Dynamic Inefficiencies in an Employment-Based Health Insurance System: Theory and Evidence∗

Hanming Fang† Alessandro Gavazza‡

April 3, 2009

Abstract

We investigate the effects of the institutional settings of the U.S. health care system on individuals’ life-cycle medical expenditures. We argue that health is a form of human capital that affects labor productivity, and that the employment-based health insurance system may lead to inefficient investment in individuals’ health care. The reason is that labor turnover and frictions in the labor market prevent an employer-employee pair from capturing the entire surplus from investment in an employee’s health. Thus, the pair underinvests in health capital, and this underinvestment increases medical expenditures during retirement.

We provide extensive empirical evidence consistent with the comparative statics predictions of our model using two datasets, the Medical Expenditure Panel Survey (MEPS) and the Health and Retirement Study (HRS). The magnitude of our estimates suggests a significant degree of inefficiency in health investment in the U.S.

JEL Classification Numbers: D84, D91, I12

∗We are grateful to Daron Acemoglu, Luis Cabral, Michael Chernew, Keith J. Crocker, Amy Finkelstein, Igal Hendel, Stephen Morris, Jörn-Steﬀen Pischke, Aloysius Siow and seminar participants at University of Connecticut, Maryland, Rice/Houston, Texas A&M, Fudan, Peking, Stanford, Yale, NBER, ERIU Conference (Ann Arbor), and IOMS Workshop (Shanghai) for many helpful comments and suggestions. Special thanks to Ray Kuntz and Michael Nolte for help with the MEPS and HRS data, respectively. We are responsible for all shortcomings.

†Department of Economics, Duke University, 213 Social Sciences Building, P.O. Box 90097, Durham, NC 27708-0097; and the NBER. Email: hanming.fang@duke.edu

‡New York University, Leonard N. Stern School of Business, 44 West 4th Street, Suite 7-81, New York, NY 10012. Email: agavazza@stern.nyu.edu
1 Introduction

The health care system in the United States differs in at least two stark ways from those of other industrialized countries. The first is its institutional organization: The United States is unique among industrialized nations in that it lacks a national health insurance system. The U.S. health insurance system is a mixture of private and public insurances, with private insurance playing a much more important role than in other industrialized countries. In particular, in the U.S., a private, employment-based system provides insurance to most of the working-age population, while a public program—i.e., Medicare—provides insurance to almost all individuals aged 65 and over. The second difference is its costs: The U.S. spends more than twice as much on health care as a fraction of G.D.P. as other developed countries. For example, in 2005, the U.S. and the U.K. spent about 17 and eight percent of their GDP, respectively, on health care (Hagist and Kotlikoff, 2005).

This paper investigates the effects of the institutional settings of the U.S. health care system on individuals’ life-cycle medical expenditures. Our premise is that health is a form of general human capital (Becker, 1964) and that health investment—medical expenditures in particular—determines the stock of health (Grossman, 1972). Hence, like all other forms of human capital, health increases labor productivity, thereby affecting the surplus generated in the employment relationship. Thus, current health expenditures are an investment that affects the current and future surplus of the employment relationship. We embed this link between health investment and employment surplus in a frictional labor market, as in Acemoglu and Pischke (1999), and derive the implications of employee turnover on the employer-employee pair’s incentives to invest in the employee’s health.

We show that employee turnover leads to an inefficiently low level of investment in the employees’ health, and that investment is lower and inefficiencies larger when employee turnover is higher. The reason is that frictions in the labor market imply that part of the surplus from the current investment in the employee’s health accrues to a future employer. Hence, the employer-employee pair does not internalize the full social surplus created by the current investment in the employee’s health. As a result, the pair underinvests in health capital. Further, we show that this inefficiently low level of medical expenditures during the working years increases medical expenditures during retirement, possibly increasing the overall expenditures.

We provide extensive empirical evidence consistent with the predictions of the model using two datasets, the Medical Expenditure Panel Survey (MEPS) and the Health and Retirement Study (HRS). Our empirical model is designed to deal explicitly with two issues that may hinder the identification of the effect of job turnover. The first is selection: Workers may select into different jobs due to unobserved characteristics, such as ability, discount factor, risk aversion, etc. that could potentially be correlated with their job turnover. The second is reverse causality: Workers’ health outcome and health expenditures could affect their job turnover. We deal with these issues using panel data to control for fixed and persistent unobservables that could affect selection into different
jobs, along with demand-side instruments—i.e., plant closures—that arguably are not affected by reverse causality.\footnote{In Appendix A, we present an alternative empirical approach that uses a measure of the importance of specific skills in each industry as a proxy for an industry’s labor turnover rate.}

The empirical analysis provides extensive evidence consistent with the predictions of our model. We find that workers with shorter job tenures spend less on health care. However, we find a stark reversal of expenditures among the elderly: Retirees who had longer job tenures spend less on health. The magnitude of our results is considerable. Workers with job tenures that are one standard deviation longer have medical expenditures about $660 higher per year. Individuals over 65 whose tenure at their main job is one standard deviation spend about $4,700 less per year on health care. Using these estimates, we can perform a back-of-the-envelope calculation to compare the lifetime medical expenditures of two workers whose only difference is their job tenures. Suppose that both individuals work 45 years and then retire for 15 years before dying, but the first individual’s job tenure is one standard deviation longer than the second individual’s. According to our estimates, during their working years, the first individual spent approximately $29,700 more on health care than the second individual did. Instead, during retirement, the first individual’s health expenditures are approximately $70,500 lower than the second individual’s. The total difference is around $40,000. This calculation suggests that one additional dollar of health expenditures during working years may lead to about 2.5 dollars of savings in retirement. Obviously, our calculation is very rough. It does not incorporate discounting, does not adjust for the inflation of medical prices, and does not adjust for differences in life qualities and life expectancies. Nonetheless, it suggests that the dynamic externality identified by our model can be large and can substantially increase the medical expenditures of the elderly.

This paper makes a number of contributions. First, it sheds light on the incentives generated by the employment-based health insurance system. We wish to emphasize at the outset that our paper is not about health insurance per se. Rather, the paper investigates how the health insurance system affects incentives to invest in health. By focusing on health investment, we do not tackle the difficult incidence issue of how much of the health insurance premium is actually paid for by the firms and workers (Gruber, 1994). Instead, we focus on whether an employment-based system internalizes the entire surplus generated by health investment. We suggest that the employment-based health care system may lead to dynamic inefficiencies.

Second, the paper suggests that different institutional arrangements of the health care system can lead to different life-cycle dynamics of health expenditures. Our analysis indicates that in the U.S., because of job turnover, the increase in expenditures is steeper when the parties—i.e., the employer-employee pair—appropriate a smaller share of the entire surplus generated by health investment. Thus, our paper suggests that an employment-based system may steepen the increase of
health expenditures over an individual’s life-cycle compared to a national health insurance system.² Moreover, by not internalizing the full long-term benefits of health investment, an employment-based health system can also increase the overall level of expenditures.

Third, taking the view that health is a form of general human capital, our paper also serves as an empirical analysis of how firms and workers invest in general human capital. In fact, we believe that health expenditures are particularly suited to studying how firms and workers jointly determine the level of general human capital investment. The reason is that health expenditures are typically well-recorded, as most health investment is provided by third-party medical professionals with well-documented charges. In contrast, for almost all other investments in general human capital, it is quite difficult to obtain a quantitative measure of total costs and each party’s contribution. In these situations, it is often the case that only firms’ general training expenditures may be recorded, while workers’ contribution to general investment is typically unobserved.

The remainder of the paper is structured as follows. Section 2 reviews some of the existing work on the relationship between health and productivity. Section 3 presents a simple model and derives its testable implications. Section 4 describes the data sets. The main empirical analysis is performed in Section 5. Section 6 discusses several alternative hypotheses. Section 7 discusses additional related literature. Section 8 concludes. Appendix A presents an extension of the model that accommodates firm-specific capital, along with an alternative empirical strategy that obtains very similar qualitative and quantitative results.

2 Background: Health and Productivity

This section reviews some of the existing work on one economic issue that lies at the heart of our paper—i.e., how health affects productivity. Because the literature is vast, we will limit our discussion to contributions that, although using different methods and different data, share the common finding that healthier individuals are more productive.³

The classical theory of human capital developed by Becker (1962, 1964) posits that investment in health capital shares all the features of investments in other forms of human capital, such as education. Grossman’s (1972) seminal paper provides a rigorous framework for studying the relationships among health, human capital and productivity. The fundamental feature of Grossman’s model is that better health provides utility directly, but also increases the amount of time individuals can devote to market activities, such as production of goods. Thus, health is an (indirect) input into an individual’s production function (i.e., health is a capital good), and past investments (i.e.,

²Systematic comparisons of the dynamics of individuals’ total health expenditures over the life-cycle across different countries is limited. One related evidence is Hagist and Kotlikoff (2005), which focuses only on public health expenditures. They report that the ratio of per capita health expenditures of the 65-69 age group relative to that of the 50-64 age group is approximately five times higher in the U.S. than in other nine OECD countries.

³For a survey on the relationship between health and productivity, see Tompa (2002).
health expenditures, exercise, nutrition) affect current health, partially offsetting the depreciation of the health stock over time.

The empirical evidence on the effects of health on productivity is rich and encompasses several fields. Several papers establish that increased life expectancy and reduced morbidity increase productivity and output. Most of these papers focus on historical trends and/or developing countries. For example, Fogel (1991, 1994) shows how improvements in health affect living standards over time in Europe and in the United States. Similarly, several empirical studies document that health has a significant and positive effect on economic growth (e.g., Barro and Sala-i-Martin, 1995; Knowles and Owen, 1995; Bloom, Canning and Sevilla, 2001).

At the individual level, a very large literature has studied the interactions between individual health and labor market outcomes. Currie and Madrian (1999) provides a very detailed overview. Many papers find that less healthy individuals are less productive, broadly defined. Among them, Haveman et al. (1994) finds that prior health limitations negatively affect work-time and have a significant negative effect on wages. Berkovec and Stern (1991) finds that bad health decreases labor market participation among the elderly. Stern (1996) finds that health limitations on the ability to work have larger effects on individual labor supply than on labor demand.

A few recent studies focus on even more-detailed micro-evidence to study the effects of health on worker productivity. For example, Nicholson et al. (2006) uses survey data from a sample of establishments and provides direct evidence that the cost or productivity loss associated with missed work is higher than the wage. Similarly, Davis et al. (2005), using a survey, finds that:

\[
\text{In 2003, an estimated 18 million adults ages 19 to 64 were not working and had a disability or chronic disease, or were not working because of health reasons. Sixty-nine million workers reported missing days due to illness, for a total of 407 million days of lost time at work. Fifty-five million workers reported a time when they were unable to concentrate at work because of their own illness or that of a family member, accounting for another 478 million days. Together, labor time lost due to health reasons represents lost economic output totaling $260 billion per year.}
\]

More broadly, an individual’s current health investment can affect his future health costs, and the individual and his current or future employer will need to pay for these future costs. Thus, an individual’s current expenditures can affect the future expected surplus of the current employment relationship through lower future health expenditures. For example, in an interesting study of diabetes management, Beaulieu et al. (2007) finds that improved diabetes care affords economic benefits to health plans, as well as valuable benefits to people with diabetes. However, some of the long-term savings from good care management are not realized because plan turnover limits the health plan’s ability to privately capture the benefits from its investments. The model in Section 3 focuses precisely on this trade-off between the short-term costs and the expected long-term gains of current health expenditures.
3 A Simple Model

In this section we present a simple model that adapts the theoretical framework of Acemoglu and Pischke (1999) to health expenditures. The goal of the model is to capture in the simplest way the effect of workers’ turnover on the incentives to invest in health. In particular, we make the simplest assumptions to focus on the dynamic externality that we described in the Introduction.

