Working paper

Modeling health shocks

Michael Freiberger (freiberger@iiasa.ac.at)
Michael Kuhn (kuhn@iiasa.ac.at)
Stefan Wrzaczek (wrzaczek@iiasa.ac.at)

WP-23-002

Approved by:

Name Albert von Jaarsveld
Program Director General and Chief Executive Officer
Date 24 January 2023

Table of contents

Modeling health shocks[*]

Michael Freiberger[1], Michael Kuhn[1], and Stefan Wrzaczek[1]

[1]*Economic Frontiers (International Institute for Applied Systems Analysis)*

January 23, 2023

Abstract

Most of the models on the life-cycle utilization of health care assume that individuals are able to foresee the development of their health perfectly. However, health shocks with significant impact (e.g. severe life-threatening conditions, the onset of chronic disease or accidents) should not be averaged into a mean value, as they have the potential to put the life-course onto a different trajectory. In this paper, we introduce a dynamic optimal control framework incorporating a stochastic health shock with individuals allocating their resources to consumption and different kinds of health care over their life-cycle. We distinguish between general health care and shock specific prevention, acute and chronic care. This set-up enables us to analyse how the health risk shapes individual behaviour with respect to the different types of health care and how health shocks change the trajectories of consumption and savings. Newly developed transformation techniques allow us to investigate the optimal decisions made in anticipation of a potential health shock and the optimal reaction to all possible shock scenarios. We are able to obtain analytic expressions for the consumption and health care utilization profiles before and after the shock and identify the driving forces. Furthermore, we extend the value of life concept to other aspects of individual health. Finally, we illustrate our findings by calculating a numerical solution calibrated to an individual facing a potential cancer diagnosis in the US.

1 Introduction

Typically, life-cycle models of health behaviour take an ex-ante stance, where a representative individual is subject to the depreciation of a health stock (Grossman, 1972), subject to some mortality process (Ehrlich, 2000; Murphy and Topel, 2006; Hall and Jones, 2007; Kuhn et al., 2015), or subject to the accumulation of deficits (Dalgaard and Strulik, 2014). Individuals fix their life-cycle decisions in perfect anticipation of these processes, regardless of whether they are deterministic or stochastic. In a similar vein, the statistical value of a life is based on an ex-ante evaluation of survival (Shepard and Zeckhauser, 1984; Rosen, 1988; Murphy and Topel, 2006; Hall and Jones, 2007; Kuhn et al., 2010; Kuhn et al., 2011; Kuhn et al., 2015).

While hugely simplifying analytical complexity, the ex-ante perspective amounts to a gross stylization. In reality, health is not developing according to a smooth process which may be shaped to some extent by health investments, but it is rather subject to smaller or greater shocks. Some of these, such as severe life-threatening diseases (e.g. heart attacks, stroke or cancer) or accidents as well as chronic diseases (e.g. diabetes or dementia), have the potential to put the entire life-course on a different trajectory. Even if individuals are anticipating the risk of such a shock, the optimization problem changes in as far as the shock subjects the individual to different constraints (e.g. a permanently lower income stream due to disability or a persistently higher mortality risk after the shock) and, thus, to a different behavioural regime. In consequence, individuals are also prone to alter their behaviour prior to a health shock by engaging in precautionary saving or in preventive actions. The nature and intensity of such actions is shaped by institutions, such as the availability (or not) of health or disability insurance, as well as by policy interventions, such as the subsidization of preventive behaviour.

In this paper, we detail a life-cycle model, allowing us to study large, singular shocks to health. Thus, we consider a life-cycle model with endogenous health and survival in the spirit of Kuhn et al. (2015) but allow for the onset of a disease or accident at some random time s. The individual is only aware of the risk, modelled as a hazard rate, that can be influenced through prevention. If the health shock materializes at s, the health of the individual is affected, implying that (i) acute life-saving and/or disability-preventing health care may be required at s; (ii) mortality may be permanently elevated in the course of a chronic disease; (iii) a state of disease/disability emerges that may subsequently affect the individual's utility and earnings; (iv) chronic health care may be required to mitigate the adverse longer-term consequences of the shock. Since we model the life of an individual over time (i.e. age) and since the health shock can set in at any time during the life-course, the health state, assets and the individual's consumption and health care choices are age and duration dependent. This goes well beyond most of the analysis to date.

In order to solve the underlying stochastic optimal control model with a random stopping rule (akin e.g. to a problem considered by Boukas et al. (1990)) we transform the model into a vintage optimal control model, following a novel approach by Wrzaczek et al. (2020). This allows us to analyse in a coherent framework the way in which the pre- and post-shock dynamics of both state and control variables are linked through anticipation and retrospection. In so doing and in contrast to most of the previous literature, we make a clear distinction between preventive health care (lowering the hazard rate directly or indirectly), acute health care (lowering the instantaneous impact of the shock on survival and/or subsequent disability) and chronic health care (lowering the disease/disability state after the shock).

Our analysis allows us to generalize the value of life (VOL) and apply it to a setting with a health shock. Specifically, we derive the value of health before and after the health shock, the value of prevention, the value of surviving the shock, and the value of morbidity. We can also calculate an ex-ante value of

health. As it turns out, these values can be used to write the first order conditions for the different forms of health care in a compact form with a straightforward interpretation. Furthermore, the valuations are instrumental in understanding the dynamics of the different forms of health care.

In order to illustrate these dynamics, we employ numerical techniques developed by Veliov (2003) to present how the development of health and survival states depends on both age and duration of the disease (following the shock). We illustrate the age-duration specific dynamics by focusing on an application of our model to the onset of cancer as one major type of a health shock. For this purpose we calibrate our model (to reasonable extent) to reflect the dynamics of survival and health care spending relating to cancer, based on US data.

The economic literature on health-related shocks to the life-cycle is disperse. The largest part of this literature analyses the impact of health-related shocks to labour productivity, to health care expenditure, and to survival. Palumbo (1999), De Nardi et al. (2010), and Kopecky and Koreshkova (2014), for instance, focus on the impact of risky health care expenditure during old age on precautionary saving; while French (2005) and French and Jones (2011) focus on the impact of health shocks on labour supply and retirement.[1] Capatina (2015) provides an overall assessment of how four channels of health-related risk (productivity, medical expenditure, time endowment, and survival) bear on income inequality and precautionary saving. Common to all of these papers is that they model health risks as a sequence of possibly state-dependent shocks over the life-cycle without emphasizing the scope for "catastrophic" shocks with a propensity to drastically shift the life-cycle trajectory. Furthermore and importantly, they do not provide scope for individuals to reduce the probability of shocks and/or mitigate their consequences by purchasing preventive and/or curative health care, or engaging in other health-related actions. Reichling and Smetters (2015) allow for possibly large shocks when showing that stochastic mortality risks may explain the (rational) abstinence from private annuities when these shocks are correlated with medical expenses. This notwithstanding, the focus of their work is again not on the impact of such shocks on the utilization of health care and resulting health outcomes.[2]

Few works so far have studied how the risk of health shocks bears on an individual's health investments, either in response to a shock or, more importantly, in anticipation of a shock. Cole et al. (2019) consider a setting where (curative) health expenditure and labour productivity are subject to health shocks, the propensity for which depends on health status. The authors examine how the individual's incentive to engage in preventive efforts aimed at improving their health status are shaped by non-discriminatory health insurance and wage-setting. There is no mortality risk and, as the authors themselves remark in their conclusions, the focus is on small (transitory) shocks to health. In Hugonnier et al. (2013) individual productivity and mortality both depend on an underlying health stock a la Grossman (1972). Individuals can invest into this stock of health, which is assumed to be subject to morbidity shocks following a Poisson process. The key distinction to our approach is that their modelling of uniform health investments and a sequence of (at least principally) transitory shocks (apart from death) does not allow them to discriminate between preventive health care (lowering the probability of a shock) and curative health care (reducing the damage from a shock). Neither do they distinguish between the valuation of preventive care as opposed to the value of curative or chronic care,[3] distinctions which will feature prominently in our work. Finally, Hugonnier et al. (2013) apply their model to understand the relationship between financial as opposed to health investments, whereas our focus lies on how the life-

---

[1] The literature on health shocks as motivation for precautionary savings ties in with a large literature on the savings response to (general) life-cycle risks (Eeckhoudt et al. (2005) and Eeckhoudt and Schlesinger (2008))

[2] Smith and Keeney (2005) examine the valuation of health in a setting where individuals face lotteries over their health and income at distinct phases of their life-cycle. While these lotteries may involve large shocks, the authors do not model the timing of these shocks; nor do they endogenize the health risks.

[3] The same applies to the models in Picone et al. (1998), Fonseca et al. (2013), Jung and Tran (2016) and Yogo (2016).

cycle allocation of preventive and curative care is shaped by the nature of the health shock. The notion of permanent health shocks is only taken up by Laporte and Ferguson (2007) who consider a version of the Grossman (1972) model in which the health stock is subject to a single irreversible shock, the arrival of which follows a Poisson process. They examine how the nature of this shock bears on the ex-ante path of health investments; however, health is assumed to bear on morbidity and, thus, on period utility but not on survival. Indeed, the length of life is assumed to be exogenous and deterministic.

This leads us to conclude that while the theoretical literature on health shocks has made considerable advances in terms of understanding the consequences of (a sequence of) small shocks for life-cycle patterns of labour supply, income, expenditure, and savings/consumption, comparatively little is known yet about the implications of shocks for the demand for health care and for health behaviours in regard to both their preventive and curative aspect. In particular, this applies to large shocks, such as the onset of severe chronic disease (diabetes, heart disease, cancer) or debilitating accidents, which induce permanent rather than transitory shifts in the mortality, morbidity and income patterns over the remaining life-cycle. These issues lie at the heart of the present work.[4]

The remainder of the paper is structured as follows. The next section contains a description of the model. Section 3 presents the analytical solution, involving in particular the derivation of various value (of health) terms from the set of relevant shadow prices (Subsection 3.1) and their subsequent employment in the first-order conditions (Subsection 3.2). Section 4 then proceeds to present the necessary foundations for the numerical analysis, with Subsections 4.1 through 4.4 setting out data, functional specifications as well as details of the solution strategy. Section 5 finally examines the numerical solution of our framework starting with a comparison of data and the calibrated model output in Subsection 5.1. Subsections 5.2 to 5.4 break down the numerical consumption, expenditure, health and survival profiles and provide an exploration of the driving forces behind their behaviour as well as illustrations of the numerical assessment of several distinguished values of health. Section 6 concludes.

2 Model

In this section we present a framework that integrates a large shock to the health of an individual into a life-cycle model. For the timing of the shock we make the following assumptions (which apply throughout the paper).

(A1) A large shock to health occurs at some age s, which is random. The probability rate of arrival is known by the individual.

(A2) The event at s occurs only once, thus the life-time of an individual can be separated into a stage before and a stage after s.

These two assumptions allow us to formulate the model as a stochastic optimal control model with a random stopping time (see Boukas et al. (1990)) and analyse it in terms of a vintage optimal control model (see Wrzaczek et al. (2020)). In both stages the individual chooses consumption and different types of health care in order to optimize (expected) life-time utility. Denoting by $t \in [0, T]$ the age of the individual, where $T > 0$ gives the maximum feasible age, we then have that $s \in [0, T]$ is the age

---

[4]From an empirical perspective, the behavioural adjustments and consequences for income and well-being of health shocks have been studied extensively. Adjustments have been studied in regard to saving and consumption (Bíró (2013)) as well as in regard to health behaviours, in particular smoking (e.g. Smith et al. (2001); Khwaja et al. (2006); Marti and Richards (2017)).

at which the model switches from stage 1 (where $t < s$) to stage 2 (where $t > s$). In the following, we introduce both stages of the model and specify the rate at which the shock arrives.

