Analyzing the Effects of Insuring Health Risks:*

On the Tradeoff between Short Run Insurance Benefits vs. Long Run Incentive Costs

Harold L. Cole
University of Pennsylvania and NBER

Soojin Kim
University of Pennsylvania

Dirk Krueger
University of Pennsylvania, CEPR and NBER

October 16, 2012

Abstract

This paper constructs a dynamic model of health insurance to evaluate the short- and long run effects of policies that prevent firms from conditioning wages on health conditions of their workers, and that prevent health insurance companies from charging individuals with adverse health conditions higher insurance premia. Our study is motivated by recent US legislation that has tightened regulations on wage discrimination against workers with poorer health status (Americans with Disability Act of 2009, ADA, and ADA Amendments Act of 2008, ADAAA) and that will prohibit health insurance companies from charging different premiums for workers of different health status starting in 2014 (Patient Protection and Affordable Care Act, PPACA). In the model, a trade-off arises between the static gains from better insurance against poor health induced by these policies and their adverse dynamic incentive effects on household efforts to lead a healthy life. Using household panel data from the PSID we estimate and calibrate the model and then use it to evaluate the static and dynamic consequences of no-wage discrimination and no-prior conditions laws for the evolution of the cross-sectional health and consumption distribution of a cohort of households as well, as ex-ante lifetime utility of a typical member of this cohort. In our quantitative analysis we find that although a combination of both policies is effective in providing full consumption insurance period by period, it lowers expected lifetime utility since it induces a more rapid deterioration of the cohort health distribution over time. Interestingly, whereas introducing each law in isolation has limited adverse dynamic incentive effects and improves welfare relative to the competitive equilibrium, a combination of both laws severely undermines the incentives to lead healthier lives. The resulting negative effects on health outcomes in society more than offset the static gains from better consumption insurance so that expected discounted lifetime utility declines as a result of introducing both policy measures in conjunction.

JEL Codes: E61, H31, I18

Keywords: Health, Insurance, Incentive

1 Introduction

Here we study the impact of social insurance policies aimed at reducing households' exposure to health related risk during their working life. These risk come through higher medical costs, higher medical premiums and lower earnings. Historically there have been major insurance efforts aimed at the elderly through Medicaid and Social Security, and the poor through Medicare and income support programs like Welfare and Food Stamps. Recently the extent of these programs and the scope of the different groups they impact on has been greatly expanded. On the insurance front, HIPPA in 1966 and the Patient Protection and Affordable Care Act in 2010 sought to increase access to health care and to prevent health insurance being differentially

*We thank Hanming Fang, Iourii Manovskii and participants of the Penn Macro Lunch and Annual Meeting for Society for Economic Dynamics 2012 for many helpful comments. Krueger gratefully acknowledges financial support from the NSF under grant SES-0820494.

priced based upon pre-existing conditions. On the income front, the 1990 Americans with Disabilities Act and its Amendment in 2009 sought to restrict the ability of employers to employ and compensate workers based upon health related reasons.

In order to analyze the impact of these policies we construct a dynamic model of health insurance with heterogeneous households. As in Grossman (1972), health for these households is a state variable. A household's health state helps to determine both their productivity at work and the likelihood that they will be subject to adverse health shocks. Our model features the two-way interaction between health and income that has emphasized in the literature. Our model of health shocks includes temporary health shocks that impact on productivity and can be offset by medical expenditures (as in Dey and Flinn 2005), and catastrophic health shocks which require health expenditure to avoid death. Our model also features persistent health shocks through the stochastic evolution of the individual's health status. This evolution is affected by the household's efforts to maintain their health which results in a moral hazard problem as health related insurance reduces households' incentives to maintain their health. We explicitly model the choice of medical expenditure and thereby endogenously determined the health insurance policy and how it responds both to the household's state in terms of health status, age and education.

The focus of our analysis is how the distributions of health status, earnings and health insurance costs will evolved under different policy choices and the impact of these choices on welfare. We consider several different policy regimes. The first a complete insurance benchmark in which the social planner can dictate both the full health insurance contract, the effort made to maintain health and the extent of redistributive transfers to ensure against all health related shocks. The second is pure competition in which workers enter into one-period employment and insurance contracts. Competition leads to these contracts partially insuring the worker against within period temporary health shocks, but not against his initial health status and the transition of this status. The second is a version of the no-prior conditions restriction on health insurance in which health insurance companies compete to offer one-period health insurance contracts in which they cannot differentially charge based upon the worker's health status. The third is a version of the no-discrimination restrictions on employment in which we assume that firms cannot differentially hire or pay workers based upon their health status. In the fourth version we consider the impact of both the no-prior conditions and the no-discrimination restrictions.

We consider both the static and the dynamic impact of these policies. One of the key aspects of the dynamic analysis is the impact these policies have on individuals' incentives to maintain their health and the feedback this creates between the health distribution of the population and the costs of health insurance and productivity of the workforce.

We evaluate the quantitative impacts of the different policies. To do so, we estimate and calibrate the model using PSID to match key aggregate statistics on wage and medical expenditure. Our results show that a combination of wage non-discrimination law and no prior conditions law provides full insurance against health risks and restores first-best allocation in the short run, but leads to a severe deterioration of health in the long run.

1.1 Institutional Background

The U.S. has a long history of policy initiatives in relation to health risk. Implicitly Welfare programs, which in date back to the 1930s and were greatly expanded by the Great Society in the 1960s, insure workers against a variety of shocks including implicitly health related shocks as they affect earnings. Since 1965 Medicare has sought to provide health insurance to the elderly and the disabled. Medicaid has sought to provide health insurance to the poor since the 1990s. The last two decades legislation in the U.S. was passed that limits the ability of employers to condition wages on the health conditions of employees, and to discriminate against applicants with prior health conditions when filling vacant positions.

1.1.1 Wage Based Discrimination

In 1990 Congress enacted the Americans with Disabilities Act (ADA) to ensure that the disabled have equal access to employment opportunities.[1] At this point a disability was interpreted as an impairment that

[1] The ADA sets the federal minimum standard of protection. States may have a more stringent level.

prevents or severely restricts an individual from doing activities that are of central importance to one's daily life. In 2009 the ADA Amendments Act (ADAAA) went into effect. This act rejected the strict interpretation of the ADA, broadening the notion of a disability. This included prohibiting the consideration of measures that reduce or mitigate the impact of a disability in determining whether someone is disabled. It also allowed people who are discriminated against on the basis of a perceived disability to pursue a claim on the basis of the ADA regardless of whether the perceived disability limits or is perceived to limit a major life activity. The ADAAA excludes from the definition of a disability those temporary or minor impairments.[2] Under the ADAAA people can be disabled even if their disability is episodic or in remission. For example people whose cancer is remission or whose diabetes is controlled by medication, or whose seizures are prevented by medication, or who can function at a high level with learning disabilities are all disabled under the act.

Before the ADA job seekers could be asked about their medical conditions and were often required to submit to a medical exam. The act prohibited certain inquiries and conducting a medical exam before making an employment offer. However, the job could still be conditioned upon successful complete of a medical exam.[3]

The ADA permits an employer to establish job-related qualifications on the basis of business necessity. However, business necessity is limited to essential functions of the job. So impairments that would only occasionally interfere with the employee's ability to perform tasks cannot be included on this list.[4] A job function is essential if the job exists to perform that function or if the limited number of employees available means that the task must be performed by this employee. Furthermore, a core requirement of the ADA is the obligation of the employer to make a reasonable accommodation to qualified disabled people.[5]

1.1.2 Insurance Cost and Exclusion Discrimination

In 1996, Congress passed the Health Insurance Portability and Accountability Act (HIPAA) which placed limits on the extent to which insurance companies could exclude people or deny coverage based upon pre-existing conditions. While insurance companies were allowed exclusions periods for coverage of pre-existing conditions, these exclusion periods were reduced by the extent of prior insurance. In particular, if an individual had at least a full year of prior health insurance and they enrolled in a new plan with a break of less than 63 days, they could not be denied coverage. However, insures were still allowed to charge higher premiums based upon initial conditions, limit coverage and set lifetime limits on benefits.[6] There is evidence that many patients with pre-existing conditions ended up either being denied coverage,[7] or having their access to benefits limited.[8]

The Patient Protection and Affordable Care Act of 2010 further extended protection against pre-existing conditions. Beginning in 2010 children below the age of 19 could not be excluded from their parents' health or denied treatment for pre-existing conditions. Beginning in 2014 this restriction will apply to adults as well. Moreover, insurance companies will no longer be able to use health status to determine eligibility, benefits or premiums. In addition, insurers will no long be able to limit lifetime or annual benefits or take away coverage because of an application mistake.[9]

---

[2] Under the ADAAA major life activities now include: caring for oneself, performing manual tasks, seeing, hearing, eating, sleeping, walking, standing , lifting, bending, speaking, breathing, learning, reading, concentrating, thinking, communicating, working, as well as major bodily functions.

[3] For example, the Equal Employment Opportunity Commision (EEOC) has ruled that an employee may be asked "how many days were you absent from work?", but not how many days were you sick?".

[4] For example, an employer cannot require a driver's license for a clerking job because it would occasionally be useful to have that employee run errands. Also qualification cannot be such that a reasonable accommodation would allow the employee to perform the task.

[5] These accomodations include: a) making existing facilities accessible and usable b) job restructuring c) part-time or modified work schedules d) reassigning a disable employee to a vacant position e) acquiring or modifying equipment or devices f) providing qualified readers or interpreters.

[6] See http://www.healthcare.gov/center/reports/preexisting.html

[7] See Kass et al. (2007).

[8] See Sommers (2006).

[9] See again http://www.healthcare.gov/center/reports/preexisting.html

1.1.3 Summary

It is our interpretation of these legislative changes that, relative to 20 years ago, it is much more difficult now for employers to condition wages on the health status of their (potential) employees and preferentially hire workers with better health. In addition, current and pending legislation will make it increasingly difficult to condition the acceptance into, and insurance premia of health insurance plans on prior health conditions.

The purpose of the remainder of this paper is to analyze the aggregate and distributional consequences of these two legislative innovations in the short and in the long run, with specific focus on their interactions.

1.2 Related Literature

Our paper incorporates health as a productive factor, and studies the effect of labor and health insurance market policies on its evolution. We allow for a two-way interaction between health shocks and earnings through worker productivity. We model medical expenditures which mitigate the impact of these health shocks. There has been a number of studies that empirically estimate the effect of health on wages. These papers (see the summary in Currie and Madrian, 1999) generally find that poor health decreases wages, both directly and indirectly through decrease in hours worked. The effect of a health shock on wages ranges from 1% to as high as 15%. Many studies consistently find that the effects on hours worked is greater than that on wages. Specifically relevant for us is Cawley (2004).

Similarly to what we do for working age individuals, Pijoan-Mas and Rios-Rull (2012), use HRS data on self-report health status to estimate a health transition function from age 50 onwards. They find that there is an important dependence in this transition function on socioeconomic status (education most importantly), and that this dependence is quantitatively crucial for explaining longevity differentials by socioeconomic group. As we do Prados (2012) models the interaction between health and earnings over the life cycle, and stresses the finding that health shocks amplify earnings inequality.[10]

A relatively small literature examines the incentive linkages between health insurance and health status. Bhattacharya et al. (2009) use evidence from a Rand health insurance experiment, which featured randomized assignment to health insurance contracts, to show that access to health insurance leads to increases in body mass and obesity. They argue that this comes from the fact that insurance, especially through its pooling effect, insulates people from the impact of their excess weight on their medical expenditures costs. Consistent with this, they find the impact of being health- insured is larger for public insurance programs than in private ones where the health insurance premium is more likely to reflect the individuals' body mass.

There are a variety of papers which have examined the macroeconomic and distributional implications of health and health insurance. These include Grossman (1972), Ehrlich and Becker (1972), Ehrlich and Chuma (1990), French and Jones (2004), Hall and Jones (2007), Ales, Hosseini and Jones (2012), Kopecky and Koreshkova (2012) and Laun (2012). Brügemann and Manovskii (2010) study the macroeconomic effects of the employer-sponsored health insurance system that is unique to the US labor market, endogenizing health. Concretely, they determine the effect of PPACA on health insurance coverage, but do not study the incentive effects of the regulation that we formalize in our model.

Several papers investigate the impact of regulation designed to limit the direct effect of health on both health insurance costs and on wages. Short and Lair (1994) consider examines how health status interacts with insurance choices. Madrian (1994) examines the lock-in effect of employer provided health care. Dey and Flinn (2005) estimate a model of health insurance with search, matching and bargaining and argue that employer provided health care insurance leads to reasonably efficient outcomes.

Related to our study of wage non-discrimination laws is the literature that studies the effect of the ADA legislation of 1990 on employment, wages and labor hours of the disabled (see DeLeire (2001) and Acemoglu and Angrist (2001), for example). Most find that it has decreased the employment of the disabled. DeLeire (2001) quantifies the effect of ADA on wage of disabled workers and reports that the negative effect of poor health on the earnings of the disabled fell by 11.3 % due to ADA.

Finally, a recent literature examines the impact of health on savings and portfolio choice in life cycle models that share elements with our framework. These include Yogo (2009), Edwards (2008) and Hugonnier et al. (2012). The latter study jointly portfolio of health and other asset choices. In their model health

[10]She also studies the impact of compulsary health insurance.

increases productivity (labor income) and decreases occurrence of morbidity and mortality shock arrival rates (as they do in our model). The paper argues that in order to explain correlations in financial and health statuses, these should be modelled jointly.

2 The Model

Time $t = 0, 1, 2, \ldots T$ is discrete and finite and the economy is populated by a cohort of a continuum of individuals of mass 1. Since we are modeling a given cohort of individuals we will use time and the age of households interchangeably. We think of T as the end of working life of the age cohort under study.

2.1 Endowments and Preferences

Households are endowed with one unit of time which they supply inelastically to the market. They are also endowed with an initial level of health h and we denote by $H = \{h_1, \ldots, h_N\}$ the finite set of possible health levels. Households value current consumption c and dislike the effort e to live a healthy life. We will assume that their preferences are additively separable over time, and that they discount the future at time discount factor β. We will also assume that preferences are separable between consumption and effort, and that households value consumption according to the common period utility function $u(c)$ and value effort according to the period disutility function $q(.)$.

We will denote the probability distribution over the health status h at the beginning of period t by $\Phi_t(h)$, and denote by $\Phi_0(h)$ the initial distribution over this characteristic.

Assumption 1 *The utility function u is twice differentiable, strictly increasing and strictly concave. q is twice differentiable, strictly increasing, strictly convex, with $q(0) = q'(0) = 0$ and $\lim_{e\to\infty} q'(e) = \infty$.*

2.2 Technology

2.2.1 Health Technology

Let ε denote the current health shock.[11] In every period households with current health h remain healthy (that is, $\varepsilon = 0$) with probability $g(h)$. With probability $1 - g(h)$ the household draws a health shock $\varepsilon \in (0, \bar{\varepsilon}]$ which is distributed according to the probability density function $f(\varepsilon)$.

Assumption 2 *f is continuous and g is twice differentiable with $g(h) \in [0, 1]$, and $g'(h) > 0, g''(h) < 0$ for all $h \in H$.*

An individual's health status evolves stochastically over time, according to the Markov transition function $Q(h', h; e)$, where $e \geq 0$ is the level of exercise by the individual and x is the amount of resources spent on health goods. We impose the following assumption on the Markov transition function Q

Assumption 3 *If $e' > e$ then $Q(h', h; e)$ first order stochastically dominates $Q(h', h; e')$.*

2.2.2 Production Technology

A individual with health status h and current health shock ε that consumes health expenditures x produces $F(h, \varepsilon - x)$ units of output.

Assumption 4 *F is continuously differentiable in both arguments, increasing in h, and satisfies $F(h, y) = F(h, 0)$ for all $y \leq 0$, and $F_2(h, y) < 0$ as well as $F_2(h, \bar{\varepsilon}) < -1$. Finally $F_{22}(h, y) < 0$ for all $y > 0$ and $F_{12}(h, y) \geq 0$.*

[11] In the quantitative analysis we will introduce a second, fully insured (by assumption) health shock to provide a more accurate map between our model and the health expenditure data.

The left panel of figure 1 displays the production function $F(h,.)$, for two different levels of the current health shock. Holding health status h constant, output is decreasing in the uncured portion of the health shock $\varepsilon - x$, and the decline is more rapid for lower levels of health ($h^* < h$). The right panel of figure 1 displays the production function as function of health expenditures x for fixed shock ε, and shows that expenditures x exceeding the shock ε leave output $F(h, \varepsilon - x)$ unaffected. A reduction in ε to ε^* thus shifts the point at which health expenditures x become ineffective to the left.

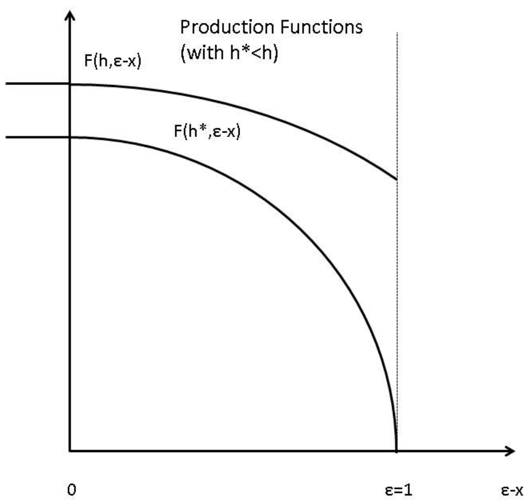

Figure 1: Production Function $F(h, \varepsilon - x)$

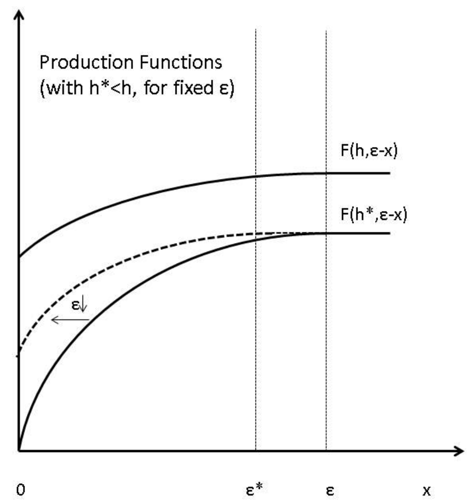

Figure 2: Production Function $F(h, \varepsilon - x)$ for fixed ε

The assumptions on the production function F imply that health expenditures can offset the impact of a health shock on productivity, but not raise an individual's productivity above what it would be if there had been no shock. In addition, the last assumption implies that the negative impact of a given net health shock y is lower the healthier a person is.[12] The assumption $F_2(h, \bar{\varepsilon}) < -1$ insures that, if hit by the worst health shock the cost of treating this health shock, at the margin, is smaller than the positive impact on productivity (output) this treatment has.

2.3 Time Line of Events

In the current period the timing of events is as follows

1. Households enter the period with current health status h.
2. Households choose e
3. Firms offer a wage $w(h)$ and health insurance contract $\{x(\varepsilon, h), P(h)\}$[13]
4. The health shock ε is drawn according to g, f
5. Resources on health $x = x(\varepsilon, h)$ are spent
6. Production and consumption takes place
7. New health status h' is drawn according to Q

[12] This is also the approach taken by Hugonnier et al. (2012) and Ehrlich and Chuma (1990).

[13] As we restrict attention to static contracts, whether firm offers contracts before or after the effort is undertaken does not matter.

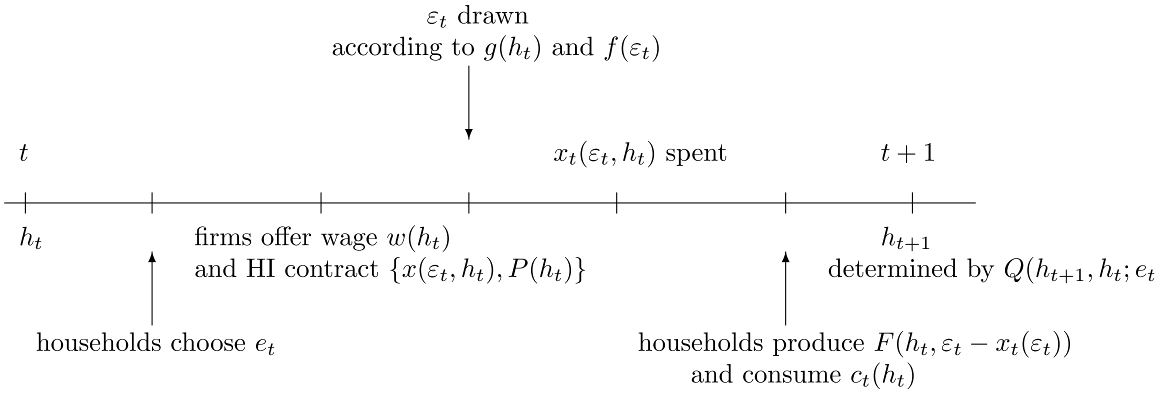

Figure 3: Timing of the Model

2.4 Market Structure without Government

There are a large number of production firms that in each period compete for workers. Firms observe the health status of a worker h and then, prior to the realization of the health shocks, compete for workers of type h by offering a wage $w^{CE}(h)$ that pools the risk of the health shocks and bundle the wage with an associated health insurance contract (specifying health expenditures $x(\varepsilon, h)$ and an insurance premium $P(h)$) that breaks even. Perfect competition for workers of type h requires that the combined wage and health insurance contract maximizes period utility of the household, subject to the firm breaking even.[14]

In the absence of government intervention a firm specializing on workers of health type h therefore offers a wage $w^{CE}(h)$ and health insurance contract $(x^{CE}(\varepsilon, h), P^{CE}(h))$ that solves

$$U^{CE}(h) = \max_{w(h), x(\varepsilon,h), P(h)} u\left(w(h) - P(h)\right) \tag{1}$$

s.t.

$$P(h) = g(h)x(0,h) + (1-g(h)) \int_0^{\bar{\varepsilon}} f(\varepsilon)x(\varepsilon,h)d\varepsilon \tag{2}$$

$$w(h) = g(h)F(h, -x(0,h)) + (1-g(h)) \int_0^{\bar{\varepsilon}} f(\varepsilon)F(h, \varepsilon - x(\varepsilon,h))d\varepsilon \tag{3}$$

Note that by bundling wages and health insurance the firm provides efficient insurance against health shocks ε, and the only source of risk remaining in the competitive equilibrium is health status risk associated with h. This risk stems both from the dependence of wages $w(h)$ as well as health insurance premia $P(h)$ on h in the competitive equilibrium, and these are exactly the sources of consumption risk that government policies preventing wage discrimination and prohibiting prior health conditions to affect health insurance premia are designed to tackle.

