In Sickness and in Health Insurance: Gender in Household Benefit Plan Choice*

Yaying Zhou†

November 9, 2022

(Click Here for Most Recent Version)

Abstract

In the United States, most households obtain health insurance from employers. I document that households subscribe to health plans offered by the husband's employer more often than the wife's, a difference which could reflect their gap in employment status, plan quality, or a per se gender effect. I identify the gender effect and quantify its welfare implications using all-payer claims data from New Hampshire. Using a novel identification strategy that holds plan quality and employment status fixed, I find the gender effect leads to a 7% higher propensity for households to take up the husband's plan, which explains one-third of the gender gap in family plan subscriptions. I further develop a structural model of household plan choice incorporating multiple dimensions of inattention. I find households only pay attention to their non-default subscriber with a 40% probability. Counterfactual simulations suggest inattention contributes to one-third of surplus loss in household health insurance subscriptions, and that additionally removing gender-based inattention yields twice the welfare benefit of a policy that focuses on plan-based inattention alone.

JEL Codes: D12, D13, D91, I13, J16
Keywords: health Insurance, inattention, gender gap, intrahousehold decision

*I am deeply grateful to my dissertation committee, Haizhen Lin, Daniel Sacks, Kosali Simon, Jeff Prince, and Jackson Dorsey for their guidance and encouragement. This work has also benefited from valuable conversations with Jason Abaluck, Michael Baye, Andrew Butters, Allan Collard-Wexler, Maura Coughlin, Rick Harbaugh, Kurt Lavetti, John Maxwell, Matthijs Wildenbeest, Mo Xiao, Ruli Xiao, and participants at the BEPP Seminar Series in Kelley School of Business and Health Policy Workshop in O'Neill School of Public and Environmental Affairs. In addition, I thank the New Hampshire Department of Health and Human Services, New Hampshire Insurance Department, along with Mary Fields and Maureen Mustard, for providing the data and sharing helpful insights. The views expressed here and all errors are my own.

†Business Economics & Public Policy, Kelley School of Business, Indiana University, 1309 E. Tenth St. Bloomington, IN 47405 (e-mail: yayizhou@iu.edu).

1 Introduction

The success of a market depends crucially on consumers' ability to choose. Consumers, however, often make suboptimal choices. For example, on health insurance, a vast literature has documented substantial choice inconsistencies for individuals in settings like Medicare Part D.[1] Yet not all health insurance decisions are made at the *individual* level. As a matter of fact, most population in the United States enroll as a *household* when they obtain health insurance through employers. However, our understanding of whether households are making sensible health insurance decisions is limited.

Due to data limitations, it has been standard in the literature to assume households as individuals who signed up for health insurance on behalf of the family.[2] However, household decisions are far more complicated as being jointly made by members with diverse knowledge, preferences, and bargaining leverage. Compared to individual choices, household decisions are additionally subject to frictions generated from members' intrahousehold bargaining. An important factor that could pin down the ultimate bargaining outcome is gender. For example, females decide more on domestic work while males have more say in financial decisions [Barber and Odean (2001)].

Gender can also affect households' health insurance decisions. In the United States, around 66% dual-earner households face the option of enrolling in the benefit plan offered by the husband's or wife's employer each year. This choice is costly since health spending takes 4% of their annual income and even 10% for low-income groups.[3] Though essentially a comparison of costs and quality, this decision could also be influenced by gender. For example, a household might only consider the husband's plans, defaulting to his employer without carefully comparing all options in their choice set. Understanding how gender affects their choice can guide policies that improve choice quality in the insurance market. For instance, if inattention to both spouses' options generates apparent lock-in, a policy that reduces switching costs wouldn't be as effective as one that raises attention to both employers in their choice set.

This paper makes three contributions to studying the role of gender in households' employer-sponsored health insurance (ESI hereafter) decisions. I first document the phenomenon that more families subscribe to the husband's ESI than the wife's. This gap might potentially be explained by the gender gap in employment status and coverage

[1]For example, Abaluck and Gruber (2011), Abaluck and Gruber (2016), Polyakova (2016), Ho et al. (2017), Bhargava et al. (2017), and Heiss et al. (2021).

[2]For example, Einav et al. (2013), Handel (2013), Handel and Kolstad (2015), Ho and Lee (2020), Marone and Sabety (2022), and Tilipman (2022).

[3]Estimates from Kaiser Family Foundation. Low-income families have an annual income below or equal to 199% of the Federal Poverty Line.

rate, or any differences in benefit generosity offered by the husband's and wife's employers. My second contribution is identifying a per se gender effect by ruling out the above two explanations with a novel identification strategy. Third, I develop and estimate a structural model to show that gender-based inattention is an important mechanism. Counterfactual simulations suggest raising attention can be an effective policy to reduce overspending in employer-sponsored health insurance.

I start by presenting national estimates on the gender of family plan subscribers from the nationally representative Medical Expenditure Panel Survey (MEPS hereafter). Around 66.3% families are enrolled in the husband's ESI versus 33.7% in the wife's, thus creating a 32.6% gender gap. Several factors could explain this. To begin with, married women have a lower employment rate than married men. I therefore focus on dual-working households where both spouses are offered ESI and find that the gender gap reduces to 18%. A second possible reason is that plans offered by the husbands' employers are systematically more generous than wives'. However, even among colleague couples who share exactly the same plan options, most households still choose the husband as their subscriber, suggesting a per se gender effect. I further test this by associating the gender gap with proxies for intrahousehold bargaining power from the literature.[4] The gap increases as husbands appear to have more say in the household, and vice versa. For example, a majority of families subscribe to the husband's plan when husbands receive more education than wives.

Motivated by the descriptive evidence, I next identify the per se gender effect and show that it can explain one-third of the gender gap from the gender differences in employment status and plan generosity. To do this, I analyzed New Hampshire's All-Payer Claims Database, which records plan enrollment and insurance claims for its ESI population from 2011 to 2021. I first confirm the gender gap in New Hampshire is large and comparable to the gap in the nationally representative data. To cleanly identify the gender effect, I compare couples that share the same pairs of employers but differ in which employer the husband and wife work for. This identification strategy infers a per se gender effect by comparing the propensity to choose plan A over plan B when the husband's employer offers plan A, relative to when the wife's employer offers it. This design automatically rules out the employment and plan generosity gap so that it is purely driven by the variation in gender. I find a 7% higher probability of enrolling in the husband's plan due to the gender effect, which explains a third of the gender gap in family plan enrollment. The weight on the husband's plan also increases with the relative decision power of husbands. Further, I detect gender-based inattention as an important mechanism from the observation that families infrequently switched subscribers even when their subscribers changed jobs.

[4]For example, Browning et al. (1994), Friedberg and Webb (2006), and Bertocchi et al. (2014).

While these tests are important to identify the gender effect, to understand its mechanisms and quantify their impacts on welfare, I develop and estimate a structural model. In the model, households maximize expected utility subject to switching costs and limited attention. The expected utility depends on the expectation and variance of plan costs, which I impute using machine learning techniques. To account for the endogeneity and measurement error of costs, I use a control function approach with insurers' reimbursement discounts as the excluded instrument. This model includes plan-based inattention developed in Ho et al. (2017), Heiss et al. (2021), and Abaluck and Adams-Prassl (2021). It also expands on prior work by incorporating both spouses' options and allowing for the possibility that households may not consider all options in their choice set even when they wake up from their default plan. My model allows for gender to influence choice in three distinct ways: First, households have an intrinsic 'gender preference' for having the husband as the subscriber; Second, after initial choices are made, they become inattentive to their default subscriber; Third, they may as well be inattentive to the default plan offered by their default subscriber and are reluctant to pay switching costs. My estimates reveal a substantial role for gender-based inattention in addition to plan-based inattention. To illustrate, households only consider options from their non-default spouse with a 40% probability, even when they 'wake up' from their default plan and start to make plan comparisons.

I take these model estimates to study the effectiveness of counterfactual interventions that aim to reduce inefficiencies by raising attention. Overall, forcing households to pay full attention significantly reduces the gender gap and improves consumer welfare by $227 (34%). Specifically, I find a policy that removes plan-based inattention can be twice as effective if combined with gender-based inattention. This analysis inspires policies that raise awareness to options in a household's entire choice set, especially from the non-default spouse. In addition, I find that households that benefit the most are those who have 'mistakenly' established their default subscriber as the one with high-cost options. This result can serve as a criterion for targeting more vulnerable households to improve their choice quality.

This paper contributes to the literature on choice inconsistencies in health insurance. Prior research has found lack of information [e.g. Kling et al. (2012)], biased information processing [e.g., Abaluck and Gruber (2011); Abaluck and Gruber (2016)], and inertia [e.g., Ho et al. (2017); Heiss et al. (2021)] as potential explanations by examining individual choices in Medicare Part D. Though these frictions are important aspects of the decision process, it is not always the case that consumers decide for themselves. This is especially true in ESI, where the unit of enrollment is usually a household, and where intrahousehold bargaining can significantly alter the decision collectively made by members. Unfortunately, intrahousehold dynamics have been assumed away in previous models of ESI choice, since they typically acquire data from one employer.[5] This paper overcomes the challenge by seeking to construct households' entire choice sets, taking advantage of a long panel of enrollment records from New Hampshire's all-payer claims data. In doing this, one can finally open the black box of households and investigate intrahousehold drivers of choice inconsistencies in health insurance and its implications on welfare. Indeed, I find households are 7% more likely to enroll in the husband's plan than the wife's, a pattern that can not be explained by their differences in plan generosity.[6] More broadly, this paper also echoes the recent work by Handel et al. (2020) on the social determinants of insurance choice. I contribute by further microfounding the gender effect into behavioral bias such as inattention. In particular, I find that gender-based inattention is responsible for 11% of the welfare loss in ESI. Understanding this is important for designing policies that address socioeconomic drivers of choice inconsistencies in the health insurance market.

The present work also builds on the literature that studies intrahousehold frictions and the role of gender. My finding on a per se gender effect serves as a test to support a bargaining model of household decision-making as opposed to a unitary model in a new setting of health insurance.[7] While recent papers [Vihriälä (2022); Choukhmane et al. (2022)] provide ex-post calculations on households' misallocation of financial accounts, my structural model quantifies intrahousehold frictions with an ex-ante perspective. Regarding the role of gender in household financial decisions, a general finding from the literature is that women are under-represented [e.g. Gu et al. (2021); Ke (2021)] and more egalitarian gender norms can raise household wealth [Guiso et al. (2021); Banerjee et al. (2021)]. My analysis confirms the gender gap in household health insurance decisions. More importantly, the structural model allows me to assess how much surplus could be achieved by eliminating specific sources of gender disparity within the household. For instance, shutting down gender-based inattention can help an average household save \$72 each year, and some as much as \$1,000.

The remainder of this paper is organized as follows. Section 2 displays descriptive evidence that motivates my work. Section 3 introduces the data. Section 4 presents the reduced-form analysis that tests the existence of the gender effect. In Section 5, I propose a consideration set choice model on household health insurance choices supplied by cost variables constructed in Section 6. Based on model estimates presented in Section

[5] These include: Einav et al. (2013); Handel (2013); Handel and Kolstad (2015); Bhargava et al. (2017); Ho and Lee (2020); Marone and Sabety (2022); Tilipman (2022).

[6] Prior work on how gender affects ESI decisions include Dushi and Honig (2003), who find ESI take-up rate for female employees is more sensitive to whether her spouse is covered than male employees.

[7] The unitary model [Samuelson (1956); Becker (1974)] assumes that members pool resources and maximize a consensus utility function, so that the distribution of resources does not affect household outcomes. On the other hand, the collective model [Chiappori (1988); Chiappori (1992)] allows members to have different preferences, which can have an impact on a household's collective choice.

7, I simulate in Section 8 counterfactual interventions that mitigate specific sources of inefficiencies related to the gender effect. Section 9 concludes.

2 Descriptive Evidence

I start the analysis by presenting descriptive evidence on the prevalence of husband subscribers in family plans. Table 1 displays the percentage of households that enroll in the husband's plan by their number of ESI offers. The first two columns are calculated from MEPS 2010-2019,[8] the last two from New Hampshire's insurance claims data, which will be described in detail below. As the first row shows, 66.3% households in the United States take up their husband's plan when having at least one offer of employer-sponsored health insurance. This number shrinks to 62.5% in the Northeast, and further to 60.7% in New Hampshire, generating a 21.4% gender gap in family plan subscriptions for the state.

What is driving the gender gap? One possible explanation is the employment gap between husbands and wives. Indeed, around 18.5% married-couple families only have husbands employed, in contrast to 6.8% wives [U.S. Bureau of Labor Statistics]. Further, even within dual-earner families, husbands are 11% more likely to be offered ESI than wives.[9] I therefore focus on dual-working households where both spouses have ESI coverage from their employers. The gender disparity turns out to decline, but not disappear in all columns and particularly in New Hampshire, to 12.8%. This reveals that the employment/coverage gap accounts for around 40% of the gender gap in the state.

Another plausible contributor to the gender gap is differences in plan generosity offered by husbands' and wives' employers (plan generosity gap). This can happen systematically when husbands tend to work for employers that offer more generous health benefits than wives. To tease out this possibility and test whether the gender gap still remains, I investigate enrollment patterns for colleague couples in Table 2 column (2). This group is particularly interesting as husbands and wives are offered exactly the same set of plans. Still, 62.4% households have husbands as their ESI subscribers, suggesting the gender gap has implications beyond the employment and plan generosity gap.[10]

In this paper, I propose that the remaining gender gap is explained by a per se *gender effect*. That is, I define the *gender effect* as the higher propensity for a dual-offered household

[8] MEPS tracks each household through a 5-round survey during a two-year period. In addition to health insurance information, MEPS also records the demographics, earnings, and employment status of each member in the household.

[9] Based on my calculation from MEPS. For simplicity, I refer to the gender disparity in both employment and ESI coverage as the 'employment gap'.

[10] Regrettably, colleague couples here appear to be older and less healthy than the general population in New Hampshire, and have larger age gap between husbands and wives.

to enroll in the husband's ESI, relative to the wife's, when holding plan generosity fixed. At this stage, I refrain from attributing it to a particular source but prefer to generalize it as an 'aggregate effect' that could be a mixture of possibilities. For example, the gender effect can reflect households' preference to stay with the spouse who has better job stability to save switching costs in the long run. It can also be that husbands are more likely to be in charge of 'major decisions' such as health insurance subscriptions for the family. Indeed, only 61.4% couples report being jointly responsible for family health insurance decisions, with 22% leaving the job to the husband, in contrast to 16.6% to the wife [Lim et al. (2022)]. The higher propensity for the husband to be the decision maker could be due to gender norms that 'money chores' should lie in the domain of men [Barber and Odean (2001); Ke (2021); Guiso et al. (2021)], or the fact that he tends to be the payer of premium bills. Even in an equally-involved dialogue, couples might still use the rule of thumb and take up the plan offered by the main earner, usually the husband, when they associate higher income with more benefits. These heuristics should be used with caution as it is not always the case that the main breadwinner will bring a better plan to the table, especially when premiums are progressive for some employers.[11] In addition, health insurance is 'horizontally differentiated' in that a plan that suits one household doesn't necessarily fit another, as their health needs may vary.