Using the index i ($i = 1, 2$) to denote the life-cycle stage that is referred to, we assume that the survival probability $S_1(t)$ in stage 1 is determined by the stage-1 mortality rate $\mu^1(\cdot)$.[5] This rate is assumed to be equal to the base mortality rate $\mu^b(t, S_1, b_1)$ which depends on age t, decreases in survival $S_1(t)$, and decreases in the quantity $b_1(t)$ of general health care subject to decreasing returns. Here, $b_1(t)$ is a generic measure of all health care that is unrelated to the condition(s) relating to the health shock. As described in Kuhn et al. (2015), the high correlation between survival and health implies that $S_1(t)$ can be interpreted as a proxy for the health status. We thus capture the negative dependency of mortality on health by including $S_1(t)$ as an argument in μ^b.[6] Altogether, the survival probability evolves with age according to

$$
\begin{aligned}
\dot{S}_1(t) &= -\mu^1(t, S_1(t), b_1(t))S_1(t) \\
&= -\mu^b(t, S_1(t), b_1(t))S_1(t), \qquad S_1(0) = 1.
\end{aligned}
\tag{1}
$$

The individual maximizes expected life-time utility

$$
\mathbb{E}_s \left[\int_0^s e^{-\rho t} S_1(t) u^1(c_1(t)) \, dt + e^{-\rho s} V^* \left(S_1(s), A_1(s), s \right) \right].
\tag{2}
$$

The first term contains the aggregated utility from birth up to s, where the period utility $u^1(c_1(t))$ from consumption $c_1(t)$ is weighted by survival $S_1(t)$ and a discount factor $e^{-\rho t}$ (the discount rate ρ is assumed to be exogenous).[7] The function $u^1(\cdot)$ is assumed to fulfil the classic assumptions of positive but diminishing marginal utility as well as the Inada conditions if consumption tends to zero or infinity. The second part of the expected utility denotes the discounted aggregated utility of the remaining life-time, given that the individual has suffered a health shock at age s. Here, $V^*(\cdot)$ denotes the optimal value (i.e. the value function) of the optimal control problem in the second stage and not only depends on the age s at the occurrence of the shock itself, but also on the survival/health state $S_1(s)$ and the assets $A_1(s)$ at the point of the shock. The expected value is built with respect to the random variable s. The likelihood of a shock occurring at a given age can be influenced by the individual through investments in preventive care h_1. Let the probability distribution of s be defined by $\mathcal{F}(t) = \mathbb{P}[s \leq t]$. The probability distribution of s can be characterized by the hazard rate η of the shock, which is generally defined by (3) and for which equation (4) holds:

$$
\eta(t) = \frac{\mathcal{F}'(t)}{1 - \mathcal{F}(t)},
\tag{3}
$$

$$
\mathcal{F}(t) = 1 - e^{-\int_0^t \eta(a)da}.
\tag{4}
$$

More specifically, we assume the hazard rate

$$
\eta(t) = \eta(t, S_1(t), h_1(t))
\tag{5}
$$

to depend on age t, to decrease in survival/health $S_1(t)$, and to decrease in the utilization of preventive

---

[5]Note that we will omit t wherever it is not of particular importance.

[6]This formulation might be surprising at first sight, but following the explanations in Freiberger and Kuhn (2020) and under some weak assumptions it is consistent with a deficit accumulation model as developed in Schuenemann et al. (2017), or a classic Grossman-type (Grossman, 1972) model with a monotonously depreciating health stock over the life-cycle.

[7]This formulation reaches back to Yaari (1965).

care $h_1(t)$ (again subject to diminishing returns). One might think, for instance, of h_1 as the propensity to invest in the vaccination against an infectious disease or as the propensity to attend precautionary screenings for cancer or heart disease.

The asset dynamics in stage 1 follow

$$\dot{A}_1(t) = (r(t) + \bar{\mu}(t))A_1(t) + w^1(t) - c_1(t) - p^b(t)b_1(t) - p^1(t)h_1(t), \tag{6}$$

$$A_1(0) = 0 \qquad \text{and} \qquad A_1(T) = 0; \tag{7}$$

where in stage 1 assets $A_1(t)$ are annuitized and generate a return $(r(t) + \bar{\mu}(t))$, with $r(t)$ and $\bar{\mu}(t)$ denoting the interest rate and the mortality risk premium on annuities[8], respectively; where $w_1(t)$ denote stage-1 earnings; where the price for consumption $c_1(t)$ is normalized to one; and where $p^b(t)$ and $p^1(t)$ denote the prices for general and specific preventive health care, $b_1(t)$ and $h_1(t)$, respectively. As usual we assume zero assets at birth and at the end of the maximum life-span T.

In stage 1, the individual chooses the (non-negative) control variables $c_1(t)$, $b_1(t)$ and $h_1(t)$ so as to maximize the objective function (2) subject to the constraints (1) and (5)–(7).

Stage 2 is modelled in a similar vein but we now consider a disease stock $E(t, s)$ as an additional state variable that bears on the individual's utility and constraints. For all variables in the second stage t and s describe the age of the individual and the age at which the health shock has occurred respectively. We assume that the condition that sets in at s may be associated with a specific mortality $\mu^m(t, s, E(t, s))$, depending on age, the time of the shock (or onset of disease), and on the disease stock. In addition, the individual continues to be subject to a base mortality $\mu^b(t, S_2(t, s), b_2(t, s))$ which can be reduced by general health investments $b_2(t, s)$. With the total mortality rate μ^2 of the individual in the second stage being the sum of μ^b and μ^m, the dynamics of stage-2 survival $S_2(t, s)$ can be written as

$$\dot{S}_2(t, s) := \frac{dS_2(t, s)}{dt} = -\mu^2(t, s, S_2(t, s), b_2(t, s), E(t, s))S_2(t, s)$$
$$= -\left[\mu^b(t, S_2(t, s), b_2(t, s)) + \mu^m(t, s, E(t, s))\right] S_2(t, s).$$

The disease stock $E(t, s)$ evolves according to

$$\dot{E}(t, s) := \frac{dE(t, s)}{dt} = f(t, s, E(t, s), h_2(t, s)) \tag{8}$$

where f depends on age, age at the time of the shock, the disease stock itself, and on disease-specific (chronic) health care $h_2(t, s)$, which is aimed at lowering the disease stock (again subject to diminishing marginal effects). Our general formulation of the disease dynamics allows for a range of different interpretations. These include, in particular, the cases of (i) an accident or acute disease at the point of the shock, which leaves the individual disabled initially but where a natural healing process, supported perhaps by health care, leads to a gradual reduction of $E(t, s)$ (understood to be the extent of disability); and (ii) a progressive disease, such as cancer, diabetes or Alzheimer dementia, where $E(t, s)$ tends to increase unless it is kept in check or lowered by the consumption of health care.

To further account for the negative consequences of the onset of disease (or disability), $E(t, s)$ is assumed to lower stage-2 earnings, w^2, and stage-2 utility u^2, i.e.

$$\frac{\partial u^2(c_2, E)}{\partial E} \leq 0 \qquad \qquad \frac{\partial^2 u^2(c_2, E)}{\partial E^2} \leq 0,$$

$$\frac{\partial w^2(t,s,E)}{\partial E} \le 0 \qquad \frac{\partial^2 w^2(t,s,E)}{\partial E^2} \le 0.$$

Finally, we assume the (initial) level of the disease state at the time of the shock $t = s$ to be a decreasing function of the general health state, as proxied by $S_1(s)$, and of one-off (acute) health care $d(s)$, i.e. $E(s,s) = B(S_1(s), d(s))$. We assume that acute care affects initial deficits in addition to the probability of surviving the health shock, as the example of cardiac arrest shows, that effective emergency care also affects the long run consequences of the heart attack (see Hassager et al. (2018)).

The dynamics of stage-2 assets $A_2(t,s)$ are similar to those in stage 1 with the following differences. First, earnings are not exogenous but depend on $E(t,s)$, as detailed above. Second, expenditures for chronic health care, purchased at price $p^2(t)$, substitute for preventive health care. Third, the initial stage-2 assets $A_2(s,s)$ (just after the shock) equal the stage-1 assets $A_1(s)$ (just before the shock) net of the expenditure for acute care which is purchased at a price $p^d(s)$). Thus,

$$\dot{A}_2(t,s) = (r(t) + \bar{\mu}(t))A_2(t,s) + w^2(t,s,E(t,s)) - c_2(t,s) - p^b(t)b_2(t,s) - p^2(t)h_2(t,s) \tag{9}$$

$$A_2(s,s) = A_1(s) - p^d(s)d(s) \qquad \text{and} \qquad A_2(T,s) = 0. \tag{10}$$

According to the last boundary condition, assets have to equal 0 at the end of life regardless of when the shock has occurred. The aggregated utility during stage 2 consists of the present value of the expected (i.e. survival weighted) utility stream over the remaining life-course

$$P(S_1(s), d(s)) \cdot \int_s^T e^{-\rho t} S_2(t,s) u^2(c_2(t,s), E(t,s)) \, dt, \tag{11}$$

where $P(S_1(s), d(s)) \in [0,1]$ is the probability that the individual survives the health shock, which similar to the initial value of the disease stock increases in the level of the stage-1 health state $S_1(s)$ and the quantity of acute health care $d(s)$ (subject to diminishing returns). Note that $P(\cdot) < 1$ captures the potential that the individual does not survive the shock (e.g. an accident, a cardiac event or a stroke), whereas $P(\cdot) = 1$ would reflect a disease that is not mortal upon its onset at s but only potentially so over time (e.g. cancer, diabetes, Alzheimer's disease). Finally, our model also embraces the case of health shocks leading to instantaneous death, $P(\cdot) = 0$. From now on, we will refer to $P(\cdot)$ as the "continuation probability".

The complete model is summarized in Equations (12) - (20).

$$\max_{c_1(t),h_1(t),b_1(t) \ge 0} \mathbb{E}_s \left[\int_0^s e^{-\rho t} S_1(t) u^1(c_1(t)) \, dt + e^{-\rho s} V^* \left(S_1(s), A_1(s), s \right) \right] \tag{12}$$

$$\dot{S}_1(t) = -\mu^1(t, S_1(t), b_1(t)) S_1(t), \tag{13}$$

$$\dot{A}_1(t) = (r(t) + \bar{\mu}(t)) A_1(t) + w^1(t) - c_1(t) - p^b(t) b_1(t) - p^1(t) h_1(t), \tag{14}$$

$$S_1(0) = 1, \qquad A_1(0) = 0, \qquad A_1(T) = 0 \tag{15}$$

where

$$V^* \left(S_1(s), \ A_1(s), s \right) := \max_{\substack{c_2(t,s),h_2(t,s), \\ b_2(t,s),d(s) \ge 0}} P(S_1(s), d(s)) \cdot \int_s^T e^{-\rho t} S_2(t,s) u^2(c_2(t,s), E(t,s)) \, dt \tag{16}$$

$$\dot{S}_2(t,s) = -\mu^2 \left(t, s, S_2(t,s), b_2(t,s), E(t,s) \right) S_2(t,s) \tag{17}$$

$$\dot{A}_2(t,s) = (r(t) + \bar{\mu}(t)) A_2(t,s) + w^2(t,s,E(t,s)) - c_2(t,s) - p^b(t) b_2(t,s) - p^2(t) h_2(t,s) \tag{18}$$

$$\dot{E}(t,s) = f(t,s,E(t,s), h_2(t,s)) \tag{19}$$

$$S_2(s,s) = S_1(s), \quad A_2(s,s) = A_1(s) - p^d(s)d(s), \quad A_2(T,s) = 0, \quad E(s,s) = B(S_1(s),d(s)) \quad (20)$$

Problem (12)-(15) can be interpreted as an optimal control model with random stopping time (see Boukas et al. (1990)). For the analysis and for the numerical solution we transform the model into a vintage optimal control model, as this offers additional economic insights as well as an established numerical solution method (see Veliov (2003)).[9] For the theoretical background and other examples of the transformation method we refer to Wrzaczek et al. (2020). As the presentation of the model in vintage optimal control form is not immediately instructive, we relegate it to appendix 7.1. Here, we only note that the vintage formulation implies that all second-stage variables are indexed by both age t and the time of the shock s, which can be interpreted as the arrival-date of a (potential) vintage of the remaining life-course (in disease). Indeed, the notation we have introduced earlier meets this criterion.

