
$$u_H \in \Big[S(H,H) - S(L,H) + HW, \quad S(H,L) + HW\Big],$$
$$v_L \in \Big[W, \quad S(H,L) - \big\{S(H,H) - S(L,H)\big\} + W\Big], \text{ and}$$
$$v_H \in \Big[S(H,H) - S(H,L) + HW, \quad S(L,H) + HW\Big]$$

Case 2-2: $S(H,H) < S(H,L) = S(L,H)$

Furthermore, as δ increases, the joint surplus is no longer increasing under comparative advantages. In other words, the joint surplus is larger when a man is matched with a low-skilled woman than with a high-skill woman. In this case, the matching pattern is the same as in Case 2-1, i.e., negative assortative matching.

Each individual's share of the joint surplus is as follows.

$$\tilde{u}_L \in \Big[0, \quad S(L,H)\Big],$$
$$\tilde{u}_H \in \Big[0, \quad S(H,L)\Big],$$
$$\tilde{v}_L \in \Big[0, \quad S(H,L)\Big], \quad \text{and}$$
$$\tilde{v}_H \in \Big[0, \quad S(L,H)\Big]$$

Each person's utility is this surplus share plus his or her utility when unmarried.

$$u_L \in \Big[W, \quad S(L,H) + W\Big],$$
$$u_H \in \Big[HW, \quad S(H,L) + HW\Big],$$
$$v_L \in \Big[W, \quad S(H,L) + W\Big], \quad \text{and}$$
$$v_H \in \Big[HW, \quad S(L,H) + HW\Big]$$

Case 3-1: $\underline{S(L,H)} < 0$ **and** $S(H,H) + S(L,L) < S(H,L)$

As δ further increases, the joint surplus becomes $\underline{S(L,H)} < 0$ and $S(H,H)+S(L,L) < S(H,L)$. Note that $S(H,H) < S(H,L)$ is a necessary condition for Case 3-1 to hold. This does not affect the world under comparative advantages; in TU matching is determined to maximize the aggregated surplus. Therefore, under social norms where this condition holds, low-skilled men and high-skilled women who generate a negative joint surplus will not be matched, but only high-skilled men and low-skilled women will be matched. Note that unmarried persons appear in this economy for the first time. In this case, the aggregated surplus generated in the marriage market is $S(H,L)$.

Each individual's share of the joint surplus is as follows.

$$\tilde{u}_L = 0,$$
$$\tilde{u}_H \in \Big[S(H,H), \quad S(H,L) - S(L,L)\Big],$$
$$\tilde{v}_L \in \Big[S(L,L), \quad S(H,L) - S(H,H)\Big], \text{ and}$$
$$\tilde{v}_H = 0$$

Each person's utility is this surplus share plus his or her utility when unmarried.

$$u_L = W,$$
$$u_H \in \Big[S(H,H) + HW, \quad S(H,L) - S(L,L) + HW\Big],$$
$$v_L \in \Big[S(L,L) + W, \quad S(H,L) - S(H,H) + W\Big], \text{ and}$$
$$v_H = HW$$

Case 3-2: $\underline{S(L,H)} < 0$ **and** $S(H,H) < 0$

Finally, as δ is very large, $S(H,H)$ also becomes negative. In this case, high-skilled women can never generate a positive joint surplus in the marriage market.

Each individual's share of the joint surplus is as follows.

$$
\begin{aligned}
\tilde{u}_L &= 0, \\
\tilde{u}_H &\in \Big[0, \quad S(H,L) - S(L,L)\Big], \\
\tilde{v}_L &\in \Big[S(L,L), \quad S(H,L)\Big], \quad \text{and} \\
\tilde{v}_H &= 0
\end{aligned}
$$

Each person's utility is this surplus share plus his or her utility when unmarried.

$$
\begin{aligned}
u_L &= W, \\
u_H &\in \Big[HW, \quad S(H,L) - S(L,L) + HW\Big], \\
v_L &\in \Big[S(L,L) + W, \quad S(H,L) + W\Big], \quad \text{and} \\
v_H &= HW
\end{aligned}
$$

Mapping of Previous Studies

Chiappori et al. (2017) and Bertrand et al. (2021) assume that the division of the housework and childcare burden among couples is determined by comparative advantage and that men always have a comparative advantage in labor compared to women.

Consider the case in which the wages of low-skilled men are higher than those of high-skilled women. In this case, for all couples, it is optimal for women to always be in charge of housework and childcare. That is, $\delta_m(e_m) = 0$ for all e_m. Therefore, the allocation of childcare burden is similar to that under the social norms when men and women have equal wages given their skills. That is, when δ is sufficiently large and in Case 3, high-skilled women leave the marriage market and only high-skilled men and low-skilled women are matched. In terms of welfare, low-skilled women gain while the rest lose. In this setting, even if we impose the constraint by social norms that women should do housework and childcare, nothing will happen. Their works implicitly assume that social norms have no effect. Therefore, in order to analyze social norms, it is necessary to move away from this assumption.

There is also an important difference between the two. In a world where men always have a comparative advantage in labor, it is efficient that high-skilled women matched with low-skilled men are responsible for housework and childcare, whereas it is inefficient to do so in a world of social norms where men and women are paid equally. In the latter case, if the constraint of social norms is removed, the aggregated surplus increases as low-skilled men take on housework and childcare.

Summary of Aggregated Surplus

Table 3.4 summarizes predictions on stable matching and the aggregated surplus by the existence of social norms and the size of δ. Only **SHH** depends on the size of δ. In Case 1, the matching pattern and the aggregated surplus are the same between comparative advantages and social norms. The high-skill man matches the high-skill woman, the low matches the low, generating the aggregated surplus, **SHH**+SLL. As δ increases but remains in Case 1, the aggregated surplus decreases as **SHH** drops.

In Case 2, stable matching and the aggregated surplus are diverged between comparative advantages and social norms. In comparative advantages, the joint surplus becomes sub-modular and the matching turns to negative assortative matching. Importantly, the aggregated surplus no longer depends on δ under comparative advantages, while remain relying on δ under social norms. Thus, as δ increases in Case 2, the aggregated surplus does not change under comparative advantages but keep decreasing under social norms.

Finally, in Case 3, although stable matching and the aggregated surplus remain the same under comparative advantages, there is only one matching between a high-skill man and a low-skill women under social norms. The aggregated surplus shrinks to SHL. Since it does not depend on δ, the aggregated surplus does not change even δ further increases.

