Explaining the Decline of the U.S. Saving Rate: the Role of Health Expenditure

Yi Chen, Maurizio Mazzocco, and Bela Szemely

Explaining the Decline of the U.S. Saving Rate: the Role of Health Expenditure[*]

Yi Chen[†]   Maurizio Mazzocco[‡]   Béla Személy[§]

Current Draft, December 2017.

Abstract

Between 1980 and 2007, the U.S. personal saving rate declined from 11 percent to 3 percent. In this paper, we provide evidence that most of the decline can be explained by rising health expenditures. We first employ reduced-form methods and exogenous variation in medical expenses generated by FDA drug approvals to document that a 1 percentage point increase in health expenditure generated a decline in the saving rate that is between 0.73 and 0.89 percentage points. Using this result, we calculate that the rise in health expenditure explains about 90 percent of the drop in the saving rate. We then develop and estimate a model of consumption, saving, and health decisions to evaluate which mechanisms are behind the decline. Using the estimated model, we document that the rise in medical expenses and drop in the saving rate are driven mainly by progress in health technology during the years 1995-2010, by the reduction in co-payment rates in the period 1986-1994, and by improvements in income processes for the years 1980-1985.

1

1   Introduction

It is well-known that the U.S. personal saving rate declined from 11 percent in the late seventies to about 3 percent in the late 2000s. Economists and policy-makers are concerned with this drop because it may signal an increased dependence on foreign investment and a future reduction in capital stock with negative consequences for labor productivity, wages, and national output. In the past twenty years economists have attempted to explain this sharp drop. An examination of the related literature indicates that the decline is still a puzzle. Parker (1999) states that "Each of the major current theories of the decline in the U.S. saving rate fails on its own to match significant aspects of the macroeconomic or household data." Guidolin and La Jeunesse (2007) review a number of arguments and theories that have been proposed and conclude that "The recent decline of the U.S. private saving rate remains a puzzle."

The main contribution of this paper is to provide evidence that the increase in medical expenditure as a share of disposable income is the major driver of the decline in the U.S. personal saving rate. The definition of the personal saving rate used by the saving literature and by the National Income and Personal Accounts (NIPA) includes total health expenses in the computation of household-sector expenditure, i.e. health expenses paid by consumers plus health expenses paid by some form of health insurance. We document that the decline is mostly driven by health expenses paid by some form of medical insurance. We provide this evidence in three steps.

In the first step, we use a simple accounting exercise to document that the share of health expenditure on its own can explain most of the drop in the U.S. saving rate that started at the end of the seventies. Papers in the literature have discarded health expenditure as a possible explanation for the decline in the U.S. saving rate because this variable has been growing since the early sixties. It has been argued that, if this variable was the main determinant of the drop, the saving rate should have started its decline in the sixties. In the paper, we confirm the finding that health expenditure has been rising since the sixties. But, we provide evidence that, in spite of this, from an accounting perspective medical expenses can be the main driver of the drop in the saving rate because of the following remarkable coincidence. From the early sixties to the end of the seventies, the social security benefits paid to households increased as a fraction of income at the same rate as the share of health expenditure, therefore offsetting the negative effect of health expenditure on the saving rate. Starting from the end of the seventies, however, the share of social

security benefits stopped increasing, whereas medical expenses continued to rise at the same rate. Medical expenditure on its own has therefore the potential to explain most of the drop in the saving rate. We confirm this hypothesis by considering a simple accounting exercise in which we set health expenditure net of social security benefits equal to a constant fraction of disposable income and compute the corresponding saving rate. Using NIPA data, we find that under this scenario the saving rate fluctuates around a constant rate from the sixties to today.

The main limitation of the accounting exercise is that it does not allow us to make causal statements about the rise in medical expenditure and the drop in the U.S. saving rate. To address this issue, as a second step, we provide evidence on a causal relationship between health expenditure and the personal saving rate using cross-state changes in our two main variables and, as an arguably exogenous source of variation in medical expenses, the approval of new drugs by the Food and Drug Administration (FDA). The idea behind the choice of this instrument is straightforward. When a new drug is approved and made available to users, medical expenses will generally increase since consumers have a new product that can help treat medical conditions. A limitation of this variable is that, in each period, it is constant for the entire U.S., whereas our analysis requires changes across states. To generate the required cross-state variation, we interact the FDA approval of new drugs with the demographic characteristics of a state. The constructed variables provide the needed variation, since the approval of a new drug produces larger increases in health expenditure in states with a larger fraction of families that make extensive use of medical products. We find that a 1 percentage point increase in health expenses generates a reduction in the saving rate of between 0.73 and 0.89 percentage points. Using our estimates and back of the envelope calculations, we conclude that about 90 percent of the drop in the U.S. saving rate can be explained by the rise in health expenditure.

Using the reduced-form results alone, it is difficult to evaluate which mechanisms are behind the rise in medical expenditure and the corresponding decline in the personal saving rate. To deal with this limitation, in the last part of the paper we develop and estimate a model of consumption, saving, and health decisions. In the model, health expenses are beneficial because they improve health status and, hence, reduce the individual mortality rate. People may be affected by two types of health conditions: non-severe, such as high blood pressure and the flu; and severe, such as cancer and heart attacks. Individuals can choose whether to undertake treatment for non-severe conditions, but must seek treatment for the severe conditions or their mortality probability

3

increases to one. Each person is covered by a private or public health insurance. Thus, incurred total medical expenses can be divided into two categories: expenses paid by the person seeking treatment; and the medical expenses paid by the private or public health insurance that covers the person. In the model, to be consistent with the definition of the saving rate used by NIPA and in the literature, both types of expenses enter the calculation of the household-sector saving rate. But only the health expenses and the health insurance premium paid by a person affect individual decisions.

The model is estimated using data from the Medical Expenditure Panel Survey (MEPS), the NIPA, the Current Population Survey (CPS), the National Health Interview Survey (NHIS), and the National Death Index (NDI) Mortality Files. The corresponding simulations indicate that the rise in medical expenditure explains most of the decline in the saving rate experienced by the U.S. economy. In our model, there are three main mechanisms that can generate the decline: progress in health technology, which induces higher medical expenses; changes in co-payment rates and the corresponding variation in the probability of seeking treatment; and changes in income with the associated changes in treatment rates. Using counterfactuals, we find that technological progress accounts for about 50 percent of the decline in the saving rate and is mainly responsible for the drop experienced by the U.S. economy between 1995 and 2010. Changes in co-payment rates explain about 25 percent of the reduction in the saving rate and account for most of its decline during the period 1986-1994, when the co-payment rates experienced a steep drop. Lastly, improvements in the income process explain the remaining 25 percent and are responsible for most of the reduction in the saving rate during the years 1980-1985 and part of the drop for the period 1986-1994.

There is one paper that is particularly relevant to understand our results. De Nardi, French, and Jones (2010) study the effect of out-of-pocket medical expenses on savings of older households and find, using the Health and Retirement Study, that higher out-of-pocket medical expenses increase their savings. At first sight, this finding appears to contradict our results. But this is not the case, since the two papers focus on the effect of two different variables on saving decisions. De Nardi, French, and Jones (2010) analyze the effect of out-of-pocket expenses, whereas we consider the effect of total health expenditure, since the aggregate saving rate is computed using all health expenses. We document that during the period 1980-2010 out-of-pocket expenses – health expenses plus health insurance premiums paid by consumers – remained approximately

constant in the aggregate. It was the rise in medical expenditure paid by some form of health insurance that produced the decline in the saving rate. Our results therefore do not contradict the findings in De Nardi, French, and Jones (2010), but complement them in understanding the impact of medical costs on saving decisions.

Our results have policy implications. They indicate that if policy makers intend to raise the saving rate, it is essential to reduce the rate at which medical expenses grow. The recent debate on health costs has focused on two sets of policies that have the potential of reducing the effect of increasing medical expenses on the U.S. economy. The first set includes policies aimed at eliminating inefficiencies in the provision of health care. This is clearly a good starting point, to the extent that it is able to generate significant reductions in medical costs. The second set of policies requires health institutions to be more selective in the adoption of newer technologies. These policies can have a large impact on health expenses and, hence, on the saving rate. But they come at a cost. First, as pointed out by Hall and Jones (2007), if health care is a luxury good, in a rich country such as the U.S. it may be welfare improving to adopt the latest health technology. Second, the constant adoption of new health technologies has the positive effect of fostering a large number of innovations in the health sector. A more frequent use of older technologies might slow this progress and could have negative welfare effects in the long run.

The rest of the paper proceeds as follows. The next section provides a discussion of related papers. In section 3, we describe the data sets and define the variables used in the paper. Section 4 uses an accounting exercise to document that most of the decline in the U.S. household saving rate can be explained by the rise in health expenditure. In section 5, we use cross-state variation to provide evidence on a causal relationship between health expenditure and the saving rate. Section 6 develops and estimates a model to evaluate which mechanisms are behind the increase in medical expenditure and the corresponding decline in the saving rate. Section 7 concludes.

2   Related Papers

A large number of studies have analyzed the decline in the U.S. saving rate. In this section, we will discuss the papers with findings that are related to ours. For a thorough review of the literature see Browning and Lusardi (1996), Parker (1999), and Guidolin and La Jeunesse (2007).

One of the first papers to address the decline in the U.S. saving rate is the work by Summers

and Carroll (1987), where they study the changes in the national and household saving rate from the fifties to 1986. Their main conclusion is that the decline in the private sector U.S. saving rate is real and not a result of measurement issues and that the most likely cause of the decline is the increase in expected income after retirement that has induced the younger cohorts to reduce the rate at which they save. In our paper, we only focus on the household saving rate which experienced most of its decline after 1986, the last year considered by Summers and Carroll.

The paper by Gokhale, Kotlikoff, and Sabelhaus (1996) is the first one to discuss the steep increase in health expenditure in the past 50 years and to suggest a possible relationship with the drop in the saving rate. They report that medical consumption as a percentage of disposable income was 3.9 in the fifties, 5.2 in the sixties, 7.3 in the seventies, 10.1 in the eighties, and 12.8 in the early nineties. This pattern suggests that it is difficult for medical expenditure to explain the decline in the saving rate. Health expenses were already growing in the sixties and seventies, whereas the saving rate started its decline in the eighties. Probably for this reason, Gokhale, Kotlikoff, and Sabelhaus (1996) do not directly explore the effect of the increase in medical consumption on household savings. Instead, they use it mainly as a motivation for decomposing the changes in the saving rate in four components: the changes generated by variations in the intergenerational distribution of resources; the effects produced by changes in the cohort-specific consumption propensities; the changes produced by modifications in the rate of government spending; and the effects of changes in demographics. This decomposition is based on a simple life-cycle model and relies on the assumptions implicit in it. Their results suggest that the decline in the saving rate is mainly a consequence of government redistribution of resources from the young generation with low propensities to consume to the old generation with higher propensities. Our results differ from theirs in three respects. First, we document that the sharp increase in health expenditure can explain on its own the decline in the U.S. saving rate if considered jointly with the evolution of social security benefits to households. Second, we use exogenous variation in medical expenditure to provide evidence on the existence of a causal relationship between health expenses and saving rates. Lastly, we estimate a model that accounts for the effects of health decisions and outcomes on saving choices to evaluate the importance of different mechanisms.

Attanasio (1998) uses the Consumer Expenditure Survey (CEX) and synthetic cohorts to study the reduction in the U.S. saving rate for the period 1981-1991. His main conclusion is that the decline was mostly generated by cohorts born between 1925 and 1939. In this paper we study

6

a longer sample period and provide evidence that the U.S. saving rate continued to decline for another two decades. This suggests that many more cohorts are responsible for the observed pattern in saving rates.

In a related paper, Parker (1999) considers the main explanations given in the literature for the decline in the saving rate, namely the effect of asset value appreciations, durable goods, changes in the intergenerational distribution of resources, financial innovations, and changes in discount factors. He then evaluates whether these hypotheses are consistent with patterns observed in micro and macro data. He concludes that "Each of the major current theories of the decline in the U.S. saving rate fails on its own to match significant aspects of the macroeconomic or household data." In particular, he rejects the hypothesis that changes in the intergenerational distribution of resources, which include the increase in medical expenditure, can explain the decline for two reasons. First, the trends in the government redistribution of resources and the increase in health expenditure predate the drop in the saving rate. Second, the ratio of consumption to income increases for all generations and not only the old ones. Both conclusions are correct. Without considering the growth in the share of social security benefits during the sixties and seventies, health expenditure cannot explain the decline in the saving rate. Moreover, the ratio of consumption to income increased for all generations, because health expenses increased for all age groups and not only for the elderly.

Gale and Sabelhaus (1999) starts from the observation that one may require different measures of aggregate savings to answer different economic questions. They then study the evolution of the U.S. saving rate using several measures of this aggregate variable. They first use the measure employed by the NIPA. In this case, the pattern displayed by the saving rate is consistent with the findings of previous papers, which document a steep decline that starts in the late seventies. They then consider an alternative measure that modifies the NIPA definition of the saving rate in a way that enables them to take into account the evolution of durable goods, retirement accounts, inflation, and tax accruals. Using this measure, the authors still report a decline in the saving rate but of a smaller magnitude. Gale and Sabelhaus consider also a third measure that adds capital gains. In this case, as one may expect, the saving rate at the end of the nineties was at the highest level in the last forty years. In our paper, we only consider the standard measure of aggregate savings provided by NIPA and we have nothing to say about the evolution of alternative measures.

The most recent survey we could find is the paper by Guidolin and La Jeunesse (2007). The first part of the paper examines whether the decline in the U.S. saving rate is real or a simple statistical artifact generated by measurement issues. Since the decline is evident in all the standard measures considered in their paper, they conclude that it cannot be easily explained by measurement issues. In the second part of the paper, Guidolin and La Jeunesse review many of the theories that have been proposed and conclude that the drop in the saving rate is still a puzzle.

We conclude this section with one remark. In this paper, we do not make statements about the optimal saving rate for the U.S. economy. We do not argue that a saving rate of 3 percent is low or that a saving rate of 11 percent is optimal. We only provide an explanation for its decline. For a discussion on the optimality of the current saving rate, one can read Lusardi, Skinner, and Venti (2001) and Scholz, Seshadri, and Khitatrakun (2006).

3 Data Description and Variable Definition

In this section we describe the main data sets used in the paper, which are the NIPA, the NHEA, retail sale data prepared by the private company Claritas, the March CPS, the SEER, the MEPS, the NHIS, and the NDI Mortality Files.

The NIPA are published by the Bureau of Economic Analysis and constitute the major source of aggregate data for the U.S. economy. We use annual data for the years 1960-2010. All files were downloaded on March 2015 from the BEA website (http://bea.gov/national/index.htm). We construct our expenditure, income, and savings variables using data from Table 2.1, Personal Income and Its Disposition, and Table 2.4.5, Personal Consumption Expenditures by Function. Specifically, we compute health expenditure as the sum of the items in the NIPA Table 2.4.5, which are listed in the first part of Table 1. Other variables used in the analysis are defined in the second part of Table 1.

The NHEA have been published since 1964 by the Department of Health and Human Services. The NHEA provide not only a comprehensive measure of total spending on health care goods and services, but also a breakdown of the sources of funds that finance these expenditures. This level of detail is important because it allows us to determine how health expenses were funded. All NHEA files were downloaded on March 2015 from the Centers for Medicare and Medicaid Services website (http://www.cms.hhs.gov). To make the NHEA data as comparable as possible

Table 1: NIPA Data

Components of Personal Health Expenditures	
Item	Line in NIPA Table 2.4.5
Therapeutic appliances and equipment	line 21
Pharmaceutical and other medical products	line 40
Health care	line 60
Net health insurance	line 93
Other Variables	
Item	Line in NIPA Tables
Personal saving rate	line 35 in NIPA Table 2.1
Social security benefits to households	line 18 in NIPA 2.1
Employer contributions for health plans	line 17 in NIPA Table 7.8
Personal taxes	line 26 in NIPA Table 2.1
Government current expenditures by function	various lines in NIPA Table 3.16

to the NIPA health data, we use the Personal Health Care data from the NHEA which tracks the personal health expenditure measures in the NIPA remarkably well. Throughout the period 1960-2009, the period for which the NHEA has data available, the ratio of the two health expenditure measures obtained using the NIPA and NHEA is between 0.94 - 1.06.[1] The variables that we use from the NHEA are defined in Table 2. The measure of total health expenditure reported by the NHEA is available yearly for the entire U.S. and also for each state separately.

Retail sales are one of the main inputs in the construction of the saving rate at the state level. We employ the most commonly used retail sale data which are prepared by the private company Claritas and can be found in the Survey of Buying Power published by the Sales & Marketing Management magazine. These data were first used by Asdrubali, Sorensen, and Yosha (1996). We use the March CPS to perform robustness checks in the cross-state regressions and to estimate the income processes required in the simulation of the model. We use the SEER to compute population estimates.

We use three data sets to estimate the health and mortality processes in our model: the MEPS, which collects health status, health conditions, and medical expenditure for a nationally representative sample of the United States; the NHIS, which gathers data on self-reported health

---

[1]The reconciliation project discussed in Sensenig and Wilcox (2001) documents that the NHEA data are generally compatible with the NIPA data.

Table 2: Definitions of Variables from NHEA

Item	Line in NHEA Data
Personal Health Care	line 73
Out of pocket expenditure	line 74
Health Institutions	line 73 minus line 74
Private Health Insurance	line 76
Private Funds	line 74+76+87+88
Public Funds	line 73 minus line above
Gov't programs	line 102
Medicare	line 77
Medicaid	line 78
Health insurance premiums paid by employers	NIPA Table 7.8, line 16
Health insurance premiums paid by households	NIPA Table 2.4.5, line 93

status and health conditions; and the NDI Mortality Files, which contain mortality data. The NHIS has the advantage over the MEPS that it can be linked to the NDI Mortality Files and, hence, can be used to estimate the relationship between health status and survival probabilities. The NHIS assigns a numerical value to each catergory of self-reported health status, with lower values assigned to healthier people (poor = 5, fair = 4, good = 3, very good = 2, and excellent = 1). In the estimation, we use these values.

4 An Accounting Exercise

In this section we use a simple accounting exercise to document that the drop in the personal saving rate that started in the late seventies can be explained by the rise in health expenditure.

We start by introducing the definition that is commonly used by NIPA and the savings literature to compute the aggregate saving rate of the household sector:

$$s_t = \frac{\text{disposable income - total expenditure}}{\text{disposable income}} = 1 - \frac{\text{non-health expenditure}}{\text{disposable income}} - \frac{\text{health expenditure}}{\text{disposable income}}, \quad (1)$$

where health expenditure is the sum of out-of-pocket expenses – health expenses and health insurance premiums paid by consumers – and health expenses paid by some form of health insurance net of the premiums paid by consumers.

From an accounting perspective, this definition makes clear how total medical expenditure affects the saving rate. Since it is included in total expenditure, an increase in total medical expenses

reduces the saving rate unless there is a corresponding decline in other types of expenditures or a corresponding increase in some of the variables included in disposable income.

In Figure 1, we illustrate the evolution of s_t and total medical expenditure as a share of disposable income from 1960 to 2009. As many other papers have documented, during the sixties and seventies the saving rate in the household sector fluctuated around 11%. But in the late seventies it started to decline until, in the second half of the last decade, it reached the unusually low level of 3%.[2] The share of health expenses followed a different trend. As documented in Gokhale, Kotlikoff, and Sabelhaus (1996), this variable increased steadily throughout the period under consideration.