The transformation includes the introduction of the following two auxiliary variables, which we will subsequently employ in our calculations and for which interpretations are straightforward. First, $Z_1(t)$ denotes the probability, that an individual has not suffered a health shock up to age t. We will also relate to $Z_1(t)$ as the survival in good health[10]. As described above, the arrival rate of the health shock $\eta(t, S_1(t), h_1(t))$ depends on age, health status and preventive health care. This implies that the development of Z_1 can be formulated through the following differential equation

$$\dot{Z}_1(t) = -\eta(t, S_1(t), h_1(t))Z_1(t) \qquad , \qquad Z_1(0) = 1. \quad (21)$$

Second, we need the auxiliary variable $Z_2(s)$, which is defined by

$$Z_2(s) = Z_1(s)\eta(s, S_1(s), h_1(s))P(S_1(s), d(s)), \quad (22)$$

and can be interpreted as the joint probability of experiencing *and* surviving a shock at age s.[11]

For further reference, Table 1 summarizes the control and state variables in the two life-cycle stages.

Control variables	*Stage 1*	*Stage 2*	*Shock time* s
Consumption	$c_1(t)$	$c_2(t,s)$	-
General Health investments	$b_1(t)$	$b_2(t,s)$	-
Prevention expenditures	$h_1(t)$	-	-
Chronic care	-	$h_2(t,s)$	-
Acute treatment	-	-	$d(s)$
State variables			
Survival probability	$S_1(t)$	$S_2(t,s)$	$S_2(s,s) = S_1(s)$
Assets	$A_1(t)$	$A_2(t,s)$	$A_2(s,s) = A_1(s) - p^d(s)d(s)$
Survival in good health	$Z_1(t)$	-	-
Joint "probability" of shock *and* survival at s	-	$Z_2(s)$	$Z_2(s) = P(S_1(s),d(s))Z_1(s)\times$ $\times\eta(s,S_1(s),h_1(s))$
Severity of health deficits	-	$E(t,s)$	$E(s,s) = B(S_1(s),d(s))$

Table 1: Summary of all state and control variables in the basic framework

---

[9]The solution process is, nevertheless, not trivial as Veliov (2003) sets up a general framework, requiring multifaceted adaptations for the solution of specific problems.

[10]It directly holds that $Z_1(t) = 1 - \mathcal{F}(t)$.

[11]The term probability is not precise as we are analysing a time-continuous model and, strictly speaking, $Z_2(s)$ defines the probability density function of s multiplied with the continuation probability P. However the term "probability" makes for more intuitive reading.

3 Analytical results

For the vintage optimal control model, we apply the Maximum Principle presented in Feichtinger et al. (2003) to arrive at a set of necessary optimality conditions. Specifically, we obtain

- a set of differential equations describing the dynamics of the adjoint variables;

- a set of transversality conditions for the adjoint variables, corresponding to state variables with a free endpoint;

- a set of first-order optimality conditions for all control variables at every point in time/age, and for the stage-2 control variables for every possible point in life at which the health shock can occur.

These conditions, together with the state equations (with initial and boundary conditions), are used to find the optimal solution. In the following section, we will employ the first order conditions and the adjoint variables (together with the corresponding differential equations) to identify and characterize the various behavioural channels of the model.

3.1 Valuations of Health

Following Rosen (1988), Murphy and Topel (2006), Hall and Jones (2007), Kuhn et al. (2015) and others we investigate the individual willingness to pay for changes in health. However, while the original works focus on a reduction in the mortality risk, i.e. the value of life (VOL), we distinguish between the willingness to pay for changes in a range of different aspects of health.

Definition 1. For the analysis of problem (12) - (20), we define the following valuations of health, which we will use throughout. Here, $\mathcal{V}$ denotes the value function of problem (12)-(15).

Value of health ψ_H^i in stage $i = 1, 2$: Willingness to pay for a reduction in the mortality rate (depreciation rate of the survival stock) μ^i in stage $i = 1, 2$. $\psi_H^i := \left(-\frac{d\mathcal{V}}{d\mu^i} \right) / \left(\frac{d\mathcal{V}}{dA_i} \right)$.

Value of prevention ψ_P: Willingness to pay for a reduction in the risk of a health shock by reducing η, $\psi_P := \left(-\frac{d\mathcal{V}}{d\eta} \right) / \left(\frac{d\mathcal{V}}{dA_1} \right)$.

Value of acute survival ψ_{AS}: Willingness to pay for an increase in the probability P of surviving the shock, $\psi_{AS} := \left(\frac{d\mathcal{V}}{dP} \right) / \left(\frac{d\mathcal{V}}{dA_2} \right)$.

Value of morbidity ψ_M: Willingness to pay for a reduction in the disease/disability stock E, $\psi_M := \left(-\frac{d\mathcal{V}}{dE} \right) / \left(\frac{d\mathcal{V}}{dA_2} \right)$.

In addition, it is convenient to define the value of second-stage life as

$$\psi_{life}^2(t, s) = \int_t^T R(t, \tau) \frac{u^2(\tau, s)}{u_{c_2}^2(\tau, s)} \, d\tau, \tag{23}$$

with

$$R(t, \tau) := \exp\left(-\int_t^\tau r(\tau') + \bar{\mu}(\tau') d\tau' \right),$$

i.e. as the present value at age t of the stream of consumer surplus over the remaining life-course in stage 2, with the return on annuities being applied as discount rate. Note the similarity to the "'conventional"' value of life in earlier works (e.g. Kuhn et al., 2015).

Based on this, Proposition 1 presents an explicit analytical formulation of the valuations defined above.

Proposition 1. *Assume the existence of optimal trajectories of consumption and (the various) health investments in both stages of the individual life-cycle model (12)-(20) together with an interior solution for the consumption profiles c_1 and c_2. The valuation terms in Definition 1 can then be written as follows.*[12]

Stage-2 valuations:

$$\psi_H^2(t,s) = \int_t^T R_H^2(t,\tau,s) \frac{u^2(\tau,s)}{u_{c_2}^2(\tau,s)} \, d\tau, \tag{24}$$

$$\psi_M^2(t,s) = \int_t^T R_M^2(t,\tau,s) \left\{ \mu_E^2(\tau,s)\psi_H^2(\tau,s) - w_E^2(\tau,s) - \frac{u_E^2(\tau,s)}{u_{c_2}^2(\tau,s)} \right\} \, d\tau. \tag{25}$$

Stage-1 valuations:

$$\psi_{AS}(t) = \frac{\psi_{life}^2(t,t)}{P(t)}, \tag{26}$$

$$\psi_H^1(t) = \int_t^T R_H^1(t,\tau) \left\{ \frac{u^1(\tau)}{u_{c_1}^1(\tau)} - \eta_{S_1}(\tau)S_1(\tau)\psi_P(\tau) + \eta(\tau)P(\tau)\frac{u_{c_2}^2(\tau,\tau)}{u_{c_1}^1(\tau)} \times \right.$$
$$\left. \times \left[\psi_H^2(\tau,\tau) + P_{S_1}(\tau)S_1(\tau)\psi_{AS}(\tau) - S_1(\tau)B_{S_1}(\tau)\psi_M(\tau,\tau) \right] \right\} d\tau, \tag{27}$$

$$\psi_P(t) = \int_t^T R_P^1(t,\tau) \left\{ \frac{u^1(\tau)}{u_{c_1}^1(\tau)} + \eta P \frac{u_{c_2}^2(\tau,\tau)}{u_{c_1}^1(\tau)}\psi_{life}^2(\tau,\tau) \right\} \, d\tau - \frac{u_{c_2}^2(t,t)}{u_{c_1}^1(t)}P\psi_{life}^2(t,t). \tag{28}$$

The various discount factors are defined by

$$R_P^1(t,\tau) := R(t,\tau)\exp\left(-\int_t^\tau \eta(\tau')P(\tau')\frac{u_{c_2}^2(\tau',\tau')}{u_{c_1}^1(\tau')}d\tau' \right), \tag{29}$$

$$R_H^1(t,\tau) := R(t,\tau)\exp\left(-\int_t^\tau \mu_{S_1}^1(\tau')S_1(\tau') + \eta(\tau')P(\tau')\frac{u_{c_2}^2(\tau',\tau')}{u_{c_1}^1(\tau')}d\tau' \right), \tag{30}$$

$$R_H^2(t,\tau,s) := R(t,\tau)\exp\left(-\int_t^\tau \mu_{S_2}^2(\tau',s)S_2(\tau',s)d\tau' \right), \tag{31}$$

$$R_M^2(t,\tau,s) := R(t,\tau)\exp\left(\int_t^\tau f_E(\tau',s)d\tau' \right). \tag{32}$$

The proof is technical and thus relegated to the Appendix 7.3. Although the above expressions look involved, the individual terms can be assigned to the forward and backward incentives that shape the qualitative behaviour of the optimal solution. Before discussing the economic interpretations of all valuations in detail we want to stress, that the prerequisites of Proposition 1 are rather weak. Besides the assumption of existence of an optimal solution, it is sufficient to have interior solutions for the consumption paths to be able to derive the presented terms. Consequently all valuations still hold in the absence of some control variables (shocks may only require some of the controls for a proper modelling) or for cases with boundary solutions concerning any types of health investments.

Value of health in the second stage: $\psi_H^2(t,s)$

The value of health in the second stage ψ_H^2 (see equation (24)) is closely related to the value of life, i.e. the discounted value of consumer surplus over the remaining life-cycle, as defined in equation (23). The

---

[12] For notational convenience we omit the state and control arguments in all functions and just indicate at which time-point the state and control values should be evaluated, e.g. $u_{c_2}^2(\tau,s) \equiv u_{c_2}^2(c_2(\tau,s), E(\tau,s))$.

only difference is that the value of health takes into additional account a term $\mu^2_{S_2} S_2$ in the discounting function R^2_H, reflecting that better health, as measured by S_2, contributes to lower mortality over the remaining life-course. Given that $\mu^2_{S_2} < 0$, the new term decreases the discount factor and, thereby, raises the value of health over and above the conventional value of life. Intuitively, this reflects the additional value of health/survival as an asset that by lowering mortality yields a return in terms of additional consumer surplus over the remaining life-course.