Summary of Share of Joint Surplus

Table 3.5 summarizes predictions on the male's share of the joint surplus, and Table 3.6 on the female's share of the joint surplus. Comparative statistics with respect to the existence of social norms and δ reveals several interesting implications. Each share of the joint surplus is represented as a set, because there are multiple stable equiliblia. Therefore, in comparative statistics we compare sets. Here, the order of sets is defined as follows. For any set $S, S' \subseteq$ **R**, we define $S \geq_s S'$ if for any $x \in S$ and $x' \in S'$, max$\{x, x'\} \in S$ and min$\{x, x'\} \in S'$ hold. Let $U_k^{i,j}$ where $i \in \{CA, SN\}$, $j \in \{1, 2, 3\}$, and $k \in \{H, L\}$ be the gain set of men with skill k in Case j in situation i. $V_k^{i,j}$ is defined similarly as the gain set of women.

Welfare of Low-skill Workers

In terms of the welfare of low-skill workers, the following relationships hold.

$$U_L^{CA,j} \geq_S U_L^{SN,j} \text{ for all } j,\ V_L^{CA,j} \geq_S V_L^{SN,j} \text{ for } j \in \{1, 2\}. \text{ and } V_L^{SN,3} \geq_S V_L^{CA,3}.$$

In Case 1, the share of joint surplus under comparative advantages and under the social norms are the same. In contrast, in Case 2, the share under comparative advantages is larger for both men and women. For Case 3, low-skill men have higher share under comparative advantages, while the share can be higher for women under social norms. This is because social norms make high-skilled women less attractive, and men no longer compete for high-skilled women, but instead compete for low-skilled women.

Welfare of High-skill Workers

In terms of the welfare of high-skill workers, the following relationships hold.

$U_H^{CA,j} = V_H^{CA,j}$ for all j. $U_H^{SN,j} \geq_S V_H^{SN,j}$ for all j. Moreover, $U_H^{SN,j} - V_H^{SN,j}$ is increasing in $j \in \{1, 2, 3\}$.

Under comparative advantages, there is a gender gap in welfare. Under social norms, male welfare is higher regardless of δ, and moreover, the gender gap is increasing in δ.

Second, the following relationships hold as well.

$U_H^{SN,j} \geq_S U_H^{CA,j}$ for $j \in \{1, 2\}$. $V_H^{CA,j} \geq_S V_H^{SN,j}$ for $j \in \{1, 2\}$.

In Case 1, since $\overline{\text{SLH}} > \underline{\textbf{SLH}}$, the upper bound for men is higher and the lower bound for women is lower under social norms than under comparative advantage. Next, consider Case 2. First, men have higher utility under social norms. The lower bound is at least as high under social norms because $\textbf{SHH} + \text{SLL} \leq \text{SHL} + \overline{\text{SLH}}$ for Case 2-1. Case 2-2 shows that the share of surplus under comparative advantage is zero under the comparative advantage, while it is non-negative under the social norm. The upper bound is strictly higher under the social norm. This is because $\text{SHL} + \underline{\textbf{SLH}} < \textbf{SHH} + \text{SLL}$.

Second, women in Case 2 have lower utility under social norms, except for the lower bound in Case 2-2. First, we can say that the lower bound of Case 2-1 is $\textbf{SHH} + \text{SLL} \geq \text{SHL} + \underline{\textbf{SLH}}$. Case 2-2, on the other hand, is higher under social norms. This reflects the fact that the joint surplus is decreasing from the perspective of high-skilled men, and thus competition for high-skilled women is no longer occurring. Considering the upper bound, we can say that in Case 2-1 and Case 2-2, $\textbf{SHH} + \text{SLL} \leq \text{SHL} + \text{SLH}$.

Finally, the following relationships hold as well.

$U_H^{CA,3} \geq_S U_H^{SN,3}$ and $V_H^{CA,3} \geq_S V_H^{SN,3}$.

This is clear from Tables 3.5 and 3.6 except for the lower bound for men in Case 3-1, which is higher under the social norm; in Case 3-1, high-skilled women can generate a positive joint surplus when matched with high-skilled men and thus remain in the marriage market, unlike Case 3- Unlike Case 2, the lower bound for high-skilled men is higher because high-skilled women remain in the marriage market if they are matched with a high-skilled man.

3.4 Policy Analysis

Family policies can affect the marriage-market outcome because they affect the size and division of childcare burden between spouses. Our theoretical framework provides predictions for how a family policy affects matching, the aggregated surplus, and the distribution of joint surplus. In this section, we analyze two policies: free childcare services and mandatory parental leave for low-skill men.

Impact of Child Care Services

Consider the effect of providing free childcare services without imposing tax. This policy can be interpreted as a decrease in δ in the theoretical framework.

Consider an economy initially in Case 3 under social norms. The policy moves it to Case 2. High-skill women return to the marriage market and are matched with high-skill men, and less educated men are matched with less educated women. The total surplus rises. The share of the surplus improves, except for the less-educated women. The welfare of low-educated women decreases because they lose the effect of increased competitiveness due to the exit of high-skill women from the marriage market. Therefore, the policy is not a Pareto improvement. However, since the aggregate surplus is rising, it could be a Pareto improvement if combined with an appropriate redistribution policy.

If the economy is in Case 2 under the social norm, moving to Case 1 does not change matching, but increases the aggregated surplus. It does not affect the share of joint surplus for less-educated men and women, but it raises the share for high-skill men and women. Note that for educated men, the increase in **SHH**, the joint surplus generated by marriages between educated men and women, due to small δ is partially offset by the decrease in SLH, the surplus generated by less-educated men and more-educated women, also due to small δ. The latter is the effect of the loss of the effect of the increased bargaining power of educated men due to the reduced competitiveness of educated women in the marriage market as a result of the large δ.

Impact of mandatory parental leave for low-skill men

Effects under comparative advantage

Let the value of δ be given and consider the case where the division of childcare burden is determined by comparative advantage. In this case, mandatory parental leave for low-skill men does not increase the aggregated surplus in any matches, because under TU households determine the allocation of childcare burden so as to maximize their utility, and matches are determined in the marriage market so as to maximize the aggregated surplus. For example, if men always have a comparative advantage in labor, the policy reduces the surplus associated with low-skill men because it would allocate time being opposite to their comparative advantage. Thus, the setting in Chiappori et al. (2017) and Bertrand et al. (2021) predict that this policy never becomes Pareto improving.

Impact under Social Norms

In contrast, if social norms dictate the division of childcare burden between spouses, mandatory parental leave for low-skill men, together with a re-distributive policy, could be Pareto improving. Consider a situation in which men and women have equal wages. Consider a policy that mandates housework and childcare for low-educated men. As an extreme example, assume that less-educated men are responsible for all housework and childcare when they get married.