2.5 Government Policies

We now describe in turn how we operationalize, within the context of our model, a policy that prevents health insurance contracts to condition on prior health conditions h, and a policy that limits the extent to which firms can pay workers that differ in their health status different wages.

[14] Note that instead of assuming that firms completely specialize by hiring only a specific health type of workers h we could alternatively consider a market structure in which all firms are representative in terms of hiring workers of health types according to the population distribution and pay according to the wage schedule $w^{CE}(h)$. In other words variation of wages and variation in health types h are perfect substitutes at the level of the individual firm in terms of supporting the competitive equilibrium allocation.

2.5.1 No Prior Conditions Law

Under this law health insurance companies are assumed to be constrained in terms of the pricing, their insurance schedule and their applicant acceptance criteria. The purpose of these constraints is to prevent the companies from differentially pricing insurance based upon health status.[15] To be completely successful, these constraints must lead to a pooling equilibrium in which all individuals are insured at the same price. The best such regulation in addition assures that the equilibrium health insurance schedule $x(\varepsilon, h)$, given the constraints, is efficient.

The first constraint is that a company must specify the total number of contracts that it wishes to issue, it must charge a fixed price independent of health status, and accept applications in their order of application up to the sales limit of the company. In this way, the insurance company cannot examine applications first and then decide whether or not to offer the applicant a health insurance contract.

The second constraint regulates the health expenditure schedule. If the no-prior conditions law is to have any bite the government needs to prevent the emergence of a separating equilibrium in which the health insurance companies (or the production firms in case they offer health insurance contracts) use the health expenditure schedule $x(\varepsilon, h)$ to effectively select the desired health types, given that they are barred from conditioning the health insurance premium P on h directly. Therefore, in order to achieve any sort of pooling in the health insurance market requires government regulation of the health expenditure schedule $x(\varepsilon, h)$. To give the legislation the best chance of being successful we will assume that the government *regulates the health expenditure schedule $x(\varepsilon, h)$ efficiently.* For the same reason, since risk pooling is limited if some household types h choose not to buy insurance, we assume that all individuals are *forced* to buy insurance.

Given this structure of regulation and a cross-sectional distribution of workers by health type and preference parameters, Φ, the health insurance premium P charged by competitive firms (or competitive insurance companies, who offer health insurance in the model), given the set of regulations spelled out above, is determined by

$$P = \sum_h \left[g(h)x(0,h) + (1-g(h)) \int f(\varepsilon)x(\varepsilon,h)d\varepsilon \right] \Phi(h) \tag{4}$$

where $x(\varepsilon, h)$ is the expenditure schedule regulated by the government. This schedule is chosen to maximize

$$\sum_h u(w(h) - P)\Phi(h)$$

with wages $w(h)$ determined by (3).

2.5.2 No Wage Discrimination Law

The objective of the government is to prevent workers with a lower health status h, and hence lower productivity, being paid less. As with the no prior conditions law, the purpose of this legislation is to help insure workers against their health status risk. However, if a production firm is penalized for paying workers with low health status h low wages, but not for preferentially hiring workers with a favorable health status (high h), then a firm can effectively circumvent the wage nondiscrimination law. Therefore, to be effective such a law must penalize *both* wage discrimination and hiring discrimination by health status.

Limiting wage dispersion with respect to gross wages $w(h)$ via legislation necessitates regulation of the health insurance market as well, in order to prevent the insurance gains from decreasing wage dispersion being undone through the adjustment of employer-provided health insurance. For example, the firm could also offer health insurance and overcharge low productivity workers and undercharge high productivity workers for this insurance, effectively undoing the illegal wage discrimination. This suggests that the government will need to limit the extent to which the cost of a worker's health insurance contracts deviates from its actuarially fair value. However, this will not be sufficient to make this policy effective.

Since the productivity of a worker depends upon the extent of his health insurance, workers whose expected productivity is below their wage will face pressure to increase their productivity through increased spending on health (and hence better health insurance coverage) while those whose productivity is above

[15] Consistent with this purpose, we will assume that the government cannot use health insurance to offset underlying differences in productivity coming from, say, education. This will prove important in the quantitative section.

their wage will have an incentive to lower their health insurance purchases. To prevent these distortions in the health insurance market and thereby achieve better consumption insurance across h types, policy makers will need to regulate the health insurance directly as well. The moderate version of health insurance regulation would be to ensure that each policy was individually optimal and actuarially fair. The most extreme version of regulation would be to combine no-wage discrimination legislation with no-prior conditions legislation and thereby achieve the static first-best, full insurance outcome. In this case health insurance would be socially efficient and actuarially fair on average.

We will analyze both cases. It will turn out that limiting wage dispersion with respect to net wages, $w(h) - P(h)$, avoids the negative incentive effects on the health insurance market. The policy of combining both no-wage discrimination and no-prior conditions can therefore be implemented through a policy of limiting net wage dispersion. The impact of the nondiscrimination law will, unfortunately, be sensitive to the way in which the law is implemented, and in particular, to the form of punishment used. If the limitation in wage variation is achieved through a policy that penalizes the firms for discriminating, then these costs are realized in equilibrium, reducing overall efficiency in the economy. If, however, the limitation on wage variation is achieved either through the threat of punishment (e.g. through grim trigger strategies in repeated interactions between firms and the government) or through the delegation of hiring in a union hiring hall type arrangement, then costs from the wage nondiscrimination law will not be realized in equilibrium.[16]

Since we wish to give the no wage discrimination law the best shot of being successful, in the main text we focus on the version of the policy in which no costs from the policy are realized in equilibrium, leaving the analysis of the alternative case to appendix B.2 and B.3. In either case we only tackle the extreme versions of these policies in which there is no wage discrimination (rather than limited wage discrimination) in equilibrium for reasons of analytic tractability. Under the policy, the firm takes as given thresholds on the size of the gap in wages or employment shares that will trigger the punishment. Assume that the wage penalty will be imposed if the maximum wage gap within the firm exceeds the threshold ε_w. Since type $h = 0$ will receive the lowest wage in equilibrium, to avoid the penalty

$$\max_h |w(h) - w(0)| \leq \varepsilon_w.$$

Letting $n(h)$ denote the number of workers of type h hired by the firm, assume that the hiring penalty will be imposed if the employment share of type h deviates from the population average by more than δ, and hence

$$\left| \frac{n(h)}{\sum_h n(h)} - \frac{\Phi(h)}{\sum_h \Phi(h)} \right| \leq \delta.$$

We will assume that the punishment is sufficiently dire, that the firm will never choose to violate these thresholds.

We analyze the more general case in appendix B.1, but here focus on the limiting case in which the thresholds ε_w and δ converge to zero. In this case, the firm will simply take as given the economy-wide wage w^* at which it can hire a representative worker. We assume that the government regulates the insurance market determining the extent of coverage by health type, $x(e, h)$, subject to the requirement that the offered health insurance contracts exactly break even, either health type by health type (in the absence of a no prior conditions law) or in expectation across health types (in the presence of the no prior conditions law).

Perfect competition drives down equilibrium profits of firms to zero which pins down the equilibrium wage rate as

$$w^* = \sum_h \left\{ g(h)F(h, -x(0, h)) + (1 - g(h)) \int_0^{\bar{\varepsilon}} f(\varepsilon) \left[F(h, \varepsilon - x(\varepsilon, h)) \right] d\varepsilon \right\} \Phi(h) \tag{5}$$

The insurance premium charged to the household is

$$P(h) = g(h)x(0, h) + (1 - g(h)) \int f(\varepsilon)x(\varepsilon, h)d\varepsilon \tag{6}$$

[16] The delegation method is similar to the structure we assumed in the insurance market since insurance companies were restricted to serving their customers on a first-come-first-serve basis. This assumption to us seems more problematic in the labor market because of the idiosyncratic nature of the benefits to the worker-firm match.

in the absence of a no-prior conditions law and

$$P = \sum_h \left[g(h)x(0,h) + (1-g(h)) \int f(\varepsilon)x(\varepsilon,h)d\varepsilon \right] \Phi(h) \tag{7}$$

in its presence. Household consumption is given by

$$\begin{aligned} c(h) &= w^* - P(h) \text{ or} \\ c &= w^* - P \end{aligned}$$

depending on whether a no prior conditions law is in place or not.

Given a cross-sectional health distribution Φ the efficiently regulated health insurance contract $x(\varepsilon, h)$ is the solution to

$$\max_x \sum_h u(w^* - P(h))\Phi(h)$$

subject to (5) and (6) if the no-prior conditions restriction is not imposed on health insurance. And, (7) with $P(h) = P$ instead of (6) if the no-prior conditions restriction is imposed.

We now turn to the analysis of the model, starting with a static version in which by construction the choice of effort is not distorted in equilibrium. We will show that in this case the competitive equilibrium implements an efficient allocation of health expenditures, but fails to provide efficient consumption insurance against prior health conditions, that is against cross-sectional variation in h. We then argue that a combination of a strict wage non-discrimination law (achieved unrealized sanctions) and a no prior conditions law result in efficient consumption insurance in the competitive equilibrium, restoring full efficiency of allocations in the regulated market economy.

3 Analysis of the Static Model

We now turn to the analysis of the static version of our model, and we will characterize both efficient and equilibrium allocations (in the absence and presence of the nondiscrimination policies). The purpose of this analysis is two-fold. First, it will result in the characterization of the optimal and equilibrium health insurance contract, a key ingredient for our dynamic model. Second, the analysis will demonstrate that in the short run (that is statically) the combination of both policies is ideally suited to provide full consumption insurance in the regulated market equilibrium, and thus restore full efficiency of the market outcome. The static benefits of these policies are then traded off against the adverse dynamic consequences on the health distribution, as our analysis of the dynamic model will uncover in the next section.

3.1 Social Planner Problem

Given an initial cross-sectional distribution over health status in the population $\Phi(h)$ the social planner maximizes utilitarian social welfare. The social planner problem is therefore given by

$$U^{SP}(\Phi) = \max_{e(h), x(\varepsilon,h), c(\varepsilon,h) \geq 0} \sum_h \left\{ -q(e(h)) + g(h)u(c(0,h)) + (1-g(h)) \int f(\varepsilon)u(c(\varepsilon,h))d\varepsilon \right\} \Phi(h)$$

subject to

$$\begin{aligned} & \sum_h \left\{ g(h)c(0,h) + (1-g(h)) \int f(\varepsilon)c(\varepsilon,h)d\varepsilon + g(h)x(0,h) + (1-g(h)) \int f(\varepsilon)x(\varepsilon,h)d\varepsilon \right\} \Phi(h) \\ \leq & \sum_h \left\{ g(h)F(h, -x(0,h)) + (1-g(h)) \int f(\varepsilon)F(h, \varepsilon - x(\varepsilon,h))d\varepsilon \right\} \Phi(h) \end{aligned}$$

We summarize the optimal solution to the static social planner problem in the following proposition, whose proof follows directly from the first order conditions and assumption 4 (see Appendix A).

Proposition 5 *The solution to the social planner problem* $\{c^{SP}(\varepsilon,h), x^{SP}(\varepsilon,h), e^{SP}(h)\}_{h\in H}$ *is given by*

$$\begin{aligned} e^{SP}(h) &= 0 \\ c^{SP}(\varepsilon,h) &= c^{SP} \\ x^{SP}(\varepsilon,h) &= \max\left[0, \varepsilon - \bar{\varepsilon}^{SP}(h)\right] \end{aligned}$$

where the cutoffs satisfy

$$-F_2(h, \bar{\varepsilon}^{SP}(h)) = 1, \tag{8}$$

and the first best consumption level is given by

$$c^{SP} = \sum_h \left[g(h)F(h,0) + (1-g(h)) \int_0^{\bar{\varepsilon}} f(\varepsilon) \left[F(h, \varepsilon - x^{SP}(\varepsilon,h)) - x^{SP}(\varepsilon,h) \right] d\varepsilon \right] \Phi(h) \tag{9}$$

The optimal cutoff $\{\bar{\varepsilon}^{SP}(h)\}$ *is increasing in* h*, strictly so if* $F_{12}(h,y) > 0$.

The social planner finds it optimal to not have the household exercise (given that there are no dynamic benefits from doing so in the static model) and to provide full consumption insurance against adverse health shocks ε, but also against bad prior health conditions as consumption is constant in h.

The optimal level of health expenditure and its implications on production is graphically presented in Figure 4. The optimal medical expenditures take a simple cutoff rule: small health shocks $\varepsilon < \bar{\varepsilon}^{SP}(h)$ are not treated at all, but all larger shocks are fully treated up to the threshold $\bar{\varepsilon}^{SP}(h)$ (Figure 4(b)). If the impact of health shocks on productivity is less severe for healthy households ($F_{12}(h,y) > 0$, as in Figure 4(a)), then the social planner finds it optimal to "insure" healthier households less, in the sense of undoing less of the negative health shocks ε through medical treatment $x(\varepsilon,h)$. This health expenditure policy function leads to production function as shown in the Figure 4(c).

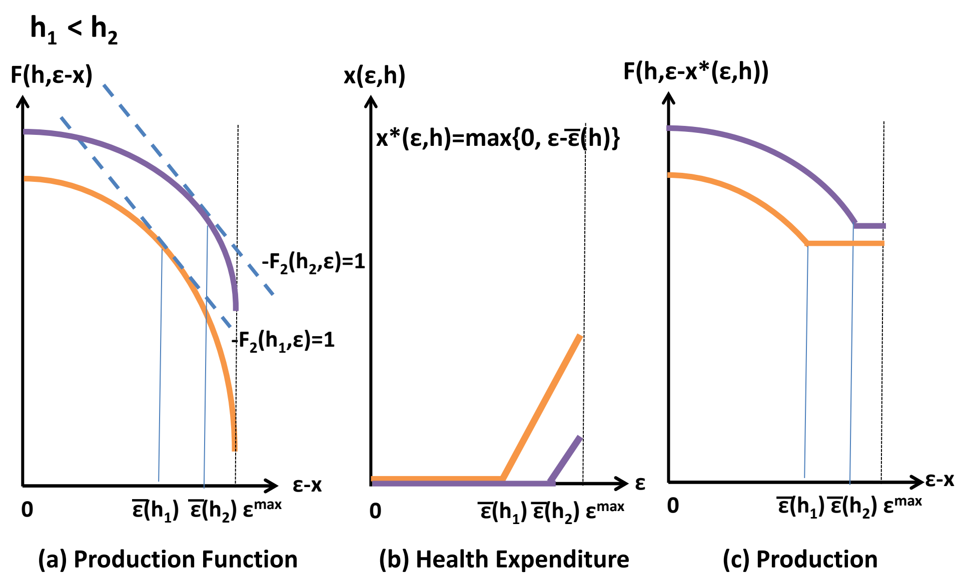

Figure 4: Optimal Health Expenditure and Production

3.2 Competitive Equilibrium

As in the social planner problem there is no incentive for households to exercise in the static model, and thus $e(h) = 0$. As described in section 2.4 the equilibrium wage and health insurance contract solves

$$U^{CE}(h) = \max_{w(h),x(\varepsilon,h),P(h)} u\left(w(h) - P(h)\right) \tag{10}$$

s.t.

$$P(h) = g(h)x(0,h) + (1-g(h)) \int_0^{\bar{\varepsilon}} f(\varepsilon)x(\varepsilon,h)d\varepsilon \tag{11}$$

$$w(h) = g(h)F(h,-x(0,h)) + (1-g(h)) \int_0^{\bar{\varepsilon}} f(\varepsilon)F(h,\varepsilon - x(\varepsilon,h))d\varepsilon \tag{12}$$

The following proposition characterizes the solution to this problem:

Proposition 6 *The unique equilibrium health insurance contract and associated consumption are given by*

$$x^{CE}(\varepsilon,h) = \max\left[0, \varepsilon - \bar{\varepsilon}^{CE}(h)\right] \tag{13}$$

$$c^{CE}(\varepsilon,h) = c^{CE}(h) = w^{CE}(h) - P^{CE}(h) \tag{14}$$

$$P^{CE}(h) = (1-g(h)) \int_{\bar{\varepsilon}^{CE}(h)}^{\bar{\varepsilon}} f(\varepsilon)\left[\varepsilon - \bar{\varepsilon}^{CE}(h)\right] d\varepsilon \tag{15}$$

$$w^{CE}(h) = g(h)F(h,0) + (1-g(h)) \int_0^{\bar{\varepsilon}} f(\varepsilon)F(h,\varepsilon - x(\varepsilon,h))d\varepsilon \tag{16}$$

and the cutoff satisfies

$$-F_2(h, \bar{\varepsilon}^{CE}(h)) = 1 \tag{17}$$

Proof. See Appendix ■

We immediately obtain the following

Corollary 7 *The competitive equilibrium implements the socially efficient health expenditure allocation since $\bar{\varepsilon}^{CE}(h) = \bar{\varepsilon}^{SP}(h)$ for all $h \in H$.*

Corollary 8 *The cutoff $\bar{\varepsilon}^{CE}(h)$ is increasing in h, strictly so if $F_{12}(h,y) > 0$.*

While it follows trivially from our assumptions that the worker's net pay, $w(h) - P(h)$, is increasing in h, it is not necessarily true that his gross wage, $w(h)$, is increasing in h as well since optimal health expenditures are decreasing in health status. We analyze the behavior of gross wages $w(h)$ with respect to health status further in Appendix C, where we provide a sufficient condition for the wage schedule to be monotonically increasing in h.

In any case, the previous results show that in the static case the *only* source of inefficiency of the competitive equilibrium comes from the inefficient lack of consumption insurance against adverse prior health conditions h. This can be seen by noting that

$$\begin{aligned} c^{SP} &= \sum_h \left\{ g(h)F(h,0) + (1-g(h)) \int_0^{\bar{\varepsilon}} f(\varepsilon)\left[F(h,\varepsilon - x^{SP}(\varepsilon,h)) - x^{SP}(\varepsilon,h)\right] d\varepsilon \right\} \Phi(h) \\ &= \sum_h \left[w^{CE}(h) - P^{CE}(h)\right] \Phi(h) = \sum_h c^{CE}(h)\Phi(h) \end{aligned}$$

In contrast to what will be the case in the dynamic model, effort trivially is not distorted in the equilibrium, relative to the allocation the social planner implements (since in both cases $e^{SP} = e^{CE} = 0$). Furthermore the equilibrium allocation of health expenditures is efficient, on account that the firm bundles the determination of wages and the provision of health insurance, and thus internalizes the positive effects of health spending $x(\varepsilon, h)$ on worker productivity.

Given these results it is plausible to expect, within the context of the static model, that policies preventing competitive equilibrium wages $w^{CE}(h)$ to depend on health status (a wage non-discrimination law) and insurance premia $P^{CE}(h)$ to depend on health status (a no prior conditions law) will restore full efficiency of the policy-regulated competitive equilibrium by providing full consumption insurance. We will show next that this is indeed the case, providing a normative justification for the two policy interventions within the static version of our model.

3.3 Competitive Equilibrium with a No Prior Condition Law

As discussed above in order to effectively implement a no prior conditions law the government has to regulate the health insurance provision done by firms or insurance company. The regulatory authority solves the problem of

$$\begin{aligned} U^{NP}(\Phi) &= \max_{x(\varepsilon,h)} \sum_h u(w(h)-P)\Phi(h) \quad (18) \\ & s.t. \\ P &= \sum_h \left[g(h)x(0,h) + (1-g(h)) \int f(\varepsilon)x(\varepsilon,h)d\varepsilon \right] \Phi(h) \quad (19) \\ w(h) &= g(h)F(h,-x(0,h)) + (1-g(h)) \int_0^{\bar{\varepsilon}} f(\varepsilon)F(h,\varepsilon - x(\varepsilon,h))d\varepsilon \quad (20) \end{aligned}$$

The next proposition characterizes the resulting regulated equilibrium allocation

Proposition 9 *The equilibrium health expenditures under a no-prior condition law satisfies, for each $\hat{h} \in H$*

$$x^{NP}(\varepsilon, \tilde{h}) = \max[0, \varepsilon - \bar{\varepsilon}^{NP}(\tilde{h})]$$

with cutoffs uniquely determined by

$$-F_2(\tilde{h}, \bar{\varepsilon}^{NP}(\tilde{h})) = \frac{\sum_h u'(w^{NP}(h) - P^{NP})\Phi(h)}{u'(w(\tilde{h}) - P^{NP})}.$$

The equilibrium wage is given by

$$w^{NP}(\tilde{h}) = g(\tilde{h})F(h,0) + (1-g(\tilde{h})) \int_0^{\bar{\varepsilon}} f(\varepsilon)[F(\tilde{h}, \varepsilon - x^{NP}(\varepsilon,\tilde{h}))]d\varepsilon.$$

Moreover, optimal cutoffs are increasing in health status.

Proof. See Appendix ■

Note that the health expenditure level is no longer efficient as the government provides partial consumption insurance against initial health status by efficiently choosing the cutoff levels $\bar{\varepsilon}^{NP}(h)$. In fact, as shown in the next proposition, it is efficient to over-insure households with *bad* health status and under-insure those with *good* health status, relative to the first-best.