Value of acute survival: $\psi_{AS}(t, s)$

The value of marginally increasing the continuation probability, ψ_{AS}, (see equation (26)) at the time of the shock directly corresponds to the conventional value of life in the second stage, ψ^2_{life}.[13] The weighting of ψ^2_{life} by $1/P(t)$ implies that the willingness to pay for acute survival increases with the risk of not surviving the health shock.

Value of morbidity: $\psi_M(t, s)$

The willingness to pay for a (marginal) reduction in the disease/disability stock, ψ_M, (see equation (25)) depends on the interaction of three effects: (i) the first part $\mu^2_E \psi^2_H$ captures the impact of the disease/disability stock on mortality or, equivalently, on the depletion of the health stock. A marginal increase in the disease stock increases the mortality risk by μ^2_E, which consequently leads to a change in second-stage survival, S_2, which is valued at the willingness to pay ψ^2_H. (ii) the part $-w^2_E$ contains the impact of the disease/disability on earnings, which directly translates into a willingness-to-pay to lower the disease/disability stock. (iii) The last part $-\frac{u^2_E}{u^2_{c_2}}$ measures the reduction in consumer surplus in the presence of disease/disability, or equivalently the marginal rate of substitution between consumption and degree of illness/disability. The greater the marginal impact of disability on utility, the greater the willingness-to-pay for reductions in morbidity. Finally, we have to adjust the standard discount rate R by accounting for the direction and speed of disease progression, as measured by the impact f_E of the disease stock on its own accumulation. Thus, the value of future changes in morbidity tends to be discounted more heavily if the disease is accelerating, i.e. $f_E > 0$, and less heavily if it is decelerating, i.e. $f_E < 0$. Intuitively, this suggests that reductions in morbidity are more valuable in the future (present) if the disease is accelerating (decelerating).

Value of prevention: $\psi_P(t)$

The willingness to pay for a reduction in the hazard rate of the health shock, ψ_P, (see equation (28)) is equivalent to the net value of remaining in (the healthy) stage 1 as opposed to transiting into (the diseased/disabled) stage 2. Accordingly, the integral term measures the value of remaining in stage 1, which in itself is composed of two distinct factors. The first part $\frac{u^1(c_1(\tau))}{u^1_{c_1}(c_1(\tau))}$ amounts to the consumer surplus for each year that continues to be spent in stage 1, whereas the second part adds the expected value of stage 2 utility should a transition occur at rate $\eta(\tau)$ in some future year $\tau > t$. Note that this value corresponds to the stage-2 value of life, ψ^2_{life}, weighted with the probability $P(\tau)$ of surviving a health shock at τ. As ψ^2_{life} is counted in units of stage-2 consumption, a conversion into units of stage-1 consumption takes place by multiplication with $\frac{u^2_{c_2}(t,t)}{u^1_{c_1}(t)}$. Notably, the discount factor R^1_P applied to the utility stream associated with remaining in stage 1 through t now takes into account the (weighted) risk of a transition into stage 2.

The value of remaining in stage 1 (integral part in equation (28)) is then offset against the value of switching to stage 2 at t (last term in (28)). This value corresponds to the stage-2 value of life, again

---

[13]The discounting term $\mu^2_{S_2} S_2$ is not present here, as a marginal change in $P(t)$ has no effect on survival S_i.

weighted with the survival probability, $P(t)$, at the time of the shock and converted into stage-1 values. Note here that the value of avoiding a deadly health shock, for which $P(\tau) = 0$ for all $\tau \in [t, T]$, exactly corresponds to the conventional stage-1 value of life.

Value of health in the first stage: $\psi_H^1(t)$

After having introduced all other valuations of health we can finally analyse the value of health in the first stage ψ_H^1, which contains multiple terms presented above (see equation (27)). In total we separate five distinct impact channels. (i) The stream of stage-1 consumer surplus, $\frac{u^1}{u_{c_1}^1}$. (ii) The value of health/survival in reducing the hazard rate and, thus, preventing the shock, $\eta_{S_1} S_1 \psi_P$. The remaining three parts capture the value of stage-1 health for reaching and living through a stage 2 life-cycle conditional on surviving a shock at age τ. Thus, all three factors are weighted with $\eta P \frac{u_{c_2}^2}{u_{c_1}^1}$, the joint probability of experiencing and surviving the health shock at τ as well as with the conversion factor. (iii) The term ψ_H^2 measures the direct value of health upon entering stage 2; (iv) the term $P_{S_1} S_1 \psi_{AS}$ captures the value of stage-1 health in enhancing acute survival following a shock; and (v) the term $B_{S_1} S_1 \psi_M$ captures the value of stage-1 health in lowering the intensity of disease/disability and, thus, morbidity at the point of the shock. We conclude by noting that the discount factor R_H^1 includes the long run impacts of survival on future mortality $\mu_{S_1}^1 S_1$ like R_H^2 as well as the (weighted) risk of entering stage 2 upon survival of a shock, $\eta P \frac{u_{c_2}^2}{u_{c_1}^1}$.

Ex-ante value of health

From an ex-ante stance the future development of the individual's health is stochastic. Thus the value of health in its general form, comparable to the well-known value of life in other contributions, should account for this uncertainty. Hence we define the (ex-ante) value of health as the expected value of the different values of health in stages 1 and 2, weighted with the corresponding probabilities. This is summarized in the following definition.

Definition 2. Assuming the existence of optimal trajectories of consumption and (the various) health investments in both stages of the individual life-cycle model (12)-(20), the (ex-ante) value of health can be defined as

$$\Psi_H(t) := Z_1(t)\psi_H^1(t) + \int_0^t Z_2(s)\psi_H^2(t, s)ds. \tag{33}$$

Following from this definition, the ex-ante value of health at age t can also be interpreted as an averaged value of health across individuals who have not experienced a shock up to age t and individuals who, at age t, have experienced different stages of disease progression following a shock experienced at an earlier age $s < t$.

Note that such a measure is useful when it comes to assessing the value of health at population level. In particular for normative purposes, it is considered unethical to distinguish individuals according to their value of life. Thus, the value of life is typically averaged across income strata, health states and often age (for an exception see Aldy and Viscusi, 2008). The ex-ante value of health would provide a value that is averaged across the possible health states at age t, including the potential health-driven inequality in earnings.

3.2 First order optimality conditions

As one crucial part of the system of optimality conditions presented in Appendix 7.2, the first-order optimality conditions give insight into the economic trade-offs between the different control variables.

Using the valuations of health presented in the previous section, we can formulate these FOCs in a compact and intuitive way.

Proposition 2. *Assume the existence of optimal trajectories of consumption and (the various) health investments in both stages of the individual life-cycle model (12)-(20) together with an interior solutions for the (various) choices of health care. The first-order optimality conditions can then be written as follows.*

$$\textbf{Stage 1:} \qquad \left[-\mu_{b_1}^1(t)\right] \cdot \psi_H^1(t) = p^b(t) \qquad (34)$$

$$\left[-\eta_{h_1}(t)\right] \cdot \psi_P(t) = p^1(t) \qquad (35)$$

$$\textbf{Stage 2:} \qquad \left[-\mu_{b_2}^2(t,s)\right] \cdot \psi_H^2(t,s) = p^b(t) \qquad (36)$$

$$\left[-f_{h_2}(t,s)\right] \cdot \psi_M(t,s) = p^2(t) \qquad (37)$$

$$\textbf{At the time of shock s:} \quad \left[-B_d(s)\right] \cdot \psi_M(s,s) + P_d(s) \cdot \psi_{AS}(s) = p^d(s) \qquad (38)$$

Consequently, the optimal allocation of health care involves that for each age/point in time t and for every possible onset of the shock s, the unit price for each type of health care equals the corresponding marginal benefit, consisting of the respective marginal effectiveness and the respective valuation of the health dimension involved. Thus, the price $p^b(t)$ for general first-stage health care, $b_1(t)$, has to equal its marginal impact on mortality $(-\mu_{b_1}^1)$ (tantamount to the depreciation rate of health) multiplied with the first-stage value of health, ψ_H^1. The interpretation for the other types of health care is analogous. Furthermore, we note that the marginal benefits of acute care, d, consist of the sum of two separate terms, as acute care does not only (potentially) increase the chances of surviving the shock but also (potentially) reduce the initial level of morbidity.

The FOCs provide immediate and intuitive information on the (relative) drivers of health care choices. Thus, the individual will demand a higher quantity of health care if it is more effective, if it has a higher value or it has a lower price. Note that the FOCs can also be read as reflecting the optimal trade-off between the different types of health care and consumption. Here, the left-hand side (LHS) of each condition reflects the marginal rate of substitution between the particular type of health care and consumption, whereas the right-hand side (RHS) gives the price ratio, with the price of the consumption good normalized to one. Thus, a higher price for health care would either have to be offset by an increase in effectiveness and/or an increase in the value of this care, the latter being reflective of greater need. The system of FOCs also allows to trace the allocation across different types of health care in light of relative effectiveness and relative valuations. Consider e.g. the condition

$$\frac{-B_d(s)}{-f_{h_2}(s,s)} + \frac{P_d(s) \cdot \psi_{AS}(s)}{[-f_{h_2}(s,s)] \cdot \psi_M(s,s)} = \frac{p^d(s)}{p^2(s)}$$

as implied by the FOCs (37) and (38). To understand the intuition, assume first a setting in which the health shock does not impose a risk to survival, implying there is no role for acute care in enhancing acute survival, i.e. $P = 1$ and $P_d(s) = 0$, as would be the case e.g. with cancer. In such a case the condition would tell us that a higher price for acute care (e.g. immediate surgery) as opposed to chronic care (e.g. pharmaceutical therapy) at the point of the onset of the disease would need to be offset by greater effectiveness of acute care in containing the disease (i.e. the progression of cancer). If acute care also bears on survival, the price ratio does not only reflect differences in the effectiveness of care in curbing morbidity but, in addition, the marginal rate of substitution between the (valued) change in acute survival for the acute treatment and the change in morbidity for the chronic treatment. Thus, a higher price for acute care at the point of the shock is supported to the extent that it not only reduces

initial morbidity but also improves survival chances by $P_d(s) > 0$. Note that this argument extends to settings, where the provision of acute care, e.g. cancer surgery, may carry a risk to survival, such that $P_d(s) < 0$. Here, the survival risk lowers the willingness to pay for acute care, implying that its utilization is lower relative to chronic care for a given price ratio and/or a given level of utilization is supported only at a lower price.[14] Similar trade-offs between other dimensions of health care, e.g. between general health care and preventive health care in stage 1 or between preventive health care in stage 1 and chronic health care in stage 2, can be constructed by appropriate combination of the relevant FOCs.

Isolating the value terms on the LHS of the FOCs, as e.g. in $\psi_P(t) = \frac{p^1(t)}{[-\eta_{h_1}(t)]}$ for preventive care or $\psi_M(t,s) = \frac{p^2(t)}{[-f_{h_2}(t,s)]}$ for chronic care allows us to interpret the FOCs in terms of the underlying dimensions of health as a final good rather than health care as an intermediate good. Thus, we find that for an optimal allocation, the value of prevention should equal the effective price of prevention, as given by the price of preventive health care adjusted for its effectiveness in curbing the arrival rate of a shock. Similarly, the value of (reducing) morbidity should equal the effective price of lowering morbidity, as given by the price of chronic health care adjusted for its effectiveness in curbing or reverting the progression of the disease. Analogous expressions can be derived for other dimensions of health care. Following Frankovic et al. (2020) who undertake this analysis in the context of survival, we can infer that medical progress that raises the effectiveness of a certain type of health care, prevention say, may be associated with a decline in the value of this dimension of health. Notably, this reflects the greater consumption of preventive health care leading to reductions in the health risk to a level for which any further reduction is less valuable (or in micro-economic terms, the decline in the effective price of prevention relative to consumption is associated with a decline in the marginal rate of substitution between prevention and consumption). We conclude with the following empirical observation: In many practical settings, the valuation of different types of health is difficult to observe. A revealed preference argument would then suggest that the willingness to pay equals the effectiveness-adjusted price of health care, which can be calculated on the basis of observable nominal prices of health care and scientific evidence on medical effectiveness.