If the economy is in a Case 3 under social norms, this policy would shift the economy to Case 3 under comparative advantages. A less-educated man and a educated woman form a couple, and the aggregated surplus rises. The share of the joint surplus improves except for the less-educated women. The less-educated women lose because they lose the effect of increased competitiveness due to the exit of the more-educated women from the marriage market. However, as noted earlier, since aggregate surplus is rising, this could be a Pareto improvement if combined with appropriate re-distributive policies.

If the economy is Case 2 under social norms, matching changes from positive assortative matching to negative assortative matching and the aggregated surplus rises. The share of joint surplus improves except for educated men. The educated men lose because they lose the effect of their own bargaining power improvement due to the lower attractiveness of educated women due to social norms. Similarly, this case can be Pareto-improved by appropriate re-distributive policies.

3.5 Conclusion

Social norms affect the division of labor within household. They have men work and women do childcare, irrespective of comparative advantages. We develop a theoretical framework to analyze the impact of social norms on the marriage-market outcome through changing the division of labor. The framework describes how the existence of social norms and the size of childcare burden affect who marries whom and the welfare of individuals.

Social norms generally decrease the aggregated surplus in an economy. Welfare implications of social norms are complex; interestingly, some of participants in the marriage market could improve their welfare by social norms. The framework provides interesting predictions about family policies, including free childcare services and mandatory parental leave for low-skill men. The framework is plain vanilla in a sense that the size of child-care burden and its allocation between spouses are simple totally exogenous. Endogenizing these aspects in the model should be important future work. This would produce richer discussions about social norms and relevant policies.

3.6 Table

Table 3.1: Test of Comparative Advantages

	-3 to -1	0 to 2	3 to 5	6 to 8	9 to 11	12 to 14
No Controls						
Base	-46.6	-208.7	-201.3	-208.1	-221.0	-230.0
	(9.3)	(9.9)	(11.2)	(12.6)	(15.0)	(17.9)
Comparative	.056	-.036	.042	.076	.071	.080
Advantages	(.054)	(.048)	(.048)	(.048)	(.050)	(.054)
Rich Controls						
Base	-35.6	-188.6	-170.6	-173.5	-167.5	-168.5
	(10.7)	(11.8)	(13.7)	(15.0)	(17.7)	(21.7)
Comparative	.048	-.066	-.020	.002	-.009	.004
Advantages	(.032)	(.033)	(.033)	(.033)	(.035)	(.040)

Table 3.2: Fraction of Unmarried between Ages 35-44

	Male			Female		
Year (Cohort)	HSG	AS	CLG	HSG	AS	CLG
1990 (C-1950)	14.0%	13.1%	10.5%	5.5%	8.0%	10.5%
2000 (C-1960)	23.3%	21.2%	17.8%	9.5%	11.6%	14.8%
2010 (C-1970)	33.4%	30.0%	24.8%	17.8%	19.2%	23.0%
2020 (C-1980)	37.7%	31.0%	26.6%	22.8%	20.5%	22.8%

Table 3.3: Cases by Social Norms and CP

	Comparative Advantages	Social Norms
Case 1	$\textbf{SHH}+\text{SLL}>\text{SHL}+\overline{\text{SLH}}$	$\textbf{SHH}+\text{SLL}>\text{SHL}+\underline{\textbf{SLH}}$
Case 2-1	$\textbf{SHH}+\text{SLL}<\text{SHL}+\overline{\text{SLH}}$	$\textbf{SHH}+\text{SLL}>\text{SHL}+\underline{\textbf{SLH}}$
Case 2-2	$\textbf{SHH}+\text{SLL}<\text{SHL}+\overline{\text{SLH}}$	$\textbf{SHH}+\text{SLL}>\text{SHL}+\underline{\textbf{SLH}}$
Case 3-1	$\textbf{SHH}+\text{SLL}<\text{SHL}+\overline{\text{SLH}}$	$\underline{\textbf{SLH}}<0$ & $\textbf{SHH}+\text{SLL}<\text{SHL}$
Case 3-2	$\textbf{SHH}+\text{SLL}<\text{SHL}+\overline{\text{SLH}}$	$\underline{\textbf{SLH}}<0$ & $\textbf{SHH}<0$ & $\text{SLL}<\text{SHL}$

Table 3.4: Aggregated Surplus

	Comparative Advantages	Social Norms
Case 1	**SHH**+SLL	**SHH**+SLL
Case 2-1	SHL+$\overline{\text{SLH}}$	**SHH**+SLL
Case 2-2	SHL+$\overline{\text{SLH}}$	**SHH**+SLL
Case 3-1	SHL+$\overline{\text{SLH}}$	SHL
Case 3-2	SHL+$\overline{\text{SLH}}$	SHL

Table 3.5: Male Share of Joint Surplus

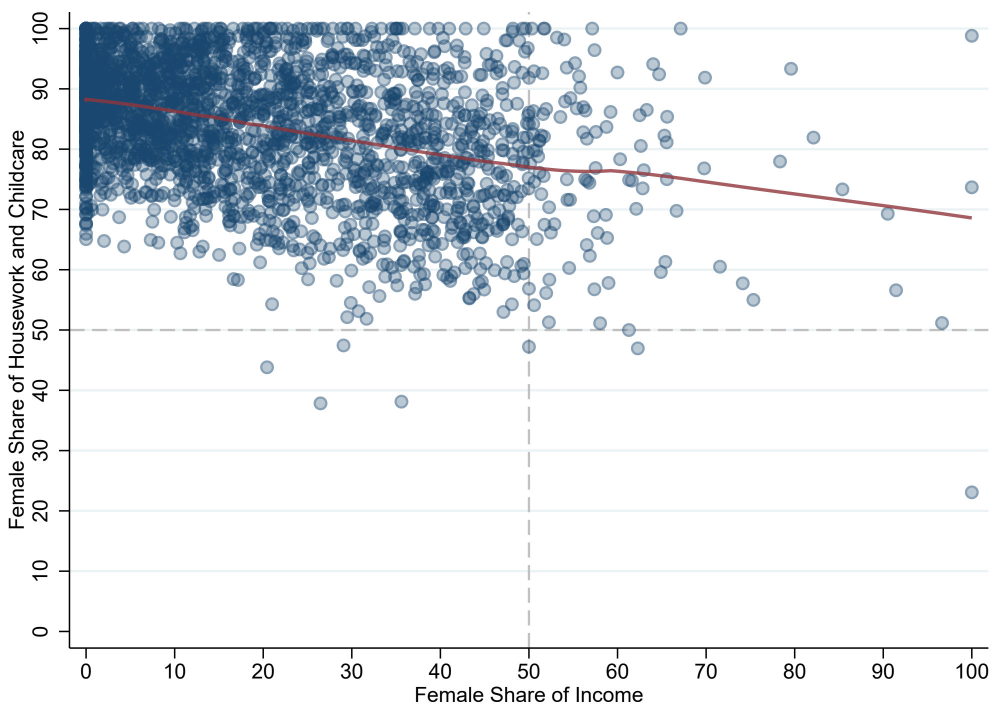
	Low-skill Men		High-skill Men	
	Comparative Advantages	Social Norms	Comparative Advantages	Social Norms
Case 1	0 - SLL	0 - SLL	SHL-SLL - **SHH**-($\overline{\text{SLH}}$-SLL)	SHL-SLL - **SHH**-($\underline{\textbf{SLH}}$-SLL)
Case 2-1	0 - $\overline{\text{SLH}}$ - (**SHH**-SHL)	0 - SLL	**SHH**-$\overline{\text{SLH}}$ - SHL	SHL-SLL - **SHH**-($\underline{\textbf{SLH}}$-SLL)
Case 2-2	0 - $\overline{\text{SLH}}$	0 - SLL	0 - SHL	SHL-SLL - **SHH**-($\underline{\textbf{SLH}}$-SLL)
Case 3-1	0 - $\overline{\text{SLH}}$	0 - 0	0 - SHL	**SHH** - SHL -SLL
Case 3-2	0 - $\overline{\text{SLH}}$	0 - 0	0 - SHL	0 - SHL -SLL