Since medical expenses started to grow in the sixties but the saving rate began to decline two decades later, health expenditure has been dismissed as a possible explanation for its decline. From an accounting perspective, however, this variable can still explain the decline because of the following coincidence. As we document in Figure 2, during the sixties and seventies the growth in the share of medical expenses was matched by a similar increase in the share of social security benefits, which NIPA include in disposable income, therefore offsetting the effect of health expenditure on the saving rate. But starting in the late seventies, the share of social security benefits remained approximately constant around 6 percent of disposable income, leaving the effect of medical expenditure on the saving rate unmatched. To better illustrate this point, in the figure, we also report the evolution of a new variable obtained by subtracting the share of social security benefits from the share of medical expenses. This variable can clearly explain the decline in the saving rate, since it was constant until the end of the seventies and grew at a steep rate since then.

To test this hypothesis, we perform a simple accounting exercise. We consider the evolution of the saving rate for a hypothetical situation in which health expenditure net of social security benefits is set equal to a constant fraction of income throughout the period. Without loss of generality, we consider the case in which the constant fraction is zero. If the increase in medical expenditure net of social security benefits is the main driver of the drop in the saving rate, in this hypothetical exercise we should observe that the saving rate fluctuates around a constant rate throughout the period. Figure 3, in which we report the share of health expenditure net of

---

[2]In the Appendix, we report the "filtered" time-series of the saving rate, which suggests that the decline started in the second part of the seventies.

social security benefits, the actual saving rate, and the hypothetical saving rate, indicates that this is the case. The hypothetical saving rate fluctuates around 15% for the entire period. It is noteworthy that with our variable we can explain even the slight decline in the saving rate that occurred in the second part of the seventies.

The accounting exercise performed in this section is informative because it suggests a potential explanation for the decline in the U.S. saving rate. It cannot be used, however, to establish whether there is a causal relationship between changes in health expenditure and changes in the saving rate, since health expenditure is, at least partially, endogenous. If health expenses increase, people optimal response may be to reduce the number of times they seek treatment or to shrink the consumption of other goods, which would diminish the effect of medical expenditure on the saving rate. To address this issue, we first account for the endogeneity of medical decisions by using cross-state variation and the FDA approval of new drugs. We then consider their endogeneity directly by developing, estimating, and simulating a model in which individuals make consumption, saving, and health decisions.

5   Cross-state Analysis

In this section, we estimate the effect of rising health expenditure on the U.S. saving rate by using variation across states in our two main variables and, as an arguably exogenous source of variation in medical expense, the FDA approval of new drugs. Because of data constraints, we can only perform the cross-state analysis for the sample period 1980-2009. Since during those years social security benefits to the household sector stayed approximately constant, in our analysis we will consider their effect on the saving rate only as part of disposable income.

5.1   Changes in Health Expenditure Across States

The idea behind using cross-state variation to determine whether there is a causal relationship between health expenditure and the saving rate is straightforward. If there is a causal relationship, states that experience larger increases in health expenditure should display larger declines in their saving rate. If this is not the case, we can reject that the relationship is causal.

The implementation of this strategy requires the construction of saving rates at the state level. The computation of state-level saving rates requires knowledge of disposable income and

total expenditure for each state. State-level disposable income can be easily measured, since data on this variable have been regularly published by NIPA since 1948. The computation of state-level total expenditure is more complicated since there is no readily available measure for this variable. To overcome this problem, we experimented with two different approaches. We first attempted to construct this variable by using micro-level data on household expenditure. Since the most reliable source of household-level expenditure in the U.S. is the Consumer Expenditure Survey (CEX), we employed this data set. There are two main issues with using the CEX to construct state-level expenditure. First, the sample size for most states is small, which makes this measure of expenditure imprecise and volatile. Second, a well-known fact in the savings literature is that the changes in the aggregate U.S. saving rate obtained using the CEX differ in a significant way from the changes constructed using NIPA data.[3] Our analysis confirms this result and therefore disqualifies this first approach as a possible way of understanding the reasons behind the recent decline in the U.S. saving rate.

A second possible method for constructing state-level expenditure is to use retail sale data. Evidence has been provided that retail sale data approximate well household expenditure at the aggregate level if one is interested in changes and not in levels.[4] Since in this section we exploit changes over time in savings rates, all we need is a state-level variable that, divided by disposable income, can replicate well the changes in the share of household expenditure. The retail sale data are, therefore, well suited for the construction of our variable. As mentioned in the data section, the retail sale data used in this paper are prepared by the private company Claritas and can be found in the Survey of Buying Power published by the Sales & Marketing Management magazine.

There are four main expenditure components that are missing from retail sale data: health expenditure, expenses related to renting a house, the rental value associated with owning a house, and expenditure on services. Unfortunately, there is no state-level measure of household expenditure on services. There exist, however, state-level measures for the other three components. We add them to retail sales to improve the ability of our measure to approximate household expenditure. Health expenditure at the state level can be constructed using data from the NHEA. The NHEA publishes two measures at the state level: by state of the provider of the health service and by residence of the consumer buying the health service. Since the data by residence are only

---

[3]See for instance Slesnick (2001), Garner et al. (2006), and Heathcote, Perri, and Violante (2010).

[4]See for instance the detailed and careful discussion in Zhou (2010).

available starting in 1991, we use as our main variable state-level health expenses by provider. The second expenditure component that is missing from retail sale data - expenses related to renting or owning a house - can be approximated using the state-level NIPA measure of housing services which is available as part of disposable income. For renters, it corresponds to the rent paid to the owner minus expenses related to housing services such as depreciation, maintenance and repairs, property taxes, and mortgage interest. For owners, it is constructed as the imputed rental value of the house minus the housing expenses described above. This state-level measure of housing services is only an approximation of the true measure because depreciation, property taxes, and mortgage interest should be included in it. Notice that expenses from maintenance and repairs are correctly subtracted from our measure of housing expenses since they are already included in the retail sale data.

In Figures 4 and 5, we provide evidence on the ability of our measure of the share of household expenditure to approximate the variable reported by NIPA. Since the NHEA started publishing health expenditure data in 1980 and the last year for which we have retail sale data is 2009, our expenditure measure covers the period 1980-2009. In Figure 4, we describe the ratio between the aggregate household expenditure obtained using our measure and using NIPA data. As expected, our measure is always below the variable constructed by NIPA. But it is noteworthy that the ratio between the two measures is approximately constant around 65% during the entire sample period. This result suggests that our measure should be able to approximate the changes in the share of household expenditure as measured by NIPA. To confirm this, in Figure 5 we plot the changes between t and $t+1$ for the NIPA saving rate and the corresponding changes computed using our measure. The graph indicates that our measure matches remarkably well the changes constructed using the NIPA data.[5]

We can now use cross-state variation to estimate the effect of a change in the share of health expenditure on the saving rate. In Figure 6, we report preliminary evidence on the relationship between these two variables by plotting them for each state and time period. The first panel reports the two variables in levels, whereas the second panel describes them using changes. In

---

[5]During our sample period, one data point is missing from the retail sale data. Until 2000 the survey recorded retail sale data for the previous calendar year. But starting from 2000, it began reporting retail sale data for the current year. The retail sale number for 1999 is therefore missing. As a consequence, when we use changes in the saving rate, two data points are missing.

both panels, we also report a line obtained by regressing the state-level saving rates on state-level health expenditures. The two scatter plots indicate that there is a negative relationship between saving rates and medical expenses both in levels and changes: states with higher levels of medical expenditure have lower saving rates; and states that experience larger increases in health expenditure display larger declines in their saving rate.

To provide more evidence, let $\bar{c}_t^h$ be the share of total health expenditure, defined as the share of health expenses paid by consumers plus the share paid by some form of health insurance. Moreover, denote with f_t and f_s time and state fixed effects, and with $X_{t,s}$ a set of control variables that vary with time and across states. We can then provide a first set of estimates on the effect of changes in the share of total health expenditure on changes in saving rates by applying Ordinary Least Squares (OLS) to the following equation:

$$s_{t,s} = \alpha_0 + \alpha_1 \bar{c}_{t,s}^h + \alpha_2 X_{t,s} + f_t + f_s + \epsilon_{t,s}. \tag{2}$$

In the estimation, we employ total medical expenses as our main explanatory variable and not just the part paid by consumers, because it is the variable that enters the aggregate saving rate of the household sector. The coefficient of interest is α_1, which measures the percentage point change in the saving rate that corresponds to a one percentage point change in the share of total health expenditure.

Standard theories indicate that saving decisions and therefore the saving rate depend on risk preferences, permanent income, and the degree of uncertainty faced by the individual. Variables that are constant over time such as risk preferences are already captured in equation (2) by the state fixed effects. To account for the time-varying variables that may affect the saving rate independently of health expenditure, we include in $X_{t,s}$ the following state-level variables. As proxies for the degree of uncertainty, we include the unemployment rate and the fraction of individuals with college or higher degree. As proxies for permanent income, we include the fraction of individuals with college or higher degree, the fraction of individuals between the ages of 30 and 60, the fraction older than 60, the fraction of African-Americans, and the fraction of Hispanics.

Before presenting the OLS results, it is important to make two remarks. First, notice that the share of health expenditure enters the right hand side of equation (2) as our main independent variable and it also affects our dependent variable, as the definition of the saving rate (1) makes clear. In spite of this, regression (2) does not simply captures a mechanical relationship between

$\bar{c}_t^h$ on the right hand side and $\bar{c}_t^h$ inside the saving rate. When there is an increase in the share of health expenses, consumers can choose to reduce the consumption of other goods if it is optimal for their saving rate to stay at a higher level. If this is the case in the data, the coefficient α_1 will be estimated to be lower than one in absolute value. Consumers may also choose to increase the consumption of goods that have some degree of complementarity with health expenses. If the increase is large enough, the coefficient α_1 will be estimated to be greater than one in absolute value.

As a second remark, note that medical expenditure entering both the dependent variable and our main explanatory variable may also affect our results if medical expenditure is measured with error. If this is the case and the measurement errors are classical, our estimated coefficient for α_1 will be biased toward zero. Our IV estimates, which we present in the next subsection, enable us to address this potential issue as long as our instruments are not affected by the same measurement errors.

The OLS results are presented in Table 3. The parameter of interest, α_1, is estimated to be equal to -0.643. This number implies that a 1 percentage point increase in the share of medical expenditure reduces the saving rate by 0.643 percentage points. This result suggests that people responded to the increase in medical expenses by partially reducing the consumption of other goods. With regard to the control variables, we find that a one percentage point increase in the share of middle-aged individuals increases the state saving rate by 0.93 percentage points. The same change in the share of Hispanics increases the saving rate by 0.40 percentage points. A one percentage point increase in unemployment rate raises the saving rate by 0.41 percentage points. Finally, the share of individuals older than 60, the share of individuals with college or higher degree, and the share of African-Americans have no significant effect on the dependent variable.

There are at least three reasons related to omitted variables that prevent a causal interpretation of our OLS results. Households tend to migrate to states with more employment opportunities and consequently higher disposable income. In addition, most of the migration is driven by individuals with high education, who generally have lower medical expenses. If the propensity to save increases with income, these two patterns generate a negative correlation between saving rates and health expenses that could bias our OLS estimate of α_1 and make it more negative. If the propensity to save decreases with income, the correlation between our two main variables will be positive and the estimated α_1 could be biased toward zero. A second potential problem is related

to economic shocks that are state specific. Since they simultaneously affect saving and health expenditure decisions, the negative estimate for α_1 may be the result of this omitted variable. Lastly, as mentioned above, measurement errors in medical expenses may bias the estimated α_1 toward zero. To deal with these possible threats to a causal interpretation of the OLS estimates, we will instrument for the state-level health expenditure using the FDA approval of new drugs interacted with the demographic structure of the corresponding state.

5.2 Instrumental Variable Strategy

To overcome the endogeneity issue affecting health expenses, we need a variable that is correlated with state-level health expenditure, but uncorrelated with the error term of equation (2). A good candidate for a variable that affects the state-level saving rate only through state-level health expenditure is the FDA approval of new drugs. This variable clearly affects health expenses, since the approval of new drugs has generally a positive effect on them. In addition, it is arguably exogenous to changes in the saving rate after the variation in medical expenditure is taken into account.

We construct the variable FDA approval of new drugs using data that are publicly available from the FDA website.[6] There are two kinds of approvals: (i) approvals of new molecular entities and (ii) approvals of new drugs composed of old molecular entities.[7] Since new molecular entities are the most likely to increase medical expenditure, we construct our approval variable using the first category. The variable is construced for the period 1980-2009, which are the years for which we have state-level health expenditure data.

A limitation of the FDA approval variable is that it does not vary across states. To address this issue, we follow the strategy adopted by other papers of interacting a global shock with local conditions (see for instance Acemoglu, Finkelstein, and Notowidigdo (2013)). Specifically,

---

[6]The data on drug approvals go back to 1940 and can be downloaded at http://www.fda.gov/Drugs/InformationOnDrugs/ucm079750.htm. These data were first used by Lichtenberg and Virabhak (2007) as one of the inputs to estimate the effect of drugs approved in different years on post-treatment health. Acemoglu and Linn (2004) also employ those data as one of the variables used to construct market size.

[7]The second category includes drugs having as ingredients new ester, new salt, or other noncovalent derivative, drugs with a new formulation, drugs with a new combination of old molecular entities, drugs with a new manufacturer, and drugs with new indications.

the federal approval of new drugs has different effects on local economies depending on their demographic structure. All else equal, states with a larger fraction of older individuals should see a steeper increase in medical expenditure after the approval of a particular drug, since older people tend to have higher demand for medical products. Similarly, states with a higher fraction of college graduates should experience a smaller increase in health expenses if more educated individuals are healthier and, hence, spend less on medical consumption. Using this insight, we construct four state-level instruments, by interacting the FDA cumulative approval of new drugs with the following state-level variables: the fraction of individuals who are older than 60; the fraction of individuals with college or higher degree; the fraction of African-Americans; and the fraction of Hispanics. Since those four variables are included in the vector of controls X, it can be shown that, if the variable FDA approvals is uncorrelated with the error term of equation (2) conditional on X, the four variables we propose as instruments are also uncorrelated with the error term.[8] Since age is the demographic variable that is likely to have the largest effects on how the FDA approvals influence medical expenditure in a state, as a robustness check, we will also report the coefficient estimates obtained by using as instrument only the approvals interacted with the fraction of individuals who are older than 60.

An exogenous shock to health expenditure may have state-level general equilibrium effects. A reduction in savings produced by a rise in health expenditure may decrease consumption in non-health goods and investments in new technologies, both of which may reduce income and hence the saving rate. The use of FDA approvals interacted with the demographic structure of a state enables us to account for both the direct and indirect effects of a shock to health expenses. Our empirical analysis ignores national general equilibrium effects, which may be important if the lack of investment in new technologies at the state level generates negative externalities for other states (Kline and Moretti (2014)).

To perform the IV estimation, we have to characterize how past FDA approvals affect current health expenditure. The simplest choice would be to assume that all past FDA approvals have

---

[8]To see this, denote with Z the variable FDA approval and with W_j, $j = 1, ..., 4$, the four state-level variables mentioned above. Then, by the law of iterated expectations and $W^j \in X$, $j = 1, ..., 4$, if $E[\epsilon Z | X] = 0$, we have,

$$E\left[\epsilon Z W^j\right] = E\left[W^j E\left[\epsilon Z | X\right]\right] = 0.$$

Hence, if $E[\epsilon Z | X] = 0$, our instruments ZW^j, $j = 1, ..., 4$, are exogenous.

the same effect on current total health expenditure. A more realistic approach should take into account that the approval of a new drug has its largest effect on total health expenditure during the years in which it has effective market exclusivity, which is defined as the time between FDA approval and the availability of the first generic version. It should also consider that, when the market exclusivity comes to an end, the effect of the new drug diminishes at some rate. This is the approach we use here.

Let M be the years of market exclusivity, ρ the subsequent rate of decline, under the assumption of a constant rate, and z_t the number of new approvals in period t. Also denote with $Z_t = \sum_{\tau=0}^{M} z_{t-\tau} + \sum_{\tau=1}^{\infty} \rho^{\tau} z_{t-M-\tau}$ the effective cumulative number of approvals in year t and with $W_{t,s}^j$, $j = 1, ..., 4$, the four state-level variables interacted with the approvals. The first stage of the IV approach can then be implemented by estimating the following equation:

$$\bar{c}_{t,s}^h = \delta_0 + \sum_{j=1}^{4} \delta_j Z_t W_{t,s}^j + \beta X_{t,s} + g_t + g_s + \eta_{t,s}, \tag{3}$$

where g_t and g_s are time and state fixed effects. Wang, Liu, and Kesselheim (2015) find that the effective market exclusivity for new molecular entities is on average 14.5 years. We will therefore use $M = 14$ and report the results for $\rho = 10\%$ and $\rho = 15\%$ (we obtain similar results using $\rho = 5\%$ and $\rho = 20\%$). In the second stage, we can then estimate equation (2) with the share of health expenditure, $\bar{c}_{t,s}^h$, instrumented using the estimates from the first stage.

Figure 7 illustrates the number of new molecular entities approved each year during the period 1980-2009 and the corresponding effective cumulative number up to a given year. The figure documents that there was a significant number of drugs that entered the market each year, with significant variation across years. The lowest number was approved in 1980 when 9 new products entered the market, whereas the highest number was recorded in 1997 with 58 new molecular entities.

The results of the first stage regressions are reported in Table 4. As mentioned above, we consider four specifications. In the first two columns, we report the estimates obtained using all instruments with $\rho = 10\%$ and $\rho = 15\%$. In the last two columns, we present the results obtained using only the interaction between approvals and age as instrument. When we use all the instruments and $\rho = 10\%$, the coefficient estimates, which are for 100 approvals, have the expected sign. The cumulative number of approvals interacted with the fraction of old people has a strong and positive effect on the share of medical expenditure in that state. The approval

19

of 100 new molecular entities increases the share of health expenditure in a given state by 0.071 percentage points relative to a state with a one percentage point lower fraction of older individuals. As expected, the effect of education is negative and statistically significant at -0.048. This means that the approval of new molecular entities in states with a larger fraction of individuals with college or higher degree has a smaller effect than in other states. The fraction of African-Americans and Hispanics have a positive and statistically significant effect on the share of medical consumption. The approval of 100 new molecular entities increases the share of health expenditure in a state with a given fraction of Blacks by 0.008 percentage points relative to a state that has a fraction of African-Americans that is one percentage point lower. For Hispanics, this percentage is 0.009. When we estimate the coefficients using a depreciation rate of 15%, the size of the estimates is similar to the case with $\rho = 10\%$ and their statistical significance remains the same. The estimated coefficient on FDA approvals interacted with age is similar but slightly larger when we use this variable as the only instrument. In all specifications, the F-test to evaluate the strength of the instruments is large with the lowest statistic being equal to 111.7 and the highest to 193.5.

The second stage results are reported in Table 5. The effect of the share of health expenditure on the saving rate is similar in the two specifications with all the instruments, with estimates equal to -0.730 when $\rho = 10\%$ and -0.746 when $\rho = 15\%$. These estimates indicate that a one percentage point increase in the share of medical expenditure reduces the saving rate by an amount that is between 0.730 and 0.746 percentage points. The estimated effects of health expenditure are similar, but slightly larger when we use only the interaction with age as the only instrument, with coefficients that are equal to -0.834 when $\rho = 10\%$ and -0.889 when $\rho = 15\%$. These results suggest that there is a small downward bias in our OLS results that may be generated by measurement errors or migration decisions.

