Estimating the Net Benefit of Customer Health Engagement Programs to Life Insurers

Hae Kang Lee†
NYU Stern School of Business
Job Market Paper
November 19, 2020

Abstract
This paper investigates the net benefit available to a life insurance company when it engages more with the policyholders to improve their health status. Using a proprietary big database of health and mortality information from a large U.S. life insurer, coupled with state-of-the-art machine learning techniques, I measure the financial benefits that will accrue to the insurer if it can change the health variables of existing policyholders. The cost of changing the health behaviors is measured from a rational addiction model, which I model and calibrate to be consistent with the vast health economics and medical literature on smoking cessation and other health interventions. The estimated net benefit available to the life insurance company from the smoking cessation program is around USD 87 million. The aggregate benefit available including other chronic conditions is around USD 872 million. I discuss the implementation issues and policy implications.

*I am extremely grateful to Stijn Van Nieuwerburgh, Ralph Koijen, Arpit Gupta, Alexi Savov, and Menachem Brenner for their excellent guidance and support. I gratefully acknowledge the data support from the Life Insurance Company. I also acknowledge that the LifeScore Labs generously allowed access to their risk models.
†Email: haekang.lee@stern.nyu.edu.
1 Introduction

This paper investigates the net benefit available to a life insurance company when it engages more with the policyholders to improve their health status. For a long period of time, the insurance industry has been known for its infrequent interactions with the existing policyholders, mainly due to the long-term nature of its business. With rapid advances in information technology, however, there are signs of change. Access to the big data, coupled with new statistical techniques such as machine learning, puts the insurance industry in a unique position to generate the insightful information relevant to the policyholders’ health and mortality. If such information can be frequently communicated and even used to incentivize the policyholders to adopt healthier lifestyles, not only will better health and longer lifespans follow, but financial benefits will accrue to the life insurance companies due to higher policy values. Capitalizing on this big trend, the insurance industry has started various health engagement programs, ranging from providing mobile applications to track healthier lifestyles, to providing subsidies for gym memberships (see Table 1 for the selected list of programs).

In this paper, I explore the viability of various customer health engagement programs and estimate the aggregate effect, by carefully investigating the financial incentives to the life insurance company and to policyholders. As my leading example, I consider a smoking cessation program. If a policyholder stops smoking, her health status improves and her expected lifespan increases. The life insurance company also benefits financially for two reasons: the delayed death-benefit payment and the longer stream of the premium income. However, smoking is addictive and there is a high mental cost associated with quitting. The insurance company can design a health engagement program by sharing some of the potential financial benefits \textit{ex ante} to incentivize the policyholders to quit smoking. The same idea can be applied more broadly to various types of health intervention programs to improve the health status of the policyholders.
To precisely measure the net benefit to the insurance company of improving a health risk variable, two components must be separately estimated: (i) the financial benefit to the life insurer by measuring the improvement in survival probabilities and the distribution of policies, and (ii) the cost to change the health behaviors. Below, I summarize the methodologies I use to estimate each component in further details.

For the benefit estimation, I rely on the extensive portfolio data provided by a major life insurance company, and a machine learning model to generate mortality predictions. The random survival forest model is the state-of-the-art machine learning prediction model that can leverage the detailed health and mortality database available to life insurance companies. In particular, the risk model can be used to generate the changes in 10-year mortality predictions due to changes in health variables, for instance cholesterol and blood pressure levels, which provides a great laboratory to test the impact of the health engagement programs. Using detailed information on outstanding contracts, the resulting improvement in longevity can be translated into dollar values (see Figure 1 for an illustrative example). This benefit amount can be calculated for each policyholder with the outstanding contract information, so I can analyze the distribution of the benefit amounts, as well as the aggregate benefit for the insurer.

For the cost estimation, I borrow from the large health economics literature and medical research data, to fit a structural model that describes smoking behavior. Specifically, I present a rational addiction model in style of Becker and Murphy (1988) and Choo (2000) as a framework to investigate smoker behavior. The rational addiction models incorporate addictive behavior, such as smoking, in the rational-agent frameworks. In these models, the rational agent fully accounts for the possibilities of being addicted to a good (e.g. cigarettes), and maximizes her lifetime utility by choosing the optimal consumption and the addiction patterns. I rely on various health economics and medical studies (such as Volpp et al (2009), Jha et al (2013) and UHDHHS (2020)) to fit a version of the rational addiction model. One change I make is to relax full rationality by assuming that some agents systematically
underestimate the risks of smoking. The calibrated model fits various aspects of smoking behavior well, such as the mean spending on smoking, quit ratio, and the relative risks compared to non-smokers. Given the good fit, I use it as a framework to calculate model-implied financial incentives required to compensate for the lost pleasure of smoking, a cash payment conditional upon the successful quitting. The minimum financial incentives needed depend on wealth and addiction status, and are calculated for each smoker in the life insurer’s portfolio.

The results highlight the promise of smoking cessation programs. On the sub-sample of policyholders who were categorized as current smokers during the underwriting process, the average benefit amount is around $18,000 in 2018 USD. Even among this sub-sample of the population, the potential aggregate benefit from the success of smoking cessation program sums up to nearly $ 540 millions USD. The distribution of the benefit amounts is right-skewed with values higher than USD 50k. Following the decomposition in Equation (4), the average benefit comprises of the $2,485 in premium effect and $15,566 in face value effect. The benefit is gradually increasing from age 20 to 60 (Panel (b)), and it is the combined effect of the increase in the face amount (Panel (a)) and the increase in the benefit-per-dollar face value (Panel (c)). Between 20 and 40, the rising face amount explains the increasing benefits, and after 40, the rising benefit-per-dollar face value explains the increasing benefits.

The cost model implies the average financial incentive needed for the policyholders to quit smoking is about $3,200. Combining the benefit estimation and the cost estimation together with the baseline long-term abstinence rate of 30%, the implication of the smoking cessation program is that about 53% policyholders will successfully quit for one year given the incentives. The net benefit amount available to the life insurance company with this program is USD 87 millions.

However, one constraint in designing the health engagement program of this type is that participating in the program should also be incentive compatible to the policyholders. For example, upfront cash payment must not induce the policyholders to quickly turn around
to another life insurer given that they have an option to lapse to negotiate a better rate by being re-classified as a better risk. Smokers are known to have higher lapsation rates than non-smokers from the data (SOA 2019) and I reflect these higher lapsation rates in the contract valuation. Thus, I price in the higher market competition for smokers and their higher price elasticities. I additionally discuss an intuitive program design to ensure the incentive compatibility to a policyholder by requiring a minimum number of years since contract issuance, where the threshold depends on the age. Long-term policyholders have more financial incentives to stay with the same insurer, as their payment tends to be front-loaded while the benefit is back-loaded, so the previous premium payments operate as a natural source of collateral that prevents lapsation.

I use the smoking cessation program as the benchmark example, but there are a list of interesting health risk variables to which I apply this framework. The applications range from obesity to various chronic conditions (e.g. hypertensions, high cholesterol), to cancer and heart conditions. For each health risk variable, I measure the benefit to the life insurance company from improving that particular health condition, holding all other health characteristics constant. I find that there are indeed large potential benefits available to the life insurance company in designing and operating the health engagement programs. In fact, improving other health conditions such as obesity, hypertension and high cholesterol bring similar benefits to smoking cessation programs (even though the prevalence of these conditions are lower than the smoking), and the aggregate benefit from the health engagement program on these conditions amounts to 872 million USD.

The rest of paper is organized as follows. I review the existing literature in the rest of this section. Section 2 presents the models that will be used as the frameworks in understanding the benefit and the cost from operating health engagement programs. Section 3 introduces data used in actual measurement. I describe additional methodological details of the measurement process in Section 4, and present the results in Section 5. In Section 6, I discuss various incentive compatibility issues for successful implementation of the programs.
I explore other health engagement programs in Section 7 and present the aggregate benefit available from these programs. Section 8 concludes.

**Literature Review**

There is a burgeoning literature studying the insurance sector and its products from a finance perspective. Koijen and Van Nieuwerburgh (2020) show that the potential financial benefits accruing to life insurers can be used *ex ante* to fund the expensive medical treatments (immunotherapy) when the existing policyholders are diagnosed with cancer. I build on their insights, but I make three contributions relative to the paper. First, the topics I am studying (health engagement programs) require more proactive roles for the life insurance companies in improving policyholders’ health, as they need to ensure that the program is incentive compatible for the policyholders to participate. This is an additional real impact the insurance sector can make, yet to be fully explored. Second, I use a more precise benefit estimation exploiting a unique database of life insurance contracts at the policyholder level. The breadth and depth of the database allows me to measure the aggregate benefits more accurately, which is important in investigating the viability of the programs. Third, I build on the vast literature of health economics and medical literature to model and estimate the cost-side of changing health behaviors, especially the addiction behaviors.

There are extensive health economics and medical literature that I build on to estimate the costs of health behavior changes. Volpp et al (2009) and Halpern et al (2015) are the two influential papers to estimate the effects of providing financial rewards to incentivize smokers to quit. Volpp et al (2009) differs from the previous papers that study similar experiments mainly by recruiting more subjects and offering larger monetary incentives (see Cahill et al 2015 for the comprehensive review). The authors run a large-scale randomized control trial that pays $750 dollars of financial incentives for those who succeed in quitting for one year, compared to the control group who only received the free smoking cessation program.
They show that financial incentives are effective and the smoking cessation success rate nearly triples. Moreover, the effect lasts even in the follow-up interviews after the incentive payments. Halpern et al (2015) run another RCT, paying $800 dollars for one-year quitters, but they test various program designs to test behavioral hypotheses. They find that the reward-based approach is more effective than the deposit-based approach, as the program take-up rates were widely different between the two methods (90 percentage points vs. 14 percentage points) even though the deposit-based method was more effective conditional on the take-up. Individual versus group reward programs do not show much differences.