Whatever the reason is, if the gender effect is caused by unequal decision-making within the household, one would expect larger gaps for male-dominated families. To test this, I correlate the propensity to enroll in a husband's plan with proxies for intra-household decision power using data on dual-offered couples from MEPS. Following the numerous literature on determinants of intrahousehold bargaining,[12] I consider the husband with more bargaining power as older, having higher income, education, claiming to be the 'head', and vice versa. As is shown in Figure 1, over 60% households pick the husband's plan when they appear to have more say in their families. However, the gender gap is less prominent when wives tend to be decision-makers. For example, when wives are the primary breadwinner, only 43.8% households end up subscribing to the husband's plan. This is not what one would expect if health plan decisions were purely driven by cost considerations. Rather, the distribution of power 'under the roof' can also play a role.

While this section presents strong evidence on the existence of the gender effect, it highlights the need for clean identification with a representative sample where husbands and wives are not colleagues. It is also helpful if one wants to infer the magnitude of the gender

[11] See an example of Indiana University, where low-wage employees contribute less to premiums for the same plan.

[12] For example, Browning et al. (1994), Friedberg and Webb (2006), and Bertocchi et al. (2014).

effect from cost differences between husbands' and wives' plan options. The identification and quantification will be detailed in Section 4 and 5 respectively, but before that, I would like to introduce the data environment demanded by these exercises.

3 Data

My primary data comes from New Hampshire Comprehensive Health Care Information System (NH CHIS) from 2011 to 2021. This all-payer claims database records ESI enrollment and medical claims for the universe of enrollees who were state residents or subscribed to an insurance policy issued in the state.[13] In particular, it has been used to supply cost estimates for NH HealthCost, a price-transparency website for consumers to shop health providers.

The claims data contains detailed information on the date of healthcare utilization, diagnoses, procedures, and the amounts charged by providers, reimbursed by insurers, and paid by patients. The enrollment data keeps track of the (encrypted) group an enrollee subscribes to each month, the insurance carrier, enrollment tier (such as single or family subscription),[14] and the plan type (like HMO).[15] It also records enrollee demographics, including age, gender, and location at the zip-code level. I supplement enrollee characteristics with the Charlson Comorbidity Index calculated from diagnoses in the claims data.[16] In addition, using zip codes, I merge several proxies for gender disparity derived from 5-year estimates of American Community Survey.[17] These include the income ratio of males over females at the median level and the gender gap in average education years.

The ideal data to identify the gender effect would be a list of couples for whom I can extract plan brochures from both the husband's and the wife's employers. To my knowledge, such data does not exist, and almost all papers on ESI assume households make insurance choices from the chosen employer's menu. Nevertheless, the data here has three merits that make it the closest to studying household ESI choices. First, each individual is identified with a unique (encrypted) ID across employers and insurers. Second, each enrollee is also attributed to a (de-identified) family. Third, for each family member, the data reports their

[13] This dataset has also been analyzed in other studies. For example, Brown (2017), Brown (2019), Wang (2018), and Ackley (2020).

[14] Enrollment tiers include EMP/ESP/ECH/FAM. EMP is short for employee-only plans; ESP for employee-plus-spouse plans; ECH for employee-plus-children plans; and FAM for family plans.

[15] Plan types include HMO/PPO/POS/EPO/IND. HMO is short for Health Maintenance Organization, PPO for Preferred Provider Organization, POS for Point of Service, EPO for Exclusive Provider Organization, and IND for Indemnity plans.

[16] The Charlson Comorbidity Index is a method to predict mortality based on 17 categories of comorbidities. It assigns each comorbidity a weight from 1 and 6 and takes the sum of all the weights as a single score for the patient. A higher Charlson Comorbidity Index indicates higher health risks.

[17] I use data from 2011 to 2019. For years 2020 and 2021, I assume the statistics follow the year 2019.

relationship to the policyholder as a subscriber, spouse, or dependent. These features allow me to trace out a couple's employment history from their enrollment records as a subscriber during the 11-year panel.

Regrettably, it is impossible to tell a spouse's employer concurrently with her enrollment as a spouse. To solve this, I construct the potential choice set for dual-working families through the following steps. First, I track the spouse's employment by examining her enrollment records as a subscriber to either a medical, drug, or dental plan.[18] Then, I assign the spouse each year to the employer to which she most closely subscribed. For example, suppose I observe individual A is the subscriber for a plan sponsored by employer 1 in 2012 and another plan by employer 2 in 2016; in 2015 when she enrolls as a spouse in her husband's plan, I treat employer 2 as A's 'true' employer. When there is a tie (year 2014), I prioritize the employer in previous years (employer 1).

Since the gender effect only concerns dual-offered households, I further select couples who have ever enrolled in both spouses' employers during the sample period.[19] This selection indicates that households in my sample might be less subject to gender norms than households who have been consistently delegating one spouse as their subscriber. Consequently, my analysis should provide a lower bound for the gender effect. Nevertheless, I am reassured to observe that household demographics, especially those that represent intra-household gender disparity, are reasonably balanced between the selected sample and the state average (see columns (1) and (3) in Table 2). In addition, I require that even when the spouse's employer did not cover her family, it still covered her colleagues. This is to avoid cases when a single-offered household gets misclassified as dual-offered when the spouse's employer stops sponsoring health insurance. Eventually, these steps filter out $11,765$ colleague couples and $47,550$ non-colleague ones from around 320 thousand couples in New Hampshire.[20]

[18] Besides medical plans, drug and dental plans also provide valuable information on the spouse's employer. For example, Handel (2013) documents that less than 15% enrollees switch dental plans when switching medical plans.

[19] Only 1/3 families are left after this step. Around 2/3 of married couples have both spouses employed [U.S. Bureau of Labor Statistics]. These numbers together suggest that around half of the dual-working households have consistently subscribed to one spouse's ESI plan.

[20] Colleague-couples are more likely to be selected into the sample because it is easier to pin down their joint employers when they unconsciously swap subscribers.

4 Identification of Gender Effect

4.1 Empirical Strategy

Motivated by the descriptive evidence in Section 2, I next identify the gender effect in family plan subscriptions with a novel reduced-form design. To illustrate, I compare families that face the same choice set but only differ in whether the husband or wife works for a particular employer. As an example, consider two households H1 and H2, and two employers E1 and E2 in Table 3. Each household contains a husband [M] and a wife [F]. Ideally, when they are randomly assigned to jobs, as in Case A, the gender effect implies both households will be more inclined with the husband's plan, that is, E2 for H1 and E1 for H2.

In reality, husbands are more likely to be sorted into high-benefit jobs. This scenario is better exhibited in Case B, where both husbands work for E1, say, a generous high-tech giant hiring a disproportionate share of male programmers; both wives work for E2, a fashion startup attracting more female designers but providing limited benefits. In this case, even when both households pick E1, it is unclear whether their choices are driven by the gender effect or the better insurance offered by E1. That is, simply checking the enrollment gap between husbands' and wives' employers, even within a specific choice set, is not sufficient to identify the gender effect, and one must separate it from any gap in plan generosity that is potentially correlated with the gender distribution in these employers. To achieve this goal, I develop an empirical design with the following regression model:

$$Choice_{ijmt} = \beta Husband_{ijt} + \xi_m + [u_{jm}] + [\zeta_{jt}] + \epsilon_{ijmt} \tag{1}$$

where $Choice_{ijmt}$ is a binary indicator for whether household i chooses employer j's plan at year t within a choice set m composed of two distinct employers. On the right-hand side, $Husband_{ijt}$ denotes whether household i's husband works for employer j. In addition to the choice set fixed effects ξ_m, I also include employer-by-choice set fixed effects $[u_{jm}]$, which rule out the plan generosity gap by fixing the relative quality of plans within a choice set. The employer-by-year fixed effects $[\zeta_{jt}]$ further address the concern that plan benefits could vary over time. This specification indicates that the share explained by the plan generosity gap is precisely the reduction in β by adding the plan fixed effects $[u_{jm}]$. The remaining inclination towards a husband's plan, as captured in β in the most stringent specification, identifies the gender effect.

4.2 Baseline Results

Before presenting the empirical results, I would like to first examine the demographics of the reduced-form sample. As is shown in Table 2, non-colleague couples are comparable to the state average in almost all dimensions, including age, health risks, and proxies for the gender disparity in intra-household bargaining. My strictest model is run on a much smaller sample that includes couples whose choice sets involve at least two families. This subset of households also shows reasonably balanced features to the full sample. Hence, I am assured of drawing on estimates from this analytic sample for all households in the state.

Table 4 reports my baseline results. Based on estimates from column (1), households are 12.7% more likely to choose the husband's plan than the wife's. However, once I control for plan quality in column (2), the coefficient of *Husband* declines but remains statistically significant. This means all else equal, the gender effect distorts household decisions towards the husband's side by around 7.1%. Further including employer-by-year fixed effects in column (3) doesn't change the results remarkably, which suggests the time-varying variation of plan quality is insufficient to affect households' enrollment decisions.

These numbers help answer the question of why dual-offered households enroll with their husbands more: 44% ($\frac{12.7-7.1}{12.7}$) of it is due to better plans sponsored by the husband's employers, while the remaining 56% comes from the gender effect. Projecting these shares to the last column of Table 1 yields the final decomposition in Table 5, where the 21.4% gender gap is attributed to 8.6% employment gap, 5.6% plan generosity gap, as well as 7.2% gender effect, each of which contributing 40.2%, 26.2%, and 33.6% to the total gap.[21]

An implicit assumption behind my empirical approach is the correct knowledge of the spouse's employer, which is impossible to know. In practice, I attribute each spouse to the closest employer where she ever signed up as a subscriber. A natural indication of this method is the closer in time when I traced out her employer, the more accurate her family's choice set will be. In Panel A of Table A1, I separate the analysis by the gap in years when I observe the spouse enrolling as a subscriber. The robustness of the estimates suggests that my approach is unlikely to be the main factor that is driving the results.

If the way I reconstruct the choice set is uncorrelated with gender, the gender effect coefficient will still be unbiased. However, more often than not, it is the wife whose employer needs to be imputed, since most subscribers are husbands. Consequently, when I 'borrow' a wife's employer from neighboring years while she is in fact unemployed, the gender effect will be overestimated. To alleviate this concern, I draw on the fact that child care is a major reason that suppresses women's employment. For instance, in 2020, the labor force

[21] Employment gap: $21.4\% - 12.8\% = 8.6\%$. Plan generosity gap: $12.8\% \times 44\% = 5.6\%$. Gender effect: $12.8\% \times 56\% = 7.2\%$.

participation rate of mothers with children under age 3, 6, and 18 are 63.3%, 65.8%, and 71.2% respectively, in contrast to 93.5%, 93.4% and 92.3% for fathers [U.S. Bureau of Labor Statistics]. I then repeat the analysis in Panel B by the presence of children in these age groups. If families with younger children have higher unemployment rates for wives, causing an exaggerated choice set, they will also come out with larger coefficients. However, this is not the case, as families with the youngest children turn out to have the least estimate.

4.3 Gender Effect and Decision Power

How does the gender effect relate to intrahousehold decision power? Recall in Figure 1, I document that husband's subscription rate increases with the husband's power. I test this formally with Model (1) by interacting $Husband$ with three measures of household bargaining power: age gap, education gap, and income ratio.[22] Unfortunately, the claims data does not provide individual demographics beyond age and gender. I therefore merge each family with American Community Survey at the zip-code level for gender-specific median income and average education years. Though not perfect, these statistics help depict the gender norms in households' neighborhoods, which are positively correlated with their own.

Table 6 presents the results. Column (1) shows that families whose husbands are older than their wives reveal a larger gender effect. In column (2), I find that the gender effect elevates by 3.1% for every extra year of education received by the husband than his wife. To interpret the results in column (3), compare a family whose husband earns twice as much as the wife to another whose income stream is equally split, the gender effect is 0.9% larger in the former case. The coefficients of the interaction terms remain largely stable when I pool them in column (4). Taken together, these results highlight the importance of intrahousehold bargaining in health insurance decisions: the higher leverage a husband has, the higher probability his family is skewed toward his plan.

4.4 Gender Effect and Inertia

The existing literature has provided ample evidence of inertia in individual health insurance decisions. For example, Handel (2013) finds that the plan switching rate for a large employer is only 11%.[23] Dual-earner households benefit from the flexibility to switch plans across employers. However, such freedom will be greatly restrained by the presence of inertia. To

[22] Age gap = Husband age - Wife age. Education gap = Husband education years - Wife education years. Income ratio = Husband income/Wife income.

[23] Households are also found to have inertia in financial decisions. For instance, Choukhmane et al. (2022) show that married couples' non-cooperation in pension accounts is persistent over time. Guiso et al. (2021) find gender parity is more relevant for households who previously suffered from attention deficits.

make matters worse, when interacting with inertia, gender will have a more persistent effect on choices as households carry over their subscriptions from one period to another.

In Table 7, I investigate the contribution of inertia to the gender effect. Starting from the baseline regression, I add in column (2) a binary variable, $Subscriber_{it-1}$, to specify whether a spouse was his/her family's subscriber last year. As expected, households are 61.5% likely to stick to their default subscriber. Moreover, once I control for inertia, the coefficient of $Husband$ reduces to 2.3%, suggesting the gender effect still exists in non-default plans. Taken together, these results point out inertia as a strong channel for the gender effect, but not all.

One might wonder if such inertia is driven by the reluctance to switch subscribers or plans. To further disentangle these two, in column (3), I interact $Subscriber_{it-1}$ with *Job switcher*. This group of households is useful as their default subscriber changed jobs from year $t-1$ to t so their default plan is no longer available. However, 58.5% of the group still end up enrolling with the previous subscriber. Notably, for these job-switcher families, switching costs would apply to whatever plan they choose in year t, so this pattern could mainly reflect the role of inattention in the decision process. That is, households have developed their own 'mindset' on who should be the subscriber of ESI, which is hard to change even when they are given a chance to switch plans. This suggests that not only 'plan inertia' but also 'gender inertia' are crucial to explaining household health insurance choices.

To summarize this section, I find the gender effect is accountable for around 34% of the gender gap in family plan subscriptions, which is partly explained by the distribution of decision power in a marriage. As a result of inattention, plan selection is highly inertial to the 'default' subscriber established by households. These findings motivate the inclusion of inattention into the choice model below.

5 Structural Model

In this section, I introduce a three-stage model of plan choice motivated by the empirical patterns in Section 4. I assume households stay inattentive to their default plan unless their attention is triggered by events such as abnormal health utilization or downgrades in plan characteristics. Once waken up from their default plan, households remain inertial to their default subscriber with some probability and just choose from the corresponding employer. Only households who stay alert from both plan-based and subscriber-based inattention manage to consider alternatives in their entire choice set, which includes two employers. Regardless of how complete their consideration set is, households evaluate each option's expected costs and risks before making a choice.