Proposition 10 *Let $\tilde{h}$ be the health status whose marginal utility of consumption is equal to the population average, i.e. for $\tilde{h}$,*

$$-F_2(\tilde{h}, \bar{\varepsilon}(\tilde{h})) = \frac{\sum_h u'(w(h) - P)\Phi(h)}{u'(w(\tilde{h}) - P)} = 1 \quad (21)$$

holds.[17] *Then,*

$$\begin{aligned} \bar{\varepsilon}^{NP}(h) < \bar{\varepsilon}^{SP}(h), &\qquad \textit{for } h < \tilde{h} \\ \bar{\varepsilon}^{NP}(h) = \bar{\varepsilon}^{SP}(h), &\qquad \textit{for } h = \tilde{h} \\ \bar{\varepsilon}^{NP}(h) > \bar{\varepsilon}^{SP}(h), &\qquad \textit{for } h > \tilde{h}, \end{aligned}$$

The cutoffs $\bar{\varepsilon}(h)$ are strictly monotonically increasing in health status h.

[17] For the purpose of the proposition it does not matter whether $\tilde{h} \in H$ or not.

Proof. See Appendix. ■

This feature of the optimal health expenditure with a no prior conditions law also indicates that mandatory participation in the health insurance contract is an important part of government regulation, since in the allocation above the healthy cross-subsidize the unhealthy *and* they are insured less than the unhealthy.

3.4 Competitive Equilibrium with a No Wage Discrimination Law

The equilibrium with a no wage discrimination law (and with costs not materialized in equilibrium) is determined by the solution to the program:

$$\begin{aligned}
U^{ND}(\Phi) &= \max_{x(\varepsilon,h)} \sum_h u(w - P(h))\Phi(h) \\
& \quad s.t. \\
P(h) &= \left[g(h)x(0,h) + (1-g(h)) \int f(\varepsilon)x(\varepsilon,h)d\varepsilon\right] \\
w &= \sum_h \left\{g(h)F(h,-x(0,h)) + (1-g(h)) \int_0^{\bar{\varepsilon}} f(\varepsilon)F(h,\varepsilon - x(\varepsilon,h))d\varepsilon\right\}\Phi(h)
\end{aligned}$$

Proposition 11 *The equilibrium health expenditures under a no-wage discrimination law alone satisfies, for each $\tilde{h} \in H$*

$$x^{ND}(\varepsilon,\tilde{h}) = \max\left[0, \varepsilon - \bar{\varepsilon}^{ND}(\tilde{h})\right]$$

with cutoffs determined by

$$-F_2(\tilde{h},\bar{\varepsilon}^{ND}(\tilde{h})) = \frac{u'(w^{ND} - P(\tilde{h}))}{\sum_h u'(w^{ND} - P(h))\Phi(h)}.$$

The equilibrium wage is given by

$$w^{ND} = \sum_h \left[g(h)\left[F(h,0)\right] + (1-g(h)) \int_0^{\bar{\varepsilon}} f(\varepsilon)\left[F(h,\varepsilon - x^{ND}(\varepsilon,h))\right]d\varepsilon\right]\Phi(h)$$

Proof. Follows directly from the first order conditions of the program. ■

Unlike in the no prior conditions case, we cannot establish monotonicity in the cutoffs $\bar{\varepsilon}^{ND}(\tilde{h})$. Note that under a no prior conditions law the regulatory authority partially insures consumption of the unhealthy by allocating higher medical expenditure to them. Under a no wage discrimination law instead, there are two opposing forces, preventing us from establishing monotonicity in cutoffs $\bar{\varepsilon}^{ND}(h)$ across health groups h. On one hand, a one unit increase in medical expenditure $P(h)$ is more costly to the unhealthy since marginal utility of consumption is higher for this group. On the other hand, production efficiency calls for higher medical expenditure for the unhealthy, given our assumption of $F_{12} > 0$ (as was the case for the no prior conditions law). Thus the cutoffs $\bar{\varepsilon}^{ND}(h)$ need not be monotone in h.

Finally, combining both a no-wage discrimination law and a no-prior conditions legislation restores efficiency of the regulated equilibrium since both policies in conjunction provide full consumption insurance against bad health realizations h. This is the content of the next

Corollary 12 *The unique competitive equilibrium allocation in the presence of both a no wage discrimination and a no prior conditions law implements the socially efficient allocation in the static model.*

Proof. The equilibrium is the solution to

$$\begin{aligned}
& \max_{x(\varepsilon,h)} \sum_h u(w^* - P)\Phi(h) \\
& \quad s.t. \\
P &= \sum_h \left[g(h)x(0,h) + (1-g(h)) \int f(\varepsilon)x(\varepsilon,h)d\varepsilon\right]\Phi(h) \\
w^* &= \sum_h \left\{g(h)F(h,-x(0,h)) + (1-g(h)) \int_0^{\bar{\varepsilon}} f(\varepsilon)F(h,\varepsilon - x(\varepsilon,h))d\varepsilon\right\}\Phi(h).
\end{aligned}$$

The result then follows trivially from the fact that this maximization problem is equivalent to the social planner problem analyzed above. The no prior conditions law equalizes health insurance premia P across health types, the no wage discrimination law implements a common wage w^* across health types, and the (assumed) efficient regulation of the health insurance market assures that the health expenditure schedule is efficient as well. ■

3.5 Summary of the Analysis of the Static Model

The competitive equilibrium implements the efficient health expenditure allocation but does not insure households against initial health conditions. Both a no-prior conditions law and a no-wage discrimination law provide partial, but not complete, consumption insurance against this risk, without distorting the effort level. The health expenditure schedule is distorted under both policies, relative to the social optimum, as the government provides partial consumption insurance through health expenditures. Only both laws in conjunction implement full consumption insurance against initial health conditions h and thus restore the first best allocation in the static model. The policies are thus fully successful in what they are designed to achieve in a static world (partially due to the fact that additional government regulation severely restricted the options of firms to circumvent the government policies).

4 Analysis of the Dynamic Model

We now study a dynamic version of our economy. Both in terms of casting the problem, as well as in terms of its computation we make use of the fact that there is no aggregate risk (due to the continuum of agents cum law of large numbers assumption). Therefore the sequence of cross-sectional health distributions $\{\Phi_t\}_{t=0}^T$ is a deterministic sequence. Furthermore, conditional on a distribution Φ_t today the health distribution tomorrow is completely determined by the effort choice $e_t(h)$ of households[18] (or the social planner), so that we can write

$$\Phi_{t+1} = H(\Phi_t; e_t(.)) \tag{22}$$

where the time-invariant function H is in turn completely determined by the Markov transition function $Q(h'; h, e)$. The initial distribution Φ_0 is an initial condition and exogenously given.

Under each policy, given a sequence of aggregate distributions $\{\Phi_t\}_{t=0}^T$ we can solve an appropriate dynamic maximization problem of an individual household for the sequence of optimal effort decisions $\{e_t(h)_{h\in H}\}_{t=0}^T$ which in turn imply a new sequence of aggregate distributions via (22). Our computational algorithm for solving competitive equilibria then amounts to iterating on the sequences $\{\Phi_t, e_t\}$. Within each period the timing of events follows exactly that of the static problem in the previous section.

4.1 Social Planner Problem

The dynamic problem of the social planner is to solve

$$V(\Phi_0) = \max_{\{e_t(h)\}} \sum_{t=0}^{T} \beta^t \left\{ U^{SP}(\Phi_t) - \sum_h q(e_t(h))\Phi_t(h) \right\}$$

where

$$\Phi_{t+1} = H(\Phi_t; e_t(h))$$

and

$$\begin{aligned} U^{SP}(\Phi) &= \max_{x(\varepsilon,h),c(\varepsilon,h)} \sum_h \left\{ g(h)u(c(0,h)) + (1-g(h)) \int f(\varepsilon)u(c(\varepsilon,h))d\varepsilon \right\} \Phi(h) \\ &= u(c^{SP}(\Phi)) \end{aligned}$$

[18] We assert here that the optimal effort in period t is only a function of the current individual health status h. We will discuss below the assumptions required to make this assertion correct.

is the solution to the static social planner problem characterized in section (3.1):

$$x^{SP}(\varepsilon,h) = \max\left[0, \varepsilon - \bar{\varepsilon}^{SP}(h)\right]$$

with cutoffs defined by

$$-F_2(h,\bar{\varepsilon}^{SP}(h)) = 1 \tag{23}$$

and consumption given by

$$c^{SP}(\Phi) = \sum_h \left[g(h)F(h,0) + (1-g(h))\int_\varepsilon f(\varepsilon)\left[F(h,\varepsilon - x^{SP}(\varepsilon,h)) - x^{SP}(\varepsilon,h)\right] d\varepsilon\right]\Phi(h).$$

We now want to characterize the optimal effort choice by the social planner, the key dynamic decision in our model both in the planner problem and the competitive equilibrium. In contrast to households in the competitive equilibrium, the social planner fully takes into account the effect of effort choices today on the aggregate health distribution and thus aggregate consumption tomorrow.

A semi-recursive formulation of the problem is useful to characterize the optimal effort choice, but also to explain the computational algorithm for the social planner problem. For a given cross-sectional distribution Φ_t at the beginning of period t the social planner solves:

$$\begin{aligned}
V_t(\Phi_t) &= u(c_t) + \max_{e_t(h)_{h\in H},}\left\{-\sum_h q(e_t(h))\Phi_t(h) + \beta V_{t+1}(\Phi_{t+1})\right\} \\
\text{s.t. } c_t &= c^{SP}(\Phi_t) \\
\Phi_{t+1}(h') &= \sum_h Q(h';h,e_t(h))\Phi(h)
\end{aligned} \tag{24}$$

In appendix D we discuss how we solve this problem numerically, iterating on sequences $\{c_t, e_t(h), \Phi_t(h)\}_{t=0}^T$ from the terminal condition $V_T(\Phi_T) = u(c_T)$. To characterize the optimal effort choice, for an arbitrary time period t we obtain the first order condition:

$$\begin{aligned}
q'(e_t(h))\Phi_t(h) &= \beta\sum_{h'}\frac{\partial V_{t+1}(\Phi_{t+1})}{\partial\Phi_{t+1}(h')}\cdot\frac{\partial\Phi_{t+1}(h')}{\partial e_t(h)} \\
&= \beta\sum_{h'}\frac{\partial V_{t+1}(\Phi_{t+1})}{\partial\Phi_{t+1}(h')}\cdot\frac{\partial Q(h';h,e_t(h))}{\partial e_t(h)}\Phi_t(h),
\end{aligned}$$

This simplifies to

$$q'(e_t(h)) = \beta\sum_{h'}\frac{\partial V_{t+1}(\Phi_{t+1})}{\partial\Phi_{t+1}(h')}\cdot\frac{\partial Q(h';h,e_t(h))}{\partial e_t(h)}. \tag{25}$$

Thus the marginal cost of extra effort $q'(e_t(h))$ is equated to the marginal benefit, the latter being given by the the change that effort has on the health distribution tomorrow[19], $\frac{\partial Q(h';h,e_t(h))}{\partial e_t(h)}$, times the benefit of a better health distribution $\frac{\partial V_{t+1}(\Phi_{t+1})}{\partial\Phi_{t+1}(h')}$ from tomorrow on. By assumption 1, $q'(0) = 0$, and assumption 3 guarantees that the right hand side of equation (25) is strictly positive. Therefore $e_t(h) > 0$ for all t and all $h \in H$.

The envelope theorem gives the benefit of a better health distribution as:

$$\frac{\partial V_t(\Phi_t)}{\partial\Phi_t(h)} = u'(c_t)\cdot\Psi(h) - q(e_t(h)) + \beta\sum_{h'}\frac{\partial V_{t+1}(\Phi_{t+1})}{\partial\Phi_{t+1}(h')}\cdot Q(h';h,e_t(h)). \tag{26}$$

Here $\Psi(h)$ gives the net of health expenditure output that, on average, an individual of health status h delivers to the social planner.[20]

[19] Note that for a suitable functional form of Q we can compute $\frac{\partial Q(h';h,e_t(h))}{\partial e_t(h)}$ analytically.

[20] Note that

$$\Psi(h) = \left[g(h)F(h,0) + (1-g(h))\int_\varepsilon f(\varepsilon)\left[F(h,\varepsilon - x^{SP}(\varepsilon,h)) - x^{SP}(\varepsilon,h)\right] d\varepsilon\right]$$

is exclusively determined by the optimal cut-off rule for health expenditures which is independent of c_t or Φ_t.

4.2 Competitive Equilibrium without Policy

In our model, since absent wage and health insurance policies households do not interact we can solve the dynamic programming problem of each household independent of the rest of society. The state variables of the household are her current health status and the households' (cohorts') preference for exercise:

$$v_t(h) = U^{CE}(h) + \max_{e_t(h)} \left\{ -q(e_t(h)) + \beta \sum_{h'} Q(h'; h, e_t(h)) v_{t+1}(h') \right\} \tag{27}$$

where

$$\begin{aligned} U^{CE}(h) &= \max_{x(\varepsilon,h),w(h),P(h)} u(w(h) - P(h)) \\ & \quad s.t. \\ w(h) &= g(h)F(h, -x(0,h)) + (1-g(h)) \int_0^{\bar{\varepsilon}} f(\varepsilon)F(h, \varepsilon - x(\varepsilon,h))d\varepsilon \\ P(h) &= g(h)x(0,h) + (1-g(h)) \int_0^{\bar{\varepsilon}} f(\varepsilon)x(\varepsilon,h)d\varepsilon \end{aligned}$$

is the solution to the static equilibrium problem in section 3.2:

$$\begin{aligned} x^{CE}(\varepsilon,h) &= \max\left[0, \varepsilon - \bar{\varepsilon}^{CE}(h)\right] \\ c^{CE}(h) &= w^{CE}(h) - P^{CE}(h) \\ P^{CE}(h) &= (1-g(h)) \int_{\bar{\varepsilon}^{CE}(h)}^{\bar{\varepsilon}} f(\varepsilon) \left[\varepsilon - \bar{\varepsilon}^{CE}(h)\right] d\varepsilon \\ w^{CE}(h) &= g(h)F(h,0) + (1-g(h)) \int_0^{\bar{\varepsilon}} f(\varepsilon)F(h, \varepsilon - x(\varepsilon,h))d\varepsilon \end{aligned}$$

with cutoff:

$$-F_2(h, \bar{\varepsilon}^{CE}(h)) = 1$$

Note again that the provision of health insurance is socially efficient in the competitive equilibrium.

In contrast to the social planner problem, and in contrast to the competitive equilibrium with no-wage discrimination or no-prior conditions law, in the unregulated competitive equilibrium there is no interaction between the maximization problems of individual households. Thus the dynamic household maximization problem can be solved independent of the evolution of the cross-sectional health distribution. It is a simple dynamic programming problem with terminal value function

$$v_T(h) = U^{CE}(h)$$

and can be solved by straightforward backward iteration.[21]

Given the solution $\{e_t(h)\}$ of the household dynamic programming problem and given an initial distribution Φ_0 the dynamics of the health distribution is then determined by the aggregate law of motion (22). The optimal choice $e_t(h)$ solves the first order condition

$$q'(e_t(h)) = \beta \sum_{h'} \frac{\partial Q(h'; h, e_t(h))}{\partial e_t(h)} v_{t+1}(h') \tag{28}$$

Note that at time t when the decision $e_t(h)$ is taken v_{t+1} is known. Furthermore, given knowledge of v_{t+1} and the optimal e_t the period t value function v_t is determined by (27). As in the social planner problem, by assumptions 1 and 3 effort $e_t(h)$ is positive for all t and h.

[21] It is certaintly straightforward to extend the terminal value function to

$$v_T(h) = U^{CE}(h) + \nu(h)$$

where $\nu(.)$ is an arbitrary function that captures the remaining lifetime utility from retirement. It would give us additional free parameters (to perhaps improve the model fit with respect to the terminal health distribution), but $\nu(.)$ is not needed to close our model.

4.3 Competitive Equilibrium with a No Prior Condition Law

As discussed above, we assume that the government in every period t *takes as given the health distribution* Φ_t and enforces the no prior condition law and regulates health insurance contracts efficiently, as in the static analysis of section 3.3. We now make explicit that the solution of the static government regulation problem (18)-(20) is a function of the cross-sectional health distribution,

$$x^{NP}(\varepsilon,\tilde{h};\Phi_t) = \max[0, \varepsilon - \bar{\varepsilon}^{NP}(\tilde{h};\Phi_t)] \tag{29}$$

with cutoffs for each $\tilde{h} \in H$ determined by

$$-F_2(\tilde{h},\bar{\varepsilon}^{NP}(\tilde{h};\Phi_t))u'(w(\tilde{h};\Phi_t) - P(\Phi_t)) = \sum_h u'(w(h;\Phi_t) - P(\Phi_t))\Phi_t(h) := Eu'(\Phi_t) \tag{30}$$

and

$$\begin{aligned} w^{NP}(h;\Phi_t) &= g(h)F(h,0) + (1-g(h))\int f(\varepsilon)[F(h,\varepsilon - x^{NP}(\varepsilon,h;\Phi_t))]d\varepsilon & (31) \\ P^{NP}(\Phi_t) &= \sum_h \left[g(h)x^{NP}(0,h;\Phi_t) + (1-g(h))\int f(\varepsilon)x^{NP}(\varepsilon,h;\Phi)d\varepsilon\right]\Phi_t(h) & (32) \end{aligned}$$

In order for the household to solve her dynamic programming problem she only needs to know the sequence of wages and health insurance premia $\{w_t(h), P_t\}$, but not necessarily the sequence of distributions that led to it. Given such a sequence the dynamic programming problem of the household then reads as

$$v_t(h) = u(w_t(h) - P_t) + \max_{e_t(h)}\left\{-q(e_t(h)) + \beta\sum_{h'} Q(h';h,e_t(h))v_{t+1}(h')\right\} \tag{33}$$

with terminal condition $v_T(h) = u(w_T(h) - P_T)$. As before the optimality condition reads as

$$q'(e_t(h)) = \beta\sum_{h'}\frac{\partial Q(h';h,e_t(h))}{\partial e_t(h)}v_{t+1}(h').$$

and thus equates the marginal cost of providing effort, $q'(e)$ with the marginal benefit of an improved health distribution tomorrow. In appendix E we describe a computational algorithm to solve the dynamic model with a no-prior conditions law.

4.4 Competitive Equilibrium with a No Wage Discrimination Law

The main difference to the previous section is that now the static health insurance contract and premium are given by health spending

$$x^{ND}(\varepsilon,\tilde{h};\Phi_t) = \max[0, \varepsilon - \bar{\varepsilon}^{ND}(\tilde{h};\Phi_t)] \tag{34}$$

with cutoffs for each $\tilde{h} \in H$ determined by

$$-F_2(h,\bar{\varepsilon}^{ND}(h))Eu'_t = u'(w^{ND}(\Phi_t) - P^{ND}(h,\Phi_t)) \tag{35}$$

where

$$Eu'_t := \sum_h u'(w^{ND}(\Phi_t) - P^{ND}(h,\Phi_t))\Phi_t(h). \tag{36}$$

The equilibrium wage is given by

$$w^{ND}(\Phi_t) = \sum_h\left\{g(h)F(h,0) + (1-g(h))\int f(\varepsilon)[F(h,\varepsilon - x^{ND}(\varepsilon,h;\Phi_t))]d\varepsilon\right\}\Phi_t(h) \tag{37}$$

and the equilibrium premium and equilibrium health insurance premium that depend on whether a no prior conditions law is in place as well. Without such policy the premia are given as

$$P^{ND}(h;\Phi_t) = P^{ND}(h) = (1-g(h)) \int f(\varepsilon) x^{ND}(\varepsilon, h) d\varepsilon \tag{38}$$

whereas with both policies in place the premium is determined by

$$P^{ND}(\Phi_t) = \sum_h \left[(1-g(h)) \int f(\varepsilon) x^{ND}(\varepsilon, h) d\varepsilon \right] \Phi_t(h) \tag{39}$$

For a given sequence of wages $\{w_t, P_t(h)\}$ the dynamic problem of the household reads as before:

$$v_t(h) = u(w_t - P_t(h)) + \max_{e_t(h)} \left\{ -q(e_t(h)) + \beta \sum_{h'} Q(h'; h, e_t(h)) v_{t+1}(h') \right\}$$

and the terminal condition $v_T(h) = u(w_T - P_T(h))$, first order conditions and updating of the value function for this version of the model are exactly the same, mutatis mutandis, as under the previous policy. In appendix E we discuss the algorithm to solve this version of the model.

If both policies are in place simultaneously, we can give a full analytical characterization of the equilibrium without resorting to any numerical solution procedure. We do so in the next

Proposition 13 *Suppose there is a no wage discrimination and a no prior condition law in place simultaneously Then*

$$e_t(h) = 0 \text{ for all } h, \text{ and all } t.$$

The provision of health insurance is socially efficient. From the initial distribution Φ_0 the health distribution in society evolves according to (22) *with $e_t(h) \equiv 0$.*

The proof is by straightforward backward induction and is given in Appendix A. In the presence of both policies there are no incentives, either through wages or health insurance premia, to exert effort to lead a healthy life. Since effort is costly, households won't provide any in the regulated dynamic competitive equilibrium, with the associated adverse long run consequences for the distribution of health in society. In contrast, with only one policy in place, as well as in the unregulated equilibrium and in the solution to the social planner problem optimal effort is positive, as we show in Appendix **??**.

Equipped with these theoretical results and the numerical algorithms to solve the various versions of our model we now map our model to cross-sectional health and exercise data from the PSID to quantify the effects of government regulations on the evolution of the cross-sectional health distribution, as well as aggregate production, consumption and health expenditures.

5 Bringing the Model to the Data

5.1 Augmenting the Model

In order to map our model successfully into the data, we augment our theoretical model in four aspects.

First, in order to capture large medical expenditure in data,[22] we introduce a health expenditure shock z. Households in the augmented model are assumed to either do not get a shock, get either a z-shock, or an ε shock, but not both. When households receive a z-shock, they *have to* spend z; otherwise, they die (or equivalently, incur a prohibitively large utility cost). We denote by $\mu_z(h)$ and $\kappa(h)$, its conditional mean and the probability of receiving a positive z-shock.[23]

[22] In the model, the only benefit of spending resources on health is to offset the negative productivity consequences of the adverse health shocks ε. This modelling assumption implies that it is never optimal to incur health expenditures larger than what the worker could produce in a given period.

[23] More detailed analysis of the change in household's wage and health insurance premium are described in Section F.1.

