An Equilibrium Analysis of the Long-Term Care Insurance Market

Ami Ko[*]

First version: November 2016

This version: March 2020

Abstract

A life-cycle model of intergenerational long-term care decisions is developed to analyze how family interactions affect the equilibrium coverage and welfare in the U.S. long-term care insurance market. I start by showing descriptive evidence that private information about children's informal care likelihood results in adverse selection: the market attracts a disproportionate number of individuals who face higher formal care utilization risk due to a lower probability of receiving informal care. Motivated by these facts, I develop and structurally estimate a dynamic intergenerational non-cooperative model featuring long-term care insurance, savings, informal care provision and employment choices. Using an equilibrium insurance market framework, I show that using family demographics in pricing long-term care insurance contracts reduces adverse selection and improves the average welfare of the family. Using the non-cooperative feature of the model, I also quantify to what extent parents forgo long-term care insurance to avoid diminishing children's informal care incentive.

[*]Georgetown University, ICC 557, Department of Economics, 37th and O Streets, N.W., Washington D.C. 20057. E-mail: ami.ko@georgetown.edu. I am extremely grateful to my advisors, Hanming Fang, Holger Sieg, and Petra Todd for their guidance and support throughout this project. I thank Naoki Aizawa, Daniel Barczyk, Minsu Chang, George-Levi Gayle, Qing Gong, Matthias Kredler, Lee Lockwood, Robert Miller, Toshihiko Mukoyama, Alexandre Poirier, John Rust, Andrew Shephard, Dongho Song, Jan Tilly, Kenneth Wolpin, and numerous seminar and conference participants for useful comments. I gratefully acknowledge financial support from the Korea Foundation for Advanced Studies and the Center for Retirement Research at Boston College (BC Grant 5002099) sponsored by the Social Security Administration.

1 Introduction

Elderly individuals face substantial risk of having functional limitations and hence requiring long-term care. In the U.S., about three fourths of 60-year-olds will have chronic conditions resulting in daily activity limitations, while the other fourth will have no such conditions until death. Formal long-term care services are expensive with the median annual cost for nursing homes exceeding $90,000. Public insurance, Medicaid, exists but is means-tested, and one has to be impoverished to be eligible for the benefits. Yet, less than 15% of the elderly own private long-term care insurance to protect against these large long-term care expenditure risks.[1] Many researchers have studied why the market for long-term care insurance is so small and identified Medicaid (Brown and Finkelstein, 2008), bequest motives (Lockwood, 2018), private information (Hendren, 2013), market power and administrative costs (Braun, Kopecky, and Koreshkova, 2019) as possible explanations. However, scant attention has been given to the role of unpaid care provided by families, usually adult children, as a substitute to formal long-term care.

This paper assesses two main mechanisms through which family interactions could affect the equilibrium of the long-term care insurance market. First, it studies whether private information about children's informal caregiving is a source of adverse selection, and if so, what the welfare consequences are. As the provision of long-term care, i.e., assistance with daily tasks, does not require much professional training, informal care provided by adult children can substitute for formal care services.[2] Since long-term care insurance companies pay only for formal care, whether a consumer has children who are likely to provide informal care is relevant for insurance companies' costs. However, premiums in the long-term care insurance market do not vary by child characteristics,[3] despite the absence of regulation that explicitly prohibits such practices.[4] This suggests that insurance companies could attract a disproportionate number of individuals who face higher formal care utilization risk due to a lower probability of receiving informal care from children.

Second, this paper studies by how much parents forgo insurance to avoid diminishing informal care incentives faced by their children. As argued in several theoretical papers, with insurance, parents cannot use bequests as an effective instrument to elicit caregiving behavior from their children (Bernheim, Shleifer, and Summers, 1985; Pauly, 1990). The reason is because the children know that even if they do not provide informal care, their inheritances will not be spent on formal care as the insurance company will pay for the cost. Under the assumption that the parents prefer

[1]The statistics for long-term care risk are computed by the author using data from the Health and Retirement Study. Formal care prices come from Genworth, https://www.genworth.com/aging-and-you/finances/cost-of-care.html.

[2]The substitutability between informal care and formal care has been documented in several papers including Houtven and Norton (2004), Charles and Sevak (2005), Houtven and Norton (2008) and Coe, Goda, and Van Houtven (2015).

[3]According to the 2015 report by Broker World (2009-2015), which surveyed major long-term care insurers who together accounted for 99% of the sales, premiums varied only by age, gender and 3 or 4 underwriting classes based on health conditions. The underwriting guideline from Genworth, the largest long-term care insurer in the U.S., also shows that no information about children is used in the underwriting process (the guideline can be found in Hendren (2013)).

[4]NAIC Long-Term Care Insurance Model Regulation, www.naic.org/store/free/MDL-641.pdf.

to be cared for by family members rather than hired strangers, the parents will rationally choose not to buy insurance to avoid distorting the incentives their children face. This strategic non-purchase of insurance has been investigated in the theoretical literature as a possibly important channel in explaining the limited size of the insurance market. This paper, to the best of my knowledge, is the first to quantify its magnitude.

I start by providing descriptive evidence that there is adverse selection in the long-term care insurance market generated by private information about children's informal care likelihood. Employing a widely used test for asymmetric information in insurance markets (Finkelstein and McGarry, 2006; Finkelstein and Poterba, 2014), I show that individuals who do not believe their children will provide informal care are more likely to have long-term care insurance and enter a nursing home in subsequent years. The negative relationship between initial beliefs about receiving informal care and ex-post formal care utilization is found even among individuals who have already purchased insurance and experience care needs. I show that there are child demographics that are highly predictive of whether a parent believes a child will provide informal care, such as the child's gender and residential proximity to the parent. I also provide descriptive evidence that the amount of parents' wealth exposed to formal care spending risk is positively correlated with caregiving behaviors from children. This suggests that bequests may be important in incentivizing children to provide informal care.

Based on these findings, I develop a dynamic intergenerational model in which an elderly parent and an adult child interact non-cooperatively from the parent's retirement to death. In the first period, the parent makes a long-term care insurance purchase decision taking into account the likelihood of the child providing informal care when needed. In each of the later periods, the parent experiences health and wealth shocks. The child allocates time to informal care provision, work and leisure. Skira (2015) also estimates a dynamic model of an adult child's informal care and work choices, but abstracts from intergenerational interactions. When hit by an adverse health shock, the parent uses formal care only when the child decides not to provide informal care. The parent pays for formal care either using her long-term care insurance, Medicaid benefits or savings. This aspect of the model is related to the literature on elderly savings and medical expenditure uncertainty (for example, Hubbard, Skinner, and Zeldes (1995), Palumbo (1999) and De Nardi, French, and Jones (2010)). In case of the parent's death, the child inherits the parent's wealth.

The model incorporates private information about the probability of receiving informal care by assuming that various family demographics, which are unpriced by insurers, affect the benefit and cost of the child's informal care provision. The child's "warm-glow" utility from providing informal care depends on the child's gender and residential proximity to the parent. The child's strategic incentive to provide care depends on the parent's wealth, which represents the child's potential inheritance. The child's cost of providing informal care is determined by the child's income function which depends on her eduction, among other things. This aspect of the paper is related to the vast empirical literature on asymmetric information in insurance markets (see Einav, Finkelstein, and Levin (2010) for a survey of the literature). My contribution lies in studying an equilibrium insurance market where I endogenize the key source of adverse selection, i.e., informal care provision by children.

The model assumes, conditional on the child not providing informal care, the required intensity of formal care usage decreases in the number of children. This is to capture, in a reduced-form way, the possibility that the child in the model may not be the only source of informal care. Owing to this assumption, the number of children also creates private information about the parent's expected formal care cost.

I estimate the intergenerational game using data from the Health and Retirement Study (HRS) 1998-2010 and actual premium data from the sample period. Using a full solution approach entails a significant computational cost as I have a dynamic game with a large state space. I overcome this issue by using a two-step Conditional Choice Probability (CCP) estimator pioneered by Hotz and Miller (1993). My estimation strategy follows Bajari, Benkard, and Levin (2007) who extend the forward simulation based CCP approach proposed by Hotz, Miller, Sanders, and Smith (1994) to dynamic games and allow for continuous choices.[5] I recover parents' preferences for care and bequests, as well as children's preferences for leisure, informal care provision and inheritances.

The estimated intergenerational game reproduces the most important features of the data including the monotonically increasing long-term care insurance ownership rate and the inverted-U pattern of informal care receipt across wealth. The model also reproduces a low correlation between long-term care insurance ownership and formal care risk found in previous studies (Finkelstein and McGarry, 2006; Braun, Kopecky, and Koreshkova, 2019). This is done by incorporating income-based advantageous selection which offsets the positive correlation induced by adverse selection based on the probability of receiving informal care. While higher-income people are healthier than lower-income people, they have higher willingness to pay for private insurance as Medicaid serves as a substantially worse substitute for them due to its means-tested nature. As the model incorporates both adverse selection and advantageous selection, in aggregate, it produces a low correlation between insurance ownership and formal care risk.

To embed the estimated intergenerational game within an equilibrium long-term care insurance market, I introduce competitive insurance companies that compete by setting prices. Using this equilibrium insurance market framework, I first show that the model replicates the magnitude of adverse selection generated by private information about children's informal care likelihood found in the descriptive analysis. Next, to reduce the adverse selection channel, I consider counterfactual risk adjustment whereby an individual's long-term care insurance premiums are adjusted based on observables that are powerful predictors of the likelihood of receiving informal care from children. I find that pricing on the presence of a daughter and a child living in close proximity as well as the number of children increases the equilibrium coverage rate and the average welfare of the family by over $6,200. In contrast, I find that pricing on consumers' gender, which was newly introduced in the long-term care insurance market in recent years, has almost no effect on the average welfare.

Finally, using the non-cooperative feature of the model, I quantify the magnitude of the strategic

[5] A series of contemporaneous papers by Aguirregabiria and Mira (2007), Pakes, Ostrovsky, and Berry (2007) and Pesendorfer and Schmidt-Dengler (2008) have recently developed estimators focusing on infinite horizon games with stationary Markov Perfect Equilibrium.

non-purchase of insurance. I find that when long-term care insurance does not crowd out children's informal care provision, the equilibrium ownership rate increases by over 7 percentage points, corresponding to a 42% increase. I find that the effect is greater among wealthier parents. This suggests that the strategic non-purchase of insurance is the most relevant for wealthy parents who have enough bequests to incentivize their children.

This paper is most closely related to the recent works by Barczyk and Kredler (2018) and Mommaerts (2016). Barczyk and Kredler (2018) use a dynamic intergenerational non-cooperative model to study elderly care arrangements. While all of the main results in my paper are about how family interactions affect the long-term care insurance market, their analysis abstracts from long-term care insurance. Their main findings are about how various government policies surrounding long-term care, such as care subsidies and Medicaid reforms, affect care arrangements and welfare of the family.[6] Mommaerts (2016) studies dynamic intergenerational interactions over insurance and care decisions. In contrast to my paper, Mommaerts (2016) uses a cooperative model with limited commitment which does not allow for the strategic non-purchase of insurance. Also, Mommaerts (2016) considers only the demand side of the long-term care insurance market and does not study how insurance selection occurs based on children's heterogeneous informal care probabilities.

The rest of this paper proceeds as follows. Section 2 presents empirical facts about long-term care in the U.S. Section 3 presents the model. Section 4 presents the data and the estimation results. Section 5 presents the main results. Section 6 concludes.

2 Empirical Facts

I start by providing empirical facts about the U.S. long-term care sector. The main data for this paper come from the HRS which has surveyed a representative sample of Americans over the age of 50 every two years since 1992. Using seven waves of the HRS 1998-2010, I provide empirical patterns that motivate the model of intergenerational long-term care decisions presented in the next section.

2.1 Background

Substantial long-term care risk. Long-term care is formally defined as assistance with basic personal tasks of everyday life, called Activities of Daily Living (ADLs) or Instrumental Activities of Daily Living (IADLs). Examples of ADLs include bathing, dressing, using the toilet and getting in and out of bed. IADLs refer to activities that require more skills than ADLs such as doing housework, managing money, using the telephone and taking medication. Declines in physical or mental abilities are the main reasons for requiring long-term care. Using individuals aged 60 and over in the HRS 1998-2010, I find that over 60% of individuals aged 85 and older need assistance

[6] For example, they find that long-term care subsidies generate large welfare gains, even when combined with a smaller Medicaid program. Fahle (2014) also uses a similar framework to evaluate various long-term care policies, but abstracts from long-term care insurance.

with daily tasks. However, not everybody develops ADL/IADL limitations towards the end of their lives. In fact, about 32% (19%) of healthy 60-year-old men (women) will never need long-term care until their death.[7] These findings suggest that elderly individuals face substantial risks about how much long-term care they would need.

Informal care as the backbone of long-term care delivery. Unpaid long-term care provided by the family - which I refer to as informal care in this paper - plays a substantial role in the long-term care sector. This is because unlike acute medical care, long-term care does not require professional training: it simply refers to assistance with basic personal tasks. Several studies have documented the importance of informal care in the U.S. long-term care sector. For example, work by Barczyk and Kredler (2018) shows that informal care accounts for 64% of all help hours received by the elderly. Table 4 presented later in Section 4 reports the average child characteristics by whether they provide informal care to parents with functional limitations.[8] Caregiving children are much more likely to be a daughter and live within a 10-mile radius of their parents. They are less likely to have college education, be married, own a home and work full-time. Only 3% of parents pay their children for help, implying that inter-vivos financial compensation for informal care is rare.

Costly formal care services. Another way to meet one's long-term care needs is to use formal long-term care services, such as nursing homes, assisted living facilities and paid home care. These formal care services are labor-intensive and expensive. In 2017, the median annual rate was $97,000 for a private room in a nursing home, $45,000 for assisted living facilities and $48,000 for paid home care.[9] Combined with substantial risks of needing long-term care, formal care is one of the largest financial risks faced by the elderly: 40% of 65-year-olds will not have any formal care expenses, while 60% will incur on average $100,000 and 5% will spend more than $300,000 during their remaining life.[10]

A very small long-term care insurance market. Private long-term care insurance provides financial protection against these large formal care risks. The U.S. long-term care insurance market is relatively young, and modern insurance products were introduced in the late 1980s (Society of Actuaries, 2014). Typical long-term care insurance contracts cover both facility care and paid home care provided by employees of home care agencies. Most do not cover informal care (Broker World, 2009-2015). For underwriting purposes, insurance companies perform cognitive tests and require medical records, blood and urine samples. According to the 2015 report by Broker World, which surveyed major long-term care insurers who together accounted for 99% of the sales, premiums varied by age, gender and underwriting class determined by health conditions; no information about applicants' children who are most likely to be primary informal caregivers was collected.[11]

[7] Author's calculation using the HRS 1998-2010. I provide details about the estimation in Section 4.2.

[8] Details on my sample selection criteria are given in Section 4.1.

[9] Genworth, https://www.genworth.com/aging-and-you/finances/cost-of-care.html.

[10] The numbers are from Kemper, Komisar, and Alecxih (2005/2006), and the monetary values have been inflated to 2017 dollars.

[11] The underwriting guideline from Genworth, the largest long-term care insurer in the U.S., also shows

Gender-based pricing was newly adopted in 2013 (Finkelstein and Poterba, 2014), despite the well-known fact that women have a higher chance of using formal care compared to men. Contracts are guaranteed renewable in the sense that an insurance company cannot cancel coverage as long as premiums are paid. They specify a constant and nominal annual premium and do not change for an individual who experiences a change in health. The average purchase age is 61 years, but most people do not use insurance until they turn 80 (Broker World, 2009-2015). Using the HRS 1998-2010, I find that among individuals aged 60 and over, only about 13% own private long-term care insurance. The coverage rate is higher at 20% when I restrict to individuals aged 60-69 who do not have any health conditions that would lead to insurance rejections, which are quite common according to Hendren (2013) and Braun, Kopecky, and Koreshkova (2019).

Medicaid as the biggest payer for formal care services. According to a report by the Kaiser Family Foundation, formal long-term care expenses totaled over $310 billion in 2013, which is close to 2% of GDP.[12] Medicaid is the biggest payer accounting for 51% of the total payments, followed by other public insurance programs (21%), out-of-pocket (19%) and private long-term care insurance (8%). In contrast to a common misconception, Medicare coverage for long-term care is very limited. Only nursing home stays following a qualified hospital stay are covered up to 100 days, and there are substantial copayments for days 21-100. Medicaid, on the other hand, provides unlimited coverage to eligible individuals with limited assets. While one has to be almost impoverished to be eligible for benefits, individuals can "spend-down" their assets until they meet Medicaid eligibility requirements, which has been identified as an important factor in explaining the limited size of the long-term care insurance market (Brown and Finkelstein, 2008).

2.2 Descriptive evidence on adverse selection

Adverse selection in the insurance purchase phase. One of the most commonly used tests for asymmetric information in insurance markets studies whether, among the set of individuals who are offered the *same* price, there exists unobserved or unused information that is highly predictive of both their ex-post risk and insurance demand (Finkelstein and McGarry, 2006; Finkelstein and Poterba, 2014). To examine if individual beliefs about receiving informal care from children serve as such a dimension of private information, using the HRS data, I estimate the following two equations:

$$NH_{i,t\sim t+5} = \beta_0 + \beta_1 B_{it}^{IC} + X_{it}'\beta_2 + error_{it} \quad \text{and} \tag{1}$$

$$LTCI_{it} = \delta_0 + \delta_1 B_{it}^{IC} + X_{it}'\delta_2 + error_{it}. \tag{2}$$

that no information about children is used in the underwriting process (the guideline can be found in Hendren (2013)).

[12] The report can be found at https://www.kff.org/medicaid/report/medicaid-and-long-term-services-and-supports-a-primer/.

$NH_{i,t\sim t+5}$ is an indicator for having a nursing home stay lasting more than 100 days over the following five-year period.[13] $LTCI_{it}$ is an indicator for current long-term care insurance ownership. B_{it}^{IC} is the key control and is an indicator for whether an individual thinks his or her children will provide informal care when needed.[14] X_{it} is a vector of individual characteristics used by insurance companies in pricing ("pricing controls"). Conditioning on X_{it} ensures that I compute predictive powers of B_{it}^{IC} among individuals faced with the same insurance price, as required by the asymmetric information test. Based on Finkelstein and McGarry (2006) and Hendren (2013), X_{it} includes age, gender and various health conditions.[15] It does not include any information about children as such information is not collected by insurance companies, as explained earlier.

The sample used to estimate Equations (1) and (2) consists of respondents who are healthy enough to buy long-term care insurance at the time of interview and old enough (ages 70-75) to develop ADL limitations in five years since the interview.[16] Columns (1) and (3) of Table 1 report the estimates of the key coefficients, β_1 and δ_1. Individuals who do not believe their children will provide informal care are 0.9 percentage points more likely to have a nursing home stay lasting more than 100 nights in the following five-year period (the mean is 1.6%) and 4.2 percentage points more likely to own long-term care insurance (the mean is 15.6%). As individuals who do not believe their children will provide informal care are (1) higher risk and (2) more likely to buy insurance, the results serve as suggestive evidence that private information about children's expected informal care provision is a source of adverse selection in the insurance market. I also verified that I obtain negative estimates of β_1 and δ_1 when I include nursing home stays lasting less than 100 days, measure subsequent nursing home utilization over a longer time horizon, or use individuals of younger ages.