Table 3.6: Female Share of Joint Surplus

	Low-skill Women		High-skill Women	
	Comparative Advantages	Social Norms	Comparative Advantages	Social Norms
Case 1	0 - SLL	0 - SLL	$\overline{\text{SLH}}$-SLL - **SHH**-(SHL-SLL)	$\underline{\textbf{SLH}}$-SLL - **SHH**-(SHL-SLL)
Case 2-1	0 - SHL - (**SHH**-$\overline{\text{SLH}}$)	0 - SLL	**SHH**-SHL - $\overline{\text{SLH}}$	$\underline{\textbf{SLH}}$-SLL - **SHH**-(SHL-SLL)
Case 2-2	0 - SHL	0 - SLL	0 - $\overline{\text{SLH}}$	$\underline{\textbf{SLH}}$-SLL - **SHH**-(SHL-SLL)
Case 3-1	0 - SHL	SLL - SHL-**SHH**	0 - $\overline{\text{SLH}}$	0 - 0
Case 3-2	0 - SHL	SLL - SHL	0 - $\overline{\text{SLH}}$	0 - 0

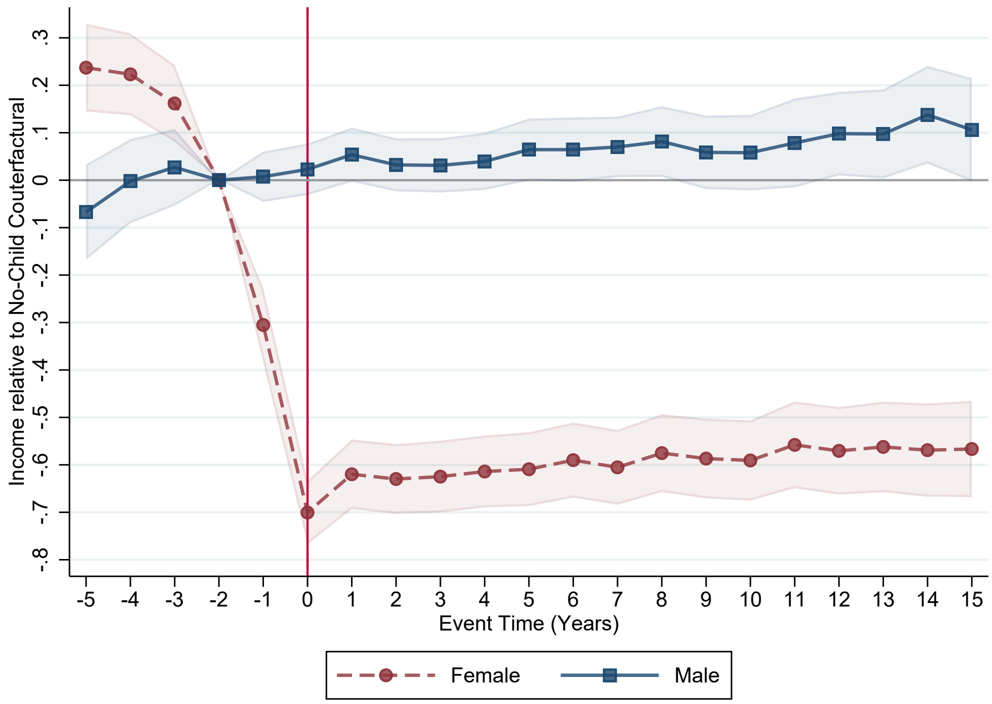
3.7 Figure

Figure 3.1: Wife's Share of Childcare Time vs Wife's Share of Income



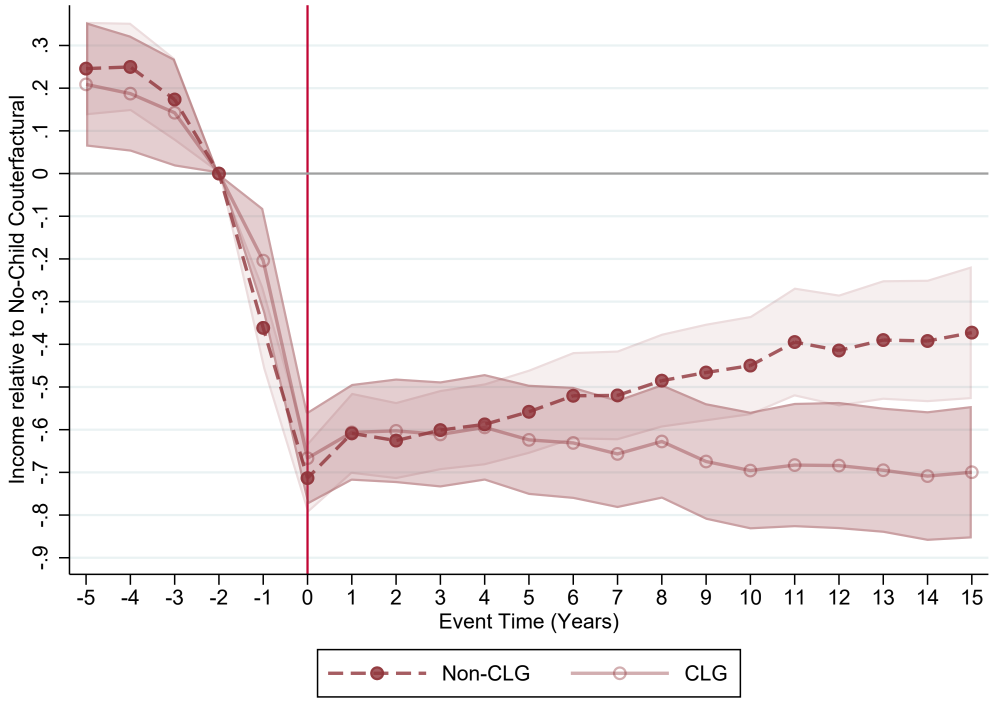
Notes: Each marker represents the wife's average share of income and time for housework and childcare over the sample periods within each household.