Second, we permit earnings in the model to depend on the age and education level of the worker. The production function now takes the cohort's age t, education $educ$, health h, and net health shock $\varepsilon - x$ as its inputs: $F(t, educ, h, \varepsilon - x)$. Note that by proceeding in this way we implicitly assume that even in the presence of a wage discrimination law individuals with higher education can be paid more, and that health insurance companies can charge differential premia to individuals with heterogeneous education levels even in the presence of a no-prior conditions law.

Third, for the model to have a change to generate the heterogeneity in exercise levels for a given health status h observed in the data we introduce preference shocks to the disutility from effort. The cost of exerting effort is now given as $\gamma q(e)$ where $\gamma \in \Gamma$ is a individual-specific preference shock that is drawn from the finite set Γ at the beginning of life and remains constant during the individual's working life.[24] The initial joint distribution over health h, education $educ$ and preferences γ is denoted by $\Phi_0(h, \gamma, educ)$. Note that since γ only affects the disutility of effort which is separable from the utility of consumption, the analysis of the static model in section 3 remains completely unchanged (and so do the optimal health insurance contracts and health expenditure allocations). In the analysis of the dynamic model, since γ is a permanent shock, all expressions involving $q(.)$ turn into $\gamma q(.)$ and the equilibria have to be solved for each γ (and each $educ$ separately), but there is no interaction between the different $(\gamma, educ)$ types.[25] In order to obtain a meaningful welfare comparison with the socially optimal allocations we also solve the social planner problem separately for each $(\gamma, educ)$ combination, therefore ruling out ex ante social insurance against bad initial $(\gamma, educ)$ draws.

As last extension of the theoretical model, since it aims at explaining household behavior during their working lives, we endow agents with a health-dependent continuation utility at period T, the last period of their planning horizon (which we interpret as the retirement age).[26]

It is again worth stressing that none of these changes affect the theoretical properties derived in the previous sections. In the rest of this section, we use our (extended) equilibrium model to find parameters to match PSID data on health, expenditure and exercise in 1999. Once the model is parameterized, in the next section we use it to analyze the positive and normative short- and long-run consequences of introducing non-discrimination legislation. In the main body of the paper, we describe the procedures concisely, relegating all detailed data description and estimation procedures to the Appendix F.

5.2 Parameter Estimation

The risk aversion parameter and discount factor are set *a priori* to $\sigma = 2$, and $\beta = 0.96$ per annum.

In order to parameterize the model we proceed in two steps. In the first step, we estimate the health transition function $Q(h'|h, e)$, the probabilities $(g(h), \kappa(h))$ of receiving health shocks ε and z, as well as the effect of the z-shocks on productivity given by ρ, directly from the PSID data. In the second step, the remaining model parameters (mainly those governing the production function F, the ε-shock distribution $f(\varepsilon)$ and preferences) are estimated by matching selected moments of the model to their empirical counterparts.

5.2.1 Parameters Estimated Directly from the Data

A. **Health Transition Function, $Q(h'|h, e)$:** PSID includes measures of light and heavy exercise levels[27] starting in 1999 which we use to estimate health transition functions.

   We denote by e^l and e^h the frequency of light and heavy exercise levels, and assume the following parametric function for health transition.

[24] It does not matter whether firms observe a worker's preference parameter γ since they engage only in short-term contracts and since h is observable (γ only affects effort and firms do not care how the employee's health evolves as they are only employed for one period.).

[25] We simply introduce heterogeneity in education and preferences in the quantitative model to account for part of the heterogeneity in outcomes observed in the data, rather than attribute all such heterogeneity to differences in age t and health h, the state variables in the theoretical model in section 4.

[26] The theoretical model implicitly assumed that this continuation utility is equal to zero, independent of health status.

[27] Number of times an individual does light physical activity (walking, dancing, gardening, golfing, bowling, etc.) and heavy physical activity (heavy housework, aerobics, running, swimming, or bicycling)

$$Q(h';h,e^l,e^h)=\begin{cases}(1+\pi(h,e^l,e^h)^{\alpha_i(h)})G(h,h'), & \text{if } h'=h+i, i\in\{1,2\}\\ (1+\pi(h,e^l,e^h))G(h,h'), & \text{if } h'=h, h>1 \text{ or } h'=h+1, h=1\\ \left(\dfrac{1-\sum\limits_{h'\geq h}Q(h'^{l},e^h)}{\sum\limits_{h'<h}G(h,h')}\right)G(h,h'), & \text{if } h'=h-1, h>1 \text{ or } h'=h, h=1\end{cases}$$

where $\pi(h,e^l,e^h)=\phi(h)(\delta e^l+(1-\delta)e^h)^{\lambda(h)}$.

Since light and heavy physical exercise can have different effects on health transition, we give weight δ on light exercise, and $(1-\delta)$ on heavy exercise. We can think of $\delta e^l+(1-\delta)e^h$ as the composite exercise level e used in the analytical part of our problem.

B. **Health Shock Probabilities, $g(h)$ and $\kappa(h)$:** In our model, $g(h)$ represents the probability of *not* having any shock, and $\kappa(h)$, probability of having a z-shock. Since we assume that households get *either* an ε or a z-shock, probability of having an ε-shock is represented by $1-g(h)-\kappa(h)$.

   From PSID, we construct the probabilities of having z-shock and ε-shock. We define households with z-shocks as those who were diagnosed with cancer, heart attack, or heart disease,[28] or who spent more on medical expenditure than their income when hit with a health shock. Households with all other heath shocks or those who missed work due to illness are categorized as having had an ε-shock.

C. **Impact of z-shock on Productivity, ρ:** Using the criterion for determining ε and z-shocks specified above, we use mean earnings of those with z-shock relative to those without health shocks to directly pin down ρ.

5.2.2 Parameters Estimated within the Model

Now, we use our model to find parameters for production function, ε and z-shock distribution, exercise preference distribution, and terminal value of health.

The structure of our model allows us to estimate the parameters in two steps. The first part of estimation which consist of finding parameters for production function and distribution of health shocks (sections A and B below), only involves static part of the model. These are independent of functions and shocks governing exercise decisions in our dynamic model. Therefore, we only use static equilibrium to match data counterparts for production and ε-shock parameters, after which we proceed to dynamic equilibrium for exercise preference distribution and terminal value of health (sections C and D).

A. **Production Function and Health Status:** We assume the following parametric function for our production technology:

$$F(t,educ,h,\varepsilon-x)=A(t,educ)h+\frac{(k-(\varepsilon-x))^{\phi(a,educ)}}{h^{\xi(a,educ)}},\quad 0<\phi(\cdot),\xi(\cdot)<1, A(\cdot)>0.$$

   The production function captures two effects of health on production: the direct effect (first term) and the indirect effect, which decreases the marginal benefit of net health shock $\varepsilon-x$ (second term). $A(t,educ)$ denotes multiplicative time and education specific effect of health on production, where $t\in\{1,2,...,7\}$ and $educ\in\{\text{less than High School, High School Grad}\}$. We also allow for differences in marginal effect of medical expenditure on production across age and education through parameters $\phi(a,educ)$ and $\xi(a,educ)$, where $a\in\{\text{Young, Old}\}$. We define *Young* as those who are between the ages of 24 and 41, which correspond to $t=1,2,3$, and the rest, *Old*. This classification divides

[28] Cancer, heart attack, and heart disease are the diseases that lead to the most medical expenditure in means compared to others.

the population in roughly half. This functional form insures that a marginal dollar spent on healthy individual is worth less than that spent on unhealthy individual ($-F_{12} < 0$).

Since income (labor earnings) in our model are direct counterparts of productivity of households with different health status, we use relative income of households across health levels $\left(\frac{w(h_2)}{w(h_1)}, \frac{w(h_3)}{w(h_1)}, \frac{w(h_4)}{w(h_1)}\right)$ and relative income of young and old to pin down $\{h_1, h_2, h_3, h_4\}$. Moreover, as $A(t, educ)$ captures the effects of age (t) and education on income, we use income conditional on age and education to pin down values of $A(t, educ)$, which are total of 14 (7×2) parameters.

We also need to find values of $\phi(\cdot)$ and $\xi(\cdot)$. As they determine cutoffs in medical expenditure for light health shocks ε conditional on health status, we use medical expenditure data to pin down these parameters. More specifically, $\phi(a, educ)$ (4 parameters) are determined to match percentage of labor earnings spent on medical expenditure for each a ($Young$ and Old) and education level, and $\xi(a, educ)$ (4 parameters), percentage of labor earnings spent on medical expenditure for each health group.

B. **Health Shock Distributions:** We need to find parameter values for ε and z shock distributions.

Under the current parametric assumptions on production function, medical expenditure on light health shocks ε is linear in shocks as $x^*(\varepsilon, h) = \max\{0, \varepsilon - \bar{\varepsilon}\}$, and the distribution of medical expenditure follows that of the shocks. French and Jones (2004) shows that the cross-sectional distribution of health care costs[29] can best be fitted by log-normal distribution (upper tail, by truncated log-normal). We assume truncated log-normal distribution for the health shocks, i.e.

$$f(\varepsilon; \mu, \sigma_\varepsilon, \underline{\varepsilon}, \bar{\varepsilon}) = \frac{\frac{1}{\epsilon\sigma_\varepsilon}\phi\left(\frac{\ln\varepsilon-\mu}{\sigma_\varepsilon}\right)}{\Phi\left(\frac{\ln\bar{\varepsilon}-\mu}{\sigma_\varepsilon}\right) - \Phi\left(\frac{\ln\underline{\varepsilon}-\mu}{\sigma_\varepsilon}\right)} \text{(Truncated Log Normal)},$$
$$\text{where } \phi \text{ and } \Phi \text{ are standard normal pdf and cdf..}$$

We find mean and standard deviation of shocks to match mean and standard deviation of medical expenditure of those with ε-shocks from data.

For mean of z-shock expenditures, we use percentage of labor income spent on catastrophic medical expenditure conditional on health status as targets.

C. **Distribution of Exercise Preference Parameters:** Given production function and ε-shock parameters, we now use dynamic model to estimate exercise preference distribution. We let $\gamma(educ) \in \{\gamma_1(educ), \gamma_2(educ)\}$, where disutility of effort is

$$\gamma q(e) = \gamma\left[\frac{1}{1-e} - e - 1\right].$$

There are total of 12 parameters that we intend to extract information from the data: values of preference parameters for each education level, $\gamma_1(educ)$ and $\gamma_2(educ)$, and measure of households with preference γ_1 conditional on education and health (which determines measure with γ_2 given health distribution from data), $p(\gamma_1|educ, h)$ (8 parameters). We use mean effort level in period 1 (ages 24-29) conditional on health (4 targets), that conditional on education (2 targets), and mean effort in period 7 (ages 60-65) conditional on education (2 targets) to find preference distribution. For values of $\gamma(educ)$, we use aggregate mean and standard deviation of effort in period 1 and measure of those with Fair and Excellent health in the last period.

D. **Marginal Value of Health at Terminal Date:** The structure of model is such that at terminal date, households have no incentive to exercise, whereas in the data, we still see a significant amount of exercise for those between the age of 60 and 65. Therefore, we let Δ_2, Δ_3, and Δ_4 be $V_T(h_2) - V_T(h_1)$, $V_T(h_3) - V_T(h_2)$, and $V_T(h_4) - V_T(h_3)$, respectively, and use them to match conditional exercise levels in the last period. The parametric assumption on our health transition function $Q(h'|h, e)$ implies that Δ_h are sufficient to obtain effort in the last period.

[29] They use HRS and AHEAD data. Health care costs include health insurance premia, drug costs and costs for hospital, nursing home care, doctor visits, dental visits and outpatient care.

The data targets and relevant parameters are summarized in Tables 9 and 10.

The parameter values are reported in Table 11, with its performance in matching calibration targets.

5.3 Model Fit

Figures 5 and 26- 29 in appendix G.1 plot exercise evolution over lifetime in data and in the model. The dotted lines are one standard deviation above and below mean in data statistics. We see that our model fits the average exercise level during the life time for Very Good and Excellent health. For Fair and Good health, the model fit is not as good as that for Very Good and Excellent, but our prediction lies within first standard deviation from the mean.[30]

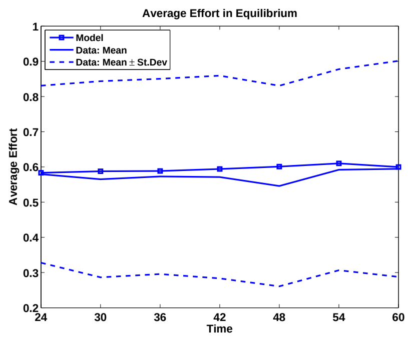

Figure 5: Average Effort in Model and Data

6 Results of the Policy Experiments: Macroeconomic Aggregates and Welfare

After having established that the model provides a good approximation to the data for the late 1990's and early 2000's in the absence of non-discrimination policies, we now use it to answer the main counterfactual question of this paper, namely, what are the effects of introducing these policies (one at a time and in conjunction) on aggregate health, consumption and effort, their distribution, and ultimately, on social welfare.

The primary benefit of the non-discrimination policy is to provide consumption insurance against bad health, resulting in lower wages and higher insurance premia in the competitive equilibrium. However, these policies weaken incentives to exert effort to lead a healthy life, and thus worsen the long run distribution of health, aggregate productivity and thus consumption. In the next two subsections, we present the key quantitative indicators measuring this tradeoff: first, the insurance benefits of policies, and second, the adverse incentive effects on aggregate production and health. Then, in subsection 6.3, we display the welfare consequences of our policy reforms. In order to keep the discussion of the results concise, in the main text we focus on the specific $(\gamma, educ)$-pair of high-school educated households with a high disutility cost for exercise (which is the largest group in the population with close to 50% mass), and present the results for all other pairs[31] as well as population weighted averages in figures 30 - 35 and tables 13 and 14 in Appendix G.2. The results there indicate that the figures we display and discuss in the main text remain substantially unchanged, qualitatively and quantitatively, when accounting for heterogeneity in education and preferences.

[30] For Fair and Good health, our model predicts higher exercise level between the ages of 30 and 54 than data. This is partly due to composition effect: in the second period of life, workers with low disutility for exercise become fair health leading to an increase in the average exercise level. One mechanical way of fixing this problem would be to let values of disutility differ across ages, i.e. γ_t, reflecting differences in taste for exercise at different stages of life.

[31] The key differences across groups are the initial distribution over health and the life cycle profile for labor income.

6.1 Insurance Benefits of Policies

Turning first to the consumption insurance benefits of both policies, we observe from figure 6 that the combination of both policies is indeed effective in providing perfect consumption insurance. As in the social planner problem, *within-group* consumption dispersion, as measured by the coefficient of variation, is zero for all periods over the life cycle if both a no-prior conditions and a no-wage discrimination are in place (the lines for the social planner solution and the equilibrium under both policies lie on top of one another and are identically equal to zero).[32] This is of course what the theoretical analysis in sections 3 and 4 predicted. Also notice from figure 6 that a wage non-discrimination law alone goes a long way towards providing effective consumption insurance, since the effect of differences in health levels on wage dispersion is an order of magnitude larger than on the corresponding dispersion in health insurance premia. Thus, although a no-prior conditions law in isolation provides some consumption insurance and reduces within-group consumption dispersion by about 30%, relative to the unregulated equilibrium, the remaining health-induced consumption risk remains large.

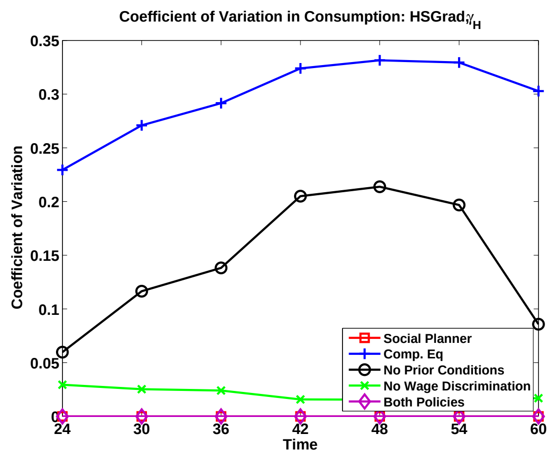

Figure 6: Consumption Dispersion: HSGrad, γ_H

Another measure of the insurance benefits provided by the non-discrimination policies is the level of cross-subsidization: workers do not necessarily pay their own competitive (actuarially fair) price of the health insurance premium or/and they are not compensated for their productivity. Under no-prior conditions policy, as established theoretically in Proposition 10, the healthy workers subsidize the premium of the unhealthy. Similarly, wages of the unhealthy workers are subsidized by the healthy, productive workers under the no-wage discrimination policy. Moreover, under both policies, there is cross-subsidization in both health insurance premium and wage.

The Figures 7 and 8 represent the degree of cross-subsidization for the excellent and fair health over the life-cycle. The plots for the premium represent the differences in the actuarially fair health insurance premium and the actual price paid under no-prior conditions policy and both policies. On the other hand, the wage plots are the differences in the productivity of the worker and the wage received under no-wage discrimination policy and both policies. Therefore, the negativity of the plots implies that the worker is paying higher premium, and is paid lower wage than they would in a competitive market without government intervention.

We see from Figure 7 that the workers with excellent health subsidize the other workers in both wage and premium. On the other hand, workers with fair health are subsidized in wage and premium under non-discrimination policies. An interesting property of the subsidies are that the level of subsidization is higher when only one of the non-disciminations are enacted, rather than both policies are. This is due to the fact that the government insures the workers with bad health through an inefficient level of medical expenditure.

[32]Due to the presence of heterogeneity in education levels and preferences the economy as a whole displays non-trivial consumption dispersion even in the presence of both policies (as it does in the solution of the restricted social planner problem). See figure ?? in the appendix G.2.

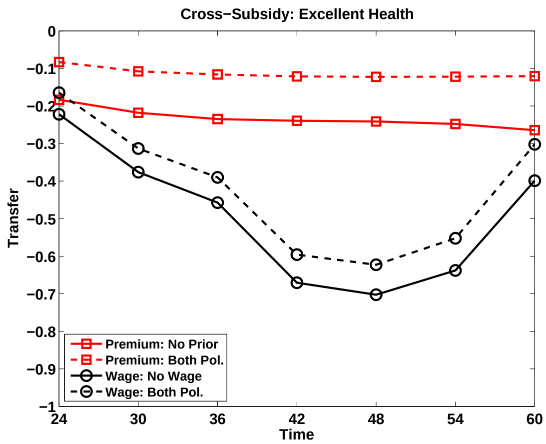

Figure 7: Cross Subsidy: Excellent Health

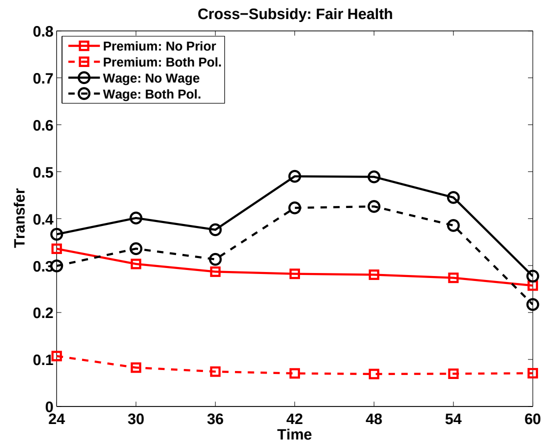

Figure 8: Cross Subsidy: Fair Health

Moreover, the degree of the insurance channel is bigger under no-prior conditions, as is clear from Figure 8: the premium subsidy received by the fair health under no-prior conditions policy is significantly larger than that under both policies.

Thus far, we discussed the insurance benefits of the non-discrimination policies. In the next subsection, we analyze the aggregate effects of the policies on production and health distribution.

6.2 Adverse Incentive Effects on Aggregate Production and Health

The associated incentive costs from each policy are inversely proportional to their consumption insurance benefits, as figure 9 shows. In this figure we plot the average[33] exerted effort over the life cycle, in the socially optimal and the equilibrium allocations. Effort is highest in the solution to the social planner problem, and positive under all policies since we allow for different level of terminal values conditional on health at retirement. Importantly, although effort is significantly smaller under the no-wage discrimination policy and both policies relative to the social planning solution and the competitive equilibrium, especially early in life, the incentives to lead a healthy life in order to reduce health insurance premia are strong enough (especially later in life) to induce effort closer to that observed in the competitive equilibrium. As a result, as figure 10 displays, while average health declines rapidly with age in the latter, the no-wage discrimination law alone induces a health distribution that is noticeably, but not fundamentally worse than that in the competitive equilibrium. The same then applies to aggregate production and aggregate consumption, as figures 12 and 13 show. Under the no-prior conditions law the reduction in incentives and thus the adverse effects on health, production and consumption are quantitatively smaller than under the no-wage discrimination legislation, since the remaining strong dependence of wages on health status in this policy environment induces an average effort over the life cycle that is only marginally smaller than that in the unregulated competitive equilibrium. See again figures 9, 12 and 13. The same plots also display the rather large losses in production and consumption emanating from the deterioration of average health levels under a combination of both policies. Finally, as figure 11 demonstrates, the decline of health levels over the life cycle manifest themselves in higher expenditures on health (insurance) later in life. Another feature of the medical expenditure across different policies are that it is the highest under the no prior conditions law. As shown in our theoretically analysis, governments use medical expenditure as a way of insuring workers of health status under the no prior conditions policy, which drive up the aggregate medical expenditure. Same holds with the no wage discrimination policy, but the extent is smaller than that under the no prior conditions policy. On the other hand, since under the both policies, efficient level of medical expenditure is chosen, and thus the aggregate medical expenditure does not diverge too much from the first-best and competitive equilibrium.

[33]As always in this section we focus exclusively on the high school completers with high γ. The plots average over households with different health status, however.