3.1 Assumptions

There are two periods with no discounting. Health is a form of general human capital and, thus, it is an input in the production function of the worker. For simplicity, we assume that health is the only input in the production function \( f(h) \), where \( f(\cdot) \) is assumed to be increasing, differentiable and concave. Workers are risk-neutral, and are endowed with an initial stock of health \( h_1 \). In the first period, workers can invest \( m_1 \) in health at a unit cost \( p \). Health evolves according to

\[
h_2 = k(h_1, m_1),
\]

where \( k \) is the health function, which we assume to be continuous and increasing in the stock of health \( h_1 \) and in the investment in health \( m_1 \)—i.e. \( \partial k/\partial h_1 > 0 \) and \( \partial k/\partial m_1 > 0 \).

In the second period, the firm and the worker separate with an exogenous probability \( q \in (0, 1) \). If they separate, the worker gets an outside wage of \( v(h_2) \), and the firm gets a surplus of zero. If they do not separate for exogenous reasons, the worker decides whether to stay with the firm and obtain the (endogenous) wage \( w_2(h_2) \) or to quit and obtain the (exogenous) outside wage of \( v(h_2) \). It is important to note that the worker’s productivity is \( f(h_2) \). However, if the worker leaves her current firm, for either exogenous or endogenous reasons, she receives a wage equal to \( v(h_2) \). Acemoglu and Pischke (1999) assumes that \( v(h_2) < f(h_2) \) to reflect labor market frictions and, more importantly, that \( v'(\cdot) < f'(\cdot) \), which they term wage compression. Acemoglu and Pischke (1999) provide a variety of mechanisms that can lead to a wedge between \( f(h_2) \) and \( v(h_2) \)—i.e., between a worker’s productivity and her wage at other firms. Similarly, they describe several mechanisms that endogenously generate wage compression.\(^4\)

We follow Acemoglu and Pischke’s (1999) full-competition regime, in which firms compete in the first period by offering a pair of wage and medical consumption \( \{w_1, m_1\} \) to workers, and in equilibrium they make zero profits.

3.2 Equilibrium

If the worker and the current firm do not separate for exogenous reasons, their employment relationship generates a surplus equal to \( f(h_2) - v(h_2) \) since \( f(h_2) \) is the output generated by the worker,

\(^4\)Several papers provide empirical evidence on wage compression: Beckmann (2001); Almeida-Santos and Mumford (2005); and Frazis and Loewenstein (2006), among others.
and \( v(h) \) is the wage that the worker gets if he quits voluntarily. We assume that this surplus 
\( f(h) - v(h) \) is divided according to the Nash Bargaining solution, in which \( \beta \in (0, 1) \) represents the worker’s bargaining power. Hence, the wage \( w_2(h) \) that the worker obtains if he/she does not quit is:

\[
w_2(h) = (1 - \beta) v(h) + \beta f(h).
\]

Thus, the firm’s expected profit in period two is:

\[
\pi_2(h) = (1 - q) \left[ f(h) - w_2(h) \right] = (1 - q) (1 - \beta) \left[ f(h) - v(h) \right];
\]

and in period one is:

\[
\pi_1(h) = f(h) - w_1 - pm_1,
\]

where \( w_1 \) is the worker’s first-period wage and \( m_1 \) is the worker’s first-period medical expenditures, both to be determined in equilibrium.

Thus, the sum of profits for the firm in the two periods (recall the no-discounting assumption for simplicity) is:

\[
\Pi = \pi_1(h) + \pi_2(h) = f(h) - w_1 - pm_1 + (1 - q) (1 - \beta) \left[ f(h) - v(h) \right]. \tag{1}
\]

Ex-ante competition among firms for the worker requires that the firm chooses \( m_1 \) and \( w_1 \) to maximize the sum of profits \( \Pi \) subject to the constraint that the worker receive as much utility as that offered by other firms \( U \)—i.e.,

\[
w_1 + (1 - q) [(1 - \beta) v(h) + \beta f(h)] + qv(h) \geq U. \tag{2}
\]

Competition for the worker among firms implies that the utility level \( U \) is high enough such that, in equilibrium, the firm makes zero profits—i.e., \( \Pi = \pi_1(h) + \pi_2(h) = 0 \).

Now, from equation (2), we have that, in equilibrium, the wage satifies:

\[
w_1 = U - (1 - q) [(1 - \beta) v(h) + \beta f(h)] - qv(h). \tag{3}
\]

Substituting the equilibrium worker’s wage (3) into the firm’s profit function (1) and maximizing with respect to the level of medical expenditures \( m_1 \), we obtain that the equilibrium level of medical expenditures \( m_1^* \) solves the following first-order condition:

\[
[qv' \left( k(h, m_1^*) \right) + (1 - q) f' \left( k(h, m_1^*) \right)] \frac{\partial k}{\partial m_1} = p. \tag{4}
\]

Equation (4) implies that investment in health is socially inefficient unless there is never separation \( (q = 0) \). To see this, note that the efficient level of health investment \( \hat{m}_1 \) solves

\[
f' \left( k(h, \hat{m}_1) \right) \frac{\partial k}{\partial m_1} = p, \tag{5}
\]
which equates the marginal social benefit of medical expenditures \( f'(k(h_1, \hat{m}_1)) \partial k/\partial m_1 \) to their marginal cost \( p \). The social benefit of health investment is given by the worker’s productivity \( f(h_2) \), which is independent of her employer, reflecting the nature of health as a form of general capital. The comparison between (4) and (5) reveals that the equilibrium health investment \( \hat{m}_1 \) is lower than the socially efficient level \( \hat{m}_1 \) because of wage compression—i.e., \( \nu'(\cdot) < f'(\cdot) \).

Moreover, Proposition 1 investigates the effect of the turnover probability \( q \) on the equilibrium level of medical expenditures \( m_1^* \), yielding the first implication that we empirically test in Section 5.1:

**Proposition 1** A decrease in the turnover rate \( q \) increases equilibrium health expenditure \( m_1^* \).

**Proof.** From equation (4), let us define \( \lambda(m_1^*, q) \) as:

\[
\lambda(m_1^*, q) = \left[q \nu'(k(h_1, m_1^*)) + (1-q) f'(k(h_1, m_1^*)) \right] \frac{\partial k}{\partial m_1} - p = 0.
\]

Using the implicit function theorem, we have:

\[
\frac{\partial m_1^*}{\partial q} = -\frac{\partial \lambda(m_1^*, q) / \partial q}{\partial \lambda(m_1^*, q) / \partial m_1^*}.
\]

From the wage compression assumption \( f' > \nu' \), we can obtain the following inequality:

\[
\frac{\partial \lambda(m_1^*, q)}{\partial q} = - \left[ f'(k(h_1, m_1^*)) - \nu'(k(h_1, m_1^*)) \right] \frac{\partial k}{\partial m_1} < 0.
\]

Moreover, the necessary second-order condition implies that:

\[
\frac{\partial \lambda(m_1^*, q)}{\partial m_1^*} = \left[q \nu''(k(h_1, m_1^*)) + (1-q) f''(k(h_1, m_1^*)) \right] \left( \frac{\partial k}{\partial m_1} \right)^2 + \left[q \nu'(k(h_1, m_1^*)) + (1-q) f'(k(h_1, m_1^*)) \right] \frac{\partial^2 k}{\partial m_1^2} < 0.
\]

### 3.3 Dynamics of Health Expenditures

We now provide a simple extension of our model to understand how health expenditures early in life affect health expenditures at older ages. In particular, we now assume that there is also a third period in which the individual is retired. In this third period, health still affects the utility of the individual—because of domestic production, for example. Formally, we assume that the utility of the individual is \( d(h_3) \), with \( d'(\cdot) > 0 \). Moreover, \( h_3 \) evolves according to the following function:

\[
h_3 = \min \left\{ k(h_2, m_2), \bar{h}_3(h_2) \right\},
\]

where \( h_2 \) is the individual’s pre-retirement health, \( m_2 \) are the medical expenditures in period two, and \( k(\cdot, \cdot) \) is the standard health production function with \( \partial k(\cdot) / \partial h_2 > 0 \) and \( \partial k(\cdot) / \partial m_2 > 0 \). The
function $\tilde{h}_3(h_2)$, with $\tilde{h}_3(\cdot) \geq 0$, captures in a reduced-form way the idea that the pre-retirement health stock $h_2$ determines the maximum level of health that can be achieved during retirement. For simplicity, we assume that all expenditures $m_2$ come at no cost to the retiree. This stark assumption captures in a simple way the fact that almost all retirees are covered by Medicare, and, thus, they do not bear the full costs of their medical expenditures.

Given these assumptions, all individuals choose medical expenditures $m^*_2$ so that their health reaches $\tilde{h}_3(h_2)$—i.e.,

$$ k(h_2, m_2^*) = \tilde{h}_3(h_2). \tag{6} $$

Applying the implicit function theorem to equation (6), we obtain that:

$$ \frac{\partial m_2^*}{\partial h_2} = \frac{\partial \tilde{h}_3/\partial h_2 - \partial k/\partial h_2}{\partial k/\partial m_2}. $$

Thus, we have $\partial m_2^*/\partial h_2 < 0$ if and only if $\partial \tilde{h}_3/\partial h_2 - \partial k/\partial h_2 < 0$. A sufficient condition for this assumption to be satisfied is that one’s health potential in retirement $\tilde{h}_3(\cdot)$ is not too sensitive to current health stock $h_2$ (i.e., $\partial \tilde{h}_3/\partial h_2$ is sufficiently small). For example, it is trivially satisfied if $\tilde{h}_3(\cdot)$ is constant.

We can now combine the above discussion with our Proposition 1 and provide the full set of empirical implications that we test in Section 5:

**Proposition 2** If $\partial \tilde{h}_3/\partial h_2 < \partial k/\partial h_2$, then workers in jobs with lower turnover rates have:

(i) a higher medical expenditures $m^*_1$ while working; and

(ii) a lower medical expenditures $m^*_2$ and better health during retirement.

4 Data

We use several distinct data sources in our empirical analysis. In particular, we use the annual Medical Expenditure Panel Survey (MEPS) and the bi-annual Health Retirement Study (HRS) to study the medical expenditures and medical care utilization of working individuals and retirees, respectively. We complement MEPS and HRS with additional variables that we use to construct instruments obtained from the Statistics of U.S. Businesses (SUSB) and a dataset on employment protection laws collected by Autor, Donohue, and Schwab (2006). Furthermore, we use the annual British Household Panel Survey (BHPS) to perform falsification tests on U.K workers. Since all these datasets are publicly available, we describe them only briefly here and refer the reader to their respective websites for a more thorough description.\(^5\)

\(^5\)The MEPS is available at [http://www.meps.ahrq.gov](http://www.meps.ahrq.gov); the HRS is available at [http://hrsonline.isr.umich.edu](http://hrsonline.isr.umich.edu); the SUSB is available at [http://www.census.gov/csd/susb](http://www.census.gov/csd/susb); and the BHPS is available at [http://www.iser.essex.ac.uk/survey/bhps](http://www.iser.essex.ac.uk/survey/bhps).
Medical Expenditure Panel Survey (MEPS)   The Medical Expenditure Panel Survey (MEPS) is a set of large-scale annual rotating (short) panel surveys of families and individuals, their medical providers, and employers across the U.S. It is designed to provide nationally representative estimates of health care use, expenditures, sources of payment, and insurance coverage for the U.S. civilian non-institutionalized population.

MEPS has several components, and the Household Component (HC) serves our purposes. HC provides data from individual households and their members, supplemented by data from their medical providers. HC collects detailed information for each person in the household on demographic characteristics, health conditions, health status, use of medical services, charges and source of payments, access to care, satisfaction with care, health insurance coverage, income, and employment. The public version of the survey reports the one-digit codes of industry and occupation of the individual.\(^6\) In our empirical analysis using synthetic cohorts reported in Section 5.1, we use MEPS data from the 1996-2006 surveys. We have deflated all monetary values using the GDP Implicit Price Deflator, with 2000 as the base year.