Figure 3.2: Impacts of Having a Child on Earnings



Notes: This shows coefficients of event time as a percentage of the counterfactual outcome without children for men and women. Event time is equal to zero when they have a baby.

Figure 3.3: Impacts of Having a Child on Earnings by Education



Notes: This shows coefficients of event time as a percentage of the counterfactual outcome without children for men and women. Event time is equal to zero when they have a baby.

Figure 3.4: Dynamics of Employment Rate by Education

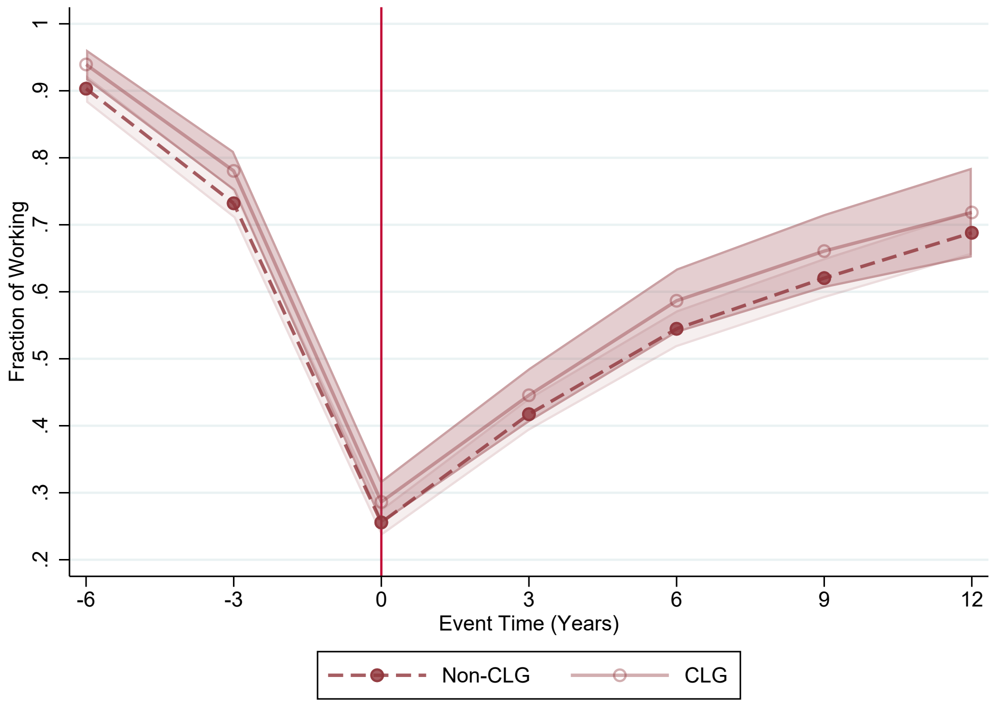


Figure 3.5: Dynamics of Fulltime Rate by Education

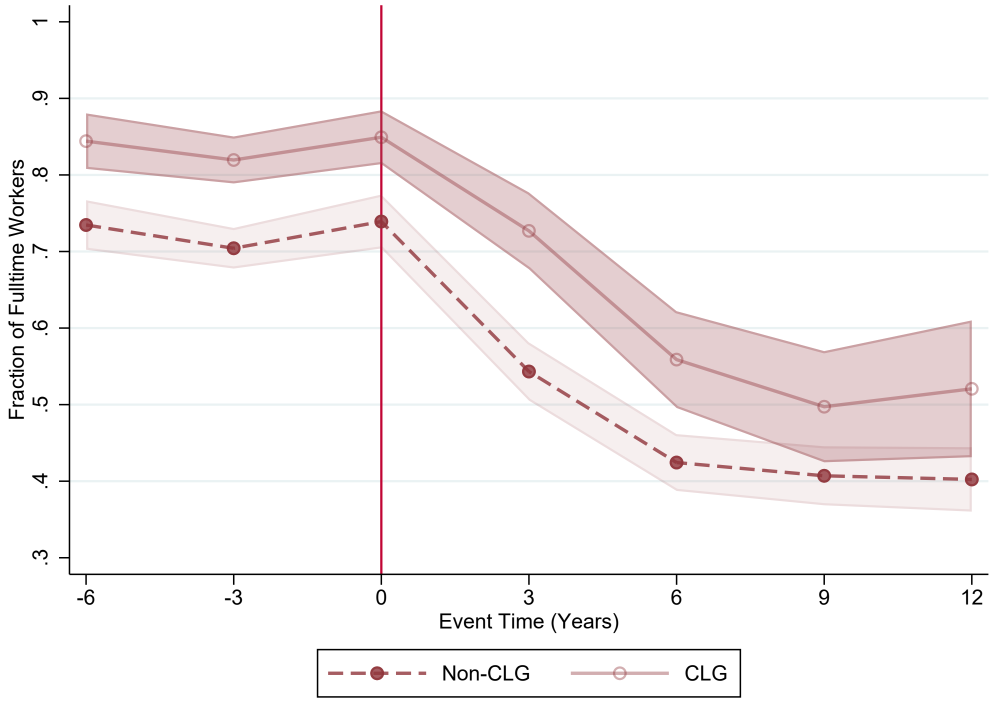
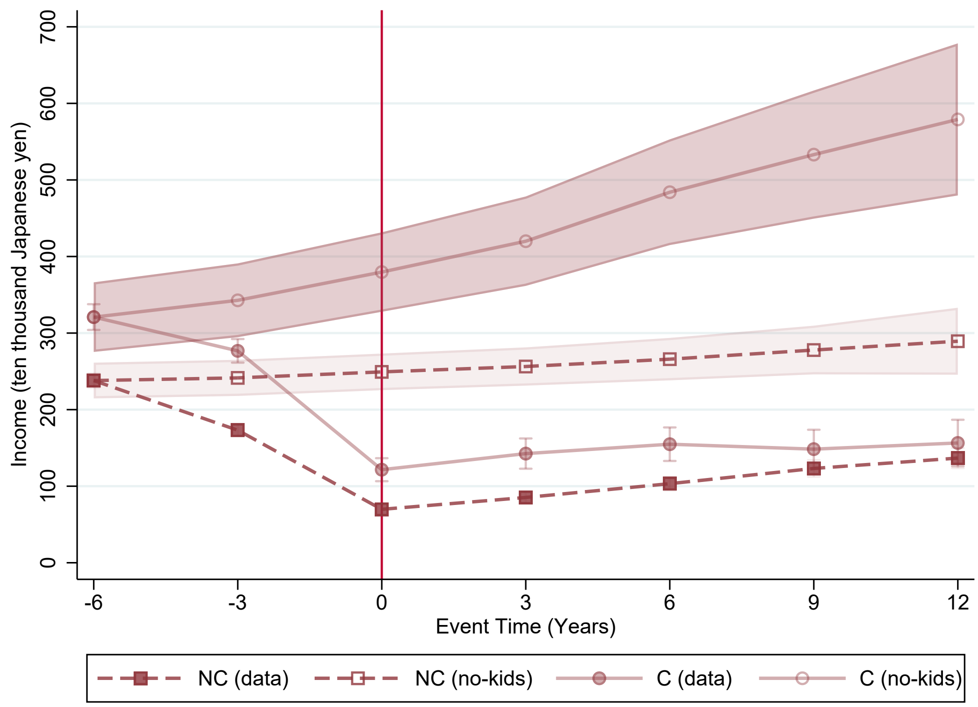