The model has three features that make it distinct from previous papers. First, compared to studies on employer-sponsored health insurance, my model expands the choice set to two employers for dual-offered households. Second, built on the recent literature on consideration set models [Ho et al. (2017); Coughlin (2019); Heiss et al. (2021); Abaluck and Adams-Prassl (2021)], my model incorporates an extra attention stage from the default subscriber in addition to the default plan. Lastly, I leverage an instrument to explicitly deal with the measurement error of plan costs. I describe these in detail below.

5.1 Wake-up from the Default Plan

$$a_{it}^* = x_{it}\alpha + e_{it} \tag{2}$$

Let a_{it}^* denote a household's latent propensity to pay attention by waking up from its default plan, which is determined by attention triggers x_{it} and an i.i.d. logistic error e_{it}. The attention triggers include health shocks to any family member in the past year and downgrades in the default plan's design. Conditional on having a $t-1$ plan (i.e., $f_{it} = 0$), household i pays attention ($a_{it} = 1$) at time t if $a_{it}^* > 0$ and remains inattentive otherwise ($a_{it} = 0$). Other households who do not have a default plan (i.e., $f_{it} = 1$) always pay attention with $Pr(a_{it} = 1) = 1$. This happens when they are newly married; or their default plan is discontinued. These together produce a closed-form expression for the wake-up probability from the default plan:

$$p_{it}^a \equiv Pr(a_{it} = 1 | f_{it}, x_{it}) = \begin{cases} \frac{1}{1+exp(-x_{it}\alpha)}, & f_{it} = 0 \\ 1, & f_{it} = 1 \end{cases} \tag{3}$$

5.2 Wake-up from the Default Gender

$$g_{it}^* = r_{it}\omega + \upsilon_{it} \tag{4}$$

After a household wakes up from its default plan, it may still rest with gender inertia and only choose plans offered by its default subscriber. Here, I model the attention to both spouses, essentially the entire choice set, in a similar fashion as Section 5.1.[24] Denote g_{it}^* as the latent propensity to wake up from the default subscriber, which is allowed to vary by gender r_{it}. Likewise, g_{it} is a dummy variable that specifies whether household i pays attention ($g_{it} = 1$) or not ($g_{it} = 0$). Conditional on having a default gender (i.e., $s_{it} = 0$), households wake up to consider all options if $g_{it}^* > 0$. On the other hand, newly married

[24] An alternative way is to model households paying attention to 1) only the husband's; 2) only the wife's; 3) or both plans with exclusive probabilities. However, 1) I show in Section 4 that path-dependencies on gender are important in my setting; 2) the current model better fits the data.

couples do not have a default subscriber ($s_{it} = 1$) and are forced to pay attention. Assuming i.i.d. logistic errors for v_{it}, the wake-up probability from the default gender is as follows:

$$p^g_{it} \equiv Pr(g_{it} = 1|s_{it}, r_{it}) = \begin{cases} \frac{1}{1+exp(-r_{it}\omega)}, & s_{it} = 0 \\ 1, & s_{it} = 1 \end{cases} \tag{5}$$

5.3 Plan Choice Stage

$$H_{it} = \begin{cases} y_{it-1j}, & a_{it} = 0 \\ G_{it-1}, & a_{it} = 1, g_{it} = 0 \\ C_{it}, & a_{it} = 1, g_{it} = 1 \end{cases} \tag{6}$$

The above two stages pin down the consideration set H_{it} among which households maximize utilities. When household i stays asleep during the open enrollment period ($a_{it} = 0$), H_{it} is its default plan y_{it-1j}. When the household wakes up from its original plan ($a_{it} = 1$) but is still unaware of another spouse's ESI ($g_{it} = 0$), H_{it} only includes plans from the default gender G_{it-1}. Only households who stay alert in both stages are able to pay attention to their full choice set C_{it}. Given consideration set H_{it}, households compare plans based on their perceived utilities:

$$\begin{aligned} u_{ijt} &= \theta_1 Husband_{ijt} - \theta_2 Cost^*_{ijt} - \theta_3 VarCost^*_{ijt} + \theta_4 Type_{jt} + \eta y_{it-1j} + \epsilon_{ijt} \\ &= \delta^*_{ijt} + \eta y_{it-1j} + \epsilon_{ijt} \end{aligned} \tag{7}$$

$$Cost^*_{ijt} = OOP^*_{ijt} + \tau^*_{jt} Premium^*_{jt} \tag{8}$$

where $Cost^*_{ijt}$ and $VarCost^*_{ijt}$ represent the expectation of costs and the variance, which correspond to the traditional view of insurance as financial products that provide risk protection in exchange for premiums. Cost is further decomposed into out-of-pocket costs (OOP hereafter) and premiums, with τ^*_{jt} denoting the share of premiums contributed by employees. In addition, I include $Type_{jt}$ to control for general differences across plan types that are not captured in the cost measure. For example, HMO plans require patients to get referrals from their primary doctor before visiting a specialist, indicating higher transaction costs.

In addition to monetary attributes, I include y_{it-1j} to mark whether a plan is the default and if yes, households will save η amount of switching costs by sticking to their default plan. Following the reduced-form specification, $Husband_{ijt}$ indicates a plan offered by the husband's employer. Conservatively speaking, this dummy can capture a mixture of the following things: 1) true utility, for example, when households value husbands' job stability; 2) gender bias, when husbands dictate health insurance; 3) any unobserved goodness of the husband's plan not currently incorporated in the model. For simplicity, I refer to this component of the gender effect as *gender preference* to reflect the fact that it is indeed a 'revealed preference' in plan choices. ϵ_{ijt} are i.i.d. type I extreme value idiosyncratic shocks.

5.4 Identification

The identification of the model relies on three distinct groups of households presented in Figure 2 and Table 8. The most straightforward group is newly-married couples. As a couple, they do not have a default plan, nor have they established a default subscriber. As active choosers, their behaviors help identify parameters in the plan choice stage. If comparable to the other households, their preferences can serve as a benchmark, and what is remaining can then be attributed to inattention or switching costs.

The second group of households is job switchers. Being in marriage for years, they have a default subscriber, but their default plan is no longer available when the default subscriber changes jobs. They are forced to wake up from the default plan but are still subject to gender-based inattention. To them, choosing either the husband's or the wife's plan now generates switching costs. Therefore, inattention is the main reason why they remain persistent with their subscriber choice. Assuming they share the same preferences as newly-married couples, their extra inclination toward the default subscriber identifies gender-based inattention.

Conditional on preference parameters and gender-based inattention being identified, the last group of households (passive choosers) helps distinguish switching costs from plan-based inattention. This identification follows Ho et al. (2017) and Heiss et al. (2021). Specifically, I model that *changes* in the default plan trigger attention, whereas the comparison on *levels* of costs across plans pins down the ultimate choice. Exclusion restrictions hold here for two reasons.[25] First, the default plan could garner attention by increasing its deductible by \$500 while still maintaining as the best plan in a household's choice set. Second, even when the default plan becomes less favorable now, the change in deductible is mapped into OOP in a nonlinear and household-specific manner. As a result, I have sufficient independence

[25]Exclusion restrictions are actually not required for identification in my case when unobserved determinants of attention and utility are not modeled as correlated. Abaluck and Adams-Prassl (2021) prove that attention probabilities can be point identified by asymmetries in cross-price derivatives of demand. When consideration depends on preferences, Barseghyan et al. (2021) show that some overlap in variables between consideration and utility does not threaten identification as long as some exclusive variables enter either stage, but not both.

between the attention and the choice stage.

5.5 Estimation

$$\mathcal{L}_i(\theta) = \prod_{t=1}^{T_i} Pr(y_{it} = j_{it}|\theta, d_i, y_{it-1}) \tag{9}$$

I simultaneously estimate all model parameters θ using simulated maximum likelihood in Train (2009). The likelihood function of household i over a sequence of choices is expressed as $\mathcal{L}_i(\theta)$. I assume choices of household i are independent over T_i conditional on their past choices y_{it-1} and observed covariates d_i, so that $\mathcal{L}_i(\theta)$ can be written as the product of probabilities each period in choosing the observed plan j_{it}. Also assuming independence across households, the objective likelihood function is then the product of $\mathcal{L}_i(\theta)$.

$$\begin{aligned} Pr(y_{it} = j_{it}|\theta, d_i, y_{it-1}) =& (1 - p^a_{it})y_{it-1j} + p^a_{it}\Big[p^g_{it}Pr(y_{it} = j|\cdot, C_{it})+ \\ & (1 - p^g_{it})y_{it-1g}Pr(y_{it} = j|\cdot, G_{it-1})\Big] \end{aligned} \tag{10}$$

The choice probability for household i in period t is built on derivations of p^a_{it} from Equation (3), p^g_{it} from Equation (5), and logit probabilities derived below in Equation (14). For newly-married couples, $p^a_{it} = p^g_{it} = 1$; so their choice probability is the same as $Pr(y_{it} = j|\cdot, C_{it})$. For job-switchers, $p^a_{it} = 1$; and their choice probability depends on whether the plan comes from the default gender y_{it-1g}. If yes, the plan will be enrolled more often when households experience gender-based inattention with probability $1 - p^g_{it}$. Passive choosers are further likelier to enroll in their default plan y_{it-1j} when they suffer from plan-based inattention with probability $1 - p^a_{it}$.

6 Cost Model

A crucial input into the structural model is plan costs. However, they are not fully observed for three reasons. First, only the chosen plan's costs are recorded, not other alternatives. Second, only ex-post spending is observed, which is a single realization of a household's ex-ante beliefs on health risks. Third, plan characteristics are not directly given in the data.

In this section, I describe the steps to approximate plan costs. I start by aggregating cost-sharing features from the claims data. Next, I assign each individual to a group of similar enrollees whose empirical distribution of spending is used to project OOP for every option in a household's choice set via a machine learning model (ML hereafter). At the end of this section, I introduce a control function approach to address the endogeneity and measurement error of plan costs. More details are disclosed in Appendix 10.2 and 10.3.

6.1 Construction of Plan Costs

Plan Characteristics My first goal is to build a dataset on the financial features of health plans. Unfortunately, the data does not directly provide plan characteristics except that I am able to merge small group plans after 2016 with CMS Health Plan Finder. On the bright side, the data records claim-level cost sharing. For example, for a physician visit, I observe the amount billed by the provider, reimbursed by the insurer, and paid out-of-pocket by the patient in the form of deductibles, copayments, and coinsurance. The sum of the insurer and patient payments reflects the negotiated price received by the provider, often called the allowed amount.

Extracting plan characteristics from insurance claims is an involved process; I discuss the details in Appendix 10.2. The goal is to aggregate the deductible, coinsurance rate, maximum out-of-pocket payment (MOOP), and copayments from a plan's claims. To do this, I take out the most frequent copay value for primary doctor and specialist visits. For the other three parameters, I fit a classic insurance line on a scatterplot of individual cumulative spending and OOP costs. The deductible-coinsurance-MOOP triplet that generates the smallest fitted error is taken as the plan's financial characteristics. Eventually, I am able to recover the menu for 56% of the plans in the data.

Out-of-pocket Costs The cost metric for an insurance plan involves two parts, premiums and OOP expenses. For the latter, I approximate how households form expectations on their OOP spending based on a plan's cost-sharing and their beliefs in future health risks. To assemble households' beliefs, I follow Abaluck and Gruber (2011) and Handel (2013) by grouping individuals into $1,048$ cells characterized by age, year, chronic conditions, and past spending levels. For each cell, I draw 500 individuals and take their realized claims as the ex-ante distribution for persons assigned in this cell. The claims for each draw are further divided into nine categories, each contributing uniquely to OOP costs.[26]

My next task is to adjust the billed amount for each claim by a negotiated price ratio PR_{sndt}, which is the average of the ratio of the allowed amount over billed amount across all claims concerning insurer s, provider n, procedure d, and year t. This adjustment comes from the observation that there is substantial price variation across providers and insurers, even across services [Clemens and Gottlieb (2017); Clemens et al. (2017); Cooper et al. (2019)].[27]

[26] The nine categories include primary doctor visits, specialist visits, inpatient stays, outpatient surgeries, emergency room visits, laboratory tests, x-rays, CT scans, and other services.

[27] For example, Cooper et al. (2019) document that around 1/5 of the total variation in hospital prices

As a consequence, before applying any cost-sharing rule, households may face different prices depending on which insurer they enrolled in.

As the last step, I exploit machine learning (Gradient Boosting) to predict individuals' OOP costs, given their category-specific spending, plan characteristics, and their two-way interactions. Overall, my machine learning model outperforms a naive approach and is robust to different algorithms. For example, the naive approach has a root mean square error (RMSE hereafter) of 529 and overpredicts OOP by about \$100. The ML model reduces RMSE to 401 and preserves the mean. With ML predictions, I then aggregate individual OOP distributions into family cost projections. Eventually, they are used to compute the mean and variance of family OOP, which are crucial inputs into the choice model. I describe in depth how I construct these measures from raw claims in Appendix 10.3.

Premiums The data also poses two challenges to recovering the true premium costs faced by households. The first is that premium information is only disclosed by about 18% New Hampshire-situs plans since 2016. The second is that only total premiums are reported, instead of the share contributed by households. According to MEPS, an average household in New Hampshire is responsible for about 26.5% of family premiums from 2011 to 2021. I thus take 26.5% as the uniform division rule while acknowledging that employers may differ in their generosity of premium contributions. Nevertheless, employees in New Hampshire do not seem to experience substantial variation in their share of premium contributions, with a standard deviation of 1.6%.

To manage the first challenge, I again leverage machine learning to predict total premiums. I focus on large groups since insurance pricing is more tailored to their risk profiles. After researching a few insurers' rating formulae, I decide to fuel the ML model with the following factors and their two-way interactions: plan characteristics; employer size and location; the distribution of age, gender, and enrollment tiers; and the group's claims history. The model with the best performance has a 0.94 correlation with observed premiums and captures 88% of its variation.

Concerns I acknowledge a few assumptions behind my cost model. First, I assume households have no private information, and that my empirical approximation truly

occurs within hospitals for the same procedure, 70% of which is driven by insurer-hospital bargaining. Hospital prices have a higher correlation within service lines than across. For physician services, Clemens and Gottlieb (2017) and Clemens et al. (2017) find that large physician groups tend to engage in service-specific bargaining. I acknowledge that not all providers adopt reimbursements as a share of total charges. As a matter of fact, Clemens et al. (2017) find that 75% physician payments follow the Medicare prospective payment schedule adjusted by a private scaling factor. This number is less than 57% for inpatient payments [Cooper et al. (2019)]. Nonetheless, my approach offers a parsimonious way to capture the differences in transaction prices across insurers without incorporating the Medicare base rates. The control function approach also deals with any measurement error from this simplification.

represents their ex-ante beliefs on health risks. Grouping individuals into delicate cells alleviates this concern, as it fully incorporates all possible diagnosed conditions from the all-payer claims data. Another way to address this is to assume households have perfect foresight and project OOP costs from their realized claims. Whatever the true theory is, I present in Section 7.3 that different models on household beliefs do not change results markedly.