There exists a potential threat to our IV strategy. If pharmaceutical companies have their headquarters and manufacturing plants in some states but not in others, the approval of new drugs may have a direct effect on state-level saving rates, even after controlling for health expenditure, through the following channel. The approval of new drugs may increase employment and income relatively more in states where the headquarter and factories of the pharmaceutical company that produces the new drug are located. If the propensity to save depends on income, the rise in employment and income will change the saving rate of that state relative to others introducing a bias in our analysis. To address this issue we proceed in two steps. First, we evaluate how serious

such a threat can be by documenting the share of workers in the pharmaceutical industry in the U.S. and the share of U.S. GDP generated by that industry. The idea behind the first step is that, if those shares are small, it is unlikely that our IV results are affected by the changes in employment and income generated by the pharmaceutical industry. According to the CPS, from 1980 to 2010, the share of workers employed in that industry fluctuated between 0.15% and 0.34%. In 2010, the share of GDP generated by the pharmaceutical industry was 0.63%. In previous years, it was of similar magnitude. These numbers suggest that the impact of this industry on state level employment and income should be small. Thus, it is unlikely that it generates significant biases in our IV estimates. However, even if unlikely, it is still possible that the shares are small at the national level, but the differences across states are sufficiently large to introduce a significant bias in our estimates. To evaluate this possibility, in the second step we use the CPS to measure the share of workers employed by the pharmaceutical industry in each state. We then re-estimate the IV coefficients after having added this variable to the set of controls. If the approval of a new drug has a direct effect on the state-level saving rate through an increase in employment and income generated by the pharmaceutical industry, the state-level share of workers in that industry is an omitted variable and should therefore have a significant effect on the state-level saving rate. The results for the IV regressions are presented in Tables 6 and 7. As expected, the coefficient on the state-level share of workers in the pharmaceutical industry is small and statistically insignificant and the coefficient on the share of medical expenditure barely changes.

We can now use the estimated coefficient on health expenditure and back of the envelope calculations to compute the fraction of the decline in U.S. saving rate that can be explained by the rise in health expenditure. From the beginning of the eighties, the share of medical expenses has increased by about 9 percentage points. If we use the average estimated coefficient on health expenses from our IV specifications, which is equal to 0.8, then a 9 percentage point increase in medical costs translates into a decline in the saving rate of about 7.2 percentage points (0.8×9). From the eighties, the personal saving rate dropped by 8 percentage points. The rise in health expenditure is therefore able to explain about 90 percent of the decline in the U.S. saving rate ($100 \times 7.2/8$).

The main finding of this section is that the rise in medical expenses had a negative and strong effect on the U.S. saving rate. The main limitation of the reduced-form results is that it is difficult to evaluate which mechanisms were responsible for the increase in medical expenditure and for the

corresponding decline in the saving rate. In the next section, we study the importance of different mechanisms by using a model of households decisions.

6 A Model of Saving and Health Decisions

In this section, we first develop a model of household choices in which saving and health decisions are linked. We then describe how the model is estimated. Lastly, we use simulations to evaluate the contribution of different factors to the decline of the saving rate.

6.1 The Model

Consider an economy populated by N cohorts. We will denote by cohort τ all individuals that are born in that year. For notational convenience, we will suppress the dependence on τ, unless it is required. Individuals in each cohort are endowed with a discount factor β, and preferences over non-health consumption c and the bequest they leave to future cohorts B. We assume that the corresponding utility function is strongly separable between those two variables and that the subutilities take the following standard forms (De Nardi, French, and Jones (2016)):

$$u\left(c\right) = \delta + \frac{c^{1-\gamma}}{1-\gamma} \qquad \text{and} \qquad v\left(B\right) = \eta\frac{\left(B+k\right)^{1-\gamma}}{1-\gamma},$$

where η determines the intensity of the bequest motive, k affects the curvature of the bequest function and, hence, whether bequests are luxury goods, and $\delta > 0$ is a utility intercept necessary to guarantee that people gain positive utility from an additional year of life.

In each year, individuals die at a rate that is a function of age, a_t, the existing health technology, θ_t, and the individual's current health status, h_t. The mortality rate function can therefore be written as follows:

$$m_t = m\left(a_t, \theta_t, h_t\right), \tag{4}$$

where the health technology θ_t changes over time following an exogenous process. The shocks that determine whether someone survives in period t, ξ_t, are i.i.d. and drawn from a logistic distribution.

The health status of a person evolves according to a function that depends on age, the health technology, the amount of resources spent on health by the person, c_t^h, the health status in the

22

previous period, h_{t-1}, and the health shocks experienced by the individual in the period, ϵ_t. The health status function takes therefore the following form:

$$h_t = h\left(a_t, \theta_t, c_t^h, h_{t-1}, \epsilon_t\right). \tag{5}$$

In each period t, individuals may suffer two types of health shocks: non-severe health shocks ϵ_t^N, with probability $P_\epsilon^N\left(h_{t-1}, a_t\right)$; and severe health shocks ϵ_t^S, with probability $P_\epsilon^S\left(h_{t-1}, a_t\right)$. Conditional on health status and age, the shocks ϵ_t^N and ϵ_t^S are drawn from independent distributions. Thus, each period a person can experience one of the following combinations of shocks: (i) no health shocks, $\epsilon_t = [0, 0]$; (ii) only non-severe health shocks, $\epsilon_t = \left[\epsilon_t^N, 0\right]$; (iii) only severe health shocks, $\epsilon_t = \left[0, \epsilon_t^S\right]$; and (iv) a combination of non-severe and a severe health shocks, $\epsilon_t = \left[\epsilon_t^N, \epsilon_t^S\right]$.

If a person is hit by a non-severe health shock, such as high blood pressure, she or he has the choice of seeking treatment and paying the corresponding medical expenses $c^{h,N}$. We will denote this decision by I_t^N. However, if the person suffers a severe shock, for instance a heart attack, she or he will always undergo treatment and pay the medical expenses $c^{h,S}$. With this feature of the model we account for the fact that, in most instances, people who experience severe health shocks either seek treatment and incur the corresponding expenses or their survival probability goes to zero – heart attacks, cancer, AIDS are examples of such shocks. Both types of medical expenses are functions of age, the current technology, the lagged health status, and a shock that determines the residual variation conditional on the observables, i.e.

$$c_t^{h,k} = c^{h,k}\left(a_t, \theta_t, h_{t-1}, \eta_t\right), \quad k = N, S. \tag{6}$$

In our model, people can affect their health expenses only by investing in their health status. With this modeling feature, we take into account that, when a health shock hits an individual, doctors choose the most appropriate tests and drugs given the available technology and patients have limited influence on the decision. A recent paper by Currie, Lin, and Zhang (2011) provides evidence in support of this hypothesis. The authors use an audit study to test whether the antibiotic abuse experienced by China in recent years is driven by patients actively asking for those products or by doctors actively prescribing them. The main finding of the paper is that the over-consumption of antibiotics is largely driven by physician behavior, providing support for our model.[9]

---

[9]There is one type of medical expenses over which people have control: preventive care. For the most part,

In the model, people can perfectly predict the evolution of health technology. We impose this assumption instead of the alternative that individuals are surprised by technological innovations, because with this restriction it is more difficult to generate the decline in the saving rate, as people can prepare for the increase in the corresponding health expenses by saving more.

In each period, a person is endowed with an amount of income y_t, which includes earnings, taxes, and transfers from the government. It is allowed to vary across ages and cohorts and it is assumed to evolve deterministically. Households can save an amount b_t using a risk-free asset with a gross return R_t, but cannot borrow (i.e. $b_t \geq 0$).

The health expenses of people older than 64 are partially covered by a public health insurance (Medicare). The health expenses of every other person are partially covered by a private health insurance, whose premium P_t^h is exogenously given and paid by the individual. The part of health expenses not covered by some form of insurance are incurred by the person. The corresponding co-payment rate, which varies by age and cohort, will be denoted by $x_\tau(a_t)$.[10]

The previous discussion implies that people face two sources of risk: health risks and mortality risks. Given this uncertainty, in each period t, individuals make three choices: how much to spend on non-health consumption; how much to save; and, if a person experiences non-severe shocks, whether to seek treatment to improve the current health status and pay the corresponding medical expenses. With the treatment choice, the model accounts for the possible existence of moral hazard in medical decisions: people are more likely to seek treatment when the co-payment rate is low. The previous discussion also indicates that in our model medical expenses and, hence, better health status, are valued by individuals because they reduce mortality risk and, conditional on age, future health expenses.[11]

---

individuals choose how much to spend on this good. But the data suggest that preventive care is only a small fraction of personal health expenditure. For instance, Maciosek et al. (2010) find that in 2006 the cost of preventive care was just 4.1% of personal health care spending, suggesting that our modeling of medical expenditure is a good approximation of actual behaviors.

[10]For employed individuals, part of the health insurance premium is paid by the employer through the employer contributions to health plans. Since NIPA includes this variable as part of income, to add it to the model, we would have to include it in the budget constraint as part of income and as part of expenditure. The contributions would therefore cancel out and have no effect on decisions. To simplify the notation, we have therefore decided not to include them.

[11]We could have allowed the health status to directly affect the individual utility. We decided against this modeling choice because the model is already complicated and we can explain the main patterns in the data without it.

We can now describe the decision process of an individual. Each person chooses how much to consume of the non-health good, how much to save, and whether to seek treatment for non-severe shocks as the solution to the following problem:

$$V_t\left(a_t, h_{t-1}, b_{t-1}\right) = \max_{c_t, b_t, I_t^N} \quad \delta + \frac{c_t^{1-\gamma}}{1-\gamma} + m_t \frac{\eta\left(b_t + k\right)^{1-\gamma}}{1-\gamma} + \left(1 - m_t\right) E_t\left[V_{t+1}\left(a_{t+1}, h_t, b_t\right)\right]$$

$$s.t. \quad c_t + P_t^h + x_t\left(I_t^N c_t^{h,N} + c_t^{h,S}\right) + b_t = y_t + R_t b_{t-1}$$

$$h_t = h\left(a_t, \theta_t, c_t^h, h_{t-1}, \epsilon_{i,t}^N, \epsilon_{i,t}^S\right) \quad \text{and} \quad b_t \geq 0.$$

where $V_t\left(a_t, h_{t-1}, b_{t-1}\right)$ is the value function at time t.

The model we have developed does not account for possible general equilibrium effects. Given that our model is already difficult to estimate and including those effects would require modeling the supply side of the economy, we leave this generalization to future research.[12]

6.2 Estimation

To evaluate which factors have the strongest effect on medical expenditure and the saving rate, we have to assign values to the parameters that characterize the model. Most of the parameters are estimated using data from MEPS, NIPA, CPS, NHIS, and the NDI Mortality Files. The remaining parameters are set equal to values used in the literature.

All the utility parameters are chosen following De Nardi, French, and Jones (2016) and reported in Table 8, except for the intercept δ. In a model in which endogenous health affects the mortality rate, it is essential to have a utility intercept that is sufficiently large to ensure that people derive positive utility from additional years of life (Rosen (1988)). Without it, the model cannot rationalize the positive investment in health observed in the data. In the simulations, we use an intercept equal to 0.0052. This value is obtained by dividing the coefficient employed in Hall and Jones (2007) first by 1,000 to account for differences in units and then by 5 to account for differences in the definition of a period (five years in Hall and Jones (2007) versus one year in our model).

---

[12]Bertrand and Morse (2016) use the CEX and cross-state variation to document that households that are not at the top of the income distribution consume a larger share of their current income when exposed to the income and consumption of households that are at the top of the distribution. They provide evidence that this result is generated mainly by social comparisons with the richest neighbors that induce the non-rich households to increase their expenditure. In our model, we abstract from the possible effect of this mechanism on saving decisions.

The mortality risk function (4) is estimated by matching the mortality data from the NDI Mortality Files to the health data from the NHIS. Using the matched data, we first construct a dummy variable equal to 1 if an individual died in a given year and a variable that describes the individual self-reported health status collected in the NHIS (poor = 5, fair = 4, good = 3, very good = 2, and excellent = 1). We then estimate a logit model in which the death dummy is the dependent variable and the independent variables are age, age squared, self-reported health status, a linear time trend to account for the effect of health technology on the mortality risk, and the interaction terms among those variables.

The most difficult function to estimate is the health status function (5). If we estimate it using OLS, we obtain a counterintuitive but easily explainable negative coefficient on medical expenditure: an increase in medical expenditure reduces the current health status. The reason for the negative estimate is that people are more likely to incur medical expenses when they have severe medical conditions and, hence, their health status is deteriorating. We deal with this reverse causality problem by estimating the health function in two steps. In the first step, we estimate the effect of medical expenditure on health status using an idea proposed by Doyle (2005). We consider the sample of individuals in MEPS who were injured in an accident and estimate the effect of health expenses by using as an instrument whether they had health insurance coverage: all else equal, people with coverage should receive more medical treatment after the accident. The assumption required for our instrument to be valid is that the insurance coverage is independent of the severity of the accident. Since accidents generally cannot be predicted, the choice of insurance coverage should be based on predicable, long-term health issues and not on the outcome of unpredictable events. In the second step, we then estimate the health function for the entire sample by regressing the self-reported health status on age, age squared, dummies equal to one if the respondent experienced severe or non-severe health shocks, time dummies, and medical expenditure, where the coefficient on medical expenditure is constrained to be equal to 0.015, the coefficient estimated in the first step. This coefficient implies that a $1,000$ increase in medical spending raises self reported health by 0.015. To illustrate the economic meaning of this parameter, note that our estimated health status function (5) implies that a non-severe shock reduces self-reported health on average by 0.15. Thus, our estimated coefficient on medical expenditure indicates that people need to spend about $10,000 to offset the negative health consequence of a non-severe shock.

The simulation of the model requires the classification of the shocks in non-severe and severe

and the estimation of the probabilities with which they occur. The MEPS data contain an event file that describes whether a person experienced a particular health condition during the year and the corresponding medical expenditure. We treat each condition as a health shock and we classify it as non-severe or severe on the basis of the medical expenditure the respondent incurred. Specifically, we compute the average per-capita expenditure for each one of the 171 conditions listed by MEPS, rank them based on that variable, and classify a condition as severe if it belongs to the top λ percent. Ranking conditions based on medical expenses is not the only way to classify health shocks. But any choice has some degree of arbitrariness and ranking the shocks based on monetary disbursements is consistent with the common insight that more severe shocks are more expensive to treat.

The parameter λ is estimated by matching the increase in total health expenditure as a share of income between the first and last year of the sample period. Simulating the model for different values of λ is computationally demanding because, for each λ, we have to re-estimate the health status function, the medical expenditure functions (6), and the probabilities for severe and non-severe conditions. We therefore use the following coarse grid search method. We compute the actual and simulated moments for $\lambda = 10\%,\ 20\%, \ldots,\ 80\%,\ 90\%$. We then use as the estimated parameter the λ that generates the lowest distance between them, which in our case is equal to 50 percent. To evaluate the robustness of our findings to the value estimated for λ, we will also report results obtained from a more extreme case in which 70% of conditions are non-severe.

For any given λ, we can easily estimate the probability that a severe condition occurs conditional on age and current health status using MEPS data. The probability of a non-severe shock $P^N(a_t, h_{t-1})$ cannot be estimated using only data, because under our assumptions we observe a non-severe condition only if a person chooses to seek treatment. We deal with this issue using the following method. The probability $P^N(a_t, h_{t-1})$ must be higher than the probability with which a person seeks medical treatment when affected by a non-severe shock $P^T(a_t, h_{t-1})$, since only a subset of people hit by a non-severe condition undergo treatment. Moreover, $P^N(a_t, h_{t-1})$ must be smaller than or equal to 1. Thus, we can write the probability of a non-severe shock as the following weighted average:

$$P^N(a_t, h_{t-1}) = P^T(a_t, h_{t-1}) + w(a_t, h_{t-1})\left(1 - P^T(a_t, h_{t-1})\right).$$

Since $P^T(a_t, h_{t-1})$ is observed in MEPS, the equation indicates that $P^N(a_t, h_{t-1})$ is also observed

if the weight $w(a_t, h_{t-1})$ can be estimated. To estimate $w(a_t, h_{t-1})$, we assume that it does not vary with age and health status and find the parameter w that matches the average treatment rate for non-severe conditions observed in the data. The value of w that best matches our moment is 0.7, which produces an average probability of a non-severe shock $P^N(a_t, h_{t-1})$ equal to 0.91. With the estimated w, we match well the average treatment rate for non-severe conditions, which is 0.69 in the data and 0.70 in the simulations.

The simulation of the model also requires knowledge of the medical expenditure function (6) for severe and non-severe shocks. They are estimated using data from MEPS for people 25 and older. Since the main objective of this section is to replicate the saving rate data measured in NIPA, the medical expenditure data from MEPS we use in the estimation must be comparable to the health expenses reported by NIPA. To make them comparable we have to address the following issues. First, when we simulate the model, we consider people that are 25 or older, whereas NIPA includes medical expenses for the entire household sector. Second, to estimate the medical expenditure function separately for non-severe and severe shocks, we must use the variable that reports health expenditure for every health condition separately. However, not all medical expenditures in MEPS are related to a condition. Without adjustment, our measure would therefore underestimate health expenses in NIPA. Lastly, MEPS data do not include some of the categories measured in NIPA data (Sing et al. (2006)). To account for these missing components, we adjust the health expenses for people 25 and older that are associated with a condition as follows. We first increase this measure by the fraction of expenses that are incurred by people younger than 25 (25 percent); we further increase the measure we obtain by the fraction that is not associated with a condition (26 percent); lastly, we increase the last measure by the share of medical expenses that are not included in MEPS but are included in NIPA (49 percent).

After the adjustments, the medical expenditure functions are estimated by regressing medical expenditure for non-severe and, separately, severe health shocks on a cubic polynomial in age, lagged self-reported health status, their interactions, and year dummies to account for technological change.[13] Because the MEPS data only covers the period 1996-2014, the year fixed effects can only be estimated for those years. But, since our objective is to simulate the saving rate from 1980 to 2010, we need year fixed effects from 1905, the year in which people aged 100 in

---

[13]It will be clear when we discuss Figure 10 that medical expenditure for severe conditions requires a cubic polynomial in age.

1980 were 25 years old, to 2085, the year in which individuals aged 25 in 2010 will be 100 years old. To address this issue, we extrapolate the year fixed effects for the period 1905-1995 using a logarithmic function to prevent negative medical expenses and using a linear function for the years 2015-2085 to avoid exponential growth. The estimated coefficients of the mortality, health status, and expenditure functions are reported in the Appendix.

We estimate the individual income process using after-tax income from the CPS for the sample period 1980-2014. For people younger than 65, we assume that the income process depends on a cohort fixed effect and a second order polynomial in age. For older individuals, the income process varies only across cohorts. The cohort fixed effects that cannot be estimated using our sample are extrapolated.

Since our preference parameters are taken from De Nardi, French, and Jones (2016), we follow their analysis and assume that the real interest rate is equal to 4 percent. The co-payment rate for medical expenses is computed for each age and cohort as the ratio between out-of-pocket medical expenses minus the health insurance premium paid by consumers and total medical expenses as reported by the NHEA, sample period 1960-2015. For years before 1960 and after 2015, the rate is extrapolated. Finally, the health insurance premium is set equal to zero for people older than 64. For younger individuals, for the period 1980-2011, it is computed as total private health insurance payments net of employer contributions as measured by NIPA divided by the population that is not covered by Medicare or Medicaid. For other years we extrapolate the insurance premium. The model is simulated using standard backward induction techniques. Individuals in cohort τ start making decision at age 25. We set their number equal to the individuals alive at age 25 in the SEER Data.