Health and Retirement Study (HRS)   The Health and Retirement Study (HRS) began as a panel survey of a nationally representative sample of people aged 51 to 61 in 1992, including their spouses, with over-samples of blacks, Hispanics and Florida residents. This original cohort (wave one) has been re-interviewed every other year since then. In 1998, the sample was supplemented with both older and younger cohorts. Eight waves are currently available.

HRS contains detailed information about current and past health status of respondents, along with rich data on their job history and information about economic and demographic variables, including education, income, and wealth. Beginning with wave three, the survey asks questions about total medical expenditures. In some waves, a continuous value is reported, while others, a series of unfolding bracket questions are asked. Based on these brackets (and some additional variables), the RAND Corporation imputes a continuous value of total medical expenditures in each waves, and this is the main dependent variable that we use in our empirical analysis on workers. We have deflated all monetary values using the GDP Implicit Price Deflator, with 2000 as the base year.

The HRS also asks questions on the individual’s employment history. A respondent is asked about past jobs at his/her first interview. From these questions, the RAND Corporation reconstruct the years of tenure at the longest reported job and the one-digit industry codes of the longest job.\(^7\) Because total medical expenditures were surveyed from 1996 (wave three), we use HRS data from 1996 to 2002 in the analysis of Section 5.2.\(^8\) Furthermore, our sample includes only individuals

\(^6\)The three-digit industry codes are contained in a version restricted from public access.

\(^7\)We construct the three-digit industry code used in Section A.2 from the restricted access HRS data ourselves employing the procedure that the RAND corporation uses for the one-digit industry code.

\(^8\)Only six waves were available when we started this project.
over 66 years of age. This age restriction is dictated by the fact that in each wave, HRS reports the medical expenditures for the previous two years, and we want our individuals to receive identical medical coverage through Medicare.

**Statistics of U.S. Businesses (SUSB)**  The Statistics of U.S. Businesses (SUSB) is a dataset extracted from the Business Register, a file of all known single- and multi-establishment employer companies maintained and updated by the U.S. Census Bureau. The Business Register is the same database that is used to produce County Business Patterns (CBP). SUSB shares some features of the CBP. It provides national and sub-national data on the distribution of economic data by size and industry, reporting the number of establishments, employment, and annual payroll for each geographic-industry-size cell.

More importantly, SUSB reports the number of establishments and corresponding employment change for births, deaths, expansions, and contractions by employment size of enterprise, industry, and state. We use data on establishment deaths to construct our instruments for job turnover in the empirical analysis of Section 5.

**Employment Protection Laws**  During the 1970s and 1980s, the majority of U.S. state courts adopted one or more common-law exceptions to the employment-at-will doctrine that limited employers’ ability to fire employees. Autor, Donohue, and Schwab (2006) presents a detailed dataset of these wrongful-discharge laws prevailing in each state and year for the period from 1972 to 1999, and investigates the effects of these employment protection laws on the labor market. We refer to Autor, Donohue, and Schwab (2006) for a precise definition of all the exceptions and for a thorough discussion of their significance. We use data on one protection—i.e., the implied contract exception—to construct our instruments for job turnover in the empirical analysis of Section 5.2.

**British Household Panel Survey (BHPS)**  The British Household Panel Survey (BHPS) is an annual panel survey beginning in 1991, following about 5,500 households and 10,300 individuals drawn from 250 areas of Great Britain. It is a data set with rich individual-level demographic, social and economic variables, as well as detailed information on health-related issues such as number of doctor visits and self-perceived health status.

### 4.1 Summary Statistics

Table 1 reports summary statistics for the main variables of the three data sets that we use in our analysis.9 Average annual individual medical expenditures are about 1,800 dollars in MEPS, and about 8,300 dollars in HRS. Obviously, individuals are younger in MEPS than in HRS (39 years old versus 75 years old), and they have shorter job tenures (the average current job tenure in MEPS

---

9The standard deviations of the variables in MEPS refer to the standard deviations of the variables used in the cohort analysis of Section 5.1. The standard deviations of individual-level variables are larger.
Table 1: Descriptive Statistics of Key Variables in MEPS and HRS Samples

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Medical Expenditure</td>
<td>1,814 1,574</td>
<td>8,327 24,707</td>
<td>... ...</td>
</tr>
<tr>
<td>Job Tenure</td>
<td>6.7 4.3</td>
<td>... ...</td>
<td>6.1 7.2</td>
</tr>
<tr>
<td>Longest Job Tenure</td>
<td>... ...</td>
<td>23.8 12.7</td>
<td>... ...</td>
</tr>
<tr>
<td>Age</td>
<td>38.9 11.6</td>
<td>75.1 6.8</td>
<td>41.1 12.3</td>
</tr>
<tr>
<td>Yrs. of Education</td>
<td>12.9 1.3</td>
<td>11.9 3.2</td>
<td>11.8 2.4</td>
</tr>
<tr>
<td>Income</td>
<td>$31,403 $11,317</td>
<td>... ...</td>
<td>£ 28,934 £ 22,129</td>
</tr>
<tr>
<td>Total Assets/10,000</td>
<td>... ...</td>
<td>3.34 8.12</td>
<td>... ...</td>
</tr>
<tr>
<td>Male</td>
<td>0.51 0.50</td>
<td>0.51 0.50</td>
<td>0.46 0.49</td>
</tr>
<tr>
<td>White</td>
<td>0.80 0.10</td>
<td>0.85 0.35</td>
<td>0.96 0.18</td>
</tr>
<tr>
<td>Black</td>
<td>0.14 0.10</td>
<td>0.13 0.33</td>
<td>... ...</td>
</tr>
<tr>
<td>Married</td>
<td>0.59 0.20</td>
<td>0.59 0.49</td>
<td>0.60 0.48</td>
</tr>
<tr>
<td>Family (Household) Size</td>
<td>3.23 0.55</td>
<td>1.94 0.93</td>
<td>3.00 1.32</td>
</tr>
<tr>
<td>Union</td>
<td>0.13 0.14</td>
<td>... ...</td>
<td>0.96 0.29</td>
</tr>
</tbody>
</table>

Notes: Medical expenditures and total assets have been deflated to correspond to 2000 dollars. Number of observations vary by variable.

is 6.7 years, while the longest pre-retirement job tenure in HRS is 23.8 years). Other individual characteristics—i.e., education, race, gender, and marital status—are roughly similar across the two data sets.

Overall, we employ multiple, distinct data sets because we are not aware of a single, publicly available data set that both follows many individuals over a long period of time and reports their employment history and their medical expenditures at several points in their lives, including retirement. MEPS and HRS are the best available data sets we know of to investigate how current and past job turnover affects the health expenditures of employed and retired individuals, respectively. In particular, these data sets report very detailed characteristics of the individuals, including the outcome variable that is the focus of our model: health expenditures. This richness of the data implies that we can control for several observed individual characteristics that are often unobserved in other studies. Moreover, the time-series dimension of the data implies that we can construct an empirical model (described in detail in Section 5) that controls for unobserved factors that may simultaneously affect individual labor market choices and health expenditures.
Table 2: A Comparison: Manufacturing versus Retailing

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Manufacturing</td>
<td>Retailing</td>
</tr>
<tr>
<td>Medical Expenditure</td>
<td>1,684</td>
<td>1,580</td>
</tr>
<tr>
<td>Job Tenure</td>
<td>8.65</td>
<td>4.91</td>
</tr>
<tr>
<td>Longest Job Tenure</td>
<td>...</td>
<td>...</td>
</tr>
</tbody>
</table>

Notes: Medical expenditures have been deflated to correspond to 2000 dollars.

4.2 An Illustrative Comparison: Manufacturing versus Retailing

Before proceeding to more-formal tests of our Propositions, this subsection presents simple illustrative patterns comparing the average medical expenditures of individuals across two one-digit industries—manufacturing and retailing—that exhibit substantial differences in average job tenure.

Although Table 2 shows simple averages, it is suggestive of the forces at work. The average tenure is longer in manufacturing jobs, both in MEPS and in HRS data (the one-sided $t$-tests have $p$-values of .00 in both MEPS and HRS data). Interestingly, MEPS data show, as Proposition 1 posits, that the average medical expenditures of people employed in manufacturing is higher than those of individuals employed in retailing: $1,684 and $1,580, respectively (the one-sided $t$-test has a $p$-value of .02). However, HRS data show, as Proposition 2 posits, that the ranking of expenditures is strikingly reversed in retirement: Average expenditures are $7,363 and $8,389, respectively (the one-sided $t$-test has a $p$-value of .04)

While this evidence is clearly not conclusive, these figures seem to uncover patterns consistent with the model. The next subsection develops more-sophisticated empirical strategies to test the implications of the model.

5 Empirical Analysis

In this section, we test the main implications of the model. The empirical analysis closely follows Propositions 1 and 2: Using MEPS data, Section 5.1 investigates how job attachment affects the medical expenditures of employed individuals, and, using HRS data, Section 5.1 analyzes how past job attachment affects the medical expenditures of retired individuals. Section 5.2 performs falsification tests using data from the UK BHPS.
5.1 Health Expenditures of Workers

To investigate the effect of job attachment on individuals’ health investment and health status, we specify the following reduced-form regression equation:

$$ y_{jit} = \beta Z_{jir} + \eta_{rt} + \epsilon_{jirt} $$

$$ = \beta_0 + \beta_T \log(\text{Job Tenure}_{jir}) + \beta_X X_{jir} + \eta_{rt} + \epsilon_{jirt}, \quad (7) $$

where the dependent variable $y_{jit}$ is one of the outcomes of interest (e.g., medical expenditures, doctor visits) for individual $j$ working in industry $i$ in region $r$ in year $t$. The main explanatory variable of interest is (the log of) Job Tenure $jir$: the number of years the individual has been employed in his/her current job; $X_{jir}$ is a large set of control variables—e.g., a cubic polynomial in age, gender, race, years of education, annual income, family size, etc.; $\eta_{rt}$ is a region $r$-year $t$ fixed effect; $\zeta_j$ is an individual fixed effect; and $\epsilon_{jirt}$ is an unobservable component.

The identification of the effect of Job Tenure in equation (7) is challenged by two main issues. The first is selection: People are not randomly allocated to jobs. Individual characteristics induce different people to select into different jobs and different industries. If these characteristics are unobserved—as is the case with ability, discount factor, risk aversion, etc.—and are correlated with Job Tenure, then the estimated coefficient $\beta_T$ will be biased. The inclusion of the fixed effects $\zeta_j$ allows us to control for any fixed, individual unobserved factor and, thus, to control for most of the effects of selection. Nonetheless, there might still be deviations from the fixed effects included in the unobservable $\epsilon_{jirt}$ that are correlated with Job Tenure. The second is reverse causality: Potentially, individuals who invest more in health could be less likely to change jobs. Or, as the literature on “job lock” has highlighted (e.g., Madrian, 1994), individuals without health insurance are more likely to change jobs, and not having health insurance may lead them to spend less on health care.

If the unobservables $\epsilon_{jirt}$ are uncorrelated over time, we can use instruments that exogenously shift Job Tenure to address both of the above issues. More precisely, we could estimate equation (7) by exploiting the panel dimension of MEPS to include individual fixed effects $\zeta_j$ or by first-differencing the data, and by instrumenting for the potentially endogenous Job Tenure.