The second assumption is no moral hazard. That is, health utilization does not depend on plan generosity. Chandra et al. (2010) summarizes the price elasticity of health care to be around -0.1 to -0.4. Given that this magnitude is mild, and considering households have to be reasonably savvy to 'predict' their moral hazard during the enrollment period, it is considered a fair assumption at least for my baseline model. Nevertheless, I show in Section 7.3 that model estimates are robust when allowing for some degree of moral hazard.

Finally, my approach is conditional on precise plan characteristics, accurate mapping on OOP costs, and true premium contributions by households, none of which is realistic due to my data constraints. I solve these by proposing the negotiated price ratio as an instrument for plan costs, which drives up the transaction price of health care but is orthogonal to the measurement error of OOP and premium expenses. I explain this in detail below.

6.2 Measurement Error

$$Cost_{ijt} = Cost^{*}_{ijt} + \mu_{ijt} \tag{11}$$

$$Cost_{ijt} = \pi_0 + \pi_1 IV_{ijt} + \pi_2 Husband_{ijt} + \pi_3 Type_{jt} + \nu_{ijt} \tag{12}$$

$$u_{ijt} = \delta_{ijt} + \eta y_{it-1j} + \lambda f(\nu_{ijt}) + \sigma \rho_{ijt} + \epsilon_{ijt} \tag{13}$$

As discussed above, $Cost_{ijt}$ is measured with error μ_{ijt}. I address this by using a control function approach following Petrin and Train (2010), Wooldridge (2015), and Hahn and Ridder (2017).[28] As the first step, I run a linear projection of $Cost_{ijt}$ on an instrument that affects plan j's utility exclusively through $Cost_{ijt}$. In the second step, I plug $f(\nu_{ijt})$, a function of the residuals from Model (12), into Model (7).[29]

The instrument that I use is the average reimbursement discounts received by neighboring households who live in the same zip code and subscribe to the same insurer as the focal

[28]The control function approach has also been used in empirical studies on health insurance such as Polyakova (2016) and Rickert (2022).

[29]$f(\nu_{ijt})$ also addresses the measurement error of $VarCost_{ijt}$, which can be considered as a squared term based on $Cost_{ijt}$.

household.[30] Recall in Section 6.1 that claim charges are firstly adjusted by negotiated discounts before applying a plan's cost-sharing rules. These discounts can thus be viewed as cost shifters that patients do not directly observe. On the other hand, households that live close tend to use the same set of providers, since distance is greatly valued in health care. Therefore, when neighboring households enjoy larger discounts with a particular insurer, it is likely that the focal household also benefits from it. Indeed, my first stage has an F-value of $4,270$ with the instrument being statistically significant at the 1% level. Notably, the strong first stage is not necessarily a mechanical outcome, as claim-level discounts are applied in computing OOP costs, but the aggregate ones are used as the instrument.

Beyond the relevance and exclusion restrictions, one additional assumption is IV_{ijt} is independent of measurement error μ_{ijt}. Another assumption requires me to specify the distribution of μ_{ijt} conditional on $f(\nu_{ijt})$.[31] Here I assume μ_{ijt} and $f(\nu_{ijt})$ are jointly normal for each j and i.i.d. over j. Then the conditional distribution $\theta_2\mu_{ijt}$ given $f(\nu_{ijt})$ is also normal with mean $\lambda f(\nu_{ijt})$ and variance σ^2. I therefore decompose the measurement error in Equation (13) into a shift in mean $\lambda f(\nu_{ijt})$ and a standard normal term ρ_{ijt} whose standard deviation σ is to be estimated. Since ρ_{ijt} is independent of the idiosyncratic term ϵ_{ijt}, I can write p_{ijt}, which is the choice probability for plan j, in a mixed logit form that integrates over the distribution of ρ_{ijt}:

$$p_{ijt} \equiv Pr(y_{it} = j | \delta_{ijt}, y_{it-1j}, \nu_{ijt}, H_{it}) = \int \frac{exp\left[\delta_{ijt} + \eta y_{it-1j} + \lambda f(\nu_{ijt}) + \sigma\rho_{ijt}\right]}{\sum_{k\in H_{it}} exp\left[\delta_{ikt} + \eta y_{it-1k} + \lambda f(\nu_{ikt}) + \sigma\rho_{ikt}\right]} d\Phi(\rho) \tag{14}$$

6.3 Endogeneity

The IV approach also allows for the case when $Cost_{ijt}$ is endogenous due to its correlation with unobserved plan characteristics [Hahn and Ridder (2017)],[32] which, when correlated with gender, would cause $Husband_{ijt}$ to capture the unobserved goodness of husband plans beyond the gender effect. Though not perfect, I propose a few ways to address this concern.

First, network information is already embedded in my construction of costs. To see this, recall that the cost measure was adjusted by negotiated price discounts specific to

[30] In case of no neighboring households, I enlarge the geographical area to be at the city or county level.

[31] Hahn and Ridder (2017) allows the measurement error to be correlated with the true value. What's more, in my setting where only the difference in utility matters for plan choice, measurement error also needs not to be mean zero.

[32] As a matter of fact, non-monetary attributes such as the deductible and coinsurance rate are often directly included into the choice model beyond their financial impact on plan costs [e.g., Abaluck and Gruber (2011); Abaluck and Gruber (2016); Heiss et al. (2021)]. This, however, contradicts the traditional view of insurance as contracts that simply transfer wealth for risk protection [Coughlin (2019)].

an insurer and provider. If a plan covers a particular provider, it will then be reflected as lower negotiated prices of that provider. The claim draws from similar enrollees thus approximate the process of drawing providers with varying quality, which, depending on a plan's bargaining leverage in and out of its network, can be picked up by the distribution of OOP costs. Second, argued in Coughlin (2019), the reason why plan attributes affect choices beyond their financial consequences is through altering consideration probabilities. I have modeled changes in these features into the wake-up probabilities from the default plan.[33]

The third point is related to the caveat of my uniform premium share measure. If husband employers are more generous in premium contributions, it will also enter into the estimated impact of $Husband_{ijt}$. I examine this by extracting from MEPS the distribution of premium contribution ratio stratified by proportions of female employees in Figure A1. It is unclear whether male-dominated employers are more generous, since the confidence intervals overlap overall. Nevertheless, I show in Section 7.3 that my results are robust to interacting premiums with employer size, which is a strong predictor of employer generosity. Still, having not completely ruled out these possibilities, I remain agnostic about what $Husband_{ijt}$ truly captures in my interpretation of the empirical results below.

7 Results

In this section, I present the main results of the structural model. I first provide descriptive statistics on the estimation sample, including the source of variation for the identification of the model. Then I show parameter estimates by progressively adding more stages to the model. Finally, I inspect the sensitivity of my results by relaxing some modeling assumptions.

7.1 Descriptive Statistics

In Table 9, I present summary statistics on households in my estimation sample whose choice environments can be backed out from the claims data. Corresponding to the identification strategy, I group them into newly-married couples, job switchers, and passive choosers. The top panel contains their demographics following Table 2. It is not surprising that newly married couples are younger and healthier than the other two groups. The proxies for intra-household bargaining are comparable for the three groups, suggesting one can be used to identify preference parameters for another. A final observation is that passive choosers have the highest propensity to enroll in the husband's ESI, with the second being job switchers,

[33] I exploit *changes* in plan attributes from the default plan (as in Ho et al. (2017) and Heiss et al. (2021)), rather than the *level* of attributes from all alternatives (as in Coughlin (2019) and Abaluck and Adams-Prassl (2021)) in order to minimize the amount of measurement error entered into the attention stage.

which confirms the mechanism in Table 7 that the gender effect amplifies with more sources of inertia.

Following Ho et al. (2017) and Heiss et al. (2021), I include two sets of attention triggers from the default plan: (i) changes in plan features from year $t-1$ to t; (ii) health shocks to any family member in year $t-1$. Since newly-married couples and job switchers do not have a default plan, I assume their attention probabilities in this stage are always one. For passive choosers, around 11% households faced an increase in deductible and 7% in copayment for a primary doctor visit.[34] On average, about 8% households hit deductible in the past year with more having at least one member visited the emergency room or had an inpatient stay.

The bottom panel lists the distribution of plan characteristics constructed in Section 6. An average household in my sample pays out-of-pocket around $\$3,600$ on premiums and $\$1,700$ on medical bills. Overall, the three groups have balanced choice sets, though job switchers tend to pay higher premiums and OOP costs. These cost variables are of course measured with error, which will be addressed with a control function in estimation.

Table 10 displays the shares of households who kept (switched) their default subscriber (plan) during the enrollment process. Among job-switchers, 6% couples switched to another spouse's ESI when the previous subscriber changed jobs. The rest 94% is crucial to identifying the inattention to the default subscriber, holding preferences fixed. For passive choosers, around 5.7% households switched plans within the same employer, which additionally contributes to identifying gender-based inattention. Another 8.9% households switched to the spouse's plan, overcoming inattention and switching costs.

7.2 Main Estimation Results

In Table 11, I present five sets of results, with each built on the previous model. In column (1), I run a conditional logit model of plan choice. In column (2), I exploit a control function to address the measurement error of plan costs. I gradually add the attention stage from the default plan in column (3) and the default gender in column (4). In column (5), I further allow intrahousehold bargaining to affect the gender effect estimate in the plan choice stage.[35]

The explanatory power of each model is given in log-likelihoods, with model (5) having the best fit. Overall, the model fits well on households' observed choices, which is detailed in Section 10.5. Also listed at the bottom of Table 11 is the shadow price of *gender preference*

[34] I acknowledge that premium increases are also an important attention trigger. However, since the premium data is sparse with the majority estimated using machine learning, I refrain from including it here to avoid measurement error.

[35] I also tried to include unobserved heterogeneity in the gender effect both in the attention and choice stages and was able to get point estimates. However, it is hard to get sensible bootstrap standard errors even using Bayesian weighting techniques, since some draws lose observations that may be crucial for identification.

revealed in the choice stage, which is the coefficient of *Husband* divided by the coefficient of *Cost*. Similarly, the coefficient of *Default Plan* can be interpreted relative to plan costs as switching costs. I also derived the average inattention propensity to the default plan across households who make passive choices, and their inattention to the default gender along with job switchers.

Comparing columns (1) and (2), the control functions alleviate the attenuation bias of the cost coefficient due to the measurement error, which explodes the dollar estimates to an unrealistic amount. Incorporating plan-based inattention stage significantly reduces switching costs from $\$3,842$ to $\$1,442$, and further including gender-based inattention to \$538. This is because inattention was mistakenly attributed to switching costs in models without an attention stage [Heiss et al. (2021); Abaluck and Adams-Prassl (2021)].

Regarding attention probabilities, households stay inattentive to their default plan with a mean probability of 58%. In addition, they are 61% inattentive to the husband subscriber versus only 4.4% less to the wife. Attention triggers used in the model fall into two groups. The first is changes in plan features from year $t-1$ to t, which are often communicated to households in the form of Plan Annual Notice of Change (ANOC) letters. The second group includes health shocks in the past year.

I interpret the magnitude of these attention triggers by their average marginal effects on the attention probabilities based on the results in column (5). The most salient attention trigger is an increase in copay from the default plan, which boosts the wake-up probability by 12%. Similarly, an increase in deductible raises awareness by 4%. To explain why copay increases trigger more attention than deductibles, Stockley (2016) finds that consumers are more responsive to copay OOP than deductible OOP, though she did not separate out the attention mechanism. Another reason could be that copays are better measured from the claims data. Health events such as hitting the deductible, having an inpatient stay, or visiting the emergency department also raise attention by less than 6%.

For estimates in the plan choice stage, I interpret them by their average marginal effects on choice probabilities conditional on full attention. Specifically, raising a plan's annual cost by $\$1,000$ reduces its choice probability by 14%, which is about a 52% decrease given that an average household chooses from 3.7 plans. The shadow price of gender preference is estimated to be \$188 each year, which again can be caused by many factors such as job stability, unobserved quality, and gender bias. To test on the last story, in column (5), I further interact *Husband* with proxies for intrahousehold bargaining. I find that comparing a family whose husband earns twice as much as the wife to another whose income is evenly distributed, the former exhibits a 1.7% higher propensity to enroll in the husband's plan, in addition to the 2.6% baseline. Apart from inattention, switching costs alone cause a family to be 8.5% more likely to stick to the default plan. Households also appear to be risk averse to plans with a high variance of costs, though it becomes insignificant once I add gender-based inattention. As a matter of fact, Abaluck and Gruber (2011) show that absolute risk aversion can be expressed as $\frac{2\beta_{Var}}{\beta_{Cost}}$. Taking estimates from column (3), the estimated risk aversion is 6×10^{-5}. For models (3)-(5), I calibrate the standard deviation of the measurement error to be \$1,462 estimated in model (2) with the assumption that its true distribution will not be affected by the model specifications.[36]

7.3 Robustness

In Table 12, I perform various robustness checks based on the baseline in Table 11 column (4). Recall that I took an ex-ante perspective in constructing OOP costs in Section 6. This approach assumes that I can approximate all the information a person has from the empirical distribution of expenditures among enrollees with similar health risks. Individuals, however, can know more than I predict, which can be reflected in their ex-post spending. I therefore take their realized claims as fixed and then map them to OOP costs. This alternative approach incorporates private information yet at the cost of assuming individuals have perfect foresight. In column (1), I replace ex-ante OOP with ex-post ones and find the estimates to be generally comparable to the baseline.

The ex-ante approach also assumed no moral hazard. That is, health utilization does not depend on plan generosity. To test how results are sensitive to this assumption, in column (2), I divide plans into three groups based on their actuarial value calculated from the claims data. I then adjust OOP to be 10% higher than their predicted values for the most generous plans and 10% less for the least. Estimates of this model are also similar to the baseline. In addition, I also show robustness to alternative machine learning algorithms in columns (3) (Elastic Net) and (4) (Random Forest).

Another concern of my model is the assumption on the uniform share of premium contributions across employers, which neglects the possibility that the husband's employer could be more generous (i.e., lower τ). First, note that the reduced-form analysis is still unaffected, since households with the same choice set will also face the same premium contribution rules. This, however, could lead to a biased interpretation of the coefficient of *Husband* as gender preference per se. I address this by allowing plan costs to interact with an employer's size in column (5).[37] This is because employer size is a strong predictor of premium contribution generosity [KFF (2020)]. Though it shows significantly positive, the

[36] This is because it becomes insignificant when jointly estimated with other parameters in my models with additional attention stages, which can affect counterfactual simulations afterward.

[37] Size is measured as the logarithm of the number of single employees.

gender preference estimate remains stable.