6.3 Results

We now use the estimated model to simulate the individual decisions for the period 1980-2010. We then aggregate them to generate total medical expenditure and the saving rate of the household sector. Each individual is simulated until her or his death. We assume that, if a person lives until age 100, she or he will die with probability one at that age.

We start by documenting the ability of our model to match the main features of the data. In Figure 8, we report total and out-of-pocket medical expenditure as observed in the NHEA and as generated by the model, where out-of-pocket expenses include both the co-payment

$x_t \left(I_t^N c_t^{h,N} + c_t^{h,S} \right)$ and the part of the health insurance premium paid by the individual P_t^h. Out-of-pocket expenses are approximately constant during the sample period in both the model and data, with the simulated variable being slightly lower than in the data. The model also matches well the increasing pattern of total medical expenditure observed in the data, with an increase of about 9 percentage points from 1980 to 2010. In the first part of the sample period, the simulated health expenses are marginally below what is observed in the data, but overall we capture well the general pattern.

Figure 9 plots the evolution of medical expenditure for severe and non-severe health shocks in the model and data for the period for which we observe these variables in the data, 1996-2010. The model can account for their increasing pattern and for the fact that medical expenses are about twice as large for severe shocks than for all other shocks. Figure 10 reports medical expenses for severe and non-severe shocks over the life-cycle. We match well the general trend for both types of shocks. The cost of treating non-severe shocks is increasing over the life-cycle in both the simulations and data. The medical expenses for severe conditions are about constant until age 45, increase steadily until age 65, and then remain approximately constant for the rest of a person's life. The cubic polynomial in age helps us explain this pattern in the data. But, even with a cubic polynomial, the actual medical cost for severe shocks rises slightly faster from age 45 to age 65 than in our model, which explains why we slightly underestimate total and out-of-pocket medical expenses over the sample period.

In Figure 11, we report the probability that a person is affected by a severe shock for the years for which we have data, 1996-2010. We also report the probability that an individual is treated for a non-severe shock which, since people can choose whether to seek treatment for this type of conditions, is the only variable we observe. We match accurately these two variables, which are approximately constant for the period considered. Figure 12 documents that the average health status in the data remains approximately constant during our sample period, a pattern that we capture well with our model.

Figure 13 reports the evolution of observed and simulated average income for our sample period and indicates that the model can replicate the increasing pattern observed in the data. In our model, all cohorts are endowed with the same income process over the life-cycle, except that they have a different cohort fixed effect that shifts their process up or down. The estimated fixed effects for different cohorts are depicted in Figure 14, where the x-axis reports the year in

which a cohort was 25 years of age. The figure documents that income improved at a similar rate for all cohorts that were 25 years of age between 1925 and 1973. CPS data indicate that the rise is driven by the increase in income that individuals of most ages experienced before 2000. Income then remained approximately constant for cohorts that were 25 in the years 1975-1995, as an increasing number of cohorts entered the period 2000-2010 when income was either stagnant or declining (CPS). Finally, the income fixed effects decline for all the remaining cohorts, driven by the decreasing income of young individuals (25-34) for the period 2000-2010 and the declining income of middle aged individuals (35-45) during the great recession, as documented by the CPS.

Figure 15 documents that our estimated mortality function is able to approximate well the average age structure observed in the data over our sample period, with the expected declining pattern that becomes steeper after age 40. In Figure 16, we report the evolution of the co-payment rate for medical insurance, which by construction is identical in the data and simulations, and the share of people seeking treatment for non-severe shocks in the simulations for the entire sample period and in the data for the period 1996-2010, the years for which we observe this variable. Our model generates a treatment rate that increases only slightly during the period 1995-2010, when the co-payment rate declines only marginally, which is consistent with the pattern observed in the data. But during the period 1986-1994, when the co-payment rate experienced a significant decline, in our model the share of people seeking treatment increases from about 43 percent to about 70 percent, as should be expected. Lower co-payment rates therefore create moral hazard in our simulations, encouraging people to increase health expenditure without reducing the consumption of non-health goods.

Lastly, in Figure 17 we report the dynamics of the observed and simulated saving rate. To make the simulated saving rate comparable to the NIPA saving rate, we include the employer contributions to health plans in the computation of the simulated variable. Our model can replicate the steep decline in the saving rate from 1980 to 2007, the year in which the great recession started. During this period, the saving rate drops by about 8 percentage points in both the data and simulations. Our model of saving and health decisions can therefore explain nearly the entire decline in the personal saving rate. During the period 2008-2010, the increase observed in the data is not matched by our model. But this is to be expected if the increase in the data was generated by the uncertainty introduced by the great recession, which we do not model.

Overall, the simulated data match well the main features of the actual data. This is a note-

worthy achievement since we do not target directly many of those patterns in the estimation. In particular, we believe that the ability of our model to produce the large decline in the saving rate is an important accomplishment, because our estimation method is not designed to match the saving decisions.

In our model, there are three main mechanisms that can generate the decline in the saving rate: progress in health technology and the corresponding increase in health expenses; reductions in the co-payment rate and the corresponding increase in the share of people seeking treatment; and improvements in income with the associated changes in treatment rates. Using counterfactuals, we now document the importance of each mechanism in explaining the decline in the saving rate.

In Figure 18, we report the simulated saving rate obtained by holding constant technological progress and the corresponding growth in medical expenditure. In our model, technological innovation affects health expenditure and, hence, the saving rate in three ways. Conditional on age and health status, as illustrated by the health expenditure function (6), technological progress θ_t raises medical costs for each person. Technological advancements also improve current health status (equation (5)), which in turn reduces health expenditure. Lastly, progress in health technology improves life expectancy through the mortality risk function (4), generally raising medical costs by increasing the number of people who need medical attention.

To evaluate these effects on the dynamics of the saving rate, we first fix the health technology at the 1980 level in the health expenditure function (6), but allow the health status and mortality rate to vary freely. The corresponding counterfactual saving rate indicates that the effect of technological progress on expenditure can explain about 50 percent of its decline, since without technological progress the saving rate decreases by 4 percentage points as opposed to the 8 percentage points observed in the base case. The counterfactual also indicates that technological progress plays a major role in the decline of the saving rate during the period 1995-2010, as the saving rate stays constant around 4% in these years when we remove its effect on medical expenditure. We then fix the health technology at the 1980 level also in the health status function (5) and mortality risk function (4). The counterfactual saving rate has similar features to the one observed in the first experiment. The only difference is that the saving rate shifts down by about 1 percentage point throughout the period due to an increase in health expenditure of older individuals.[14]

---

[14]Since technological progress is identified in the data by time dummies, we cannot distinguish it from other

The third experiment reveals that the steep decline in co-payment rates experienced by the U.S. economy from 1986 to 1994 (Figure 16) explains most of the drop in saving rate during this period. As we document in Figure 19, when we hold constant the co-payment rate at the 1980 level in addition to fixing the health technology, the saving rate remains approximately constant from 1986 to 2010, with a small decline of about 1 percentage point during the period 1986-1994. Moreover, in the experiment, from 1980 to 2010 the saving rate declines by only 2 percentage points, indicating that technological progress and the evolution of co-payment rates account for about 75 percent of the total decline (6 out of the 8 percentage points).

The rest of the decline, particularly the reduction from 1980 to 1985, is generated by improvements in the income process. When we fix health technology, co-payment rates, and the income process at the 1980 level, the saving rate in Figure 19 is approximately constant around 6 percent.

The actual and simulated data presented in this section enable us to understand why in our model technological progress is the main factor behind the decline in the saving rate for the period 1995-2010, the co-payment rate is the leading mechanism for the period 1986-1994, and income for the years 1980-1985. As shown in Figure 8, total health expenditure increased by about 5 percentage points during the period 1995-2010, whereas out-of-pocket expenses remained approximately constant. Figure 16 documents that in those years the co-payment rate declined only marginally and, hence, the treatment rate remained approximately constant. It cannot therefore explain the increase in medical expenditure during this period. Lastly, as we document in Figure 14, cohorts that entered the economy in this period had declining income processes, which implies that income cannot explain the increase in medical expenditure. The rise in health costs in the years 1995-2010 can therefore only be explained by technological progress.

An important question is: Why does technological progress and the corresponding increase in health expenditure reduces the saving rate? In our model, people can perfectly predict advancements in health technology. If they choose, individuals can increase their saving rate to insure themselves against higher future health expenses. Our simulations indicate that in our model this is not the case. The explanation for this result hinges on the NIPA definition of the saving rate, which is the one we use in our simulations. The saving rate is computed as household income minus household non-health consumption minus total health expenditure, everything divided by household income, where total health expenditure includes both out-of-pocket expenses and med-

---

time-varying variables we do not observe, such as inflation in prices of medical products or increasing inefficiencies.

ical expenses paid by some form of health insurance. During the years 1995-2010, total medical expenses doubled both in the data and in our simulations. But out-of-pocket expenses, the part of medical expenditure that affects household decisions, remained approximately constant between 3 and 4 percent. Consequently, households had limited incentives to reduce their consumption as a share of income and increase their saving rate. This is what we document in Figure 20, where simulated consumption as a share of income declines only marginally over this period. The figure provides evidence that also the non-health consumption's share of income, as measured by NIPA, remained relatively stable during our sample period, with the exception of the recession of the early nineties and the great recession.

The data can also be used to illustrate why the decline in the co-payment rate is the main factor behind the drop in the saving rate for the years 1986-1994. In Figure 21, we report the simulated share of people who seek treatment for non-severe conditions over the life-cycle. In our model, individuals start to undergo treatment in case of non-severe shocks around the age of 35 and then the share grows until the age of 65, after which it stays approximately constant at 0.85. Since medical treatment is a normal good in our model, this implies that a rise in income increases the treatment rate for individuals who are 35 and older. The estimated income fixed effects that we report in Figure 14 indicate that many cohorts that were between the ages of 35 and 65 in the period 1986-1994 had similar income processes. Changes in income can therefore explain only a small part of the increase in treatment rates. Moreover, the increase in treatment rates generated by declining co-payments is sufficient to generate the entire rise in medical expenditure during this period. It is therefore the major factor in our model behind the decline in the saving rate during these years.

It is relatively simple to explain why in our model a reduction in co-payment rates generates the decline in the saving rate. With lower co-payments, more people seek treatment for non-severe conditions to improve their health status. As long as health status is not perfectly substitutable with non-health consumption, which is the case in our model, the higher health expenditure generated by higher treatment rates produces a decline in the saving rate.

During the period 1980-1985 the co-payment rate remained approximately constant. But, according to Figure 14, many cohorts that were between 35 and 65 years old in this period had increasing income processes. In our model, this improvement in income generates an increase in treatment rates from about 30 to 40 percent. Since this mechanism can explain on its own the

34

increase in total medical expenses during the years 1980-1985, income is the main factor underlying the decline in the saving rate for this period.

Our results indicate that most of the drop in the saving rate was generated by the increase in medical expenses paid by some form of insurance. To corroborate them, in Figure 22, we report the actual and simulated saving rate computed by including in total expenditure only out-of-pocket health expenses. Both variables are much more stable than the saving rate computed using the NIPA definition, with a small decline of less than 2 percentage points. This finding confirms our previous conclusion that people had limited incentives to reduce their non-health consumption because their budget constraint was barely affected by the rise in medical expenditure.

We conclude this section by providing evidence that our results are not driven by using 50 percent as the fraction of health conditions that are severe. In Figure 23, we report the simulated saving rate and the share of medical expenditure for the base case and a scenario in which just 30 percent of conditions are severe. Total medical expenditure is slightly higher for most of the sample period when 30 percent of conditions are severe. As a consequence, the saving rate is slightly lower in that case. But the differences are small. At first sight, this result may appear counterintuitive. But an analysis of the simulated data explain why there are only small differences between the two cases. All else equal, a reduction in the share of severe conditions makes people more responsive to increases in medical expenditure, because they can choose whether to seek treatment in more instances. This effect should reduce the responsiveness of the saving rate to health cost increases. This is not, however, the only change that takes place. When the share of severe conditions declines to 30 percent, the estimated costs of treating severe as well as non-severe conditions increase, because the bottom 20 percent of conditions that were initially considered severe are now treated as non-severe. This change goes in the opposite direction of the first effect by generating larger increases in medical expenditure. Lastly, the reduction in the share of severe conditions to 30 percent reduces the estimated probability that a person is hit by a severe health shock and increases the corresponding probability for a non-severe shock. As a consequence, the share of people that seek treatment for a non-severe condition may rise even if the cost is higher just because more people are hit by that type of shock, which goes against the first effect. The overall impact of reducing the share of severe shocks can therefore increase or decrease medical expenditure and the saving rate depending on which effects dominate.

6.4 How Was the Increase in Health Expenditure Paid For?

The findings of this section indicate that the decline in the saving rate was mostly driven by the increase in total health expenditure. Our results also reveal that individuals did not pay directly for this increase, since out-of-pocket expenses remained approximately constant during the sample period. Then, how was the rise in medical expenditure paid for?

To answer this question, it is helpful to divide total health expenditure into two parts: public health expenditure, which is covered by some form of public health insurance (mainly Medicare or Medicaid); and private health expenditure, which is covered by some form of private health insurance. As documented in Figure 24, both private and public health expenses contributed to the increase in total health expenditure, with private health expenditure doubling from 4.9% in 1960 to 10.5% in 2009 and public health expenditure experiencing an increase of about 7 percentage points.

We start by providing evidence on how the rise in private health expenditure was funded. Since we have already documented that consumers did not finance part of the growth through an increase in their out-of-pocket health expenses, its rise may have been financed in two possible ways. Health insurance firms may have paid for the increase through a reduction in their underwriting gains, defined as the difference between the premiums received and the benefits paid. Or, employers may have funded a portion of the rise with an increase in their contributions to the employees' health plans.

With regard to health insurance firms, it has been documented that in the past several decades private insurers have experienced a sequence of underwriting cycles, defined as three consecutive years of underwriting gains followed by three consecutive years of losses.[15] This pattern suggests that the increase in private health expenditure was not funded by a reduction in underwriting gains.

Figure 25 reports the amount of private health expenditure that was paid by employers as insurance premiums divided by the household-sector disposable income. This variable increased at a steep rate, going from 0.9% in 1960 to 5.2% in 2009. As reported in Figure 26, the premiums paid by consumers also grew as a share of income throughout the period. But, during those years,

---

[15]These cycles have been documented, for instance, in Reed, Robert, and Maule (1989), Gabel et al. (1991), and in the executive report prepared in 2003 by the consulting firm Milliman USA (www.aha.org/aha/content/2003/pdf/MillimanReport030410.pdf).

consumers enjoyed a reduction in co-payments that offset the increase in premiums. These results suggest that the rise in private health expenditure was financed almost entirely by an increase in the share of employer contributions to health plans.

There are two ways the corporate sector could have dealt with the increase in employer contributions: by reducing their profits or by transferring the increase to workers through a reduction in their earning. Figure 27 uses NIPA data to provide suggestive evidence that the corporate sector transferred, at least partially, the rise in employer contributions to workers through a reduction in wages and salaries. The figure reports the employer contributions to health plans, income earned by employees, and the sum of those two variables, all as a share of disposable income. The figure indicates that at the time the share of employer contributions increased, the share of earned income declined significantly. The household sector may have therefore funded part or all of the increase in private health expenditure. Since in our model the income process is estimated directly from the data, our simulations and the corresponding decline in the saving rate account for this possible way of funding the rise in private health expenditure.

With regard to the increase in public health expenditure, there are four main alternatives the government may have used to finance its expansion: (i) by reducing the expenditure on other public goods, (ii) by raising taxes on the household sector, (iii) by raising taxes on production and imports, on corporate income, and/or taxes from the rest of the world, or (iv) by increasing its debt.

Figure 28 describes the evolution of expenditure on the four public goods with the highest outlays as a share of disposable income using NIPA data. It provides evidence that in the period 1980-2010, while public health expenses increased from 4.8% to 9.8% of disposable income, the share of military outlays declined from 7% to 5.7% and economic affairs from 3.7% to 2.7%. These patterns suggest that about half of the increase in public health expenditure was financed by a reduction in expenditure on other government items. Our model does not account for the existence of these public goods. But if individuals have preferences that are separable between non-health consumption and these public goods, the main results obtained by simulating the model are valid even if the increase in public health expenditure came at the cost of lower public consumption.

Figure 29 documents that, because taxes on the household sector as a share of disposable income remained approximately constant during our sample period, they were not used to finance the increase in public health expenditure. In the same figure, we report the evolution of total

government revenues and total government expenditure. There are three patterns that are worth discussing. First, total government revenues as a fraction of disposable income fluctuated around a constant rate of 40%. As a consequence the rise in health expenditure was not funded by an increase in other taxes. Second, during the period considered, government expenditure as a fraction of income increased by 12 percentage points to 48.6%. Lastly, from 1970 government expenditure has been above government revenues in all years except for the period 1997-2001. These last two patterns suggest that the remaining half of the rise in public medical expenses was financed by issuing debt.

In the model, we implicitly assume that, if the government chooses to fund the increase in public health expenditure by issuing debt, people believe that the government will not raise taxes during their lifetime to repay it. Government debt has therefore no effect on individual decisions. Given that since 1960 personal taxes as a share of income have fluctuated around the same level of 14 percent, we believe this is a reasonable assumption. But, to evaluate the robustness of our results to different beliefs, we have also simulated the model under the beliefs that the government will raise taxes to repay the existing debt twenty years (five political terms) after the debt was issued. With the new beliefs, the simulated saving rate presented in Figure 30 is about 1 percent higher than the rate obtained under the beliefs that the debt will not be repaid by the current generation. This is to be expected because in the new environment individuals react to higher medical expenses by saving a larger fraction of their income in expectation of the higher taxes they will have to pay in the future. However, the increase in saving rate relative to the base case is constant throughout the sample period. Since medical expenses rise at about the same rate starting from the sixties, people increase their saving rate by a similar percentage in each year. Thus, our main finding that medical expenditure can explain most of the decline in the saving rate is still valid under the new beliefs.

7 Conclusions

In this paper, we study the relationship between the evolution of the U.S. personal saving rate and changes in health expenditure. We find that a 1 percentage point increase in health expenditure produces a decline in the saving rate of between 0.73 and 0.89 percentage points. This finding combined with back of the envelope calculations implies that the increase in health expenditure

can explain about 90% of the decline in the saving rate. We provide this evidence using first an accounting exercise and then cross-state variation in saving rates and health expenditures combined with FDA approvals of new drugs.

To evaluate which mechanisms generate the increase in health expenditure and the corresponding decline in the saving rate, we develop and estimate a model of consumption, saving, and health decisions. We find that progress in health technology explains about 50% of the decline in the saving rate and that this mechanism plays a major role during the period 1995-2010. The decline in co-payment rates accounts for about 25% of the reduction in the saving rate. This mechanism is particularly important in explaining the drop in the years 1986-1994. The last 25% is explained by the rise in income and the corresponding increase in treatment rates. This factor is behind most of the increase experienced by the U.S. economy during the period 1980-1985.

The results documented in this paper present a difficult dilemma for policy makers. They indicate that, if the U.S. intends to increase the saving rate of the household sector, the growth rate of health expenditure has to decline. But the reduction in medical costs may have negative consequences for household welfare and innovation as discussed in the introduction. The U.S. will have to choose on which type of capital to invest and the optimal choice may be to invest in health capital. If this is the case, the U.S. economy will have to cope with low personal saving rates for the foreseeable future.