A large economics literature studies the dynamic consumer behavior with addiction. Stigler and Becker (1977) were first to point out that a stable utility function can explain addictive behaviors under the standard rational agent framework. Other papers formulate addiction as a solution to a dynamic optimization problem with different specifications of utility functions (Boyer (1978, 1983), Spinnnewyn (1981), Iannaccone(1984, 1986)), often building on the habit persistence literature (Pollak (1970, 1976), Ryder and Heal (1973)) in analyzing the relationships between the intertemporal choices. In their seminal paper, Becker and Murphy (1988) formulate the theory of rational addiction in a general setting, and provide a comprehensive discussion of the assumptions, the stability of the equilibria, and the implications of the model. At the core of a rational addiction model is the complementarity between the consumption choices across different periods ("adjacent complementarity"), which is modeled as the complementarity between the current consumption choice and the addiction stock level (history of the past consumption behavior) in their model. The rational agent factors in this complementarity in her optimization, which induces addictive consumption behaviors.\footnote{Note that there is a literature focusing on the time-inconsistent behaviors of the agents. Orphanides and Zervos (1995) and Suranovic et al (1999) model addiction while relaxing the perfect rationality assumptions, and Gruber and Köszegi (2001) and Machado and Sinha (2007) study different policy implications under the hyperbolic discounting agents.} Chaloupka (1991), Becker et al (1994), and Gruber and Köszegi (2001) test the rational addiction model to find empirical support for the forward-looking
nature of the agents. See Chaloupka and Warner (2000) for the comprehensive review of the literature. Choo (2000) presents a dynamic model of both smoking and cessation behavior under the rational addiction framework. He incorporates the health state variable of the agent to explain abrupt quitting (“cold turkey”) behavior, and calibrate the model using data from an external RCT. More recent IO and marketing literature use the dynamic discrete choice framework (Arcidiacono et al (2007), Chen et al (2009), Gordon and Sun (2015)) to structurally model smoker behavior. My modeling of smoking behaviors applies the rational addiction modeling in a lifecycle framework compared to the usual settings with an infinitely-lived agent, and I incorporate the extensive health economics and medical research to fit the model.

2 Model

To measure the net benefit available to the life insurance company by running health engagement programs, the benefit side and the cost side have to be separately measured. I discuss each component in the following sections, using the smoking cessation program as the primary example.

2.1 Financial Benefits to Life Insurers

I provide a simple economic framework of the potential financial benefits available to the life insurance company when a policyholder’s health status improves. Similar to Koijen and Van Nieuwerburgh (2020), I measure the financial benefits accruing to the life insurers from the existing life insurance contracts (policies, interchangeably), due to the longer expected lifespans.

Consider a life-insurance policyholder $i$ with the set of personal characteristics $x_i$ that affect the probability distribution of his survival, conditional on being alive today. It includes basic demographic variables, such as age (currently at age $\tau$), gender, and health information
such as being a smoker or not. I denote by \( \{ \pi_{\tau,t}(x_i) \}_{t=\tau+1,...,T} \) the conditional survival probability until the age \( t \), where the agent surely dies when she turns age \( T + 1 \). Note that this is equivalent to specifying the conditional survival probability from age \( t \) to \( t + 1 \), i.e., \( \{ \pi_{t,t+1}(x_i) \}_{t=\tau+1,...,T} \).

The dependence on the policyholder’s personal characteristics can be thought as the “pricing formula” of the life insurance company, given all the available information. The simplest form of pricing (mortality assessment) is just reading \( \pi_{t,t+1} \) from the national life tables\(^2\) which means that the life insurance company conditions only on the age and gender information in its pricing.

Two main terms specified in the existing life insurance contract are the annual premium \( (p_i) \) and the face value \( (F_i) \), or the death benefit. The timing assumption of this contract is as follows. When the agent enters a period as age \( t \), she first decides whether to lapse (abandon) the policy or not. Suppose that the lapse rate during age \( t \) is modeled as \( 1 - k_{t,t+1}(x_i) \) (that is, \( k_{t,t+1} \) is the conditional probability of “keeping” the policy during age \( t \)). If she has decided to keep the policy, she needs to pay the period premium \( p_i \) to the life insurance company. During the period \( t \), a random mortality shock hits and the policyholder will die with probability \( \mu_{t,t+1}(x_i) = 1 - \pi_{t,t+1}(x_i) \). In case the agent dies during the period \( t \), the life insurance company will pay the full face value of the insurance contract, \( F_i \), at the end of the period. At the beginning of the period when the agent \( i \) becomes age \( \tau \), the insurance company will price this contract on its book as (for notational convenience, the dependence

\[\begin{align*}
\text{The standard relationship between } \{ \pi_{\tau,t}(x_i) \} \text{ and } \{ \pi_{t,t+1}(x_i) \} \text{ holds, i.e., } \prod_{j=\tau}^{t-1} \pi_{j+1,j+1} = \pi_{\tau,t}. \text{ I use the same subscription notation for } k_{\tau,t} = \prod_{j=\tau}^{t-1} k_{j+1,j+1} \text{ later.}
\end{align*}\]

of $\pi$ and $k$ on $x_i$ is omitted): \[ V_i = \sum_{t=\tau}^{T} \frac{1}{R^{t-\tau}} \cdot k_{\tau,t} \cdot \left( \pi_{\tau,t} \cdot p_i - (\pi_{\tau,t} - \pi_{\tau,t+1}) \cdot \frac{F_i}{R} \right) \] \[ = \left( \sum_{t=\tau}^{T} \frac{k_{\tau,t} \pi_{\tau,t}}{R^{t-\tau}} \right) \cdot p_i + \left( \sum_{t=\tau}^{T} \frac{k_{\tau,t} \pi_{\tau,t} (\pi_{t,t+1} - 1)}{R^{t-\tau+1}} \right) \cdot F_i \] \[ (1) \]

The first term indicates the discounted value (with the discount rate $R^{-1} < 1$) of the expected future premium that the life insurer will receive. The per-period premium is constant at $p_i$, while the length of the actual payment crucially depends on the survival probability function $\pi_{\tau,t}$. The second term is the discounted value of the future death benefit payout when the policyholder dies. The benefit is paid only once, so the timing of the death is crucial in valuing this payment leg.

Now suppose that the insurance company has successfully changed one health variable (smoking) variable in $x_i$, so the personal variables become $\bar{x}_i$. The insurance company will use the same pricing function $\{ \pi_{t,t+1}, k_{t,t+1} \}$ to value the contract, so it can be written as: \[ \bar{V}_i = \left( \sum_{t=\tau}^{T} \frac{k_{\tau,t}(\bar{x}_i) \pi_{\tau,t}(\bar{x}_i)}{R^{t-\tau}} \right) \cdot p_i + \left( \sum_{t=\tau}^{T} \frac{k_{\tau,t}(\bar{x}_i) \pi_{\tau,t}(\bar{x}_i) (\pi_{t,t+1}(\bar{x}_i) - 1)}{R^{t-\tau+1}} \right) \cdot F_i \] \[ = \left( \sum_{t=\tau}^{T} \frac{\bar{k}_{\tau,t} \bar{\pi}_{\tau,t}}{R^{t-\tau}} \right) \cdot p_i + \left( \sum_{t=\tau}^{T} \frac{\bar{k}_{\tau,t} \bar{\pi}_{\tau,t} (\bar{\pi}_{t,t+1} - 1)}{R^{t-\tau+1}} \right) \cdot F_i \] \[ (3) \]

where I use $\bar{k}_{\tau,t} = k_{\tau,t}(\bar{x}_i)$ and $\bar{\pi}_{\tau,t} = \pi_{\tau,t}(\bar{x}_i)$ for notational convenience.

The impact of the change in a health variable on the lapsing behavior is an empirical question. I discuss the detailed issues with lapsation in Section 6. For now, I assume that $\bar{k}_{\tau,t} = \bar{k}_{\tau,t}, \forall t$ (going forward I will refer to this assumption as the *independence lapsation assumption*). Under the independence lapsation assumption, the financial benefit accrued
to the life insurance company can be written as:

\[
\tilde{V}_i - V_i = \left( \sum_{t=\tau}^{T} \frac{k_{\tau,t}}{R^{t-\tau}} \cdot (\tilde{\pi}_{\tau,t} - \pi_{\tau,t}) \right) \cdot p_i + \left( \sum_{t=\tau}^{T} \frac{k_{\tau,t}}{R^{t-\tau+1}} \cdot (\tilde{\pi}_{\tau,t}(\tilde{\pi}_{t,t+1} - 1) - \pi_{\tau,t}(\pi_{t,t+1} - 1)) \right) \cdot F_i
\]

(4)

Since the change from \(x_i\) to \(\tilde{x}_i\) is health-improving, the conditional survival probability in each period weakly goes up, that is, \(\tilde{\pi}_{j,j+1} \geq \pi_{j,j+1}, \forall j = \tau, \ldots, T\). I additionally assume that health is strictly improving in at least one time period, i.e., \(\tilde{\pi}_{j,j+1} > \pi_{j,j+1}\) for some \(j\).

The two terms in equation (4) can be interpreted as the premium effect and the face value effect. I show that the valuation of the contract indeed increases with improved health of the policyholder (hence increased conditional survival probability). Start with the premium effect term. Since it immediately follows from the above health-improving assumption that \(\tilde{\pi}_{\tau,t} \geq \pi_{\tau,t}\) for all \(t\), every term in the premium effect is non-negative, with at least one term being positive. As a result, it is clear that the entire premium effect is positive. The economic intuition is that as the policyholder is expected to live longer due to improved health, the expected future premium income also increases.

On the other hand, the sign of the face value effect is less clear because for some \(t\), the difference term \(((\tilde{\pi}_{\tau,t}(\tilde{\pi}_{t,t+1} - 1) - \pi_{\tau,t}(\pi_{t,t+1} - 1))\) can become negative. The following proposition shows that under the independent lapsation assumption described above, we can show that the face value effect is also positive.

**Proposition 1.** Under the independent lapsation assumption, i.e., when \(k_{j,j+1} = \tilde{k}_{j,j+1}, \forall j\), the face value effect in equation (4) is positive.

**Proof.** See Appendix.

The intuition for the positive face-value effect is also straightforward, as illustrated in Figure 1. Note that the face value payout is the liability leg to the life insurance company in
this contract. When the policyholder becomes healthier and is expected to live longer, the timing of the the death benefit payout is delayed to the future. Since $F_i$ is typically large, the face value effect is still large after taking the NPV.

**Proposition 2.** Under the independent lapsation assumption, the valuation of the contract increases when $x_i$ to $x_{i'}$ is health-improving, i.e., $V_i > V_{i'}$.

**Proof.** In equation (4), the change in the value can be decomposed into the premium effect and the face value effect. Both terms are positive according to the previous discussion in the text and Proposition 1. Hence, the valuation of the contract to the life insurer strictly increases.