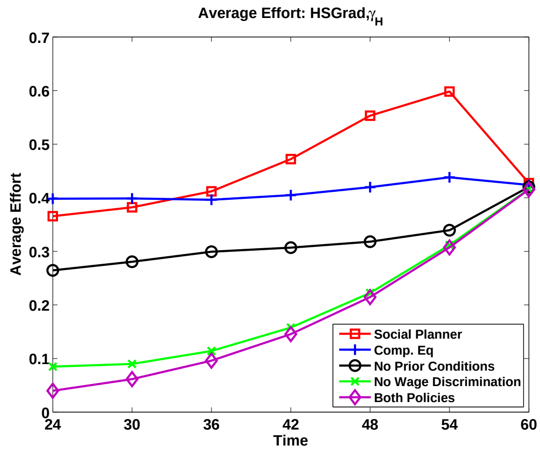

Figure 9: Effort: HSGrad, γ_H

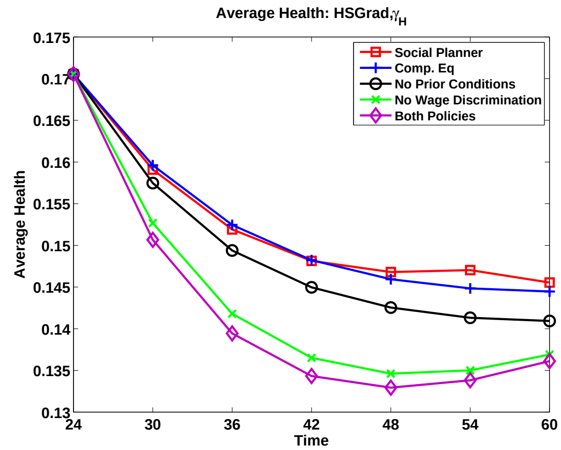

Figure 10: Average Health: HSGrad, γ_H

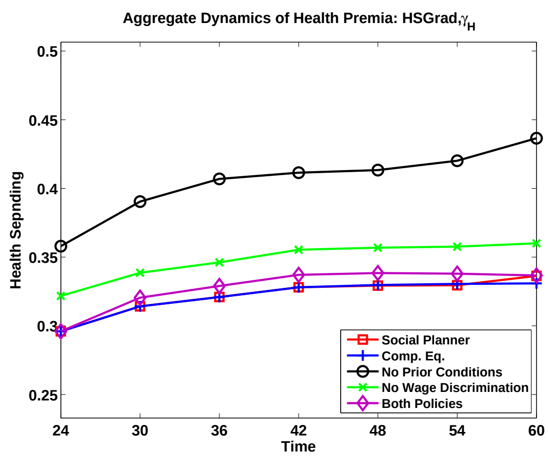

Figure 11: Health Spending: HSGrad, γ_H

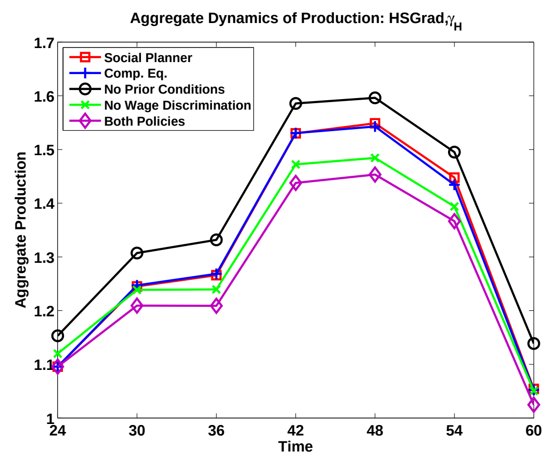

Figure 12: Production: HSGrad, γ_H

Overall, the effect on aggregate effort, health, production and consumption suggests a quantitatively important trade-off between consumption insurance and incentives. Within the spectrum of all policies, the unregulated equilibrium provides strong incentives at the expense of risky consumption, whereas a policy mix that includes both policies provides full insurance at the expense of a deterioration of the health distribution. The effects of the no-prior conditions law on consumption insurance and incentives are modest, relative to the unregulated equilibrium. In contrast, implementing a no wage discrimination law or both policies insures away most of the consumption risk, but significantly reduces (although does not eliminate) the incentives to keep up health by exercising, especially early in the life cycle. In the next subsection we will now document how these effects translate into welfare consequences from the hypothetical policy reforms.

6.3 Welfare Implications

In this section we quantify the welfare impact of the policy innovations studied in this paper. For a fixed initial distribution $\Phi_0(h)$ over health status, denote by $W(c, e)$ the expected lifetime utility of a cohort member (where expectations are taken prior to the initial draw h of health) from an arbitrary allocation of consumption and effort over the life cycle.[34] Our consumption-equivalent measure of the welfare consequences

[34] That is, using the notation from section 4, for the socially optimal allocation

$$W(c^{SP}, e^{SP}) = V(\Phi_0)$$

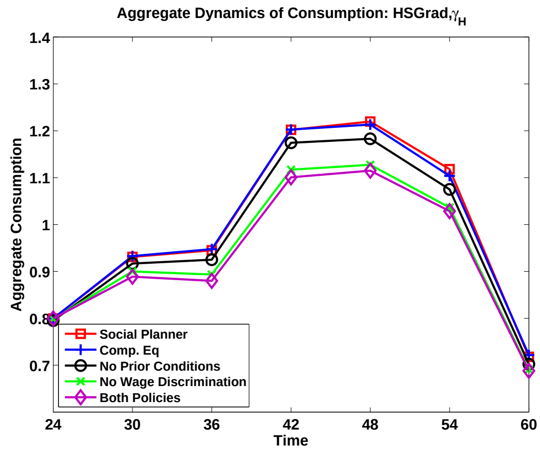

Figure 13: Consumption: HSGrad, γ_H

of a policy reform is given by

$$W(c^{CE}(1+CEV^i), e^{CE}) = W(c^i, e^i)$$

where i represents the four policy scenarios that we consider, that is $i \in \{SP, NP, NW, Both\}$. Thus CEV^i is the percentage reduction of consumption in the solution of the competitive equilibrium required to make households indifferent (ex ante) between the competitive equilibrium[35] and going through their work lives in an equilibrium with policy regime i.

In order to emphasize the importance of the dynamic analysis in assessing welfare we also report the welfare consequences of the same policy reforms in the static version of the model in section 3. We compute the static consumption-equivalent loss (relative to the social optimum) as

$$U(c^{CE}(1+SCEV^i)) = U(c^i)$$

where $U(c)$ is the expected utility from period 0 *consumption*[36], under the cross-sectional distribution Φ_0, and thus determined by the static version of the model.[37] Therefore $SCEV^i$ provides a clean measure of the static gains from better consumption insurance induced by the policies against which the dynamic adverse incentive effects have to be traded off.

From Table 1 we observe that the introduction of the non-discrimination policies improves social welfare, relative to the competitive equilibrium. Focusing on the results based on lifetime utility (the last column of table **??**), the welfare gains of implementing a no-prior conditions policy amount to about 7.05% of lifetime consumption, relative to the laissez-faire competitive equilibrium. A no-wage discrimination law would result in a more modest, but still sizeable gain of 7.03%.

Finally, and most crucially for the main point of this paper, if both policies are introduced jointly, the welfare consequences are less favorable than under each policy in isolation.

and for equilibrium allocations, under policy i,

$$W(c^i, e^i) = \int v_0^i(h) d\Phi_0.$$

[35] Recall that the social planner problem is solved for the specific $(\gamma, educ)$ group and thus even the social planner therefore does not provide ex-ante insurance against unfavorable $(\gamma, educ)$-draws.

[36] In the static version of the model effort is identically zero in the social planner problem and in the equilibrium under all policy specifications, and therefore disutility from effort is irrelevant.

[37] Thus, using the notation from section 3

$$U(c^{CE}) = U^{CE}(\Phi_0) \text{ for } i \in \{SP, NP, NW, Both\}$$

and

$$U(c^{CE}) = \int U^{CE}(h) d\Phi_0.$$

	Static	Dynamic
Social Planner	5.1777	13.5151
Competitive Equilibrium	0.0000	0.0000
No Prior Conditions Law	4.4918	7.0531
No Wage Discrimination Law	4.8780	7.0303
Both Policies	5.1777	6.8849

Table 1: Aggregate Welfare Comparisons

	Fair	Good	Very Good	Excellent
Social Planner	58.5369	12.8954	16.1256	3.5705
Competitive Equilibrium	0.0000	0.0000	0.0000	0.0000
No Prior Conditions Law	43.4772	9.3671	5.9972	-3.4125
No Wage Discrimination Law	53.6017	12.7906	4.4938	-7.6006
Both Policies	56.0514	11.5087	4.8639	-9.3338

Table 2: Welfare Comparison in the Dynamic Economy Conditional on Health

We also document in Table 2, the welfare measures conditional on health in the dynamic economy. As expected, the fair health benefits the most with these policies, while the workers with excellent health prefer the competitive equilibrium to the introduction of any of the policies.

7 Conclusion

In this paper, we studied the effect of labor and health insurance market regulations on evolution of health and production, as well as welfare. We showed that both a no-wage discrimination law (an intervention in the labor market), in combination with a no-prior conditions law (an intervention in the health insurance market) provides effective consumption insurance against health shocks, holding the aggregate health distribution in society constant. However, the dynamic incentive costs and their impact on health and medical expenditures of both policies, if implemented jointly, is large. Introduction of one of the policies has higher welfare gain than under both policies. Moreover, although non-wage discrimination law provides more insurance, no-prior conditions law is better as it preserves most of the incentives to stay healthy through dependence of wages on health. We therefore conclude that a complete policy analysis of health insurance reforms on one side and labor market (non-discrimination policy) reforms cannot be conducted separately, since their interaction might prove less favorable despite welfare gains from each policy separately.

Our analysis of health insurance and incentives over the working life has ignored several potentially important avenues through which health and consumption risk affects welfare. First, the benefits of health in our model are confined to higher labor productivity, and thus we model the investment motives into health explicitly. It has abstracted from an explicit modeling of the benefits better health has on survival risk, although the positive effect of health h on the continuation utility after retirement captures this effect in our model, albeit in a fairly reduced from. Similarly, better health might have a direct effect on flow utility during working life. As we argue in appendix H at least in one extension of the model introducing a direct consumption benefit from better health the effort exerted of all households is shifted upwards by a constant (so that effort is not zero any longer even in the both-policy scenario), without affecting the statement about the effects and desirability of the policies considered here. Finally, in our analysis labor income risk directly translates into consumption risk, in the absence of household private saving. We conjecture that the introduction of self-insurance via precautionary saving against this income risk weakens the argument in favor of the policies studied in this paper. Future work has to uncover whether such an extension of the model also affects, quantitatively or even qualitatively, our conclusions about the relative desirability of these policies.

References

[1] Acemoglu, D. and Joshua D. Angrist (2001), "Consequences of Employment Protection? The Case of the Americans with Disabilities Act." *Journal of Political Economy,* Volume 109, No. 5, 915-957.

[2] Ales, L., R. Hosseini, and L. Jones (2012), "Is There "Too Much" Inequality in Health Spending Across Income Groups?" *Working Paper.*

[3] Battacharya, J., K. Bundorf, N. Pace and N. Sood (2009), "Does Health Insurance Make You Fat?," NBER *Working Paper* 15163.

[4] Brügemann, B. and I. Manovskii (2010), "Fragility: A Quantitative Analysis of the US Health Insurance System," *Working paper*, University of Pennsylvania

[5] Cawley, J. (2004), "The Impact of Obesity on Wages," *Journal of Human Resources*, 39(2), 451-474

[6] Currie, J. and B. Madrian (1999), "Health, Health Insurance and the Labor Market." In *Handbook of Labor Economics,* Volume 3C, edited by David Card and Orley Ashenfelter, Amsterdam: North Holland, 3309-3407.

[7] DeLeire, T. (2001). "Changes in Wage Discrimination against People with Disabilities: 1984-93." *Journal of Human Resources,* Volume 35, No. 1, 144-158.

[8] M. Dey and C. Flinn (2005), "An Equilibrium Model of Health Insurance Provision and Wage Determination," *Econometrica* 73, 571-627.

[9] Edwards, R. (2008), "Health Risk and Portfolio Choice," *Journal of Business and Economic Statistics*, 26, 472-485.

[10] Ehrlich, I. and G. Becker (1972), "Market Insurance, Self-Insurance, and Self-Protection," *Journal of Political Economy*, 80, 623-648.

[11] Ehrlich, I. and H. Chuma (1990), "A Model of the Demand for Longevity and the Value of Life Extension," *Journal of Political Economy*, 98, 761-782.

[12] French, E. and J. Jones (2004), "On the Distribution and Dynamics of Health Care Costs," *Journal of Applied Economics*, 19, 705-721.

[13] Grossman, M. (1972), "On the Concept of Health Capital and the Demand for Health," *Journal of Political Economy*, 80(2), 223-255.

[14] Hall, R. and C. Jones (2007), "The Value of Life and the Rise in Health Spending," *Quarterly Journal of Economics*, 122(1), 39-72.

[15] Hugonnier, J., F. Pelgrin and P. St-Amour (2012), "Health and (Other) Asset Holdings," *Working Paper.*

[16] Kass, N., A. Medley, M. Natowicz, S. Chandros Hull, R. Faden, L. Plantinga and L. Gostin (2007), "Access to Health Insurance: Experiences and Attitudes of Those with Genetic versus Non-genetic Medical Conditions," *American Journal of Medical Genetic,* 143A, 707–717.

[17] Kopecky, K. and T. Koreshkova (2012), "The Joint Impact of Social Security and Medicaid on Incentives and Welfare," *Working Paper.*

[18] Laun, T. (2012), "Optimal Social Insurance with Endogenous Health," *Working Paper.*

[19] Madrian, B. (1994), "Employment-Based Health Insurance and Job Mobility: Is There Evidence of Job-Lock?," *Quarterly Journal of Economics*, 109, 27-54.

[20] Pijoan-Mas, J. and V. Rios-Rull (2012), "Heterogeneity in Expected Longevities," *Working Paper.*

[21] Prados, M. (2012). "Health and Earnings Inequality," *Working Paper*.

[22] Sommers, A. (2007), "Access to Health Insurance, Barriers to Care, and Service Use among Adults with Disabilities," *Inquiry*, 43, 393-405.

[23] Yogo, M. (2009), "Portfolio Choice in Retirement: Health Risk and the Demand for Annuities, Housing, and Risky Assets," NBER *Working Paper* 15307.

A Proofs of Propositions

Proposition 5

Proof. Since exercise does not carry any benefits in the static model, trivially $e^{SP} = 0$. Attaching Lagrange multiplier $\mu \geq 0$ to the resource constraint, the first order condition with respect to consumption $c(\varepsilon)$ is

$$u'(c(\varepsilon, h)) = \lambda$$

and thus $c^{SP}(\varepsilon, h) = c^{SP}$ for all $\varepsilon \in E$ and $h \in H$. Thus, not surprisingly, the social planner provides full consumption insurance to households. The optimal health expenditure allocation maximizes this consumption

$$c^{SP} = \max_{x(\varepsilon,h)} \sum_h \left\{ g(h) \left[F(h, -x(0,h)) - x(0,h) \right] + (1 - g(h)) \int f(\varepsilon) \left[F(h, \varepsilon - x(\varepsilon, h)) - x(\varepsilon, h) \right] d\varepsilon \right\} \Phi(h)$$

Denoting by $\mu(\varepsilon, h) \geq 0$ the Lagrange multiplier on the constraint $x(\varepsilon, h) \geq 0$, the first order condition with respect to $x(\varepsilon, h)$ reads as

$$-F_2(h, \varepsilon - x(\varepsilon, h)) + \mu(\varepsilon, h) = 1$$

Fix $h \in H$. By assumption 4 $F_{22}(h, y) < 0$ and thus either $x(\varepsilon, h) = 0$ or $x(\varepsilon, h) > 0$ satisfying

$$-F_2(h, \varepsilon - x(\varepsilon, h)) = 1$$

for all ε. Thus off corners $\varepsilon - x(\varepsilon, h) = \bar{\varepsilon}^{SP}(h)$ where the threshold satisfies

$$-F_2(h, \bar{\varepsilon}^{SP}(h)) = 1. \tag{40}$$

Consequently

$$x^{SP}(\varepsilon, h) = \max \left[0, \varepsilon - \bar{\varepsilon}^{SP}(h) \right].$$

The fact that $\bar{\varepsilon}^{SP}(h)$ is increasing in h, strictly so if $F_{12}(h, y) < 0$, follows directly from assumption 4 and (40). ■

Proposition 6

Proof. Attaching Lagrange multiplier $\mu(h)$ to equation (11) and $\lambda(h)$ to equations (12) the first order conditions read as

$$\begin{aligned}
u'(w(h) - P(h)) &= \lambda(h) = -\mu(h) && (41)\\
\lambda(h) F_2(h, -x(0,h)) &\leq \mu(h) && (42)\\
&= \text{ if } x(0,h) > 0 &&\\
\lambda(h) F_2(h, \varepsilon - x(\varepsilon, h)) &\leq \mu(h) && (43)\\
&= \text{ if } x(\varepsilon, h) > 0 &&
\end{aligned}$$

Thus off corners we have

$$F_2(h, \hat{\varepsilon} - x(\hat{\varepsilon}, h)) = F_2(h, \varepsilon - x(\varepsilon, h)) = K \tag{44}$$

for some constant K. Thus off corners $\varepsilon - x(\varepsilon, h)$ is constant in ε and thus medical expenditures satisfy the cutoff rule

$$x^{CE}(\varepsilon, h) = \max \left[0, \varepsilon - \bar{\varepsilon}^{CE}(h) \right]. \tag{45}$$

Plugging (45) into (43) and evaluating it at $\varepsilon = \bar{\varepsilon}^{CE}(h)$ yields

$$\lambda(h) F_2(h, \bar{\varepsilon}^{CE}(h)) = \mu(h). \tag{46}$$

Using this result in the second part of (41) delivers the characterization of the equilibrium cutoff levels

$$F_2(h, \bar{\varepsilon}^{CE}(h)) = -1 \text{ for all } h \in H$$

which are unique, given the assumptions imposed on F. Wages, consumption and health insurance premia then trivially follow from (11) and (12). ■

Proposition 9

Proof. Let Lagrange multipliers to equations (19) and (20) be μ and $\lambda(h)$, respectively. Then, the first order conditions are:

$$\begin{aligned}
\sum_h u'(w(h)-P)\Phi(h) &= \mu \\
u'(w(h)-P)\Phi(h) &= \lambda(h) \\
(1-g(h))f(\varepsilon)[-F_2(h,\varepsilon-x(\varepsilon,h))]\lambda(h) &\leq \mu(1-g(h))f(\varepsilon)\Phi(h) \\
&= \text{ if } x(\varepsilon,h)>0 \\
g(h)[-F_2(h,-x(0,h))]\lambda(h) &\leq \mu g(h)\Phi(h) \\
&= \text{ if } x(0,h)>0
\end{aligned}$$

Thus, off-corners we have

$$F_2(h,\varepsilon-x(\varepsilon,h)) = F_2(h,\hat{\varepsilon}-x(\hat{\varepsilon},h)) = K$$

for some constant K and the cutoff rule is determined by

$$u'(w(h)-P)[-F_2(h,\bar{\varepsilon}^{NP}(h))] = \sum_h u'(w(h)-P)\Phi(h). \tag{47}$$

Moreover, let us take the derivative of (47) with respect to h.

$$\begin{aligned}
& u''(w(h)-P)\frac{\partial w(h)}{\partial h}F_2 + u'(w(h)-P)\left\{F_{12}+F_{22}\frac{\partial\bar{\varepsilon}^{NP}(h)}{\partial h}\right\} = 0 \\
& u''(w(h)-P)\frac{\partial\bar{\varepsilon}^{NP}(h)}{\partial h}\frac{\partial w(h)}{\partial\bar{\varepsilon}^{NP}(h)}F_2 + u'(w(h)-P)\left\{F_{12}+F_{22}\frac{\partial\bar{\varepsilon}^{NP}(h)}{\partial h}\right\} = 0 \\
\Rightarrow\quad & \frac{\partial\bar{\varepsilon}^{NP}(h)}{\partial h}\left\{u''(w(h)-P)F_2\frac{\partial w(h)}{\partial\bar{\varepsilon}^{NP}(h)} + u'(w(h)-P)F_{22}\right\} = -u'(w(h)-P)F_{12}
\end{aligned}$$

Note that as $\bar{\varepsilon}$ increases $w(h)$ decreases, since $F(h,\varepsilon-x(\varepsilon,h))$ is decreasing for $\varepsilon<\bar{\varepsilon}$, and constant for $\varepsilon\geq\bar{\varepsilon}$. Thus, we have

$$\frac{\partial\bar{\varepsilon}^{NP}(h)}{\partial h} > 0.$$

■

Proposition 10

Proof. From (21), we immediately obtain

$$-F_2(h,\bar{\varepsilon}^{NP}(h)) = \frac{\sum u'(w(h)-P)\Phi(h)}{u'(w(h)-P)} \quad \begin{matrix} <1 \\ =1 \\ >1 \end{matrix} \quad \Rightarrow \quad \begin{matrix} \bar{\varepsilon}^{NP}(h)<\bar{\varepsilon}^{SP}(h) \\ \bar{\varepsilon}^{NP}(h)=\bar{\varepsilon}^{SP}(h) \\ \bar{\varepsilon}^{NP}(h)>\bar{\varepsilon}^{SP}(h) \end{matrix}$$

as $-F_2(h,\bar{\varepsilon}^{SP}(h)) = 1$.