References

Acemoglu, Daron, Amy Finkelstein, and Matthew J. Notowidigdo. 2013. "Income and Health Spending: Evidence from Oil Price Shocks." *The Review of Economics and Statistics* 95 (4): 1079–1095.

Acemoglu, Daron, and Joshua Linn. 2004. "Market Size in Innovation: Theory and Evidence from the Pharmaceutical Industry." *The Quarterly Journal of Economics* 119 (3): 1049–1090 (August).

Asdrubali, Pierfederico, Bent E Sorensen, and Oved Yosha. 1996. "Channels of Interstate Risk Sharing: United States 1963-1990." *The Quarterly Journal of Economics* 111 (4): 1081–1110 (November).

Attanasio, Orazio P. 1998. "Cohort Analysis of Saving Behavior by U.S. Households." *Journal of Human Resources* 33 (3): 575–609.

Bertrand, Marianne, and Adair Morse. 2016. "Trickle-Down Consumption." *The Review of Economics and Statistics* 98 (5): 863–879.

Browning, Martin, and Annamaria Lusardi. 1996. "Household Saving: Micro Theories and Micro Facts." *Journal of Economic Literature* 34 (4): 1797–1855.

Currie, Janet, Wanchuan Lin, and Wei Zhang. 2011. "Patient knowledge and antibiotic abuse: Evidence from an audit study in China." *Journal of Health Economics* 30 (5): 933 – 949.

De Nardi, Mariacristina, Eric French, and John B. Jones. 2010. "Why Do the Elderly Save? The Role of Medical Expenses." *Journal of Political Economy* 118 (1): 39–75.

De Nardi, Mariacristina, Eric French, and John Bailey Jones. 2016. "Medicaid Insurance in Old Age." *American Economic Review* 106 (11): 3480–3520 (November).

Doyle, Joseph J. 2005. "Health Insurance, Treatment and Outcomes: Using Auto Accidents as Health Shocks." *The Review of Economics and Statistics* 87 (2): 256–270.

Gabel, Jon, Roger Formisano, Barbara Lohr, and Steven DiCarlo. 1991. "Tracing the cycle of health insurance." *Health Affairs* 10 (4): 48–61.

Gale, William G., and John Sabelhaus. 1999. "Perspectives on the Household Saving Rate." *Brookings Papers on Economic Activity* 1999 (1): 181–224.

Garner, Thesia I, George Janini, William Passero, Laura Paszkiewicz, and Mark Vendemia. 2006. "The CE and the PCE: A Comparison." *Monthly Labor Review*, September, 20–46.

Gokhale, Jagadeesh, Laurence J. Kotlikoff, and John Sabelhaus. 1996. "Understanding the Postwar Decline in U.S. Saving: A Cohort Analysis." *Brookings Papers on Economic Activity* 1996 (1): 315–407.

Guidolin, Massimo, and Elizabeth A. La Jeunesse. 2007. "The Decline in the U.S. Personal Saving RateL Is It Real or Is It a Puzzle?" *Federal Reserve Bank of St. Louis Review* 89 (6): 491–514.

Hall, Robert E, and Charles I Jones. 2007. "The Value of Life and the Rise in Health Spending." *The Quarterly Journal of Economics* 122 (1): 39–72 (02).

Heathcote, Jonathan, Fabrizio Perri, and Giovanni L. Violante. 2010. "Unequal We Stand: An Empirical Analysis of Economic Inequality in the United States: 1967-2006." *Review of Economic Dynamics* 13 (1): 15–51 (January).

Kline, Patrick, and Enrico Moretti. 2014. "Local Economic Development, Agglomeration Economies, and the Big Push: 100 Years of Evidence from the Tennessee Valley Authority *." *The Quarterly Journal of Economics* 129 (1): 275–331.

Lichtenberg, Frank R., and Suchin Virabhak. 2007. "Pharmaceutical-embodied technical progress, longevity, and quality of life: drugs as 'Equipment for Your Health'." *Managerial and Decision Economics* 28 (4-5): 371–392.

Lusardi, Annamaria, Jonathan Skinner, and Steven Venti. 2001. "Savings puzzles and saving policies in the United States." *Oxford Review of Economic Policy* 17 (1): 95–115.

Maciosek, Michael V., Coffield Ashley B., Flottemesch Thomas J., Edwards Nichol M., and Solberg Leif I. 2010. "Greater use of preventive services in US health care could save lives at little or no cost." *Health Affairs* 29 (9): 16561660.

Parker, Jonathan. 1999. "Spendthrift in America? On Two Decades of Decline in the U.S. Saving Rate." *NBER Macroeconomics Annual* 14:317–370.

Reed, James L., Stanley A. Robert, and Robert G. Maule. 1989. "Health Profit/Loss Cycles Hold on the Market." *The National Underwriter*, vol. 1.

Rosen, Sherwin. 1988. "The value of changes in life expectancy." *Journal of Risk and Uncertainty* 1 (3): 285–304 (Sep).

Scholz, John Karl, Ananth Seshadri, and Surachai Khitatrakun. 2006. "Are Americans Saving Optimally for Retirement?" *Journal of Political Economy* 114 (4): 607–643.

Sensenig, Arthur, and Ernest Wilcox. 2001. "National Health Accounts/National Income and Product Accounts Reconciliation - Hospital Care and Physician Services." In *Medical Care Output and Productivity*, edited by David M. Cutler and Ernst R. Berndt. University of Chicago Press.

Sing, Merrile, Jessica S. Banthin, Thomas M. Selden, Cathy A. Cowan, and Sean P. Keehan. 2006. "Reconciling Medical Expenditure Estimates from the MEPS and NHEA, 2002." *Health Care Financing Review* 28 (1): 25–40.

Slesnick, Daniel T. 2001. *Consumption and Social Welfare: Living Standards and Their Distribution in the United States*. Cambridge University Press.

Summers, Lawrence, and Chris Carroll. 1987. "Why Is U.S. National Saving So Low?" *Brookings Papers on Economic Activity* 18 (2): 607–642.

Wang, Bo, Jun Liu, and Aaron S. Kesselheim. 2015. "Variations in time of market exclusivity among top-selling prescription drugs in the united states." *JAMA Internal Medicine* 175 (4): 635–637.

Zhou, Xia. 2010. "Essays on U.S. State-level Financial Wealth Data and Consumption Data." *Johns Hopkins University Dissertation*.

Figure 1: Personal Saving Rate from NIPA and the Share of Health Expenditure from NHEA
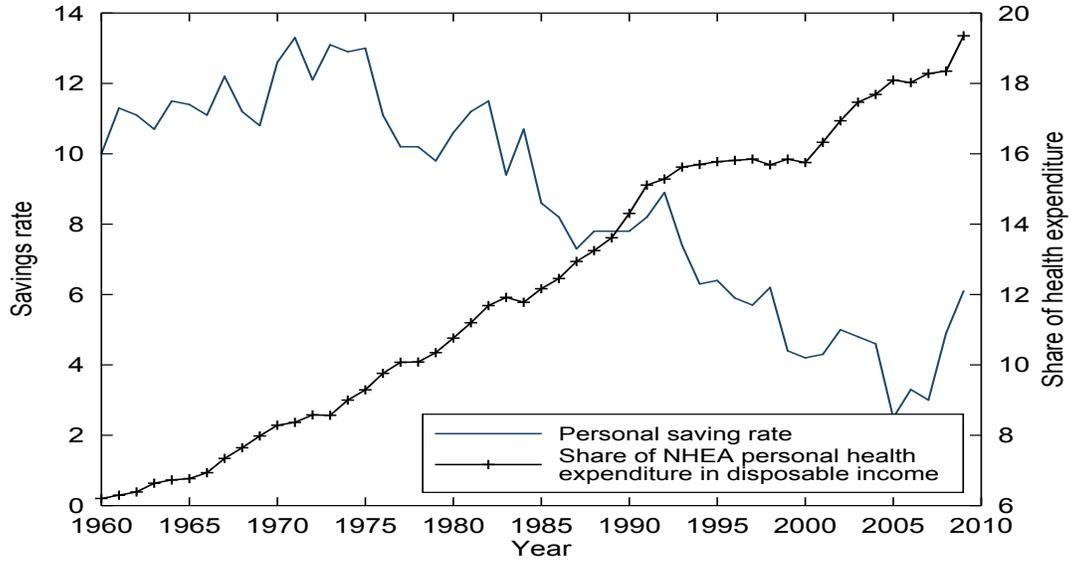
Figure 2: Health Expenditure and Social Security Benefits as a Percentage of Income, NIPA
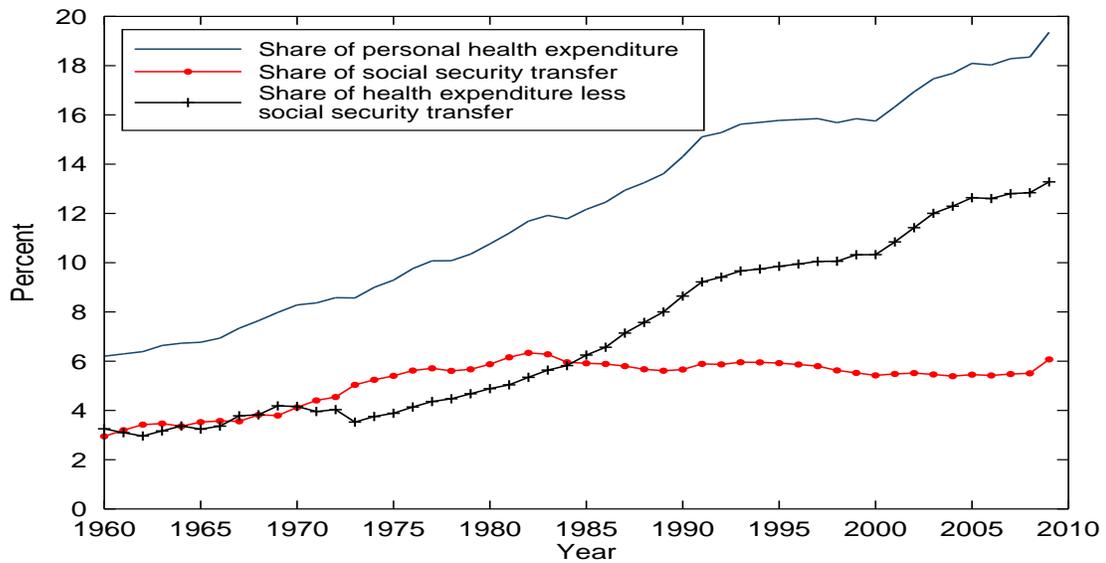
Figure 3: Saving Rates and Health Expenditure Net of Social Security Benefits, NIPA
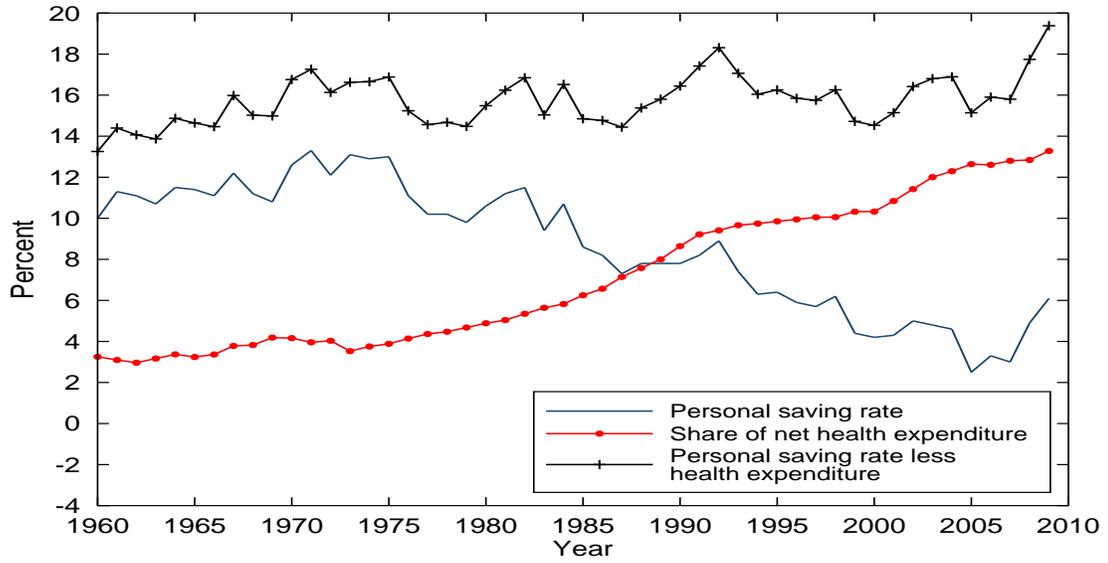
Figure 4: Ratio Between Constructed and NIPA Household Expenditure
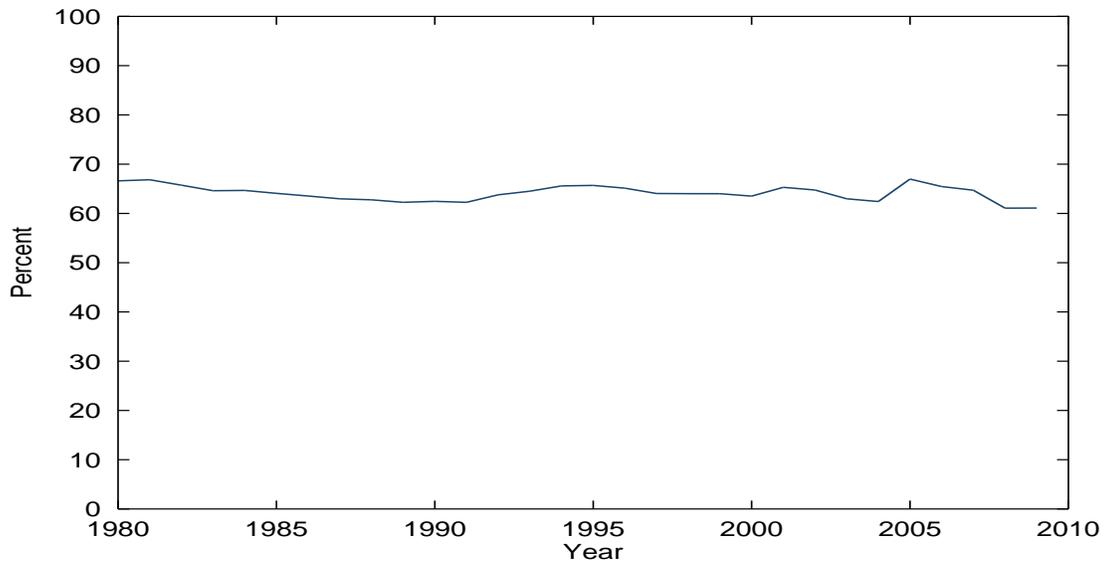
Figure 5: Changes for Constructed and NIPA Aggregate Saving Rates
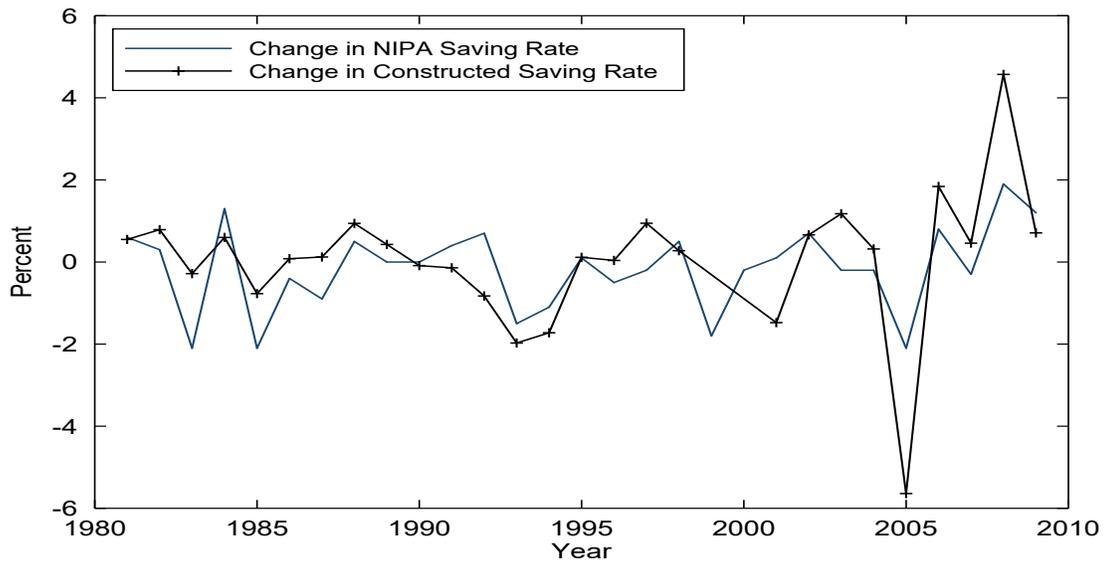
Figure 6: State-level Saving Rates and Health Expenditure

$$y = 0.504 - 0.728\ x + e$$
$$(0.008)\ \ (0.051)$$

(a) Levels

$$y = 0.005 - 1.416\ x + e$$
$$(0.001)\ \ (0.216)$$
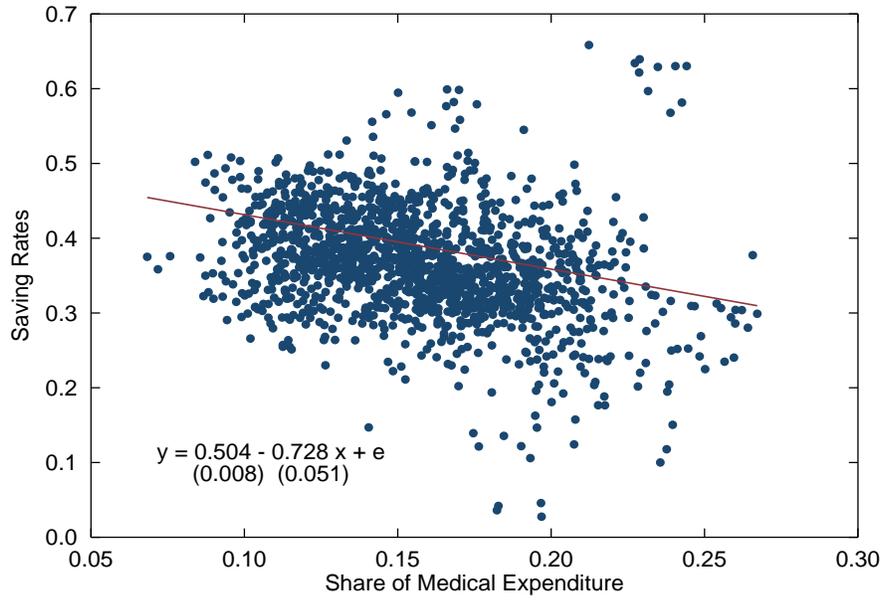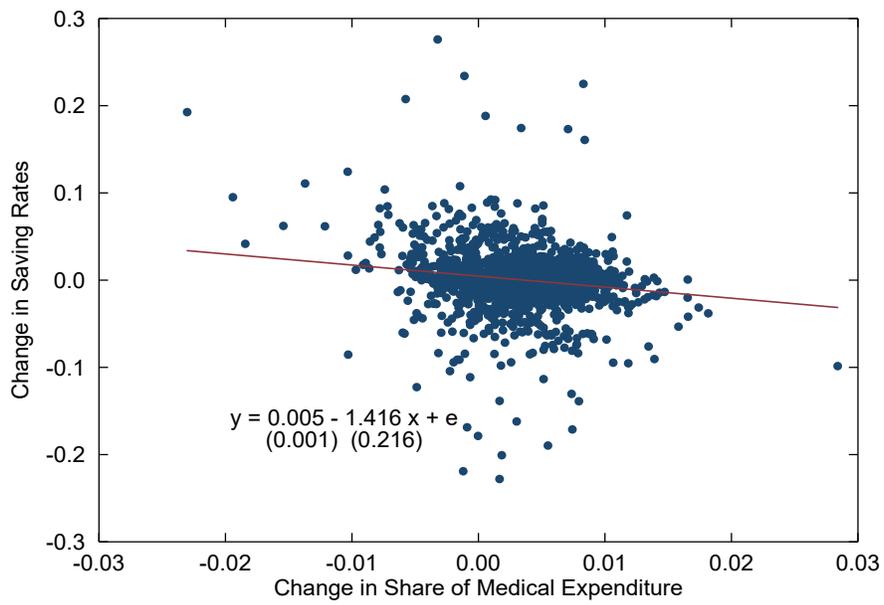
(b) Changes