Figure 3.6: Observed and Counter-factual Dynamics of Earnings



Bibliography

Acland, Dan, and Matthew R Levy. 2015. "Naiveté, Projection Bias, and Habit Formation in Gym Attendance." *Management science* 61 (1): 146–160.

Aguiar, Mark, Erik Hurst, and Loukas Karabarbounis. 2013. "Time Use during the Great Recession." *The American economic review* 103 (5): 1664–1696.

Allcott, Hunt, and Judd B Kessler. 2019. "The Welfare Effects of Nudges: A Case Study of Energy Use Social Comparisons." *American economic journal. Applied economics* 11 (1): 236–276.

Allcott, Hunt, and Todd Rogers. 2014. "The Short-Run and Long-Run Effects of Behavioral Interventions: Experimental Evidence from Energy Conservation." *The American economic review* 104 (10): 3003–3037.

Andersen, Torben M, and Michael Svarer. 2011. "State dependent unemployment benefits." *The Journal of risk and insurance* 78 (2): 325–344.

Anderson, Michael, Carlos Dobkin, and Tal Gross. 2012. "The Effect of Health Insurance Coverage on the Use of Medical Services." *American Economic Journal: Economic Policy* 4 (1): 1–27.

Andresen, Martin Eckhoff, and Emily Nix. 2022. "What Causes the Child Penalty? Evidence from Adopting and Same-Sex Couples." *Journal of labor economics* 40 (4): 971–1004.

Aron-Dine, Aviva, Liran Einav, and Amy Finkelstein. 2013. "The RAND Health Insurance Experiment, three decades later." *Journal of Economic Perspectives* 27 (1): 197–222.

Baicker, Katherine, and Dana Goldman. 2011. "Patient cost-sharing and healthcare spending growth." *Journal of Economic Perspectives* 25 (2): 47–68.

Baicker, Katherine, Sendhil Mullainathan, and Joshua Schwartzstein. 2015. "Behavioral Hazard in Health Insurance." *The Quarterly Journal of Economics* 130 (4): 1623–1667.

Baicker, Katherine, Sarah L Taubman, Heidi L Allen et al. 2013. "The Oregon experiment–effects of Medicaid on clinical outcomes." *The New England Journal of Medicine* 368 (18): 1713–1722.

Baily, Martin Neil. 1978. "Some aspects of optimal unemployment insurance." *Journal of public economics* 10 (3): 379–402.

Barcellos, Silvia Helena, and Mireille Jacobson. 2015. "The Effects of Medicare on Medical Expenditure Risk and Financial Strain." *American Economic Journal: Economic Policy* 7 (4): 41–70.

Becker, Gary S. 1973. "A Theory of Marriage: Part I." *The journal of political economy* 81 (4): 813–846.

Becker, Gary S, and Kevin M Murphy. 1988. "A Theory of Rational Addiction." *The journal of political economy* 96 (4): 675–700.

Bertrand, Marianne, Patricia Cortes, Claudia Olivetti, and Jessica Pan. 2020. "Social Norms, Labour Market Opportunities, and the Marriage Gap Between Skilled and Unskilled Women." *The Review of economic studies* 88 (4): 1936–1978.

Bilal, Adrien. 2021. "The Geography of Unemployment." September.

Brot-Goldberg, Zarek C, Amitabh Chandra, Benjamin R Handel, and Jonathan T Kolstad. 2017. "What does a deductible do? The impact of cost-sharing on health care prices, quantities, and spending dynamics." *The Quarterly Journal of Economics* 132 (3): 1261–1318.

Card, David, Carlos Dobkin, and Nicole Maestas. 2004. "The Impact of Nearly Universal Insurance Coverage on Health Care Utilization and Health: Evidence from Medicare." *NBER Working Paper*.

Card, David, Carlos Dobkin, and Nicole Maestas. 2008. "The Impact of Nearly Universal Insurance Coverage on Health Care Utilization: Evidence from Medicare." *The American Economic Review* 98 (5): 2242–2258.

Card, David, Carlos Dobkin, and Nicole Maestas. 2009. "DOES MEDICARE SAVE LIVES?" *The Quarterly Journal of Economics* 124 (2): 597–636.

Chandra, Amitabh, Evan Flack, and Ziad Obermeyer. 2021. "The Health Costs of Cost-Sharing." February.

Chandra, Amitabh, Jonathan Gruber, and Robin McKnight. 2010. "Patient Cost-Sharing and Hospitalization Offsets in the Elderly." *The American Economic Review* 100 (1): 193–213.