Columns (2) and (4) of Table 1 estimate Equations (1) and (2), respectively, with one additional dimension of private information: an individual's self-assessed probability of entering a nursing home over the next five-year period, denoted by B_{it}^{NH}. I include this term to compare the importance of private information about informal care options (B_{it}^{IC}) to other dimensions of private information, such as unobserved health, that may be captured in B_{it}^{NH}.[17] The inclusion of B_{it}^{NH} has no effect on the economic magnitude and statistical significance of the relationship between

[13]To differentiate nursing home stays that are partially paid by Medicare, I restrict to nursing home stays lasting more than 100 days. The reason why I measure subsequent nursing home utilization over the following five-year period is because below, I include an individual's self-assessed probability of entering a nursing home over the next five-year period as an additional control in Equations (1) and (2). As I will mention later, the results are robust to including shorter nursing home stays or measuring nursing home risk over a longer time horizon.

[14]The specific question asked in the HRS is "Suppose in the future, you needed help with basic personal care activities like eating or dressing. Will your daughter/son be willing and able to help you over a long period of time?" If the answer is positive for any of the respondent's children, B_{it}^{IC} is set to one and zero otherwise.

[15]The health conditions used as controls are cognitive score and indicators for having a psychological condition, diabetes, lung disease, arthritis, heart disease, cancer and high blood pressure.

[16]I follow Hendren (2013) to identify rejection conditions and exclude individuals who have ADL/IADL limitations, have experienced a stroke, or have used nursing homes or paid home care in the past.

[17]Several studies have used the self-assessed probability of entering a nursing home B_{it}^{NH} to construct a measure of private information about formal care risk (Finkelstein and McGarry, 2006; Hendren, 2013).

Table 1: Results from the Asymmetric Information Test

	(1)	(2)	(3)	(4)
Dependent variable, Y:	Use NH		LTCI	
B^{IC}	-0.009**	-0.009**	-0.042***	-0.037***
	(0.004)	(0.004)	(0.011)	(0.011)
B^{NH}		-0.008		0.219***
		(0.011)		(0.033)
Pricing controls, X	Yes	Yes	Yes	Yes
Mean of Y	0.016	0.016	0.156	0.156
Observations	5,739	5,739	5,739	5,739

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors are clustered at the household level and reported in parentheses. The sample consists of individuals aged 70-75 who are healthy enough to purchase long-term care insurance based on underwriting guidelines in Hendren (2013). B^{IC} is an indicator for whether an individual thinks his or her children will provide informal care in the future. B^{NH} is the individual's self-assessed probability of nursing home entry over the following five-year period. The dependent variable in Columns (1) and (2) is an indicator for having a nursing home stay lasting more than 100 days over the following five-year period. The dependent variable in Columns (3) and (4) is an indicator for current long-term care insurance ownership. A linear probability model is used in all four regressions. All four regressions include buyer characteristics used by insurers in pricing as controls: gender, age, cognitive score and indicators for having a psychological condition, diabetes, lung disease, arthritis, heart disease, cancer and high blood pressure (Finkelstein and McGarry, 2006; Hendren, 2013).

informal care beliefs and subsequent nursing home risk. What is worth noting is that B_{it}^{NH} has no power in predicting subsequent nursing home use conditional on informal care beliefs and pricing controls: the relationship is indeed negative and statistically insignificant.[18] If B_{it}^{NH} reflects private information about one's health, the insignificant relationship suggests that the amounts of residual private information about health are small.[19]

To examine if there are observables that are highly predictive of whether a parent believes a child will provide informal care, I regress the parent's beliefs about receiving informal care from the child on the child's characteristics, parental assets and individual characteristics used by long-term care insurers in pricing. The results are presented in Table A.2 of Appendix A. It shows that several child demographics, especially whether the child is a daughter and lives within a 10-mile radius to the parent, are powerful predictors of the parent's beliefs about receiving informal care from the child.

[18]This result is consistent with Hendren (2013) who finds little predictive power of beliefs about nursing home entry among individuals who are eligible to buy long-term care insurance. The fact that beliefs about informal care provision by children have predictive power while beliefs about nursing home entry do not suggests individuals' imperfect ability to incorporate all relevant information in forming these beliefs. As argued in Finkelstein and McGarry (2006), if B_{it}^{NH} is a sufficient statistic for private information about nursing home use, then conditional on B_{it}^{NH}, all other individual information (including B^{IC}) should have no power in predicting nursing home use.

[19]I also verified that I obtain statistically insignificant relationship between B_{it}^{NH} and subsequent nursing home utilization when I include nursing home stays lasting less than 100 days, measure subsequent nursing home use over a longer time horizon, or use individuals of younger ages.

Adverse selection in the post-purchase utilization phase. One potential concern in using the results in Table 1 as descriptive evidence for adverse selection based on consumers' beliefs about receiving informal care from children is that the negative relationship between initial informal care beliefs and subsequent nursing home use might be non-linear in long-term care insurance ownership. For example, one might think that when a parent with long-term care insurance is hit by an adverse health shock, regardless of his or her initial beliefs about receiving informal care, the parent will always use formal care as it is de facto free.

I provide descriptive evidence against such a concern by showing that beliefs about receiving informal care result in adverse selection even after the purchase of insurance. To this end, I reestimate Equation (1) on a sample of long-term care insurance owners who currently have limitations in performing at least two out of the five ADLs asked in the HRS.[20]. The exact equation I estimate is:

$$NH_{i,t-2\sim t} \quad = \quad \beta_0 + \beta_1 B_{i0}^{IC} + X_{i0}' \beta_2 + error_{it}. \tag{3}$$

The dependent variable $NH_{i,t-2\sim t}$ measures the disabled insurance owner's nursing home use in the last two years.[21] The key control is B_{i0}^{IC} which represents whether the individual thought *initially* that he or she would receive informal care from children. This variable, together with pricing controls X_{i0}, are measured when the individual was healthy and observed to own long-term care insurance for the first time. Therefore, B_{i0}^{IC} represents the residual private information about receiving informal care that the individual possessed when he or she selected into insurance.

Column (1) in Table 2 shows that sick long-term care insurance owners who a priori thought they would receive informal care from children are 8.7 percentage points less likely to have a nursing home stay lasting more than 100 nights. This is substantial as the mean nursing home utilization rate is 17.5%. Column (3) uses the number of nights spent in a nursing home in the past two years as the dependent variable. It shows that individuals who initially thought their children would provide informal care spend about 42 less nights, which is considerable given the mean nursing home night of 85.[22]

To address another potential concern that the predictive power of B_{i0}^{IC} might only hold for long-term care insurance owners with relatively minor long-term care needs and hence negligible formal care expenses, using again the sample of long-term care insurance owners with two or more ADL limitations, I reestimate Equation (1) allowing the predictive power of B_{i0}^{IC} to vary by the severity of long-term care needs.[23] Individuals with 2-3 ADL limitations are categorized as having

[20] The five ADLs asked in the HRS are bathing, dressing, eating, getting in/out of bed and walking across a room. Note that the vast majority of long-term care insurance contracts specify that for care to be reimbursable, the individual must need assistance in at least two ADLs (Brown and Finkelstein, 2007).

[21] As I will discuss later, the results are robust to measuring nursing home utilization over a longer time horizon.

[22] Note that standard errors are relatively larger in Columns (3) and (4). This may be due to the fact that the dependent variable includes short nursing home stays that are intended for acute medical recovery rather than long-term care.

[23] The exact regression that I estimate is: $NH_{i,t-2\sim t} = \beta_0 + \beta_1 \mathbb{I}[H_{it} = \text{severe}] + \beta_2 \mathbb{I}[H_{it} = \text{minor}] B_{i0}^{IC} +$

Table 2: Predictive Power of Initial Informal Care Beliefs Among Disabled Insurance Owners

	(1)	(2)	(3)	(4)
Dependent variable, Y:	Use NH		NH nights	
B^{IC}	-0.087*		-41.904*	
	(0.048)		(25.592)	
$\mathbb{I}[H = \text{minor}] \times B^{IC}$		-0.012		-7.548
		(0.046)		(22.531)
$\mathbb{I}[H = \text{severe}] \times B^{IC}$		-0.233**		-102.920*
		(0.112)		(59.604)
Pricing controls, X	Yes	Yes	Yes	Yes
Means				
$E(Y)$	0.175	0.175	85.493	85.493
$E(Y \mid H = \text{minor})$	0.085	0.085	38.477	38.477
$E(Y \mid H = \text{severe})$	0.391	0.391	197.891	197.891
Observations	217	217	217	217

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors are clustered at the household level and reported in parentheses. The sample consists of long-term care insurance owners who currently have two or more ADL limitations. Individuals with 2-3 ADL limitations are categorized as having minor long-term care needs (H = minor), while those with 4-5 ADL limitations are categorized as having severe needs (H = severe). B^{IC} is an indicator for whether an individual initially thought his or her children would provide informal care, measured when the individual did not have any ADL limitations and was first observed to own long-term care insurance. Columns (1) and (2) use an indicator for having a nursing home stay lasting more than 100 days in the past two years as the dependent variable. They report estimates from linear probability models. Columns (3) and (4) use the number of nights spent in a nursing home in the past two years as the dependent variable. They report estimates from OLS regressions. All four regressions include buyer characteristics used by insurers in pricing as controls.

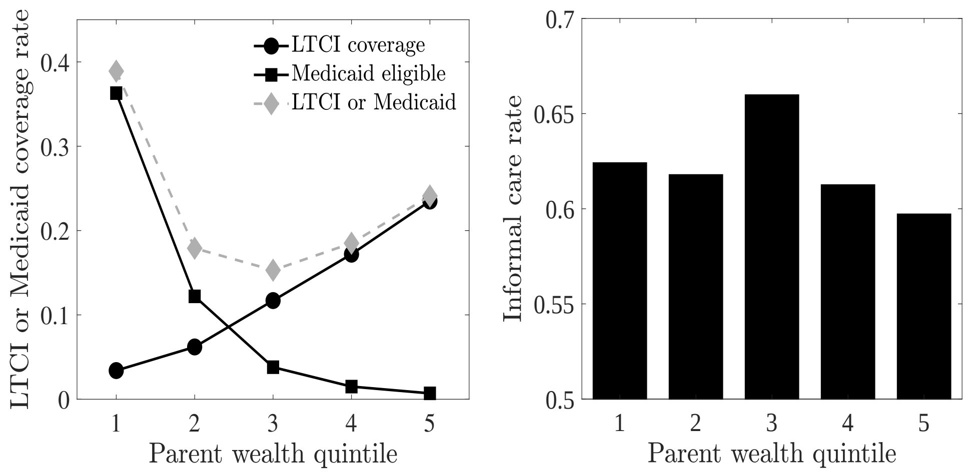
minor long-term care needs (H_{it} = minor), while those with 4-5 ADL limitations are categorized as having severe needs (H_{it} = severe).

Columns (2) and (4) of Table 2 show the estimation results. Among individuals with severe conditions, those with initial beliefs $B_{i0}^{IC} = 1$ are 23.3 percentage points less likely to have a long nursing home stay and spend 103 less nights in a nursing home. These are considerable as the mean nursing home utilization rate is 39.1%, and the mean nursing home night is 198 for this group of severely disabled individuals. For individuals with minor conditions, the estimated correlation between initial informal care beliefs and formal care use is also negative but it lacks statistical significance. As formal care expenses for this health group are small (the mean nursing home night is only 38), it does not change the result that initial beliefs about receiving informal care from children generate substantial adverse selection. I also verified that the results are robust to measuring nursing home utilization over a longer time horizon.

To sum, the results reported in Tables 1 and 2 provide descriptive evidence that private information about the likelihood of receiving informal care from children results in adverse selection not just in the initial stage of insurance purchase, but also in the post-purchase utilization phase.

$\beta_3 \mathbb{I}[H_{it} = \text{severe}] B_{i0}^{IC} + X_{i0}' \beta_4 + error_{it}$. The key coefficients of interest are β_2 and β_3.

Figure 1: Insurance Coverage and Informal Care Receipt



Notes: The left panel reports the long-term care insurance coverage rate and the share of Medicaid eligibles. The right panel reports the share of individuals with long-term care needs who receive informal care from their children.

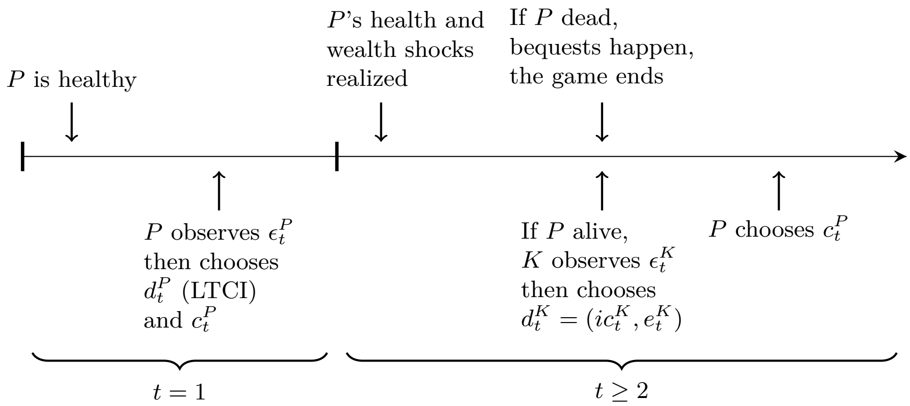
Furthermore, the fact that parents forgo almost free long-term care when informal care is likely implies that they may prefer informal care to formal care.[24]

2.3 Informal care and bequests

Several theoretical papers like Bernheim, Shleifer, and Summers (1985), Pauly (1990), Zweifel and Struwe (1996) and Courbage and Zweifel (2011) have argued that parents may forgo insurance because it reduces children's inheritances exposed to risk, thereby reducing the effectiveness of bequests in eliciting favorable behaviors from children. If children provide informal care in part to protect their inheritances from formal care, then one could expect a positive relationship between children's informal care provision and out-of-pocket prices of formal care that their parents face. The left panel in Figure 1 reports the long-term care insurance coverage rate and the share of Medicaid eligibles by wealth quintile. The long-term care insurance coverage rate increases in wealth while the share of Medicaid eligibles decreases in wealth. Individuals in the middle of the wealth distribution face the largest out-of-pocket costs of formal care as the share covered by either long-term care insurance or Medicaid is the lowest. Indeed, the right panel in Figure 1 shows that there is an inverted-U pattern of informal care receipt, with middle-wealth parents receiving most informal care from children. Such data patterns suggest that inheritances exposed to formal care spending risk may be important in shaping children's informal care decisions.

[24] I verified that this pattern again emerges among Medicaid beneficiaries who also face free long-term care as the government pays all the cost: Medicaid beneficiaries who thought a priori that their children would provide informal care use substantially less formal care when hit by a severe health shock.

Figure 2: Timing of Events



Notes: P denotes the parent and K denotes the kid. d_t^P is the parent's once-and-for-all long-term care insurance (LTCI) purchase, and ϵ_t^P represents associated preference shocks. c_t^P is the parent's consumption. d_t^K comprises the child's informal care provision (ic_t^K) and employment (e_t^K), and ϵ_t^K represents associated preference shocks.

3 Intergenerational Game

The model presented in this section describes interactions between an elderly parent and an adult child from the parent's retirement to death. In the first period, the parent makes a long-term care insurance purchase decision taking into account the likelihood of the child providing informal care when needed. In each of the later periods, the parent experiences health and wealth shocks. The child allocates time to informal care provision, work and leisure. When the parent's long-term care needs are realized, the parent uses formal care only when the child decides not to provide informal care. The parent pays for formal care either using her long-term care insurance, Medicaid benefits or savings. In case of the parent's death, the child inherits the parent's wealth. Figure 2 summarizes the timing of events. For now, I abstract from the supply side of the long-term care insurance market and assume standard long-term care insurance policies are sold at a given price. I explicitly introduce the supply side and define the insurance market equilibrium in Section 5 where I present counterfactuals.

3.1 Environment

Variables related to the parent will have superscript P, and variables related to the child will have superscript K. Time, indexed by t, is discrete and finite. A period corresponds to two years as the HRS interviews are conducted biennially. There are uncertainties about the parent's health, $h_t^P \in \{0, 1, 2, 3\}$, which is defined based on the parent's long-term care needs and mortality: the parent can be healthy ($h_t^P = 0$), have light long-term care needs ($h_t^P = 1$), have severe long-term care needs ($h_t^P = 2$), or be dead ($h_t^P = 3$). In the first period, the parent is healthy and is 60 years old. The game ends when the parent dies, and the parent dies for sure at age 100.

Choices. In the first period, the parent is assumed to be healthy, and the parent decides whether to buy long-term care insurance once-and-for-all, $d_t^P \in \{0,1\}$, and how much to consume, c_t^P.[25] From the second period on, the parent's health shock is realized and observed by both the parent and the child. If the parent is alive, the child moves first by choosing a discrete choice vector $d_t^K = (ic_t^K, e_t^K)$ comprising informal care provision, ic_t^K, and employment, e_t^K. The child's informal care provision decision is binary, $ic_t^K \in \{0,1\}$, and I assume the child never provides informal care when the parent is healthy.[26] The child's employment $e_t^K \in \{0,1,2\}$ can take three values: no work ($e_t^K = 0$), part-time work ($e_t^K = 1$) and full-time work ($e_t^K = 2$). The feasible choice set for the child is therefore $\{d_t^K = (ic_t^K, e_t^K)|(0,0),(0,1),(0,2)\}$ if the parent is healthy, $h_t^P = 0$, and $\{d_t^K = (ic_t^K, e_t^K)|(0,0),(0,1),(0,2),(1,0),(1,1),(1,2)\}$ if the parent is sick, $h_t^P \in \{1,2\}$. From the second period on, the parent only decides how much to consume, and the parent makes the consumption decision after observing the child's informal care and employment choices. Note that the parent's formal care use is effectively determined by the child's informal care provision decision: when the parent is hit by an adverse health shock, the parent uses formal care and pays the relevant cost only if the child decides not to provide informal care.[27]

Preferences. The child's flow utility while the parent is alive is:

$$\pi^K(c_t^K, l_t^K, ic_t^K; h_t^P, ic_{t-1}^K, X^K) + \epsilon_t^K(d_t^K) \tag{4}$$

where

$$\pi^K(c_t^K, l_t^K, ic_t^K; h_t^P, ic_{t-1}^K, X^K) = \theta_c^K \frac{(c_t^K)^{1-\rho_c^K}-1}{1-\rho_c^K} + \theta_l^K \frac{(l_t^K)^{1-\rho_l^K}-1}{1-\rho_l^K} + \omega^K(ic_t^K; h_t^P, ic_{t-1}^K, X^K) \tag{5}$$

and

$$\omega^K(ic_t^K; h_t^P, ic_{t-1}^K, X^K) = \begin{cases} 0 & \text{if} \quad ic_t^K = 0, \\ \theta_{h_t^P}^K + \theta_{male}^K male^K + \theta_{far}^K far^K + \theta_{start}^K \mathbb{I}[ic_{t-1}^K = 0] & \text{if} \quad ic_t^K = 1. \end{cases} \tag{6}$$

The child has additively separable preferences for consumption c_t^K, leisure l_t^K and informal care provision ic_t^K. $\epsilon_t^K(d_t^K)$ is an additive preference shock associated with discrete choice vector d_t^K and follows an *i.i.d.* Type I extreme value distribution with scale one. The child privately observes ϵ_t^K before she chooses d_t^K. The child's consumption and leisure preferences follow a constant

[25]The assumption that the decision to buy long-term care insurance is once-and-for-all is empirically grounded. First, the average age of buyers is 61, and 80% of the sales are made to consumers aged between 50 and 69, implying that elderly Americans typically make a decision about whether to purchase insurance as they enter retirement (Broker World, 2009-2015). Second, most insurers do not sell contracts to individuals older than 70: sales made to individuals aged 70+ account for less than 5% (Broker World, 2009-2015). Third, lapses are very rare: according to Genworth, the biggest long-term care insurer in the U.S., the lapse rate is only 0.7% per year (https://www.latimes.com/business/story/2019-07-24/long-term-care-insurance-disaster).