Figure 7: FDA Approvals and Cumulative Approvals of New Molecular Entities
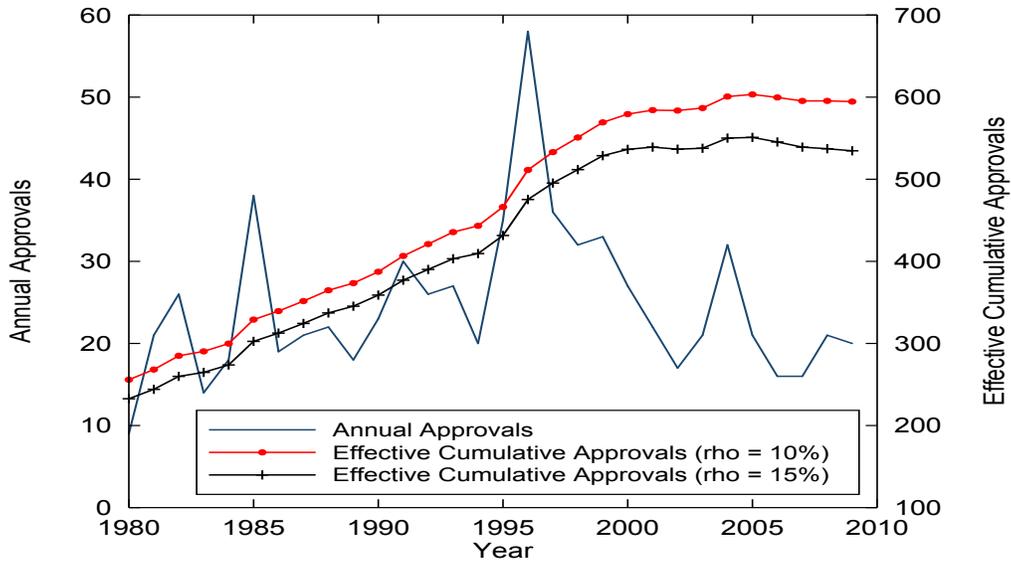
Figure 8: Simulated and Actual Data, Medical Expenditure
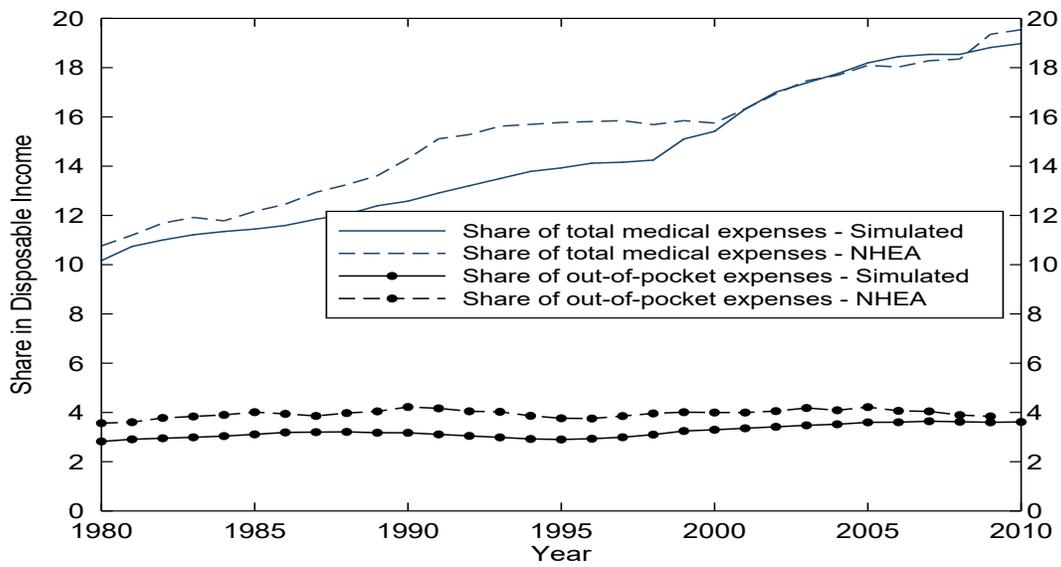
Figure 9: Simulated and Actual Data, Medical Expenditure by Type of Shock
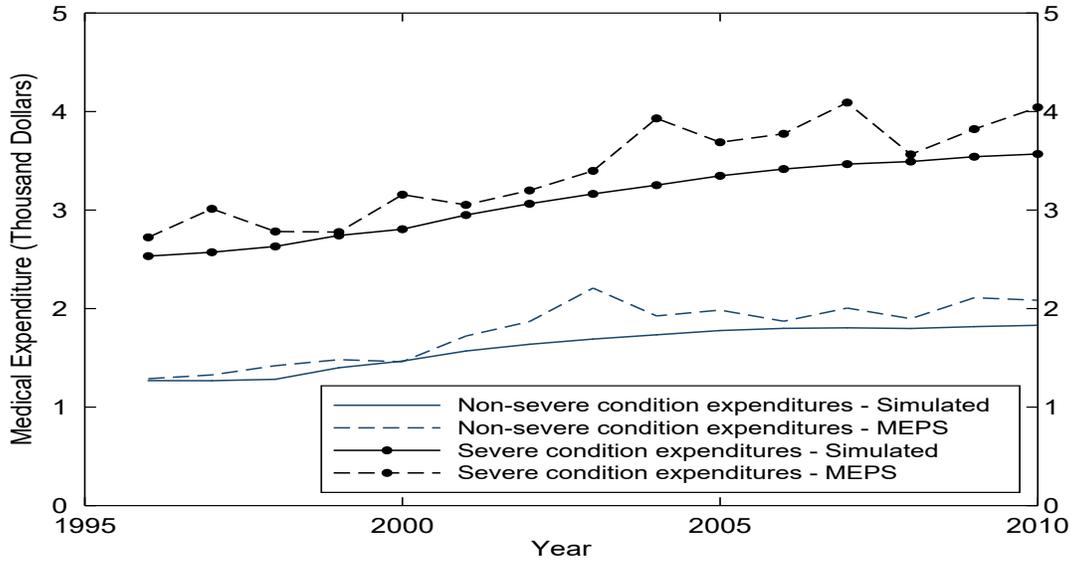
Figure 10: Simulated and Actual Data, Medical Expenditure Over the Life-cycle
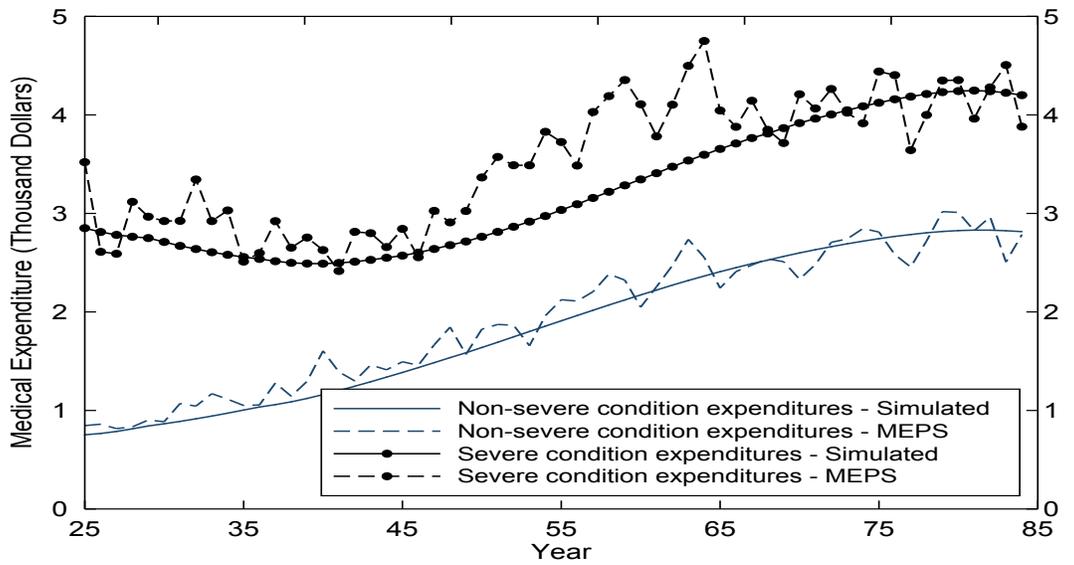
Figure 11: Simulated and Actual Data, Probability of Severe and Non-severe Shocks
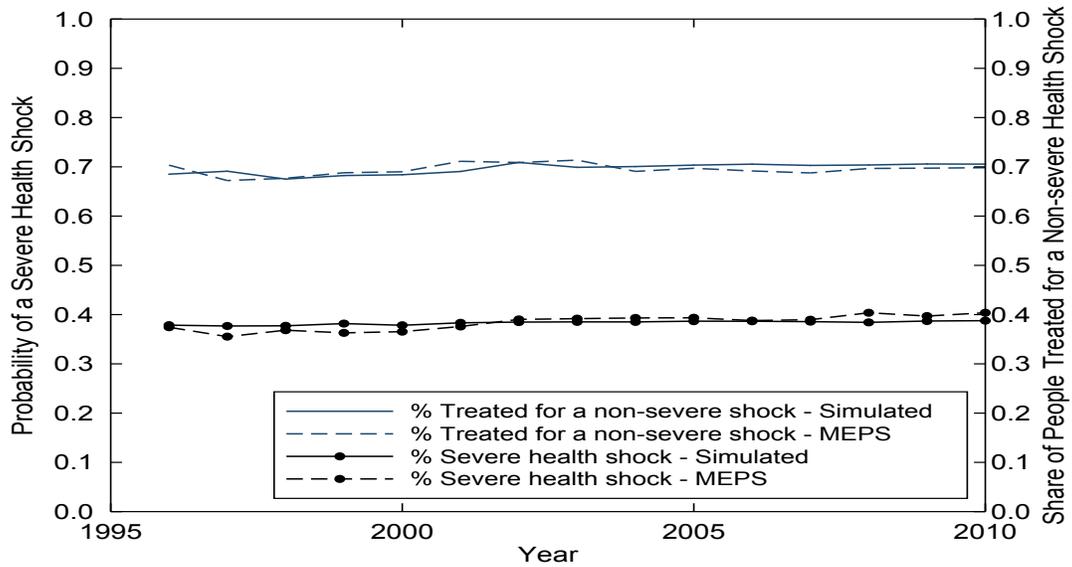
Figure 12: Simulated and Actual Data, Health Status
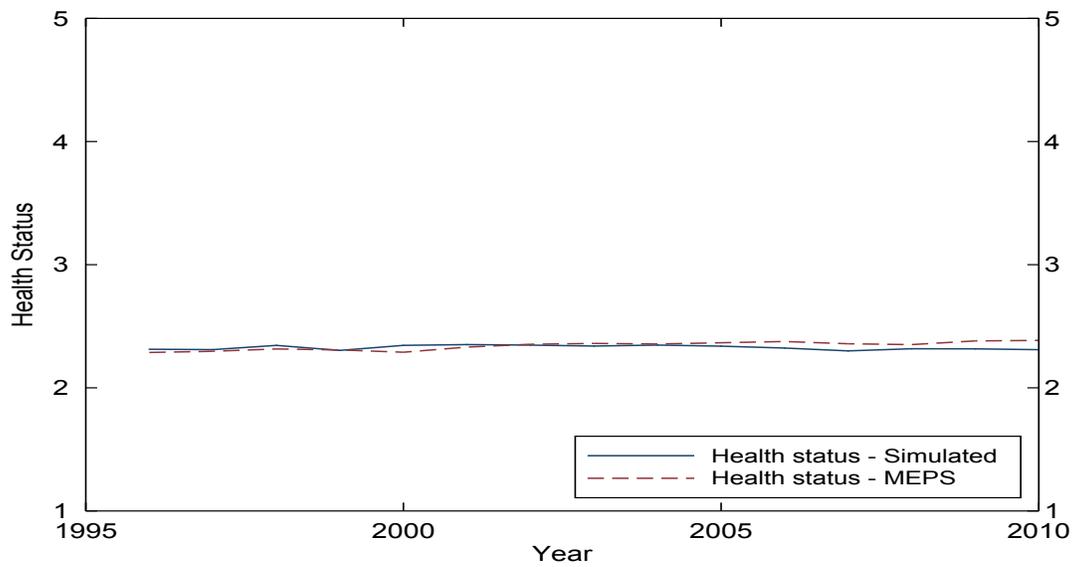
Figure 13: Simulated and Actual Data, Income
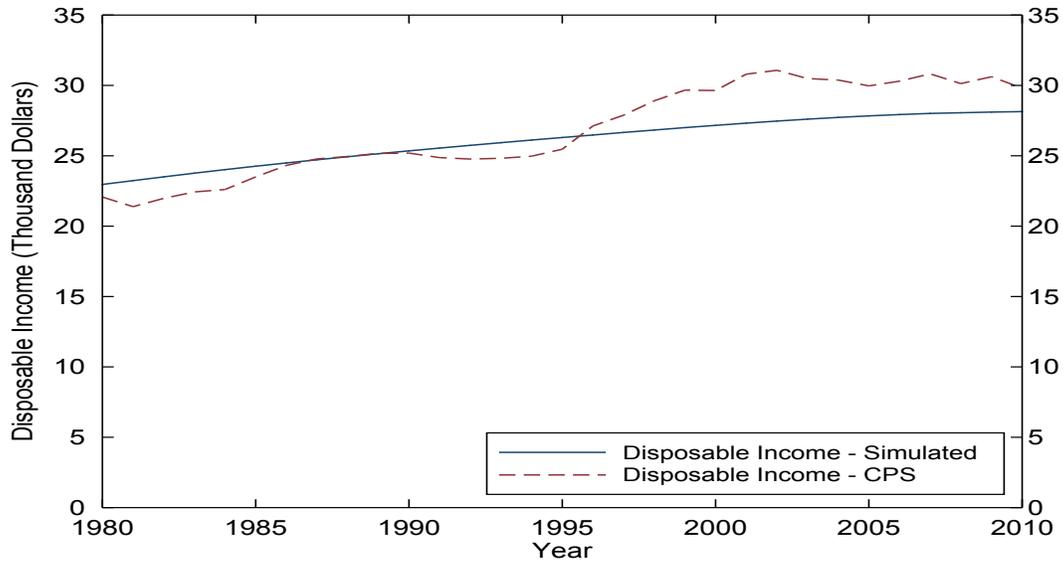
Figure 14: Income Fixed Effects
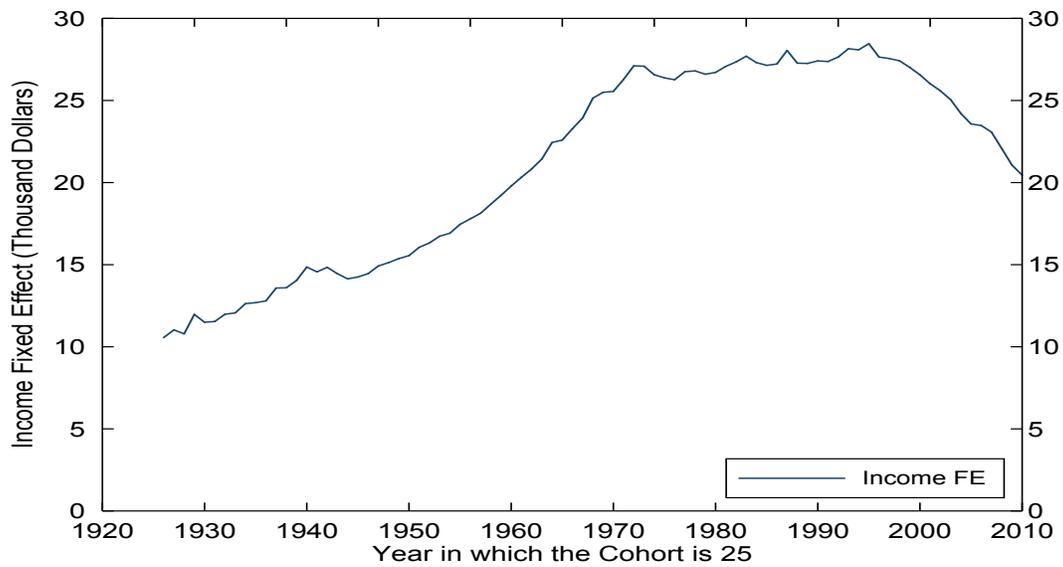
Figure 15: Age Structure
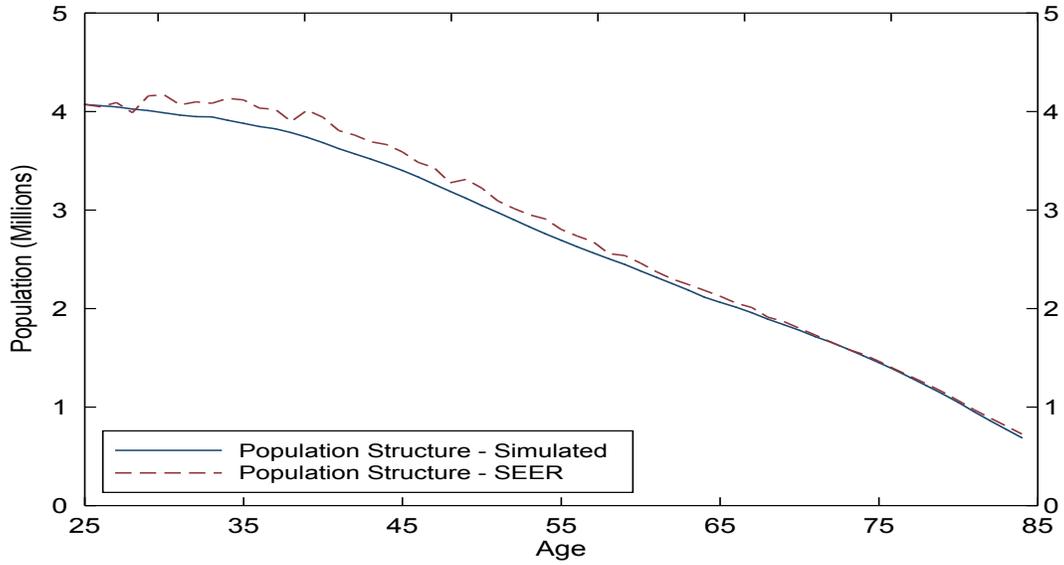
Figure 16: Share of Treatment and Co-payment Rate
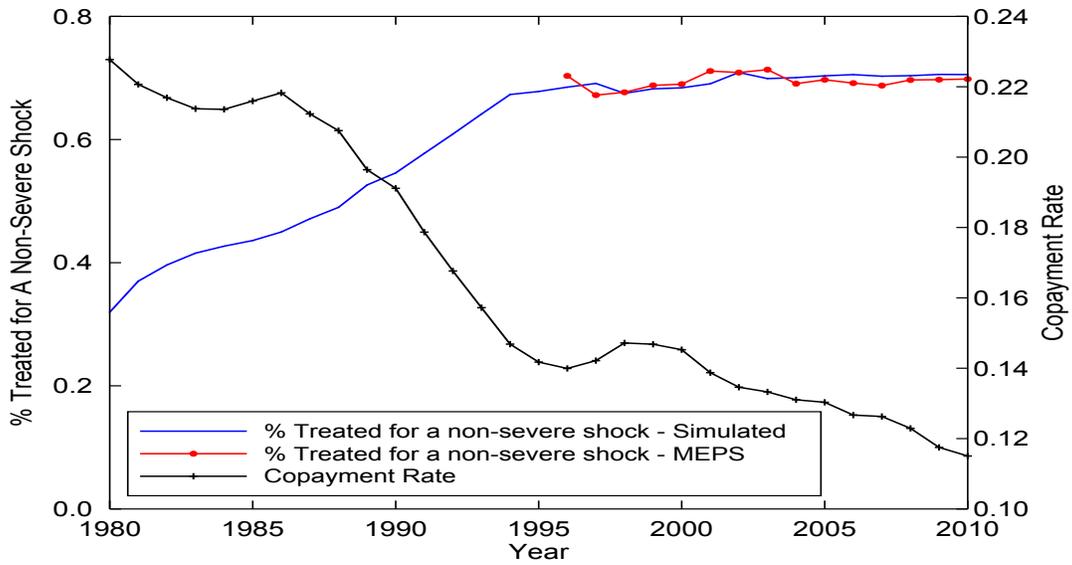
Figure 17: Simulated and Actual Data, Saving Rate
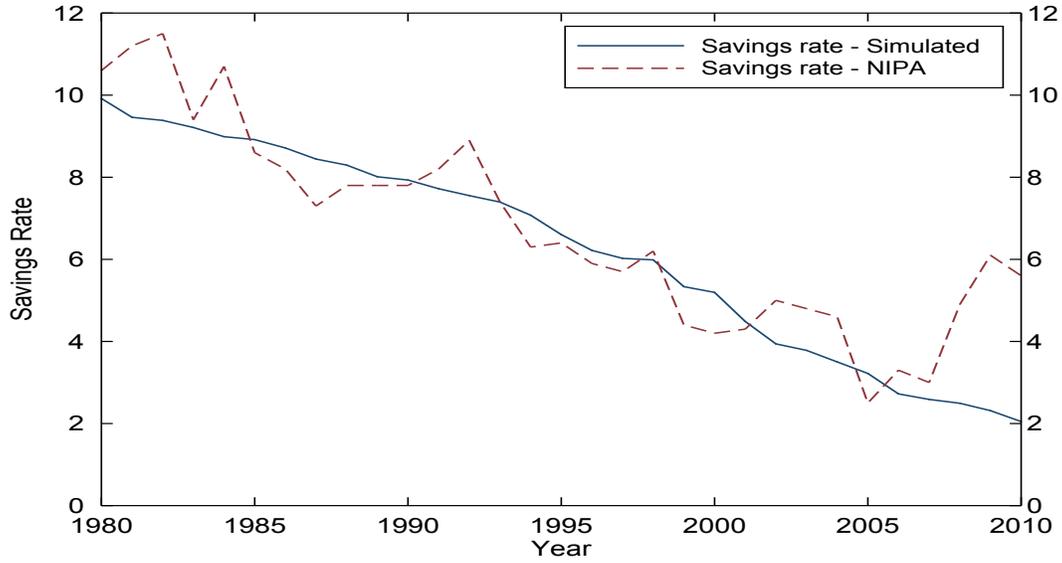
Figure 18: Counterfactual Experiments
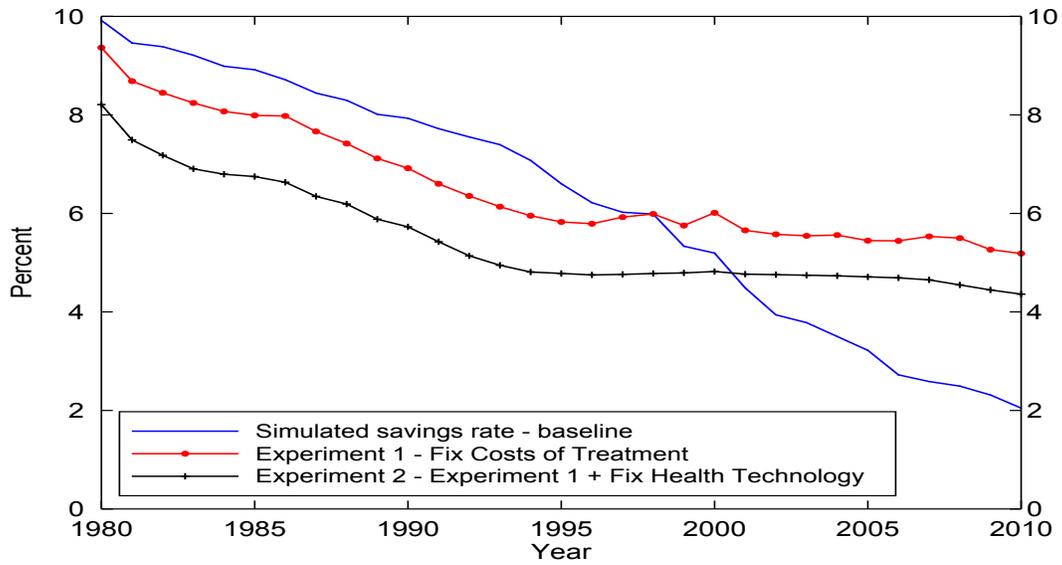
**Figure 19:** Counterfactual Experiments, Continuation
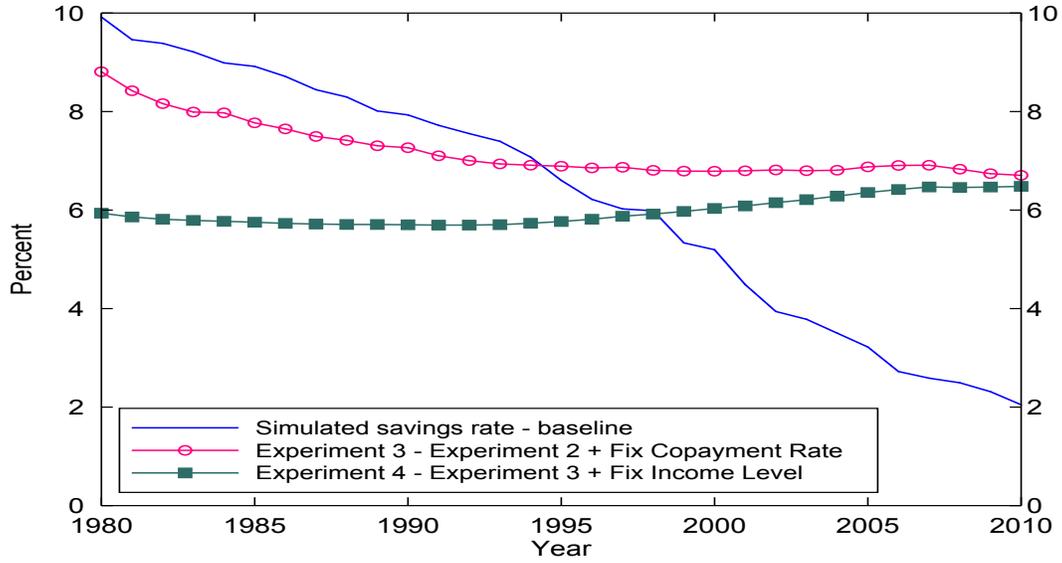
**Figure 20:** Non-health Consumption as a Share of Income
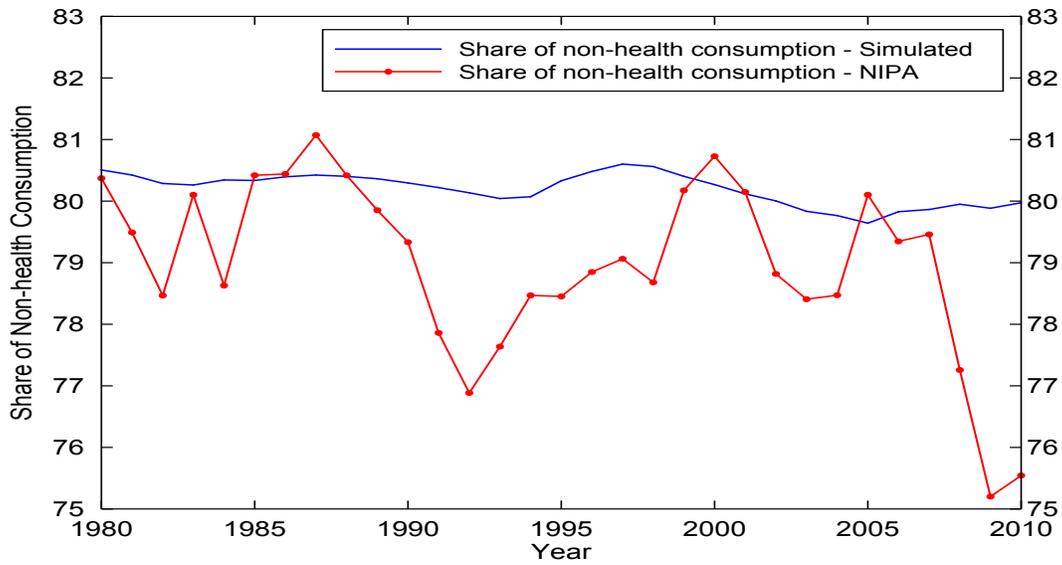
**Figure 21:** Life-cycle Treatment Probability for Non-severe Shocks
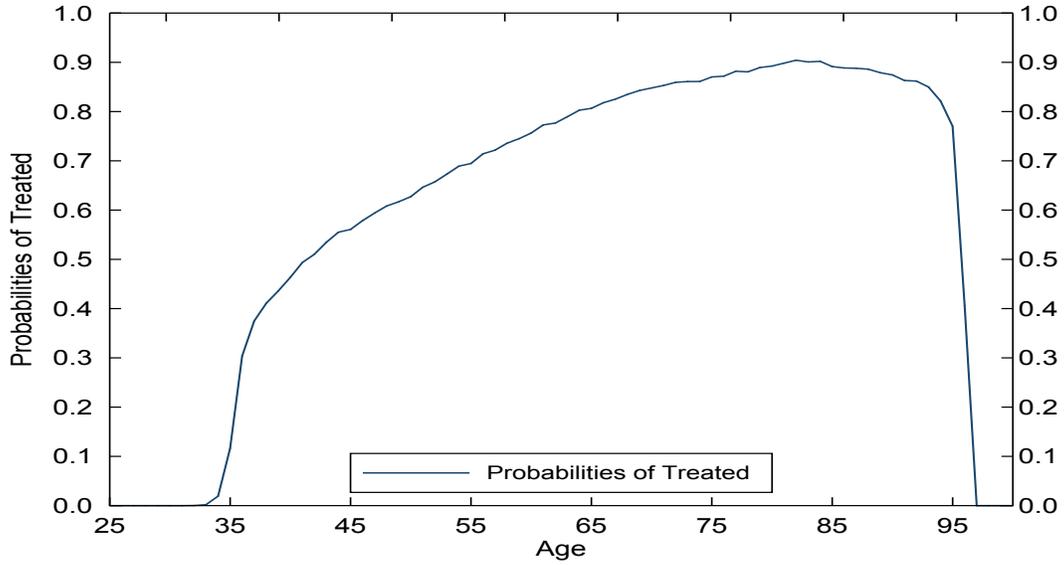
**Figure 22:** Saving Rate with Only Out-of-pocket Health Expenditure
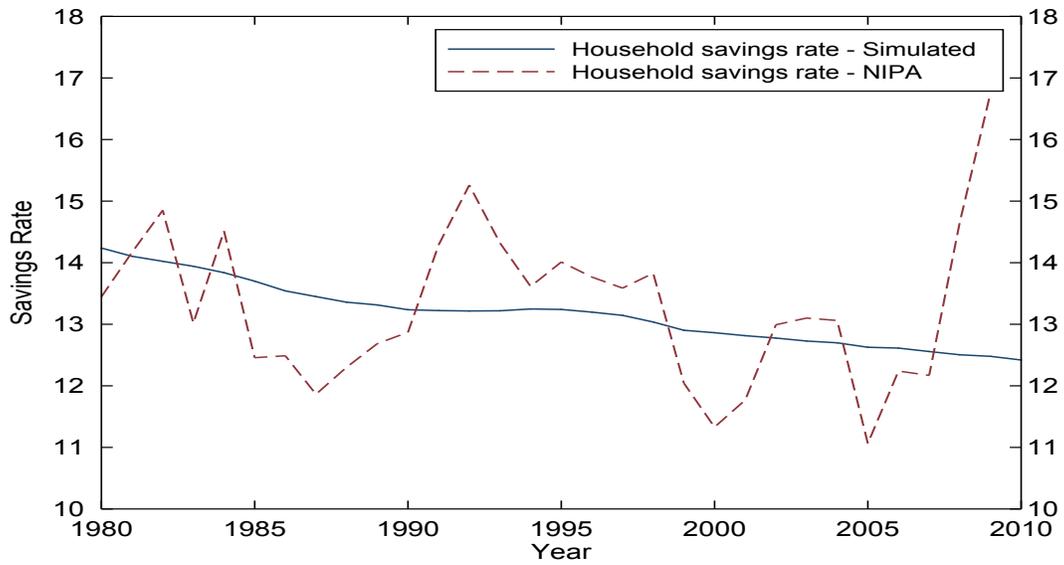
Figure 23: Simulated Share of Medical Expenditure and Saving Rate with 30% of Severe Conditions
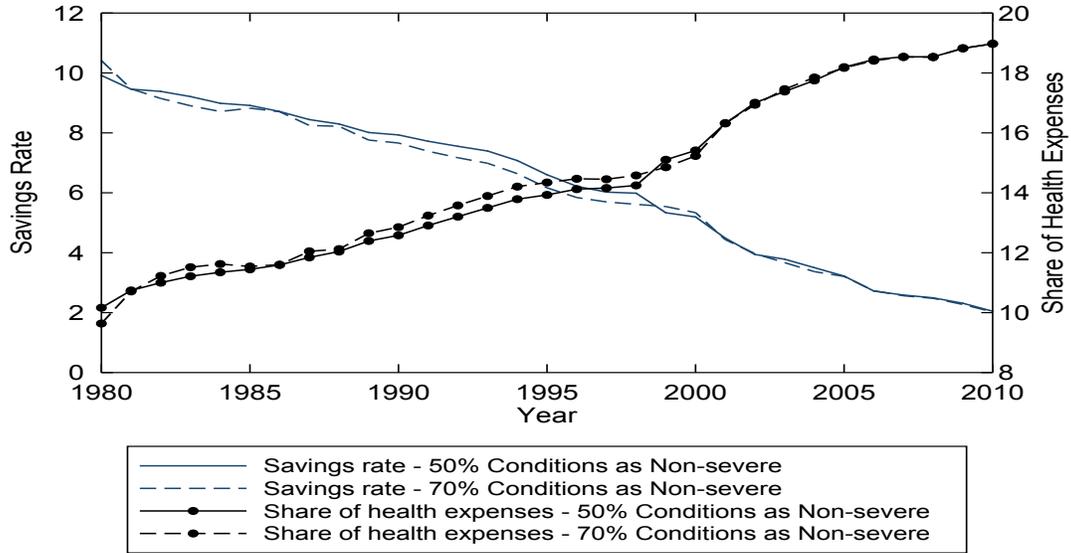