Given the multi-dimensionality of health, we can extend the previous argument to examine the relative valuation of different types of health care. Drawing on the FOCS (36) and (37), for instance, we can write

$$\frac{\psi_M(t,s)}{\psi_H^2(t,s)} = \frac{p^2(t)/\left[-\mu_{b_2}^2(t,s)\right]}{p^b(t)/\left[-f_{h_2}(t,s)\right]}.$$

Thus, the value of lower morbidity relative to the value of survival can be understood to reflect the relationship between the effective price of morbidity relative to the effective price of survival. This has important repercussions from a practical point of view, where many studies in medical evaluation seek to probe into consumer/patient assessment of improvements to the health-related quality of life relative to survival (see e.g. Rowen and Brazier, 2011:for a survey). While such studies are often carried out in the context of an abstract decision-framework, we note that when resulting from market assessments (e.g. when asking expert physicians to state these trade-offs), differences (across individuals, populations or over time) in the relative valuations of health may be as reflective of differences in the underlying preferences as of differences in the relative effectiveness of the different types of health care. Here, again revealed preference analysis may be brought to bear to deduce relative valuations.

---

[14]Recall here our assumption that all controls exhibit diminishing returns.

4 Numerical analysis: an application to cancer

Within the next two sections, we illustrate insights from our model by conducting a calibration exercise and calculating optimal individual behaviour numerically. To provide a concrete context, we study the impact of a (potential) cancer diagnosis on the life-cycle allocation of health care and consumption.

In order to avoid adding the substantial complexity that is involved with the hidden progression of cancer up to the point of a diagnosis, we assume that (i) a diagnosis at age/time s coincides with the onset of cancer. Thus, we disregard the building-up of the cancer-stock prior to s, which is a reasonable assumption when assuming that a diagnosis coincides with the point of first symptoms. As there is no clear empirical correlation[15] between the stage of cancer at the time of diagnosis and the general health status (or the respective age), we suppose (ii) that the individual enters with a constant positive disease stock $E(s,s) = \phi_0 > 0$ independent of its stock of health, $S_1(s)$, at the point of diagnosis. While it is further reasonable to assume that (iii) a diagnosis at s is not associated with an acute risk to survival, such that $P(S_1, d) \equiv 1$, we also assume for the purpose of this analysis that (iv) there is no role of acute care at the time of diagnosis, i.e. $d(s) \equiv 0$. This is justified when assuming that the diagnosis is early enough to rule out a significant role for surgery. To avoid the complexity involved with the timing of diagnosis, we abstract from (v) screening measures. Finally, in order to limit the number of states and controls we assume (vi) that healthy and unhealthy behaviours such as smoking, drinking, eating habits or exercising that affect the risk of cancer are captured by "'generic"' investments in health, b_1. Note that these feed into the hazard rate $\eta(t, S_1(t), h_1(t))$ through their impact on the health stock/survival and thereby generate health benefits beyond cancer, a reasonable assumption. We, thus, assume (vii) that there is no further role for specific cancer prevention, such that $h_1(t) \equiv 0$. We also assume that age does not have a direct bearing on the hazard rate, implying that we work with $\eta(S_1(t))$.

Altogether, we obtain a slightly reduced model, which nevertheless enables us to analyze the impacts of cancer on the utilization of health care before and after the diagnosis, including a detailed characterization of the utilization pattern of chronic care, which from now on we refer to as cancer care.

4.1 Data sources

We calibrate our model to the data of an individual in the United States in the 2010s, using a number of data sources for the input parameters.[16] We assume that an individual receives earnings from age 20, which is also the age at which the individual begins to make life-cycle decisions.

Input parameters

One of the few fully exogenous inputs is the base earnings profile $\{w_1(t) \mid t \in [20, 110]\}$, which in our model we assume to be the average earnings age-profile for the US in 2011, as taken from the National Transfer Accounts (NTA) database[17]. We account for the fact, that there are no public pensions in our model and all health care expenditures are paid out of pocket, by adding the net health and pension transfer profiles (which are also contained in the NTA database) to the base working income to obtain a profile of disposable income.

We abstain from trying to fit the (non-health) consumption profile from the NTA database, as it shows the typical hump-shape over the life-cycle. Hansen and İmrohoroğlu (2008) have shown that

---

[15] Goodwin et al. (1986) find a positive correlation between the stage of cancer and the age at diagnosis for some types of cancer, while for other types this correlation turns out to be negative. Even for breast cancer alone the results turn out to be unclear. While Satariano et al. (1986) and Mandelblatt et al. (1991) find a positive relationship between age and severity, Yancik et al. (2001) find it to be non-significant.

[16] Due to limitations in data availability we are unable to use all data from the same year.

[17] www.ntaccounts.org, see Lee and Mason (2011) and United Nations (2013) for details.

such a pattern can only be explained consistently within a life-cycle model when assuming that annuity markets are absent (or severely incomplete).[18] However, we quite deliberately prefer to model a complete annuity market for the explicit purpose of establishing a flat consumption profile as a benchmark against which to assess the impact of health shocks on consumption and identify the underlying channels. Here, an imperfect annuity market would obscure some of the effects. This notwithstanding, we aim for a level of consumption that is in line with the average consumption over the life-cycle.[19]

In further pursuit of establishing a flat benchmark consumption profile, we opted to set two of the other exogenous inputs, the interest rate $r(t)$ and the discount rate ρ, both equal to 3%. If we eliminated the possibility of a health shock in our model, this would then lead to a completely flat consumption profile, with the annuity rate perfectly covering the mortality risk in this scenario. Therefore $r(t) = \rho$ enables us to directly identify the impact of the existence of a potential health shock in the consumption profile before and after the cancer diagnosis.

Calibration goals

We use health expenditure data from the NTA-Database (available for the year 2011) and combine it with age-specific information about the share of cancer specific health care expenditures from `healthdata.org`[20] to construct age profiles for (i) the general (non cancer) health expenditure and (ii) cancer specific expenditures. We compare these against the average general health expenditure and average cancer care expenditure profiles generated by our model and aim to match the latter with their empirical counterparts.[21]

Furthermore, we use the age-specific mortality profile for the US in 2011 from the human mortality database[22] as a basis for establishing survival profiles. Using the mortality rate directly to calculate the corresponding survival profile within our model, we obtain the equivalent of the average survival $\overline{S}(t) := Z_1(t)S_1(t) + \int_0^t Z_2(s)S_2(t,s)ds$ in our model. To obtain the appropriate data against which to compare first-stage survival in our model, we take age-specific cancer mortality rates from the SEER-database[23] and subtract them from the corresponding general mortality rates. From the resulting age-profile of non-cancer mortality we construct a cancer-free survival profile as the appropriate comparison for the S_1 profile in the model.

We then employ the cancer-free survival profile from the data together with age-specific rates of cancer incidence[24], again taken from the SEER-database, to calibrate the hazard rate $\eta(S_1)$, which we assume to depend only on the survival state. Finally, we use information from the SEER-database about cancer-specific survival depending on the duration after the cancer diagnosis to calibrate the cancer specific mortality rates for four different (rough) age-groups over the first ten years after the diagnosis. For further details about the estimation and calibration strategy for the general and cancer-specific mortality rates (μ^b and μ^m) and the cancer incidence rate (η) we refer to appendix 7.6.

---

[18]From the Euler-equation (42) we see that the absence of annuities ($\bar{\mu} = 0$) implies that increasing mortality μ^i with age ultimately leads to a decline in consumption. Annuities eliminate this risk (or, as in our case, even overcompensate it), so there is no significant force anymore that would shift consumption to younger ages, implying a hump-shaped pattern cannot be obtained.

[19]The US exhibit a significant life-cycle deficit, implying that life-time consumption is significantly higher than life-time income. Thus, the raw data contradicts our assumption of zero assets at the end of life. To fulfil the equivalent condition of life-time consumption and life-time income being equal, we have decided to raise income by 15% for our calibration.

[20]Institute for Health Metrics and Evaluation (IHME) (2016) (Accessed 2020-01-13)

[21]We define average health expenditure at age t as $\bar{b}(t) := Z_1(t)b_1(t) + \int_0^t Z_2(s)b_2(t,s)ds$ and, similarly, average cancer care expenditure as $\overline{h_2}(t) := \int_0^t Z_2(s)h_2(t,s)ds$.

[22]*Human Mortality Database* (accessed 2019-10-04)

[23]*SEER*Explorer: An interactive website for SEER cancer statistics [Internet]*

[24]We take average incidence rates between 2012 and 2016.

4.2 Remarks on the annuity market

In most models annuity markets are assumed to be either absent or complete with the equilibrium annuity rate being equal to the mortality rate. Our setting is more complex, as it contains two mortality rates reflecting different health regimes: one with cancer and one without. Thus the rate of return for annuities $\bar{\mu}(t)$ crucially depends on the way the annuity market is structured.

In this calibration exercise we assume that the insurer has no information about whether or when an individual has been diagnosed with cancer. As a result the annuity rate at time t turns out to be the expected (or averaged in a population context) mortality rate, which can be calculated as

$$\bar{\mu}(t) = \frac{Z_1(t)(-\dot{S}_1(t)) + \int_0^t Z_2(s)(-\dot{S}_2(t,s))ds}{Z_1(t)S_1(t) + \int_0^t Z_2(s)S_2(t,s)ds} = \frac{Z_1 S_1 \mu^1 + \int_0^t Z_2 S_2 \mu^2 ds}{Z_1 S_1 + \int_0^t Z_2 S_2 ds}. \tag{39}$$

The nominator adds up all deaths across the two groups with and without cancer and relates them to the total population in the denominator.[25]

4.3 Solution strategy

The numerical solution of a two-stage optimal control problem with random switching time is a far-from-trivial problem. As already indicated in the problem introduction the transformation into a vintage-structured optimal control problem is the first step in the solution process. The transformation (see Wrzaczek et al. (2020) for further details) allows our numerical strategy to rest on an existing gradient-based optimization algorithm, as described by Veliov (2003). However, particular features of our model, e.g. the asset end-point constraints and the balanced annuity market, as well as variation by orders of magnitude across some of the gradients of the controls[26] made further non-trivial adaptations to the numerical method necessary.

4.4 Functional specifications

To generate a numerical solution of our model, as summarized in equations (12) - (20), we need to specify the following functional forms.

Utility

Following Hall and Jones (2007) and many others, we employ an adjusted CRRA-utility function for the instantaneous utility from consumption:

$$u^1(c) = \frac{c^{1-\sigma}}{1-\sigma} + \bar{u}, \qquad 0 < \sigma \neq 1 \qquad .$$

Here, we assume $\bar{u}$ to be a sufficiently large constant, which guarantees $u^1(c) > 0$ for all reasonable consumption levels. For the second period, we assume that the cancer stock affects the utility of

---

[25]One of multiple alternative assumptions would be that the insurer has perfect information about a cancer diagnosis. In this case the annuity rate would be equal to the mortality rate for each individual. Investigating the implications of different annuity markets is an interesting task in its own right that goes beyond the scope of the present paper.