However, if the unobservables have a persistent component—i.e., $\epsilon_{jt} = \rho \epsilon_{jt-1} + \nu_{jt}$—then fixed-effects or first-differences are not sufficient to eliminate the persistent component of the error term, as $\Delta y_{jt} = \beta \Delta Z_{jt} + \Delta \eta_{rt} + \Delta \epsilon_{jt}$. However, the Arellano and Bond (1991) procedure is specifically designed to handle persistent unobservables in panel data. Arellano and Bond suggest subtracting $\rho y_{jt-1}$ from $y_{jt}$ to eliminate $\epsilon_{jt} - \rho \epsilon_{jt-1}$, leaving only the innovation $\nu_{jt}$ of the unobservable:

$$ y_{jt} = \rho y_{jt-1} + \beta Z_{jt} - \rho \beta Z_{jt-1} + (1 - \rho) \zeta_i + \eta_{rt} - \rho \eta_{rt-1} + \nu_{jt}. \quad (8) $$

Taking first-differences, the following equation obtains:

$$ \Delta y_{jt} = \rho \Delta y_{jt-1} + \beta \Delta Z_{jt} - \rho \beta \Delta Z_{jt-1} + \Delta \eta_{rt} - \rho \Delta \eta_{rt-1} + \Delta \nu_{jt}. \quad (9) $$
In the differenced form, however, the new errors $\Delta \nu_{jkit}$ are correlated with the differenced lagged dependent variable $\Delta y_{jkit-1}$ by construction, and potentially with the variables $\Delta Z_{jkit}$ and $\Delta Z_{jkit-1}$, as well. Therefore, a vector $W$ of instruments is required to construct moments $E(\Delta \nu_{jkit} * W)$, and to estimate equation (9) via GMM. Arellano and Bond use the lagged values $y_{jkit-h}$ and $Z_{jkit-h}$ with lags $h \geq 2$ as instruments for $\Delta y_{jkit-1}$ and $\Delta Z_{jkit-l}$ $l = 0, 1$, respectively, as the new error term $\Delta \nu_{jkit}$ is uncorrelated by construction with lags of order higher than two.\footnote{First-differencing the data introduces serial correlation in the new errors $\Delta \nu_{jkit}$. Arellano and Bover (1995) suggests an alternative procedure that does not introduce serial correlation in the new errors. The procedure—called Orthogonal Deviations—consists of constructing the deviation for each observation from the average of future observations in the sample for the same panel-id. However, this approach does not work with autocorrelated errors, as in equation (9).} These instruments, $y_{jkit-h}$ and $Z_{jkit-h}$ with lags $h \geq 2$, are “mechanically” correlated with the potentially endogenous variables $\Delta y_{jkit-1}$ and $\Delta Z_{jkit-1}$. Hence, following Arellano and Bond, we can use $y_{jkit-h}$ with lags $h \geq 2$ as instruments for the lagged endogenous variable $\Delta y_{jkit-1}$. Moreover, Arellano and Bover (1995) and Blundell and Bond (1998) suggest adding the original equation (8) in levels to the GMM criterion, instrumenting the endogenous variables in levels with first-differences of the instruments.

The SUSB data provide us with instruments for the main endogenous variable—JOB TENURE—that have a stronger economic content than Arellano and Bond’s instruments—i.e., instruments that shift the endogenous variable for plausibly exogenous reasons. In particular, the SUSB data set reports the number and the rate of deaths of establishments in industry $i$ and region $r$, and the number and the rate of workers that lost their jobs due to establishment deaths in industry $i$ and region $r$. Hence, we can use lags (or order higher than two) of these variables as instruments for JOB TENURE. The idea of these instruments is that they are clearly correlated with average JOB TENURE in industry $i$ and region $r$. However, establishment deaths do not directly affect voluntary separations, so reverse causality is not a concern. Moreover, lags of the instruments $Z_{jkit-h}$ with lags $h \geq 2$ purge any undesired correlation with $\Delta \nu_{jkit}$, the first difference in the innovation in the unobservables.\footnote{As Blundell and Bond (1998) demonstrates, the Arellano and Bond procedure does not work well if the dependent variables are very persistent. However, this does not appear to be a concern in our case, as our dependent variables exhibit year-to-year variation. Similarly, our main explanatory variable, JOB TENURE, and its instruments (death rate of establishments and the fraction of workers that lost their jobs due to establishment deaths) exhibit substantial year-to-year variation.}

Unfortunately, the panel component of MEPS is too limited (two years) to use the Arellano-Bond procedure on individual data. Since the procedure has the very attractive feature that it allows us to control for persistent unobserved heterogeneity, we use MEPS data to construct synthetic panels. As in all papers that use synthetic panels, the definition of a cohort is arbitrary. In our case, we are constrained by the sample size of each MEPS survey and by the limited geographic and industry information available in the public version of the MEPS. As a result, we choose to define...
cohorts by grouping people by sex, decade of birth, one-digit industry, and Census Region. With a slight abuse of notation, we can, thus, write the cohort version of the empirical model defined by equations (7) and (9) as:

\[ y_{jt} = \beta_0 + \beta_T \text{Job Tenure}_{jt} + \beta_X X_{jt} + \eta_{rt} + \zeta_j + \epsilon_{jt} \]

\[ \Delta y_{jt} = \rho \Delta y_{jt-1} + \beta \Delta Z_{jt} - \rho \beta \Delta Z_{jt-1} + \Delta \eta_{rt} - \rho \Delta \eta_{rt-1} + \Delta \nu_{jt}, \]

where the subscript \( j \) now denotes a cohort, for which industry \( i \) and region \( r \) are fixed over time. The dependent variable \( y_{jt} \) is again one of the outcomes of interest for cohort \( j \) (working in industry \( i \) in region \( r \)) in year \( t \). Similarly, \( \text{Job Tenure}_{jt} \) is the average number of years individuals in cohort \( j \) have been employed in their current firm. \( X_{jt} \) is now the cohort-average of a large set of control variables: the average age of individuals in the cohort, age squared, age cubed, average education, annual income, annual income squared, size of the family, fraction of whites, and fraction of blacks. \( \eta_{rt} \) is, as before, a year fixed effect for each region \( r \). \( \zeta_j \) is now a fixed effect for cohort \( j \) (again, working in industry \( i \) in region \( r \)). \( \epsilon_{jt} \) is an unobservable, autoregressive component with innovation \( \nu_{jt} \)—i.e., \( \epsilon_{jt} = \rho \epsilon_{jt-1} + \nu_{jt} \).

In summary, our empirical model deals with selection through the use of the Arellano and Bond (1991) procedure and with reverse causality through the use of demand-side instruments. Moreover, the use of these demand-side instruments implicitly assumes that the employer-employee pair forms expectations about plant closures that are correlated with their realizations, so that expected turnover and actual turnover generated by these plant closures covary.

**Results on Medical Expenditures.** Panel A in Table 3 presents the results for the (log of) medical expenditures. The specification of column (1) assumes that the unobservables \( \epsilon_{jt} \) are serially uncorrelated. The specification of column (2) employs the specification of equation (11) that explicitly takes into account serial correlation in the unobservables \( \epsilon_{jt} \).

The results reported in Columns (1) and (2) in Panel A are similar. The coefficients of \( \log(\text{Job Tenure}) \) in both specifications are positive (and significant at the one-percent level). Thus, these coefficients are consistent with the idea that an employer-employee pair invests more in the employee’s health when the employee’s expected turnover is lower, as predicted by Proposition 1. Moreover, the economic significance of the effects is rather large. The estimated coefficients of \( \log(\text{Job Tenure}) \) in specifications (1) and (2) imply that increasing \( \text{Job Tenure} \) by ten percent increases the annual medical expenditures of workers by about seven percent, a rather large effect.

**Results on Doctor Visits.** We further investigate whether individuals are more likely to visit a doctor when job turnover is lower. These additional regressions have three goals. First, total medical expenditures include many different types of expenditures, and it is useful to check whether the results apply also to a more-narrow and basic category of health care. Second, the price of
Table 3: The Relationship Between Workers’ Job Tenure and Medical Expenditures (Panel A) and Doctor Visits (Panel B)

<table>
<thead>
<tr>
<th></th>
<th>Panel A: Log Medical Expenditures</th>
<th>Panel B: Fraction Not Visiting Doctors</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
</tr>
<tr>
<td>Log (Job Tenure)</td>
<td>0.698***</td>
<td>0.686***</td>
</tr>
<tr>
<td></td>
<td>(0.248)</td>
<td>(0.256)</td>
</tr>
<tr>
<td>Age</td>
<td>-0.385***</td>
<td>-0.464*</td>
</tr>
<tr>
<td></td>
<td>(0.065)</td>
<td>(0.273)</td>
</tr>
<tr>
<td>Age Squared</td>
<td>0.009***</td>
<td>0.010</td>
</tr>
<tr>
<td></td>
<td>(0.002)</td>
<td>(0.007)</td>
</tr>
<tr>
<td>Education</td>
<td>0.236***</td>
<td>0.190***</td>
</tr>
<tr>
<td></td>
<td>(0.018)</td>
<td>(0.023)</td>
</tr>
<tr>
<td>Income/10,000</td>
<td>0.005</td>
<td>0.007</td>
</tr>
<tr>
<td></td>
<td>(0.060)</td>
<td>(0.064)</td>
</tr>
<tr>
<td>Male</td>
<td>-1.465</td>
<td>-1.115***</td>
</tr>
<tr>
<td></td>
<td>(0.042)</td>
<td>(0.059)</td>
</tr>
<tr>
<td>Married (Fraction)</td>
<td>0.445</td>
<td>0.208***</td>
</tr>
<tr>
<td></td>
<td>(0.118)</td>
<td>(0.133)</td>
</tr>
<tr>
<td>Family Size</td>
<td>-0.160***</td>
<td>-0.182***</td>
</tr>
<tr>
<td></td>
<td>(0.035)</td>
<td>(0.041)</td>
</tr>
<tr>
<td>Union (Fraction)</td>
<td>0.054</td>
<td>0.295</td>
</tr>
<tr>
<td></td>
<td>(0.164)</td>
<td>(0.192)</td>
</tr>
<tr>
<td>ρ</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td></td>
<td>(0.030)</td>
<td>...</td>
</tr>
<tr>
<td># Obs</td>
<td>4652</td>
<td>3930</td>
</tr>
<tr>
<td>Panels</td>
<td>620</td>
<td>586</td>
</tr>
</tbody>
</table>

Notes: (1). Columns 1 and 3 report Arellano-Bond IV regressions assuming iid errors, while Columns 2 and 4 assume AR(1) errors. (2). All regressions also contain Age Cubed, Income Squared, Firm Size, Race and year fixed effects (not reported). (3). Standard errors in parentheses are calculated by applying the finite sample correction proposed by Windmeijer (2005) and are robust to autocorrelation and heteroskedasticity of unknown form. (4). *, **, *** denote significance at ten, five and one percent, respectively.
medical expenditures may differ across individuals, and doctor visits offer a *quantity* of services acquired. Third, in Section 5.4, we compare patterns of health investment between the U.S. and the U.K., and the UK data—like data from many countries with a national health system—report only the quantity of medical services acquired, not the expenditures.

Thus, we estimate a few regressions in which the dependent variable is equal to the fraction of people in the cohort who report that they *did not* visit a doctor in the last year—i.e., they had zero doctor visits. Panel B of Table 3 reports the estimated coefficients. The specification of column (3) assumes that the unobservables $\epsilon_{jt}$ are serially uncorrelated. Column (4) employs the specification of equation (9) that explicitly takes into account serial correlation in the unobservables $\epsilon_{jt}$. The estimates in Columns (3) and (4) are, again, very similar. The estimated coefficients imply that increasing job tenure by ten percent decreases the probability of visiting a doctor at least once per year by 1.1 percentage points, which represents an approximately three-percent decrease from the average sample probability of not visiting a doctor, equal to .39.

Overall, the results reported in Table 3 provide strong evidence consistent with the prediction of Proposition 1.