Chandra, Amitabh, Benjamin Handel, and Joshua Schwartzstein. 2019. "Behavioral economics and health-care markets." In *Handbook of Behavioral Economics*, edited by Bernheim, B Douglas, Stefano DellaVigna, and David Laibson Volume 2. 459–502, North-Holland.

Charness, G, and U Gneezy. 2009. "Incentives to exercise." *Econometrica: journal of the Econometric Society* 77 (3): 909–931.

Chetty, Raj. 2006. "A general formula for the optimal level of social insurance." *Journal of public economics* 90 (10): 1879–1901.

Chetty, Raj. 2008. "Moral Hazard versus Liquidity and Optimal Unemployment Insurance." *The journal of political economy* 116 (2): 173–234.

Chiappori, Pierre-André, Bernard Salanié, and Yoram Weiss. 2017. "Partner Choice, Investment in Children, and the Marital College Premium." *The American economic review* 107 (8): 2109–2167.

Choudhry, Niteesh K, Jerry Avorn, Robert J Glynn et al. 2011. "Full Coverage for Preventive Medications after Myocardial Infarction." *The New England journal of medicine* 365 (22): 2088–2097.

Cullen, Julie Berry, and Jonathan Gruber. 2000. "Does Unemployment Insurance Crowd out Spousal Labor Supply?" *Journal of labor economics* 18 (3): 546–572.

Ehrlich, Isaac, and Gary S Becker. 1972. "Market Insurance, Self-Insurance, and Self-Protection." *The journal of political economy* 80 (4): 623–648.

Einav, Liran, and Amy Finkelstein. 2017. "Moral Hazard in Health Insurance: What We Know and How We Know It." November.

Einav, Liran, Amy Finkelstein, and Paul Schrimpf. 2015. "The Response of Drug Expenditure to Nonlinear Contract Design: Evidence from Medicare Part D." *The quarterly journal of economics* 130 (2): 841–899.

Ellis, Randall P. 1986. "Rational Behavior in the Presence of Coverage Ceilings and Deductibles." *The Rand Journal of Economics* 17 (2): 158–175.

Ellis, Randall P, Bruno Martins, and Wenjia Zhu. 2017. "Health care demand elasticities by type of service." *Journal of Health Economics* 55 232–243.

Elsby, Michael W L, Bart Hobijn, and Ayşegül Şahin. 2015. "On the importance of the participation margin for labor market fluctuations." *Journal of monetary economics* 72 64–82.

Engelhardt, Gary V, and Jonathan Gruber. 2011. "Medicare Part D and the Financial Protection of the Elderly." *American Economic Journal: Economic Policy* 3 (4): 77–102.

Finkelstein, Amy, Neale Mahoney, and Matthew J Notowidigdo. 2018. "What does (formal) health insurance do, and for whom?" *Annual Review of Economics* 10 (1): 261–286.

Finkelstein, Amy, and Robin McKnight. 2008. "What did Medicare do? The initial impact of Medicare on mortality and out of pocket medical spending." *Journal of Public Economics* 92 (7): 1644–1668.

Finkelstein, Amy, Sarah Taubman, Bill Wright et al. 2012. "THE OREGON HEALTH INSURANCE EXPERIMENT: EVIDENCE FROM THE FIRST YEAR." *The Quarterly Journal of Economics* 127 (3): 1057–1106.

Fujita, Shigeru, and Garey Ramey. 2009. "The cyclicality of separation and job finding rates." *International economic review* 50 (2): 415–430.

Fukushima, Kazuya, Sou Mizuoka, Shunsuke Yamamoto, and Toshiaki Iizuka. 2016. "Patient cost sharing and medical expenditures for the Elderly." *Journal of Health Economics* 45 115–130.

Goodman-Bacon, Andrew. 2021. "The Long-Run Effects of Childhood Insurance Coverage: Medicaid Implementation, Adult Health, and Labor Market Outcomes." *The American Economic Review* 111 (8): 2550–2593.

Grembi, Veronica, Tommaso Nannicini, and Ugo Troiano. 2016. "Do Fiscal Rules Matter?" *American economic journal. Applied economics* 8 (3): 1–30.

Gross, Tal, and Matthew J Notowidigdo. 2011. "Health insurance and the consumer bankruptcy decision: Evidence from expansions of Medicaid." *Journal of Public Economics* 95 (7): 767–778.

Hall, Robert E. 2005. "Employment Fluctuations with Equilibrium Wage Stickiness." *The American economic review* 95 (1): 50–65.

Haltiwanger, John C, Henry R Hyatt, Lisa B Kahn, and Erika McEntarfer. 2018. "Cyclical Job Ladders by Firm Size and Firm Wage." *American Economic Journal: Macroeconomics* 10 (2): 52–85.

Han, Hsing-Wen, Hsien-Ming Lien, and Tzu-Ting Yang. 2020. "Patient Cost-Sharing and Healthcare Utilization in Early Childhood: Evidence from a Regression Discontinuity Design." *American Economic Journal: Economic Policy* 12 (3): 238–278.

Hsu, Joanne W, David A Matsa, and Brian T Melzer. 2014. "Positive Externalities of Social Insurance: Unemployment Insurance and Consumer Credit." July.

Hu, Luojia, Robert Kaestner, Bhashkar Mazumder, Sarah Miller, and Ashley Wong. 2018. "The Effect of the Affordable Care Act Medicaid Expansions on Financial Wellbeing." *Journal of Public Economics* 163 99–112.

Hussam, Reshmaan, Atonu Rabbani, Giovanni Reggiani, and Natalia Rigol. 2022. "Rational Habit Formation: Experimental Evidence from Handwashing in India." *American economic journal. Applied economics* 14 (1): 1–41.

Iizuka, Toshiaki, and Hitoshi Shigeoka. 2021. "Asymmetric Demand Response when Prices Increase and Decrease: The Case of Child Healthcare." *The Review of Economics and Statistics, forthcoming* 1–30.

Iizuka, Toshiaki, and Hitoshi Shigeoka. 2022. "Is Zero a Special Price? Evidence from Child Healthcare." *American Economic Journal: Economic Policy, forthcoming*.

Ikegami, Naoki, Byung-Kwang Yoo, Hideki Hashimoto et al. 2011. "Japanese universal health coverage: evolution, achievements, and challenges." *The Lancet* 378 (9796): 1106–1115.

Ito, Koichiro, Takanori Ida, and Makoto Tanaka. 2018. "Moral Suasion and Economic Incentives: Field Experimental Evidence from Energy Demand." *American Economic Journal: Economic Policy* 10 (1): 240–267.