Figure 24: Public and Private Health Expenditure As Percentage Of Disposable Income
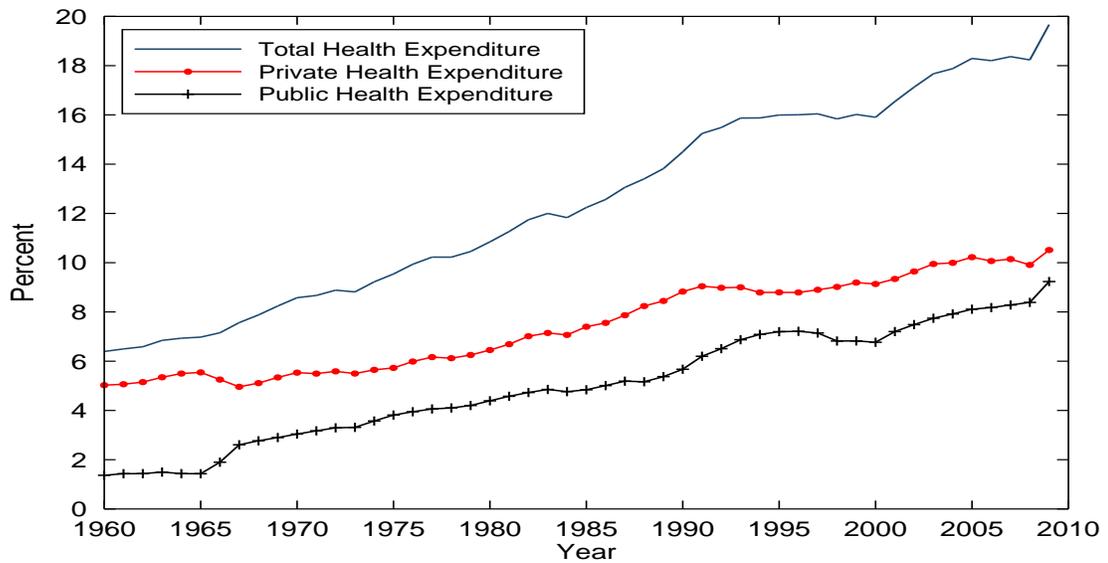
Figure 25: Contributions By Employers To Health Expenditure As Percentage Of Disposable Income