### 5.2 Health Expenditures of Retirees

The analysis of retirees’ medical expenditures follows as closely as possible the previous analysis of workers’ medical expenditures. Nonetheless, some slight modifications are necessary because of the different data, and we now describe them in detail. Our reduced-form equation now reads:

$$y_{jt} = \beta_0 + \beta_T \log(PAST\ TENURE_{jt}) + \beta_X X_{jt} + \eta_t + \zeta_j + \epsilon_{jt},$$  \hspace{1cm} (12)

where the dependent variable $y_{jt}$ is one of the outcomes of interest (e.g., medical expenditures, health status) of individual $j$ in year $t$. The main explanatory variable of interest is (log of) $PAST\ TENURE_{jt}$: the number of years of the longest job tenure. $X_{jt}$ is again a set of control variables: a cubic polynomial in the age of the individual; years of education; total assets; total assets squared; size of the family; gender; and race. $\eta_t$ is a year fixed effect, $\zeta_j$ is an individual effect, and $\epsilon_{jt}$ is an unobservable component.

Since the HRS is a panel data set, we could potentially control for unobserved heterogeneity with individual fixed-effects. However, the HRS sample is composed mainly of retired individuals, and, thus, the main variable of interest—i.e., $PAST\ TENURE$—has almost no within-panel variation. Hence, we cannot include individual fixed-effects. Nevertheless, since almost all individuals in the HRS are retired, their labor market histories are clearly determined many years before the sample period (1996-2002). This temporal lag implies that period-$t$ shocks that may affect medical expenditures should not be correlated with past labor market history. Moreover, this temporal lag

---

12We had to drop the fixed effect for the fifth wave because it was creating problems in the estimation.
helps us in the construction of the instruments and, thus, in the identification of the effect of Past Tenure on medical expenditures in equation (12).

In particular, the HRS reports the Census division of birth of each individual. Our instruments exploit the variation in employment protection across Census divisions and the variation in the effects of protection across workers with heterogenous, predetermined characteristics—i.e., education, gender, age. More precisely, Autor, Donohue, and Schwab (2006) investigates the labor-market impacts of wrongful-discharge protections adopted by U.S. state courts between 1972-1999. The authors find that one doctrine—i.e., the implied contract exception—reduced employment and that the short-term impact was most pronounced for demographic subgroups that change jobs most frequently: females, younger, and less-educated workers. Thus, arguably, the implied contract exception affected Past Tenure, and the effect was different according to individual characteristics, such as education, gender, age. Autor, Donohue, and Schwab (2006) constructs an annual panel, reporting whether each state implemented the implied contract exception. Since the HRS reports the Census division, we take the average of all states within a Census division for the years 1980 and 1990, and we further construct interactions of these two instruments with: the age of the individual in 1980 and 1990, respectively; a binary indicator equal to one if the individual reports more than 13 years of education, and zero otherwise; and a binary indicator equal to one if the individual is a male, and zero otherwise.

**Results on Medical Expenditures.** Panel A of Table 4 presents the results for the (log of) medical expenditures of individuals in the HRS sample. These results are remarkable. As Proposition 2 predicts, the coefficient of log(Past Tenure) in column (1) is negative and significant. Moreover, the economic significance of the coefficient is large. The magnitude of the coefficient of log(Past Tenure) in column (1) means that increasing Past Tenure by ten percent decreases the annual medical expenditures by approximately 7.5 percent.

**Results on Health Status.** We further investigate whether retirees with higher past job attachment report better health. Panel B of Table 4 reports the coefficient of a regression in which the dependent variable is equal to one if the retiree reported being in the lowest two categories of self-reported health (i.e., fair and poor), and zero otherwise. The coefficient estimate of log(Past Tenure) is negative and statistically significant at one-percent level. The magnitude of the coefficient indicates that the probability that a retiree reports poor health decreases by approximately four percentage points when his job tenure prior to retirement increases by ten percent.

In summary, our analysis indicates that retirees with a longer past job tenure have a lower medical expenditures and are in better health, consistent with the predictions of Proposition 2 of our model.
Table 4: The Relationship Between Retirees’ Past Job Tenure and Medical Expenditures (Panel A) and Health Status (Panel B)

<table>
<thead>
<tr>
<th>Panel A: Log Medical Expenditure (1)</th>
<th>Panel B: Health Status (2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Log (Job Tenure)</td>
<td>-0.746** (0.356)</td>
</tr>
<tr>
<td>Age</td>
<td>0.222 (0.442)</td>
</tr>
<tr>
<td>Age Squared</td>
<td>-0.001 (0.005)</td>
</tr>
<tr>
<td>Education</td>
<td>0.033*** (0.010)</td>
</tr>
<tr>
<td>Total Assets/1,000,000</td>
<td>-0.052 (0.033)</td>
</tr>
<tr>
<td>Male</td>
<td>0.441*** (0.169)</td>
</tr>
<tr>
<td>Married</td>
<td>-0.066** (0.033)</td>
</tr>
<tr>
<td>Household Size</td>
<td>0.011 (0.016)</td>
</tr>
<tr>
<td># Obs</td>
<td>27,229</td>
</tr>
<tr>
<td>Panels</td>
<td>10,395</td>
</tr>
</tbody>
</table>

Notes: (1) The dependent variable for Column 2 is a dummy variable taking value 1 if the retiree reported being in the lowest two categories of self-reported health (i.e., fair and poor). (2). The estimated equations also contain: Age Cubed; Total Assets Squared; Year, Race, and Census Division fixed effects. Their coefficients are not reported. (3). Standard errors in parentheses. (4) *, **, *** denote significance at ten, five, and one percent, respectively.

5.3 Assessing the Magnitude of Life-Cycle Inefficiency

Our previous analysis shows that the differences in medical expenditures among workers and among retirees with different job tenures are large. Moreover, the analysis reveals a stark reversal in medical expenditures: Individuals with higher expenditures during their working years have lower expenditures during retirement. In this section, we combine the previous estimates of workers’ and retirees’ expenditures and seek to quantify the dynamic externality that lies at the heart of this paper. More precisely, we wish to compare the lifetime expenditures of two workers A and B whose only difference is their job tenures. We want to be very clear that calculating the exact size of the externality implied by our regressions is very complicated, in particular because our main variable of interest, job tenure, is measured at two different points in time. Nevertheless, we try to perform a simple back-of-the-envelope calculation.

Suppose that both individuals work for 45 years and then retire for 15 years before dying. Individual A works in a job in which mobility is high, while individual B works in a job in which mobility is low. For example, let us assume that individual B’s job tenure is one standard deviation higher than that of individual A. In MEPS, one standard deviation of log tenure is equal to .52. Multiplying it by the coefficient of log tenure in the MEPS regressions, we obtain .7*.52=36 percent.
At the average of MEPS medical expenditures ($1,814), this implies that that individual A has expenditures lower than B’s by approximately $660 per year.

Let us now consider both individuals’ medical expenditures during retirement. In particular, let us assume that the cross-sectional difference in job tenure in MEPS data carries over and, thus, one standard deviation in log tenure in MEPS translates into one standard deviation in log tenure in HRS. One standard deviation of log tenure in the HRS data is equal to .77. Multiplying it by the coefficient of log tenure in the HRS regressions, we obtain .74*.77=56 percent. At the average of HRS medical expenditure ($8,327), this implies that individual A has expenditures higher than B’s by approximately $4,700 per year.

Thus, if individuals A and B work for 45 years and then retire for 15 years, non-discounting their expenditures, we have that, during their working years, individual A’s health expenditures are approximately $29,700 lower than individual B’s. And, during retirement, individual A’s health expenditures are approximately $70,500 higher than individual B’s. The total difference is around $40,000, a rather large difference. This calculation suggests that one additional dollar of health expenditures during the working years may lead to about 2.5 dollars of savings in retirement. Obviously, this is a very rough calculation that neglects many important factors, such as mobility between low-attachment to high-attachment jobs. Moreover, it comes from two different data sets, and not from a single, long panel. In addition, on one side, it neglects discounting, but on the other side, it also neglects that the price of medical care has been rising more than the interest rate. Furthermore, it neglects any effect on quality of life and mortality. In summary, we believe that this calculation neglects many other factors that are important in assessing the full effect of the externality. Nonetheless, we believe it describes in a very simple way the externality we have in mind, and its magnitude in the data.

5.4 Falsification Test: UK Workers

In this section, we investigate whether individuals in the U.K.—a country with a national health system—exhibit the same patterns that we documented for individuals in the U.S. We believe that this is a useful comparison: While the employment-based health care system is unique to the U.S., wages and turnover patterns are very similar across many developed countries (Katz, Loveman and Blanchflower, 1995). Hence, we conduct “falsification” tests by replicating as closely as possible some of the analysis of Section 5.1 using data from the British Household Panel Survey (BHPS), a data set that reports quite-detailed information on individual labor market histories, along with some information on health-related issues.

Unfortunately, we cannot directly investigate the relationship between medical expenditures and job tenure, as in Panel A of Table 3. The reason is that BHPS does not report total medical expenditures at the individual level. Indeed, we are not aware of any non-U.S. data set compa-
rable to MEPS—i.e., a dataset that collects information about total medical expenditures at the individual level. Nonetheless, the BHPS reports the number of doctor visits for a sample of U.K. individuals. This allows us to conduct a falsification test by replicating the analysis of doctor visits for U.S. individuals, as we reported in Panel B of Table 3. If the employment-based health care system in the U.S. is responsible for the relationship we documented in Table 3, then we should not expect the number of doctor visits of U.K. workers to have the same relationship with job tenure.

We implement this falsification test in a panel regression in which the dependent variable is a binary variable equal to one if the individual did not visit a doctor in the last year, and zero otherwise. The specification is as close as possible to the specification of Panel B of Table 3, with the additional advantage that we can directly use individual panel techniques since the BHPS is a long panel data set. Moreover, the BHPS reports rich individual-level variables that allow us to construct instruments for the potentially endogenous variable \textit{job tenure}. More precisely, the BHPS reports the district of birth of the individual, so that we use as instruments for (the log of) \textit{job tenure} of individual \(i\) (the log of) the average tenure of all individuals of the same sex born in the same five-year window and in the same district as individual \(i\). The basic idea of the instruments is that the district of birth is obviously exogenous to the individual, as is its industrial composition, for example. However, continuing the example, the industrial composition of the district of birth affects the skills that the individual accumulates (say through inter-generation transmission of human capital or the type of schooling). Therefore, individuals are more likely to work in industries that are popular in the labor market of their area of birth. Through such a mechanism, the longest job tenure of each individual is correlated with the job tenure of all individuals born in the same district. However, the instruments are plausibly uncorrelated with individual ability, the main unobserved individual effects that may simultaneously determine health investment and health expenditures, as we discussed in Section 5.1. Moreover, since the career choices of men and women are quite different, we separate the instruments by gender.

Table 5 reports the results of several different specifications. All specifications show that individuals with a longer job tenure are not more likely to visit a doctor in the U.K., in sharp contrast with the U.S. evidence reported in Panel B of Table 3. Moreover, this result is robust to several different methodologies—instruments with fixed effects (column (1)); instruments with first-differences (column (2)); Arellano-Bond procedure with i.i.d. residuals (column (3)); and Arellano-Bond procedure with autoregressive residuals (column (4)). The coefficient of the log of job tenure is not statistically significant in any of these specifications.

fraction of total medical expenditures in a national health insurance system.
Table 5: Falsification Test: The Relationship Between Workers’ Job Tenure and Doctor Visits in the U.K.