Let us take $h_L<\tilde{h}<h_H$, and suppose

$$-F_2(h_L,\bar{\varepsilon}^{NP}(h_L)) > 1 > -F_2(h_H,\bar{\varepsilon}^{NP}(h_H)), \tag{48}$$

i.e.

$$\begin{aligned}
\bar{\varepsilon}^{NP}(h_H)<\bar{\varepsilon}^{SP}(h_H) &\quad\Rightarrow\quad w^{NP}(h_H)>w^{SP}(h_H) \\
\bar{\varepsilon}^{NP}(h_L)>\bar{\varepsilon}^{SP}(h_L) &\quad\Rightarrow\quad w^{NP}(h_L)<w^{SP}(h_L),
\end{aligned}$$

where $w^{SP}(h) = g(h)F(h,0)+(1-g(h))\int f(\varepsilon)F(h,\varepsilon-x(\varepsilon,h))d\varepsilon$. Then, we have

$$u'^{NP}(h_H)-P) < u'^{SP}(h_H)-P) < u'^{SP}(h_L)-P) < u'^{NP}(h_L)-P),$$

where the second inequality follows from (51). This result, in combination with (48) implies

$$u'^{NP}(h_L) - P)[-F_2(h_L, \bar{\varepsilon}^{NP}(h_L))] > u'^{NP}(h_H) - P)[-F_2(h_H, \bar{\varepsilon}^{NP}(h_H))],$$

a contradiction to (21). ■

Proposition 13

Proof. Is by backward induction. Trivially $e_T(h) = 0$. In period T, since both policies are in place, the wage and health insurance premium of every household is independent of h. Thus

$$v_T(h) = u(w_T - P_T) = v_T$$

and therefore the terminal value function is independent of h. Now suppose for a given time period t the value function v_{t+1} is independent of h. Then from the first order condition with respect to $e_t(h)$ we have

$$q'(e_t(h)) = \beta v_{t+1} \sum_{h'} \frac{\partial Q(h'; h, e)}{\partial e}$$

But since for every e and every h, $Q(h'; h, e)$ is a probability measure over h' we have $\sum_{h'} \frac{\partial Q(h';h,e)}{\partial e} = 0$ and thus $e_t(h, \gamma) = 0$ for all h, on account of our assumptions on $q'(.)$. But then

$$v_t(h) = u(w_t - P_t) + \left\{ -0 + \beta v_{t+1} \sum_{h'} Q(h'; h, 0) \right\} = u(w_t - P_t) + \beta v_{t+1} = v_t$$

since $\sum_{h'} Q(h'; h, 0) = 1$ for all h. Thus v_t is independent of h. The evolution of the health distributions follows from (22), and given these health distributions wages and health insurance premia are given by (37) and (39). ■

B Further Analysis of the No-Wage Discrimination Case

B.1 Health Insurance Distortions with No-Wage Discrimination

The firm's break-even condition is

$$\sum_h \left\{ g(h)F(h, 0) + (1 - g(h)) \int_0^{\bar{\varepsilon}} f(\varepsilon)[F(h, \varepsilon - x^{NP}(\varepsilon, h))]d\varepsilon - w(h) \right\} \Phi(h) = 0,$$

and hence on average the production level of a worker will equal his gross wage. Taking $\varepsilon_w > 0$ and $\delta > 0$ as given, workers for whom the wage limits, $\max_{h,h'} |w(h) - w(h')| \leq \varepsilon_w$, bind will be paid either more or less than their production level depending on whether the wage discrimination bound binds from above or below. The firm will optimally choose to hire less than the population share of any health type h whose wage is above their production level, and hence some of these workers will be unemployed. Since we have assume that there is no cost to working and workers pay for their own insurance, competition over health insurance will lead these workers to increase their health insurance, $x(e, h)$, so that their productivity is within ε_w of their wage $w(h)$. In the limit as $\varepsilon_w \to 0$, this implies that

$$w(h) = g(h)F(h, 0) + (1 - g(h)) \int_0^{\bar{\varepsilon}} f(\varepsilon)[F(h, \varepsilon - x^{NP}(\varepsilon, h))]d\varepsilon, \tag{49}$$

holds and they are fully employed, or $w(h) - P(h) = 0$. On the flip side, there will be excess demand for workers whose expected production is more than $w(h)$, they will therefore find it optimal to either lower their insurance, and in the limit as $\varepsilon \to 0$ either (49) holds they or set $x(e, h) = 0$ if they end up at corner with respect to health insurance. Assuming that neither corner binds, this implies that the no-wage discrimination policy will be undone by adjustments in the health insurance market. This motivated our assumption that the government will choose to regulate the health insurance market to prevent this outcome as part of the no-wage discrimination policy.

For health types for which the bounds do not bind, market clearing implies that

$$w(h) = g(h)F(h,0) + (1-g(h))\int_0^{\bar{\varepsilon}} f(\varepsilon)[F(h,\varepsilon - x^{NP}(\varepsilon,h))]d\varepsilon$$

while actuarial fairness implies that

$$P(h) = (1-g(h))\int_0^{\bar{\varepsilon}} f(\varepsilon)x^{NP}(\varepsilon,h))d\varepsilon.$$

Hence, an efficient health insurance contract for this type will maximize $w(h) - P(h) = w^{CE}(h) - P^{CE}(h)$. Since $w^{CE}(h) - P^{CE}(h)$ is increasing in h, it follows that the wage bound binds for the lowest and highest health types.

B.2 No-Wage Discrimination with Realized Penalties in Equilibrium

Here we assume that the firm must pay a cost for having wage dispersion conditional on health type or for having the health composition of its work force differ from the population average. The wage variation penalty is assumed to take the form

$$C\sum_h [w(h) - w(0)]^2\, n(h),$$

since health type 0 will have the lowest wage in equilibrium, and where C is the penalty parameter and $n(h)$ is measure of type h workers the firm hires. Note that with this penalty function the penalty will apply to all workers with health $h > 0$.[38] The penalty from having one's composition deviate from the population average is given by

$$\sum_h D\left[\frac{n(h)}{\sum n(h)} - \frac{\Phi(h)}{\sum \Phi(h)}\right]^2.$$

Since these penalties are small for small deviations, it will turn out that penalty costs will be realized in equilibrium. Since both of these penalties are real we need to subtract them from production. We will assume that there too the government will regulate the insurance market to prevent workers low health status workers raising their productivity by over-insuring themselves against health risks and high health status workers lowering their productivity by under-insuring themselves.

We begin analyzing this case by assuming that the penalties for wage discrimination C and hiring discrimination D are both finite and then we examine the equilibrium in the limit as they become large. The firm takes as given the health policy of the worker and the equilibrium wage $w(h)$ and chooses the measure of each health type to hire $n(h)$ so as to maximize

$$\max_{n(h)}\sum_h \left[g(h)\left[F(h,-x(0,h)) - x(0,h)\right] + (1-g(h))\int_0^{\bar{\varepsilon}} f(\varepsilon)\left[F(h,\varepsilon - x(\varepsilon,h)) - x(\varepsilon,h)\right]d\varepsilon - w(h)\right]n(h)$$
$$-C\sum_h [w(h) - w^*)]^2\, n(h) - \sum_h \left[\frac{n(h)}{\sum n(h)} - \frac{\Phi(h)}{\sum \Phi(h)}\right]^2,$$

[38] If we have assumed that the form of the penalty was

$$C\int_h [w(h) - w^*]^2\, \psi(h)dh,$$

where w^* is the average wage, this would mean that low productivity workers are more costly and less productive, which will discourage hiring them. Hence, with this form the low productivity workers will only be employed because of the compositional penalty, which means that the hiring penalty must bind at the margin. Hence the less than average productivity workers will be in positive net supply in equilibrium, which will complicate the analysis because some of these workers will be employed and some will not be.

where w^* is taken here to mean the lowest wage. Trivially, the firm will want to hire more than the population share of any type h for whom

$$
\begin{aligned}
N(h) \quad \equiv \quad & \left[g(h)\left[F(h,-x(0,h))-x(0,h)\right]+(1-g(h))\int_0^{\bar{\varepsilon}} f(\varepsilon)\left[F(h,\varepsilon-x(\varepsilon,h))-x(\varepsilon,h)\right]d\varepsilon - w(h)\right] \\
& -C\left[w(h)-w^*)\right]^2
\end{aligned}
$$

is positive and less that the population share if $N(h)$ is negative. Since all firms share this condition, they will all choose the same relative shares of each type of worker. Since workers are willing to work so long as $w(h)-P(h)>0$, it follows that $w(h)$ cannot be more than w^* if $N(h)$ is not positive. To see this note that there would be excess supply of type h workers and hence the labor market would not clear. Moreover, a firm would rather hire a worker of type h at $w^*-\varepsilon$ than for w^* for ε small. Hence, if $w(h)=w^*$, then $N(h)=0$ so long as $w^*-P(h)>0$. Hence, for the labor market to clear for each health type, either $N(h)=0$ for type h or $N(h)>0$ but $w(h)-P(h)=0$. This implies the following proposition.

Proposition 14 *If C and D are positive but finite, and $w(h)-P(h)>0$ for all h, then in equilibrium all households are hired, all firms are representative, and the wage $w(h)$ is equal to a worker's productivity less the cost of paying him.*

Since the government can set $x(\varepsilon,h)=0$ which implies that $P(h)=0$, we assume that $w(h)-P(h)>0$ for all health types.

B.3 Realized Penalties with Both Policies

Since all that workers care about is their net wage $\tilde{w}(h)$, which is also equal to their consumption, it follows that workers are indifferent over contracts that offer combinations of a gross wage $w(h)$ and medical costs $P(h)$ for which $\tilde{w}(h)=w(h)-P(h)$ is constant. Hence, it is natural to assume that the firm takes the equilibrium *net wage* function $\tilde{w}(h)$ as given and chooses the measure of each health type to hire, $n(h)$, and its health plan, $x(\varepsilon,h)$, to solve the following problem

$$
\begin{aligned}
& \max_{n(h),x(\varepsilon,h)} \sum_h \left[g(h)\left[F(h,-x(0,h))-x(0,h)\right]+(1-g(h))\int_0^{\bar{\varepsilon}} f(\varepsilon)\left[F(h,\varepsilon-x(\varepsilon,h))-x(\varepsilon,h)\right]d\varepsilon - \tilde{w}(h)\right] n(h) \\
& -C\sum_h \left[\tilde{w}(h)-\tilde{w}(0)\right]^2 n(h) - \sum_h D\left[\frac{n(h)}{\sum n(h)}-\frac{\Phi(h)}{\sum\Phi(h)}\right]^2 .
\end{aligned}
$$

Proposition 15 *If C and D are positive but finite, then in equilibrium all households are hired, all firms are representative, the net wage $\tilde{w}(h)$ is equal to a worker's productivity less the cost of paying him more than $\tilde{w}(0)$, and $\tilde{w}(0)=w^{CE}(0)-P(0)$. The firm optimally sets $x(\varepsilon,h)=x^{CE}(\varepsilon,h)$. As $C\to\infty$, $\tilde{w}(h)\to\tilde{w}(0)$.*

Proof. The optimality condition for $x(h,\varepsilon)$ if $\varepsilon=0$ is

$$F(h,-x(0,h))-1\le 0$$

and if $\varepsilon>0$ is

$$F(h,\varepsilon-x(\varepsilon,h))-1\le 0 \text{ w. equality if } x(\varepsilon,h)>0.$$

These are the same conditions as in the competitive equilibrium.

Next, we show that $\tilde{w}(h)$ has to be increasing in h and hence $\tilde{w}(0)$ is the lowest paid type. The wage penalty is w.r.t. to the lowest paid worker type, which we denote by w^*. Given that optimum insurance is the same as in the competitive equilibrium, it follows that the net earnings per worker is $w^{CE}(h)-P^{CE}(h)-\tilde{w}(h)$, and from before $w^{CE}(h)-P^{CE}(h)$ is increasing in h. Hence, for the firm to break even

$$
\begin{aligned}
& \sum_h \left[w^{CE}(h)-P^{CE}(h)-\tilde{w}(h)\right]n(h) \\
- \quad & C\sum_h \left[\tilde{w}(h)-w^*\right]^2 n(h) - \sum_h D\left[\frac{n(h)}{\sum n(h)}-\frac{\Phi(h)}{\sum\Phi(h)}\right]^2 = 0,
\end{aligned}
$$

and the optimality condition for $n(h)$ is

$$\begin{aligned} & \left[w^{CE}(h) - P^{CE}(h) - \tilde{w}(h)\right] - C\left[\tilde{w}(h) - w^*\right]^2 \\ - \quad & D\left[\frac{n(h)}{\sum n(h)} - \frac{\Phi(h)}{\sum \Phi(h)}\right]\left[1 - \frac{n(h)}{\sum n(h)}\right]\frac{1}{\sum n(h)} = 0. \end{aligned}$$

This condition implies that a firm will hire more that the population share of any type h for whom

$$\tilde{N}(h) \equiv w^{CE}(h) - P^{CE}(h) - \tilde{w}(h) - C\left[\tilde{w}(h) - w^*\right]^2 > 0,$$

and less than the population share if the reverse is true. However any health type h that are not fully employed in equilibrium would have excess members who would be happy to be hired any positive wage. Hence, either type h is paid the lowest equilibrium wage or they are fully employed. Hence, any type h for whom $w(h) > w^*$ are fully employed. Any type receiving the lowest wage must be fully employed since the firm would be willing to hire more of these workers if we lowered the bottom wage by ε. Since all workers are fully employed, it follows that all firms will choose to be representative to avoid the hiring penalty, and that $\tilde{w}(0) = w^{CE}(0) = w^*$ and $\tilde{w}(h)$ is increasing h. Finally, since the marginal penalty for a deviation in a type's net wage from the economy-wide lowest type's wage is given by

$$-C\left[\tilde{w}(h) - \tilde{w}(0)\right]^2,$$

and since this cost goes to infinity as $C \to \infty$ for any positive wage gap, it follows that as C becomes large $\tilde{w}(h) \to \tilde{w}(0)$, and all of the workers are paid as if they were the lowest health status type and all of their productivity gap is absorbed by the cost of discriminating on wages. *Q.E.D.* ■

The fact that the productivity advantage of higher health status individuals is completely absorbed by the discrimination costs means that the society as a whole gets no gain from their productivity advantage. So the health expenditures that raise their productivity above the lowest type are inefficient. In addition, expenditure on the lowest health type relaxes the wage discrimination penalty on other types. So this equilibrium outcome is not socially efficient.

C Wages in the Competitive Equilibrium

To understand the implications of proposition 6 for the behavior of equilibrium wages, note that our results imply that the equilibrium competitive wage is given by

$$\begin{aligned} w^{CE}(h) &= g(h)F(h,0) + (1-g(h))\int_0^{\bar{\varepsilon}^{CE}(h)} f(\varepsilon)F(h,\varepsilon - x(\varepsilon,h))d\varepsilon \\ &\quad +(1-g(h))\int_{\bar{\varepsilon}^{CE}(h)}^{\bar{\varepsilon}} f(\varepsilon)F(h,\bar{\varepsilon}^{CE}(h))d\varepsilon. \end{aligned}$$

Hence

$$\begin{aligned} \frac{dw^{CE}(h)}{dh} &= g'(h)\left[\begin{array}{c} F(h,0) - \int_0^{\bar{\varepsilon}^{CE}(h)} f(\varepsilon)F(h,\varepsilon - x(\varepsilon,h))d\varepsilon \\ - \int_{\bar{\varepsilon}^{CE}(h)}^{\bar{\varepsilon}} f(\varepsilon)F(h,\bar{\varepsilon}^{CE}(h))d\varepsilon \end{array}\right] \\ &\quad +g(h)F_1(h,0) + (1-g(h))\int_0^{\bar{\varepsilon}^{CE}(h)} f(\varepsilon)F_1(h,\varepsilon - x(\varepsilon,h))d\varepsilon \\ &\quad +(1-g(h))\int_{\bar{\varepsilon}^{CE}(h)}^{\bar{\varepsilon}} f(\varepsilon)F_1(h,\bar{\varepsilon}^{CE}(h))d\varepsilon \\ &\quad +(1-g(h))\int_{\bar{\varepsilon}^{CE}(h)}^{\bar{\varepsilon}} f(\varepsilon)F_2(h,\bar{\varepsilon}^{CE}(h))\frac{d\bar{\varepsilon}^{CE}(h)}{dh}d\varepsilon, \end{aligned}$$

since net effect of the change in the integrand bounds generated by $\frac{d\bar{\varepsilon}^{CE}(h)}{dh}$ is zero. Next note that our optimality condition for $\bar{\varepsilon}^{CE}(h)$, (17), implies that

$$F_{12}(h,\bar{\varepsilon}^{CE}(h))dh + F_{22}(h,\bar{\varepsilon}^{CE}(h))d\bar{\varepsilon}^{CE}(h) = 0,$$

and hence

$$\frac{d\bar{\varepsilon}^{CE}(h)}{dh} = \frac{-F_{12}(h,\bar{\varepsilon}^{CE}(h))}{F_{22}(h,\bar{\varepsilon}^{CE}(h))}.$$

This result, along with (17), implies that

$$\begin{aligned} \frac{dw^{CE}(h)}{dh} &= g'(h)\left[\begin{array}{c} F(h,0)-\int_0^{\bar{\varepsilon}^{CE}(h)} f(\varepsilon)F(h,\varepsilon - x(\varepsilon,h))d\varepsilon \\ -\int_{\bar{\varepsilon}^{CE}(h)}^{\bar{\varepsilon}} f(\varepsilon)F(h,\bar{\varepsilon}^{CE}(h))d\varepsilon \end{array}\right] \\ &\quad +g(h)F_1(h,0)+(1-g(h))\int_0^{\bar{\varepsilon}^{CE}(h)} f(\varepsilon)F_1(h,\varepsilon - x(\varepsilon,h))d\varepsilon \\ &\quad +(1-g(h))\int_{\bar{\varepsilon}^{CE}(h)}^{\bar{\varepsilon}} f(\varepsilon)F_1(h,\bar{\varepsilon}^{CE}(h))d\varepsilon \\ &\quad -(1-g(h))\int_{\bar{\varepsilon}^{CE}(h)}^{\bar{\varepsilon}} f(\varepsilon)F_2(h,\bar{\varepsilon}^{CE}(h))\frac{F_{12}(h,\bar{\varepsilon}^{CE}(h))}{F_{22}(h,\bar{\varepsilon}^{CE}(h))}d\varepsilon. \end{aligned} \tag{50}$$

All of the terms in (50) are trivially positive except the last, which is negative since $F_{22} < 0$. However, so long as the spillover ratio F_{12}/F_{22} evaluated at $(h,\bar{\varepsilon}^{CE}(h))$ is not too negative then, then wages will vary positive with health status. Note that this is trivially implied if the direct effect of the change in health status offsets the spillover, or

$$F_1(h,\bar{\varepsilon}^{CE}(h)) - F_2(h,\bar{\varepsilon}^{CE}(h))\frac{F_{12}(h,\bar{\varepsilon}^{CE}(h))}{F_{22}(h,\bar{\varepsilon}^{CE}(h))} > 0. \tag{51}$$

Note that this is a condition purely on the fundamentals of the economy since $\bar{\varepsilon}^{CE}(h)$ is given by an (implicit) equation that depends only on exogenous model elements. We summarize our results in the following proposition:

Proposition 16 *The competitive wage is increasing in h if (50) is positive.*

D Computation of the Social Planner Problem

The idea to solve the problems in (24) is to iterate on sequences $\{c_t, e_t(h), \Phi_t(h)\}$, using the first order condition (25) for the optimal effort choice and the envelope condition (26). To initialize the iterations, note that

$$\begin{aligned} V_T(\Phi_T) &= u(c_T) \\ \frac{\partial V_T(\Phi_T)}{\partial \Phi_T(h)} &= u'(c_T)\cdot\left[g(h)F(h,0)+(1-g(h))\int_\varepsilon f(\varepsilon)\left[F(h,\varepsilon - x^{SP}(\varepsilon,h)) - x^{SP}(\varepsilon,h)\right]d\varepsilon\right] \\ &\equiv u'(c_T)\cdot\Psi(h) \end{aligned} \tag{52}$$

For these expressions we only need to know c_T, the term $\Psi(h)$ is just a number that depends on h and is known once we have solved the static insurance problem. This suggests the following algorithm to solve the dynamic social planner problem:

Algorithm 17

1. *Guess a sequence $\{c_t\}_{t=0}^T$*
2. *Determine $\frac{\partial V_T(\Phi_T)}{\partial \Phi_T(h)}$ from (52)*
3. *Iterate on t to determine $\{e_t(h)\}_{t=0}^{T-1}$*
   (a) *For given $\frac{\partial V_{t+1}(\Phi_{t+1})}{\partial \Phi_{t+1}(h')}$ use (25) to determine $e_t(h)$.*
   (b) *Use $c_t, e_t(h)$, $\frac{\partial V_{t+1}(\Phi_{t+1})}{\partial \Phi_{t+1}(h')}$ and (26) to determine $\frac{\partial V_t(\Phi_t)}{\partial \Phi_t(h)}$*

4. Use the initial distribution Φ_0 and $\{e_t(h)\}_{t=0}^{T-1}$ to determine $\{\Phi_t\}_{t=0}^{T}$ and thus $\{c_t^{new}\}_{t=0}^{T}$.

5. If $\{c_t^{new}\}_{t=0}^{T} = \{c_t\}_{t=0}^{T}$ we are done. If not, set $\{c_t\}_{t=0}^{T} = \{c_t^{new}\}_{t=0}^{T}$ and go to 1.

This algorithm is straightforward to implement numerically, since we only have to iterate on the aggregate consumption sequence, not on the sequence of distributions. In particular, the only moderately costly operation comes in step 2a) but even there we only have to solve one nonlinear equation in one unknown (although we have to do it $T * card(H)$ times per iteration).

E Computation of the Equilibrium with a No-Prior-Conditions Law and/or a No-Wage Discrimination Law

The algorithm to solve this version of the model shares its basic features with that for the social planner problem, but differs in terms of the sequence of variables on which we iterate:

Algorithm 18 *1. Guess a sequence*[39] $\{Eu'_t, P_t\}_{t=0}^{T}$.

2. Given the guess use equations (29)-(32) to determine health cutoffs and wages $\{\bar{\varepsilon}_t^{NP}(h), w_t(h)\}$.

3. Given $\{w_t(h), P_t\}$, solve the household dynamic programming problem (33) for a sequence of optimal effort policies $\{e_t(h)\}_{t=0}^{T}$.

4. From the initial health distribution Φ_0 use the effort functions $\{e_t(h)\}_{t=0}^{T}$ to derive the sequence of health distributions $\{\Phi_t\}_{t=0}^{T}$ from equation (22).

5. Obtain a new sequence $\{Eu_t'^{new}, P_t^{new}\}_{t=0}^{T}$ from (31) and (32).

6. If $\{Eu_t'^{new}, P_t^{new}\}_{t=0}^{T} = \{Eu'_t, P_t\}_{t=0}^{T}$ we are done. If not, go to step 1. with new guess $\{Eu_t'^{new}, P_t^{new}\}_{t=0}^{T}$.