8 Counterfactual Simulations

The structural model points out a few sources for the aggregate gender effect that I found with my reduced-form design: 1) gender preference on the husband's ESI, 2) switching costs, 3) inattention to the default plan, and 4) gender-based inattention. In this section, I simulate counterfactuals that aim to eliminate the 'inefficient' components among the above sources to understand their contribution to the gender effect and their welfare implications. At this stage, I refrain from removing the 'gender preference' channel, since it is potentially commingled with explanations such as unobserved premium advantages of the husband's plan. For the same reason, I am also reluctant to remove switching costs, which can suggest psychological costs, paperwork costs, and acclimation costs, each has different implications on welfare. Rather, I will focus on the two channels of inattention and show how inattention alone can explain a sizable fraction of overspending in ESI. Certainly, these experiments are not unrealistic. Removing plan-based inattention echoes policies that raise consumer attention through personal reminders or media campaigns [e.g., Kiss (2015)].

Resolving gender-based inattention is another option. After households make their initial decisions, they are constrained by the norms on 'who wears the pants' in health insurance subscriptions. Consequently, families won't adjust their choices flexibly in the following years in response to changes in their health conditions or the choice environment. Motivated by this, I run a counterfactual simulation by forcing households to pay attention to both spouses while still allowing them to be inattentive to their default plan. This simulation can also have implications in real life. For instance, policymakers could launch household-level enrollment campaigns to raise awareness of options coming from both the husband's and wife's employers. A regulation to uniform the open enrollment period across employers will have the benefit that when households start to pay attention, they know they should pay attention to their entire choice set.

$$u^N_{ijt}(\rho) = \delta_{ijt} - \theta_1 Husband_{ij} + \lambda f(\nu_{ijt}) + \sigma\rho_{ijt} \tag{15}$$

$$FW_{ijt}(\rho) = \frac{1}{\theta_2}(u^*_{it}(\rho) - u^N_{ijt}(\rho)) \tag{16}$$

$$E(FW_{it}) = \int \sum_{j\in C_{it}} FW_{ijt}(\rho)p_{ijt}(\rho)d\Phi(\rho) \tag{17}$$

$$E(Husband_{it}) = \int \sum_{j \in C_{it}} Husband_{ijt} p_{ijt}(\rho) d\Phi(\rho) \tag{18}$$

To evaluate these counterfactual interventions, I propose two metrics. The first is average foregone welfare across households. Unlike total costs, a household's normative utility [u^N_{ijt}] used here takes into account risk protections and network quality while omitting 'gender preference' and switching costs if one believes they only matter heuristically [Abaluck and Gruber (2016)]. Let $u^*_{it}(\rho) = max_{j \in C_{it}} u^N_{ijt}(\rho)$, the foregone welfare when plan j is chosen is expressed by its difference in u^N_{ijt} with the utility-maximized plan normalized by the cost coefficient [Equation (16)]. The average foregone welfare for household i in year t is then weighted by choice probabilities over the distribution of measurement error [Equation (17)]. In practice, I simulate p_{ijt} to be discrete to allow for a dynamic impact of past choices on current ones (through inattention or switching costs) before averaging over 500 decision paths for each household. To study the gender effect, I compute the average percentage of households that enroll in the husband's ESI in a similar way by interacting the simulated choice probabilities with a dummy indicating the husband's plan [Equation (18)].

In Table 13, I present the average shares of husband subscriptions and foregone welfare under four scenarios: the baseline [C0], removing gender-based inattention [C1], removing plan-based inattention [C2], and removing both sources of inattention [C3], which is essentially forcing full attention to all options in a household's choice set. Among these, the most salient single contributor to the gender gap is gender-based inattention (2.2%), with plan-based inattention also explaining a 1.1%. Removing both sources of inattention has a reinforced effect that eliminates most of the gap.[38] Overall, these counterfactual simulations imply that inattention alone could explain most of the gender gap, with gender-based inattention being the main driver.

Turning to the second column, my estimates suggest that an average household leaves \$669 dollar equivalent of welfare on the table each year.[39] Among the counterfactual scenarios, removing gender-based inattention can save \$72 (11%) from the baseline, though smaller than the magnitude of withdrawing plan-based inattention (\$116). Benefiting from the reinforced effect, eliminating both sources of inattention can achieve sizable reductions

[38] This might seem contradictory to my deduced-form decomposition that plan generosity and gender effect play an equal role in explaining the gender gap. However, the estimation sample used here is limited to employees in relatively large employers whose plan characteristics can be traced out from the claims data. Because of this, the plan generosity gap here may not be as large as the reduced form. Nevertheless, the goal of this exercise is to help us better understand the gender effect.

[39] Compared to the literature, Handel (2013) finds the average money lost in a large employer ranges from \$300 to \$450. Bhargava et al. (2017) finds an average single employee could have saved \$352 – \$375. My estimate here is different for two reasons: 1) my measure is for households with at least two adult members; 2) my choice set involves two employers.

of \$227, which is around 34% of the baseline.[40] In general, this analysis sheds light on the importance of incorporating gender-based inattention in doubling the welfare savings for policies that fight plan-based inattention. For example, employers could add a note in their annual enrollment reminders nudging households to pay attention to both spouses who could be potentially covered by ESI. This policy, if working effectively, can save 6.8 billion dollars from around 30 million dual-working households in the United States each year.

In Figure 3, I plot the distribution of changes in foregone welfare for the scenario that forces full attention [C3] relative to the baseline. Though the average saving is around \$230, a significant number of households can benefit from \$500 to \$1,000 with some as much as \$2,000. Intuitively, these households were hurt the most because they had set the 'wrong' defaults. In Figure 4, I plot the relationship between the welfare change and husbands' and wives' cost differences among households having husbands as their default subscribers. Not surprisingly, welfare savings increase monotonically with the cost differences, suggesting households who have the least generous husbands' plans could have saved the most. Similar patterns are also found with premium differences in Figure A2. Though quite intuitive, these analyses provide an easy criterion to filter out the vulnerable population for choice inconsistencies. That is, households could simply compare premium differences between two employers to decide their need for correcting inattention.

9 Conclusion

In this paper, I utilize insurance enrollment data to document households' higher propensity to take up the husband's benefit plans. My identification strategy relies on comparing families with the same choice set but only differ in whether the husband or wife works for a particular employer. I find a pure gender effect causes households to enroll in the husband's plan 7% more, which explains a sizable fraction of the gender gap in family plan subscriptions in the United States. Though I took a conservative stance on what the gender effect truly represents, my analysis that associates it with intrahousehold bargaining power suggests at least part of it could be resolved through better coordination within the household. A natural direction for future work is to investigate the microfoundations of the gender effect by supplementing detailed household survey data.

The structural model sought to separate out inattention as a source of inefficiency in the decision-making process. In addition to plan-based inattention as frequently documented in

[40] This result can be compared to Model II in Heiss et al. (2021), where an individual saves \$46 with forced attention in Medicare Part D. Abaluck and Gruber (2011) finds inertia explains less than one-fourth of foregone welfare.

the literature, households also exhibit a substantial amount of gender-based inattention. As a matter of fact, counterfactual simulations suggest removing gender-based inattention can double the welfare savings from removing plan-based inattention alone. Such policy could be designed to increase the exposure of ESI offerings from the neglected spouses, for example, through employer reminders and media campaigns. Another relevant policy is to uniform the open enrollment period across employers so that households expect to pay attention to their entire choice set at the same time each year.

This study speaks more generally to decision-aid tools used to help consumers make better insurance choices. Specifically, I stress that these policies should be implemented at the household level rather than the employer level to benefit from the complete choice set for dual-offered households. What's more, counterfactual exercises underline the specific population to target: households who have deep-rooted stereotypes regarding 'who wears the pants' on insurance subscriptions but have been suffering from high costs with their defaults. A takeaway from this is, for example, a website that provides personalized decision support could filter the need by asking whether the two spouses decide on benefit plans together.

Admittedly, my analyses have several caveats. First, I did not model the general equilibrium in which premiums adjust with households' enrollment decisions. Such features are important to consider as gender-neutral enrollment would imply a higher premium contribution for employers with more female workers. On the other hand, male-dominated employers would face smaller insurance pools and higher premiums, which causes households to favor the wife's ESI even more. Female-dominated employers could adjust their premium contribution as a response, thus canceling out the welfare improvement of this policy.

The second shortcoming of my analysis is I did not observe the spouse's employer but imputed it from her enrollment records. Future studies could better build households' choice sets from, for instance, tax filings data. Thirdly, I did not incorporate the unobserved heterogeneity of the gender effect in my model due to a lack of variation for estimation. I therefore leave it to future research to better model households' circumstances by potentially linking to survey data. Lastly, the claims data does not directly provide plan characteristics, nor does it contain households' share of premium contributions. With the development of all-payer claims data, it is countable that researchers can better approximate plan costs without the need to deal with measurement error. Despite these caveats, my work opens up possibilities for research exploring intra-household dynamics in health insurance choice. With more and more women joining the labor force,[41] research in this strand will benefit an increasing number of families.

[41] Women's labor force participation rate increased from 33.4% in 1949 to 57.6% in 2019 [U.S. Bureau of Labor Statistics].

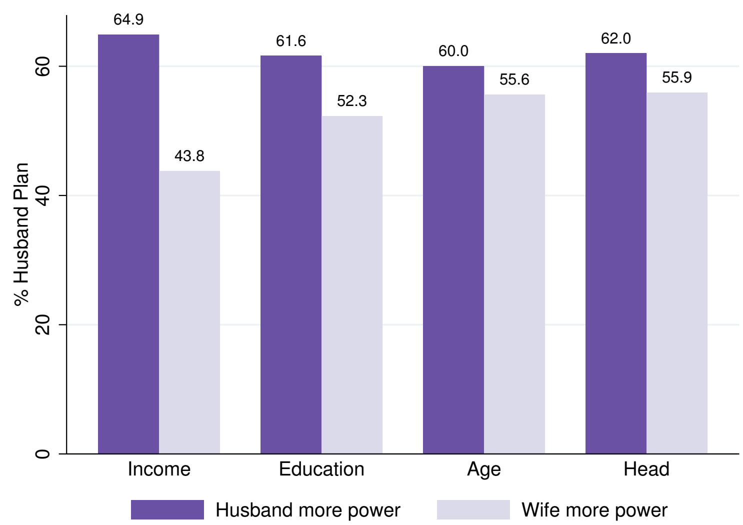


Figure 1: Health Plan Choice and Household Bargaining Power

Notes: Statistics are calculated from MEPS. Households are weighted by FAMWT to form a nationally-representative sample.

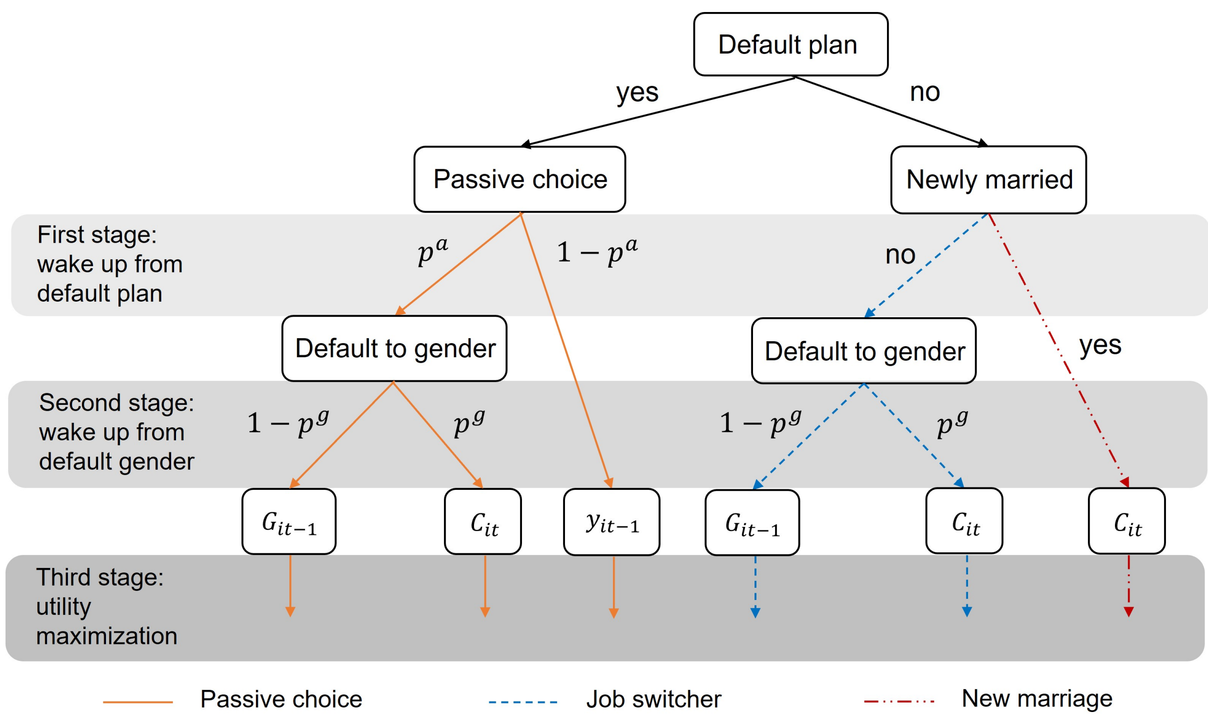


Figure 2: Stages of the Model

Notes: New marriage refers to couples who get married for the first time. Job switchers are couples whose previous subscribers changed jobs. Passive choosers are couples with a valid default plan.

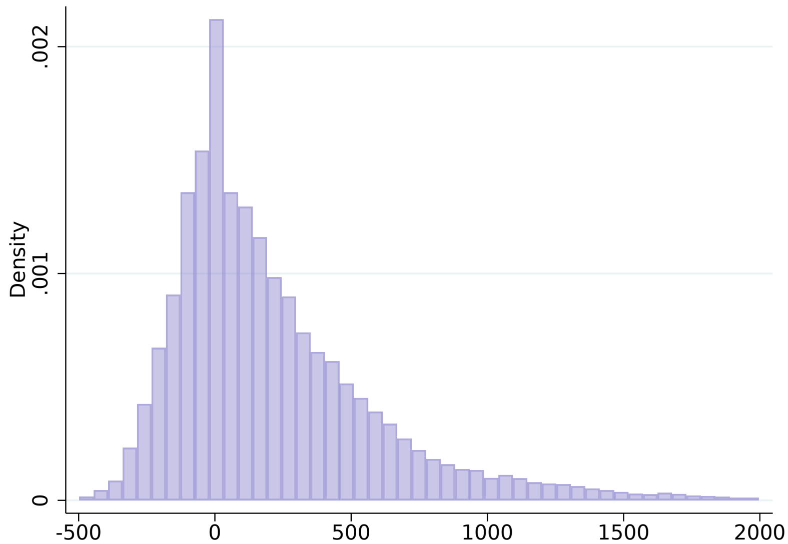


Figure 3: Distribution of Change in Foregone Welfare ($): Full Attention Scenario

Notes: This graph plots the distribution of change in foregone welfare for a counterfactual scenario that assumes full attention (C3 in Table 13) relative to the baseline. Welfare is defined as total costs plus the dollar equivalent value of risk protection and network qualities estimated in Table 11 column (4). Foregone welfare is computed as the welfare difference between the chosen plan and the welfare-maximizing plan.