<table>
<thead>
<tr>
<th></th>
<th>IV with Fixed Effect</th>
<th>IV with First Difference</th>
<th>Arellano-Bond with iid Residual</th>
<th>Arellano-Bond with AR(1) Residual</th>
</tr>
</thead>
<tbody>
<tr>
<td>Log (Job Tenure)</td>
<td>-0.011</td>
<td>-0.021</td>
<td>-0.023</td>
<td>0.111</td>
</tr>
<tr>
<td></td>
<td>(0.019)</td>
<td>(0.021)</td>
<td>(0.078)</td>
<td>(0.101)</td>
</tr>
<tr>
<td>Age</td>
<td>0.016*</td>
<td>0.025</td>
<td>0.008</td>
<td>0.021</td>
</tr>
<tr>
<td></td>
<td>(0.009)</td>
<td>(0.020)</td>
<td>(0.006)</td>
<td>(0.048)</td>
</tr>
<tr>
<td>Age Squared</td>
<td>0.000</td>
<td>-0.000</td>
<td>-0.000</td>
<td>-0.000</td>
</tr>
<tr>
<td></td>
<td>(0.001)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.001)</td>
</tr>
<tr>
<td>Education</td>
<td>...</td>
<td>...</td>
<td>(0.003)</td>
<td>(0.002)</td>
</tr>
<tr>
<td>Income/10,000</td>
<td>-0.002</td>
<td>-0.001</td>
<td>0.004***</td>
<td>-0.001</td>
</tr>
<tr>
<td></td>
<td>(0.001)</td>
<td>(0.002)</td>
<td>(0.002)</td>
<td>(0.003)</td>
</tr>
<tr>
<td>Male</td>
<td>...</td>
<td>...</td>
<td>0.156***</td>
<td>0.134***</td>
</tr>
<tr>
<td></td>
<td>...</td>
<td>...</td>
<td>(0.005)</td>
<td>(0.005)</td>
</tr>
<tr>
<td>Married</td>
<td>0.000</td>
<td>0.007</td>
<td>-0.001</td>
<td>-0.006</td>
</tr>
<tr>
<td></td>
<td>(0.007)</td>
<td>(0.010)</td>
<td>(0.005)</td>
<td>(0.009)</td>
</tr>
<tr>
<td>Household Size</td>
<td>0.002</td>
<td>-0.002</td>
<td>0.002</td>
<td>-0.000</td>
</tr>
<tr>
<td></td>
<td>(0.002)</td>
<td>(0.003)</td>
<td>(0.002)</td>
<td>(0.003)</td>
</tr>
<tr>
<td>Union</td>
<td>-0.007</td>
<td>-0.005</td>
<td>-0.007</td>
<td>-0.012</td>
</tr>
<tr>
<td></td>
<td>(0.006)</td>
<td>(0.005)</td>
<td>(0.008)</td>
<td>(0.011)</td>
</tr>
<tr>
<td>$\rho$</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>0.130***</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.007)</td>
</tr>
<tr>
<td># Obs</td>
<td>93,709</td>
<td>95,955</td>
<td>94,015</td>
<td>75,955</td>
</tr>
<tr>
<td>Panels</td>
<td>15,931</td>
<td>15,760</td>
<td>16,237</td>
<td>15,760</td>
</tr>
</tbody>
</table>

Notes: (1). The dependent variable in each specification is a dummy variable taking value one if the individual did not visit a doctor in the last year, and zero otherwise. (2). All regressions also include Age Cubed, Income Squared, Race, year fixed effects and geographic region fixed effects. Their coefficients are not reported. (3). Standard errors in parentheses; and *, **, *** denote significance at ten, five and one percent, respectively.
6 Alternative Hypotheses

The results of the previous empirical analysis provide strong evidence that job turnover affects health expenditures, as Propositions 1 and 2 predict. We now consider several alternative hypotheses. For most of these alternatives, we discuss how our empirical model allows us to distinguish the implications of our model against other plausible explanations.

**Good Jobs versus Bad Jobs.** Several papers document true wage differentials across industries and jobs and a negative correlation between wage differentials and quit rates (e.g., Pencavel, 1970; Krueger and Summers, 1988; Gibbons and Katz, 1992). If workers are less likely to leave “good jobs” than “bad jobs”—good jobs offer higher wages and richer benefits, including health insurance—and if health insurance and health expenditures are related—in the data they are related—then differences between good jobs and bad jobs could imply a positive correlation between job attachment and health expenditures.

However, as we described in detail in Section 5.1, our empirical model on workers’ medical expenditures is designed to precisely control for fixed and for persistent unobserved effects that may induce different workers to select into different jobs/industries. Moreover, the instruments that we employ in the empirical analysis on workers’ medical expenditures exploit demand-side (i.e., firms) variation in turnover across regions and industries, precluding any reverse causality hypothesis based on supply-side (i.e., workers) variation in quits. Similarly, the instruments we use in the empirical analysis on retirees’ health expenditures exploit variations in wrongful-discharge protections across individuals’ regions of birth, and the heterogenous effect of these protections across individuals with heterogenous, exogenous characteristics. As Autor, Donohue, and Schwab (2006) documents, these employment protections appear to have shifted the demand for labor, rather than the supply.

**What about job-lock?** An influential literature shows that the employment-based health insurance system provides inefficiently low separation between mismatched workers and firms (Madrian, 1994; Gruber and Madrian, 1994, 1997, 2002). This “job-lock” literature postulates that workers are less likely to leave jobs that offer health insurance than comparable jobs without health insurance. If health insurance and health expenditures are related—and in the data they are related—these hypotheses could also imply a positive correlation between job attachment and health expenditures.

However, we emphasize that our empirical analysis is designed to circumvent any reverse causality hypothesis. Moreover, if job-lock were the only mechanism at work in the data, we would expect individuals with worse health to select jobs with more generous health benefits since less-healthy workers presumably benefit more from generous health benefits. Thus, in steady state, we should expect to find less-healthy worker in jobs with lower turnover. Instead, Panel B of Table 4 showed us that the opposite is true: Healthier individuals were working in lower turnover jobs.
In summary, we believe that wage differentials and job-lock are well-suited to addressing the question of why mobility differs across individuals and jobs. However, we think that wage differentials and job-lock cannot explain the empirical patterns in health care expenditures that are the focus of our paper. Most likely, they are valid explanations for complementary facts, but do not provide alternative interpretations of all the empirical findings in this paper.

Is health more important in jobs that have also higher attachment? The empirical relationship between job attachment and health expenditures could simply be due to the fact that health is more important in industries that have also higher job attachment.

However, the evidence from U.K. workers does not substantiate this claim. The results reported in Table 5 indicate that U.K. workers with longer job tenures are not more likely to visit a doctor, in sharp contrast to the results for U.S. workers reported in Panel B of Table 3. This difference in health care utilization between U.K. and U.S. is also in stark contrast with many labor-market patterns—i.e., wages and inequality—that are remarkably similar in the two countries (Katz, Love- man and Blanchflower, 1995; Gosling and Lemieux, 2004).

What about myopic workers? A potential explanation of our empirical findings has to do with different wage profiles. If, in high turnover jobs, wage profiles are flatter (high earlier, lower later), more myopic people choose these jobs, attracted by the higher initial wage. These people are likely to have a different intertemporal discount—i.e., they value today much more than tomorrow. This could explain why their health expenditures are lower.

However, this explanation is in sharp contrast with current theories of human capital. General human capital steepens wage-tenure profiles because workers must pay, in the form of lower wages, for any training that is general and, thus, transferable across employers. Early in their career, workers receive investment in human capital and lower wages. When human capital begins to increase productivity later in their career, workers have higher earnings. Because general human capital is transferable, firms must pay workers their full marginal product in the post-investment period. Conversely, any type of specific human capital flattens wage-tenure profiles because the firm makes a specific investment, but recoups its investment later, once the workers are locked in. Indeed, this is exactly what the extension of our model presented in Appendix A.1 predicts.

Moreover, in high turnover jobs, the relative importance of specific human capital is presumably lower than that of general human capital. As a result, we should expect high-turnover jobs to have steeper, not flatter, wage profiles (as this alternative explanation needs). Indeed, Crocker and Moran (2003) provides empirical evidence consistent with these predictions of human capital theories. In the wage regressions reported in Table (2) of their paper, they find that returns to tenure are higher in high-turnover jobs.
Is it a pure wealth effect? Another potential alternative explanation is a pure wealth effect. If wages are higher in low-turnover jobs, then a simple wealth effect might explain why health expenditures are higher in low-turnover jobs. Indeed, Hall and Jones (2006) argues that the growth of health spending in the past half-century is a rational response to the growth of income per person. According to their model, health spending is a superior good with an income elasticity well above one.\textsuperscript{14}

Clearly, our explanation and that of Hall and Jones are not mutually exclusive. Hall and Jones focus on the growth of expenditures in the last 20 years, while we focus on the intertemporal profile of expenditures. However, we believe that the wealth effect cannot fully explain a number of our cross-sectional results. First, all our regressions on workers’ medical expenditures include workers’ current income and the best proxy for permanent income—i.e., education. Moreover, in the regressions on retirees’ medical expenditures, we find that the coefficient of total assets is negative, and not significant. Thus, we have no evidence that wealthier retired individuals spend more on health.

7 Related Literature

Our paper is related to several strands of the literature. The first is the literature on the interactions between health care markets and labor markets. Several papers examine how employer-provided health insurance may lead workers to keep jobs they would rather leave, for fear of losing insurance coverage for preexisting conditions (Madrian, 1994; Gruber and Madrian, 1994, 1997, 2002). Our paper complements these studies by focusing on a related, but conceptually different link between the health care market and the labor market in the U.S.

In particular, as in Grossman’s (1972) seminal contribution and a number of more recent papers (Murphy and Topel, 2005; Hall and Jones, 2007), our paper focuses on the consumption/investment aspect of health care versus the insurance aspect. Standard health-care contracts bundle regular medical care—i.e., care for frequent and common treatment—with pure insurance—i.e., protection against low-probability and high-cost events.\textsuperscript{15} However, in contrast to most papers in this strand of the literature, our paper delves deeper into the incentives generated by the institutional arrangements that govern health care, especially employer-provided health insurance, and the interaction between private and public insurance.

\textsuperscript{14}On the other side, Acemoglu, Finkelstein and Notowidigdo (2009) use oil price shocks and cross-sectional variation in the oil reserves across different areas of the U.S. and find that the income elasticity of health expenditures is almost always less than one.

\textsuperscript{15}The recent introduction of health savings accounts (HSA), in part, breaks the link between consumption and insurance. HSAs are tax-favored savings accounts that can be used to pay for medical expenditures, combined with high-deductible health insurance plans. According to the U.S. Department of the Treasury (\url{http://www.ustreas.gov/offices/public-affairs/hsa/}), in 2005, 3.2 million individuals were already covered by HSA-type insurance, and the figure is projected to be 25 to 30 million people in the year 2010.
The second strand of the literature related to our paper is on general and specific human capital. The connection to this literature is immediate because we consider health as a form of general human capital. Becker’s (1962, 1964) classical theory of human capital distinguishes between investments in general and specific human capital based on the transferability of the acquired skills when a worker switches employers. To the extent that healthy workers are more productive in all firms, health investment is quite plausibly a form of general investment (Grossman, 2000).

One of the most celebrated results in the classical theory of human capital is that in a frictionless and competitive labor market, workers capture all the returns from their general human capital investments (Becker, 1962, 1964). Thus, workers pay for the entire costs of general human capital investments, as employers obtain no return from them, and invest the efficient amount. The empirical observation that firms seem to pay for general training of their employers—in contrast to the predictions of the classical human capital theory—has stimulated a few recent theoretical explanations. Acemoglu and Pischke (1998, 1999) show that, when labor-market frictions lead to “wage compression,” then firms may pay for investments in the general skills of their employees. The compression in the wage structure transforms the “technologically” general skills into de facto “specific” skills, thus providing firms with incentives to invest in their workers’ general skills. Even though Acemoglu and Pischke’s theoretical models also yield testable predictions about the level of general human capital investment, most of the literature has focused exclusively on why firms share the costs of general training.

The third strand of the literature related to our paper examines the interactions between public and private insurance (Cutler and Gruber, 1996; Brown and Finkelstein, 2008). Most of these papers focus on how public insurance programs crowd out the demand for private insurance. Thus, while these papers considers the contemporaneous interaction between the public and private insurances, we focus on the intertemporal interactions and on health investment rather than on insurance.