[26]In the data, almost no child provides care to parents without any daily activity limitations.

[27]This assumption implies that informal and formal care are substitutes, as found in empirical studies like Charles and Sevak (2005) and Coe, Goda, and Van Houtven (2015).

relative risk aversion utility function. The function ω^K, which is largely based on Skira (2015), represents the child's preference for informal care provision and captures the child's impure altruism towards the parent. This warm-glow utility is normalized to zero for no caregiving. When the child does provide care, the warm-glow utility depends on the severity of the parent's health condition $h_t^P \in \{1,2\}$, $male^K$ which is an indicator for whether the child is a male and is contained in the child demographic vector X^K, far^K an indicator for whether the child lives outside a 10-mile radius to the parent also contained in X^K, and $\mathbb{I}[ic_{t-1}^K = 0]$ an indicator for whether the child did not provide care in the previous period. The possible dependence upon ic_{t-1}^K is to incorporate costs associated with initiating informal care provision such as adjusting to new environments, learning how to take care of disabled parents and changing one's schedule in a substantial manner. The inclusion of $male^K$ and far^K is motivated by the data pattern that children's informal care behaviors vary substantially by gender and residential proximity to parents.[28] As these child characteristics are not used by insurers in pricing, they generate private information about the informal care likelihood.

The parent's flow utility while the parent is alive is

$$\pi^P(c_t^P, ic_t^K; h_t^P) + \epsilon_t^P(d_t^P) \tag{7}$$

where

$$\pi^P(c_t^P, ic_t^K; h_t^P) = \theta_c^P \frac{(c_t^P + c_{nh}\mathbb{I}[h_t^P = 2 \text{ and } ic_t^K = 0])^{1-\rho_c^P} - 1}{1-\rho_c^P} + \theta_{fc}^P \mathbb{I}[h_t^P \in \{1,2\} \text{ and } ic_t^K = 0]. \tag{8}$$

The parent has additively separable preferences for consumption and long-term care.[29] $\epsilon_t^P(d_t^P)$ is the insurance choice-specific preference shock realized only in the first period and follows an *i.i.d.* Type I extreme value distribution with scale one. It is privately observed by the parent before she makes the insurance decision. The parent's preference for consumption follows a constant relative risk aversion utility function. When the parent has severe long-term care needs ($h_t^P = 2$) and uses formal care, which is assumed to be inevitable when $ic_t^K = 0$, then the parent receives a certain amount of consumption from the formal care use denoted by c_{nh}. The assumption here is that the type of formal care used in the health status $h_t^P = 2$ is nursing home care, which provides food and housing. θ_{fc}^P represents the parent's utility from using formal care in the health status $h_t^P \in \{1,2\}$. If the parent prefers informal care over formal care, then θ_{fc}^P will be negative.

Terminal values. When the parent dies ($h_t^P = 3$), the game ends, and the child inherits the parent's remaining wealth w_t^P. To close the model, I assume when the parent dies, the child does not provide informal care, chooses no work and optimally consumes her inheritance over the next T_0 periods.[30] Such assumptions are sufficient to derive a closed-form terminal value for the child

[28] Table 4 in Section 4 presents these differences. Furthermore, Table A.2 in Appendix A shows that the child's gender and residential proximity to the parent are highly predictive of whether the parent believes the child will provide informal care.

[29] The parent's flow utility is not a function of leisure because I assume that (1) the parent is retired and spends all available time on leisure, and (2) her leisure utility is additively separable.

[30] Making assumptions about terminal values in a finite life-cycle model where an economic agent does

denoted by $\pi_d^K(w_t^P)$.[31] As the parent does not incur any formal care expenses while receiving informal care from the child, the child may be strategically motivated to provide informal care to increase her inheritance utility. The parent's bequest utility exhibits pure altruism in the sense that it is a function of the child's inheritance utility:

$$\pi_d^P(w_t^P) = \theta_d^P \pi_d^K(w_t^P). \tag{9}$$

This specification is based on Kaplan (2012) who also uses such pure altruism setup to study intergenerational transfers between non-elderly parents and young adult children.[32]

Long-term care insurance. I consider one standardized private long-term care insurance policy which pays benefits for formal care expenses only when the parent is unhealthy ($h_t^P \in \{1, 2\}$), has a maximal per-period benefit cap b and provides coverage for life.[33] The per-period premium is p and is paid in every period when the parent is not receiving benefits from the insurance company. During the sample period, premiums varied only by age and health. As all parents are assumed to be healthy in the first period, which is the only period where they are able to purchase insurance, the premium p is the same for all buyers.

Budget constraints. The child's consumption and leisure are determined according to the following budget and time constraints, respectively:

$$c_t^K = y^K(e_t^K; \mathbb{I}[e_{t-1}^K = 2], age_t^K, X^K), \tag{10}$$

$$l_t^K = T_{total} - T_{ic_t^K, h_t^P} - T_{e_t^K}. \tag{11}$$

Equation (10) states that the child does not save,[34] and her consumption is equal to her income

not die in the terminal period is also used in Kaplan (2012).

[31] The exact specifications of $\pi_d^K(w_t^P)$ are found in Appendix B.

[32] One alternative way to model the parent's bequest utility is to use impure altruism and specify the bequest utility as a direct function of w_t^P with some assumptions about its functional forms. Such impure altruism is used by De Nardi (2004), De Nardi, French, and Jones (2010) and Lockwood (2018) in the context of a single agent elderly life-cycle savings model. As my model is a game in which both the parent and the child value bequests, it is more beneficial to incorporate the parent's bequest utility as pure altruism to avoid making additional assumptions about functional forms.

[33] These features are based on typical long-term care insurance products sold during my sample period (Brown and Finkelstein, 2007; Broker World, 2009-2015).

[34] Note that the HRS does not provide information about children's assets. On the one hand, abstracting from child savings could in theory underpredict informal care provision. With savings allowed, the child could self-insure against possible earnings loss resulting from reducing work to care for the disabled parent. That is, the caregiving child could mitigate the effect of the earnings loss by drawing down his/her assets. However, Skira (2015) finds that there is no descriptive evidence that caregiving children experience significantly different changes in assets or savings than non-caregiving children. On the other hand, abstracting from child savings could overpredict informal care provision as it precludes the case where children with more savings substitute away from informal care and help their parents pay for formal care. However, I find that among disabled parents who use formal long-term care, the mean financial transfer received from children is merely $300 annually. These empirical facts suggest that in reality, children's informal care decisions are not much affected by their assets.

y^K whose log value is equal to

$$\underbrace{\gamma_1 + \gamma_2 age_t^K + \gamma_3 (age_t^K)^2 + \gamma_4 home^K + \gamma_5 mar^K}_{\text{non-labor income}} + \underbrace{\gamma_6 \mathbb{I}[e_t^K = 1]}_{\text{part-time labor income}}$$
$$\underbrace{+ \mathbb{I}[e_t^K = 2] * \left\{ \gamma_7 + \gamma_8 age_t^K + \gamma_9 (age_t^K)^2 + \gamma_{10} edu^K + \gamma_{11} \mathbb{I}[e_{t-1}^K = 2] \right\}}_{\text{full-time labor income}}. \tag{12}$$

The non-labor income depends on the child's age age_t^K, age squared, home ownership $home^K$ and marital status mar^K. The child's full-time labor income depends on the child's age, age squared, education edu_t^K and whether the child worked full-time in the previous period. The variables $home^K$, mar^K and edu_t^K are contained in the child's demographic vector X^K. The dependence of the full-time income upon $\mathbb{I}[e_{t-1}^K = 2]$ is to capture possible penalties for being out of the workforce in the previous period.[35] Note that while the child characteristics that enter the income function affect the child's cost of providing informal care, none are used by insurers in pricing insurance contracts.

The child's leisure hours are residually determined by the time constraint in Equation (11) where T_{total} is the child's total endowed time, $T_{ic_t^K, h_t^P}$ is the associated help hours for informal care choice ic_t^K when the parent's health is h_t^P, and $T_{e_t^K}$ is the required work hours for employment choice e_t^K.

The model incorporates means-tested Medicaid as a consumption floor for the parent. If the parent's net assets after paying the insurance premium and incurring out-of-pocket formal care expenses falls below a certain threshold, the government provides transfers. The parent's wealth after receiving government transfers (if any) is

$$\hat{w}_t^P = \max\left\{ w_t^P + y^P - p - \left(x_{h_t^P, n^P} - \min\left\{ b, x_{h_t^P, n^P} \right\} \right), \bar{w}_g \right\}. \tag{13}$$

y^P is permanent income. $x_{h_t^P, n^P}$ represents formal care expenses before any transfers from insurance. As mentioned earlier, the parent incurs formal care expenses ($x_{h_t^P, n^P} > 0$) only when the parent is disabled ($h_t^P \in \{1, 2\}$) *and* the child decides not to provide informal care. I assume the parent's severity of long-term care needs determines the type of formal care used: the parent uses paid home care when $h_t^P = 1$ and nursing home care when $h_t^P = 2$. How intensively the parent uses formal care depends on the number of children the parent has, n^P. This is to incorporate, in a reduced-form way, the possibility that the child in the model may not be the only source of informal care. This is motivated by the empirical pattern that among non-childless individuals with long-term care needs, the number of nights spent in a nursing home is lower by almost 50% for individuals with 4 or more children compared to those with 3 or less. Private insurance benefits

[35] This modeling choice captures potential dynamic costs of providing care: when a child quits full-time work to care for the parent, she incurs not only static costs (current forgone wages) but also dynamic costs in the form of lower future wages when she returns back to the workforce. Skira (2015) argues that such dynamic considerations are important in accurately capturing children's trade-off between informal care provision and labor supply.

b are strictly positive only when the parent has private insurance, is sick and uses formal care. Insurance premium payment p is strictly positive only when the parent has insurance and is not receiving any insurance benefits ($b = 0$). $\bar{w}_g$ represents the level of consumption floor ensured by the government where

$$\bar{w}_g = \begin{cases} \bar{w}_{\text{low}} & \text{if } h_t^P = 2 \text{ and } ic_t^K = 0, \text{ i.e., in a nursing home,} \\ \bar{w}_{\text{high}} - p & \text{otherwise.} \end{cases} \tag{14}$$

I assume $\bar{w}_{\text{low}} < \bar{w}_{\text{high}}$ based on Medicaid's stringent restrictions on assets for nursing home residents. For non-nursing home residents, the level of the floor depends on insurance premiums as social insurance does not pay for private insurance (Lockwood, 2018). It is important to note that Medicaid is a secondary payer in the sense that long-term care insurance must pay the benefits first.

The parent's wealth at the beginning of the next period is given by

$$w_{t+1}^P = \max\left\{0, (1+r)\left(\hat{w}_t^P - c_t^P\right) - m_{t+1}^P\right\} \tag{15}$$

where r is the real per-period interest rate, and m_{t+1}^P is an *i.i.d.* wealth shock realized at the beginning of the next period for which the parent is liable up to $(1+r)\left(\hat{w}_t^P - c_t^P\right)$. There is no borrowing, and the parent's consumption is constrained by $c_t^P \leq \hat{w}_t^P$.

Health transitions. The parent's health transition probabilities follow a Markov chain and depend on the parent's current health, gender, age and permanent income. This suggests that the parent's health transition process is treated as exogenous and does not depend on the receipt of informal or formal care. This is based on previous studies that find the evolution of long-term care needs and mortality is largely unaffected by the receipt of care, and the primary role of long-term care lies in reducing discomfort experienced by the elderly with everyday task limitations (Byrne, Goeree, Hiedemann, and Stern, 2009).

State space. The set of state variables that are commonly observed by the parent and child at the beginning of period t is:

$$s_t = (age_t^P, h_t^P, w_t^P, ltci_t^P, age_t^K, ic_{t-1}^K, \mathbb{I}[e_{t-1}^K = 2]; X^P, X^K).$$

age_t^P and age_t^K denote the age of the parent and child, respectively. $ltci_t^P$ is an indicator for whether the parent has private long-term care insurance. It is determined in the first period as a result of the parent's long-term care insurance decision and is fixed from then on. As explained earlier, ic_{t-1}^K enters the child's informal care utility to capture possible costs associated with initiating informal care provision. $\mathbb{I}[e_{t-1}^K = 2]$ enters the child's income function to capture possible dynamic costs of quitting work to provide informal care. X^P is a vector of parental demographics including the parent's gender, permanent income and number of children. X^K is a vector of child demographics including the child's gender, education, marital status, home ownership and residential proximity to the parent. All variables in s_t evolve deterministically except for the parent's health and wealth.

3.2 Equilibrium of the intergenerational game

Strategy profile. To define equilibrium decision rules of the family, I first define a strategy profile $\sigma = (\sigma^K, \sigma^P)$ comprising a set of decision rules for the child and parent. $\sigma^K = \{\sigma^K(s_t, \epsilon_t^K)\}$ is a mapping from the set of common states and child preference shocks to the set of feasible informal care and employment choices $\mathbb{C}^K(h_t^P)$, which depends on the parent's health. As the child can provide informal care only when the parent is sick, $\mathbb{C}^K(h_t^P = 0) = \{d_t^K = (ic_t^K, e_t^K) \mid$ (0,0), (0,1), (0,2)$\}$ and $\mathbb{C}^K(h_t^P \in \{1,2\}) = \{d_t^K = (ic_t^K, e_t^K) \mid$ (0,0), (0,1), (0,2), (1,0), (1,1), (1,2)$\}$.

$\sigma^P = (\sigma^{P,d}, \sigma^{P,c})$ consists of the parent's insurance decision rule $\sigma^{P,d}$ which is relevant only in $t=1$ and consumption decision rule $\sigma^{P,c}$. $\sigma^{P,d} = \{\sigma^{P,d}(s_t, \epsilon_t^P)\}$ is a mapping from the set of common states and parent preference shocks to the insurance choice set, $\{0, 1\}$. The parent's consumption decision rule $\sigma^{P,c} = \{\sigma^{P,c}(s_t, d_t^K, d_t^P)\}$ is a mapping to $(0, \hat{w}_t^P]$ where $\hat{w}_t^P$ is the parent's net assets before consumption and is defined earlier in Equation (13). Note that while the consumption decision rule is specified generally as a function of (s_t, d_t^K, d_t^P), in $t=1$, d_t^K is irrelevant as the child does not make decisions, and in $t \geq 2$, d_t^P is irrelevant as insurance decisions are made only in the first period.

Child's value functions. Let $\tilde{V}^K(s_t, \epsilon_t^K; \sigma)$ denote the child's value conditional on state s_t and the realization of her private preference shocks ϵ_t^K if she behaves optimally today and in the future when the parent behaves according to her decision rules specified in σ. In states where the parent is dead, with a slight abuse of notation, define $\tilde{V}^K(\cdot) = \pi_d^K(w_t^P)$ where $\pi_d^K(w_t^P)$ is the child's inheritance utility. In each period while the parent is alive, the child's value function is represented as

$$\tilde{V}^K(s_t, \epsilon_t^K; \sigma) = \max_{d_t^K \in \mathbb{C}^K(h_t^P)} \left\{ \pi^K(c_t^K, l_t^K, ic_t^K; h_t^P, ic_{t-1}^K, X^K) + \epsilon_t^K(d_t^K) + \beta E\left[\tilde{V}^K(s_{t+1}, \epsilon_{t+1}^K; \sigma) \middle| s_t, d_t^K; \sigma \right] \right\} \quad (16)$$

where β is the discount factor, and the expectation is over the parent's health and wealth shocks and child's preference shocks of the next period. Define $V^K(s_t; \sigma)$ as the child's expected value function $V^K(s_t; \sigma) = \int \tilde{V}^K(s_t, \epsilon_t^K; \sigma) g(\epsilon_t^K) d\epsilon_t^K$ where g is the PDF of ϵ_t^K. Define the child's choice-specific value function, $v^K(s_t, d_t^K; \sigma)$, as the per-period payoff of choosing d_t^K minus the preference shock plus the expected value function:

$$v^K(s_t, d_t^K; \sigma) = \pi^K(c_t^K, l_t^K, ic_t^K; h_t^P, ic_{t-1}^K, X^K) + \beta E\left[V^K(s_{t+1}; \sigma) \middle| s_t, d_t^K; \sigma \right]. \quad (17)$$

Parent's value functions. Let $\tilde{V}^P(s_t, d_t^K; \sigma)$ denote the parent's value conditional on state s_t and the child's choice vector d_t^K if the parent behaves optimally today and in the future when the child behaves according to her decision rules specified in σ. Again, with a slight abuse of notation, define $\tilde{V}^P(\cdot) = \pi_d^P(w_t^P)$ for $h_t^P = 3$ where $\pi_d^P(w_t^P)$ is the parent's bequest utility. For $t \geq 2$, $\tilde{V}^P(s_t, d_t^K; \sigma)$

is defined as

$$\tilde{V}^P(s_t, d_t^K; \sigma) = \max_{c_t^P \in (0, \hat{w}_t^P]} \left\{ \pi^P(c_t^P, ic_t^K; h_t^P) + \beta E\left[\tilde{V}^P(s_{t+1}, d_{t+1}^K; \sigma) \middle| s_t, d_t^K, c_t^P; \sigma \right] \right\} \tag{18}$$

where the expectation is over the parent's wealth and health shocks and child's preference shocks of the next period. Define $V^P(s_t; \sigma) = E\left[\tilde{V}^P(s_t, d_t^K; \sigma)\right]$ as the parent's expected value function before the parent observes the child's actions. The expectation is over the child's preferences shocks of the current period ϵ_t^K, which determine the child's choice vector d_t^K according to the child's decision rule $\sigma^K(s_t, \epsilon_t^K)$. For $t = 1$, in addition to choosing c_t^P, the parent makes a once-and-for-all insurance choice $d_t^P \in \{0, 1\}$. I denote the parent's insurance choice-specific value function as $v^P(s_t, d_t^P; \sigma)$, and it is defined as the parent's per-period payoff of choosing insurance choice d_t^P minus the preference shock plus her expected value function,

$$v^P(s_t, d_t^P; \sigma) = \pi^P(c_t^P, ic_t^K; h_t^P) + \beta E\Big[V^P(s_{t+1}; \sigma)\Big| s_t, d_t^P, c_t^P; \sigma\Big] \tag{19}$$

where $c_t^P = \sigma^{P,c}(s_t, d_t^K, d_t^P)$, i.e., the first period consumption according to σ, and $ic_t^K = 0$ as there is no informal care provision in the first period.

Markov-perfect equilibrium. A strategy profile $\sigma^* = (\sigma^{*K}, \sigma^{*P})$ is a Markov perfect equilibrium (MPE) of the intergenerational game if and only if all of the following conditions are satisfied.