In this section, I have presented the framework that will be used to calculate the financial benefit accruing to the life insurance, when the health of the policyholder improves. The benefit can be decomposed into two different sources, the longer premium cashflow streams in the future, and the delayed payout of the death benefits. Two steps are crucial to actually calculate the benefit equation (4). First, from the life insurance company’s pre-existing portfolio, I need to observe the distribution of $(x_i, p_i, F_i)$. Second, using the mortality model, I need to measure the survival probability, $\pi_{t,t+1}(x_i)$. Especially, one of the covariates in the pricing function must be smoking (or other variable of interest in running a health engagement program) so that we can measure both $\pi_{t,t+1}$ and $\tilde{\pi}_{t,t+1}$. The difference, which is the marginal effect from the health improvement, are the key inputs in the benefit equation (4). The details of these steps will be further discussed in Section 4.1.

### 2.2 Cost Model

I build on the rational addiction model of Becker and Murphy (1988) and Choo (2000) to model the smoking and cessation decisions, and then use the model to estimate the cost of behavior changes.

---

4The average size of a purchased life insurance policy was 163 thousands in 2017 (ACLI 2018, p.67).
There are two goods in the economy, cigarettes and regular goods. Unlike the regular good, smoking is addictive and the past consumptions of cigarettes affect the current (and the future) period’s marginal utility through the “addiction stock” variable, $a_t$. This unobservable state variable summarizes the past smoking behavior of the agent. In the beginning of period $t$, the agent observes her state variable, the addiction stock ($a_t$) and then chooses the optimal consumption of both goods: cigarettes for smoking ($s_t$) and regular good ($c_t$). All the prices are normalized using the price of regular goods, and denote the relative price of cigarettes by $p_t$. The agent earns $y_t$ amount of income (also normalized at the price of regular goods), and allocates the spending between two goods.

The agent $i$ enters the economy at time zero, as age zero (this corresponds to age 20 in the data, which is the average legal smoking age) and for sure die at age $T + 1$ (I use 1-year period and assume $T = 80$, i.e. all the agents die when they turn 100). The agent $i$ maximizes her lifetime utility,

$$V_{i,0} = \max_{\{c_t, s_t\}_{t=0}^{T}} \sum_{t=0}^{T} \beta^t \pi_{0,t}(x_i) E[u(c_t, s_t, a_t; x_i)]$$

subject to a simple budget constraint

$$w_t + y_t(x_i) \geq c_t + p_t s_t$$

$$w_{t+1} = R(w_t + y_t - (c_t + p_t s_t))$$

and in consideration of his survival projection, $\pi_{0,t}$, which will be modeled below.

Rewriting the consumer optimization problem in the Bellman equation form (I omit the
dependence on the personal characteristics $x_i$ to simplify notation)

$$V_t(a_t, w_t) = \text{Max}_{c_t, s_t} u(c_t, s_t, a_t) + \beta \pi_t(a_t) EV_{t+1}(a_{t+1}, w_{t+1})$$  \hspace{1cm} (7)

subject to: $w_{t+1} \geq 0,$

$$w_{t+1} = R(w_t + y_t - (c_t + p_t s_t))$$  \hspace{1cm} (8)

$$a_{t+1} = \begin{cases} \min(a_t + s, \bar{a}), & \text{w.p. } q_t \\ \min((1 - \delta)a_t + s, \bar{a}), & \text{w.p. } 1 - q_t \end{cases}$$  \hspace{1cm} (9)

The per-period utility function $u$ depends not only on the consumption levels of both goods $(c_t, s_t)$, but the level of the addiction stock $(a_t)$, following the rational addiction literature that started with Sigler and Becker (1977) and more clearly formalized in Becker and Murphy (1988). I assume the standard properties of the utility function, $u_c > 0$, $u_{cc} < 0$, $u_s > 0$, $u_{ss} < 0$, $u_{sa} > 0$. The last condition corresponds to the "adjacent complementarity" condition in Becker and Murphy (1988), which captures the reinforcement effect of engaging into addictive consumption (smoking; $s_t$), which is at the core of the addiction modeling. In short, the higher the addiction level, the higher is the marginal utility drawn from additional smoking. In addition, $u(0, 0, a_t) = b > 0$ is the crucial assumption for modeling the value-of-life, following the extensive discussion presented in Hall and Jones (2007).

To be more specific, I use the following utility function as the baseline specification:

$$u(c, s, a) = b + \frac{c^{1-\gamma}}{1-\gamma} + \alpha \ln(1 + \sigma s) \cdot f(a; \theta)$$  \hspace{1cm} (10)

where $f \geq 1$ is an increasing function of the habit stock parameterized by $\theta$, and $\gamma > 1$ and $b, \alpha, \sigma > 0$ to be consistent with the discussions above. Note that the overall structure of the utility function combines the utility specification used in Hall and Jones (2007) with the

---

5Substituting the Equation (10) of Hall and Jones (2007) with $\sigma = 1$ yields the equation (10) except for the addiction stock adjustment term.
utility specification from Choo (2000) for the addictive goods. For the choice of \( f \), instead of the simple linear function in Choo (2000), I use \( f(a; \theta) = 1 + \min(a, \theta) \). This modification ensures that the cigarette consumption does not vary too much over the lifecycle, which is observed in the actual data (see Figure A4 in Appendix to see that the vast majority of daily cigarette smoking is under 2 packs a day, regardless of the age or smoking history). Note that the specification satisfies the three hallmarks of the rational addition modeling.

Modeling addictive behavior should encompass three well-documented properties of addiction: (1) reinforcement effect, which means the higher addiction status, the higher the marginal utility of the addictive consumption, (2) tolerance, which means the higher the addiction status, the lower the utility given the same addictive consumption in the current period, and (3) withdrawal, which means quitting has negative psychological effect to the agent. Even though (1) and (3) are obtained instantaneously from the current specification, the tolerance property may not seem obvious, given \( u_a > 0 \). However, unlike the traditional rational addiction modeling with infinitely-lived agents, Choo (2000) explicitly models the mortality hikes coming from the addiction, hence the tolerance effect is modeled through the dependence of the transitional probability (per-period mortality) on the addiction level, i.e. through the equation (11).

The addiction stock follows a modified depreciation equation specified in Equation (9), with the random depreciation rate between 1 and \( \delta \), where the smoking in current period strengthens the addiction. I parametrize \( q_t = \exp(-q_x \cdot t) \) to match the quickly declining relapse rates across age groups, which is widely shown in many medical research. The fact that \( q_t \) exponentially declines over age implies that younger populations are more vulnerable to addictive substances and hence build the addiction stock at a faster pace and experience more withdrawal effects. I additionally assume that \( a \) stays on the fixed support, \( 0 \leq a \leq \bar{a} \), similar to the set up in Choo (2000). The agent also cares about the health consequence of smoking, as shown in the Bellman equation with the term \( \pi_t \) that specifies the one-period survival probability. I model the mortality function (one minus the survival probability) as
the affine function of the addiction level $a$ and age $t$ (and gender $g$),

$$
\mu_t(a) = 1 - \pi_t(a) = \exp(\kappa_t \cdot a + \gamma_0 + \gamma_1 \cdot t) \quad (11)
$$

where the parameters $\gamma_0, \gamma_1$ are separately estimated for each gender. To simplify notation, I omit the dependence on $g$ going forward.

There is no explicit modeling of the health status in the model. However, we can consider the negate of the term in (11), $-\kappa_t \cdot a - \gamma_0 - \gamma_1 \cdot t$, as the implicit health status. Higher health status implies lower mortality. We can rewrite the mortality $\mu$ of the agent $i$ as:

$$
\mu_t(a) = \exp(\kappa \cdot a_t + \gamma_0 + \gamma_1 \cdot t)
$$

where $\bar{\mu}_t = \exp(\gamma_0 + \gamma_1 \cdot t)$ can be considered as the baseline mortality of the agent when $a = 0$, i.e., never-smokers.\(^6\) The history of smoking, summarized in the state variable $a_t$, scales the current-period mortality by age-dependent factor $\kappa_t$, relative to the never-smokers.

\section*{2.2.1 Heterogeneity in Risk Perception}

With the extensive theoretical discussions and the empirical tests developed on Becker and Murphy (1988), the rational addiction model is widely accepted as the framework to describe the addictive behaviors. However, when I combine the rational addiction model with the lifecycle model framework, there is one puzzling fact. Under a rational lifecycle model, the agent has to reduce the smoking consumption rapidly later in her life, because the mortality due to smoking becomes too costly, given the high baseline mortality. Anticipating this, the agent in the model usually starts reducing smoking from late 40s, quits during 50s to reach

\(^6\)Note that such parametrization a simpler version of Gompertz-Makeham function, which is regarded as a good approximation of the population mortality function. Koijen and Van Nieuwerburgh (2020) also uses the approximation.
the low optimal addiction level in her 60s. However, the actual data in panel (a) of Figure A4 indicates that the smokers reach their peak smoking in their 60s, and reduces the amount of smoking only slightly afterwards.\footnote{Note that this puzzle is essentially the same one depicted in Viscusi and Hersch (2008), where the authors find that the VSL of smokers do not decline in age, which is inconsistent with the usual implication from the lifecycle models, for example, in the Figure 3 of Shepard and ZeckHauser (1984).}

To correctly capture this lifecycle pattern in the data, I make a modification to the model to partially relax the perfect rationality. Instead of the perfect risk prediction of $\kappa$, I assume that some agents systematically underestimate the relative risk of smoking (for this type, I assume $\kappa_{UE} \geq 0$ instead of $\kappa$) with the probability $p_{UE}$. Two pieces of evidence support this modeling choice. First, Figure A6 plots the survey results from Krosnick et al (2017) where the authors directly ask smokers and quitters about their relative risk perception. Compared to the actual relative mortality risks from lung cancer, more than 50% of smokers massively underestimate the relative risk. In fact, about 15% of the respondents answer that the risks of smokers are equal (or even smaller) compared to those of non-smokers. Given that the lung cancer is the disease that is most widely publicized to be caused by smoking, this massive underestimation in relative risk is surprising, but supports my assumption of underestimation in overall mortality.\footnote{On the other hand, Viscusi (1990) uses similar survey results to conclude that the smokers overestimate the risk of lung cancer. However, the actual wording in the questionnaire is crude – it does not provide any context such as age or smoking status (e.g. for former smokers), and it asks about the absolute risks. The questions and the preceding contexts are much carefully designed in Krosnick et al (2017).} Second, Weinstein et al (2005) show that agents display over-confidence in evaluating their own health risks, which in turn affects the smoking and cessation decisions. The effect of overconfidence is effectively the same as the underestimation of the smoking relative risks, and so it provides more support for the modeling choice of heterogeneous risk perception.
3 Data

3.1 Insurance Database

The main dataset is coming from an anonymous life insurance company (referred to the Life Insurance Company, or simply as the life insurer, in this paper). The database includes the individual-level insurance holdings data, and their detailed health information when they were applying for life insurance products. I have access to the de-identified version of this proprietary database,\(^9\) and it provides a valuable link between the policyholders’ health information to their insurance contracts. From the joint distributions of the actual insurance contracts and the policy holder covariates (demographics and health information), I calculate the potential financial benefit from running a health engagement program, which I further describe in Section 4.