The algorithm for no-wage discrimination is a slight modification of that for no-prior conditions. The algorithm iterates over $\{Eu'_t, w_t\}_{t=0}^{T}$. In Step 1 given the guess use equations (34)-(38) to determine health cutoffs and premia $\{\bar{\varepsilon}_t^{NP}(h), P_t(h)\}$. In Step 4 obtain a new sequence $\{Eu_t'^{new}, w_t^{new}\}_{t=0}^{T}$ from (37) and (36). With both policies, equation (39) replaces (38) in all expressions.

F Data and Calibration

F.1 Augmented Model Analysis

We assume that households *must* incur the cost z, when z-shock is hit. This assumption and the fact that households are risk averse imply that the z-shock will be fully insured in the competitive equilibrium under any policy (and of course by the social planner).

Moreover, we assume that households receiving a z-shock can still work, but that their productivity is only ρ times that of a healthy worker. Therefore, in a competitive equilibrium, the wage of a worker with health status h is given by

$$w(h) = g(h)F(h,0) + \rho\kappa(h)F(h,0) + (1 - g(h) - \kappa(h)) \int F(h, \varepsilon - x(\varepsilon, h)) f(\varepsilon) d\varepsilon$$

and the health insurance premium is determined as

$$P(h) = (1 - g(h) - \kappa(h)) \int x(\varepsilon, h) f(\varepsilon) d\varepsilon + \mu_z(h)$$

Given our assumptions there is no interaction between the z-shocks and the health insurance contract problem associated with the ε-shock since it is prohibitively costly by assumption not to bear the z-expenditures. The

[39] Instead of $\{Eu_t\}$ one could iterate on $\{w_t(h)\}$ which is more transparent, but significantly increases the dimensionality of the problem.

role of the z-expenditures is to soak up the most extreme health expenditures observed in the data associated with catastrophic illnesses, but to otherwise leave our theory from the previous sections unaffected.

The static analysis goes through completely unchanged in the presence of the z-shocks. In the dynamic analysis the benefits of higher effort e and thus a better health distribution $\Phi_t(h)$ now also include a lower probability $\kappa(h)$ of receiving a positive z-shock and a lower mean expenditure $\mu_z(h)$ from that shock with better health h. This extension of the model leads to straightforward extensions of the expressions derived in the analysis of the dynamic model in section 4, and does not change any of the theoretical properties derived in sections 3 and 4

F.2 Descriptive Statistics of the PSID Data

Before we proceed to descriptive statistics of the PSID data, we summarize, in Table 3, the mapping between variables in our model and data.

Table 3: Mapping between Data and Model

Model	Description	Data	
		PSID Variable	Actual Data Used
x, μ_z	Medical Expenditure	Average of total expenditure reported in 1999, 2001, 2003	1997-2002
w	Earning	Average of total labor income reported in 1999, 2001, 2003	1998,2000,2002
h	Health Status	Self-reported Health in 1997	1997

Since our model period is six years, we take average of reported medical expenditure and wages over six year periods that we observe. Moreover, we use health status data from 1997 (rather than 1999) to capture the effect of health on wages and medical expenditure.

Table 4 documents descriptive statistics of key variables from the 1999 PSID data that we use in our analysis.

Table 4: Descriptive Statistics of Key Variables in PSID

	Mean	Std. Dev.	Min	Max
Age	41	10	23	65
Labor Income	30,170	40,573	0	1,153,588
if Labor Income > 0	32,076	41,097	0.55	1,153,588
Excellent	38,755	55,406	0	940,804
Very Good	32,768	40,351	0	1,153,588
Good	25,516	25,908	0	384,783
Fair	12,605	13,926	0	81,300
Medical Expenditure	1,513	4,624	0	127,815
Excellent	1,234	2,374	0	28,983
Very Good	1,647	5,812	0	127,815
Good	1,486	4,283	0	93,298
Fair	1,792	4,950	0	65,665
Health Status	2.77	0.95	1	4
Physical Activity: fraction(number) of days in a year				
Light	0.63 (230.99)	0.39 (142.28)	0	1 (365)
Heavy	0.29 (105.69)	0.35 (126.85)	0	1 (365)

In the PSID, each individual (head of household) self-reports his health status in a 1 to 5 scale, where 1 is Excellent, 2, Very Good, 3, Good, 4, Fair, and 5 is Poor. Even with large number of observations, only about 1% of total individuals report their health status to be poor. Thus, for our analysis, we will use four levels of health status (merge poor and fair together).[40] Since PSID reports household medical expenditure, we control for family size using modified OECD equivalence scale. [41]

As we model working-age population, each household starts his life as a 24 year old and makes economic decisions until he is 65 years old. Our model time period is 6 years and thus they live for 7 time periods. We choose six year time period to capture the effect of exercises on health transition. Since exercises tend to have positive longer-term effects than do medical expenditure, by allowing for a medium-term time period, we are able to quantify the impact of exercises in a more reliable way.

Table 5 presents transition matrix of health status over six years. We see that health status is quite persistent.

Table 5: Health Transition over 6 years

	Excellent	Very Good	Good	Fair	Total
Excellent	1,286	904	335	92	2,617
	49.14 %	34.54 %	12.80 %	3.52 %	100 %
Very Good	482	1,844	1,217	274	3,817
	12.63 %	48.31 %	31.88 %	7.18 %	100 %
Good	187	712	1,592	637	3,128
	5.98 %	22.76 %	50.90 %	20.36 %	100 %
Fair	36	109	358	957	1,460
	2.47%	7.47 %	24.52 %	65.55 %	100 %
Total	1,991	3,569	3,502	1,960	11,022
	18.06 %	32.38 %	31.77%	17.78 %	100 %

Physical Activity Data Here, we report some statistics on physical activity.

- *Variation of Physical Activity and Its Impact on Health Transition*

  Density of light and heavy physical activity levels by health are summarized in Figures 14 and 15. From variations in health evolution by physical activity and initial health status, we find that about 30% of variance in health status in the future is explained by health status today, whereas, light and physical activity explains about 8% and 14%, respectively. Moreover, both initial health status and light (heavy) exercise explains 46% (41%) of variance in future health outcome.[42]

- *Physical Activity Over Time*

  Light physical activity has steadily decreased over time, whereas heavy physical activity decreased for a while, but started increasing in 2005 (Figures 16 and 17).

F.3 Health Shocks, Distribution of Medical Expenditure, and Discussion on Categorization of Health Shocks

Before going into discussing the medical expenditure distribution in data, we briefly discuss the appropriate counterparts of data moments for our model. In our model, households do not consume medical care when they do not get a health shock (although, they can choose not to spend any in case of health shock, since $x^*(h,\varepsilon) = \max\{0, \bar{\varepsilon}(h)\}$). Therefore, in data, we are interested in the distribution of medical expenditure conditional on having gotten any health shocks (which we have some information in PSID).

The Table 6 summarizes medical expenditure by shock. Note that all numbers reported are yearly average taken over six years (1997-2002).

[40] Labor income and medical expenditure data for fair health in Table 4 include poor (5) in data.

[41] Each additional adult gets the weight of 0.5, and each child, 0.3.

[42] From the law of total variance, we know

$$var(Y) = \mathbb{E}\left(var(Y|X)\right) + var\left(\mathbb{E}(Y|X)\right),$$

where the former is the *unexplained* and the latter, *explained* component of the variance.

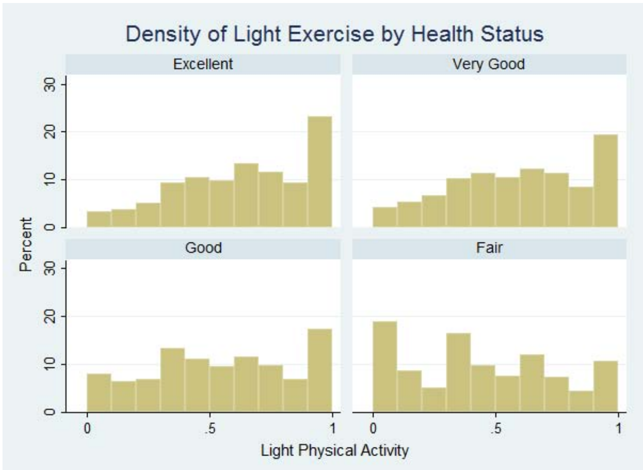

Figure 14: Density of Light Physical Activity

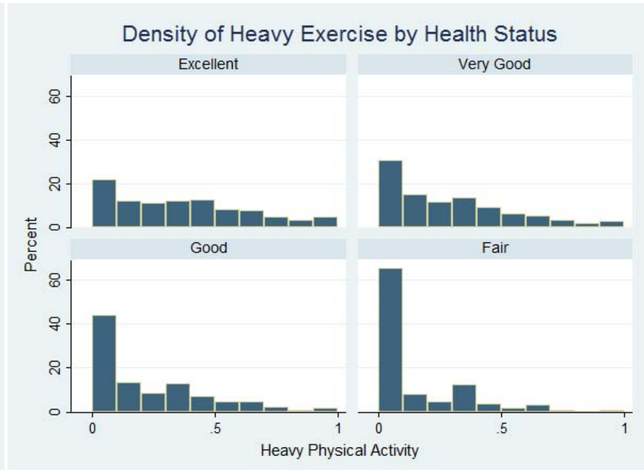

Figure 15: Density of Heavy Physical Activity

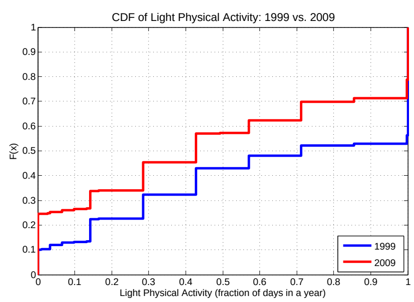

Figure 16: Light Activity

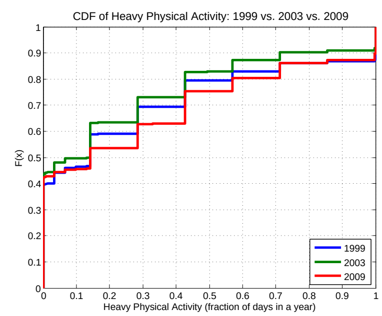

Figure 17: Heavy Activity

Table 6: Average Medical Expenditure by Health Shock Categories

	Obs	Mean	Std. Dev.	Min	Max
All	4,226	1,513	4,624	0	127,815
No Shock	1,419	1,350	4,447	0	101,952
Any Shock	2,807	1,595	4,710	0	127,815
Catastrophic Disease Shock	168	3,745	9,363	0	93,298
Cancer	51	5,210	15,134	0	93,298
Heart Attack	46	3,334	4,705	0	27,161
Heart Disease	94	3,382	5,535	0	38,500
Light Shock	2,767	1,585	4,732	0	127,815
Diabetes	183	2,088	7,196	0	93,298
Stroke	33	2,200	4,905	0	27,161
Arthritis	322	1,684	3,166	0	38,500
Hypertension	566	1,825	6,143	0	93,298
Lung Disease	63	1,705	2,476	0	12,595
Asthma	61	1,135	1,444	0	7,170
Ill	2,351	1,637	5,040	0	127,815
z-shock	297	4,704	12,834	0	127,815
ε-shock	2,510	1,227	2,023	0	32,909

We see that cancer, heart attack, and heart disease incur the most medical expenditure, and thus we categorize them to be *catastrophic* shocks (z-shocks). Although the diseases PSID specifically reports information on are those that are common, they are not, by all means, exhaustive of the kind of health diseases that one can be diagnosed with. And this is hinted when we look at the medical expenditure statistics for those who report to have missed work due to illness. The maximum amount of medical expenditure they spend exceeds those of the others, and this might be due to some severe diseases for which they had to be treated.

Therefore, in addition to cancer, heart attack, and heart disease, we categorize those who have spent more than their labor income on medical expenditure as having had a *catastrophic (z) health shock*.[43] Those who had a health shock that were not cancer, heart attack, or heart disease, and who spent less than their income on medical expenditure is considered to have had ε-shock.[44]

Figures 18 - 21, plot logs of medical expenditure distribution for all population, for those with ANY health shock, those with z-shock, and those with ε-shock. By definition, mean medical expenditure of z-shock households are higher than those of ε-shock, and so are standard deviations.

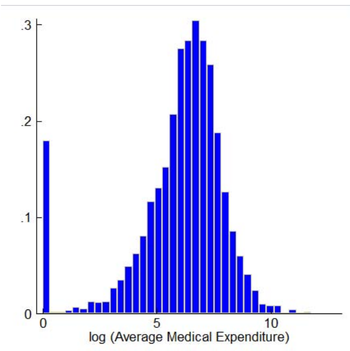

Figure 18: Average Medical Expenditure Distribution

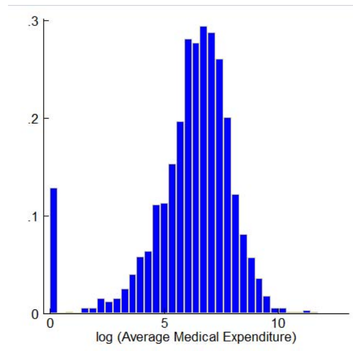

Figure 19: Average Expenditure with Health Shock

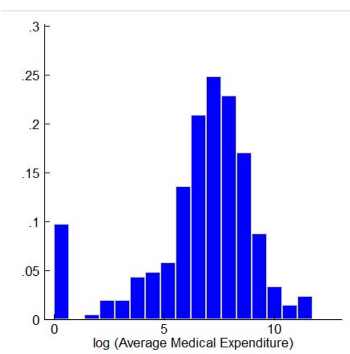

Figure 20: Average Expenditure w/ z-shock

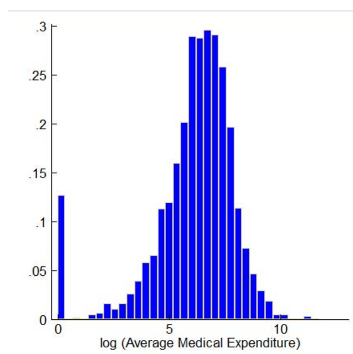

Figure 21: Average Expenditure w/ ε-shock

F.4 Estimation Results (Outside the Model)

Health Transition Using the functional form described in the main body of the paper, we estimate the health transition function in the following way.

[43]Categorizing catastrophic health shocks using expenditures as percentage of income is not new. There has been discussion on insuring catastrophic health shocks, and they mostly refer to *high* amount of expenditure as percentage of income.

[44]In PSID sample, median of percentage of labor income spent on medical expenditure is 2%, and the mean, 132%. Only about 5% of households with health shocks spend medical expenditure in excess of their labor income.

Let set of parameters to be estimated be $\boldsymbol{\theta} = \left\{ \{G(h,h')\}, \delta, \phi(h), \lambda(h), \alpha_1(h), \alpha_2(h) \right\}$. We use Generalized Method of Moments to estimate these parameters.

We first determine the exercise intervals and assign each individual initial health status and exercise level bins, k.[45] Using the transition from the data $\mathbb{E}(q^k(h'))$, we minimize the distance between our estimated transition function and data, i.e.

$$\boldsymbol{\theta} = \arg\min_{\boldsymbol{\theta}} \left(\frac{1}{K} \sum_{k=1}^{K} \left[Q(h'_k; \boldsymbol{\theta}) - \mathbb{E}(q^k(h')) \right] \right) \hat{\mathbf{W}} \left(\frac{1}{K} \sum_{k=1}^{K} \left[Q(h'_k; \boldsymbol{\theta}) - \mathbb{E}(q^k(h')) \right] \right),$$

where K denotes the total number of groups and $\hat{\mathbf{W}}$, weighting matrix. Here, we use the efficient weighting matrix.

With exercise step size of nine,[46] we get the following estimated parameter values ($h = 1, 2, 3, 4$ corresponds to health being fair, good, very good, and excellent, respectively, i.e. the higher the h the better one's health status.).

$$\begin{aligned}
\hat{G}(h,h') &= \begin{bmatrix} 0.8742 & 0.0927 & 0.0230 & 0.0101 \\ 0.6597 & 0.2547 & 0.0609 & 0.0249 \\ 0.1404 & 0.3949 & 0.3442 & 0.1204 \\ 0.0850 & 0.3170 & 0.5406 & 0.0573 \end{bmatrix} \\
\delta &= 1 \\
\phi &= [2.2796, 1.1063, 0.5179, 8.4123] \\
\lambda &= [0.3308, 0.0193, 0.5939, 0.1878] \\
\alpha_1 &= [1.3274, 12.8747, 7.0260] \\
\alpha_2 &= [0.8035, 5.8693]
\end{aligned}$$

The estimated transition functions are plotted in Figures 22 - 25. In the figures, the smoothed functions are estimated transition, whereas the straight lines represent the data. We see that our functional form fits the data quiet well.

Table 7: Health Shock Probabilities by Health Status

	Observations	Any Health Shock $1-g(h)$	z Shock $\kappa(h)$	ε Shock $1-g(h)-\kappa(h)$
All	4,226	0.66	0.07	0.59
Fair	458	0.66	0.21	0.45
Good	1,139	0.71	0.07	0.63
Very Good	1,618	0.68	0.05	0.62
Excellent	1,143	0.60	0.03	0.57

Health Shock Probabilities As seen in Table 7, there are about 7% of households who receive z-shocks over six years, and the probabilities are decreasing in health status. However, probability of getting *any* health shock is not the highest for the Fair health individuals (from Good to Excellent, it is monotone). This might be due to the fact that given that the health status is already bad, probabilities that one would get other minor adverse health shocks (ε shocks in the model) are not very high.

[45] In data, there are households with more than 2 steps of differences in health status. We included them as having transited to either one step above or below their initial health status in calculating the transition probabilities from data.

[46] PSID has exercise data from 1999 to 2009. The total number of observations for 6 year transition is 11,022.

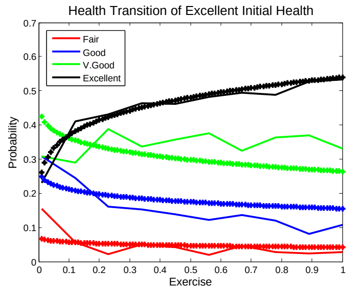

Figure 22: GMM: Transition of Excellent Health

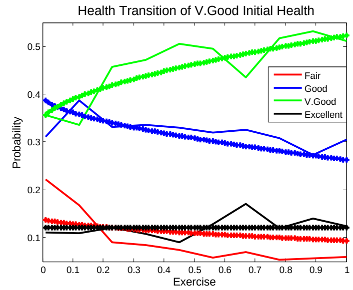

Figure 23: GMM: Transition of Very Good Health

Figure 24: GMM: Transition of Good Health

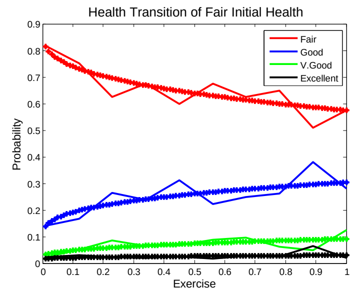

Figure 25: GMM: Transition of Fair Health

Effect of Health Shock on Productivity In Table 8, we summarize working hours and labor income reported by those with different health shock categories.

The six year average hours worked of those with z-shocks are about half that of the ones who did not get any shock (and worked) and they earn about half on average. Therefore, we take $\rho = 0.4235$, which is the percentage of labor income earned by those with z-shock, compared to those who have worked and did not experience any health shock (since we denote earnings of those with z-shock as $\rho F(h, 0)$).