[26]The gradients of stage 1 are initially weighted way higher than that of stage 2, what results in the effects of the gradient in the second stage being levered out. Without adjustments this would imply a faster convergences of the controls in the first stage and could lead to premature termination of the algorithm as either (i) the gradients of the second stage are directly below a reasonable numerical threshold or (ii) the improvements in the objective value for adjustment of second-stage-controls are swallowed by numerical inaccuracies due to the small directional gradient step combined with their resp. small weights in the objective function.

consumption in a multiplicative form:

$$u^2(c, E) = u^1(c)v(E) = u^1(c) \exp \left\{ \kappa_0 \cdot E^{\kappa_1} \right\},$$

with $v(E) \in [0, 1]$ and $v'(E) < 0$ for $\kappa_0 < 0$. The mixed derivative $u_{cE}^2 = u_c^1(c)v'(E)$ is negative, implying that a higher cancer stock reduces the marginal utility of (non health care) consumption. Note that this is in line with empirical evidence (see Finkelstein et al. (2013)).

Mortality

For this numerical presentation, we assume that the non-cancer mortality rate μ^b does not depend on the state of survival, enabling us to better disentangle cancer specific mortality in the second stage. The parameters γ_i and α_i will be calibrated to reproduce the survival profile without a cancer diagnosis (see appendix 7.6).

$$\mu^b(t, S_1, b_1) = \mu^b(t, b_1) = g(t)b_1^{\varepsilon(t)} \qquad g(t) = \exp \left\{ \gamma_0 + \gamma_1 t + \gamma_2 t^2 + \gamma_3 t^3 \right\}$$
$$\varepsilon(t) = \alpha_0 + \alpha_1 t \quad (< 0)$$

Cancer incidence

Comparing the survival data adjusted for cancer mortality and the cancer incidence rates from the SEER-database we find that the function in equation (40) delivers the best fit, while still using just three parameters.

$$\eta(t, S_1) = \eta(S_1) = \frac{\beta_0}{1 + \beta_1 \left(\frac{1 - S_1}{S_1} \right)^{\beta_2}} \tag{40}$$

Cancer stock

For the progression of the cancer stock, we decided to draw on the biological development and spread of cancerous cells. Talkington and Durrett (2015) present multiple established processes for the untreated spread of cancerous cells. We use the simplest version (which is sufficient for our purposes) and hence assume that the number of cancerous cells grows over time at the constant rate δ_0. Following the work of Heuser et al. (1979), who estimated the average doubling time for breast cancer cells to be 327 days, we propose a rate of $\delta_0 = 0.774$ (per year). In a normative step, we set the initial number of cancerous cells when diagnosed constant to 1, i.e. $B(S, d) = \phi_0 = 1.0$.

For the effects of cancer care we lean on the development of health deficits introduced by Dalgaard and Strulik (2014). Depending on the intensity of cancer care h_2 (which exhibits diminishing returns modelled with $\delta_2 \in (0, 1)$) and the available technology (captured by δ_1), we propose that the growth can be slowed down or also turned negative to hopefully eradicate cancerous cells in the long-run.

$$f(t, s, E, h_2) = \delta_0 E - \delta_1 h_2^{\delta_2} E^{\delta_3}.$$

However, in contrast to Dalgaard and Strulik (2014) we propose that the effectiveness of cancer care also increases with the number of cancerous cells and is therefore multiplied by E^{δ_3} with $\delta_3 > 0$.

Cancer specific mortality

For the cancer specific mortality function $\mu^m(t, s, E)$, we eliminate the dependence on age t and assume that the whole mortality process is driven by the number of cancerous cells E and the age at which cancer was diagnosed. As we find in the data, mortality increases with the age at diagnosis. Hence we

propose the following functional form:

$$\mu^m(t, s, E) = \mu^m(s, E) = \psi_0 \cdot E \cdot \exp\left\{\psi_1 \cdot \left(\frac{s}{T}\right)^{\psi_2}\right\} \tag{41}$$

We estimate these parameters ψ_i using the cancer incidence and survival data as described in appendix 7.6.

Prices for health care

After the abstraction from preventive and acute care for this numerical exercise, there remain two prices for health care goods and services (each expressed in units of the consumption good):

- *Price of general health care p^b*
  As we will detail in the appendix, we write mortality directly as a function of health care expenditures, $p^b b_1$ for the purpose of this calibration. We can therefore set the price equal to one $\implies p^b = 1$.

- *Price of cancer care p^2*
  Due to the multiple types and combinations of cancer treatment, it turns out to be most practical to also measure cancer care directly in its monetary units. Hence we have decided to set the price for cancer care equal to one $\implies p^2 = 1$.

Table 2 summarizes the functional forms and parameter choices. We wish to stress, that we did not attempt to obtain a full calibration of the model, since this would have required the introduction of further state and control variables to represent the involved nature of cancer risk, prevention, development and treatment in the first place. Instead we aimed for a reduced model formulation, which nevertheless replicates key patterns of cancer progression and cancer care and thereby enables exceptional new insights into the behavioural patterns regarding health care (general and cancer-specific) and consumption.

Note that some parameter values shown in Table 2 are not the result of an automated calibration process, but are manually chosen to improve the calibration fit. The high complexity of the model and intricacy of the solution process in addition to relatively long computation times for one set of parameters did not allow us to conduct standard calibration methods. Hence the parameters $\delta_1, \delta_2, \delta_3$ and κ_0, κ_1 are not chosen through a process minimizing a strictly defined quality of calibration criterion (e.g. minimizing the maximum absolute difference in the cancer mortality rates), but are manually set to obtain a qualitatively appropriate fit of the several profiles presented.

$u^1(c) = \frac{c^{1-\sigma}}{1-\sigma} + \overline{u}$	σ	1.13^+
	$\overline{u}$	11.355^+
$\varepsilon(t) = \alpha_0 + \alpha_1 \cdot t$	α_0	-0.2700^*
	α_1	0.0667^*
$g(t) = e^{\gamma_0 + \gamma_1 t + \gamma_2 t^2 + \gamma_3 t^3}$	γ_0	-6.690^*
	γ_1	-0.982^*
$\left(\mu^{\mathrm{b}}(\mathrm{t},\mathrm{b}) = \mu_0(\mathrm{t}) \cdot \mathrm{g}(\mathrm{t}) \cdot \mathrm{b}^{\varepsilon(\mathrm{t})} \right)$	γ_2	15.846^*
	γ_3	-7.560^*
Interest rate	r	0.03
Discount rate	ρ	0.03
$\eta(t,S) = \frac{\beta_0}{1 + \beta_1 (\frac{1-S_1}{S_1})^{\beta_2}}$	β_0	0.0243^{**}
	β_1	0.0087^{**}
	β_2	-2.020^{**}
$B(S,d) = \phi_0$	ϕ_0	1.0
$f(t,s,E,h_2) = \delta_0 E - \delta_1 h_2^{\delta_2} E^{\delta_3}$	δ_0	0.7740^+
	δ_1	2.0
	δ_2	0.03
	δ_3	1.3
$\mu^m(t,s,E) = \psi_0 \cdot E \cdot \exp\left\{ \psi_1 \cdot \left(\frac{s}{T}\right)^{\psi_2} \right\}$	ψ_0	0.1705^{**}
	ψ_1	1.8475^{**}
	ψ_2	1.5574^{**}
$v(E) = \exp\left\{ \kappa_0 \cdot E^{\kappa_1} \right\}$	κ_0	$\ln(0.7)$
	κ_1	1.2

Table 2: Summary of functional specifications and parameters in the model. Parameters indicated by ** are estimated before the solution process using only the empirical data. Parameters indicated with * result from the calibration process within the solution process. Parameters marked with $^+$ are chosen from the literature. Parameters without indications are either educated guesses (ρ, r, κ_i) or are manually chosen to improve the calibration ($\delta_1, \delta_2, \delta_3$).

5 Numerical results

Turning to the numerical results, we will first show to which degree we were able to replicate the data and thereby meet our calibration goals (Section 5.1). In a next step we discuss and compare the profiles of consumption before and after a cancer diagnosis and use the Euler-equations to identify how these profiles are impacted by the different aspects of the (potential) diagnosis (Section 5.2). Section 5.3 presents the profiles of general health investments and cancer care together with the resulting developments of the health state in both stages and the disease stock capturing the progression of cancer. In so doing, we also examine the numerical evaluations of the various valuations of health presented in Section 3.1 and their respective decomposition. Finally Section 5.4 presents the Euler-type equations for the different types of health care both in the most general formulation as well as their numerical evaluations for the optimal solution. These calculations allow us (as for the consumption profiles) to pin down the several impacts a potential diagnosis has on the individual's decision making. Throughout the whole section we summarise the key outcomes in corollaries 1–6.

5.1 Expenditure and survival patterns: Model vs. data

Figures 1 and 2 compare for a range of variables the outcomes of our model with data from the US, as discussed in the sections above. More specifically, Figure 1 summarizes the different types of expenditures in our model. The upper left panel shows the expected values for general (non-cancer) health expenditures as well as for the spending on cancer care. Overall, the expected expenditure profiles follow the data reasonably well and the corresponding spending shares, shown in the upper right panel, paint a similar picture. As was discussed before, for the purpose of better isolating transmission channels, we

have quite deliberately opted for a model with an annuity market. For this reason, our model is not suited to replicate the US consumption profile from the data, but the lower left panel shows that the expected consumption profile from the model meets our goal of matching average consumption. Finally, the lower right panel illustrates the composition of the age-profiles of income (labour and transfers).
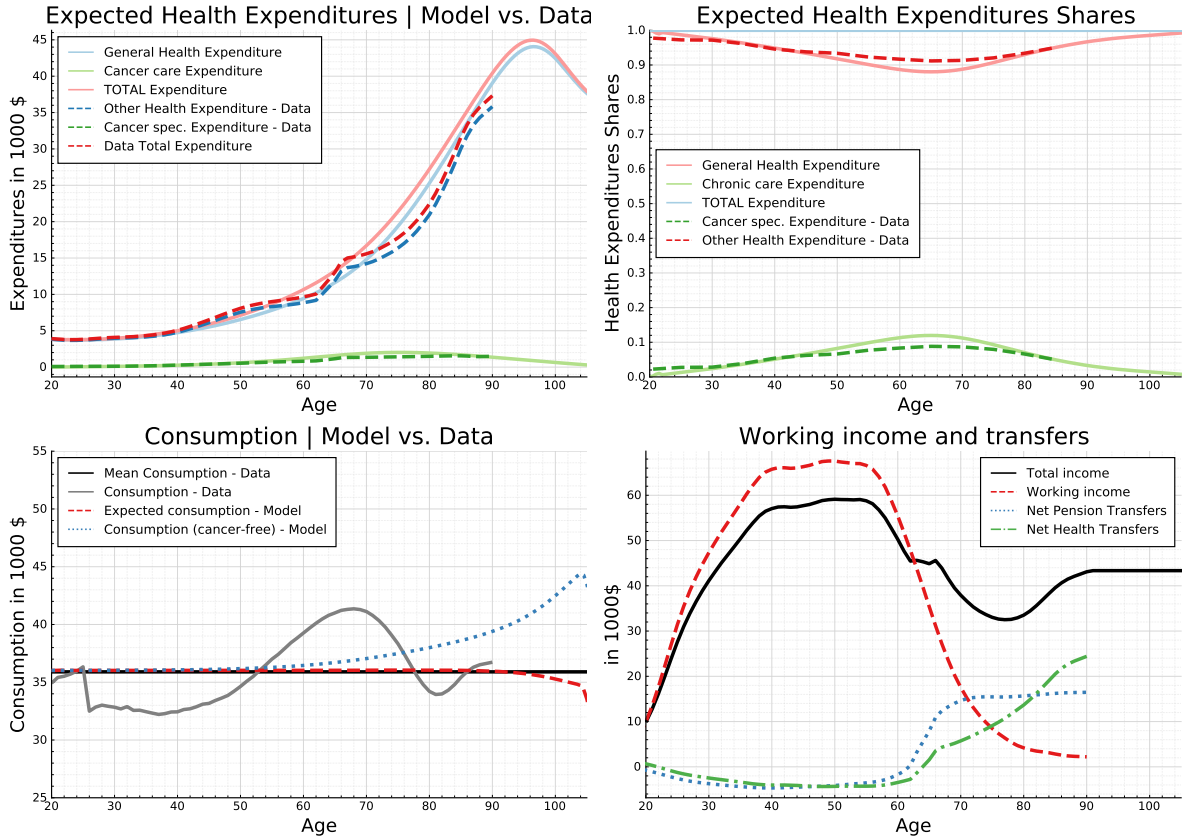