Table 2 describes the summary statistics of the database. The insurance contracts table and the health information table separately exist, and my analysis is limited to the intersection of the two tables, for those who I can access both the contract information and the health information. I additionally limit my analysis only to the whole life and the term life products, while universal life or variable life insurance are excluded. For these reasons, the result I present in this paper only provides a lower bound available to the life insurance companies. I later discuss how to extrapolate these values to the population, and the potential limits existing in the database.

3.2 LifeScore Labs\(^{\text{SM}}\) Risk Models

The LifeScore Labs\(^{10}\) helps life insurance carriers to simplify the underwriting process by assessing the insurance risks more systematically. The company achieves this goal by de-

---

\(^9\)Given the sensitive nature of the information, an outside vendor encrypted various columns to systematically de-identify the database, to address potential privacy concerns.

\(^{10}\)https://www.lifescorelabs.com
veloping the risk scoring models that utilize the ever-growing amount of consumer data available to the carriers. LifeScore Med360SM, one of the main products of the company, predicts the mortality risks of policyholders based on the standardized application data, mainly the health variables. The mortality risk prediction is based on a random survival forest model (Ishwaran et al 2008, Maier et al 2019), which is one of the latest machine learning techniques in survival data analysis.

I use a version of the LifeScore Labs model, and apply it to the database of the Life Insurance Company. This allows me to retrieve the precise mortality risk predictions based on high-dimensional health information available in the database.

### 3.3 National Adult Tobacco Survey

National Adult Tobacco Survey\(^{11}\) (NATS) is jointly conducted by the U.S. Centers for Disease Control and Prevention (CDC) and the U.S. Food and Drug Administration (FDA) to survey the state of the smoking behaviors among the adult population in the U.S. Given its large sample size (over 75,000 interviews conducted over the telephone calls) it is nationally representative of the non-institutionalized adult (18 years or older) population. The survey questions cover the current smoking behaviors, the history of cigarette usage, addiction status (cravings), cessation attempts, purchasing behaviors, in addition to some demographic information variables.

The survey is used to calculate the widely-cited national smoking prevalence rate, which is currently 14%. See Appendix figures to see the various descriptive statistics among the smokers calculated from 2013-14 survey.

4 Methodologies

4.1 Benefit Measurement

The theoretical framework to measure the financial benefits to the life insurer is described in section 2.1. I utilize the data from the life insurance company, and apply the state-of-the-art machine learning techniques to measure the financial benefits that they can use to incentivize the policyholders to quit smoking. The first step is to estimate the extended lifespan due to quitting smoking, which comes from the difference $\pi_{\tau,t}(x_i)$ and $\pi_{\tau,t}(\tilde{x}_i)$ in Section 2.1. The empirical challenge comes from the potential interactions between the smoking decision (and history) and other personal covariates, such as age, gender, income and education level, and other health status variables. The health effects of smoking is known to be complicated in nature, and modeling the relationship using a simple linear-based model is insufficient at best.

The recent development in machine learning techniques provides an effective and robust way to control these various covariates and establish the underlying relationship. Especially, I use the fitted random survival forest model (Ishwaran et al 2008, Maier et al 2019) results of the LifeScore Labs model to effectively control for all the nonlinear variables and the interactions among the variables. Even though the detailed fitting is proprietary to the insurance company, the basic process of fitting the model to the big data is similar to other machine learning models. Given the initial heath information (lab test results required to anyone who wants to apply for the life insurance underwriting) at the application stage, coupled with the mortality events observations of the following years, the model is fitted by finding the “best predictor” of the mortality events. Figure 2 describes the overall process of generating and fitting the random survival forest model.

For the scope of this project, I have utilized the fitted version of the model, which produces the cumulative 10-year mortality predictions based on the health information provided with
the initial applications. Specifically, for each policyholder, I find the baseline mortality and survival curves to match the initial model prediction. When a health covariate gets improved (e.g. a smoker becomes a quitter, even though the smoking history remains), the fitted model updates the 10-year mortality predictions reflecting the improvement in health. This effectively generates \( \tilde{\pi}_{0,10}, \pi_{0,10} \) for each policyholder. To accurately value the contracts, the 10-year survival prediction is not enough, and the lifetime mortality curve must be predicted. Therefore, I combine the machine learning results with the mortality curves from the 2015 Valuation Basic Table published by the Society of Actuaries (SOA 2018). I generate the individual lifetime mortality curves in a two-step process, to be consistent with the general contract valuation practices in the industry. To be more specific, for each \((g,s)\) demographic group for (gender, smoking status) pairs (for example male-smokers), I first estimate the scale factor for the mortality curves from the 2015 Valuation Basic Table to match the average 10-year mortality of the demographic group,

\[
c_{g,s} = \arg\min_c \left| \frac{1}{|N_{g,s}|} \sum_{p \in N_{g,s}} \left( \prod_{j=0}^{9} (1 - \lambda_{p_a+j,p_a+j+1}^{g,s} \cdot c) - \pi_{0,10}(p) \right) \right|
\]

where \(N_{g,s}\) denotes the set of policyholders in the demographic group \((g,s)\), \(p_a\) denotes the age of the policyholder \(p\), and \(\lambda_{p_a+j,p_a+j+1}^{g,s}\) numbers are directly read from the 2015 VBT table,\(^{12}\) and \(\pi_{0,10}(p)\) is the predicted mortality from the model for the policyholder \(p\). Note that this process mimics the common valuation process used in the industry to match the average Actual-to-Estimated ratio, of the pooled population, at the demographic group level.

As the second step, I further account for the policyholder-level variations in risks by finding the within-group scaling factor \(c_p\) that matches the 10-year mortality predictions. Suppose the policyholder \(p\) has gender \(g\), age \(a\), and has the 10-year survival prediction \(\pi_{0,10}(p)\) as a smoker, and \(\tilde{\pi}_{0,10}(p)\) assuming she quits. Then I can find the pre-quitting and

\(^{12}\)Between the two sets (ANB/ALB) mortality curves provided in the Valuation Basic Table, I use the ANB (At-the-Nearest-Birthday) as the baseline, as ALB table is derived from the ALB table. Switching to ANB curves does not materially impact the analysis.
post-quitting scaling factors \( c_p \) and \( \tilde{c}_p \) that satisfy the following equations

\[
\pi_{0,10}(p) = \prod_{j=0}^{9} (1 - \lambda_{a+j,a+j+1}^{\text{ smoker}} \cdot c_{g,\text{ smoker}} \cdot c_p)
\]

\[
\tilde{\pi}_{0,10}(p) = \prod_{j=0}^{9} (1 - \lambda_{a+j,a+j+1}^{\text{ non-smoker}} \cdot c_{g,\text{ non-smoker}} \cdot \tilde{c}_p)
\]

Solving for \( c_p \) and \( \tilde{c}_p \) effectively generates the lifetime mortality predictions of \( \{\tilde{\pi}_{j,j+1}, \pi_{j,j+1}\} \) for the policyholder \( p \), and hence we can use the Equation (4) to calculate the dollar value of potential benefit when the policyholder gets the improved health from the health status engagement programs.

Equation (4) has the lapsation terms \( k_{\tau,t} \), which is the fraction of policyholders who forego the premium payments to cancel the life insurance. I read the lapsation numbers from the joint study of SOA and LIMRA (SOA 2019), conditional on the product type (term life vs. whole life) and the smoking status (tobacco vs. non-tobacco). I consider the term structures of the lapsation behavior, so that the lapsation is higher in the early years and stabilizes over time. See Figure A2 for the term structures of the lapsation used in valuation. I will discuss the effect of these lapsation rates on the benefit calculation in a great detail, in Section 6.1.

### 4.2 Cost Measurement

The cost of changing the health behavior is measured with the framework discussed in Section 2.2. Given the three state variables in the model – age \( t \), addiction level \( a \), and wealth \( w \) – the model provides the optimal value function \( V_t(a,w) \) that is attained by the rational agent, by choosing optimal amount of consumption \( c \) and smoking \( s \).

Consider a similar, but more constrained dynamic optimization problem of the same rational agent, who just decided to quit smoking for the current period (i.e. for one year).\(^{13}\)

\(^{13}\)The assumption here is that the insurance company can assure that the agent is indeed smoke-free. There exist cost-effective lab tests that can be used for this purpose. Note that Volpp et al (2009) and Halpern et al (2015) utilize such lab tests guarantee the compliance of the subjects with the incentive program.
In the model, this is equivalent to set all the current and future smoking variable \( s_t \) as zero, thus the addiction naturally decays at the rate \( 1 - \delta \) with probability \( 1 - q_t \). Given the set of state variables \((a, t, w)\), the agent only chooses the optimal consumption (and the saving) levels by solving the following dynamic problem,

\[
\tilde{V}_t(a_t, w_t) = \max_{c_t} \left( u(c_t, 0, a_t; x_t) + \beta \pi_t(a_t) EV_{t+1}(a_{t+1}, w_{t+1}) \right)
\]

subject to: \( w_{t+1} \geq R(w_t + y_t - c_t) \)

\[
a_{t+1} = \begin{cases} 
  a_t, & w.p. \ q_t \\
  (1 - \delta) a_t, & w.p. \ 1 - q_t
\end{cases}
\]

Since this is a constrained version of the optimization problem in (7), the value function is lower than the original value function by construction, i.e., \( V_t(a, w) \geq \tilde{V}_t(a, w), \forall a, w, t \). Now, the insurance company can compensate for the loss of utility from smoking cessation (the agent loses all the pleasure of smoking) by paying \( \tilde{w} \) amount of money that satisfies:

\[
\tilde{V}_t(a_t, w_t + \tilde{w}(a_t, w_t, t)) = V_t(a_t, w_t)
\]

By definition, \( \tilde{w} \) is the lower bound of the successful smoking cessation incentives, as the agent will prefer to “switch” and attain \( \tilde{V} \) utility with the additional wealth. Note that this compensation amount depends on all the state variables, namely \( a, t, \) and \( w \). The monetary required to incentivize the agent to quit crucially depends on her age, addiction stats (smoking history), and the wealth level. The heterogeneity in \( \tilde{w} \) implied by the model will be fully presented in Section 5.