Table 8: Hours Worked and Labor Income by Health Shock

		Obs	Mean	Std. Dev.	Min	Max
Hours Worked	All	4,226	1,823	856	0	5,300
	Positive Hours	3,903	1,974	704	7	5,300
	No Shock, Positive Hours	1,259	1,987	781	14	4,732
	z-shock	297	998	1,033	0	3,640
	ε-shock	2,639	1,892	763	0	5,300
Labor Income	All	4,226	30,171	40,573	0	1,153,588
	Positive Hours	3,903	32,362	41,364	0	1,153,588
	No Shock, Positive Hours	1,259	32,606	49,358	0	940,804
	z-shock	297	13,809	25,470	0	253,560
	ε-shock	2,639	31,163	36,883	0	1,153,588

F.5 Estimation/Calibration Results (Within the Model)

Table 9: Data Targets

	Parameters	Data Targets
Health Status	$\{h_i\}_{i=1,2,3,4}$	Income of h_i relative to h_1 $\log \frac{w(h_2)}{w(h_1)} = 0.2739$ $\log \frac{w(h_3)}{w(h_1)} = 0.4691$ $\log \frac{w(h_4)}{w(h_1)} = 0.5948$ Income of Old relative to $Young$ $\log \frac{w(O)}{w(Y)} = 0.1114$
Production Function	$A(t, educ)$	Income in t of less than HS relative to Income of $Young$ and Fair health $t = 1, < HS : -0.0042$ $t = 2, < HS : 0.1449$ $t = 3, < HS : 0.1715$ $t = 4, < HS : 0.1980$ $t = 5, < HS : 0.0907$ $t = 6, < HS : -0.0969$ $t = 7, < HS : -0.1112$ Income in t of HS Grad relative to Income of $Young$ and Fair health $t = 1, HS : 0.2980$ $t = 2, HS : 0.4738$ $t = 3, HS : 0.5082$ $t = 4, HS : 0.5988$ $t = 5, HS : 0.6060$ $t = 6, HS : 0.5395$ $t = 7, HS : 0.2406$
	$\phi(a, educ)$	% Income spent on Med Exp. by Health $\frac{\mathbb{E}(x\|h_1)}{\mathbb{E}(w\|h_1)} = 0.0525$ $\frac{\mathbb{E}(x\|h_2)}{\mathbb{E}(w\|h_2)} = 0.0429$ $\frac{\mathbb{E}(x\|h_3)}{\mathbb{E}(w\|h_3)} = 0.0353$ $\frac{\mathbb{E}(x\|h_4)}{\mathbb{E}(w\|h_4)} = 0.0308$
	$\xi(a, educ)$	% Income on Med Exp. by Education and Age($a \in \{Y, O\}$) $\frac{\mathbb{E}(x\|Y, < HS)}{\mathbb{E}(w\|Y, < HS)} = 0.0386$ $\frac{\mathbb{E}(x\|Y, HS)}{\mathbb{E}(w\|Y, HS)} = 0.0348$ $\frac{\mathbb{E}(x\|O, < HS)}{\mathbb{E}(w\|O, < HS)} = 0.0428$ $\frac{\mathbb{E}(x\|O, HS)}{\mathbb{E}(w\|O, HS)} = 0.0356$

Table 10: Data Targets (continued)

	Parameters	Data Targets
ε-shock Distribution	μ, σ_ε	Mean and St.Dev of Agg. Med. Exp. on Light Shocks $\frac{\mathbb{E}(x)}{\mathbb{E}(w)} = 0.0362$ $\frac{\sigma(x)}{\mathbb{E}(x)} = 1.6462$
z-shock Distribution	$\mu_z(h)$	% of Income Spent on Catastrophic Shock by Health $\frac{\mathbb{E}(z\|h_1)}{\mathbb{E}(w\|h_1)} = 0.4664$ $\frac{\mathbb{E}(z\|h_2)}{\mathbb{E}(w\|h_2)} = 0.2234$ $\frac{\mathbb{E}(z\|h_3)}{\mathbb{E}(w\|h_3)} = 0.2681$ $\frac{\mathbb{E}(z\|h_4)}{\mathbb{E}(w\|h_4)} = 0.1261$
Exercise Disutility	$\{\gamma_1(educ), \gamma_2(educ)\}$	Mean and St.Dev of Exercise in $t = 1$ $\mathbb{E}(e_{t=1}) = 0.5735$ $\sigma^2(e_{t=1}) = 0.2828$ Measure of Fair and Excellent in $t = 7$ $\Phi_{t=T}(h_1) = 0.1944$ $\Phi_{t=T}(h_4) = 0.1618$
Preference Distribution	$p(\gamma\|educ, h)$	Mean Exercise in $t = 1$ by Health $\mathbb{E}(e_{t=1}\|h_1) = 0.5030$ $\mathbb{E}(e_{t=1}\|h_2) = 0.5235$ $\mathbb{E}(e_{t=1}\|h_3) = 0.5950$ $\mathbb{E}(e_{t=1}\|h_4) = 0.6087$ Mean Exercise by Education in $t = 1, 7$ $\mathbb{E}(e_{t=1}\| < HS) = 0.5303$ $\mathbb{E}(e_{t=1}\|HS) = 0.5956$ $\mathbb{E}(e_{t=7}\| < HS) = 0.5517$ $\mathbb{E}(e_{t=7}\|HS) = 0.6159$
Terminal (Marginal) Value	$\{\Delta_2, \Delta_3, \Delta_4\}$	Exercise in the Last Period by Health $\mathbb{E}(e_{t=T}\|h_1) = 0.4641$ $\mathbb{E}(e_{t=T}\|h_2) = 0.6092$ $\mathbb{E}(e_{t=T}\|h_3) = 0.6535$

Table 11: Calibrated Parameters

Parameters	Description	Value	Statistics	Data	Model
h_1	Health Status	0.0720	Relative log Wages in Health and Age	0.2739	0.2967
h_2		0.1302		0.4691	0.5185
h_3		0.2012		0.5948	0.6083
h_4		0.2197		0.1114	0.1064
$A(t=1,<HS)$	Effect of Age, Education on Productivity	0.9957	Relative log Wages in Time and Education	-0.0042	-0.0048
$A(t=2,<HS)$		1.9560		0.1449	0.1428
$A(t=3,<HS)$		2.2002		0.1715	0.1689
$A(t=4,<HS)$		2.6227		0.1980	0.1985
$A(t=5,<HS)$		1.9144		0.0907	0.0905
$A(t=6,<HS)$		0.8189		-0.0969	-0.0972
$A(t=7,<HS)$		0.7141		-0.1112	-0.1176
$A(t=1,HS)$		2.5295		0.2980	0.2954
$A(t=2,HS)$		3.7442		0.4738	0.4186
$A(t=3,HS)$		4.0988		0.5082	0.4329
$A(t=4,HS)$		6.1605		0.5988	0.6230
$A(t=5,HS)$		6.3582		0.6060	0.6308
$A(t=6,HS)$		5.6297		0.5395	0.5565
$A(t=7,HS)$		2.8878		0.2406	0.2417
$\phi(Y,<HS)$	Effect of Med. Exp. on Productivity	0.7991	% Income on Med.Exp. by Health,Education,Age	0.0525	0.0562
$\phi(O,<HS)$		0.8739		0.0429	0.0441
$\phi(Y,HS)$		0.8612		0.0353	0.0332
$\phi(O,HS)$		0.9056		0.0308	0.0268
$\xi(Y,<HS)$		0.0110		0.0386	0.0401
$\xi(O,<HS)$		0.0050		0.0348	0.0383
$\xi(Y,HS)$		0.0119		0.0428	0.0454
$\xi(O,HS)$		0.0083		0.0356	0.0368
μ_ε	Mean of health shock	0.6638	Mean Medical Expenditure	0.0362	0.0417
σ_ε	St. Dev. of health shock	0.0710	St.Dev Medical Expenditure	1.6462	2.4454
$\mu_z(h_1)$	Mean of z-shock	0.3529	Income spent on Catastrophic Shock	0.4664	0.4760
$\mu_z(h_2)$		0.2204		0.2234	0.2209
$\mu_z(h_3)$		0.3356		0.2681	0.2695
$\mu_z(h_4)$		0.1721		0.1261	0.1263
$\gamma_1(<HS)$	Disutility by Education	0.0008	Mean of Exercise, $t=1$	0.5735	0.5834
$\gamma_2(<HS)$		0.1380	St.Dev Exercise, $t=1$	0.2828	0.2773
$\gamma_1(HS)$		0.0001	Measure of Fair in $t=T$	0.1944	0.2294
$\gamma_2(HS)$		0.0796	Measure of Ex. in $t=T$	0.1618	0.1295
$p(\gamma_1 \| <HS,h_1)$	Pop. with γ_1 by Health,Education	0.0419	Conditional Mean Effort by Health, Education in $t=1,7$	0.5030	0.4964
$p(\gamma_1 \| <HS,h_2)$		0.4361		0.5235	0.5220
$p(\gamma_1 \| <HS,h_3)$		0.1462		0.5950	0.6129
$p(\gamma_1 \| <HS,h_4)$		0.8856		0.6089	0.6254
$p(\gamma_1 \| HS,h_1)$		0.0565		0.5303	0.5465
$p(\gamma_1 \| HS,h_2)$		0.0557		0.5956	0.5936
$p(\gamma_1 \| HS,h_3)$		0.5409		0.5517	0.5525
$p(\gamma_1 \| HS,h_4)$		0.3705		0.6159	0.6128
Δ_2	Marginal value of health in $t=T$	0.0018	Conditional Mean Effort in $t=T$	0.4641	0.4895
Δ_3		1.8571		0.6092	0.6015
Δ_4		2.4752		0.6535	0.6627

G Quantitative Results

G.1 Model Fit

Figures 26–29 represent the model fit for average effort of each health level.

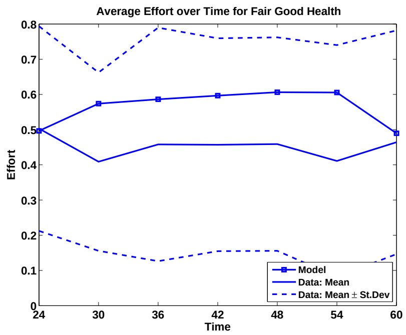

Figure 26: Average Effort: Fair

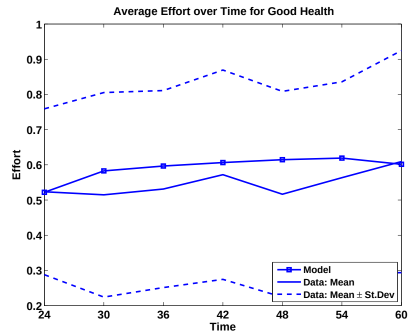

Figure 27: Average Effort: Good

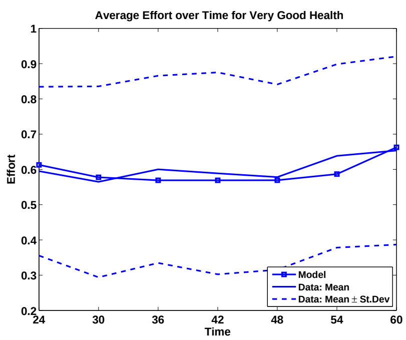

Figure 28: Average Effort: Very Good

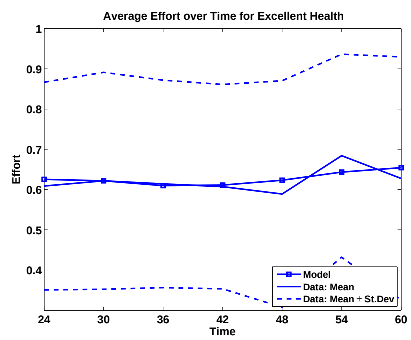

Figure 29: Average Effort: Excellent

G.2 Policy Implications

In the main body of our paper, we presented aggregate and welfare consequences of policies for those who are high school graduates and have high disutility of exercise. Here, we report the aggregate consequences, as measured by weighted averages[47] of households with different education, health, and preference for exercise.

Insurance Benefits The Table 12 presents the weighted-averages (across education and exercise preference) of the cross-subsidies by health level under different policy regimes. We measure cross subsidies in premium by the differences between the actuarially fair health premium and premium paid under policies; and cross subsidies in wage by the differences between the aggregate wage and productivity of the worker (of a given health level). As discussed in the main text, the negative cross-subsidy implies that the worker is paying higher premium than the actuarially fair price and/or getting paid less in wages than he produces.

Since under no-prior conditions law, only premium is subsidized, and under no-wage discrimination law, only wage is subsidized, we report cross-subsidies of premium and wages under each law. The second row under each health level reports separately the subsidies of premium and wage, under both policies.

[47] The education and health measures are obtained directly from data, and the measures of low and high disutility for each education, health group from calibration exercise described in Section .

Table 12: Cross-Subsidy by Health Level under Different Policy Regimes

Health	Policy	24–29		30–35		36–41		42–47		48–53		54–59		60–65	
		Prem.	Wage	Prem.	Wage	Prem.	Wage	Prem.	Wage	Prem.	Wage	Prem.	Wage	Prem.	Wage
Fair	One Policy	0.267	0.289	0.286	0.368	0.277	0.356	0.279	0.461	0.267	0.452	0.247	0.399	0.237	0.258
	Both Policies	0.087	0.234	0.081	0.306	0.075	0.296	0.072	0.395	0.071	0.388	0.071	0.338	0.072	0.197
Good	One Policy	0.040	0.079	0.016	0.063	-0.016	0.026	-0.031	0.023	-0.046	0.013	-0.048	0.005	-0.041	-0.019
	Both Policies	-0.030	0.098	-0.048	0.087	-0.055	0.056	-0.058	0.058	-0.059	0.050	-0.059	0.045	-0.058	0.026
Very Good	One Policy	0.038	0.022	-0.027	-0.099	-0.051	-0.159	-0.070	-0.270	-0.067	-0.277	-0.060	-0.228	-0.061	-0.096
	Both Policies	0.067	-0.024	0.065	-0.137	0.059	-0.194	0.055	-0.310	0.054	-0.320	0.054	-0.272	0.054	-0.143
Excellent	One Policy	-0.163	-0.185	-0.201	-0.322	-0.218	-0.390	-0.235	-0.549	-0.233	-0.560	-0.228	-0.500	-0.235	-0.333
	Both Policies	-0.088	-0.124	-0.104	-0.249	-0.110	-0.312	-0.114	-0.466	-0.115	-0.477	-0.116	-0.414	-0.115	-0.240

Aggregate Effects Figures 30 - 35 represent the aggregate variables under different policies. The aggregate variables are generated by taking the weighted averages across distribution of education and exercise disutility.

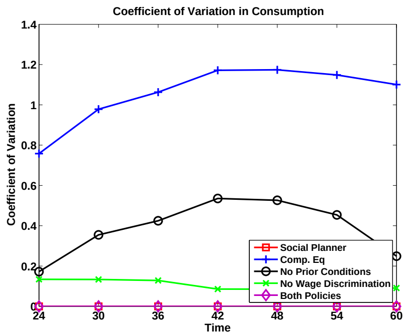

Figure 30: Consumption Dispersion: Weighted Averages

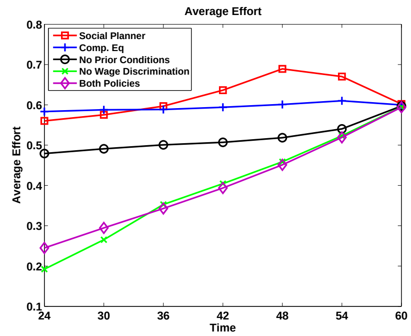

Figure 31: Exercise

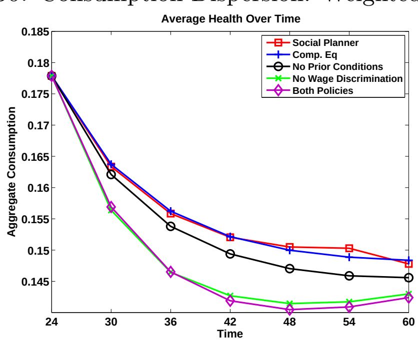

Figure 32: Average Health

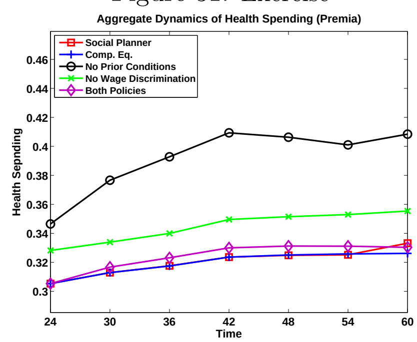

Figure 33: Medical Expenditure

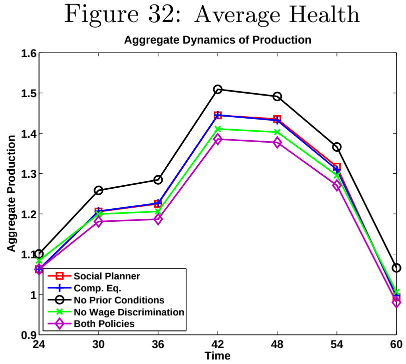

Figure 34: Production: HSGrad, γ_H

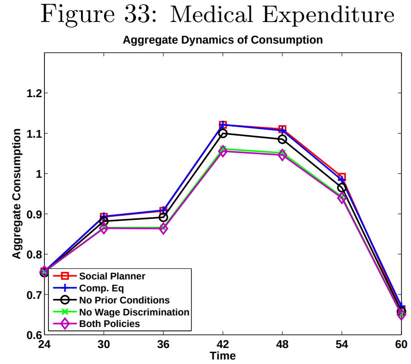

Figure 35: Consumption

Weafare Implications Tables 13 and 14 shows consumption equivalent variation for all groups as well as the aggregates.

	(< HS,γ_L)	(< HS,γ_H)	(HS Grad,γ_L)	(HS Grad,γ_H)	Aggregate
Social Planner	3.5281	4.8273	2.6091	6.8525	5.1777
Competitive Equilibrium	0.0000	0.0000	0.0000	0.0000	0.0000
No Prior Conditions Law	3.0924	4.1973	2.3988	5.8712	4.4918
No Wage Discrimination Law	3.1616	4.3915	2.3655	6.5729	4.8780
Both Policies	3.5281	4.8273	2.6091	6.8525	5.1777

Table 13: Welfare Comparisons in Static Economy

	(< HS,γ_L)	(< HS,γ_H)	(HS Grad,γ_L)	(HS Grad,γ_H)	Aggregate
Social Planner	8.1733	10.4419	13.5278	15.1429	13.5151
Competitive Equilibrium	0.0000	0.0000	0.0000	0.0000	0.0000
No Prior Conditions Law	5.6262	6.0006	6.1596	8.0140	7.0531
No Wage Discrimination Law	5.2831	5.8135	6.0787	8.1140	7.0303
Both Policies	5.1933	4.9675	7.4836	7.3411	6.8849

Table 14: Welfare Comparisons in Dynamic Economy

H Sensitivity Analysis

In this subsection, we check the robustness of our quantitative results with respect to race and gender.

PSID asks questions on ethnicity[48], and among them, we take those who answered to be of a national origin (47% of the total sample in 1997) to test robustness. We also restrict our sample to males (about 77%) for the second robustness check.

The health transition function and production funcion related parameters are the key driving forces of our quantitative results. Therefore, we provide evidece for the similarity in health transition and the labor earnings over the life cycle between the total population and the subsamples.

For the health transition function $Q(h'|h,e)$, we obtain a measure of differences in the estimated probabilities and the data moments, i.e.,

$$\chi^2 = \sum_{i=1,N} \frac{q^{data}(h'|h,e^i) - Q^{est}(h'|h,e^i)}{Q^{est}(h'|h,e^i)},$$

where the $q^{data}(h'|h,e^i)$ and $Q^{est}(h'|h,e^i)$ are the actual data and the estimated probability of a worker with initial health status h with exercise level e^i[49] ending up being health status of h' in the next period. The χ^2 value for the health transition is 1.16 and 1.02 for whites and males, where the $\chi^2_{49,0.05}$[50] is 79.

With regards to the production function, we provide in Table 15, the data moments associated with the subsamples, in comparison with the full sample. The qualitative features of the moments are similar across different samples: although the absolute numbers for the changes in income over the life-cycle vary in their levels, the gradients over the life cycle are similar. Thus our quantitative results are robust to restricting our samples to white and males.

H.1 Benefits of Effort Not Related to Labor Productivity

So far the only benefit of effort e consisted in probabilistically raising health in the future which in turn impacts positively future wages and health insurance premia. As a result, a combination of both policies reduces optimal effort to zero. We now briefly argue that our main results do not necessarily hinge on this assumption. Suppose that the net cost of providing effort is given by

$$\gamma \left[q(e) - \theta e \right].$$

Our previous specification is a special case with $\theta = 0$, and $\gamma\theta$ measures the direct utility benefit from one unit of exercise. In the absence of any other benefits from exercise (say, from higher wages or lower health insurance premia), as in the economy with both laws in place, the optimal effort level e^{BP} now solves

$$q'(e^{BP}) = \theta$$

and thus $e^{BP} > 0$ if and only if $\theta > 0$. Thus for a given function q the parameter θ governs the minimal effort level that each household will provide, and thus a lower bound below which no policy can distort effort levels.

The equations determining optimal effort levels (equation (25) for the social planner problem and equation (28) for the competitive equilibrium under the various policies) with preference shocks γ and direct utility benefits from exercising $\gamma\theta e$ now become

$$\begin{aligned} q'(e_t(h)) &= \theta + \frac{\beta}{\gamma} \sum_{h'} \frac{\partial V_{t+1}(\Phi_{t+1})}{\partial \Phi_{t+1}(h')} \cdot \frac{\partial Q(h';h,e_t(h))}{\partial e_t(h)} \\ q'(e_t(h)) &= \theta + \frac{\beta}{\gamma} \sum_{h'} \frac{\partial Q(h';h,e_t(h))}{\partial e_t(h)} v_{t+1}(h') \end{aligned}$$

[48] The exact choices are American (5%); Hyphenated American (e.g., African-American, Mexican-American) (14%); National origin (e.g., French, German, Dutch, Iranian, Scots-Irish) (47%); Nonspecific Hispanic identity (e.g., Chicano, Latino) (2%); Racial (e.g., white or Caucasian, black) (29%) and; 6 Religious (e.g., Jewish, Roman, Catholic, Baptist).

[49] We divide the population into five exercise bins, and use them to evaluate the differences, as we do in our estimation procedure. The only difference is that due to the shortage of observations (since we only use half the total sample), instead of nine bins (in the full model), we use five bins.

[50] The degrees of freedom is 49, as the number of observations are $4 \times 4 \times 5$ ($\sharp$ Health Today $\times$ $\sharp$ Health Tomorrow $\times$ $\sharp$ Exercise Bins), and the number of parameters, 30 (80-1-30). Using the full sample, the χ^2 value is 0.9986.

Moments Description		All	Whites	Male
Income by Age of Less than HS	$t = 1$	-0.0042	-0.0892	0.0320
	$t = 2$	0.1449	0.2026	0.1214
	$t = 3$	0.1715	0.2464	0.1653
	$t = 4$	0.1980	0.2901	0.2091
	$t = 5$	0.0907	0.0014	0.0183
	$t = 6$	-0.0969	-0.3306	-0.1409
	$t = 7$	-0.1112	-0.0970	-0.0742
Income by Age of HS Grad.	$t = 1$	0.2980	0.3019	0.2990
	$t = 2$	0.4738	0.5867	0.4835
	$t = 3$	0.5082	0.6073	0.5522
	$t = 4$	0.5988	0.6274	0.6027
	$t = 5$	0.6060	0.6500	0.6216
	$t = 6$	0.5395	0.5276	0.5366
	$t = 7$	0.2406	0.1792	0.3376
% Income Spent on Med. Exp.	Fair	0.0525	0.0573	0.0482
	Good	0.0429	0.0428	0.0395
	Very Good	0.0353	0.0376	0.0346
	Excellent	0.0308	0.0320	0.0290
% Income Spent on Med. Exp by Less than HS	Young	0.0386	0.0350	0.0373
	Old	0.0348	0.0357	0.0376
% Income Spent on Med. Exp. by HS Grad	Young	0.0428	0.0465	0.0495
	Old	0.0356	0.0379	0.0447

Table 15: Moments for the Subsample of Population

and for any given initial health level h, any preference shock γ and any policy the optimal effort level is simply shifted upwards. Our theoretical results remain intact, and importantly, effort is still lowest under a combination of both policies. Quantitatively however, the dynamics of the health distribution in the presence of both policies is altered by the fact that now effort is uniformly positive.