Optimality of the child's informal care and employment choices:

$$\sigma^{*K}(s_t, \epsilon_t^K) = \underset{d_t^K \in \mathbb{C}^K(h_t^P)}{\text{argmax}} \left\{ v^K(s_t, d_t^K; \sigma^*) + \epsilon_t^K(d_t^K) \right\} \text{ for any } (s_t, \epsilon_t^K) \text{ where } t \geq 2, \tag{20}$$

Optimality of the parent's insurance choice:

$$\sigma^{*P,d}(s_t, \epsilon_t^P) = \underset{d_t^P \in \{0,1\}}{\text{argmax}} \left\{ v^P(s_t, d_t^P; \sigma^*) + \epsilon_t^P(d_t^P) \right\} \text{ for any } (s_t, \epsilon_t^P) \text{ where } t = 1, \tag{21}$$

Optimality of the parent's consumption choice:

$$\sigma^{*P,c}(s_t, d_t^K, d_t^P) = \underset{c_t^P \in (0, \hat{w}_t^P]}{\text{argmax}} \left\{ \pi^P(c_t^P, ic_t^K; h_t^P) + \beta E\Big[V^P(s_{t+1}; \sigma^*)\Big| s_t, d_t^K, d_t^P, c_t^P; \sigma^*\Big] \right\} \tag{22}$$

for any (s_t, d_t^K, d_t^P).

As the model is finite, and within each period there are sequential moves by the players (the child moves first followed by the parent), the model has a unique equilibrium. What is worth emphasizing is that in equilibrium, the parent makes the long-term care insurance decision based on her beliefs about receiving informal care from the child in the future, which are given by the child's optimal informal care choice probabilities.

3.3 Model discussions

Adverse selection based on children's informal care likelihood. The model incorporates asymmetric information about the probability of receiving informal care by assuming that there are family demographics that affect the child's informal care provision decisions but are not priced by insurers. The child's informal care utility $\omega^K(\cdot)$ and opportunity cost of care captured in the income function $y^K(\cdot)$ depend on various child characteristics, as shown in Equations (6) and (12), respectively. To provide descriptive evidence that these "unpriced" characteristics are the key source of private information about children's informal caregiving, in Table A.2 of Appendix A, I show that the child characteristics, especially whether a child is a daughter and lives close to the parent, are powerful predictors of whether the parent believes the child will provide informal care.[36] As insurance premiums do not vary by child demographics, the parent with a high probability of receiving informal care (and hence lower formal care risk) forgoes insurance rendering the market adversely selected.[37]

Advantageous selection based on income. Finkelstein and McGarry (2006) argue that both adverse selection and advantageous selection exist in the long-term care insurance, which have offsetting effects on the correlation between insurance ownership and nursing home risk. To replicate the low correlation, the model incorporates income-based advantageous selection. The estimated health transition probabilities presented later in Table 5 in Section 4 show that higher-income individuals are less likely to require long-term care. Nevertheless, they could have higher willingness to pay for private insurance than lower-income individuals since Medicaid serves as a substantially worse substitute for them due to its means-tested nature. Selection of higher-income individuals into insurance is advantageous because they are healthier, and private insurance always pays first due to Medicaid's secondary payer status. This is consistent with Finkelstein and McGarry (2006) who show that individuals with higher wealth are less likely to go into a nursing home but are more likely to have long-term care insurance. If income-based selection is advantageous enough, then it will offset the positive correlation between insurance ownership and formal care risk induced by adverse selection.

Strategic non-purchase of insurance. The model allows for the strategic non-purchase of insurance by assuming that the parent can affect the informal care incentives that the child faces but not their actions. The parent with long-term care insurance can no longer use bequests to induce caregiving behaviors from her child. The reason is because the child knows that even if she does not provide informal care, her inheritances will not be spent on formal care as the insurance company will pay for the cost. Under the assumption that the parent prefers informal care to

[36] Another reason for capturing heterogeneous informal care probabilities through these child characteristics is because I want to use the model to simulate realistic counterfactuals that could reduce the amounts of private information about the probability of receiving informal care. This is not possible if such private information is inherently unobserved.

[37] In Table A.1 of Appendix A, I show descriptive evidence that parents indeed expect to use less formal care when they believe their children will provide informal care in the future.

formal care, the parent might choose not to buy insurance to avoid distorting the incentives faced by the child. The parent's relative preference for informal care serves as a vital ingredient of the strategic non-purchase of insurance and will be structurally estimated.

Non-cooperative interactions. The primary reason for choosing the non-cooperative framework over a cooperative one is to allow for the strategic non-purchase of insurance, investigated by several theoretical papers as an important explanation for the limited size of the long-term care insurance market (for example, Pauly (1990)). There are also modeling issues that arise in using cooperative models. The distribution of assets between family members is indeterminate in cooperative models with full commitment. In the presence of Medicaid eligibility, which considers only parents' assets, these models predict unrealistic outcomes where children own all of their parents' assets. This prediction does not hold in the data as a substantial share of the parents hold on to their assets until death. Mommaerts (2016) overcomes this problem by using a cooperative model with limited commitment where parents and children own separate assets which affect their outside option of non-cooperation. But her cooperative framework does not allow for the strategic non-purchase of insurance. Furthermore, in my estimation sample, about 87% of elderly parents and adult children belong to separate households. This suggests that transaction costs associated with cooperation and asymmetric information about one's resources are more of a concern. Lundberg and Pollak (1993) and Castilla and Walker (2013) show that in such cases, intra-household allocations can default to a non-cooperative equilibrium.

The number of children and required formal care usage intensity. As the model has only one adult child who could potentially provide care, it does not capture interactions of multiple siblings.[38] The model, however, does account for the statistical difference in formal care risk by the number of children. This is motivated by the empirical pattern that among non-childless individuals with long-term care needs, the number of nights spent in a nursing home is lower by almost 50% for individuals with 4 or more children compared to those with 3 or less. How this assumption affects informal care decisions of the child in the model is worth a discussion. The child with more siblings is less strategically motivated to provide informal care, because her inheritances are exposed to smaller formal care costs. As a result, the parent with more children faces (1) a higher probability of using formal care, but (2) lower formal care usage intensity. In the estimated model, the second channel is dominant, and consequently, the model replicates the empirical pattern that parents with more children have smaller formal care expenses.

Timing of long-term care decisions. The model assumes in each period where the parent is sick, the child moves first by deciding whether to provide informal care, and the parent uses formal

[38] For example, children may share the burden of caregiving, free ride on other siblings' efforts or compete to secure a greater share of the bequest. Byrne, Goeree, Hiedemann, and Stern (2009) study strategic interactions among multiple children and parents surrounding long-term care arrangements in a static model. Groneck (2017) finds that providing informal care has a positive effect on the amount of bequests received relative to non-caregiving siblings. Brown (2006) finds that a child's caregiving status is positively correlated with expected bequests. Investigating how multiple children and parents interact over insurance and long-term care decisions in a dynamic framework is an interesting direction for future work.

care only when the child decides not to provide care. Under an alternative timing assumption, the parent would move first by making a formal care utilization decision, and the child would decide whether to provide informal care if the parent in the first stage decided not to use formal care. This alternative timing assumption is not very plausible provided that parents prefer informal care to formal care as found in Mommaerts (2016) and Barczyk and Kredler (2018).[39] If parents get a higher utility from receiving informal care than using formal care, it is unlikely that the parents will forgo the chance to receive preferable informal care by choosing to enter a nursing home first, before their children have a chance to decide whether they want to provide informal care.[40] As a robustness check, Appendix D uses a simplified version of the model and shows that the main mechanisms of the model do not change significantly under the alternative timing assumption.

Inter-vivos financial transfers. The model assumes the parent cannot make inter-vivos financial transfers to elicit informal care provision from the child. This assumption is based on existing studies that find that exchange-motivated inter-vivos transfers from parents to children are infrequent and small in the HRS data.[41] Using the HRS 1998-2010 data, I also find that among respondents who receive informal care from children, only 3% say that they pay their children for providing informal care. The model also does not allow the child to make financial transfers to help the parent pay for formal care. In the HRS, among disabled parents who use formal long-term care, only 9% receive any financial assistance from their children, and the mean transfer amount is merely $300 per year.[42] As the average nursing home cost was over $200 per day during the sample period, it suggests that it is reasonable to abstract from children's financial assistance.

4 Estimation of the Intergenerational Game

4.1 Data and sample selection procedure

To estimate the intergenerational game, I use data from the HRS which has surveyed a representative sample of Americans over the age of 50 every two years since 1992. For the estimation, I use seven interview waves which happened biannually from 1998 to 2010.[43] All monetary values

[39] As I noted above, the parent's relative preference for different types of care will be structurally estimated.

[40] While one might think that parents using formal care could still receive informal care from children, this happens infrequently in the data: among families where either informal or formal care is used, only 15% use both types of care.

[41] McGarry and Schoeni (1997) find that there is no evidence that parents provide financial assistance to their children in exchange for caregiving. Brown (2006) reports that 14% of respondents aged 69 and older receive regular care from their children, while only 1% pay a child for informal care. Groneck (2017) finds that the correlation between informal care provision and inter-vivos transfers made in the previous wave before a parent's death is not statistically significant. Barczyk and Kredler (2018) find that the 90th percentile of financial transfers made from living parents to caregiving children is only $500 annually.

[42] Even when I further restrict to parents who have neither long-term care insurance nor Medicaid benefits, the mean transfer amount from children is still around $300 per year.

[43] I exclude the first two waves (1992 and 1994) because most of the key variables I use, including children's informal care provision, are reported starting with the third wave (1996). I exclude the third wave as it has inaccurate asset data (Lockwood, 2018).

presented henceforth are in 2013 dollars, unless otherwise noted.

From 12,177 respondents who were aged 60 and over in 1998 and do not miss any interviews, I restrict to respondents (1) who were single in either 1998 or 2000 which reduces the number of respondents, N, to 5,144, (2) who were retired in either 1998 or 2000 ($N = 4,779$), (3) who had at least one child aged 21 and over in 1998 and alive while the respondent was alive ($N = 4,009$), and finally, (4) who do not have missing values for any of the variables needed to estimate the model ($N = 3,195$).[44]

As the model describes informal care decisions of one adult child, I apply the following rules to select a child when a respondent reports having multiple children who were aged 21 and over in 1998 and alive while the respondent was alive. For a family where informal care was provided in any of the waves, I pick the major caregiving child.[45] For a family where no informal care was provided through out the sample period, I randomly select one child.[46] The final estimation sample consists of 3,195 families and 12,703 family-year observations.[47]

To measure long-term care needs, I use information about ADL limitations, cognitive impairment and use of either informal or formal care. The HRS asks respondents whether they have a difficulty carrying out each of five ADLs (bathing, dressing, eating, getting in/out of bed and walking across a room) and conducts various tests designed to measure cognitive ability.[48] I categorize a respondent as cognitively impaired if she is in the bottom 10% of the cognitive score distribution. I classify a respondent as healthy ($h_t^P = 0$) if she does not receive any long-term care, or she has 0-1 ADL limitation without cognitive impairment. Among individuals who receive some long-term care, I classify an individual as having light long-term care needs ($h_t^P = 1$) if the individual has 2-3 ADL limitations without cognitive impairment, and as having severe long-term care needs ($h_t^P = 2$) if the individual has either 4-5 ADL limitations or cognitive impairment.

The model assumes a parent with long-term care needs receives either informal or formal care, but not both. In my estimation sample, about 15% of the parents classified as having long-term care needs ($h_t^P = 1, 2$) receive both types of care. If a parent reports having used both informal care and nursing home care, I compare the number of informal care days with the number of days spent in a nursing home and assign the type of care with a longer usage. If the parent reports

[44]The last restriction excludes individuals who are observed for just one period as I cannot compute their change in wealth.

[45]The HRS asks respondents which child provides informal care most. I pick the child who is reported as the major caregiving child for most waves.

[46]The model generates heterogeneity in informal care provision by incorporating various child characteristics. If I selected a child with characteristics that highly predict informal care provision for a family where no informal care happens throughout the entire sample period ("no informal care family"), the model would understate heterogeneity in informal care provision. As a small number of parents from these no informal care families experience adverse health shocks during the sample period, the random selection rule has a limited effect in widening the difference in child characteristics by informal care choice in a given period.

[47]For each family, I exclude their last observed interview as I cannot compute the change in wealth during that period.

[48]The cognitive tests include word recall, subtraction, backward number counting, object naming, date naming and president naming.

Table 3: Parent Estimation Sample

	Mean	Median
Age	77	
Female	0.81	
Have 4+ children	0.38	
Wealth ($)	265,297	90,805
Annual income ($)	19,509	15,640
Have light LTC needs	0.11	
Have severe LTC needs	0.08	
Informal care rate		
: among those with light LTC needs	0.61	
: among those with severe LTC needs	0.53	
LTCI ownership rate		
: among healthy individuals aged 60-69	0.21	
Observations	12,703	

Notes: The table reports summary statistics of parents in the estimation sample constructed from the HRS 1998-2010. The sample size is 3,195 families and 12,703 family-year observations. Monetary values are in 2013 dollars.

having used both informal care and paid home care, I assume the type of care is informal.[49]

I measure parent wealth as the net value of total assets less debts, which includes real estate, housing, vehicles, businesses, stocks, bonds, checking and savings accounts and other assets. For the parent's permanent income, I use the sum of employer pension, annuity income, social security retirement income and other income. As the model assumes the parent's income is time-invariant, for each parent in the sample, I compute the average income over the sample period.

The HRS does not ask respondents about their consumption behaviors, but a subsample of the HRS respondents have been selected at random and surveyed about their consumption behaviors in the Consumption and Activities Mail Survey (CAMS). About 25% of my estimation sample is found in the CAMS data. I use these respondents' reported consumption to obtain empirical consumption decision rules in the first stage of the CCP estimation, which I will further describe in Section 4.3.

To obtain data on long-term care insurance choices, I use respondents aged 60-69 in the estimation sample who do not have any conditions that would lead to insurance rejections. Excluding individuals who would be rejected by insurers is important because while the model abstracts from insurance rejections, in reality, a non-trivial fraction of elderly individuals cannot purchase insurance due to rejections (Hendren, 2013; Braun, Kopecky, and Koreshkova, 2019). As my model studies insurance choice of individuals who are able to purchase insurance, I use the insurance coverage rate among individuals who would not be rejected by insurers.

Table 3 presents the summary statistics of parents in the estimation sample. The mean age is 77, about 80% of the parents are female, and about 40% have four or more children. The mean

[49]The HRS does not ask about the intensity of paid home care utilization.

Table 4: Child Estimation Sample

	All	Children not providing care to sick parents	Children providing care to sick parents
Age	50	54	53
Female	0.54	0.46	0.70
Live within 10 mi of the parent	0.50	0.43	0.84
Have some college education	0.49	0.44	0.41
Married	0.63	0.66	0.55
Homeowner	0.66	0.67	0.51
Work full-time	0.67	0.55	0.52
Work part-time	0.08	0.07	0.10
Paid by parent for providing care	-	-	0.03
Observations	12,703	1,075	1,459

Notes: The table reports summary statistics of children in the estimation sample. The sample size is 3,195 families and 12,703 family-year observations. Monetary values are in 2013 dollars. Column (1) uses all children from pooled HRS 1998-2010. Column (2) uses children who do not provide informal care to parents with long-term care needs, while Column (3) uses children who do.

wealth is $265,297, and the mean annual income is $19,509. About 11% of the parents have light long-term care needs, and 8% severe long-term care needs. The share receiving informal care from children is 61% among parents with light long-term care needs, while that share is lower at 53% among parents with severe conditions. The long-term care insurance coverage rate from the entire estimation sample is 14%, but when restricted to individuals aged 60-69 who are healthy enough to purchase insurance, the coverage rate is higher at 21%.

Table 4 presents the summary statistics of children in the estimation sample. Caregiving children are much more likely to be a daughter and live close to their parents. They are less likely to have college education, be married, own a home and work full-time. Only about 3% of caregiving children are paid by their parents for providing informal care.

4.2 Empirical specification

This section describes parameters of the model that are estimated or calibrated outside the model. The model assumes the parent's health transition probabilities follow an exogenously given Markov process where the next period's health is determined by the parent's current health, age, gender and permanent income. I estimate the health transition probabilities by maximum likelihood estimation using a flexible logit. Table 5 reports simulated long-term care risk and life expectancy for healthy 60-year-olds. Long-term care risk decreases in income suggesting that all else equal, selection of higher-income individuals into insurance is advantageous.

To estimate the parent's wealth shock distribution, I compute the residual asset fluctuations in the data using the model's wealth accumulation law specified in Equations (13) and (15). I assume the wealth shock follows a normal distribution with an estimated mean of $9,203 and a standard

Table 5: Simulated Long-Term Care Risk and Life Expectancy for Healthy 60-year-olds

	Years with any LTC needs	Years with severe LTC needs	Life expectancy
Permanent income			
Low	6.69	2.73	76.71
Middle	5.56	1.84	80.20
High	4.23	1.20	81.76
Gender			
Male	4.04	1.31	77.21
Female	6.46	2.33	81.12

Notes: The table reports simulated long-term care risk and life expectancy for healthy 60-year-olds. The simulation sample consists of healthy 60-year-olds from the HRS 1998-2010. The health transition probabilities are estimated as a flexible function of current health, age, gender and permanent income.

deviation that is about seven times larger.[50]

For formal care prices, I use the average rates in 2008 which was \$230 per day for nursing home care and \$21 per hour for paid home care (MetLife, 2008). The model assumes formal care usage intensity depends on the number of children. For parents with less than 4 children, I assume they use paid home care for 21 hours per week if their health is $h_t^P = 1$, and nursing home care for the entire period if their health is $h_t^P = 2$. For parents with 4 or more children, the intensity of formal care usage is reduced by 50% in each health state $h_t^P \in \{1, 2\}$, consistent with what is observed in the data.[51]

In the model, there is one standard long-term care insurance policy that the healthy parent can purchase at age 60. Based on the data collected by Broker World in their survey of major long-term care insurance companies, I assume the standard policy has a per-period benefit cap that is equivalent to 70% of nursing home costs (i.e, $b = 0.70 \times \$230 \times 365 \times 2$) and provides coverage for life.[52] From Brown and Finkelstein (2007), I obtain the average premium which was \$3,195 per year in 2002. In estimating the model, I assume this is the annual premium that all parents uniformly pay if they purchase insurance in the first period.[53]

I set the Medicaid threshold for nursing home residents to zero (Lockwood, 2018). This is

[50]In estimating the standard deviation of the wealth shock, I allow for the possibility that some of the residual asset fluctuations in the data might have stemmed from measurement error. Based on numerous simulations, I find that the estimated model produces the best fit when most of the residual asset fluctuations are attributed to wealth shock. This suggests that measurement error may not be severe in my asset data. This is plausible as the HRS has included an asset verification section, called Section U, since the 2002 wave in which respondents are asked to correct or confirm current or previous reports about wealth components if the difference between them is large. The verification was very successful in bringing down the standard deviation of wave-to-wave change in wealth (Hurd, Meijer, Moldoff, and Rohwedder, 2016).

[51]In my estimation sample, the mean nursing home night is lower by almost 50% for disabled parents with 4 or more children compared to those with 1-3 children.

[52]During my sample period, about 75% of policies offered such lifetime coverage options (Broker World, 2009-2015).

[53]As mentioned earlier, during the sample period of 1998-2010, premiums varied only by age and health. This means that all healthy 60-year-olds paid the same price regardless of their other demographics.

consistent with Medicaid's stringent restrictions on assets for nursing home residents. I set the Medicaid threshold for paid home care users at $9,156 following Brown and Finkelstein (2008). The consumption value of nursing home services is also set to $9,156.