4.2.1 Model Calibration

Table 3 presents the summary of the calibrated model parameters. I use the risk aversion parameter \( \gamma = 2 \) for the consumption good, which is standard in finance and economics.
\[ R = 1.03 \text{ and } \beta = 0.97 \] is also standard, matching the long-term average of the interest rates. The coefficient \( b = 1.2 \) is chosen to match the value of life\(^{14}\) of Age 40 male individual at 4 millions U.S. dollars, in a similar spirit of Hall and Jones (2007). The coefficients \( \alpha = 0.03 \text{ and } \sigma = 10 \) are chosen to match the annual spending on smoking of USD 3k a year, which is the average number of cigarette smoking per day (16), at the 10 dollars a pack price, while keeping the value of life at the target. \( \theta = 1.0 \) is chosen to limit the variance in the lifetime smoking consumption, while keeping the average level at the target level. The lifecycle income profile \( y_t \) at different age is adopted from the website DQYDJ (https://dqydj.com/income-percentile-by-age-calculator/), which is the average calculated from the original data source Census 2018 Annual Society and Economic Supplement survey data, and I apply the cubic interpolation. For the age above 65, I set the post-retirement income at 80\% of the average income between 20 and 65, which implies a 66\% income replacement ratio. I assume \( p = 1 \) to give the same relative price of consumption good and cigarettes over time, and abstract from the price expectation.\(^{15}\)

The decay rate of smoking effect \( \delta \) and the relative mortality scaling factors \( \kappa, \kappa_{UE} \) are calibrated from the medical research reviews presented in the Surgeon General Report on Smoking Cessation (USDHHS 2020). Regarding \( \delta \), the 1990 Surgeon general report is again cited for the famous results “the decline in risk of death compared with continuing smokers begins shortly after quitting and continues for at least 10 to 15 years. After 10 to 15 years of abstinence, the risk of all-cause mortality returns nearly to that of persons who never smoked (UHDHHS 2020, p.461).” I pick \( \delta = 0.3 \), which gives a reasonable rate of decay after 10 to 15 years. Moreover, the 2020 Surgeon General Report presents the up-to-date evidence gathered from the large-scale CPS II (Cancer Prevention Study II), where

---

\(^{14}\)The estimation of value of life spans a wide range, anywhere from 2 million to 10 million dollars. Please refer to Hall and Jones for the extensive discussions and the rationale. I abstract from this issue in the current paper.

\(^{15}\)Quick Analysis from BLS finds that this is not exactly true. Whether smokers correctly forecast the increase in cigarette prices is an empirical question. I will discuss this issue in the robustness section.
the relative mortality risks based on the smoking status and age \(^{16}\) are presented (2020 USDHHS Table 5.5a and Table 5.5b). To match these relative risk mortality by age, \(\kappa\) is jointly estimated with the parameters that scale the average spending on smoking (which in turn affects the level of addiction from the model). \(p_{UE}\) is set to be 0.4, which is the fraction of quitters among the ever-smokers (quit ratio) in 50s. \(\kappa_{UE}\) is calibrated to match the expected years gained in life from quitting at different ages, presented in Jha et al (2013).\(^{17}\)

Figure 5 presents the life cycle model solution simulation paths, solved with the calibrated parameters in Table 3. Three panels in the left column is for the agent with correct \(\kappa\), and the right panels are for the agent with \(\kappa_{UE}\). The first row shows that the life cycle smoking patterns implied by the model. The average spending on smoking is 0.18 in the model, compared to the 0.15 target from the data. The second row shows the random paths of addiction evolution. Because of the \(q_t\) randomness assumptions, addiction increases at younger age, and gradually decline as the agent gets older. The last row presents the simulated mortality curves of the agents, where the red line is the baseline mortality of never-smokers.

5 Results

5.1 Benefits

Figure 3 and Table 5 present the distribution of the benefit amounts of the policyholders, when they successfully quit smoking. The average benefit is 18,049, in 2018 USD, which is sizable compared to the average face value of USD 540k. The benefit is heavily right-skewed, where the median is about USD 9,100 and the inter-quartile range is from USD

\(^{16}\)To be precise, the duration of smoking instead of the age must be considered. However, “nearly 9 out of 10 smokers first try smoking by 18 years of age, with 99% of smokers doing so by age 26” (2020 USDHHS p.540) guarantees that the age generates a reasonable estimate of past smoking duration.

\(^{17}\)Note that Jha et al (2013) is peculiar in its definition of quitters. The authors use a strict condition of past 5-year full abstinence, which excludes the effect of relapse. I calculate the year gains in the same way to match the paper.
4,900 to 18,600. Using the decomposition in Equation(4), the benefit can be decomposed to USD 2,485 of the premium effect and USD 15,566 of the face value effect.

The benefit numbers can be further conditioned to understand the heterogeneity. Figure 4 shows presents the statistics in 5-year age groups, where each policyholder age is rounded to the nearest 5-year age groups. The first panel shows that the average face amount gradually increases with age until early 40s, and then stays flat. The second panel shows that the average benefits from successful smoking cessation gradually increases with age. The third panel shows the normalized benefit, as a fraction of the face value amount. Even though the second panel shows gradual increase in the benefit amount, this is in fact a combined effect of the first and the third panel. Appendix figure A1 presents the full distribution of the benefit amounts by age groups. The policyholders are highly concentrated in ages between 25 and 45, which is the typical time to start buying life insurance policies due to changes in family structure and increased income. At each age group, the benefit amounts are all highly skewed to the right.

Table 4 confirms the intuitions with a regression. I regress the log of the benefit amount for on log of the face value, policyholder age and then the product type indicator. Other conditions being equal, the policy size linearly scales the benefit amounts, so the coefficient close to one is not surprising. This alone explains more than the 60% of the variations in the log benefit amounts. After controlling for the size, I add the policyholder age and the whole life variables, which are the two most important variables in explaining the size-controlled benefit variations. On average, 10-year increase in the policyholder age increases the benefit amount by about 14%. This confirms the intuition from Figure 4. Finally, compared to the term life product, being a whole life product increases the benefit by 40% on average. This confirms the intuition presented in Figure 1c. Intuitively the benefit is the weighted present value of the gaps between blue and orange cures. When the maturity of the product increases (the median length of the term life is 20 years in the database), the benefit also tends to increase.
5.2 Costs and the Net Benefits

Using the Equation (13), I calculate the model-implied incentives requirement needed for one-year smoking cessation decision for the agent. At different age, addiction and wealth level, different incentive is required to change the health behavior of the agents. Using the simulated life cycle addiction and wealth paths from the model solutions, I calculate the required incentive amounts and then take the average to get the estimation of the mean incentives. However, carefully accounting for the agent risk perception type is important, because the two types of agents behave very differently. For example, the underestimation type consists of roughly 40% of the observed current smokers at young ages. However, if I observe a current smoker with age greater than 60, I assume that the smoker has $\kappa_{UE}$ (the underestimation type) in the Bayesian sense. Using the model solutions and the simulations of the solution paths, I explore various hypothetical program designs.

An important component in measuring the incentives is the possibilities of relapse, or the quitters’ tendency to return to the smoking habit. Nicotine is a highly addictive substance and relapse patterns are widely reported in medical research, especially during in the short-term. It is also reported that young population are much more likely to relapse. There are abundant empirical evidence that the cumulative long-term abstinence rate becomes stable after the first 3 to 5 years, conditional on being abstinent for a full year. Therefore, I calculate the simulation-implied five-year abstinence rate, conditional on the success of the 1-year incentive programs, and will use the rate for the lower bound for the calculation. For example, Figure 7c second bar charts show that the 5-year model implied abstinence rate is nearly 40%. The actual number from the Volpp et al (2009) is slightly higher (9.4% out of 14.7%), but this is because the RCT only tracks the subjects up to 6-month follow ups, and

18 Among many research results on the smokers’ relapse, see García-Rodríguez et al (2013), Hughes et al (2008), Krall et al (2002) and Hwakins et al (2010) for wide range of evidence of convergence after 3-5 years. The converging long-term abstinence rate is shown to be between 40% and 80% depending on the subject demographics and study settings. In medical research, 1-year abstinence is often referred as long-term abstinence, and it predicts the success of long-term abstinence.
the model is matching the relapse pattern described in the literature quite well.

Utilizing all these components, I analyze the net benefit from the smoking cessation program in two steps. I first calculate the conditional probability of a current smoker with age \( a \) being the underestimation type (\( \kappa = \kappa_{UE} \)). This is calculated from the fraction of observed smokers in the simulation, adjusted for \( p_{UE} \) and the evolving survival probability. Since the type of each policyholder cannot be precisely observed, this is the best estimate of the costs for one-year smoking abstinence, as shown in Figure 6, Panel (c). The next step is to discount the benefit numbers with proper long-term abstinence rates, conditional on successful one-year abstinence. Based on the previous discussion on relapse, I use the baseline assumption of the long-term conditional abstinence rate (conditional on being abstinent for one year) of 30%, which is in the lower end of the range surveyed in the literature. I also exclude the small group of policyholders under age 25 for this exercise. After pricing in the relapse probabilities, 48% of the sample population has the higher benefit estimation than the cost. Figure 8 shows the joint distribution filtered for the policies that have higher potential benefits than the costs. The average benefit of these policies are USD 30k, and the average cost is around USD 2,700, which results in the aggregate net benefit coming from operating is around USD 87 millions.

6 Discussion on the Program Implementation

6.1 Lapsation and Incentive Compatibility

One important component in life insurance contract valuation is lapsation, or the policyholders’ option to forgo the premium payments to void the contract. Similar to other financial contracts with optionalities, such as mortgage loans, the projected lapsation behavior has to be priced in valuing the contracts, as shown in Equation (2). In the life insurance markets, lapsation can come from three main sources. First, many group policies, i.e. the life
insurance provided in relation with employment, are automatically terminated when the policyholder changes or loses the job, or retires. These natural lapsations are peculiarities of the U.S life insurance market. Second, during the economic downturns, periodic premium payments become burdensome to the policyholders, thereby leading to higher lapsation rates. Industry-level aggregate data (SOA 2019) clearly shows this trend during the 2007-2009 financial crisis. Third, lapsation shows the level of competition in the market. Life insurance policy markets are highly competitive, and the products are standardized across the providers. Therefore, the price competition is fierce and the policyholders can switch to other life insurers, especially during the early holding years before building the policy values.

With the possibility of lapsation, the smoking cessation program cannot share the full benefit amount with the policyholder. Receiving the present value of the full benefit will incentivize the smokers to shop around for another life insurance policy, now that her health status has improved from smoker to former smoker. One program design choice I make to address this issue is to analyze the incentives for one-year abstinence, not for the multi-year abstinence.\(^\text{19}\) Even though one-year abstinence inevitably leads to partial relapse, the overall cost is much lower compared to the long-term programs.