The fourth strand of literature is on the dynamic inefficiency in the insurance market (Hendel and Lizzeri, 2003; Crocker and Moran, 2003; Finkelstein, McGarry and Sufi, 2005, Herring, 2006; and Cebul et. al., 2008). Hendel and Lizzeri (2003) Crocker and Moran, (2003), and Finkelstein, McGarry and Sufi (2005) consider a different inefficiency from the one that we highlight. In particular, these papers suggest that inefficient risk-sharing arises when parties do not commit to long-term insurance contracts since short-term contracts cannot insure the reclassification risk.

---

*16* In contrast, employers and employees share the returns and the costs of investment in firm-specific skills. See, for example, Hashimoto (1981).

*17* Acemoglu and Pischke (1998, 1999) considers many potential forms of market frictions, including search friction, asymmetric information, complementarity between general and specific skills, etc.

*18* Recent papers by Balmaceda (2005) and Kessler and Lülfesmann (2006) show that, under some surplus sharing rules, specific and general human capital endogenously interact. Thus, even if the labor market is competitive, an employer may choose to contribute to workers’ general training.
that arises from a change in insurees’ risk type.\textsuperscript{19} To some extent, the dynamic inefficiency in our analysis is also related to the inability of workers to commit to long-term employment with the firm. Perhaps the closest paper to ours is Crocker and Moran (2003), which argues that workers in industries with higher specific-skill requirements are more committed to their firms, thereby allowing their employers to provide more complete insurance of health risks. In contrast, we focus on health investment and health outcomes.

Herring (2006) and Cebul \textit{et. al.} (2008) consider a type of inefficiency similar to the one that we focus on. Herring (2006) argues that insurees’ turnover may reduce an insurer’s incentives to provide the socially-optimal level of preventive care. Using data from the Community Tracking Study’s Household Survey, Herring finds that insurers’ turnover has a negative effect on the utilization of preventive services. Cebul \textit{et. al.} (2008) studies the effect of search frictions in the market for employer-based health insurance and makes the point that frictions in insurance markets may reduce incentives to invest in future health. While clearly complementary, there are several crucial differences between our paper and Herring (2006) and Cebul \textit{et. al.} (2008). First, we focus on employees’ turnover rates, while Herring (2006) and Cebul \textit{et. al.} (2008) focus on enrollees’ turnover among insurance companies. Employers and workers enjoy most of the costs and benefits of medical expenditures, so we believe it is appropriate to focus on them. Moreover, almost half of all firms and more than 80 percent of firms with more than 5,000 employees are self-insured (Barr, 2007, pp. 84). For these self-insured firms, insurers only administer employers’ claims. Thus, it seems that workers’ turnover is what crucially affects the incentives to invest in health. Second, we examine the dynamics of medical expenditures over an individual’s life-cycle. Specifically, we investigate how retirees’ medical expenditures and health status are related to their turnover rates prior to retirement. This allows us to understand the order-of-magnitude of the dynamic externality that we focus on.

\section*{8 Conclusion}

In this paper, we investigated how the employment-based health insurance system in the U.S. affects individuals’ life-cycle health-care decisions. We take the viewpoint that health is a form of human capital that affects workers’ on-the-job productivity, and derive implications of employees’ turnover on the incentives to undertake health investment. Our model suggests that employees’ turnover leads to dynamic inefficiencies in health investment. In particular, it suggests that the employment-based health insurance system may lead to an inefficient, low level of individual health investment during individuals’ working lives. Moreover, we show that underinvestment in health is higher

\textsuperscript{19}Diamond (1992) mentioned that the lack of long-term health insurance is an important market failure. Cochrane (1995) showed that time-consistent health insurance contracts \textit{with severance payments} can fully insurance consumers with the reclassification risk with strings of short-term contracts that may provide coverage for reclassification risk with short-term contracts. See, also, Pauly, Kunreuther and Hirth (1995).
when workers’ turnover rate is higher, and it increases medical expenditures during retirement.

We present a model that makes this process explicit and then investigate its empirical relevance using data from the Medical Expenditure Panel Survey and the Health and Retirement Survey. We document a large number of empirical patterns, all consistent with our model. Moreover, the magnitude of our estimates suggests a significant degree of intertemporal inefficiencies in health investment in the U.S. Our back-of-the-envelope calculations suggest that, on average, one dollar of medical expenditures during the working years may decrease medical expenditures during retirement by about 2.5 dollars.

APPENDIX

A Specific Capital

This Appendix has two goals: 1) to extend the model of Section 3 to allow for endogenous turnover; and 2) to present the results of an additional empirical strategy that closely follows the extension of the model. This empirical strategy differs substantially from the analysis of Sections 5.1 and 5.2. Nonetheless, the qualitative and quantitative results are remarkably similar.

A.1 An Extended Model

In the model of Section 3, the turnover probability $q$ was exogenously fixed. Obviously, in many cases, employees decide to voluntarily leave employers and, thus, turnover is endogenous. We now consider a simple extension of the model that delivers endogenous turnover. The main new mechanism is firm-specific human capital. This extension is also particularly useful because in the next section, A.2, we use a measure of industry-specific human capital provided by the Department of Labor as the main proxy for job turnover.\(^{20}\)

We assume that there is a continuum of jobs/industries and that jobs/industries differ in the importance of specific capital. In job $i$, the production function of a worker is

$$y_i = f(h, s_i),$$

where $s_i$ are skills partially specific to job $i$. More precisely, a worker moving to a different job can transfer only a fraction $(1 - \iota)$ of his skills $s_i$, so that a higher indexed job $i$ has more-specific skills. For simplicity, assume that the employee acquires the level of skills $s_i$ during the first period with

\(^{20}\)Related models in which specific capital and turnover rates are endogenously modeled can be found in Chang and Wang (1995, 1996). They focus on the role of asymmetric information where current employers are assumed to know more about workers’ productivity than potential employers.
the employer via a learning mechanism as in Jovanovic (1979), and that the level \( s_i \) is equal across all jobs.\(^{21}\)

To obtain endogenous turnover in the model, we assume that, in the second period, the worker can approach another firm at no cost. The new firm and the worker draw a match-specific productivity shock \( \epsilon \) from the distribution \( G(\epsilon) \). Production in the new firm \( y_2^n \) is equal to

\[
y_2^n = f(h_2, (1 - i)s_i) + \epsilon.
\]

In the new firm, the worker and the employer divide the surplus according to the Nash bargaining solution, so that, at the new firm, the worker gets a wage equal to:

\[
w_2^n(y_2^n, y_2^o) = (1 - \beta)w_2^o(y_2^o, y_2^n) + \beta y_2^n,
\]

(A1)

where \( w_2^o(y_2^o, y_2^n) \) and \( y_2^o = f(h_2, s_i) \) are the wage and the production in the old firm, respectively. Similarly, at the old firm, the worker gets a wage equal to:

\[
w_2^o(y_2^o, y_2^n) = (1 - \beta)w_2^n(y_2^n, y_2^o) + \beta y_2^o.
\]

(A2)

Solving the system of equations (A1) and (A2), we obtain

\[
w_2^n(y_2^n, y_2^o) = \frac{(1 - \beta)y_2^o + y_2^n}{2 - \beta} \quad \text{and} \quad w_2^o(y_2^o, y_2^n) = \frac{(1 - \beta)y_2^n + y_2^o}{2 - \beta}.
\]

The worker leaves his old firm whenever the new firm offers him a higher salary. Thus, the probability that the worker leaves the old firm in the second period is equal to:

\[
Pr(w_2^n(y_2^n, y_2^o) \geq w_2^o(y_2^o, y_2^n)) = Pr(y_2^n \geq y_2^o) = 1 - G(f(h_2, s_i) - f(h_2, (1 - i)s_i)),
\]

which is decreasing in \( i \). Thus, the introduction of firm-specific human capital makes turnover endogenous. We state this result as a modified version of Proposition 1:

**Proposition 3** Workers in jobs with more-specific skills (higher \( i \)) have a lower turnover rate and, thus, higher health investment.

**A.2 An Alternative Empirical Strategy**

In this section, we report results from an empirical strategy that differs from the strategy of Sections 5.1 and 5.2. More precisely, we construct a proxy of current (for employed individuals) and past (for retirees) job attachment at the three-digit industry level using data from the 1991 Dictionary of Occupational Titles (DOT). We then match this proxy to MEPS and HRS data.

We present the results of this alternative empirical strategy that more closely follow the results of Sections 5.1 and 5.2. In Fang and Gavazza (2007), we provide several additional tests, all consistent with the results reported here.

\(^{21}\)We can allow specific skills \( s_i \) to be endogenously accumulated at some cost. All results go through, at the cost of additional assumptions and calculations. Details are available from the authors upon request.
Dictionary of Occupational Titles (DOT) and Average Specific Vocational Preparation (ASVP). The Dictionary of Occupation Titles compiled by U.S. Department of Labor (1991) provides information about the training specificity required in various occupations. The variable, known as “Specific Vocational Preparation” (SVP) is defined as “the amount of time required to learn the techniques, acquire information, and develop the facility needed for average performance in a specific job-worker situation.” It is based on the nine numerical categories of vocational preparation, ranging from “Short demonstrations only” (category 1) to “Over 10 years” (category 9) [see U.S. Department of Labor (1991) for more details].

The Employee Retirement Income Security Act of 1974 (ERISA) requires employers to provide the same menu of health-care benefits to all workers in order for these benefits to be tax-exempt. Thus, firms presumably use the average job attachment of all workers in the firm as the relevant measure of job attachment when deciding what health benefits to offer to workers. Unfortunately, the data do not provide us with detailed information on the entire workforce of each individual firm. Thus, we focus on differences across industries in our analysis. More precisely, we follow the procedure described in Crocker and Moran (2003) and construct an Average Specific Vocational Preparation (ASVP) index for all the three-digit industry codes. Specifically, for each industry, we construct ASVP by taking the weighted average of the SVPs of workers’ occupations, where the weights are given by the five-percent sample from the 1990 Census.

The industries with the three lowest values of ASVP are “Services to dwellings” (industry code 722), “Services to private households” (industry code 761) and “Taxicab service” (industry code 402)—all industries in which intuition suggests that specific human capital is not important. Industries with the highest values of ASVP are “Legal services” (industry code 841), “Engineering, architectural, and surveying services” (industry code 882) and “Miscellaneous professional and related services” (industry code 892). Intuition suggests that a higher ASVP value would be associated with a higher importance of industry-specific human capital, and indeed Crocker and Moran show that a higher industry ASVP value is a strong predictor of longer job tenure at the firm level.

We match the constructed industry ASVP variable with individual-level data from MEPS (1998), HRS (2002) and BHPS (1997) in our analysis below. Table A1 shows the average ASVP for individuals’ current industry in MEPS (1998) and BHPS (1997) and individuals’ pre-retirement industry with longest job tenure in HRS (2002).

Empirical Specification. We match the ASVP index to the three-digit industry of each worker in the MEPS to investigate how current job attachment affects workers’ current medical expenditures. We further match each individual’s three-digit industry of the longest reported job in the HRS to investigate how past job attachment affects retirees’ medical expenditures. The basic
Table A1: Means and Standard Deviations of the ASVP Variable

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>5.19</td>
<td>0.83</td>
<td>5.23</td>
<td>0.82</td>
</tr>
</tbody>
</table>

The analysis is based on the following regression:

\[ y_i = \alpha \text{ASVP}_i + \beta \text{X}_i + \epsilon_i, \tag{A3} \]

where \( y_i \) is one of the several outcomes considered for individual \( i \), such as total health expenditures, doctor visits, health status, etc.; \( \text{X}_i \) is a large set of control variables including individual \( i \)'s age (also squared and cubed), education, gender, etc. The coefficient of \( \text{ASVP}_i \), \( \alpha \), measures the average effect of our proxy for job attachment on the outcome \( y_i \), after controlling for a large number of factors included in the vector \( \text{X}_i \).