I estimate the coefficients of the child's income function specified in Equation (12) outside the model using all children in my estimation sample. The implicit assumption here is that the opportunity costs of informal care are the same for caregiving and non-caregiving children conditional on the observables included in the income function. Given that the model does not incorporate unobserved types, and I observe the child's income regardless of the child's labor supply decisions, I can estimate the income function outside the model without dealing with a selection issue. Details about the estimation are discussed in Appendix B.2.

The child's total endowed time is set to 112 hours per week. Based on the mean informal care hours conditional on parental health, I assume providing informal care requires 21 hours per week if the parent has light long-term care needs, and 40 hours per week if the parent has severe long-term care needs. Full-time employment requires 35 hours per week, and part-time employment requires 18 hours per week.

I assume a coefficient of relative risk aversion of 3 for the parent's consumption utility function and calibrate the parent's consumption scale parameter (θ_c^P) to 1.65e+9. I also use a coefficient of relative risk aversion of 3 for the child's consumption and leisure utility functions and calibrate the child's consumption scale (θ_c^K) to 5.07e+9.

Following Brown and Finkelstein (2008), I use 3% time preference rate per year ($\beta = \frac{1}{1.06}$) and 3% annual real interest rate ($r = 0.06$). I consider three values of permanent income which correspond to the 20th, 55th and 80th percentiles of parents' income distribution in the sample. I assume the child is 29 years younger than the parent, which is the average age difference between parents and children in the estimation sample.

4.3 Two-step CCP estimation

I estimate the rest of the structural parameters within the model.[54] To reduce the computational cost of estimating a dynamic game with a large state space, I use a two-step CCP estimation methodology pioneered by Hotz and Miller (1993). Specifically, for the estimation of policy functions and value functions, I follow Bajari, Benkard, and Levin (2007) who extend the forward simulation based CCP approach proposed by Hotz, Miller, Sanders, and Smith (1994) to dynamic games and allow for continuous choices. In the first stage, I obtain empirical estimates of the equilibrium decision rules, which involves regressing observed choices on state variables. Using the policy function estimates, I use forward simulation to estimate the agents' value functions. As the agents' preferences are linear in structural parameters that I estimate, averaging over multiple simulated paths is performed only once, which greatly reduces the computational cost. In the second stage, I use the value function estimates to construct a pseudo likelihood function and search for structural parameter values that maximize the likelihood.

[54]See Table 6 for the list of the internally estimated structural parameters.

Policy function estimation. Associated with a strategy profile $\sigma = (\sigma^K, \sigma^{P,d}, \sigma^{P,c})$, let

$$P_\sigma^K(d_t^K|s_t) \quad = \quad \int \mathbb{I}\left\{\sigma^K(s_t, \epsilon_t^K) = d_t^K\right\} g(\epsilon_t^K) d\epsilon_t^K, \tag{23}$$

$$P_\sigma^P(d_t^P|s_t) \quad = \quad \int \mathbb{I}\left\{\sigma^{P,d}(s_t, \epsilon_t^P) = d_t^P\right\} g(\epsilon_t^P) d\epsilon_t^P. \tag{24}$$

As ϵ_t^K and ϵ_t^P have a Type I extreme value distribution, I can express the agents' relative choice-specific value functions as

$$\tilde{v}^K(s_t, d_t^K; \sigma) \quad := \quad v^K(s_t, d_t^K; \sigma) - v^K(s_t, d_t^{K'}; \sigma) = \ln P_\sigma^K(d_t^K|s_t) - \ln P_\sigma^K(d_t^{K'}|s_t), \tag{25}$$

$$\tilde{v}^P(s_t, d_t^P; \sigma) \quad := \quad v^P(s_t, d_t^P; \sigma) - v^P(s_t, d_t^{P'}; \sigma) = \ln P_\sigma^P(d_t^P|s_t) - \ln P_\sigma^P(d_t^{P'}|s_t) \tag{26}$$

where $d_t^{K'}$ and $d_t^{P'}$ are anchor choices for the child and parent, respectively. So once I have estimates of choice probabilities $\hat{P}_\sigma^i$ for $i \in \{K, P\}$, I have estimates of the relative choice-specific value functions, which are sufficient to recover discrete choice decision rules:

$$\hat{\sigma}^K(s_t, \epsilon_t^K) \quad = \quad \underset{d_t^K \in \mathbb{C}^K(h_t^P)}{\operatorname{argmax}} \left\{\ln \hat{P}_\sigma^K(d_t^K|s_t) - \ln \hat{P}_\sigma^K(d_t^{K'}|s_t) + \epsilon_t^K(d_t^K)\right\}, \tag{27}$$

$$\hat{\sigma}^{P,d}(s_t, \epsilon_t^P) \quad = \quad \underset{d_t^P \in \{0,1\}}{\operatorname{argmax}} \left\{\ln \hat{P}_\sigma^P(d_t^P|s_t) - \ln \hat{P}_\sigma^P(d_t^{P'}|s_t) + \epsilon_t^P(d_t^P)\right\}. \tag{28}$$

With unlimited data, policy functions could be estimated non-parametrically. As I have a large state space and limited sample size, I make parametric assumptions on the form of the policy functions, as often done in the application of CCP estimators (Arcidiacono and Ellickson, 2011). I use a logit regression to estimate the child's informal care and employment choice probabilities and the parent's insurance purchase probabilities. Following Bajari, Benkard, and Levin (2007), the parent's consumption policy function is directly estimable from the data on consumption.[55] I observe consumption for about 25% of my estimation sample in the CAMS data. Using this subsample, I employ a linear regression to estimate the consumption policy function. I denote the resulting policy function estimates by $\hat{\sigma} = (\hat{\sigma}^K, \hat{\sigma}^{P,d}, \hat{\sigma}^{P,c})$.

Value function estimation. As in Bajari, Benkard, and Levin (2007), I use the policy function estimates $\hat{\sigma}$ to forward simulate the model and estimate value functions by directly summing up per-period utilities. One useful observation is that for each agent, both the flow utility while the parent is alive and terminal utility when the parent dies are linear in the structural parameters θ that I estimate. As a result, given a strategy profile σ, each agent's value function is also linear in θ and can be represented by $V^i(s_t; \sigma; \theta) = W^i(s_t; \sigma) \cdot \theta$ where W^i does not depend on unknown θ. So once I have an estimate of W^i from the forward simulation procedure based on policy function estimates $\hat{\sigma}$, I can simply scale it by θ to obtain value function estimates at different parameter values, i.e., $\hat{V}^i(s_t; \hat{\sigma}; \theta) = \hat{W}^i(s_t; \hat{\sigma}) \cdot \theta$. Appendix C.1 provides more details about how I take

[55] This requires the assumption that the consumption policy function is a monotonic function of the wealth shock.

advantage of this linearity in the forward simulation procedure, which allows a low-cost estimate of W^i.

Pseudo maximum likelihood (PML) estimation. I use the estimated value functions to construct the pseudo likelihood function as in Aguirregabiria and Mira (2007) and search for the parameters that maximize this function. Appendix C.2 describes the procedure. Standard errors are computed using bootstrapping.

4.4 Identification

I first provide identification arguments for the child's preference parameters. As children with healthy parents only make employment decisions, their choices are helpful in identifying the child's leisure preference parameter separately from the child's informal care utility. The parameters that govern the child's informal care utility are identified from variation in informal care choices by parent health, child demographics and whether informal care was provided in the previous period. These parameters are separately identified from the child's inheritance preference parameter based on informal care choices of children whose parents have almost no assets. As these children do not expect to receive inheritances, their informal care choices are largely driven by the warm-glow informal care utility. Children with insured parents are also not strategically motivated to provide informal care as insurance companies would protect their inheritances from formal care costs in case of no informal care. Their informal care choices therefore also help identify the parameters governing the child's warm-glow informal care utility. For strong identification of the child's inheritance preference parameter, expected inheritances should vary sufficiently by informal care choices. Substantial formal care prices and the assumption that informal care receipt eliminates the need for formal care result in enough variation in expected inheritances by informal care choices.

Insurance choices and savings of uninsured parents help identify how much parents prefer informal care to formal care. The reason is because these choices influence children's incentive to provide informal care by affecting their inheritances exposed to formal care risk. The parent's bequest preference parameter, which represents the parent's altruism, is separately identified from the care preference parameter based on savings of parents who already own insurance. This is because insured parents cannot influence their children's actions using bequests: their children know that their inheritances are protected against formal care expenditures, even if they do not provide care. Therefore, savings of insured parents are informative about parents' altruism toward children.

4.5 Estimation results

Table 6 reports the estimates of the parameters that are structurally estimated by the CCP estimator. Panel A reports the child's preference parameter estimates. The psychological burden of providing care varies substantially by child demographics. Sons find provision of informal care more burdensome than daughters, and children who do not live within 10 miles of their parents experience higher utility costs than children who do. There is a substantial cost in initiating informal

Table 6: Structurally Estimated Parameters

Parameter	Notation	Estimate	Standard error
Panel A: Child's preferences			
Leisure scale	θ_l^K	2.64e+7	2.93e+6
Informal care utility			
$h_t^P = 1$	$\theta_{h_t^P=1}^K$	1.49	0.09
$h_t^P = 2$	$\theta_{h_t^P=2}^K$	1.23	0.11
Male	θ_{male}^K	-0.45	0.06
Live outside 10 mi radius	θ_{far}^K	-0.94	0.08
Initiate caregiving	θ_{start}^K	-1.35	0.11
Inheritance scale	θ_d^K	1.43e+09	4.44e+8
Panel B: Parent's preferences			
Formal care utility	θ_{fc}^P	-12.62	3.91
Bequest utility	θ_d^P	0.60	0.56

Notes: The table reports results from the two-step CCP estimation. Standard errors are computed using 50 bootstrap samples.

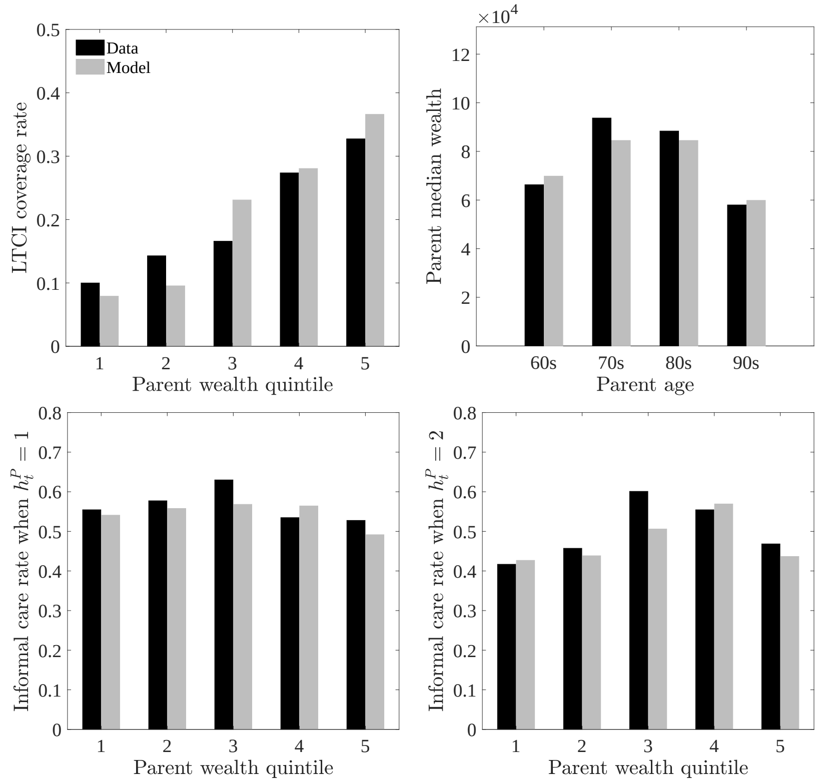
care which may reflect switching or adjustment costs. As a result, the model generates persistence in informal care, consistent with Skira (2015). The child values receiving inheritance, implying that the child is strategically motivated to provide informal care. Panel B reports the parent's preference parameter estimates. The parent's utility from using formal care is negative. As I have normalized the parent's preference for informal care to zero, the negative estimate implies that the parent has a distaste for formal care relative to informal care, consistent with previous studies (Mommaerts, 2016; Barczyk and Kredler, 2018). The altruism factor θ_d^P which determines the strength of the parent's bequest motive has an estimate of 0.6.

I now discuss the fit of the model reported in Figure 3. The model replicates the empirical pattern that the long-term care insurance purchase rate increases monotonically in wealth. It also does a decent job of matching the parent's wealth evolution over the life-cycle. The model also reproduces the inverted-U pattern of informal care provision across parent wealth, although the predicted pattern is slightly shifted to the right compared to the empirical counterpart.

Table 7 reports how well the model matches the average characteristics of non-caregiving and caregiving children. The model reproduces the empirical pattern that caregiving children are much more likely to be a daughter and live closer to their parents. It also replicates the pattern that they are less likely to have college education, be married, own a home and work full-time.

I examine whether the model replicates the low correlation between insurance ownership and nursing home risk found in previous studies. Finkelstein and McGarry (2006) show that conditional on pricing controls, there is a negative, albeit statistically insignificant, correlation between insurance ownership and nursing home entry. Using a sample of healthy 60-year-olds, I compare the individuals' simulated lifetime nursing home risk by their simulated insurance choices in the first period. As reported in Table 8, the model is able to reproduce the negative correlation: about

Figure 3: Model Fit (1)



Notes: Gray bars represent model simulated moments, and black bars represent their empirical counterparts.

Table 7: Model Fit (2)

	Children not providing care to sick parents	Children providing care to sick parents
Female	0.46	0.70
	[0.43]	[0.67]
Live within 10 mi of the parent	0.43	0.84
	[0.31]	[0.78]
Have some college education	0.44	0.41
	[0.48]	[0.41]
Married	0.66	0.55
	[0.66]	[0.63]
Homeowner	0.67	0.51
	[0.66]	[0.60]
Work full-time	0.55	0.52
	[0.58]	[0.53]
Work part-time	0.07	0.10
	[0.26]	[0.29]

Notes: Gray numbers reported in brackets represent model simulated moments. Black numbers represent empirical moments. The sample is restricted to children whose parents have long-term care needs.

Table 8: Simulated Correlation between LTCI Ownership and Nursing Home Utilization

	LTCI owners	Non-owners
Share ever using nursing home care before death	0.288	0.302

Notes: The table reports simulated nursing home risk of healthy 60-year-olds over the remainder of their lives, conditional on their simulated long-term care insurance choices in the first period.

28.8% of insurance owners enter a nursing home at some point before their death, while 30.2% of non-owners do. The difference is comparable to Braun, Kopecky, and Koreshkova (2019) who use a structural model of long-term care insurance and find that about 36.9% of insurance owners in the very old stage of life enter a nursing home while 40.7% of non-owners enter. As discussed earlier in Section 3.3, private information about children's informal care provision increases the correlation between insurance ownership and formal care risk, while income-based selection reduces the correlation. The results show that the latter has a slightly larger effect leading to a negative correlation in aggregate.

5 Counterfactuals

To quantify the effects of family interactions on the insurance market equilibrium, I embed the estimated intergenerational game in an equilibrium long-term care insurance market. To incorporate the supply side of the market, I assume there are competitive risk-neutral insurance companies.[56] They sell the standard long-term care insurance policy with features described in Section 3.1 to healthy 60-year-olds and compete by setting premiums. The equilibrium premium is such that insurers break even subject to a loading factor. Specifically, it satisfies

$$p^* = \min\{p : AC(p) = (1 - load) \times AR(p)\} \tag{29}$$

where $AC(p)$ is firms' average present-discounted lifetime claims from consumers who buy insurance when the premium is p. $AR(p)$ is firms' average present-discounted lifetime premium payments from their policyholders. The term *load* is a loading factor which captures administrative costs, and based on previous studies, I use an 18% load.[57]

[56] According to a 2016 report prepared by the Center for Insurance Policy and Research for National Association of Insurance Commissioners, since the late 1990s, there have been about a dozen insurance companies accounting for more than 80% of the sales in the long-term care insurance market (the report can be found at https://www.naic.org/documents/cipr_current_study_160519_ltc_insurance.pdf). Therefore, one can think of the break-even premium I compute as the lower bound on the equilibrium premium which may or may not be higher due to market concentration. Note that the market could still have competitive prices in equilibrium even with a few insurers if they are playing a Bertrand-like pricing competition. As long-term care insurance contracts are essentially financial contracts that specify reimbursement amounts for formal care utilization episodes, the degree of product differentiation is relatively low.

[57] Braun, Kopecky, and Koreshkova (2019) report that administrative expenses associated with underwriting and claims processing were 20% of present-value premium in 2000 and 16% of present-value premium in 2014.

I apply the following algorithm to compute the insurance market equilibrium: (1) for a given price of long-term care insurance, I solve the intergenerational game backward using the structural parameter estimates, (2) I use optimal decision rules of the family to forward simulate the model, (3) using simulated insurance choices and formal care utilization, I compute insurance companies' average revenue and cost, and (4) I repeat the steps (1)-(3) until I find the premium that satisfies the break-even condition in Equation (29).[58]

I build the simulation sample by selecting healthy 60-year-olds from the HRS 2000-2002.[59] I do not restrict the sample to single individuals because during the sample period, all healthy 60-year-olds paid the same price regardless of their marital status.[60] For each parent, I select one child using the same strategy employed in the construction of the estimation sample. I make 100 duplicates for each parent-child pair to increase the sample size.

The break-even premium is computed as $4,707 per year, and the resulting coverage rate is 17.8%. Henceforth, I will refer to this equilibrium as the benchmark equilibrium. Note that in estimating the model, I used the average premium of the standard policy over the sample period, which was $3,195 in 2002. This is substantially lower than the model implied break-even premium. Indeed, in the last decade, long-term care insurance companies reported huge losses due to underpriced policies from older blocks of sales, and almost all insurance companies sought approvals from the state governments to increase premiums on existing policies.[61] For example, Genworth, the biggest long-term care insurer, reports that their accumulated losses on policies sold in the 1990s were $3.6 billion through 2005 alone.[62] In 2015, they requested rate increases of 80-85% on policies sold before 2011 in most states.[63] Some of the most frequently given explanations for such underpricing include lower-than-expected lapse rates. Insurers anticipated policyholders would abandon their policies at a rate of about 5% per year, but according to Genworth, the realized lapse rate is only 0.7%.[64] One other explanation is insurers' lack of experience in predicting long-term costs, which is relatively a new class of risks.[65] As my model assumes zero lapse rates and predicts insurers' costs taking into account possible adverse selection due to the probability of receiving informal care, I obtain a break-even premium that is higher than the empirical premium which seems to have been underpriced.

[58] Appendix B.3 describes the numerical method used to solve the model.

[59] As described in Section 4.2, I use the average premium in 2002 in estimating the intergenerational game. To compare the model predicted equilibrium premium to this, I use potential long-term care insurance buyers from 2000-2002.

[60] As the model is estimated using single parents, the estimated model may overpredict informal care from children for married individuals. However, this issue is mitigated by the fact that (1) long-term care needs are late-life risks, and (2) the share of singles increases sharply with age.

[61] https://www.latimes.com/business/story/2019-10-01/long-term-care-insurance.