The joint study of SOA and LIMRA (SOA 2019) presents details of the lapsation behavior conditioned on many policy characteristics. Especially, it contains the separate lapsation statistics of smokers and non-smokers. The term structure of lapsation rates across different product types (whole life or term life with varying maturities) are presented in Figure A2, and smokers have higher lapsation rates than non-smokers across the board. This clearly indicates that pricing competition is fierce for smokers, because they usually pay annual premiums that are multiples of what non-smokers pay. This price elasticity of smokers and the higher price competition is already reflected in the benefit numbers I calculate in Section 5. Even though the database from only one insurer does not allow the full empirical investigation of the higher lapsation behavior after the health improvement, the baseline contract valuation

\(^{19}\text{In my knowledge, there is no RCT testing the financial incentives for multi-year smoking cessations.}\)
is pricing in the higher market competition for the smokers, which addresses the incentive-compatibility of the policyholders.

One intuitive way to ensure the incentive compatibility to a policyholder is to require a minimum policy age to be eligible to participate, where the threshold depends on the age level. Long-term policyholders have more financial incentives to stay with the same insurer, as their payment tends to be front-loaded while the benefit is back-loaded, so the previous premium payments operate as natural collateral to prevent lapsation. At different ages I can calculate the precise minimum holding period requirement, that will build enough policy values in the early years while increasing the market-quouted premium rates due to higher ages. This condition ensures the strict incentive compatibility of the policyholders not to switch to another insurer.

6.2 Sensitivity of Benefits to the Lapsation Assumptions

When the zero-lapsation assumption is used, the average benefit level is much higher at around the USD 38K, which is more than twice as large as the baseline results. Such a big discrepancy is puzzling at first glance, especially compared to the robustness discussion presented in Koijen and Van Nieuwerburgh (2020), where the authors show much less dramatic impacts of different lapsation assumptions.

Figure 9 illustrates the economic intuition behind the differences in lapsation sensitivities between the cases. The first panel shows the health improvement from the smoking cessation program, which is the primary example used in this paper (the mortality event numbers are unadjusted from SOA 2015 VBT). The second panel shows the effect of the immunotherapy discussed in Koijen and Van Nieuwerburgh (2020) on a hypothetical late-stage cancer with annual flat mortality rate 20%.20 The face value effect, which is the main component of the

---

20Even though this mortality assumption seems to be very high compared to the smoking or other health conditions discussed in this paper, it is on the benign side among the late-stage cancers presented in Koijen and Van Nieuwerburgh (2020).
benefit is calculated from how much the mortality event curve is shifted to the right, and then discounted accordingly including the lapsation assumptions. Since the late-stage cancer has a highly front-loaded mortality event distribution, immediate death benefit payout to the policyholder with cancer is insensitive to the discount rate assumption and the lapsation assumption. Compared to the PV of the payout leg at the improved health status (blue curve), the absolute value of the immediate payout is much larger, which causes the insensitivities of the benefit amount to the lapsation assumption. Compared to the late-stage cancer, the benefit from the smoking cessation (and other chronic diseases) happen farther in the future. Therefore, pricing in this higher lapsation is crucial to correctly measure the benefit from these health engagement programs.

7 Extensions to Other Health Intervention Programs

There are many health variables in the database, and the current smoking status variable is only one of them. Using the same methodology, I calculate the potential benefits from other health intervention programs, assuming the existing health conditions can be improved. I focus on widely prevalent health conditions, which are listed in Table 6. The first row reiterates the average around 18k USD benefit from smoking cessation programs, already discussed in the previous sections. I also investigate the prevalence and the potential benefits from improving obesity, high cholesterol, hypertension, heart condition, and cancer. In determining the threshold for detecting abnormal range of the health variable, I adopt widely accepted thresholds for each health conditions, except that I use much higher (300) than the usual 240 threshold for the high-cholesterol. Refer to the second column of Table 6 for the threshold I use to filter the policyholders with each health condition, and the health improvement assumptions I use to calculate the benefit. Figure 10 shows the overall distribution of these health variables, where the threshold line is drawn in red.

The result shows some interesting variations in the average benefit numbers across the
health conditions. Obesity, high cholesterol and hypertension have the average benefit around 15k, 18k, and 24k respectively. These benefit levels are similar to the effect of smoking cessation programs. Indeed, managing the hypertension to the normal level can accrue more financial benefits to the life insurance company than the smoking cessation programs. To explore the comprehensive viability of these health engagement programs, the cost estimations must be modeled to be consistent with the medical literature, but these benefit numbers provide a good starting point. The aggregate financial benefit available from running these health engagement programs on the chronic conditions (or metabolic diseases) successfully is around 872 million U.S. dollars. Cancer and the heart conditions are binary variables in the database, but I can run similar exercises to measure the benefit when the variable is assumed to be turned off by proper treatments. The average benefit for cancer and heart conditions are around 10k and 24k USD, respectively. Including the cancer and heart conditions, the aggregate benefit is as large as 2 billion U.S. dollars. 21

8 Conclusion

In this paper I present the net benefit available to a life insurance company when it engages more with the policyholders to improve their health status. A unique, detailed database of health and mortality information from a large U.S. life insurer and the predictions from the state-of-the-art machine learning techniques, allow me to precisely measure the financial benefits of various health engagement programs. I present the smoking cessation program as the primary example, which results in USD 18k benefit per policyholder on average. To complete the cost-benefit analysis of operating the smoking cessation incentive programs, I build

21 Note that all different types of cancers are grouped as one category here, without further conditioning on the sites or stages, which will average out wide variabilities in mortality impacts. While Koijen and Van Nieuwerburgh (2020) fully explores these variability across the cancer types and stages, here the data does not allow me to observe the details of the cancers. I use 872 million USD number as the baseline, because these binary variables for cancer and heart condition lose too much health information, and managing these conditions will cost much more compared to managing other metabolic conditions.
and calibrate the rational addiction model of smoking, which I match to the key findings in health economics and medical research. The net benefit estimation of the smoking cessation program to the life insurance company is 87 millions USD under the baseline assumptions. Various implementation issues are discussed, especially in relation to the lapsation of the policyholders. I apply the same methodologies to calculate the aggregate benefit available to the life insurance company from running various health engagement programs on chronic conditions, and the aggregate benefit is as large as USD 872 million. The large benefits suggest that various health engagement programs ought to be offered to a targeted group of policyholders. Life insurers can play a socially important role in promoting better health of the policyholders while financially benefiting from the success of the programs.
References


Cahill, Kate, Jamie Hartmann-Boyce, and Rafael Perera. 2015. “Incentives for Smoking Cessation.” Cochrane Database of Systematic Reviews (5).


of Cigarette Consumption.” Working Paper, Yale University.


21st-Century Hazards of Smoking and Benefits of Cessation in the United States


This figure provides an illustrative example of the financial benefits accrued to the life insurance company (see Section 2.1 for details) using the case of 30-year old male. In all three panels, the red dotted curves show the previous health status (smoker), and the blue curves are the improved health status (quitter). When the health engagement program successfully improves the health status of the policyholder, the mortality curve (panel (a)) shifts downwards. As a result, the conditional survival probability curve shifts up (panel (b)), and the probability distribution of the mortality events shifts to the right (panel (c)). Panel (b) is related to the premium effect of equation (4) from the additional premium income streams, while panel (c) is related to the face value effect from the delayed death benefit payment.
Figure 2. **Random Survival Forest Model**

This figure describes the process of running the random survival forest (RSF) model of Ishwaran et al (2008). The overall structure is similar to the widely popular random forest model (Breiman 2001), but instead of the categorization or regression purpose, the random forest model is fitted to the survival data that is right-censored. Similar procedures apply: (1) reproducing bootstrapped samples (trees), (2) repeatedly branching out by selecting the split variables that give the highest difference via log-rank test (3) averaging (*ensembling*) the predictions from each trees to generate the model prediction (4) tuning the hyperparameters through cross-validation.
Figure 3. **Distributions of the Potential Benefit from Smoking Cessation.**

This figure presents the potential financial benefits (in 2018 USD) available to the Life Insurance Company when the smoking cessation program successfully helps the policyholders to quit. When a policyholder stops smoking, her health status clearly improves and the expected lifespan increases, so the life insurance company also benefits financially for two reasons: the delayed death-benefit payment and the longer stream of the premium income. See Equation (4) for the theoretical framework and Section 4.1 for the detailed measurement procedures. The average benefit per policyholder is around USD 18k, and the distribution is skewed to the right.
Figure 4. **Heterogeneity in the Benefits from Smoking Cessation.**

This figure presents the heterogeneity in the benefit amounts according to 5-year age groups. Each policyholder age is rounded to the nearest 5-year age groups. The first panel shows that the average face amount increases with age until early 40s, and then stays flat. The second panel shows that the average benefits from successful smoking cessation gradually increases with age. This third panel shows the normalized benefit as a fraction of the face value amount.
Figure 5. **Cost Model Solution.**

This figure presents the sample paths generated by the lifecycle model solution described in Section 4.2. There are two types of agents in the model. Both types maximize their lifetime utility by optimally choosing consumptions on the regular good and the cigarette, but they have different beliefs on the relative risks of smoking. The first type correctly predicts the actual relative risks documented by RCTs (orange dotted line), while the second type (blue line) underestimates the relative mortality risk of smoking ($\kappa_{UE} = 0.1$). The second type is crucial in explaining the existence of smokers in later stages in their lives. $\kappa_{UE} = 0.1$ is calibrated to the additional gain in years, presented in Jha et al (2013) (see Appendix Figure A7).
Figure 6. **Calibrated Model to Data.**

This figure presents three main target moments for the calibration. Panel (a) shows the increase in life expectancy at different quitting ages from Jha et al (2013). Panel (b) shows that the model is well matching the target 4 million USD value of life at age 40. Panel (c) matches the large-scale RCT from Volpp et al (2009), to pay $750 (about $1,000 in current value) to quit smoking for one year. The left bars show the quit rate among all the intent-to-treat samples, while the right bars show the conditional long-term abstinence ratio after the financial incentive program is over. Data from Volpp et al (2009) shows 6-month follow-up abstinence, and I present 5-year long-term abstinence rate here.
Figure 7. Required Financial Incentives.

This figure presents the required financial incentives for the smoking cessation of one year, calculated by Equation (13) and the calibrated model. Baseline estimations are calculated using the simulated model solution paths of smoking addiction and wealth. Panel (a) and (b) present the required incentive costs for the two types of agents with different risk perceptions. Panel (c) is the estimated costs, conditional only on the age of the current smoker.