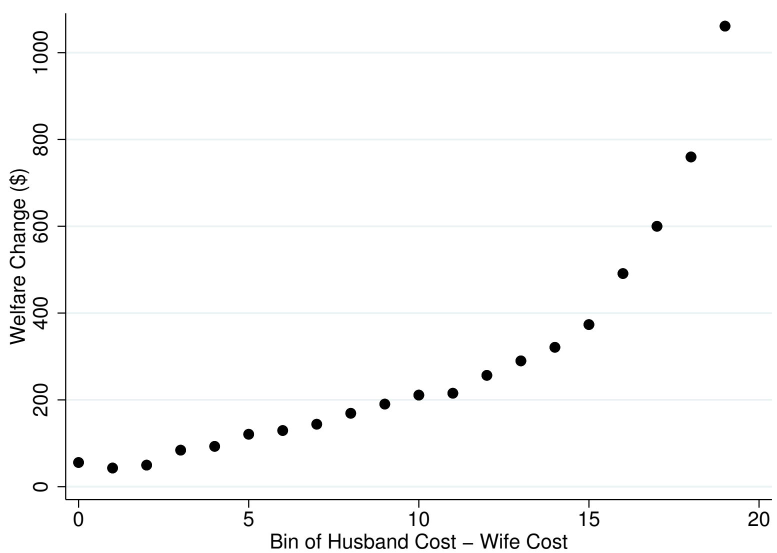


Figure 4: Change in Foregone Welfare ($) by Cost Difference: Full Attention Scenario

Notes: This graph plots the average change in foregone welfare among households who have their husbands as the default subscriber for a counterfactual scenario that assumes full attention (C3 in Table 13) relative to the baseline by bins of the cost differences within the household. The cost differences are defined by the total costs of the least-expensive plan offered by the husband minus that of the wife. A higher value indicates that the husband's option is more pricey than the wife's.

Table 1: Descriptive Evidence on Gender Gap

Husband plan %	MEPS	Northeast	NH	NH Gap
At least one ESI	66.3%	62.5%	60.7%	21.4%
Both ESI	59.0%	53.7%	56.4%	12.8%

Notes: Statistics in the first two columns are derived from MEPS. Households are weighted by FAMWT to form a nationally-representative sample. Statistics in the last two columns are calculated from New Hampshire All-payer Claims Data.

Table 2: Summary Statistics: Reduced-Form Sample

	Full Sample	Colleague	Non-colleague	Non-colleague FE
Husband	60.7	62.4	56.4	57
Age	46.4	51.7	44.6	45.1
Health risk	0.3	0.5	0.3	0.3
Age gap	1.8	2.1	1.8	1.7
Income ratio (ACS)	1.5	1.5	1.5	1.5
Education gap (ACS)	-0.1	-0.1	-0.1	-0.1
# Couple	319,632	11,765	47,550	13,841
# Couple-year	1,178,311	29,138	128,404	37,118

Notes: The unit of observation is a couple-year. Age measures the average age between husbands and wives. Health risk is the Charlson Comorbidity Index, with a higher value indicating larger health risks. Age gap = Husband age - Wife age. Income ratio = Median male income/Median female income at the zip code level. Education gap = Average male education years - Average female education years at the zip code level. Both income and education data come from 5-year estimates of the American Community Survey (ACS).

Table 3: Identification Strategy: Reduced Form

	E1	E2
H1	F	M
H2	M	F

(a) Case A

	E1	E2
H1	M	F
H2	M	F

(b) Case B

Notes: A row is a household, and a column is an employer. M denotes husband, and F denotes wife.

Table 4: Gender Effect: Baseline Results

	(1)	(2)	(3)
Husband	0.127***	0.0713***	0.0708***
	(0.0120)	(0.00974)	(0.00953)
N	89308	89308	89308
Adj R2	-0.0365	0.442	0.445
Choice Set FE	X	X	X
Employer-Choice Set FE		X	X
Employer-Year FE			X

Notes: *** Significant at the 1% level. ** Significant at the 5% level. * Significant at the 10% level. Standard errors are shown in parentheses and cluster at the choice set level.

Table 5: Decomposition of Gender Gap

Source	Gap	Share
Employment	8.6%	40.2%
Plan generosity	5.6%	26.3%
Gender effect	7.2%	33.5%

Notes: Based on estimates from Table 1 and 4.

Table 6: Gender Effect: Household Bargaining

	(1)	(2)	(3)	(4)
Husband	0.0676***	0.0754***	0.0596***	0.0640***
	(0.00976)	(0.00979)	(0.0113)	(0.0121)
Husband × Age gap	0.00181**			0.00185**
	(0.000885)			(0.000883)
Husband × Education gap		0.0309***		0.0279**
		(0.0119)		(0.0122)
Husband × Log(Income ratio)			0.0295*	0.0201
			(0.0174)	(0.0178)
N	89308	89308	89308	89308
Adj R2	0.445	0.445	0.445	0.445

Notes: *** Significant at the 1% level. ** Significant at the 5% level. * Significant at the 10% level. Standard errors are shown in parentheses and cluster at the choice set level. Age gap = Husband age - Wife age. Income ratio = Median male income/Median female income at the zip code level. Education gap = Average male education years - Average female education years at the zip code level. Both income and education data come from 5-year estimates of the American Community Survey (ACS).

Table 7: Gender Effect: Inertia

	(1)	(2)	(3)
Husband	0.0708***	0.0226***	0.0227***
	(0.00953)	(0.00515)	(0.00514)
$Subscriber_{it-1}$		0.615***	0.618***
		(0.00624)	(0.00626)
$Subscriber_{it-1}\times$ Job switcher			-0.0329***
			(0.00976)
N	89308	89308	89308
Adj R2	0.445	0.677	0.677

Notes: *** Significant at the 1% level. ** Significant at the 5% level. * Significant at the 10% level. Standard errors are shown in parentheses and cluster at the choice set level.

Table 8: Identification Strategy: Model

Sample	Preference	Gender inattention	Plan inattention	Switching costs
New marriage	X			
Job switcher	X	X		
Passive chooser	X	X	X	X

Notes: New marriage refers to couples who get married for the first time. Job switchers are couples whose previous subscribers changed jobs. Passive choosers are couples with a valid default plan. 'Preference' denotes parameters in the choice stage except for switching costs. 'Gender inattention' means the inattention to the default subscriber. 'Plan inattention' means the inattention to the default plan.

Table 9: Summary Statistics: Estimation Sample

	New Marriage		Job Switcher		Passive	
Demographics						
Husband [%]	53	(49.9)	54.9	(49.8)	56.1	(49.6)
Age	38.6	(10.8)	44.9	(10.3)	45.5	(10.1)
Health	0.2	(0.7)	0.3	(0.9)	0.3	(0.9)
Age gap	1.9	(4.7)	1.9	(4.1)	1.8	(4.1)
Income ratio (ACS)	1.5	(0.3)	1.5	(0.3)	1.5	(0.4)
Education gap (ACS)	-0.1	(0.3)	-0.1	(0.3)	-0.1	(0.3)
Attention shocks						
Plan raise deductible	/		/		0.11	(0.31)
Plan raise PC copay	/		/		0.07	(0.26)
Hit deductible	/		/		0.08	(0.27)
Has inpatient visit	/		/		0.09	(0.29)
Has ER visit	/		/		0.26	(0.44)
Plan characteristics						
Deductible [$1000]	1.5	(1.4)	1.5	(1.4)	1.3	(1.3)
Coinsurance [%]	15.1	(12.8)	13.8	(13.3)	15.1	(12.8)
PC copay	21.6	(7.8)	20.9	(8.2)	19.3	(8.6)
Premium [$1000]	3.6	(1.1)	3.8	(1.3)	3.6	(1.2)
OOP [$1000]	1.5	(1.1)	1.9	(1.4)	1.7	(1.3)
Variance of OOP [($1000)^2]	1.9	(2.1)	2.4	(2.7)	2	(2.4)
# Household	1,741		5,346		20,205	
# Choice	1,741		6,177		43,943	

Notes: Means with standard deviations in parentheses. New marriage refers to couples who get married for the first time. Job switchers are couples whose previous subscribers changed jobs. Passive choosers are couples with a valid default plan. Demographics and attention shocks are summarized at the household-year level, while plan characteristics are at the plan-year level. Income ratio = Median male income/Median female income at the zip code level. Education gap = Average male education years - Average female education years at the zip code level. Both income and education data come from 5-year estimates of the American Community Survey (ACS).

Table 10: Switching Propensity by Household Type

	Same Gender	Same Plan	Share (%)
Job switcher	1	0	94
Job switcher	0	0	6
Passive	1	1	85.3
Passive	1	0	5.7
Passive	0	0	8.9

Notes: This table shows the percentage of households who keep (switch) their default gender (plan) during enrollment for job switchers and passive choosers.

Table 11: Model Estimates

	(1)	(2)	(3)	(4)	(5)	$\frac{\partial p^a}{\partial x}$
Stage 1: Wake-up from Default Plan						
Constant			-0.589	-0.437	-0.429	
			(0.059)	(0.059)	(0.059)	
Deductible increase			-0.014	0.165	0.166	0.041
			(0.058)	(0.068)	(0.068)	
Copay increase			0.314	0.482	0.487	0.120
			(0.070)	(0.086)	(0.087)	
Hit deductible			0.155	0.239	0.241	0.059
			(0.068)	(0.078)	(0.078)	
Inpatient visit			0.122	0.109	0.109	0.027
			(0.063)	(0.070)	(0.071)	
ER visit			0.076	0.084	0.085	0.021
			(0.039)	(0.044)	(0.044)	
Stage 2: Wake-up from Default Gender						
Costant				-0.268	-0.267	
				(0.050)	(0.050)	
Default husband				-0.177	-0.181	
				(0.081)	(0.081)	$\frac{\partial p_j}{\partial x}$
Stage 3: Utility Maximization						
Cost [$1000]	-0.168	-0.752	-0.890	-0.969	-0.982	-0.141
	(0.008)	(0.122)	(0.165)	(0.281)	(0.283)	
Husband	0.179	0.284	0.327	0.182	0.182	0.026
	(0.014)	(0.025)	(0.033)	(0.069)	(0.070)	
Husband × Age gap					-0.014	-0.002
					(0.009)	
Husband × Education gap					0.080	0.012
					(0.126)	
Husband × Log(Income ratio)					0.387	0.056
					(0.178)	
Variance of OOP [($1000)^2]	-0.014	-0.023	-0.029	-0.001	-0.001	0.000
	(0.005)	(0.005)	(0.007)	(0.010)	(0.010)	
Default plan	2.617	2.888	1.282	0.521	0.534	0.085
	(0.015)	(0.036)	(0.081)	(0.086)	(0.086)	
$\hat{\nu}$		0.559	0.645	0.622	0.636	
		(0.123)	(0.166)	(0.282)	(0.284)	
$\hat{\nu^2}$		0.020	0.026	0.031	0.031	
		(0.002)	(0.003)	(0.003)	(0.003)	
SD(ρ)		1.462	1.462	1.462	1.462	
		(0.112)	/	/	/	
Log-Likelihood	-32,602	-32,502	-32,275	-29,625	-29,617	
$ Gender preference	1,069	378	368	188	185	
$ Switching costs	15,605	3,842	1,442	538	544	
Inattention: Default plan	/	/	62.8%	58.2%	58.0%	
Inattention: Default husband	/	/	/	60.9%	61.0%	
Inattention: Default wife	/	/	/	56.7%	56.6%	

Notes: Coefficients with bootstrapped standard errors in parentheses. The last column calculates the average marginal effects on the probability of waking up from the default plan and the choice probabilities conditional on full attention. Both income and education data come from 5-year estimates of the American Community Survey (ACS).

Table 12: Model Robustness Checks

	(1)	(2)	(3)	(4)	(5)
Stage 1: Wake-up from Default Plan					
Constant	-0.453	-0.446	-0.465	-0.441	-0.434
	(0.060)	(0.059)	(0.059)	(0.059)	(0.059)
Deductible increase	0.168	0.164	0.205	0.197	0.166
	(0.067)	(0.068)	(0.068)	(0.069)	(0.068)
Copay increase	0.495	0.477	0.469	0.484	0.483
	(0.086)	(0.086)	(0.085)	(0.086)	(0.087)
Hit deductible	0.256	0.240	0.244	0.242	0.233
	(0.077)	(0.078)	(0.077)	(0.078)	(0.078)
Inpatient visit	0.109	0.107	0.103	0.107	0.109
	(0.070)	(0.070)	(0.070)	(0.070)	(0.070)
ER visit	0.079	0.084	0.083	0.086	0.085
	(0.043)	(0.044)	(0.043)	(0.044)	(0.044)
Stage 2: Attention from Default Gender					
Costant	0.283	0.267	0.293	0.291	0.267
	(0.050)	(0.050)	(0.050)	(0.050)	(0.050)
Default husband	0.158	0.180	0.159	0.166	0.181
	(0.082)	(0.081)	(0.081)	(0.081)	(0.081)
Stage 3: Utility Maximization					
Cost [$1000]	-1.215	-0.940	-1.066	-0.924	-0.994
	(0.305)	(0.279)	(0.321)	(0.293)	(0.279)
Husband	0.213	0.164	0.157	0.131	0.202
	(0.068)	(0.066)	(0.066)	(0.063)	(0.070)
Variance of OOP [($1000)^2]	-0.075	-0.017	0.060	0.049	0.000
	(0.009)	(0.010)	(0.010)	(0.013)	(0.010)
Default plan	0.503	0.507	0.485	0.516	0.524
	(0.089)	(0.086)	(0.088)	(0.087)	(0.086)
Cost $\times$ Log(Size)					0.006
					(0.002)
$\hat{\nu}$	0.976	0.592	0.676	0.518	0.626
	(0.306)	(0.281)	(0.323)	(0.295)	(0.280)
$\hat{\nu^2}$	0.030	0.035	0.031	0.034	0.031
	(0.002)	(0.004)	(0.004)	(0.004)	(0.003)
Log-Likelihood	-29708	-29634	-29671	-29644	-29622
$ Gender Preference	176	175	148	142	203
$ Switching Costs	414	539	455	559	527
Inattention: Default plan	58.6%	58.4%	58.8%	58.2%	58.2%
Inattention: Default husband	60.8%	61.0%	61.1%	61.2%	61.0%
Inattention: Default wife	57.0%	56.6%	57.3%	57.2%	56.6%

Notes: Coefficients with bootstrapped standard errors in parentheses. The standard deviation of measurement error is calibrated to be $1,462, the same as Table 11. Column (1) replaces ex-ante OOP with ex-post OOP. Column (2) adjusts OOP to be 10% higher for plans with the top 1/3 actuarial value and 10% lower for the least. Column (3) replaces the ML algorithm to be Elastic Net. Column (4) replaces the ML algorithm to be Random Forest. Column (5) adds the interaction term between plan costs and employer size.