**Figure 26:** Decomposing Consumer Health Expenditure As Percentage Of Disposable Income
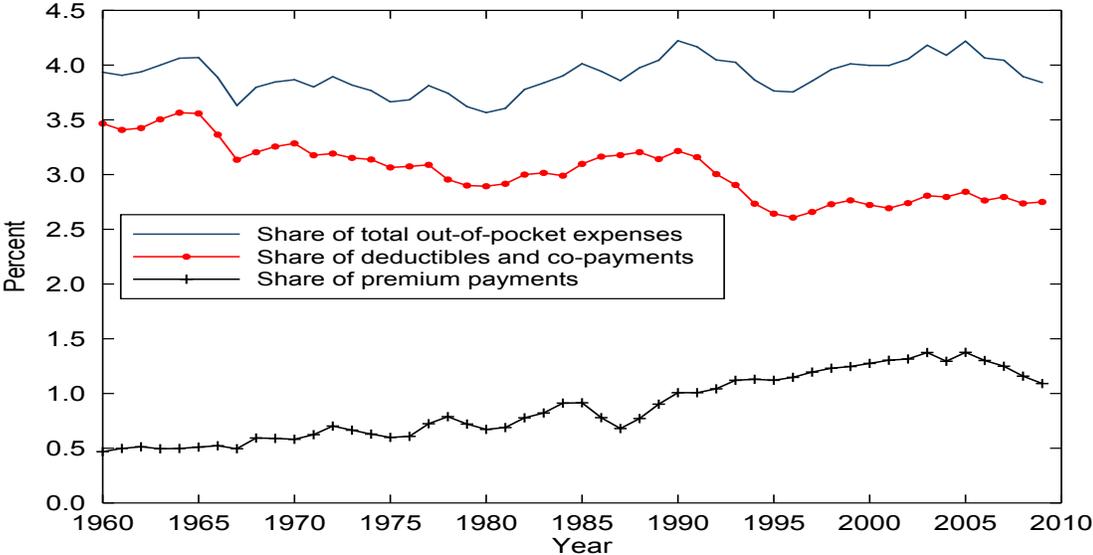
Figure 27: Share of Wages and Salaries and Share of Employer Contributions to Health Plans
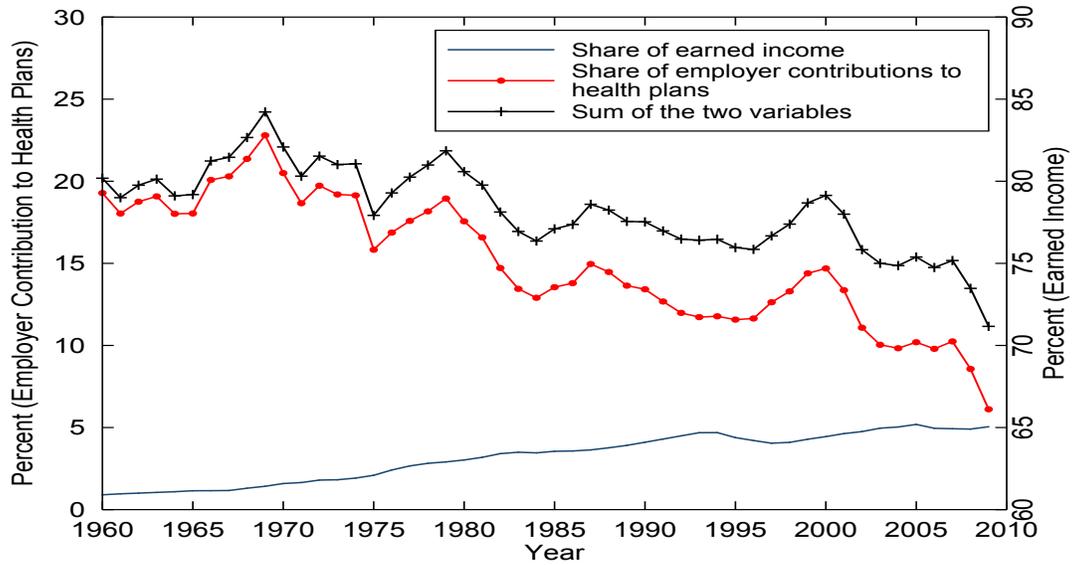

Figure 28: Public Expenditure Components As a Share of Disposable Income
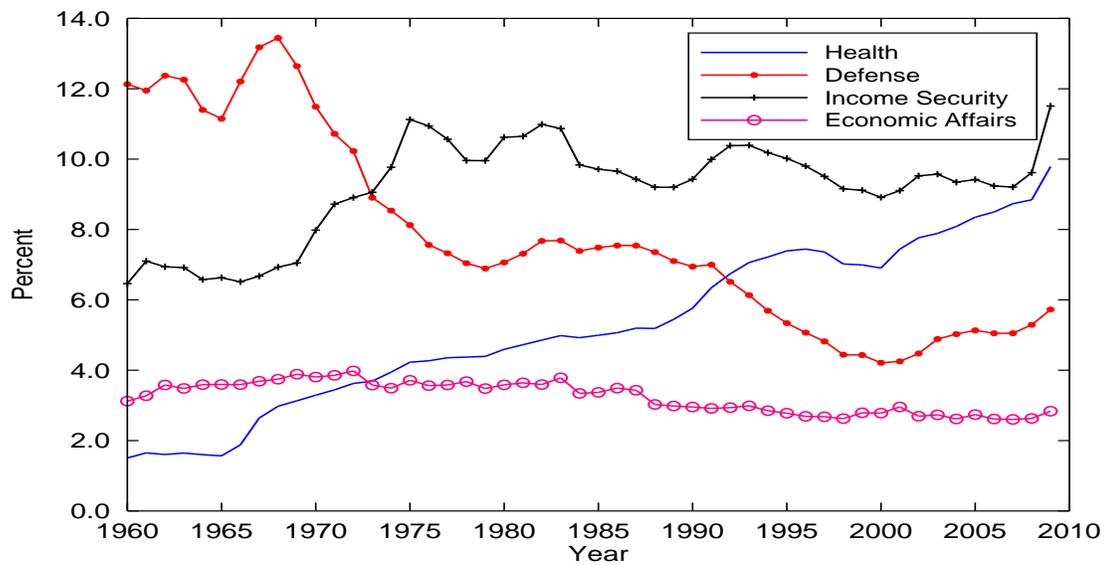
Figure 29: Personal Taxes, Government Revenue and Expenditure As a Share of Disposable Income
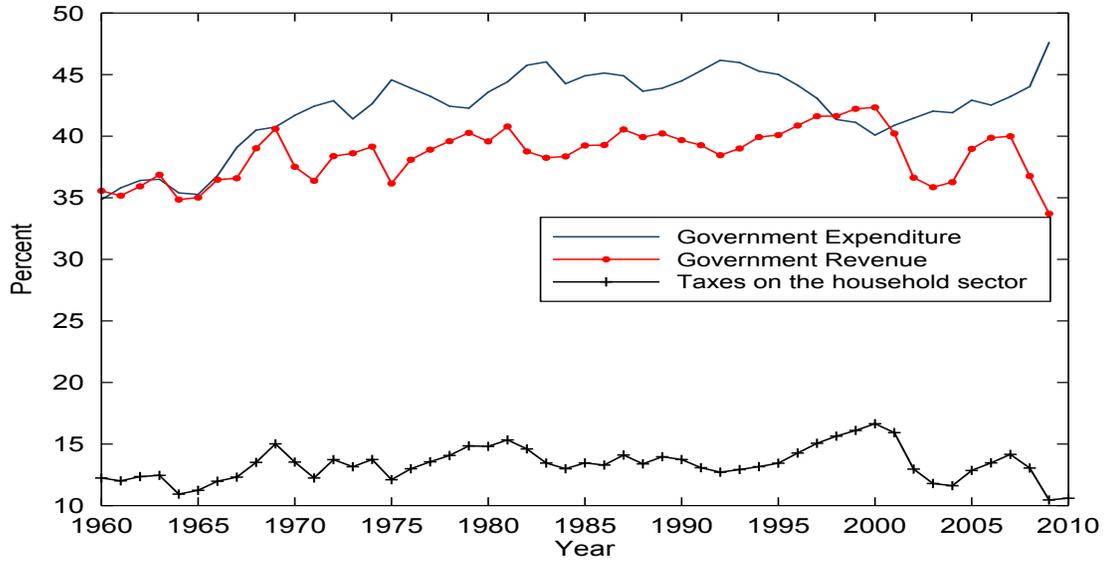
Figure 30: Simulated Saving Rate with Expected Debt Repayment
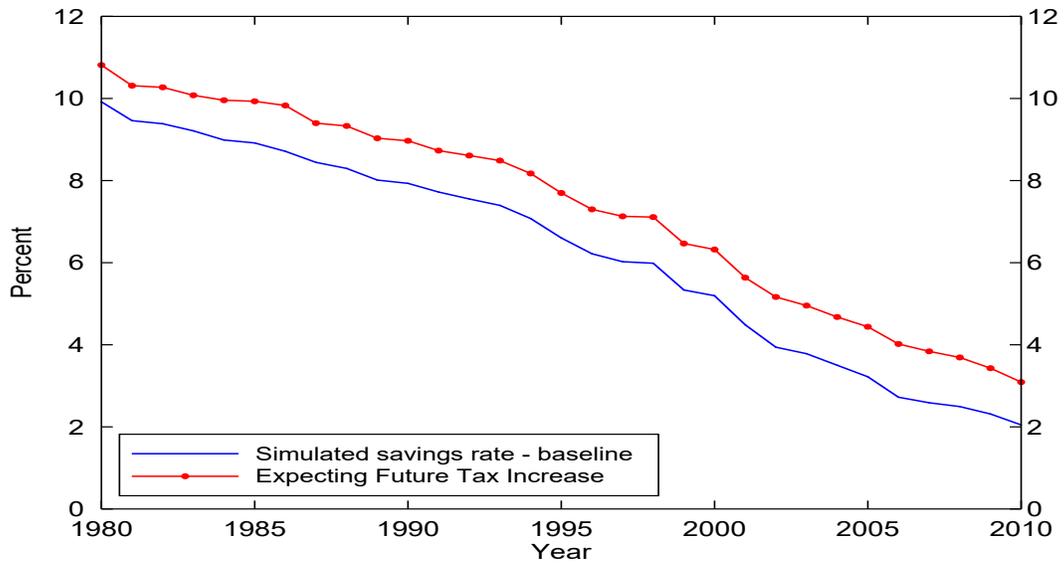
Table 3: Cross-state Regression

Dependent Variable: State-level Saving Rate	
Share of Medical Expenditure	**-0.643*****
	(0.135)
Share of Old Individuals	0.220
	(0.222)
Share of Middle-aged Individuals	0.933***
	(0.241)
Share with College or Higher Degree	0.062
	(0.056)
Share of African-Americans Individuals	0.182
	(0.194)
Share of Hispanic Individuals	0.404***
	(0.093)
State Unemployment Rate	0.410***
	(0.085)
N	1275
R-squared	0.845

State-level saving rates constructed using data from NIPA, Retail Sale Data, and NHEA. Time and state fixed effects are included as regressors. * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 4: Instrumental Variable Regressions

	First Stage Regressions. Dependent Variable: Share of Medical Expenditure			
	$\rho = 10\%$	$\rho = 15\%$	$\rho = 10\%$	$\rho = 15\%$
	all instruments	all instruments	age	age
Approval * Share Old	0.0708***	0.0752***	0.0899***	0.0955***
	(0.0072)	(0.0077)	(0.0065)	(0.0069)
Approval * Share with College or Higher Degree	-0.0479***	-0.0516***		
	(0.0051)	(0.0055)		
Approval * Share Blacks	0.0082***	0.0089***		
	(0.0019)	(0.0021)		
Approval * Share Hispanic	0.0086**	0.0099**		
	(0.0043)	(0.0046)		
N	1275	1275	1275	1275
F-test	113.0	111.7	193.5	189.8

All regressions include as control variables the fraction of individuals between the ages of 30 and 60, the fraction older than 60, the fraction with college or higher degree, the fraction of Blacks and Hispanics, and the unemployment rate, time and state fixed effects. * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 5: Instrumental Variable Regressions

	$\rho = 10\%$ all instruments	$\rho = 15\%$ all instruments	$\rho = 10\%$ age	$\rho = 15\%$ age
Second Stage Regressions. Dependent Variable: State-level Saving Rates				
Share of Medical Expenditure	**-0.730****	**-0.746****	**-0.834****	**-0.889****
	(0.281)	**(0.281)**	**(0.430)**	**(0.433)**
Share Old	0.171	0.173	0.183	0.189
	(0.221)	(0.221)	(0.230)	(0.230)
Share Middle Age	0.661***	0.661***	0.661***	0.661***
	(0.232)	(0.232)	(0.234)	(0.234)
Share with College or Higher Degree	0.061	0.061	0.060	0.059
	(0.054)	(0.054)	(0.054)	(0.054)
Share Blacks	0.304	0.305	0.310*	0.313*
	(0.188)	(0.188)	(0.188)	(0.187)
Share Hispanics	0.331***	0.327***	0.303**	0.288*
	(0.119)	(0.119)	(0.147)	(0.148)
State Unemployment Rate	0.344***	0.347***	0.366***	0.379***
	(0.104)	(0.104)	(0.135)	(0.135)
N	1275	1275	1275	1275

State-level saving rates constructed using data from NIPA, Retail Sale Data, and NHEA. Time and state fixed effects are included as regressors. * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 6: Instrumental Variable Regressions With Share of Workers in Pharmaceutical Industry

First Stage Regressions. Dependent Variable: Share of Medical Expenditure				
	$\rho = 10\%$	$\rho = 15\%$	$\rho = 10\%$	$\rho = 15\%$
	all instruments	all instruments	age	age
Approval * Share Old	0.0704***	0.0748***	0.0902***	0.0958***
	(0.0072)	(0.0078)	(0.0065)	(0.0069)
Approval * Share with College or Higher Degree	-0.0481***	-0.0518***		
	(0.0052)	(0.0056)		
Approval * Share Blacks	0.0082***	0.0090***		
	(0.0019)	(0.0020)		
Approval * Share Hispanic	0.0086**	0.0100**		
	(0.0043)	(0.0046)		
N	1275	1275	1275	1275
F-test	112.9	111.5	194.3	190.6

All regressions include as control variables the fraction of individuals between the ages of 30 and 60, the fraction older than 60, the fraction with college or higher degree, the fraction of Blacks and Hispanics, the unemployment rate, the share of workers in the pharmaceutical industry, time and state fixed effects. * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 7: Instrumental Variable Regressions With Share of Workers in Pharmaceutical Industry

	$\rho = 10\%$ all instruments	$\rho = 15\%$ all instruments	$\rho = 10\%$ age	$\rho = 15\%$ age
Second Stage Regressions. Dependent Variable: State-level Saving Rates				
Share of Medical Expenditure	**-0.725****	**-0.741****	**-0.860****	**-0.914****
	(0.281)	**(0.281)**	**(0.431)**	**(0.435)**
Share Old	0.172	0.174	0.187	0.193
	(0.221)	(0.221)	(0.231)	(0.231)
Share Middle Age	0.647***	0.647***	0.647***	0.647***
	(0.233)	(0.233)	(0.235)	(0.236)
Share with College or Higher Degree	0.056	0.056	0.055	0.054
	(0.054)	(0.054)	(0.054)	(0.054)
Share Blacks	0.307	0.308	0.315*	0.318*
	(0.189)	(0.189)	(0.189)	(0.188)
Share Hispanics	0.333***	0.329***	0.296*	0.281*
	(0.119)	(0.120)	(0.147)	(0.148)
State Unemployment Rate	0.336***	0.339***	0.365***	0.378***
	(0.105)	(0.104)	(0.135)	(0.136)
State Share of Workers in Pharma. Ind.	0.644	0.645	0.649	0.651
	(0.589)	(0.588)	(0.588)	(0.588)
N	1275	1275	1275	1275

State-level saving rates constructed using data from NIPA, Retail Sale Data, and NHEA. Time and state fixed effects are included as regressors. * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 8: Main Parameters

Parameter	Value	Source
Discount Factor β	0.994	De Nardi, French, and Jones (2016)
Curvature Parameter γ	2.83	De Nardi, French, and Jones (2016)
Bequest Intensity η	39.7	De Nardi, French, and Jones (2016)
Bequest Curvature k	13.0	De Nardi, French, and Jones (2016)
Utility Intercept δ^a	0.0052	Hall and Jones (2007)
Share of Severe Conditions λ	0.50	MEPS
Weight w in $P^N(a_t, h_{t-1})$	0.70	MEPS

Notes: [a]The intercept is obtained by first dividing the value used in Hall and Jones (2007) by 1,000 to account for differences in units and then by dividing the result by 5 account for the different definition of a period: 5 years in Hall and Jones (2007) versus 1 year in this paper.

A Appendix

Figure A.1: Actual and Filtered Saving Rate
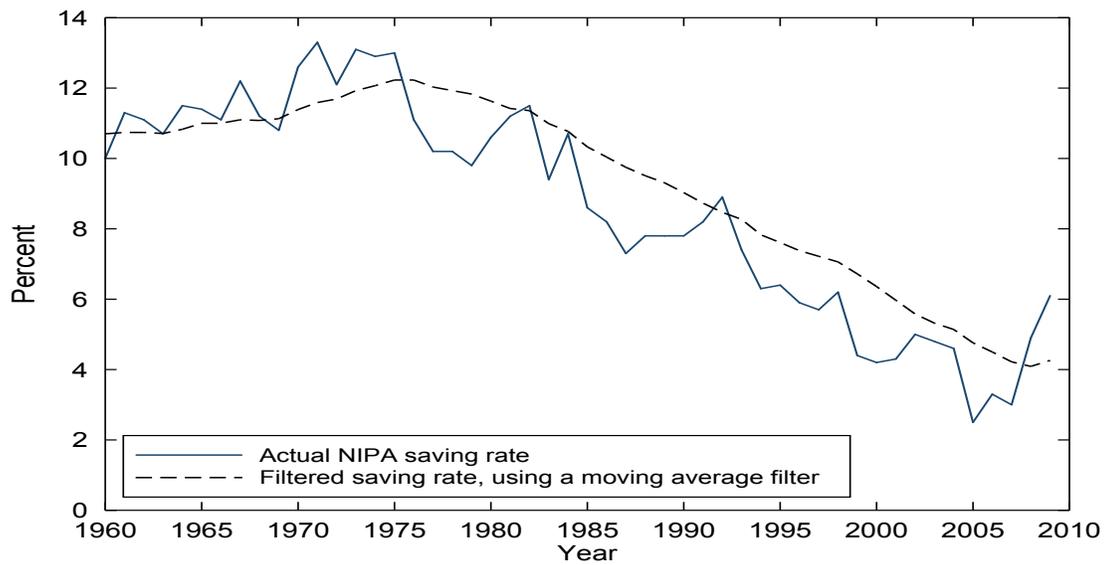
Table A.1: Mortality, Health, and Expenditure Functions

Mortality		Health Status		Expenditure		
					Non-severe	Severe
Age	9.123***	Health	0.606***	Age	0.082	0.302**
	(3.390)		(0.002)		(0.061)	(0.145)
Age^2	-7.4248***	Age	0.0080***	Age^2	-0.215*	-0.854***
	(2.786)		(0.0002)		(0.120)	(0.277)
Health	59.207*	Expenses	-0.015	Age^3	0.176**	0.640***
	(31.431)		–		(0.073)	(0.168)
Year	0.111**	ϵ^N	0.151***	Health	1.021**	3.523***
	(0.049)		(0.006)		(0.432)	(1.183)
Age*Year	-0.005***	ϵ^S	0.227***	Age*Health	-0.061**	-0.239***
	(0.002)		(0.006)		(0.027)	(0.068)
Age^2*Year	0.004***			Age^2*Health	0.173***	0.565***
	(0.001)				(0.052)	(0.125)
Age*Health	-2.025*			Age^3*Health	-0.128***	-0.381***
	(1.036)				(0.031)	(0.073)
Age^2*Health	1.515*					
	(0.827)					
Health*Year	-0.030*					
	(0.016)					
Age*Year*Health	0.0010**					
	(0.0005)					
Age^2*Year*Health	-0.0008*					
	(0.0004)					
N	1,563,828	N	166,308	N	237,276	136,192

Time fixed effects are included as regressors in the estimation of the health status and health expenditure functions.

* significant at 10%; ** significant at 5%; *** significant at 1%.