(a) Financial Incentives Required for type $\kappa = 0.5$

(b) Financial Incentives Required for type $\kappa_{UE} = 0.1$

(c) Financial Incentives conditional on being Current Smoker
Figure 8. **Joint Distribution of Costs and Benefits.**

This figure presents the joint distribution of the potential benefits that are greater than the required costs, with the baseline long-term abstinence rate assumption of 30%. About 48% of the policies satisfy that the potential benefit is greater than the costs (policyholders under age 25 are excluded in the calculation). For these subsample of policies, the average benefit is 30k, while the required cost is 2,700 USD. The aggregate net benefit is about 87 million USD.
Figure 9. **Illustration of the Lapsation Sensitivities in Cancer vs Smoking.**

This figure illustrates the discrepancies in the lapsation sensitivities in the benefit amounts in two health improvement cases. The first panel shows the health improvement from the smoking cessation program, which is the primary example used in this paper (the mortality event numbers are unadjusted from SOA 2015 VBT). The second panel shows the effect of immunotherapy discussed in Kojien and Van Nieuwerburgh (2020) on a hypothetical late-stage cancer with annual flat mortality rate 20%. The face value effect, which is the main component of the benefit is calculated from how much the mortality event curve is shifted to the right, and then discounted accordingly including the lapsation assumptions. The late-stage cancer has a highly front-loaded mortality event distribution, which mostly dilutes the effect of lapsation.
Figure 10. **Distributions of Other Target Health Variables.**

This figure presents the distributions of other significant health variables used in the mortality prediction models. Similar to the smoking cessation program, I assume the improvement in these target variables and calculated the potential financial benefits from the improved health. For diabetes, I assume that the glucose levels are improved to the (age, sex, smoking status) group means. For obesity, I assume that the BMI levels are improved to 25. For high cholesterol, I assume that the cholesterol levels are improved to 240. For hypertension, I assume that the blood pressures are improved to 80 (diastolic) and 120 (systolic).
Table 1. **Selected List of Health Engagement Initiatives**

Life insurance industry is actively exploring the opportunities in health engagement programs for its policyholders to utilize the low-cost information and mobile technology and the newly available statistical techniques for big database. Below table is the selected list from the industry report (SCOR 2020) of the wide range of health engagement programs in the industry.

<table>
<thead>
<tr>
<th>Insurer</th>
<th>Health Engagement Initiative</th>
<th>Year</th>
<th>Category</th>
</tr>
</thead>
<tbody>
<tr>
<td>AXA</td>
<td>Partnered with global fitness aggregator ClassPass to provide wellness solutions that promote both physical and mental wellbeing to its customers.</td>
<td>2020</td>
<td>Gym &amp; Workout Programs</td>
</tr>
<tr>
<td>Gen Re</td>
<td>Looking to help insurance companies take advantage of Monsenso’s mHealth solution, which can assist in the prevention, early intervention and treatment of mental illness.</td>
<td>2017</td>
<td>B2B</td>
</tr>
<tr>
<td>Mass Mutual</td>
<td>Policyholders can access Aaptiv and Timehifter as part of Haven Life Plus, a rider to the Haven term policy that offers benefits beyond coverage.</td>
<td>2019</td>
<td>Gym &amp; Workout Programs, Sleep Wellness</td>
</tr>
<tr>
<td>Mutual of Omaha</td>
<td>Offers Mutually Well, a healthy lifestyle program for Medicare Supplement customers.</td>
<td>2019</td>
<td>Holistic Wellness, Seniors</td>
</tr>
<tr>
<td>John Hancock</td>
<td>Launched its HealthyFood program to reward customers for healthy eating by giving discounts or cash back on foods designated as healthy.</td>
<td>2016</td>
<td>Weight Management</td>
</tr>
<tr>
<td>John Hancock</td>
<td>John Hancock Aspire offers customers living with diabetes life insurance paired with a technology-enabled program that provides coaching, clinical support, education, and rewards to help manage and improve their health.</td>
<td>2019</td>
<td>Chronic Care</td>
</tr>
<tr>
<td>Munich Re</td>
<td>Introduced Wellgage, a connected health and wealth platform.</td>
<td>2017</td>
<td>Activity Wearables</td>
</tr>
<tr>
<td>Prudential</td>
<td>Launched Pulse, a health management mobile application that provides AI-powered self-help tools and real-time information.</td>
<td>2019</td>
<td>Holistic Wellness</td>
</tr>
<tr>
<td>RBC Insurance</td>
<td>Its digital wellness program allows users to identify health risks, get recommendations, track progress, measure impact and earn rewards.</td>
<td>2018</td>
<td>Holistic Wellness</td>
</tr>
<tr>
<td>State Farm</td>
<td>Working on an Alexa to improve senior care.</td>
<td>2020</td>
<td>Social Wellness, Seniors</td>
</tr>
</tbody>
</table>
Table 2. **Summary Statistics of the Insurance Database**

This table presents the summary statistics of the insurance database provided by the Life Insurance Company. The sample population is limited to whole life or term life insurance policies issued with the available health information of the policyholders during a specific period of time. Panel (a) presents the summary statistics of the variables at the contract level, and Panel (b) presents the summary statistics of the health variables collected during the underwriting process.

(a) **Contract Variables**

<table>
<thead>
<tr>
<th>Names</th>
<th>Mean</th>
<th>Std. Dev</th>
<th>Median</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Face Value</td>
<td>539,982</td>
<td>1,039,537</td>
<td>250,000</td>
<td>467,965</td>
</tr>
<tr>
<td>Is Whole Life</td>
<td>0.375</td>
<td></td>
<td></td>
<td>467,965</td>
</tr>
<tr>
<td>Is Term Life</td>
<td>0.625</td>
<td></td>
<td></td>
<td>467,965</td>
</tr>
<tr>
<td>Term Life Maturity</td>
<td>17.7</td>
<td>5.0</td>
<td>20.0</td>
<td>292,544</td>
</tr>
<tr>
<td>Policyholder Age</td>
<td>38.9</td>
<td>10.4</td>
<td>37.0</td>
<td>467,965</td>
</tr>
</tbody>
</table>

(b) **Policyholder Health Variables**

<table>
<thead>
<tr>
<th>Names</th>
<th>Mean</th>
<th>Std. Dev</th>
<th>Median</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Is Current Smoker</td>
<td>0.064</td>
<td></td>
<td></td>
<td>467,965</td>
</tr>
<tr>
<td>Has Cancer</td>
<td>0.043</td>
<td></td>
<td></td>
<td>467,965</td>
</tr>
<tr>
<td>Has Heart Condition</td>
<td>0.082</td>
<td></td>
<td></td>
<td>467,965</td>
</tr>
<tr>
<td>BMI</td>
<td>26.6</td>
<td>4.9</td>
<td>25.9</td>
<td>467,965</td>
</tr>
<tr>
<td>Glucose</td>
<td>83.4</td>
<td>14.8</td>
<td>83.0</td>
<td>467,965</td>
</tr>
<tr>
<td>Blood Pressure (Diastolic)</td>
<td>72.8</td>
<td>6.8</td>
<td>73.0</td>
<td>467,965</td>
</tr>
<tr>
<td>Blood Pressure (Systolic)</td>
<td>115.0</td>
<td>9.3</td>
<td>115.0</td>
<td>467,965</td>
</tr>
<tr>
<td>Cholesterol</td>
<td>188.8</td>
<td>37.7</td>
<td>186</td>
<td>467,965</td>
</tr>
</tbody>
</table>
Table 3. Rational Addiction Model Parameter Calibration

This table shows the summary of the calibrated parameters of the rational addiction model described in Section 4.2.1. The parameters are calibrated using various sources in medical and health economics literature.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Calibrated Value</th>
<th>Rationale</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\beta$</td>
<td>0.97</td>
<td>Standard in Literature</td>
</tr>
<tr>
<td>$R$</td>
<td>1.03</td>
<td>Standard in Literature</td>
</tr>
<tr>
<td>$\gamma$</td>
<td>2</td>
<td>Standard in Literature</td>
</tr>
<tr>
<td>$b$</td>
<td>1.2</td>
<td>Value of life (USD 4 million at age 40), Average spending on smoking</td>
</tr>
<tr>
<td>$\alpha$</td>
<td>0.03</td>
<td>Value of life (USD 4 million at age 40), Average/Max spending on smoking</td>
</tr>
<tr>
<td>$\sigma$</td>
<td>10</td>
<td>Average spending on smoking</td>
</tr>
<tr>
<td>$\theta$</td>
<td>1.0</td>
<td>Average spending on smoking, low smoking consumption variance</td>
</tr>
<tr>
<td>$\delta$</td>
<td>0.3</td>
<td>Effect of successful cessation of 10-15 years (USDHHS 2020)</td>
</tr>
<tr>
<td>$\kappa$</td>
<td>0.5</td>
<td>Relative risks from Table 5.5 of USDHHS (2020)</td>
</tr>
<tr>
<td>$\kappa_{UE}$</td>
<td>0.1</td>
<td>Gains in Life (6, 4 years) from quitting at age 50, 60 (Jha et al 2013)</td>
</tr>
<tr>
<td>$p_{UE}$</td>
<td>0.4</td>
<td>Quit Ratio at Age 50-59</td>
</tr>
<tr>
<td>$\gamma_{c}$</td>
<td>0.0167</td>
<td>Life cycle pattern of smoking</td>
</tr>
<tr>
<td>$\gamma_0$</td>
<td>-10.08</td>
<td>SOA 2015 VBT</td>
</tr>
<tr>
<td>$\gamma_1$</td>
<td>0.085</td>
<td>SOA 2015 VBT</td>
</tr>
</tbody>
</table>
Table 4. **Heterogeneity in Smoking Cessation Benefit Amounts**

This table presents the heterogeneity in the estimated benefit amounts from the smoking cessation program. The first column shows the regression results of the log benefit amount on the log size of the insurance policy. The second column includes the Policyholder Age as a regressor, and the third column includes the product type indicator (being a whole life policy compared to the term life). Standard errors are reported in parentheses, and all the coefficients are significant at the 1% level.