Jessoe, Katrina, and David Rapson. 2014. "Knowledge Is (Less) Power: Experimental Evidence from Residential Energy Use." *The American economic review* 104 (4): 1417–1438.

Just, David R, and Joseph Price. 2013. "Using Incentives to Encourage Healthy Eating in Children." *The Journal of human resources* 48 (4): 855–872.

Keeler, E B, J P Newhouse, and C E Phelps. 1977. "Deductibles and the Demand for Medical Care Services: The Theory of a Consumer Facing a Variable Price Schedule under Uncertainty." *Econometrica* 45 (3): 641–655.

Keeler, E B, and J E Rolph. 1988. "The demand for episodes of treatment in the Health Insurance Experiment." *Journal of Health Economics* 7 (4): 337–367.

Kleven, Henrik. 2022. "The Geography of Child Penalties and Gender Norms: Evidence from the United States." June.

Kleven, Henrik, Camille Landais, Johanna Posch, Andreas Steinhauer, and Josef Zweimüller. 2019a. "Child Penalties across Countries: Evidence and Explanations." *AEA Papers and Proceedings* 109 122–126.

Kleven, Henrik, Camille Landais, and Jakob Egholt Søgaard. 2019b. "Children and Gender Inequality: Evidence from Denmark." *American economic journal. Applied economics* 11 (4): 181–209.

Kleven, Henrik, Camille Landais, and Jakob Egholt Søgaard. 2021. "Does Biology Drive Child Penalties? Evidence from Biological and Adoptive Families." *American Economic Review: Insights* 3 (2): 183–198.

Kondo, Ayako, and Hitoshi Shigeoka. 2013. "Effects of universal health insurance on health care utilization, and supply-side responses: Evidence from Japan." *Journal of Public Economics* 99 1–23.

Kroft, Kory, and Matthew J Notowidigdo. 2016. "Should Unemployment Insurance Vary with the Unemployment Rate? Theory and Evidence." *The Review of economic studies* 83 (3): 1092–1124.

Krueger, Alan B, and Bruce D Meyer. 2002. "Chapter 33 Labor supply effects of social insurance." In *Handbook of Public Economics*, Volume 4. 2327–2392, Elsevier.

Krueger, Alan B, and Andreas Mueller. 2010. "Job search and unemployment insurance: New evidence from time use data." *Journal of public economics* 94 (3): 298–307.

Krusell, Per, Toshihiko Mukoyama, Richard Rogerson, and Ayşegül Şahin. 2017. "Gross Worker Flows over the Business Cycle." *The American economic review* 107 (11): 3447–3476.

Kuhn, Moritz, Iourii Manovskii, and Xincheng Qiu. 2021. "The Geography of Job Creation and Job Destruction." October.

List, John A, and Anya Savikhin Samek. 2015. "The behavioralist as nutritionist: Leveraging behavioral economics to improve child food choice and consumption." *Journal of health economics* 39 135–146.

Loewenstein, George, Joseph Price, and Kevin Volpp. 2016. "Habit formation in children: Evidence from incentives for healthy eating." *Journal of health economics* 45 47–54.

Mazumder, Bhashkar, and Sarah Miller. 2016. "The Effects of the Massachusetts Health Reform on Household Financial Distress." *American Economic Journal: Economic Policy* 8 (3): 284–313.

McGuire, Thomas G. 2011. "Demand for health insurance." *Handbook of Health Economics* 2 317–396.

Mueller, Andreas I. 2017. "Separations, Sorting, and Cyclical Unemployment." *The American economic review* 107 (7): 2081–2107.

Mukoyama, Toshihiko, Christina Patterson, and Ayşegül Şahin. 2018. "Job Search Behavior over the Business Cycle." *American Economic Journal: Macroeconomics* 10 (1): 190–215.

Nilsson, Anton, and Alexander Paul. 2018. "Patient cost-sharing, socioeconomic status, and children's health care utilization." *Journal of Health Economics* 59 109–124.

Pissarides, Christopher A. 2000. *Equilibrium Unemployment Theory, second edition*. MIT Press.

Royer, Heather, Mark Stehr, and Justin Sydnor. 2015. "Incentives, Commitments, and Habit Formation in Exercise: Evidence from a Field Experiment with Workers at a Fortune-500 Company." *American economic journal. Applied economics* 7 (3): 51–84.

Sánchez, Juan M. 2008. "Optimal state-contingent unemployment insurance." *Economics letters* 98 (3): 348–357.

Schmieder, Johannes F, and Till von Wachter. 2016. "The Effects of Unemployment Insurance Benefits: New Evidence and Interpretation." *Annual review of economics* 8 (1): 547–581.

Schmieder, Johannes F, Till von Wachter, and Stefan Bender. 2012. "The Effects of Extended Unemployment Insurance Over the Business Cycle: Evidence from Regression

Discontinuity Estimates Over 20 Years." *The quarterly journal of economics* 127 (2): 701–752.

Schofield, Heather, George Loewenstein, Jessica Kopsic, and Kevin G Volpp. 2015. "Comparing the effectiveness of individualistic, altruistic, and competitive incentives in motivating completion of mental exercises." *Journal of health economics* 44 286–299.

Shigeoka, Hitoshi. 2014. "The Effect of Patient Cost Sharing on Utilization, Health, and Risk Protection." *The American Economic Review* 104 (7): 2152–2184.

Shimer, Robert. 2004. "Search Intensity." https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.122.4693&rep=rep1&type=pdf, Accessed: 2022-2-12.

Shimer, Robert. 2005. "The Cyclical Behavior of Equilibrium Unemployment and Vacancies." *The American economic review* 95 (1): 25–49.

Shimer, Robert. 2012. "Reassessing the ins and outs of unemployment." *Review of economic dynamics* 15 (2): 127–148.

Stigler, George J, and Gary S Becker. 1977. "De Gustibus Non Est Disputandum." *The American economic review* 67 (2): 76–90.

Volpp, Kevin G, Leslie K John, Andrea B Troxel, Laurie Norton, Jennifer Fassbender, and George Loewenstein. 2008. "Financial incentive-based approaches for weight loss: a randomized trial." *JAMA: the journal of the American Medical Association* 300 (22): 2631–2637.

Volpp, Kevin G, Andrea B Troxel, Mark V Pauly et al. 2009. "A Randomized, Controlled Trial of Financial Incentives for Smoking Cessation." *The New England journal of medicine* 360 (7): 699–709.

Wherry, Laura R, Sarah Miller, Robert Kaestner, and Bruce D Meyer. 2018. "Childhood Medicaid Coverage and Later-Life Health Care Utilization." *The Review of Economics and Statistics* 100 (2): 287–302.