Table 13: Counterfactual Results

Scenario	Description	% Husband	$ Foregone Welfare
C0	Baseline	55.7%	669
C1	No gender-based inattention	−2.2%	-72
C2	No plan-based inattention	−1.1%	-116
C3	No gender-& plan-based inattention	−4.7%	-227

Notes: This table shows the percentage of households in the husband's ESI and the dollar value of foregone welfare by counterfactual scenarios. Welfare is defined as total costs plus the dollar equivalent value of risk protection and network qualities estimated in Table 11 column (4). Foregone welfare is computed as the welfare difference between the chosen plan and the welfare-maximizing plan. Counterfactuals C1 to C3 are reported relative to the baseline. The no gender-based inattention scenario sets the attention probabilities to both the husband's and the wife's ESI to be one when households wake up from their default plan. The no plan-based inattention scenario sets the wake-up probability from the default plan to be one while allowing for gender-based inattention. C3 combines C1 and C2, which assumes full attention to all plans in a household's choice set.

References

Abaluck, J., Adams-Prassl, A., 2021. What Do Consumers Consider before They Choose? Identification from Asymmetric Demand Responses. *The Quarterly Journal of Economics* 136, 1611–1663.

Abaluck, J., Gruber, J., 2011. Choice Inconsistencies among the Elderly: Evidence from Plan Choice in the Medicare Part D Program. *American Economic Review* 101, 1180–1210.

Abaluck, J., Gruber, J., 2016. Evolving Choice Inconsistencies in Choice of Prescription Drug Insurance. *American Economic Review* 106, 2145–84.

Ackley, C.A., 2020. Essays on Insurance Design and the Demand for Medical Care. Ph.D. thesis. Boston University.

Banerjee, A.N., Hasan, I., Kumar, K., Philip, D., 2021. The Power of a Financially Literate Woman in Intra-household Financial Decision-Making. *Available at SSRN: 3246314* .

Barber, B.M., Odean, T., 2001. Boys Will be Boys: Gender, Overconfidence, and Common Stock Investment. *The Quarterly Journal of Economics* 116, 261–292.

Barseghyan, L., Molinari, F., Thirkettle, M., 2021. Discrete Choice under Risk with Limited Consideration. *American Economic Review* 111, 1972–2006.

Becker, G.S., 1974. A Theory of Social Interactions. *Journal of Political Economy* 82, 1063–1093.

Bertocchi, G., Brunetti, M., Torricelli, C., 2014. Who Holds the Purse Strings within the Household? The Determinants of Intra-Family Decision Making. *Journal of Economic Behavior & Organization* 101, 65–86.

Bhargava, S., Loewenstein, G., Sydnor, J., 2017. Choose to Lose: Health Plan Choices from a Menu with Dominated Option. *The Quarterly Journal of Economics* 132, 1319–1372.

Brown, Z.Y., 2017. An Empirical Model of Price Transparency and Markups in Health Care. University of Michigan Working Paper .

Brown, Z.Y., 2019. Equilibrium Effects of Health Care Price Information. *Review of Economics and Statistics* 101, 699–712.

Browning, M., Bourguignon, F., Chiappori, P.A., Lechene, V., 1994. Income and Outcomes: A Structural Model of Intrahousehold Allocation. *Journal of political Economy* 102, 1067–1096.

Chandra, A., Gruber, J., McKnight, R., 2010. Patient Cost-Sharing and Hospitalization Offsets in the Elderly. *American Economic Review* 100, 193–213.

Chiappori, P.A., 1988. Rational Household Labor Supply. *Econometrica* , 63–90.

Chiappori, P.A., 1992. Collective Labor Supply and Welfare. *Journal of political Economy* 100, 437–467.

Choukhmane, T., Goodman, L., O'Dea, C., 2022. Efficiency in Household Decision-Making: Evidence from the Retirement Savings of US Couples. Retrieved from https://tahachoukhmane.com/wp-content/uploads/2022/07/CGO-couples-401k-Slides-july-2022.pdf.

Clemens, J., Gottlieb, J.D., 2017. In the Shadow of a Giant: Medicare's Influence on Private Physician Payments. *Journal of Political Economy* 125, 1–39.

Clemens, J., Gottlieb, J.D., Molnár, T.L., 2017. Do Health Insurers Innovate? Evidence from the Anatomy of Physician Payments. *Journal of Health Economics* 55, 153–167.

Cooper, Z., Craig, S.V., Gaynor, M., Van Reenen, J., 2019. The Price Ain't Right? Hospital Prices and Health Spending on the Privately Insured. *The Quarterly Journal of Economics* 134, 51–107.

Coughlin, M., 2019. Insurance Choice with Non-monetary Plan Attributes: Limited Consideration in Medicare Part D. Cornell University Working Paper .

Dushi, I., Honig, M., 2003. Price and Spouse's Coverage in Employee Demand for Health Insurance. *American Economic Review* 93, 252–256.

Einav, L., Finkelstein, A., Ryan, S.P., Schrimpf, P., Cullen, M.R., 2013. Selection on Moral Hazard in Health Insurance. *American Economic Review* 103, 178–219.

Friedberg, L., Webb, A., 2006. Determinants and Consequences of Bargaining Power in Households. NBER Working Paper no. 12367 .

Gu, R., Peng, C., Zhang, W., 2021. The Gender Gap in Household Bargaining Power: A Portfolio-Choice Approach. University of Cambridge Working Paper .

Guiso, L., Zaccaria, L., et al., 2021. From Patriarchy to Partnership: Gender Equality and Household Finance. Einaudi Institute for Economics and Finance Working Paper .

Hahn, J., Ridder, G., 2017. Instrumental Variable Estimation of Nonlinear Models with Nonclassical Measurement Error Using Control Variables. *Journal of Econometrics* 200, 238–250.

Handel, B., Holmes, J., Kolstad, J., Lavetti, K., 2018. Insurer Innovation and Health Care Efficiency: Evidence from Utah. Retrieved from http://kurtlavetti.com/innovation˙hhkl.pdf.

Handel, B.R., 2013. Adverse Selection and Inertia in Health Insurance Markets: When Nudging Hurts. *American Economic Review* 103, 2643–82.

Handel, B.R., Kolstad, J.T., 2015. Health Insurance for "Humans": Information Frictions, Plan Choice, and Consumer Welfare. *American Economic Review* 105, 2449–2500.

Handel, B.R., Kolstad, J.T., Minten, T., Spinnewijn, J., 2020. The Social Determinants of Choice Quality: Evidence from Health Insurance in the Netherlands .

Heiss, F., McFadden, D., Winter, J., Wuppermann, A., Zhou, B., 2021. Inattention and Switching Costs as Sources of Inertia in Medicare Part D. *American Economic Review* 111, 2737–81.

Ho, K., Hogan, J., Scott Morton, F., 2017. The Impact of Consumer Inattention on Insurer Pricing in the Medicare Part D Program. *The RAND Journal of Economics* 48, 877–905.

Ho, K., Lee, R.S., 2020. Health Insurance Menu Design for Large Employers .

Ke, D., 2021. Who Wears the Pants? Gender Identity Norms and Intrahousehold Financial Decision-Making. The Journal of Finance 76, 1389–1425.

KFF, 2020. Employer Health Benefits 2020 Annual Survey.

Kiss, A., 2015. Salience and Switching.

Kling, J.R., Mullainathan, S., Shafir, E., Vermeulen, L.C., Wrobel, M.V., 2012. Comparison Friction: Experimental Evidence from Medicare Drug Plans. *The Quarterly Journal of Economics* 127, 199–235.

Lim, H., Shin, S.H., Wilmarth, M.J., Park, N., 2022. Who Decides? Financial Decision-Making among Older Couples. *Journal of Family and Economic Issues* 43, 310–337.

Marone, V.R., Sabety, A., 2022. When Should There Be Vertical Choice in Health Insurance Markets? *American Economic Review* 112, 304–42.

Petrin, A., Train, K., 2010. A Control Function Approach to Endogeneity in Consumer Choice Models. *Journal of Marketing Research* 47, 3–13.

Polyakova, M., 2016. Regulation of Insurance with Adverse Selection and Switching Costs: Evidence from Medicare Part D. *American Economic Journal: Applied Economics* 8, 165–95.

Rickert, D., 2022. When Switching Costs Benefit Consumers. Empirical Evidence from a Market with Lock-in Effects. PSL University Working Paper .

Samuelson, P.A., 1956. Social Indifference Curves. *The Quarterly Journal of Economics* 70, 1–22.

Stockley, K., 2016. Evaluating Rationality in Responses to Health Insurance Cost-Sharing: Comparing Deductibles and Copayments. Harvard University Working Paper .

Tilipman, N., 2022. Employer Incentives and Distortions in Health Insurance Design: Implications for Welfare and Costs. *American Economic Review* 112, 998–1037.

Train, K.E., 2009. *Discrete Choice Methods with Simulation*. Cambridge University Press.

Vihriälä, E., 2022. Intra-Household Frictions, Anchoring, and the Credit Card Debt Puzzle, in: Aalto University Working Paper.

Wang, R., 2018. Price Transparency and Healthcare Costs: The Case of the New Hampshire Healthcare Market. Tulane University Working Paper .

Wooldridge, J.M., 2015. Control Function Methods in Applied Econometrics. *Journal of Human Resources* 50, 420–445.

10 Appendix

10.1 Additional Figures

Table A1: Reduced-form Robustness Checks

	(1) Gap ≤ 1	(2) Gap ≤ 3	(3) Gap ≤ 5
Husband	0.0586***	0.0646***	0.0680***
	(0.0100)	(0.00972)	(0.00965)
N	55268	76923	85343
Adj R2	0.443	0.433	0.438
	(1) Kid ≤ 3	(2) Kid ≤ 6	(3) Kid ≤ 18
Husband	0.105***	0.113***	0.101***
	(0.0344)	(0.0265)	(0.0153)
N	10470	17244	43228
Adj R2	0.552	0.534	0.498

Notes: *** Significant at the 1% level. ** Significant at the 5% level. * Significant at the 10% level. Standard errors are shown in parentheses and cluster at the choice set level. 'Gap' denotes the gap in years for backing out the spouse's employer. 'Kid' denotes the presence of children under a specific age.

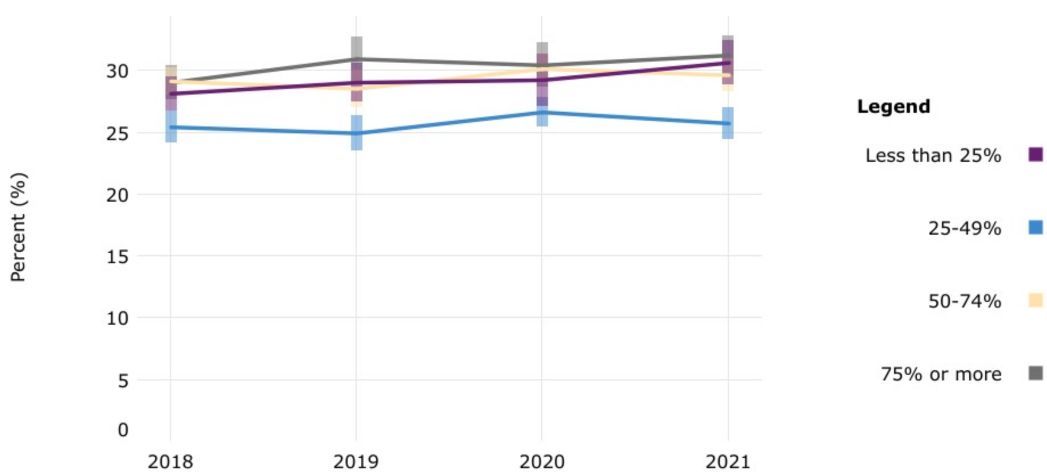


Figure A1: % Employee Premiums in Family Plans by % Females

Notes: Statistics from MEPS. This graph plots the percent of total premiums contributed by employees enrolled in family coverage at private-sector establishments that offer health insurance by the share of female employees from 2018 to 2021 with 95% confidence intervals.

Table A2: Counterfactual Welfare with Alternative Measures

Counterfactual	$ Foregone Welfare	+ Gender 'Preference'	+ Switching Costs
Baseline	669	656	486
No gender-based inattention	-72	-67	-38
No plan-based inattention	-116	-114	-8
No gender- & plan-based inattention	-227	-218	-56

Notes: This table shows various measures of foregone welfare for counterfactual scenarios C0-C3 in Table 13. The first column uses the baseline measure, which is total costs plus the dollar equivalent value of risk protection and network qualities. The second column includes gender 'preference' in the measure of welfare, and the third column incorporates switching costs.

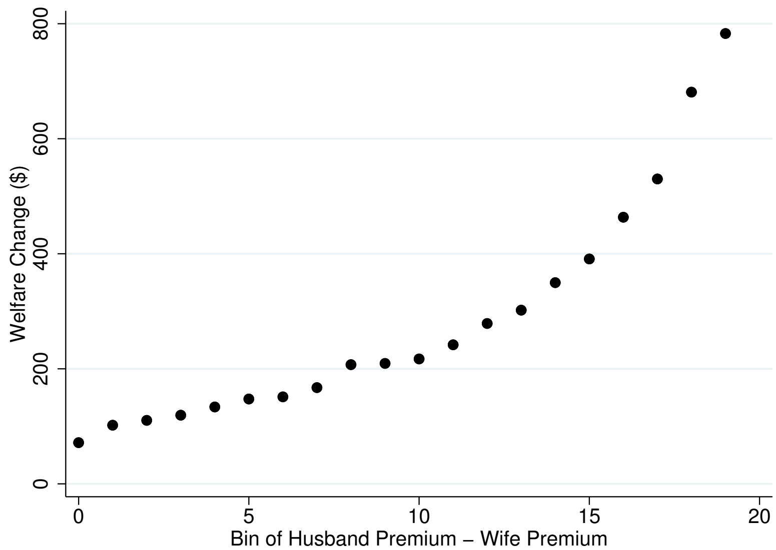
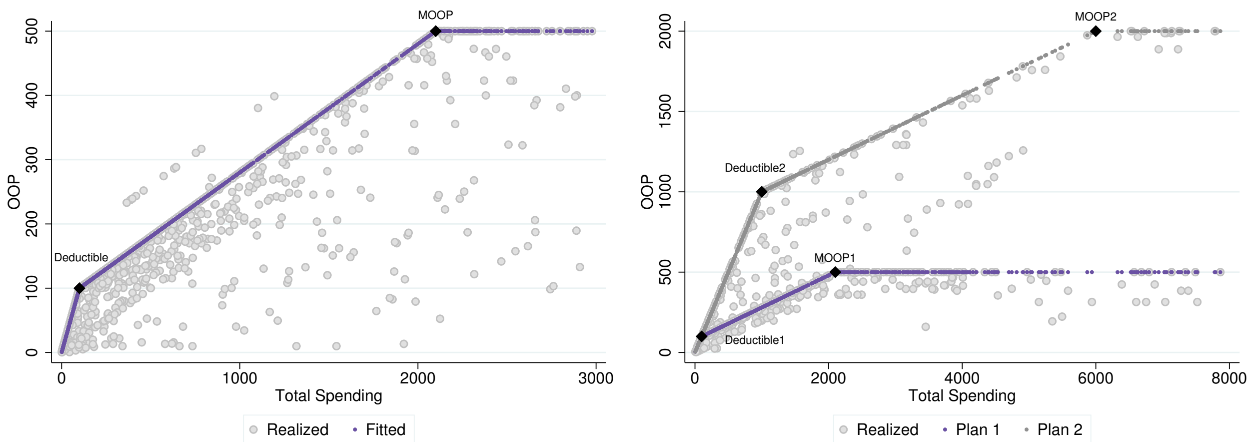


Figure A2: Change in Foregone Welfare ($) by Premium Difference: Full Attention Scenario

Notes: This graph plots the average change in foregone welfare among households who have their husbands as the default subscriber for a counterfactual scenario that assumes full attention (C3 in Table 13) relative to the baseline by bins of the premium differences within the household. The premium differences are defined by the premiums of the least-expensive plan offered by the husband minus that of the wife. A higher value indicates that the husband's option is more pricey than the wife's.