<table>
<thead>
<tr>
<th></th>
<th>$\ln(\text{Benefit})$</th>
<th>$\ln(\text{Benefit})$</th>
<th>$\ln(\text{Benefit})$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\ln(\text{FaceAmt})$</td>
<td>0.8575</td>
<td>0.8566</td>
<td>0.9003</td>
</tr>
<tr>
<td></td>
<td>(0.004)</td>
<td>(0.004)</td>
<td>(0.004)</td>
</tr>
<tr>
<td>Policyholder Age</td>
<td>0.0136</td>
<td>0.0140</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.000)</td>
<td>(0.000)</td>
<td></td>
</tr>
<tr>
<td>$\text{Is Whole Life}$</td>
<td></td>
<td></td>
<td>0.4027</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.007)</td>
</tr>
<tr>
<td>Constant</td>
<td>-1.3027</td>
<td>-1.8192</td>
<td>-2.5943</td>
</tr>
<tr>
<td></td>
<td>(0.049)</td>
<td>(0.050)</td>
<td>(0.049)</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.609</td>
<td>0.626</td>
<td>0.662</td>
</tr>
<tr>
<td>$N$</td>
<td>29,865</td>
<td>29,865</td>
<td>29,865</td>
</tr>
</tbody>
</table>
Table 5. **Summary of the Costs and the Benefits from Smoking Cessation**

This table presents the summary statistics of the costs and benefits from the smoking cessation program. The average benefit amount is around 18k in 2018 USD. The distribution is highly skewed to the right, and more than half of the benefit estimation is below USD 10k. Using Equation (4), the average benefit amount can be decomposed into 86% of the face value effect, and 14% of the premium effect.

<table>
<thead>
<tr>
<th></th>
<th>Benefit</th>
<th>Cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>18,049</td>
<td>3,186</td>
</tr>
<tr>
<td>Std. Dev</td>
<td>3,938</td>
<td>1,720</td>
</tr>
<tr>
<td>Median</td>
<td>9,112</td>
<td>3,665</td>
</tr>
<tr>
<td>Fraction of Face Value Effect</td>
<td>0.86</td>
<td></td>
</tr>
<tr>
<td>Fraction of Premium Effect</td>
<td>0.14</td>
<td></td>
</tr>
<tr>
<td>Fraction Higher than Cost*</td>
<td>0.53</td>
<td></td>
</tr>
</tbody>
</table>

*After adjusting for relapse
Table 6. **Benefits for Different Health Conditions**

This table presents the potential dollar benefits from improving other health conditions of the policyholders. I indicate how many policyholders within the sample population are subject to each health engagement programs. The first row presents the same benefit number as in Table 5 from successful smoking cessations. For diabetes, I assume that the glucose, fructosamine and HbA1c levels are improved to the average levels of the same age, sex, and current smoking status group. For obesity, I assume that the BMI levels are improved to 25. For high cholesterol, I assume that the level is improved to 240. For hypertension, I assume that the blood pressures are improved to 80 (diastolic) and 120 (systolic). Heart and cancer conditions are only flagged as a binary variable, so I assume that the existing conditions are cured.

<table>
<thead>
<tr>
<th>Health Condition</th>
<th>Original / Improved Conditions</th>
<th>% of Sample</th>
<th>Average Benefit (in 2018 USD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Smoking</td>
<td>No / Yes</td>
<td>6.4%</td>
<td>18,049</td>
</tr>
<tr>
<td>Diabetes</td>
<td>Glucose &gt; 125 / Group Mean*</td>
<td>1.3%</td>
<td>12,080</td>
</tr>
<tr>
<td>Obesity</td>
<td>BMI &gt; 40 / BMI 25</td>
<td>1.2%</td>
<td>15,011</td>
</tr>
<tr>
<td>High Cholesterol</td>
<td>Cholesterol &gt; 300 / Cholesterol 240</td>
<td>0.7%</td>
<td>17,533</td>
</tr>
<tr>
<td>Hypertension</td>
<td>D &gt; 90 or S &gt; 140 / D 80 and S 120</td>
<td>1.0%</td>
<td>24,381</td>
</tr>
<tr>
<td><strong>Aggregate</strong></td>
<td></td>
<td></td>
<td><strong>USD 872 million</strong></td>
</tr>
<tr>
<td>Heart Condition</td>
<td>Yes / No</td>
<td>8.2%</td>
<td>23,505</td>
</tr>
<tr>
<td>Cancer</td>
<td>Yes / No</td>
<td>4.3%</td>
<td>9,912</td>
</tr>
<tr>
<td><strong>Agg. incl. Heart &amp; Cancer</strong></td>
<td></td>
<td></td>
<td><strong>USD 1,975 million</strong></td>
</tr>
</tbody>
</table>

* Group mean of glucose, fructosamine, HbA1c of the non-diabetic patients within the same age group (5yr), sex, and current smoking status.
Appendix

A.1. Proof of Proposition 1.

**Proposition 1.** Under the independent lapsation assumption, i.e., when $k_{j,j+1} = \tilde{k}_{j,j+1}, \forall j$, the face value effect in equation (4) is positive.

**Proof.** We want to show that the face value effect, which is the second term of the equation (4),

$$FVE_i = \left( \sum_{t=\tau}^{T} \frac{k_{\tau,t}}{R^{t-\tau+1}} \cdot ((\tilde{\pi}_{\tau,t}(\tilde{\pi}_{t,t+1} - 1) - \pi_{\tau,t}(\pi_{t,t+1} - 1)) \right) \cdot F_i \quad (A1)$$

is positive.

Denote by $\delta_t = \frac{k_{\tau,t}}{R^{t-\tau+1}}$ for $t = \tau, \tau + 1, \ldots, T$. By construction, $\delta_\tau = 1/R < 1$, and for any $\tau \leq t \leq T - 1$, $\delta_{t+1} = \delta_t \cdot \frac{k_{t,t+1}}{R} < \delta_t$ as $R > 1$ and $k_{t,t+1} < 1$. So $\{\delta_t\}_{t=\tau,\tau+1,\ldots,T}$ is a strictly decreasing sequence between 0 and 1.

Now define $x_t = \tilde{\pi}_{\tau,t}(\tilde{\pi}_{t,t+1} - 1)$ and $y_t = \pi_{\tau,t}(\pi_{t,t+1} - 1)$ for $t = \tau, \ldots, T$. First observe that $x_\tau - y_\tau = \tilde{\pi}_{\tau,\tau+1} - \pi_{\tau,\tau+1} \geq 0$ by health-improving assumption. Similarly, for any $\tau + 1 \leq t \leq T$, $\sum_{s=\tau}^{t} x_s = \tilde{\pi}_{\tau,t} - 1$, and $\sum_{s=\tau}^{t} y_s = \pi_{\tau,t} - 1$, so

$$\left( \sum_{s=\tau}^{t} x_s \right) - \left( \sum_{s=\tau}^{t} y_s \right) = \tilde{\pi}_{\tau,t} - 1 - (\pi_{\tau,t} - 1) \geq 0 \quad (A2)$$

Rewriting the equation (A1) with the above notations:

$$FVE_i = \left( \sum_{t=\tau}^{T} \delta_t (x_t - y_t) \right) \cdot F_i \quad (A3)$$

$$= \left( \sum_{t=\tau}^{T} (\delta_t - \delta_{t+1}) \sum_{s=\tau}^{t} (x_s - y_s) \right) \cdot F_i \quad (A4)$$

where the equation (A4) is a simple algebraic manipulation with $\delta_{T+1} = 0$. Since $\{\delta_t\}_{t=\tau,\tau+1,\ldots,T}$
is a strictly decreasing sequence, the final expression in equation (A4) is non-negative when applying the equation (A2). Since we have assumed that at least one period exists where the conditional survival strictly increases (the same assumption we used in the premium effect), the face value effect becomes positive.
A.2. Additional Figures and Tables

Figure A1: Distribution of Benefit Amounts by Age Groups

This figure presents the full distribution of the benefit amounts from smoking cessation program, for the sample. Figure 4 presents the mean numbers by age groups, while the actual number of policyholders per age groups is important for designing the program. The age is highly concentrated around 40, and the benefit is right-skewed at all age groups.
Figure A2: Term Structures of Lapsation Rates

This figure presents the term structures of lapsation rates for smokers and non-smokers. The lapsation rates are from the industry survey (SOA 2019). The first panel shows the lapsation rates for the whole life policies, the second panel for the 20-year term life policies, the third panel for the 15-year term life policies, and the last panel for the 10-year term life policies. Across the product types and policy ages, smokers have higher lapsation rates compared than non-smokers.
Figure A3: Distribution of Number of Cigarette Smoking per Day

This figure presents the distribution of the number of cigarette smoking per day, plotted from the National Adult Tobacco Survey 2013-14. The sample is all the everyday smoker (filtered with Smokstatus_r code 1). The survey responses are bunched at 10, 20, 30, 40 and 60 as one pack of cigarette contains 20 cigarettes in them. The average number of smoking is 16.
Figure A4: **Average Cigarette Smoking per Day at Different Ages**

This figure presents the average number of cigarette smoking per day at different, 5-year age groups, plotted from the National Adult Tobacco Survey 2013-14 using the sample of all the everyday smoker (filtered with `Smokstatus_r` code 1). There are two interesting facts that the lifecycle rational addiction model captures. First, the number of smoking increases in the early years, as the smoker gets more addicted to smoking. Second, after hitting a peak, the smoker reduces the amount of smoking, as she is increasingly more concerned about the health effect.

(a) Distribution of Smoke per Day for Current Smokers

(b) Distribution of Smoke per Day for Current and Former Smokers
Figure A5: Composition by Smoking Status at Different Ages

This figure presents the composition of the current smokers and the quitters (former smokers) within each age groups, calculated from the National Adult Tobacco Survey 2013-14. The fraction of the lighter-colored bar is often referred as the quit ratio in medical research, since it shows what fraction of ever-smokers (defined as someone who smokes more than 100 cigarettes in her lifetime). This plot indicates that the quit ratio declines with ages.
Figure A6: **Relative Risk Perception of Lung Cancer Mortality**

This figure plots the distribution of the relative risk perception of lung cancer, using the survey data from Table 3 in Krosnick et al (2017). The grey line indicates the actual relative risks of lung cancer mortality, of 40-year old male who has been smoking one pack a day since 20 years old. Note that roughly 15% of the respondents actually indicate that the relative risk of dying from lung cancer is equal or lower for smokers, compared to the non-smokers. Overall distribution shows the massive underestimation of the relative risk among the current and the former smokers.
Figure A7: Empirical Benefits of Smoking Cessation

This figure reproduces one of the main result figures in Jha et al (2013), which is a large-scale meta-analysis of several cohort studies to measure the benefits of smoking cessations. This is regarded as the most up-to-date results on smoking cessation program results, and also widely cited and communicated in the general media, hence I use these results as major moments that I target in calibrate the cost model.