### A.2.1 Medical Expenditures of Workers

Table A2 presents the results of several regressions that investigate employees’ medical expenditures (Panel A) and doctor visits (Panel B). Columns (1) and (2) report the coefficients of Tobit regressions in which the dependent variables are the level and the log of an individual’s (annual) total medical expenditures, respectively. We employ Tobit regressions since the dependent variables are right-censored at zero expenditures. Columns (3) and (4) report the coefficients of negative binomial regressions in which the dependent variable is the number of office-based visits and the number of physician visits, respectively.\textsuperscript{22}

Columns (1) and (2) show that individuals working in high-ASVP industries—i.e., industries with low turnover rates—have higher medical expenditures. The marginal effect calculated from the Tobit regression coefficient on ASVP reported in Column 1 implies that a unit increase in ASVP increases annual medical expenditures by around $113 dollars, or about six percent of the average medical expenditure, a rather large effect. The coefficient of ASVP reported in Column 2 is much bigger: It implies that a unit increase in ASVP increases annual medical expenditures by about 15 percent.\textsuperscript{23}

\textsuperscript{22}The number of office visits is the sum of visits to physicians and nonphysicians. MEPS classifies the following categories as nonphysicians: chiropractors, midwives, nurses and nurse practitioners, optometrists, podiatrists, physician’s assistants, physical therapists, occupational therapists, psychologists, social workers, technicians, and receptionists/clerks/secretaries.

\textsuperscript{23}The difference between the two coefficients suggests that individuals in low-turnover industries have higher average medical expenditures and a lower variance of medical expenditures. This is an implication of Jensen’s inequality due to the log transformation of the dependent variable. See, also, Santos Silva and Tenreyro (2006).
Table A2: Relationship Between Industry ASVP and Total Medical Expenditures (Panel A) and Doctor Visits (Panel B)

<table>
<thead>
<tr>
<th>Variables</th>
<th>Panel A: Total Medical Expenditures</th>
<th>Panel B: Doctor Visits</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Level (1)</td>
<td>Log (2)</td>
</tr>
<tr>
<td>ASVP</td>
<td>199.8***</td>
<td>0.22***</td>
</tr>
<tr>
<td></td>
<td>(92.2)</td>
<td>(0.06)</td>
</tr>
<tr>
<td>Age</td>
<td>-135.5*</td>
<td>-0.24***</td>
</tr>
<tr>
<td></td>
<td>(72.6)</td>
<td>(0.05)</td>
</tr>
<tr>
<td>Age Squared</td>
<td>3.8**</td>
<td>0.006***</td>
</tr>
<tr>
<td></td>
<td>(1.7)</td>
<td>(0.001)</td>
</tr>
<tr>
<td>Education</td>
<td>74.5***</td>
<td>0.13***</td>
</tr>
<tr>
<td></td>
<td>(27.9)</td>
<td>(0.02)</td>
</tr>
<tr>
<td>Male</td>
<td>-1257.7***</td>
<td>-1.63***</td>
</tr>
<tr>
<td></td>
<td>148.4)</td>
<td>(0.11)</td>
</tr>
<tr>
<td>Income/1000</td>
<td>-38.3*</td>
<td>-0.04***</td>
</tr>
<tr>
<td></td>
<td>(21.4)</td>
<td>(0.01)</td>
</tr>
<tr>
<td>Family Size</td>
<td>-239.4***</td>
<td>-0.19***</td>
</tr>
<tr>
<td></td>
<td>(36.9)</td>
<td>(0.02)</td>
</tr>
<tr>
<td>Union</td>
<td>512.0***</td>
<td>0.43***</td>
</tr>
<tr>
<td></td>
<td>(178.6)</td>
<td>(0.11)</td>
</tr>
</tbody>
</table>

Notes: (1). Panel A reports the coefficient estimates from Tobit regressions with the total medical expenditure (level) or Log (1+Total Medical Expenditures) as the dependent variable. Panel B reports the coefficient estimates from negative binomial regressions where the dependent variables are “Number of Office Based Visits” and “Number of Visits to Physicians,” respectively (2). All regressions include a constant and additional controls for Firm Size, Race, Census Region and MSA dummies, as well as Age Cubed. Their coefficient estimates are not shown. (3). Robust standard errors clustered at the industry level are in parenthesis; (4). *, **, *** denote significance at ten, five and one percent, respectively.
Columns (3) and (4) show that workers in high-ASVP industries visit medical providers more frequently. The magnitudes of the coefficients imply that a unit increase in ASVP is associated with an increase in the annual number of medical providers’ visits and physician visits of about five percent, a rather large effect.

Overall, the results of Table A2 are consistent with Proposition 1 (and its extension, Proposition 3) of our model.

A.2.2 Medical Expenditures of Retirees

We now, using HRS data, investigate how past job attachment affects retirees’ medical expenditures and health status. More precisely, the HRS reports individuals’ longest job, along with its three-digit industry code. Thus, we match our ASVP proxy to the industry in which the individual had his longest job.

Column (1) in Table A3 presents the results of a Tobit regression that investigates how past ASVP affects retirees’ current medical expenditure. Column (2) presents the results of an ordered Probit regression that investigates how past ASVP affects retirees’ current health status. The dependent variable is a categorical indicator of self-reported health status, with 1 indicating “Excellent,” 2 “Very Good,” 3 “Good,” 4 “Fair,” and 5 “Poor.”

Column (1) shows that medical expenditures are lower for individuals who worked in high-ASVP industries prior to retirement. The marginal effect from the estimated Tobit coefficients on ASVP shows that a one-unit increase in pre-retirement industry ASVP is associated with a decrease of $1,037 in annual medical expenditures, a substantial effect. Column (2) of Table A3 shows that the coefficient of ASVP is negative and statistically significant, indicating that individuals who worked in high ASVP industries prior to retirement have better self-reported health in retirement. This is particularly interesting since Column (1) shows that these same individuals have lower medical expenditures. Overall, these findings are consistent with Proposition 2 of our model.

A.2.3 Assessing the Magnitude of Life-Cycle Inefficiency

We now combine the estimates of the previous two sections to try to calculate the magnitude of the externality implied by our new set of regressions, in a parallel way to our calculations of Section 5.3.

Thus, suppose that both individuals A and B work for 45 years and then retire for 15 years before dying. Individual A works in an industry in which ASVP is low (i.e., mobility is high), while individual B works in an industry in which ASVP is high (i.e., mobility is low). More precisely, individual A’s ASVP is one unit lower than individual B’s. The coefficient of ASVP in the regressions of Table A2 on MEPS data implies that individual A’s expenditures are lower than B’s by $113 per year. Instead, the coefficient of ASVP in the regressions of Table A3 on HRS data
Table A3: Relationship Between Retirees’ Total Medical Expenditures and Perceived Health Status and the ASVP of their Pre-retirement Industries

<table>
<thead>
<tr>
<th>Variables</th>
<th>Total Medical Expenditures</th>
<th>Perceived Health Status</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
</tr>
<tr>
<td>ASVP</td>
<td>-1,012.5* (537.3)</td>
<td>-0.045** (0.020)</td>
</tr>
<tr>
<td>Age</td>
<td>19,889 (23,447)</td>
<td>-0.046 (0.515)</td>
</tr>
<tr>
<td>Age Squared</td>
<td>-268.3 (307.4)</td>
<td>-0.001 (0.007)</td>
</tr>
<tr>
<td>Age Cubed</td>
<td>1.22 (1.34)</td>
<td>0.000 (0.000)</td>
</tr>
<tr>
<td>Education</td>
<td>132.1 (125.9)</td>
<td>-0.069*** (0.006)</td>
</tr>
<tr>
<td>Male</td>
<td>1766.0*** (730.6)</td>
<td>0.072*** (0.029)</td>
</tr>
<tr>
<td>Assets/10,000</td>
<td>-56.2*** (21.7)</td>
<td>-0.013*** (0.004)</td>
</tr>
<tr>
<td>Family Size</td>
<td>-100.3 (350.9)</td>
<td>0.034** (0.014)</td>
</tr>
<tr>
<td>Married</td>
<td>-839.9 (791.3)</td>
<td>-0.178*** (0.037)</td>
</tr>
<tr>
<td>No. of Obs.</td>
<td>5,583</td>
<td>6,730</td>
</tr>
</tbody>
</table>

Notes: (1). Column 1 reports the estimates from a Tobit regression with the total medical expenditures as the dependent variable. Column 2 reports the estimates from an ordered Probit regression with “Perceived health status 1: Excellent; 2: Very Good; 3: Good; 4: Fair; 5 Poor” as the dependent variable; (2). Both regressions include Race, Census Region and MSA dummies and their coefficient estimates are not shown; (3). Robust standard errors clustered at the industry level are in parenthesis; and (4). *, **, *** denote significance at ten, five and one percent, respectively.

implies that individual A has higher medical expenditures than individual B by $1,037 per year.

Thus, if individuals A and B work for 45 years and then retire for 15 years, non-discounting their expenditures, we have that individual A has approximately $5,000 less in medical expenditures per year than individual B when working, but approximately $15,000 more in medical expenditures when retired. This calculation suggests that one additional dollar of medical expenditures during the working years may lead to about three dollars of savings during retirement. Again, we wish to emphasize that this calculation is very rough, as we already noted in Section 5.3. Nonetheless, it describes in a very simple way the externality we have in mind and its magnitude in the data. Moreover, it is quite interesting to note that the magnitude of the externality is quantitively close to the one that we calculated in Section 5.3, obtained from a very different empirical methodology.24

24 The levels of the expenditures do not correspond to the levels reported in Section 5.3 because one unit of ASVP does not translate into one standard deviation of the log of job tenure.
A.2.4 Falsification Test: U.K. Workers

We now present the results of a falsification test that uses data from the U.K. BHPS, similar to the test we performed in Section 5.4. Specifically, we use the 1997 wave of British Household Panel Survey (BHPS) to investigate the relationship between ASVP and doctor visits for U.K. workers.

Column (1) in Table A4 reports the results from a negative binomial regression in which the dependent variable is the number of annual doctor visits. The estimated coefficient is almost zero, and is not statistically significant. This shows that our proxy for job attachment ASVP does not significantly affect the frequency of doctor visits, in sharp contrast to the evidence for the U.S. reported in Panel B of Table A2. Column (2) further reports the results of an ordered Probit regression that investigates the relationship between ASVP and a five-categorical indicator of self-reported health. Column (2) shows that the coefficient of ASVP does not statistically differ from zero. This provides additional evidence in favor of the mechanism identified by our model.
Table A4: Falsification Test: Relationship Between Industry ASVP and Doctor Visits and Perceived Health Status for U.K. Workers.

<table>
<thead>
<tr>
<th>Variables</th>
<th>Doctor Visits (1)</th>
<th>Perceived Health Status (2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ASVP</td>
<td>-0.007</td>
<td>-0.035</td>
</tr>
<tr>
<td>Age</td>
<td>-0.031***</td>
<td>-0.006</td>
</tr>
<tr>
<td>Age²</td>
<td>0.0004***</td>
<td>0.000</td>
</tr>
<tr>
<td>Education</td>
<td>-0.001</td>
<td>-0.030**</td>
</tr>
<tr>
<td>Male</td>
<td>-0.502***</td>
<td>-0.129***</td>
</tr>
<tr>
<td>Income/10,000</td>
<td>-0.026***</td>
<td>-0.043**</td>
</tr>
<tr>
<td>Family Size</td>
<td>-0.006</td>
<td>0.013</td>
</tr>
<tr>
<td>Union</td>
<td>0.110**</td>
<td>0.052</td>
</tr>
</tbody>
</table>

Notes: (1). Column 1 reports the coefficient estimates from a negative binomial regression where the dependent variable is the “Number of Annual Doctor Visits”; Column 2 reports the coefficient estimates from an ordered Probit regression with “Perceived health status 1: Excellent; 2: Very Good; 3: Good; 4: Fair; 5 Poor” as the dependent variable; (2). Robust standard errors clustered at the industry level are in parenthesis; and (3). *, **, *** denote significance at ten, five and one percent, respectively.
References