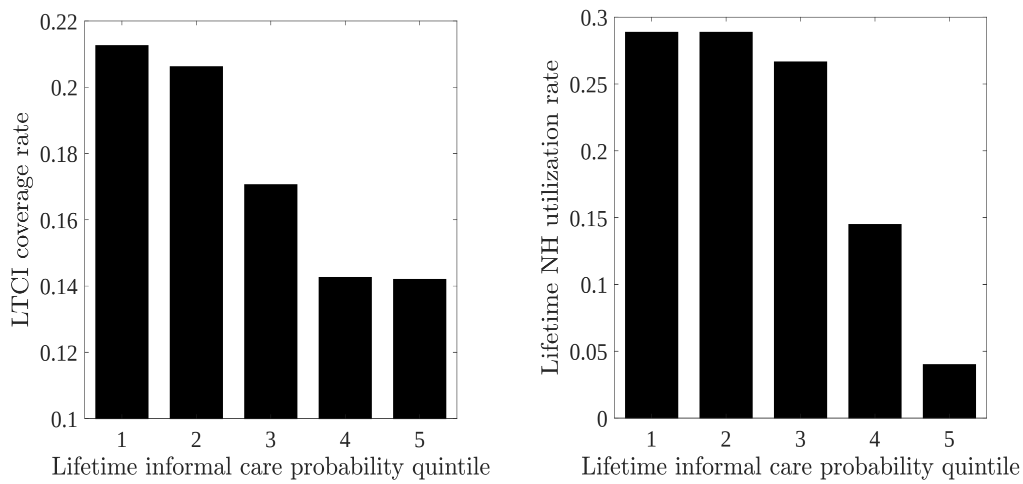
[62] https://www.latimes.com/business/story/2019-07-24/long-term-care-insurance-disaster.

[63] https://www.nytimes.com/2015/09/03/your-money/managing-the-costs-of-long-term-care-insurance.html.

[64] See footnote 62.

[65] See footnote 62.

Figure 4: Simulated Adverse Selection Based on Children's Informal Care Likelihood



Panel A: Insurance purchase phase Panel B: Post-purchase utilization phase

Notes: The simulation sample consists of healthy 60-year-olds. In both panels, the horizontal axis represents individuals' simulated probability of receiving informal care in periods where they have long-term care needs over the life-cycle (see the text for details). In Panel A, the vertical axis represents the healthy 60-year-olds' simulated insurance purchase rate. For Panel B, I only use individuals in the simulation sample who purchase insurance in the first period. In Panel B, the vertical axis represents these insurance owners' simulated nursing home (NH) use frequency in periods where they have long-term care needs over the life-cycle.

5.1 Adverse selection based on children's informal care likelihood

The next section presents the paper's main counterfactual which investigates the welfare effect of reducing private information about children's informal care likelihood in the long-term care insurance market. The validity of this exercise depends on how accurately the model captures adverse selection generated by private information about children's informal caregiving. This section shows that the model is able to replicate the magnitude of this adverse selection channel found in the descriptive analysis in Section 2.

I start by constructing a model-based measure that could capture how likely a parent is to receive informal care over the life-cycle. To this end, I simulate the model assuming no parent can purchase insurance.[66] For each family in the simulation sample, I compute how frequently informal care is provided in simulation periods where the parent has long-term care needs. For example, if a parent in the simulation sample is hit by an adverse health shock ($h_t^P \in \{1, 2\}$) in four periods over the life-cycle and receives informal care in one, then the frequency is computed as $\frac{1}{4} = 25\%$. I treat this as the model-implied measure for the *lifetime* probability of receiving informal care. This measure is used as the horizontal axis in Panels A and B of Figure 4.

Next, I simulate the benchmark equilibrium and examine how a parent's simulated insurance

[66]The reason why I remove long-term care insurance is because I want to quantify parents' beliefs about receiving informal care when they do not yet own long-term care insurance.

choice is correlated with his or her lifetime probability of receiving informal care, computed above. Panel A in Figure 4 shows the correlation. The negative slope confirms that parents who expect a low probability of receiving informal care over the life-cycle are more likely to select into insurance in the model. Reduced-form evidence, reported earlier in Table 1, shows that the insurance coverage rate for individuals who believe their children will provide informal care is lower by 4.2 percentage points than individuals who do not believe their children will provide help. As the insurance coverage rate of the sample in Table 1 is 15.6%, one can think of the ratio $\frac{4.2}{15.6} = 27\%$ as the empirical importance of the informal care likelihood on insurance ownership. To compare the model-based results to this, I classify parents whose simulated lifetime informal care probability is in the bottom 50% as "not believe children will help" and parents whose lifetime informal care probability is in the top 50% as "believe children will help". The simulated insurance take-up rate for the latter group is lower by 5.7 percentage points. As the equilibrium coverage rate is 17.8%, the model counterpart to the empirical ratio $\frac{4.2}{15.6} = 27\%$ is computed as $\frac{5.7}{17.8} = 32\%$, which is very similar. This implies that the model is able to replicate the relationship between the likelihood of receiving informal care and insurance ownership observed in the data.

Finally, I examine if the model is able to replicate the negative correlation between initial beliefs about receiving informal care and nursing home use among disabled individuals who have already purchased insurance, as documented in Table 2. To this end, I restrict to individuals in the simulation sample who purchase insurance in the first period. For each of these simulated insurance owners, I compute how frequently nursing home care is used in simulation periods where the individual experiences long-term care needs. For example, if the insurance owner is hit by an adverse health shock in four periods over the life-cycle and uses nursing home care in one period, then the frequency is computed as $\frac{1}{4} = 25\%$. Panel B in Figure 4 shows how this simulated nursing home utilization rate over the life-cycle is correlated with the insurance owner's lifetime probability of receiving informal care. Consistent with reduced-form evidence reported in Table 2, there is a strong negative correlation. Classifying again those in the bottom 50% of the lifetime informal care probability distribution as "not believe children will help" and those in the top 50% as "believe children will help", the simulated nursing home utilization rate over the life-cycle is lower by 16 percentage points for the latter group (the mean is 20%). This is broadly consistent with reduced-form evidence in Table 2 which shows that the probability of having a lengthy nursing home stay over a two-year period is lower by about 9 percentage points for disabled insurance owners who initially believed their children would provide informal care (the mean is 17.5%).[67]

To sum, the simulation results reported in Figure 4 confirm that the model is able to reproduce the quantitative importance of private information about children's informal care likelihood on parents' insurance decisions as well as their formal care utilization decisions after having purchased insurance.

[67]Note that in the HRS 1998-2010, the number of insurance owners for whom I observe their lifetime nursing home risk occurrences is very small because (1) the length of the sample period is only 12 years, and (2) few individuals own insurance. This is why I instead rely on nursing home use in the past two years to measure disabled insurance owners' risk occurrences.

5.2 Countefactual risk adjustment

To reduce adverse selection generated by private information about the probability of receiving informal care, I now consider counterfactual risk adjustment whereby an individual's long-term care insurance premiums are adjusted based on observables that are predictive of expected informal care provision from children. Note that there is no direct regulation on individual covariates that could be used to set insurance prices in this market,[68] and gender-based pricing was newly introduced in 2013 (Finkelstein and Poterba, 2014).

As discussed in Section 2.2, reduced-form evidence suggests that a child's gender and residential proximity to the parent are powerful predictors of whether the parent says the child will provide informal care. Consistent with this fact, the estimated model predicts that daughters and children living close to their parents are much more likely to provide informal care, as shown in Table 7. Furthermore, by assuming that the required formal care usage intensity is lower for parents with four or more children, the model replicates the empirical pattern that these parents incur substantially less formal care expenses compared to parents with three or less children.

I therefore consider counterfactual risk adjustment where premiums are adjusted based on (1) whether a parent has a daughter, (2) whether the parent has a child living in a 10-mile radius of the parent,[69] and (3) whether the parent has four or more children.[70] Under this risk adjustment, there will be $2^3 = 8$ market segments. I divide the simulation sample into 8 groups accordingly and compute the insurance market equilibrium for each of the 8 market segments.

I measure parents' welfare as the initial wealth transfer needed to make a parent under default pricing, where all healthy 60-year-olds face the same price, indifferent to counterfactual pricing. For the computation of children's welfare, I cannot use the same approach as the model does not incorporate children's savings. Instead, I calculate by how much their parents' wealth, which represents the children's inheritances, needs to increase under default pricing to make the children equally well off as they would be under counterfactual pricing. This measure is able to capture the children's welfare because their expected value increases in their parents' wealth due to bequests. It also has an easy interpretation when one compares this counterfactual pricing to alternative policies where the government makes long-term care related transfers to the elderly population.

[68] While the regulation on long-term care insurance rates varies from state to state, most states' regulations are based on the Long-Term Care Insurance Model Regulation established by National Association of Insurance Commissioners (NAIC) which regulate both initial rates and rate increases based on minimum expected loss ratio (NAIC Long-Term Care Insurance Model Regulation, www.naic.org/store/free/MDL-641.pdf).

[69] While children's residential proximity to parents is a key predictor of their informal care provision, if insurance prices depended on this characteristic, then it might be subject to a strategic change. For example, a potential buyer might live with her child only until she purchases long-term care insurance. I suspect such strategic responses will be minor as there are already several insurance markets that use place of residence in pricing contracts (e.g., the U.K. annuity market and U.S. automobile insurance market). Furthermore, insurance companies could add a contractual provision that policyholders will be subject to a premium revision in case of a change in priced characteristics.

[70] As long as premiums of insurance contracts do not alter the way the family interact over long-term care, the statistical difference in formal care risk by the number of children can be treated as policy invariant. This is plausible because what really matters for children's informal care decisions is whether their parents have long-term care insurance rather than how much their parents have paid for it.

Table 9: Equilibrium Effects of Counterfactual Risk Adjustment

Priced observables	LTCI take-up rate	Average annual premium	Average cost	Average parent welfare	Average child welfare
Benchmark	0.178	$4,707	$51,532	$0	$0
Gender	0.177	$4,723	$51,007	-$18	$147
Child demographics	0.210	$3,928	$43,874	$4,270	$1,956

Notes: The first row reports the benchmark insurance market equilibrium where all healthy 60-year-olds pay the same price. The second row (Gender) reports the market equilibrium where prices are conditional on the gender of a consumer. The third row (Child demographics) reports the market equilibrium where prices are conditional on whether the consumer has four or more children, a daughter and a child living in a 10-mile radius. Except for the first row where there is a single market segment, "Average annual premium" represents the average of break-even annual premiums of multiple market segments. "Average cost" represents insurers' mean present-discounted lifetime claims from all individuals who purchase insurance.

Table 9 summarizes the results. To make a comparison to recently introduced gender-based pricing, the table also reports the market equilibrium when prices vary by parent gender. As shown in the third row, adjusting premiums based on child demographics increases the equilibrium coverage rate from 17.8% to 21%. The average welfare gain is almost $4,300 for parents and $2,000 for children. As parents with a high probability of receiving informal care select into insurance, insurance companies' average cost goes down. The average premium across the 8 market segments is $3,928, which is substantially lower than the benchmark equilibrium premium of $4,707. The results show that pricing based on family observables that are highly predictive of the informal care likelihood reduces adverse selection and generates welfare gains.

Note that gender-based pricing has almost no effect on the equilibrium coverage rate. When prices vary by consumers' gender, the equilibrium premium goes down for men while it goes up for women. This is because women have a higher probability of developing long-term care needs as shown earlier in Table 5. However, women do not have substantially higher willingness to pay for insurance because they are more likely to spend down to Medicaid eligibility due to larger formal care spending risk. This is consistent with Brown and Finkelstein (2008) who argue that the extent to which private insurance is redundant of benefits that Medicaid would otherwise have paid (which they call Medicaid's "implicit tax") is higher for women as they face greater long-term care risk. As a result, when women are faced with a higher premium under gender-based pricing, many of them drop coverage, and this effect cancels out all of the increase in insurance take-up from men.

It should be noted that using individual characteristics in setting insurance prices may be costly to insurance companies in a way that is not captured in my analysis. For example, insurance companies may be concerned that using child demographics in pricing could lead to regulatory response or consumer backlash.[71] Unpacking insurance companies' decision of which observables to price on is an interesting direction for future work.

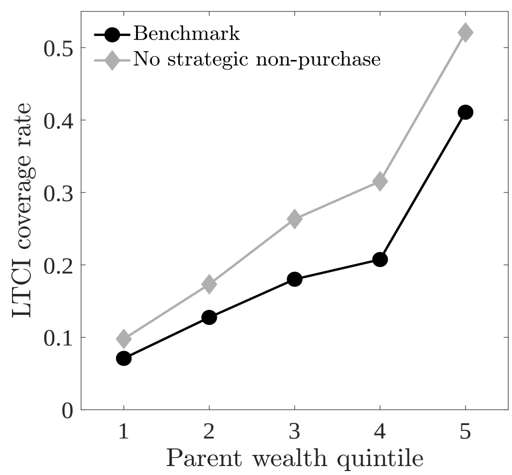
[71] See Finkelstein and Poterba (2014) for a discussion about potential explanations for "unused observables" in insurance markets.

Table 10: Strategic Non-Purchase of Insurance and the Insurance Market Equilibrium

	LTCI take-up rate	Annual premium	Average cost
Benchmark equilibrium	0.178	$4,707	$51,532
Equilibrium without strategic non-purchase	0.252	$4,512	$49,516

Notes: The first row reports the benchmark insurance market equilibrium. The second row reports the counterfactual equilibrium where long-term care insurance does not crowd out children's informal care provision, and hence there is no strategic non-purchase of insurance. Specifically, children whose parents own insurance are forced to make the same informal care choices as they would when their parents did not own insurance. "Average cost" represents insurers' mean present-discounted lifetime claims from all individuals who purchase insurance.

Figure 5: Strategic Non-Purchase of Insurance by Parent Wealth



Notes: The black line represents the insurance coverage rate in the benchmark equilibrium. The gray line represents the insurance coverage rate in the counterfactual equilibrium where long-term care insurance does not crowd out children's informal care provision, and hence there is no strategic non-purchase of insurance.

5.3 Strategic non-purchase of insurance

Using the non-cooperative feature of the model, I provide the first estimate on the strategic non-purchase of insurance where parents forgo insurance because they are concerned about its crowd-out effect on children's informal care provision. To do this, I simulate the model assuming the purchase of long-term care insurance does not reduce children's informal care provision. Specifically, I force the child whose parent purchases long-term care insurance to make the same informal care choices as the child would when the parent did not purchase insurance.[72]

Table 10 summarizes the results. The first row reports the benchmark insurance market equilib-

[72] Adding the assumption that "the parent hides the purchase of insurance from the child" to the intergenerational game does not fully eliminate the crowd-out effect of insurance on the child's informal care provision. This is because by the Bayes' rule, the child will correctly infer the parent's insurance purchase probability in equilibrium.

rium where the child is allowed to show behavioral responses to the parent's purchase of insurance. As said earlier, the equilibrium coverage is 17.8%, and annual premium is $4,707. The second row reports the counterfactual equilibrium where long-term care insurance does not crowd out children's informal care provision, and hence there is no strategic non-purchase of insurance. The equilibrium coverage rate increases to 25.2%, corresponding to an almost 42% increase. Figure 5 presents the coverage increase by parental wealth at age 60. It shows that most of the increase comes from relatively higher-wealth individuals. This is expected as the strategic non-purchase of insurance is the most relevant for parents who have enough wealth to incentivize their children using bequests. The average cost to insurance companies decreases when insurance does not reduce children's informal care provision as insured parents are more likely to receive informal care. The equilibrium premium is adjusted to reflect the reduction in the average cost: it drops from $4,707 to $4,512.

To sum, there is quantitatively meaningful strategic non-purchase of insurance, reducing the equilibrium long-term care insurance coverage rate from 25.2% to 17.8%. This is the first estimate on the effect of strategic bequests on elderly parents' insurance choices and provides empirical evidence for relevant theoretical studies such as Bernheim, Shleifer, and Summers (1985), Pauly (1990), Zweifel and Struwe (1996) and Courbage and Zweifel (2011).

6 Conclusion

Using a dynamic intergenerational game between an elderly parent and an adult child, this paper studies how family interactions over long-term care affect the long-term care insurance market equilibrium. I find that private information about children's informal care likelihood results in adverse selection where the market attracts a disproportionate number of individuals who face higher formal care utilization risk due to a lower probability of receiving informal care from children. I show that pricing based on family observables that are highly predictive of the informal care likelihood reduces adverse selection and generates welfare gains. Using the non-cooperative feature of the model, I also show that there is quantitatively meaningful strategic non-purchase of insurance where parents forgo insurance to avoid diminishing children's informal care incentive.

Challenges in the long-term care sector, such as the aging of the baby boom generation, increasing burdens of informal caregivers and growing Medicaid spending on formal care, have triggered various policy recommendations. They include the government providing family care subsidies and insurance companies paying cash to informal caregivers. Such recommendations are non-market-based which could lead to even bigger efficiency costs, or involve drastic changes in the structure of the insurance products and raise doubts about the practicality. In contrast, my proposal of using family demographics in pricing is market-based and is already in momentum: the fact that insurance companies have started to price on consumer gender makes my proposal well-grounded.

References

Aguirregabiria, V., and P. Mira, 2002, Swapping the Nested Fixed Point Algorithm: A Class of Estimators for Discrete Markov Decision Models, *Econometrica* 70, 1519–1543.

Aguirregabiria, V., and P. Mira, 2007, Sequential Estimation of Dynamic Discrete Games, *Econometrica* 75, 1–53.

Arcidiacono, P., and P.B. Ellickson, 2011, Practical Methods for Estimation of Dynamic Discrete Choice Models, *Annual Review of Economics* 3, 363–394.

Bajari, P., C. L. Benkard, and J. Levin, 2007, Estimating Dynamic Models of Imperfect Competition, *Econometrica* 75, 1331–1370.

Barczyk, D., and M. Kredler, 2018, Evaluating Long-Term-Care Policy Options, Taking the Family Seriously, *Review of Economic Studies* 85, 766–809.

Bernheim, B.D., A. Shleifer, and L.H. Summers, 1985, The Strategic Bequest Motive, *Journal of Political Economy* 93, 1045–1076.

Braun, R. A., K. A. Kopecky, and T. Koreshkova, 2019, Old, Frail, and Uninsured: Accounting for Features of the U.S. Long-Term Care Insurance Market, *Econometrica* 87, 981–1019.

Broker World, 2009-2015, Long Term Care Survey, Broker World.

Brown, J. R., and A. Finkelstein, 2007, Why Is the Market for Long-Term Care Insurance So Small?, *Journal of Public Economics* 91(10), 1967–1991.

Brown, J. R., and A. Finkelstein, 2008, The Interaction of Public and Private Insurance: Medicaid and the Long-Term Care Insurance Market, *American Economic Review* 98(3), 1083–1102.

Brown, M., 2006, Informal Care and the Division of End-of-Life Transfers, *Journal of Human Resources* 41, 191–219.

Byrne, D., M.S. Goeree, B. Hiedemann, and S. Stern, 2009, Formal Home Health Care, Informal Care, and Family Decision Making, *International Economic Review* 50, 1205–1242.

Castilla, C., and T. Walker, 2013, Is Ignorance Bliss? The Effect of Asymmetric Information between Spouses on Intra-Household Allocations, *American Economic Review: Papers & Proceedings* 103, 263–268.

Charles, K.K., and P. Sevak, 2005, Can Family Caregiving Substitute for Nursing Home Care?, *Journal of Health Economics* 24, 1174–1190.

Coe, N. B., G. S. Goda, and C. H. Van Houtven, 2015, Family Spillovers of Long-Term Care Insurance, NBER Working Paper No. 21483.