Figure 1: Calibration results: Comparison between expenditures predicted by the model and expenditure data

In Figure 2 we focus on the fit of the survival and mortality data. The upper left panel shows that we meet remarkably well both average survival, $\overline{S}(t) = Z_1(t)S_1(t) + \int_0^t Z_2(s,t)S_2(s,t)ds$ and survival conditional on remaining cancer-free, $S_1(t)$. Furthermore we can see that the cancer-free survival, $Z_1(t)$, profile of our model meets the corresponding profile derived from the data relatively well, even if slightly underestimating it.[27] The lower left panel shows cancer-specific mortality data over the first 10 years after a diagnosis for 4 different age-groups. Note that we plot the logarithm of the mortality rate to account for the strong differences in magnitude of these values. The model and data profiles match very well despite some discrepancy for the youngest and oldest age groups around the point of diagnosis. Finally, the lower right panel shows the total mortality rates before and after a cancer diagnosis.[28] The good fit of mortality before a cancer diagnosis is reflecting the fit of the survival profiles.

---

[27]This could be due to numerical inaccuracies, as data on cancer incidence was only available in 5-year intervals.

[28]We again took a logarithmic scale as mortality rates towards the end of life are in a different order of magnitude compared to mortality rates in younger ages.
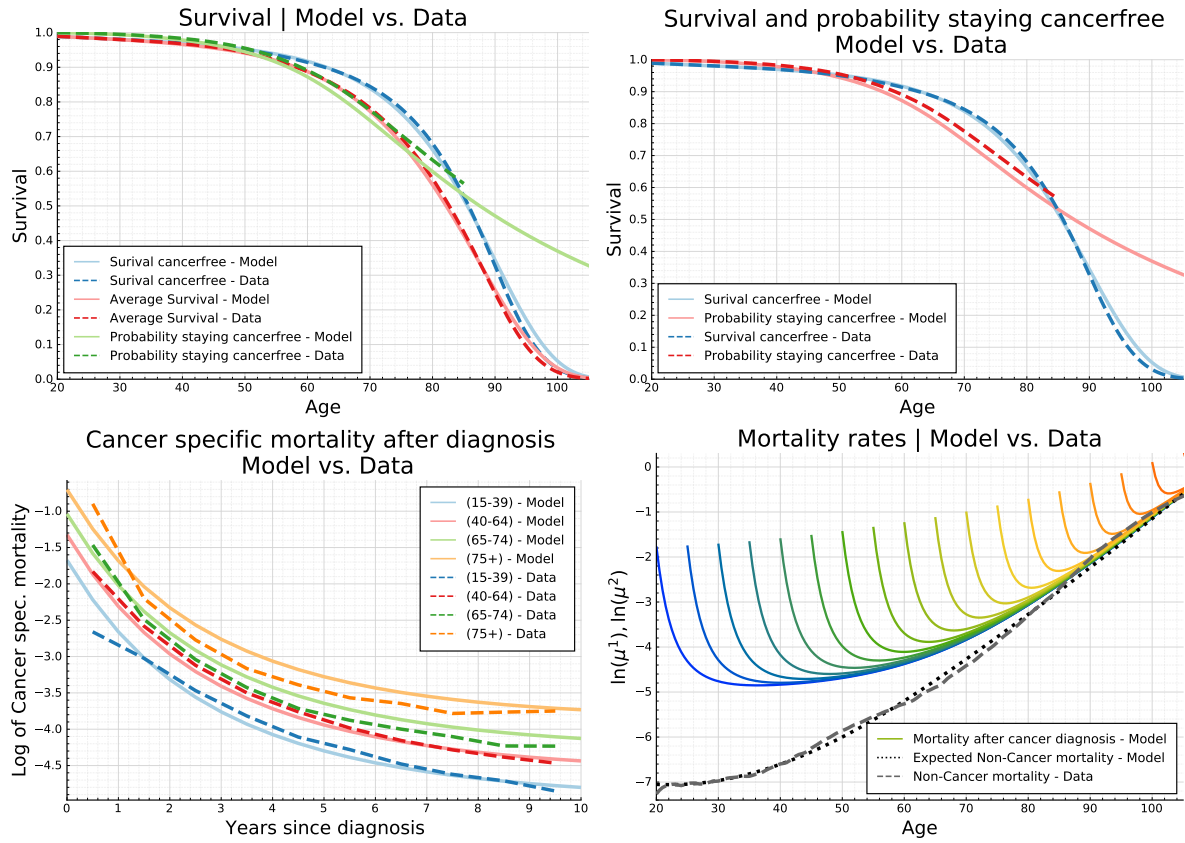
Figure 2: Calibration results: Comparison of survival and mortality profiles between the model and the data

5.2 Consumption profiles

Consider now, the consumption profiles, as plotted in Figure 3. The grey dashed line depicts the consumption profile in the absence of cancer. The solid lines represent the consumption profile following a cancer diagnosis at a given age, where for illustrative purposes we move the age at diagnosis in five year steps. Finally, the black dotted line illustrates the consumption level chosen immediately after a diagnosis, the distance between the dotted and dashed line representing the instantaneous consumption change following a cancer diagnosis.

The consumption profile in the absence of cancer is relatively flat and only increases towards the end of life from 35000$ per year up to around 45000$ in the very late stages of life. In contrast, the optimal choice of consumption following a diagnosis reveals strong fluctuations. Right after the diagnosis individuals decrease their consumption if the diagnosis occurs relatively early in life. This immediate drop is followed by an increase over the next two years which reduces the gap to consumption levels of cancer-free individuals of the same age to insignificant levels. This increase is followed by a slightly U-shaped pattern over the remaining life-course (partially increasing the consumption difference again). For early diagnosed individuals (approximately before age 70 in our setting) the levels of consumption vary between about 28000$ and 35000$ per year. If the diagnosis occurs after the age of 75, individuals boost their consumption directly after the cancer diagnosis and the consumption path follows a constant decline over the remaining life-course. The upward jump is increasing with the age at diagnosis and reaches up to levels of 55000$ per year. Towards the end of life consumption decreases to 26000$ per year.
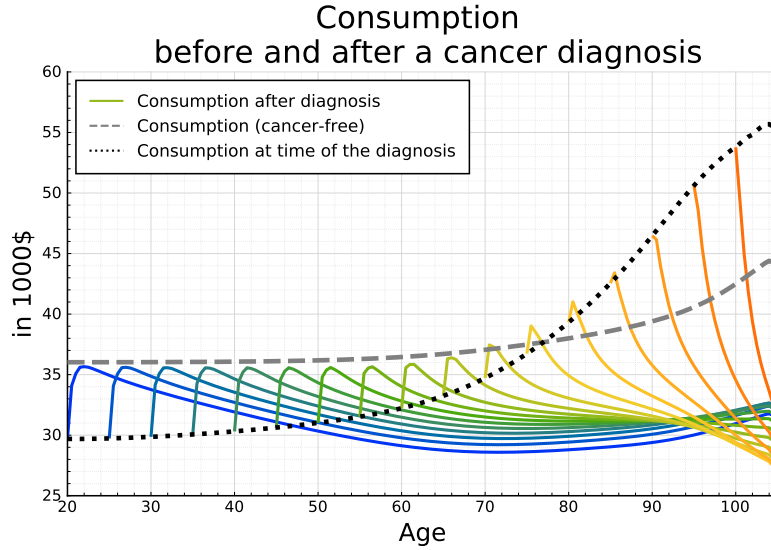
Figure 3: Consumption profiles in the first and in the second stage

To understand the widely different consumption patterns of cancer-free, early, and late diagnosed individuals, we derive the Euler-type equations (42) and (43) for the dynamics of consumption.[29] For the derivations we use the first-order optimality conditions for consumption together with the dynamics of the adjoint variables of assets (see appendix 7.5).

$$\frac{\dot{c_1}}{c_1} = \frac{u^1_{c_1}}{-u^1_{c_1 c_1} c_1} \left[r - \rho + \bar{\mu} - \mu^1 - \eta + \eta P(S_1, d) \frac{u^2_{c_2}}{u^1_{c_1}} \right], \tag{42}$$

$$\frac{\dot{c_2}}{c_2} = \frac{u^2_{c_2}}{-u^2_{c_2 c_2} c_2} \left[r - \rho + \bar{\mu} - \mu^2 + \frac{u^2_{c_2 E}}{u^2_{c_2}} f \right] \tag{43}$$

For both stages we are able to derive the elements of a standard Euler equation in the presence of partially insured mortality risk. These involve the difference between the interest rate, r, and discount rate, ρ, which drops out in our calibration. Additionally, the mortality risk, μ^i, shifts consumption toward younger ages in both stages to the extent it is not offset by annuity returns, $\bar{\mu}$.[30] In the intuitive case of higher mortality following a cancer diagnosis (see lower right panel of Figure 2) and under our assumption of a non-state contingent annuity premium $\bar{\mu}(t)$, as defined in equation (39), it likely holds that:

$$\mu^1(t, S_1, b_1) \leq \bar{\mu}(t) \leq \mu^2(t, s, S_2(t, s), b_2(t, s), E(t, s))$$

Hence the annuity rate overcompensates the mortality rate before a diagnosis leading to a deferral of consumption, while it only partially compensates the force of mortality after a diagnosis leading to an advancement of consumption immediately after a shock. However, as Figure 5 will show for the case of cancer, the second consideration may not hold over the whole life-cycle. When the cancer stock is reduced dramatically after the first few years the total mortality rate is returning back to the cancer-free mortality levels, while the average mortality $\bar{\mu}$ is still high, due to the higher total mortality of more

---

[29]Note that all Euler equations (here and in the following sections) are presented in their general form, and the interpretations also apply to a general health shock. Here, we discuss them in the context of the numerical results based on our calibration for cancer and discuss which terms are non-existent in our specified set-up.

[30]As already indicated in Section 4 in absence of the risk of a cancer diagnosis this would imply completely flat (and identical) consumption profiles before and after the shock. Hence we can attribute all variation in consumption over the life-course to different aspects of the potential health shock/cancer diagnosis.

recently diagnosed cancer patients. Hence, survivors of cancer have an incentive yet again to defer consumption in their late years.