10.2 Plan Characteristics



(a) One Option

(b) Two Options

Figure A3: Plan Financial Characteristics from Claims

I first back out plan characteristics from the claims data, which are crucial inputs into the cost model below. To do this, I separate claims into two types: one applies copay while the other counts into deductible and coinsurance. For office visits that apply copay, I differentiate primary doctor visits from specialist visits, and take their most frequent copay values respectively. A notable feature for many plans in New Hampshire is free lab tests at independent facilities and designated hospitals.[42] I therefore set a plan's lab tests to be free if over 50% of its lab claims are free in a given year.

For the rest of the services that apply deductible and coinsurance, I aggregate individuals' cumulative monthly spending and OOP costs, and then create a scatter plot for each policy-year as in Figure A3 Panel (a). This graph is used to fit the closest insurance line that is characterized by a deductible-coinsurance-MOOP triplet. Specifically, I choose from discrete values of the three parameters and pick a combination that minimizes the mean squared distance between the line and all the points.[43] Similar methods have been used in Handel et al. (2018) and Marone and Sabety (2022).

While implementing this, I find some policies show up as two options in Figure A3 Panel (b). I therefore repeat this exercise with two lines for a subset of policies that didn't fit well with one.[44] I determine the group to have two options when its MSE is significantly lower than the one-option case, and both lines fit close to a non-negligible share of participants. 4.7% groups in my data end up with two options.

[42]For example, Anthem Site of Service, Harvard Pilgrim Low-Cost Provider, and Tufts Freedom Plan. For details please see Ackley (2020).

[43]To better fit the line, I allow for higher weights for points whose deductible/MOOP is divisible by \$250/\$500 or whose coinsurance rate is divisible by 10%.

[44]I restrict the deductible gap between two options to be at least \$500 and the coinsurance rate to be the same, since it is typical for an employer to offer a high-deductible and low-deductible option for otherwise similar plans.

Lastly, using HIOS plan ID, I merge small group plans after 2016 with CMS Health Plan Finder Data to fill out their plan characteristics. I then impute a plan's missing attributes from its neighboring years. This final dataset on plan characteristics is further used to construct households' choice sets and calculate OOP costs below.

10.3 Cost Model in Detail

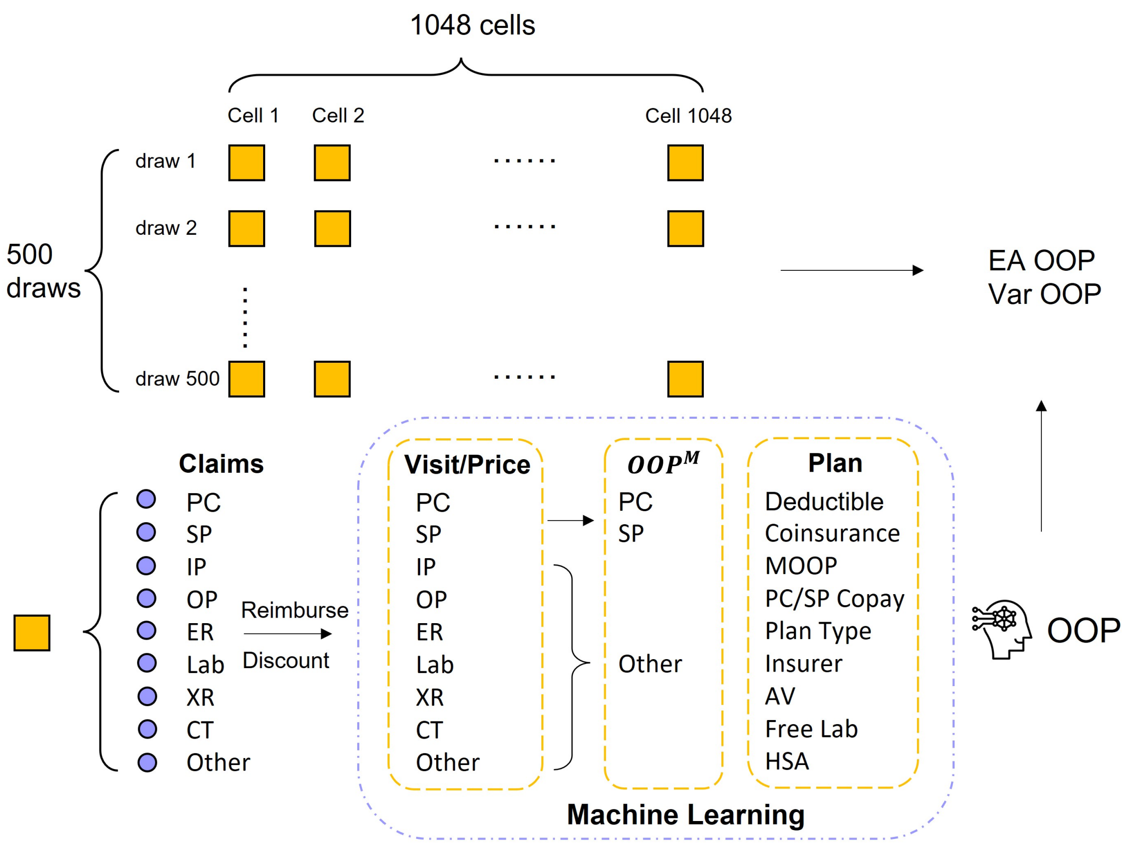


Figure A4: Mapping Out-of-pocket Costs

This section discusses the steps to build the distribution of OOP costs, F_{ijt}, which is specific to household i, plan j, and year t. It is then used as an important input to the choice model to quantify the shadow price of gender preference. I predict F_{ijt} in a way that takes into account (i) enrollee demographics, health risks, and past spending; (ii) the empirical distribution of claims among 'identical' individuals defined in (i); (iii) price negotiation between insurers and providers; (iv) a machine learning algorithm that converts category-specific medical claims into OOP expenses.

I start by assigning individual i^n in household i into one of $1,048$ cells based on year (11), age group (7), chronic conditions (17) (from the Charlson Comorbidity Index), and past spending levels (8). I combine a cell with neighboring cells if it has fewer than 500 members. Then I draw 500 individuals from each cell, and take their realized claims as the ex-ante distribution for each person (represented by k) in the cell, including individual i^n.

$$\Phi_{kj} : H_k \xrightarrow{PR_{sndt}} AH_{kj} \tag{19}$$

The next part of the cost model maps the distribution of claims H_k into out-of-pocket costs OOP_{kj} for each plan j in household i's choice set. This is implemented by first adjusting the billed amount for each claim in H_k by a negotiated price ratio PR_{sndt}. As a bargaining outcome between insurers and providers, PR_{sndt} is calculated as the average of $\frac{Reimbursed}{Billed}$ over claims specific to insurer s, provider n, procedure d and year t.[45] This measure captures the fact that households who live close to an insurer's 'realm' and utilize its better-negotiated providers/procedures pay less on average than other households.

$$\Omega_{kj} : AH_{kj} \xrightarrow{CS_{jt}} OOP_{kj} \tag{20}$$

The third step projects the adjusted distribution of claims AH_{kj} into out-of-pocket costs OOP_{kj} with a plan's cost-sharing features CS_{jt}. I use machine learning to account for the complex nature of medical plans by training the model on realized OOP from the chosen plans and predicting OOP for alternative plans. I start by dividing each draw of H_k into 9 categories: primary doctor visits, specialist visits, inpatient admission, outpatient surgeries, emergency room visits, laboratory tests, x-rays, CT scans, and other services. For each category, I pull out the sampled individual's total number of visits and the adjusted billed amount. These category-specific records are taken as the first set of predictors in the machine learning model.

The second set of features includes insurer brands, plan types, and plan characteristics imputed in Appendix 10.2. In addition, I include a plan's actuarial value (AV) calculated from the claims data, and mark whether it features a Health Savings Account (HSA).[46] The last set of inputs is manually calculated OOP costs. For categories that are carved out with copays, $OOP = copay \times \#visits$. For the remaining categories that count into the deductible, I aggregate their adjusted spending and check if it meets the deductible before applying the coinsurance rate. OOP^M_{kj} is then the sum of OOP costs over all categories for the sampled individual k with plan j. I also construct an AV-based OOP costs: $OOP^{AV}_{kj} = TotalSpending_k \times AV_{jt}$.

[45] I require a valid PR_{sndt} to have at least 5 claims, which covers around 70% claims. For the remaining, I fill it with insurer-provider-year average price ratios (18%), insurer-city-procedure-year average (4%), insurer-procedure-year average (4%), provider-year average (2%), procedure-year average (1%), and insurer-city-year average (1%).

[46] HSA plans often feature high deductibles and apply deductibles to office visits.

Table A3: Performance on Machine Learning OOP

	Corr	R2	RMSE	Mean
Realized OOP	/	/	/	575
Naïve	0.872	0.760	529	680
Elastic Net	0.893	0.798	457	580
Random Forest	0.904	0.817	447	608
Gradient Boosting	0.919	0.844	401	588

The three groups of features, together with their two-way interactions, provide the fuel for machine learning. I leverage three algorithms: Elastic Net, Random Forest, and Gradient Boosting. Their performance on a 20% test sample is listed in Table A3, in addition to the naive approach that uses OOP_{kj}^{M}. The three algorithms produce similar predictions, and all of them outperform the naive approach, which overpredicts OOP by $80. The ML-predicted OOP achieves an over 90% correlation with realized OOP, and explains 81% of its variation while preserving the mean. In the main analysis, I use the OOP predicted using Gradient Boosting and show robustness to the other two methods.

$$F_{ij} : OOP_{kj} \rightarrow OOP_{ij} \tag{21}$$

To create the final object of interest, I sum up OOP_{kj} across all members in household i to get OOP_{ij}, and repeat for the 500 draws to determine its distribution F_{ij}. Using F_{ij}, I finally compute the mean and variance of OOP costs used in the choice model.

10.4 Premium Model

There are a few sources where I could trace out premiums for a given plan. First, premium information is provided for plans with New Hampshire situs after 2016 (18%). Second, small-group plans with a valid HIOS Plan ID after 2016 can be merged to CMS Health Plan Finder Data for premiums (5%). Lastly, I exploit a similar machine learning approach as detailed in Appendix 10.3 for large group plans (77%). Following a few insurers' group rating guidelines, I include three sets of predictors in the machine learning model below.

The first category includes plan characteristics from Appendix 10.2. I additionally include whether the group is self-insured and the number of distinct hospitals/physicians visited by the group's enrollees as proxies for network generosity. The second category includes group basics such as size, 3-digit zip code, and the distribution of age, gender, and enrollment tiers.[47] The last category involves the group's claims history in year $t-1$ and $t-2$. In addition, I attach two-way interactions for predictors within and across each category.

[47] I assign an employer's location to be the 3-digit zip code where most employees live. I divide enrollees into 6 age groups and compute shares of enrollees in a specific age group+gender. Similarly, I calculate the share of employees that enroll in EMP/ESP/ECH/FAM tiers.

Table A4: Performance on Machine Learning Premium

	Corr	R2	RMSE	Mean
Premium	/	/	/	5,655
Elastic Net	0.650	0.422	1482	5654
Random Forest	0.907	0.823	981	5663
Gradient Boosting	0.938	0.881	722	5654

The performance of premium models on a 20% test sample is listed in Table A4. Again, the three machine learning algorithms produce similar predictions, with Gradient Boosting having the highest performance. The correlation between predicted and realized premiums is 57.5%, which absorbs about one-third of its variance. The weaker performance on premiums relative to OOP costs is due to a smaller sample size, since the prediction here is made at the plan-year level.

10.5 Model Fit

Table A5: % Households in Husband's ESI

	Total	Active	Job Switcher	Passive
Data	55.9%	53.0%	54.9%	56.1%
Model	55.7%	51.9%	53.6%	56.2%

Notes: The first row shows the percentage of households that are enrolled in the husband's ESI from the estimation data by active choosers, job switchers, and passive choosers defined in Table 8. The second row shows the corresponding model predictions.

Table A5 reports the performance of the structural model by comparing the simulated percentage of households enrolling in their husbands' ESI to the estimation data. Overall, the model's prediction is only 0.2% off and preserves the relation that the gender gap is the largest among passive choosers, then job switchers, and lastly active choosers. At the plan level, my simulated choice probabilities have a 0.78 correlation with actual choices. Also shown in Figure A5, my model successfully separates chosen plans from unchosen options.

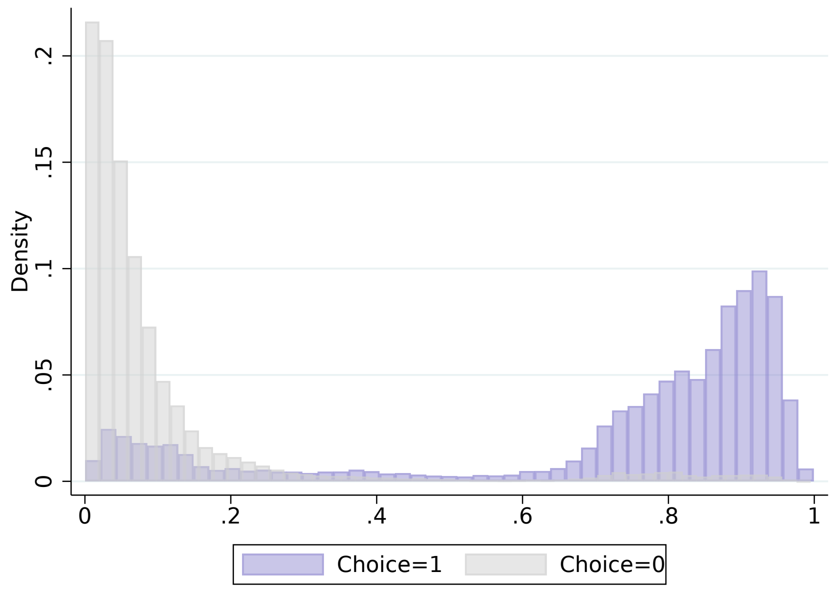


Figure A5: Fitted Choice Probabilities

Notes: This graph plots the distribution of fitted choice probabilities from the model by whether or not a plan was actually chosen in the data.