Courbage, C., and P. Zweifel, 2011, Two-Sided Intergenerational Moral Hazard, Long-Term Care Insurance, and Nursing Home Use, *Journal of Risk and Uncertainty* 43, 65–80.

De Nardi, M, 2004, Wealth Inequality and Intergenerational Links, *Review of Economic Studies* 71, 743–768.

De Nardi, M., E. French, and J.B. Jones, 2010, Why Do the Elderly Save? The Role of Medical Expenses, *Journal of Political Economy* 118, 39–75.

Einav, L., A. Finkelstein, and J. Levin, 2010, Beyond Testing: Empirical Models of Insurance Markets, *Annual Review of Economics* 2, 311–336.

Fahle, S., 2014, Harnessing the Potential of Family Caregivers: A Solution to the U.S. Long-Term Care Crisis?, Working paper.

Finkelstein, A., and K. McGarry, 2006, Multiple Dimensions of Private Information: Evidence from the Long-Term Care Insurance Market, *American Economic Review* 96, 938–958.

Finkelstein, A., and J. Poterba, 2014, Testing for Asymmetric Information Using "Unused Observables" in Insurance Markets: Evidence from the U.K. Annuity Market, *Journal of Risk and Insurance* 81, 709–734.

Groneck, M., 2017, Bequests and Informal Long-Term Care: Evidence from the HRS Exit-Interviews, *Journal of Human Resources* 52, 531–572.

Hendren, N., 2013, Private Information and Insurance Rejections, *Econometrica* 81, 1713–1762.

Hotz, V.J., and R.A. Miller, 1993, Conditional Choice Probabilities and the Estimation of Dynamic Models, *Review of Economic Studies* 60, 497–529.

Hotz, V.J., R.A. Miller, S. Sanders, and J. Smith, 1994, A Simulation Estimator for Dynamic Models of Discrete Choice, *Review of Economic Studies* 61, 265–289.

Houtven, C. H. Van, and E. C. Norton, 2004, Informal Care and Health Care Use of Older Adults, *Journal of Health Economics* 23, 1159–1180.

Houtven, C. H. Van, and E. C. Norton, 2008, Informal Care and Medicare Expenditures: Testing for Heterogeneous Treatment Effects, *Journal of Health Economics* 27, 134–156.

Hubbard, R. G, J. Skinner, and S.P. Zeldes, 1995, Precautionary Saving and Social Insurance, *Journal of Political Economy* 103, 360–399.

Hurd, M.D., E. Meijer, M.B. Moldoff, and S. Rohwedder, 2016, Improved Wealth Measures in the Health and Retirement Study, *RAND Working Paper*.

Kaplan, G., 2012, Moving Back Home: Insurance against Labor Market Risk, *Journal of Political Economy* 120, 446–512.

Kemper, P., H. Komisar, and L. Alecxih, 2005/2006, Long-Term Care Over an Uncertain Future: What Can Current Retirees Expect?, *Inquiry* 42, 335–350.

Lockwood, L., 2018, Incidental Bequests and the Choice to Self-Insure Late-Life Risks, *American Economic Review* 108, 2513–2550.

Lundberg, S., and R. A. Pollak, 1993, Separate Spheres Bargaining and the Marriage Market, *Journal of Political Economy* 101, 988–1010.

McGarry, K., and R.F. Schoeni, 1997, Transfer Behavior within the Family: Results from the Asset and Health Dynamics Study, *Journals of Gerontology Series B* 52B, 82–92.

MetLife, 2008, National Survey of Nursing Home and Assisted Living Costs, MetLife Mature Market Institute.

Mommaerts, C., 2016, Long-Term Care Insurance and the Family, Working paper.

Pakes, A., M. Ostrovsky, and S. Berry, 2007, Simple Estimators for the Parameters of Discrete Dynamic Games (with Entry/Exit Examples), *Rand Journal of Economics* 38, 373–399.

Palumbo, M.G., 1999, Uncertain Medical Expenses and Precautionary Saving Near the End of the Life Cycle, *Review of Economic Studies* 66, 395–421.

Pauly, M.V., 1990, The Rational Nonpurchase of Long-Term Care Insurance, *Journal of Political Economy* 98, 153–168.

Pesendorfer, M., and P. Schmidt-Dengler, 2008, Asymptotic Least Squares Estimators for Dynamic Games, *Review of Economic Studies* 75, 901–928.

Skira, M., 2015, Dynamic Wage and Employment Effects of Elder Parent Care, *International Economic Review* 56, 63–93.

Society of Actuaries, 2014, An Overview of the U.S. LTC Insurance Market (Past and Present): The Economic Need for LTC Insurance, the History of LTC Regulation & Taxation and the Development of LTC Product Design Features, .

Zweifel, P., and W. Struwe, 1996, Long-Term Care Insurance and Bequests as Instruments for Shaping Intergenerational Relationships, *Journal of Risk and Uncertainty* 12, 65–76.

Appendix

A More Descriptive Evidence

A.1 Subjective nursing home entry probability and beliefs about receiving informal care

The model makes a prediction that the the parent expects informal care provision by the child to lower her future formal care risk when she makes a long-term care insurance purchase decision. To provide descriptive evidence for this, I estimate the correlation between an individual's self-assessed probability of entering a nursing home in the next five years, B_{it}^{NH}, and whether the individual thinks his/her children will provide informal care, B_{it}^{IC}, conditional on buyer characteristics used by long-term care insurers in pricing, X_{it}. The sample consists of individuals who are healthy enough to purchase insurance and old enough to have ADL limitations in the next five years (the same sample used to do the asymmetric information test reported in Table 1). Table A.1 presents the regression results. The estimated correlation between B_{it}^{NH} and B_{it}^{IC} conditional on X_{it} is indeed negative and statistically significant: individuals who think their children will provide informal care in the future have a self-assessed probability of entering a nursing home in the next five years that is lower by 2.6 percentage points.

Table A.1: Relationship between Subjective Nursing Home Entry Probability and Beliefs about Receiving Informal Care

Dependent variable:	B^{NH}
B^{IC}	-0.026***
	(0.005)
Pricing controls, X	Yes
Mean of dependent variable	0.111
Observations	5,739

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors are clustered at the household level and reported in parentheses. The sample consists of individuals aged 70-75 who are currently healthy enough to purchase long-term care insurance based on underwriting guidelines in Hendren (2013). The dependent variable B^{NH} is an individual's self-assessed probability of entering a nursing home over the following five-year period. The key control B^{IC} is an indicator for whether the individual thinks his or her children will provide informal care in the future. Other controls include buyer characteristics used by insurers in pricing (X).

A.2 Predictors of beliefs about receiving informal care

To justify the model's assumption that "unpriced" child characteristics may be the key source of private information about the probability of receiving informal care, I regress whether a healthy individual believes a particular child will provide informal care on the child's characteristics, parental assets and individual characteristics used by long-term care insurers in pricing. Results presented in Table A.2 show that whether the child is a daughter and lives within a 10-mile radius to the parent have by far the largest economic significance. This is why in the model, these characteristics directly enter the child's informal care utility function, $\omega^K(\cdot)$.

Table A.2: Predictors of Beliefs about Receiving Informal Care

Dependent variable:	Parent believes his/her child will provide informal care	
Child is female	0.126***	(0.005)
Child lives within a 10-mi. radius	0.225***	(0.006)
Child has college education	0.005	(0.006)
Child's age	-0.001**	(0.000)
Child is married	0.042***	(0.006)
Child owns a home	0.040***	(0.006)
Parent has four or more children	0.011	(0.008)
Parent is in the bottom wealth quintile	-omitted-	
2nd wealth quintile	0.025*	(0.014)
3rd wealth quintile	0.009	(0.014)
4th wealth quintile	-0.002	(0.014)
top wealth quintile	-0.026*	(0.015)
Parent is in the bottom income group	-omitted-	
middle income group	-0.008	(0.009)
top income group	0.007	(0.011)
Pricing controls, X	Yes	
Mean of dependent variable	0.346	
Observations	66,144	

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors are clustered at the household level and reported in parentheses. The sample consists of parent-child pairs in which the parent is aged 60-75 and healthy enough to buy long-term care insurance. Estimates are from a linear probability model. The regression also includes buyer characteristics used by insurers in pricing (X).

B Model Details

B.1 Child's terminal value

To derive the child's terminal value, I assume that when the parent dies, the child does not work, and optimally chooses to consume her inheritance w_t^P over the next T_0 periods. Given that the child is risk-averse, she will allocate the inheritance equally over the next T_0 periods. Let x denote the equally allocated amount. Using $\beta = \frac{1}{1+r}$, I obtain $x = \frac{1-\beta}{1-\beta^{T_0}} w_t^P$. The child's terminal value is computed as the discounted sum of the consumption utilities over the next T_0 periods:

$$\pi_d^K(w_t^P) = \theta_d^K \frac{1-\beta^{T_0}}{1-\beta} \frac{x^{1-\rho_c^K} - 1}{1-\rho_c^K} \tag{B.1}$$

where θ_d^K is the inheritance scale parameter. For the empirical specification of the model, I set $T_0 = 6$.

B.2 Child's income function

This section describes how I obtain estimates of γ's in Equation (12). The HRS reports children's annual family income as bracketed values: below \$10K, \$10K-35K, \$35K-70K, above \$35K, and above \$70K. I put children in the "above \$35K" bracket into the "\$35K-70K" bracket. As each period is two years in my model, I double the threshold values and define $\hat{y}_i^K$ as

$$\hat{y}_i^K = \begin{cases} 1 & \text{if below \$20K,} \\ 2 & \text{if between \$20K-70K,} \\ 3 & \text{if between \$70K-140K,} \\ 4 & \text{and if above \$140K.} \end{cases} \tag{B.2}$$

I assume there is an underlying continuous family income, $\tilde{y}_i^K$, which is defined as

$$\log(\tilde{y}_i^K) = x_i^K \gamma + \eta_i \tag{B.3}$$

where

$$\begin{aligned} x_i^K \gamma = \quad & \underbrace{\gamma_1 + \gamma_2 age_t^K + \gamma_3 (age_t^K)^2 + \gamma_4 home^K + \gamma_5 mar^K}_{\text{non-labor income}} + \underbrace{\gamma_6 \mathbb{I}[e_t^K = 1]}_{\text{part-time labor income}} \\ & \underbrace{+ \mathbb{I}[e_t^K = 2] * \left\{ \gamma_7 + \gamma_8 age_t^K + \gamma_9 (age_t^K)^2 + \gamma_{10} edu^K + \gamma_{11} \mathbb{I}[e_{t-1}^K = 2] \right\}}_{\text{full-time labor income}}. \end{aligned} \tag{B.4}$$

I assume η_i follows an *i.i.d.* normal distribution with mean zero and variance σ_η^2. The log likelihood function is

$$\log L(\gamma, \sigma_\eta | \hat{y}^K, x^K) = \sum_i \log P(\hat{y}_i^K | x_i^K; \gamma, \sigma_\eta) \tag{B.5}$$

where

$$\begin{aligned} P(\hat{y}_i^K = 1 | x_i^K) &= \Phi_{\sigma_\eta}(\log(20\text{K}) - x_i^K \gamma | x_i^K), \\ P(\hat{y}_i^K = 2 | x_i^K) &= \Phi_{\sigma_\eta}(\log(70\text{K}) - x_i^K \gamma) - \Phi_{\sigma_\eta}(\log(20\text{K}) - x_i^K \gamma), \\ P(\hat{y}_i^K = 3 | x_i^K) &= \Phi_{\sigma_\eta}(\log(140\text{K}) - x_i^K \gamma) - \Phi_{\sigma_\eta}(\log(70\text{K}) - x_i^K \gamma), \quad \text{and} \\ P(\hat{y}_i^K = 4 | x_i^K) &= 1 - \Phi_{\sigma_\eta}(\log(140\text{K}) - x_i^K \gamma | x_i^K). \end{aligned}$$

Φ_{σ_η} is the CDF of η_i. The estimation sample consists of children aged between 21 and 59 in the HRS 1998-2010. The estimates of γ are reported in Table B.1.

Table B.1: Child Family Income Estimates

	Estimate
Constant	7.601
Age	0.110
Age^2	-0.001
Home	0.452
Married	0.534
Full-time	1.691
Full-time×age	-0.067
Full-time×age^2	0.001
Full-time×college	0.337
Full-time×full-time$_{-1}$	0.268
Part-time	0.197
σ_η	0.522

B.3 Solution method

Let $P_{\sigma^*} := (P^K_{\sigma^*}, P^P_{\sigma^*}, \sigma^{*P,c})$ denote the associated choice probabilities of the MPE σ^*. The terminal values have known functional forms: π^K_d for the child and π^P_d for the parent. I proceed backward in time and apply the following steps for each period t:

(1) I obtain the parent's optimal consumption policy function $\sigma^{*P,c}$ by solving Equation (22).

(2) For $t \geq 2$, I obtain the child's optimal informal care and employment choice probabilities $P^K_{\sigma^*}$:

$$P^K_{\sigma^*}(d^K_t|s_t) = \frac{\exp\left(v^K(s_t, d^K_t; \sigma^*)\right)}{\sum_{d^{K'}_t \in \mathbb{C}^K(h_t)} \exp\left(v^K(s_t, d^{K'}_t; \sigma^*)\right)}. \tag{B.6}$$

(3) For $t = 1$, I obtain the parent's optimal insurance choice probabilities $P^P_{\sigma^*}$:

$$P^P_{\sigma^*}(d^P_t|s_t) = \frac{\exp\left(v^P(s_t, d^P_t; \sigma^*)\right)}{\sum_{d^{P'}_t \in \{0,1\}} \exp\left(v^P(s_t, d^{P'}_t; \sigma^*)\right)}. \tag{B.7}$$

I discretize the parent's wealth into a fine grid and use interpolation for wealth points not contained in the grid. As wealth shocks are assumed to be normally distributed, I use Gauss-Hermite quadrature for numerical integration.

C Two-Step CCP Estimation Details

C.1 Simulation-based value function estimation

This section describes the forward simulation procedure used to estimate value functions. Let me denote each agent's per-period utility by $\tilde{\pi}^i(\cdot)$ which encompasses the agent's flow utility while the parent is alive and terminal utility when the parent dies:

$$\tilde{\pi}^K(d_t^K, s_t, \epsilon_t^K; \theta) = \begin{cases} \pi^K(c_t^K, l_t^K, ic_t^K; h_t^P, ic_{t-1}^K, X^K; \theta) + \epsilon_t^K(d_t^K) \text{ if parent alive,} \\ \pi_d^K(w_t^P; \theta) \text{ if parent dead} \end{cases} \tag{C.1}$$

and

$$\tilde{\pi}^P(c_t^P, d_t^K, s_t, \epsilon_t^P; \theta) = \begin{cases} \pi^P(c_t^P, ic_t^K; h_t^P; \theta) + \epsilon_t^P(d_t^P) \text{ if parent alive,} \\ \pi_d^P(w_t^P; \theta) \text{ if parent dead.} \end{cases} \tag{C.2}$$

As shown in Section 3.1 of the main text and Appendix B.1, for each agent, both the flow utility while the parent is alive and inheritance/bequest utility when the parent dies are linear in the structural parameters that I estimate.[73] Therefore, I can rewrite each agent's per-period utility as

$$\tilde{\pi}^K(d_t^K, s_t, \epsilon_t^K; \theta) \quad = \quad \phi^K(d_t^K, s_t, \epsilon_t^K) \cdot \theta \tag{C.3}$$

and

$$\tilde{\pi}^P(c_t^P, d_t^K, s_t, \epsilon_t^P; \theta) \quad = \quad \phi^P(c_t^P, d_t^K, s_t, \epsilon_t^P) \cdot \theta \tag{C.4}$$

where ϕ^i is a vector of "basis functions" for each agent's per-period utility. As each agent's per-period utility is linear in θ, so too will be their value functions associated with given strategy profile σ:

$$\begin{aligned} V^K(s_t; \sigma; \theta) &= E\left[\sum_{\tau=t}^{T} \beta^{\tau-t} \phi^K\left(\sigma^K(s_\tau, \epsilon_\tau^K), s_\tau, \epsilon_\tau^K\right) \middle| s_t\right] \cdot \theta \\ &= W^K(s_t; \sigma) \cdot \theta \end{aligned} \tag{C.5}$$

and

$$\begin{aligned} V^P(s_t; \sigma; \theta) &= E\left[\sum_{\tau=t}^{T} \beta^{\tau-t} \phi^P\left(\sigma^{P,c}(s_\tau, \sigma^K(s_\tau, \epsilon_\tau^K), \sigma^{P,d}(s_\tau, \epsilon_\tau^P)), \sigma^K(s_\tau, \epsilon_\tau^K), s_\tau, \epsilon_\tau^P\right) \middle| s_t\right] \cdot \theta \\ &= W^P(s_t; \sigma) \cdot \theta. \end{aligned} \tag{C.6}$$

For $i \in \{K, P\}$, W^i is the expected discounted sum of basis functions which does not depend on unknown parameters θ. So once W^i is estimated, I can simply scale it by different parameter values

[73] See Table 6 for the list of the structural parameters estimated within the model.

to obtain value function estimates. A single simulated path based on the policy function estimates $\hat{\sigma} = (\hat{\sigma}^K, \hat{\sigma}^{P,d}, \hat{\sigma}^{P,c})$ is obtained by taking the following steps.

Step 1. For each state s_t, I draw agents' preference shocks.

Step 2. Using the policy function estimates $\hat{\sigma}^K$ and $\hat{\sigma}^{P,d}$, I determine the agents' discrete choices. The parent's consumption is given by $\hat{\sigma}^{P,c}$.

Step 3. Using the determined choices and drawn shocks, I compute the basis functions of each agent's per-period utility, ϕ^i.

Step 4. I draw a new state s_{t+1} using parent wealth and health shock distributions.

Step 5. I repeat Steps 1-4 until the parent dies in which case I sum over the discounted basis functions of each agent's per-period utility.

I draw S simulated paths and average the discounted sum of basis functions over the S simulated paths. This gives me an estimate of $\hat{W}^i$ which is multiplied by θ to result in value function estimates $\hat{V}^i(s_t;\hat{\sigma};\theta) = \hat{W}^i(s_t;\hat{\sigma})\cdot\theta$.

C.2 Pseudo maximum likelihood estimation

The data available for estimation consist of $\{s_{t_{n\tau}}, d^K_{t_{n\tau}}, d^P_{t_{n\tau}}; n = 1, ..., N, \tau = 1, ..., T_n\}$ where N is the number of parent-child pairs, and T_n is the number of interviews in which the nth parent-child pair is observed.[74] Before I define the pseudo likelihood function as in Aguirregabiria and Mira (2007), I first define the likelihood function, which can be obtained from fully solving the model. The likelihood function is given as

$$L^*(\theta) = \prod_{n=1}^{N}\prod_{\tau=1}^{T_n-1} P^K_{\sigma^*}(d^K_{t_{n\tau}}|s_{t_{n\tau}};\theta)P^P_{\sigma^*}(d^P_{t_{n\tau}}|s_{t_{n\tau}};\theta)f(w^P_{t_{n,\tau+1}}|s_{t_{n\tau}}, d^K_{t_{n\tau}}, d^P_{t_{n\tau}}, \sigma^{*P,c}(s_{t_{n\tau}}, d^K_{t_{n\tau}}, d^P_{t_{n\tau}};\theta)) \quad (C.7)$$

where $P_{\sigma^*} = (P^K_{\sigma^*}, P^P_{\sigma^*}, \sigma^{*P,c})$ are the optimal decision rules obtained from solving the model backward as outlined in Appendix B.3 at a vector of parameter values θ. The function $f(\cdot)$ is the conditional density of the parent's wealth, and from Equation (15), it is computed as

$$f(w^P_{t+1}|s_t, d^K_t, d^P_t, c^P_t) = f_m\left((1+r)(\hat{w}^P_t - c^P_t) - w^P_{t+1}\right)^{\mathbb{I}(w^P_{t+1}>0)}\left(1 - F_m\left((1+r)(\hat{w}^P_t - c^P_t)\right)\right)^{\mathbb{I}(w^P_{t+1}=0)} \quad (C.8)$$

where f_m and F_m are the PDF and CDF of the parent's wealth shock, respectively.[75]

The pseudo likelihood function uses an approximation of $P_{\sigma^*} = (P^K_{\sigma^*}, P^P_{\sigma^*}, \sigma^{*P,c})$ by using the value function estimates from the first-stage and thereby avoiding the need to solve the model. The

[74] For pseudo maximum likelihood estimation, I do not use consumption variables which are available only for 25% of my estimation sample. I instead use parents' wealth transition to incorporate consumption choices.

[75] Note that the parent's net assets available for consumption, $\hat{w}^P_t$, depend on the child's informal care choice as it determines formal care costs and also on d^P_t as it determines premium payments.

pseudo likelihood function is given as

$$L(\theta;\hat{\sigma}) = \prod_{n=1}^{N}\prod_{\tau=1}^{T_n-1} \Psi^K(d^K_{t_{n\tau}}|s_{t_{n\tau}};\hat{\sigma};\theta)\Psi^{P,d}(d^P_{t_{n\tau}}|s_{t_{n\tau}};\hat{\sigma};\theta)f(w^P_{t_{n,\tau+1}}|s_{t_{n\tau}},d^K_{t_{n\tau}},d^P_{t_{n\tau}},\Psi^{P,c}(s_{t_{n\tau}},d^K_{t_{n\tau}},d^P_{t_{n\tau}};\hat{\sigma};\theta)) \tag{C.9}$$

where $\Psi = (\Psi^K, \Psi^{P,d}, \Psi^{P,c})$ is called the policy iteration operator or policy improvement mapping as it updates the first-stage policy function estimates $\hat{\sigma} = (\hat{\sigma}^K, \hat{\sigma}^{P,d}, \hat{\sigma}^{P,c})$ by embedding the agents' optimizing behaviors of the current period (Aguirregabiria and Mira, 2002). The policy iteration operator $\Psi = (\Psi^K, \Psi^{P,d}, \Psi^{P,c})$ is computed as the following:

$$\Psi^K(d^K_t|s_t;\hat{\sigma};\theta) = \frac{\exp\left(\hat{v}^K(s_t, d^K_t;\hat{\sigma};\theta)\right)}{\sum_{d^{K'}_t \in \mathbb{C}^K(h^P_t)} \exp\left(\hat{v}^K(s_t, d^{K'}_t;\hat{\sigma};\theta)\right)} \tag{C.10}$$

where $\hat{v}^K(\cdot)$ is defined as in Equation (17) with value function estimates $\hat{V}^K(s_{t+1};\hat{\sigma};\theta) = \hat{W}^K(s_{t+1};\hat{\sigma})\cdot\theta$ in place for $V^K(s_{t+1};\sigma;\theta)$,

$$\Psi^{P,d}(d^P_t|s_t;\hat{\sigma};\theta) = \frac{\exp\left(\hat{v}^P(s_t, d^P_t;\hat{\sigma};\theta)\right)}{\sum_{d^{P'}_t \in \{0,1\}} \exp\left(\hat{v}^P(s_t, d^{P'};\hat{\sigma};\theta)\right)} \tag{C.11}$$

where $\hat{v}^P(\cdot)$ is defined as in Equation (19) with value function estimates $\hat{V}^P(s_{t+1};\hat{\sigma};\theta) = \hat{W}^P(s_{t+1};\hat{\sigma})\cdot\theta$ in place for $V^P(s_{t+1};\sigma;\theta)$, and

$$\Psi^{P,c}(s_t, d^K_t, d^P_t;\hat{\sigma};\theta) = \underset{c^P_t \in (0,\hat{w}^P_t]}{\operatorname{argmax}} \left\{\pi^P(c^P_t, ic^K_t; h^P_t;\theta) + \beta E\left[\hat{V}^P(s_{t+1};\hat{\sigma};\theta)\middle| s_t, d^K_t, d^P_t, c^P_t;\hat{\sigma}\right]\right\}. \tag{C.12}$$

The two-step CCP estimator, denoted by $\hat{\theta}$, maximizes the pseudo likelihood function:

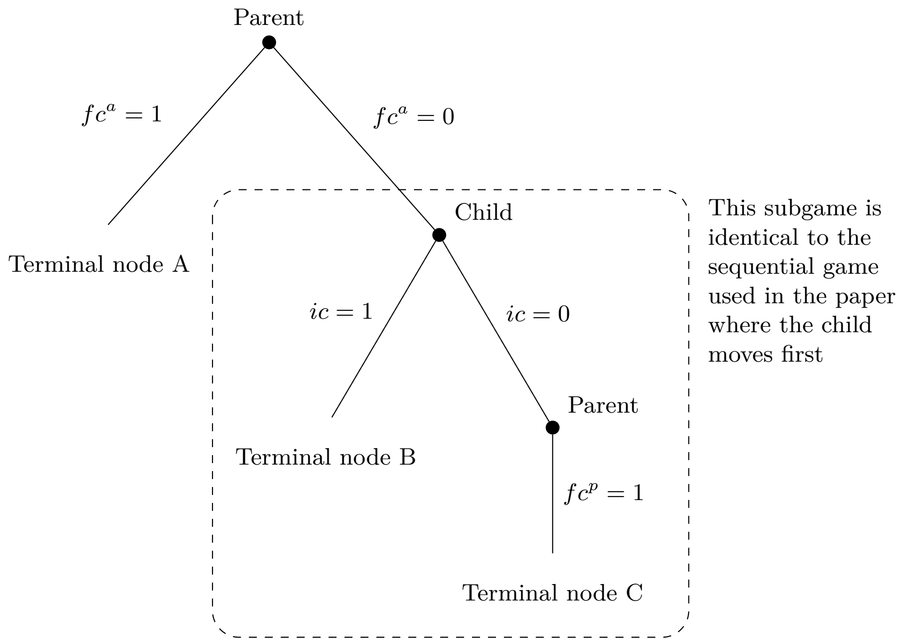
$$\hat{\theta} = \underset{\theta\in\Theta}{\operatorname{argmax}}\ L(\theta;\hat{\sigma}). \tag{C.13}$$

To compute standard errors, I use bootstrapping as in Bajari, Benkard, and Levin (2007).

D Alternative Timing of the Family's Long-Term Care Decisions

The model assumes in each period where the parent is sick, the child moves first by deciding whether to provide informal care, and the parent uses formal care only when the child decides not to provide care. Under an alternative timing assumption, the parent would move first by making a formal care utilization decision, and the child would decide whether to provide informal care only when the parent in the first stage decided not to use formal care. This appendix shows that the family's long-term care arrangements do not change significantly under the alternative timing assumption.

Figure D.1: Alternative Sequential Game



Notes: This game tree represents the sequential game of the family's long-term care decisions under the alternative timing assumption. In the first stage, the sick parent decides whether to use formal care, $fc^a \in \{0, 1\}$. In the second stage, the child decides whether to provide informal care, $ic \in \{0, 1\}$, only if the parent decided not to use formal care in the first stage ($fc^a = 0$). In the third stage, the sick parent "passively" uses formal care ($fc^p = 1$) with probability one, if the child in the second stage decided not to provide informal care. The subgame starting after the parent in the first stage chooses not to use formal care ($fc^a = 0$) is identical to the sequential game used in the paper where the child moves first.

Figure D.1 presents the sequential game over long-term care decisions between a sick parent and his/her child under the alternative timing assumption. In the first stage, the parent decides whether to use formal care $fc^a \in \{0, 1\}$ after observing her preference shocks associated with each formal care choice ($\varepsilon_{fc^a=0}, \varepsilon_{fc^a=1}$). If the parent chooses not to use formal care, then the child in the second stage decides whether to provide informal care $ic \in \{0, 1\}$ after privately observing her preference shocks associated with each informal care choice. When the child decides not to provide informal care, it is implausible to think that the sick parent stays untreated without any care for the entire period, unless there is an incredibly long delay in the child's informal care provision decision. Therefore, I assume if the child does not provide informal care in the second stage, then the sick parent "passively" uses formal care in the third stage ($fc^p = 1$) with probability one.[76] Note that the subgame starting when the parent in the first stage chooses not to use formal care ($fc^a = 0$) is identical to the sequential game over long-term care decisions used in the paper where the child moves first. The implication is that as long as the parent in the first stage has a high

[76]As a robustness check, I have verified that the results do not change significantly when I remove the third stage.

probability of staying away from formal care, the alternative sequential game will regress to the original sequential game where the child moves first.

For the purpose of illustrating the effect of different timing assumptions, I treat the game presented in Figure D.1 as the entire game. That is, I analyze how timing assumptions affect long-term care decisions of a sick parent and his/her adult in a given period and abstract from other features of the model including the child's employment decisions, the parent's savings, bequests, Medicaid etc.

To fix ideas, for a given value of the parent's initial wealth $w \geq 0$ and private long-term care insurance ownership status $ltci \in \{0, 1\}$, let me define the payoff of the sick parent in this simplified model based on the the paper's parametric assumptions on preferences:

$$\tilde{\pi}(fc^a, ic) = \theta_c \frac{c^{1-\rho} - 1}{1 - \rho} + \theta_{fc} \underbrace{\mathbb{I}[fc^a = 1 \text{ or } fc^a + ic = 0]}_{\text{indicator for using formal care}} + \varepsilon_{fc^a} \tag{D.1}$$

where

$$c = w - x_{fc}\mathbb{I}[fc^a = 1 \text{ or } fc^a + ic = 0] \cdot \mathbb{I}[ltci = 0]. \tag{D.2}$$

As assumed in the paper, the parent's utility is additively separable in consumption and preference for long-term care. The preference for informal care has been normalized to zero, and θ_{fc} captures the parent's preference for formal care. The parent's consumption c is equal to her wealth w minus the out-of-pocket cost of formal care. The parent using formal care incurs the formal care cost x_{fc} if and only if she does not own private long-term care insurance, i.e., $ltci = 0$. The parent's payoff for each terminal node in Figure D.1 is then computed as:

$$\text{A:} \;\; \tilde{\pi}_{fc^a=1} := \tilde{\pi}(fc^a=1, ic=\varnothing) = \theta_c \frac{c^{1-\rho} - 1}{1 - \rho} + \theta_{fc} + \varepsilon_{fc^a=1} \text{ where } c = w - x_{fc}\mathbb{I}[ltci = 0] \tag{D.3}$$

$$\text{B:} \;\; \tilde{\pi}_{ic=1} := \tilde{\pi}(fc^a=0, ic=1) = \theta_c \frac{c^{1-\rho} - 1}{1 - \rho} + \varepsilon_{fc^a=0} \text{ where } c = w \tag{D.4}$$

$$\text{C:} \;\; \tilde{\pi}_{ic=0} := \tilde{\pi}(fc^a=0, ic=0) = \theta_c \frac{c^{1-\rho} - 1}{1 - \rho} + \theta_{fc} + \varepsilon_{fc^a=0} \text{ where } c = w - x_{fc}\mathbb{I}[ltci = 0] \tag{D.5}$$

Assume the child's optimal informal care provision probability is given by

$$Pr(ic = 1) \in [0, 1]. \tag{D.6}$$

The parent's expected utility conditional on choosing no formal care in the first stage is computed as

$$\tilde{\pi}_{fc^a=0} \;\; := \;\; Pr(ic = 1) \cdot \tilde{\pi}_{ic=1} + (1 - Pr(ic = 1)) \cdot \tilde{\pi}_{ic=0}. \tag{D.7}$$

It is easy to see that $\varepsilon_{fc^a=0}$ is an additive component of $\tilde{\pi}_{fc^a=0}$. Imposing the assumption that $(\varepsilon_{fc^a=0}, \varepsilon_{fc^a=1})$ follow an *i.i.d.* Type I extreme value distribution with scale one, the parent's optimal formal care choice probability in the first stage is computed as

$$Pr(fc^a = 1) = \frac{\exp(\pi_{fc^a=1})}{\exp(\pi_{fc^a=0}) + \exp(\pi_{fc^a=1})} \tag{D.8}$$

where

$$\pi^P_{fc^a=0} = \tilde{\pi}^P_{fc^a=0} - \varepsilon_{fc^a=0} \tag{D.9}$$

$$\pi^P_{fc^a=1} = \tilde{\pi}^P_{fc^a=1} - \varepsilon_{fc^a=1} \tag{D.10}$$

The probability of the parent ending up with formal care, either as a result of the parent's active choice in the first stage or as the last resort in the third stage, is computed as

$$Pr(fc \text{ is used}|\text{parent moves first}) = \underbrace{Pr(fc^a = 1)}_{\substack{\text{prob. of parent} \\ \text{“actively” using} \\ \text{formal care}}} + \underbrace{(1 - Pr(fc^a = 1)) \cdot (1 - Pr(ic = 1))}_{\substack{\text{prob. of parent} \\ \text{“passively” using formal care}}}. \tag{D.11}$$

Under the original timing assumption used in the paper, the child moves first by deciding whether to provide informal care to the sick parent, and the sick parent uses formal care if and only if the child decides not to provide informal care. The probability of the family using formal care under this original timing assumption is therefore

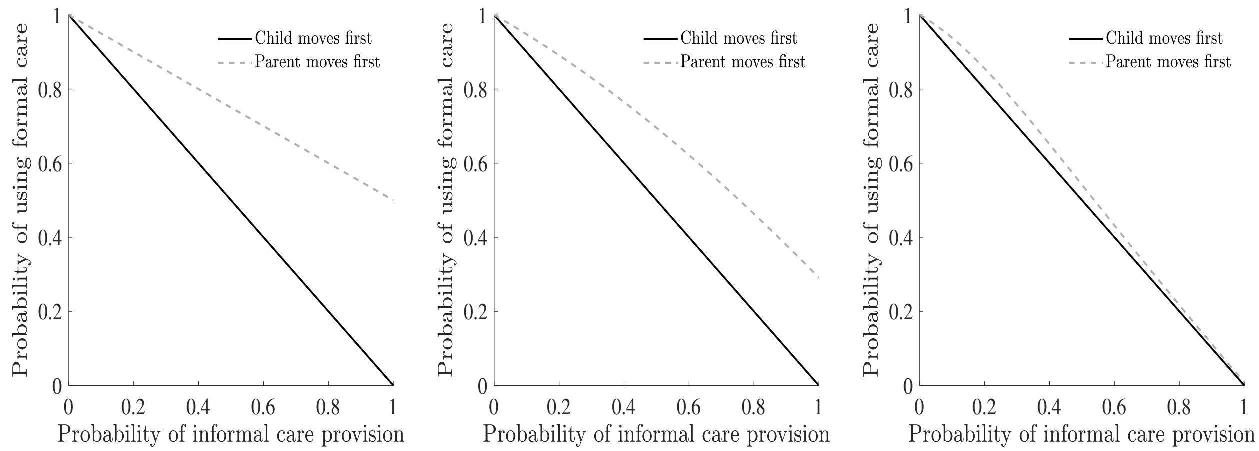
$$Pr(fc \text{ is used}|\text{child moves first}) = 1 - Pr(ic = 1). \tag{D.12}$$

Figure D.2 shows how the relationship between children's informal care provision likelihood (x-axis) and parents' formal care utilization probability (y-axis) is affected by different timing assumptions among families *with* insurance. The three panels in Figure D.2 use different values of the parent's relative preference for formal care: Panel A uses $\theta_{fc} = 0$ which means the parent is indifferent between informal care and formal care, Panel B uses $\theta_{fc} = -0.89$ which is three standard errors *above* the estimated value of θ_{fc} in the paper, and Panel C uses $\theta_{fc} = -4.80$ which is two standard errors *above* the estimated value.

In each panel, the black line represents $Pr(fc \text{ is used}|\text{child moves first})$ as a function of $Pr(ic = 1)$ which has a slope of minus one by construction for all values of θ_{fc}: when the child moves first, the parent uses formal care if and only if the child decides not to provide informal care. The gray dashed line represents $Pr(fc \text{ is used}|\text{parent moves first})$ as a function of $Pr(ic = 1)$.[77] In all panels, this gray dashed line has a strictly negative slope which means that the child's informal care likelihood is negatively correlated with the parent's formal care risk in a substantial way. In other words, even when the parent is the first mover, private information about children's informal care likelihood will be important in predicting consumers' risk in the long-term care insurance market, one of the paper's main results.

[77] Note that when the parent has insurance, $Pr(fc \text{ is used}|\text{parent moves first})$ depends only on θ_{fc} and the child's informal care provision probability.

Figure D.2: The Effect of Timing Assumptions for Families with Insurance



Panel A: $\theta_{fc} = 0$ Panel B: $\theta_{fc} = -0.89$ Panel C: $\theta_{fc} = -4.80$

Notes: This figure is based on the simplified model illustrated in Figure D.1 which describes long-term care decisions between a sick parent and his/her child. Each panel reports, for the indicated value of θ_{fc}, the relationship between children's informal care provision probability (x-axis) and insured parents' formal care utilization probability (y-axis), conditional on who moves first. In each panel, the black line represents $Pr(fc \text{ is used}|\text{child moves first})$, and the gray dashed line represents $Pr(fc \text{ is used}|\text{parent moves first})$. Note that when the parent has insurance, $Pr(fc \text{ is used}|\text{parent moves first})$ depends only on θ_{fc} and the child's informal care provision probability. Panel A uses $\theta_{fc} = 0$ which means the parent is indifferent between informal care and formal care, Panel B uses $\theta_{fc} = -0.89$ which is three standard errors above the estimated value of θ_{fc} in the paper, and Panel C uses $\theta_{fc} = -4.80$ which is two standard errors above the estimated value.

Note that in Panel A, the slope of the gray dashed line is much flatter than that of the black line. This means that when informal care is *not* preferable, the current timing assumption ("child moves first") could overpredict the importance of the informal care likelihood in the long-term care insurance market compared to the alternative timing assumption. However, as soon as I assume that the parent prefers informal care to formal care, the effect of different timing assumptions quickly disappears. The intuition is the following: when the parent prefers informal care to formal care, even if the parent moves first by making the formal care utilization decision, the parent will choose to stay "untreated" and wait for the child to make the informal care decision, hoping to receive preferable informal care. In other words, in the game tree illustrated in Figure D.1, the parent in the first stage will most likely choose $fc^a = 0$, and the alternative sequential game will regress to the original sequential game where the child moves first.

I have also replicated Figure D.2 for families *without* insurance. As suspected, the effect of timing assumptions are even *smaller* based on reasonable values of other parameters needed to compute $Pr(fc \text{ is used}|\text{parent moves first})$ when the parent does not have insurance.[78] The intuition is

[78]These parameters are $(w, x_{fc}, \rho, \theta_c)$. w is set to the parent's mean wealth in the data, and (x_{fc}, ρ, θ_c) are set to the values used in the paper. I have also experimented with other values of these parameters and the results are robust.

simple: uninsured parents in the first stage are even more likely to choose $fc^a = 0$ as they have to pay for formal care, increasing the likelihood of being back in the original game where the child moves first.