The remaining factors in equations (42) and (43) differ and we analyse them separately along some numerical illustrations in Figures 4 and 5. These figures show the different terms $F_i(t)$ in the Euler-equations aggregated from the beginning of the life-cycle (resp. from the time of diagnosis) up to each point in time t. Considering, for instance, stage-1 consumption $c_1(t)$ and defining $\frac{\dot{c}_1(t)}{c_1(t)} =: \sum F_i(t)$ we can write $c_1(t) = c_1(t_0) \exp\{\int_{t_0}^t \sum F_i(s) ds\}$ and, thus, $\log(c_1(t)) - \log(c_1(t_0)) = \sum \int_{t_0}^t F_i(s) ds$. Hence, the aggregated terms are direct expressions of the (log) difference between current consumption $c_1(t)$ and initial consumption $c_1(t_0)$. We apply the same reasoning to all other control variables in the subsequent sections.

First stage

In the first stage the uninsured risk of suffering a health shock, η, shifts consumption towards younger ages, but in contrast to the mortality risk, there is an additional effect associated with a health shock. The term $\eta P \frac{u_{c_2}^2}{u_{c_1}^1}$ shows that the desire for consumption smoothing tends to shift consumption back to later ages. The term contains the product of the hazard rate, η, the probability of surviving the shock, P, and the ratio of the marginal utilities relating to the consumption levels $c_2(t,t)$ and $c_1(t)$ immediately after and before the shock. The rationale behind this shift lies in the incentive to accumulate precautionary savings which can be used to maintain a given level of consumption following the shock.

Three thought experiments can illustrate how the two offsetting effects combine. If the health shock has no effect on the household at all, $P = 1$ and $c_2(t,t) = c_1(t)$ hold. As a result, $\eta = \eta P(S_1, d) \frac{u_{c_2}^2}{u_{c_1}^1}$ and, as expected, the two terms cancel out in the c_1-dynamics. For a second experiment, we propose a health shock with no chance of survival ($P = 0$). Here, the hazard rate η has the same impact as the mortality rate μ^1 as there is no more desire for consumption smoothing. Finally, suppose that a health shock lowers the marginal utility of consumption such that $u_{c_2}^2 < u_{c_1}^1$ for $c_2(t,t) = c_1(t)$. For any $P \leq 1$ we then have that $\eta > \eta P(S_1, d) \frac{u_{c_2}^2}{u_{c_1}^1}$, implying that the incentive to advance consumption dominates the precautionary savings motive. However, our numerical analysis for the case of cancer shows that in general a drop in consumption after the diagnosis is possible (e.g. for early diagnosed individuals). In combination with the negative impact of the cancer stock on the marginal utility, this implies ambiguity for $u_{c_2}^2 \gtreqless u_{c_1}^1$. Hence we are unable to make general statements about the combined effect and the desire for precautionary saving could potentially be more pronounced than the incentive to advance consumption.

In Figure 4, we illustrate the different parts of equation (42) aggregated up to each time-point t. We can see that the desire to advance consumption resulting from the term $-\eta$ (dashed purple line) dominates the incentive for precautionary savings, $\eta P(S_1, d) \frac{u_{c_2}^2}{u_{c_1}^1}$ (dash-dotted green line). However the difference between the two only becomes significant after the ages around 50, but strongly increases towards the end of life. As $P = 1$, the combined effect can be simplified to $\eta \cdot (u_{c_2}^2/u_{c_1}^1 - 1)$. Hence the increasing cancer incidence rate η predominantly defines the magnitude of the combined effect, while its sign is determined by the divergence of the marginal utility ratio $\frac{u_{c_2}^2}{u_{c_1}^1}$ from 1. Thus the combined effect being non-positive for all ages indicates, that the impact of the cancer stock on the marginal utility still implies $u_{c_2}^2 < u_{c_1}^1$, although $c_2(t,t) < c_1(t)$ for diagnoses in younger ages. Consequently, desire for precautionary savings is curbed compared to the health shock risk and we would obtain a tendency for consumption to decline over the cancer-free life-cycle (from this combined effect alone). However, as Figure 4 shows, the growing saving incentive that arises from the increasing gap between the annuity return and mortality risk for individuals who remain cancer-free (dotted blue line) overcompensates this

effect and leads to an overall increasing consumption profile (black solid line, which we also have already seen in Figure 3).[31] We summarise our findings in Corollary 1.
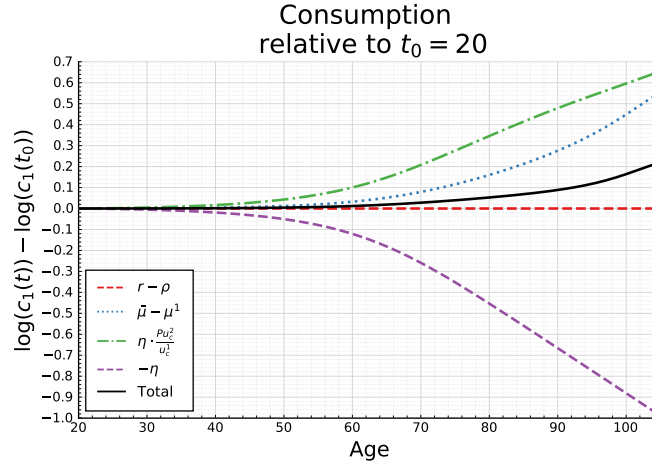

Figure 4: Illustrations of the impacts of the different parts of the Euler equation for consumption before the diagnosis

Corollary 1 (Consumption before diagnosis). *Compared to an individual not facing the risk of a cancer diagnosis, its introduction implies an increasing consumption profile for cancer free individuals. This results from the incentive to defer consumption (as the annuity rate overcompensates their mortality risk) dominating the desire to advance consumption (as the marginal utility of consumption is higher when cancer-free).*

Second stage

In the second stage there is no risk of a further shock, but the individual takes into account the development of the shock-specific deficit stock E, i.e. the cancer stock. This leads to the term $\frac{u^2_{c_2 E}}{u^2_{c_2}} f$ in the growth rate of consumption in equation (43). In the more intuitive case[32] of $u^2_{c_2 E} < 0$, where the marginal utility of consumption declines with the deficit stock, two possible scenarios can arise: (i) consumption is shifted towards earlier ages where the marginal utility is still high if the deficit stock increases over time $f > 0$, or (ii) consumption is deferred if the deficit stock decreases after the shock $f < 0$.

For illustration, we consider in Figure 5 the consumption profiles following a diagnosis at ages 30 and 70, respectively. As we have already seen (lower right panel in Figure 2), the mortality risk after the diagnosis is significantly increased for several years and consequently is only partially offset by the annuity rate. This implies advanced consumption right after the diagnosis as depicted by the dotted blue line in both panels in Figure 5. However, individuals diagnosed early in life can have an incentive to defer consumption if they survive long enough for their mortality risk to reapproach that of cancer-free individuals, such that the term $\overline{\mu} - \mu^2$ becomes positive. Altogether this results in a U-shaped consumption pattern illustrated by the dotted blue line in the left panel.

As we will present later (see Figure 7), the cancer stock decreases in a convex pattern regardless of the age at diagnosis, hence $f < 0$ during the first years after the diagnosis. Therefore, $\frac{u^2_{c_2 E}}{u^2_{c_2}} f > 0$

---

[31] This analysis for the consumption profile already highlights how the introduction of a stochastic health shocks affects the individual on multiple levels and how intricate the effort to disentangled the different driving forces can be.

[32] The marginal utility of consumption is higher for lower levels of the deficit stock. This holds for the multiplicative specification of the stage-2 utility function in our calibration and is empirically supported by Finkelstein et al. (2013).

captures the incentive of individuals to postpone their consumption in line with their recovery from cancer (dash-dotted green line). In our calibration this would imply a steep increase in consumption in the first few years after the diagnosis, followed by a fairly constant profile as the deficit stock is significantly decreased and has limited impact (similar for early and late diagnoses).

Adding up both terms to obtain the total effect, we are able to explain why the consumption profiles of early and late diagnosed individuals feature such pronounced qualitative differences. In the left panel we see, that for $s = 30$ the desire to defer consumption dominates the incentive to advance consumption in the first two years resulting in an initial increase in consumption in total (solid black line). As the impact of the cancer stock vanishes, the difference between annuity returns and mortality risk adds the U-shape for the consumption profile over the remaining life-course. For $s = 70$, the above average mortality risk after a cancer diagnosis becomes the dominant driver after the first 6 months. Following an initial boost, consumption then declines until the end of life (solid black line), as post-cancer mortality never decreases to levels, where it falls short of the annuity rate.
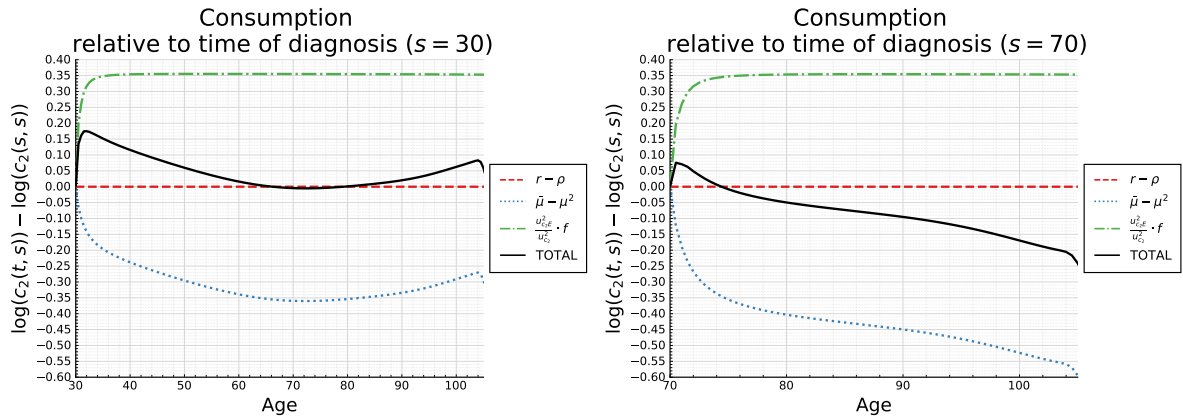

Figure 5: Illustrations of the impacts of the different parts of the Euler equations for consumption after the diagnosis.

Corollary 2 (Consumption after diagnosis)**.** *The increased mortality risk following a cancer diagnosis implies, that individuals strongly advance their consumption. This effect can be dampened or even overcompensated by the incentive to postpone consumption to ages where a reduction in the intensity of cancer allows for a higher marginal utility of consumption.*

5.3 Health expenditure and state profiles

In this section we focus on the health expenditure choices and the resulting profiles of the health/survival stock and cancer deficits. We also calculate, for our numerical specification, the valuations of health from Definition 1 and decompose them according to Proposition 1, noting that they will play a crucial role in the analysis of the health care Euler-type equations in Section 5.4.

Health expenditure profiles

For the life-cycle allocation of general health expenditures, as shown in the upper left panel in Figure 6, a cancer diagnosis does not imply a great difference regarding the qualitative shape over the life-cycle. Quantitatively, however, we obtain some significant differences. For early diagnoses up to age 70 health expenditures initially drop by around 2000$ per year, which amounts to around 50% of the expenditures before a diagnosis for the youngest groups. This gap persists for the remaining lifetime following a cancer

diagnosis and even widens after the age of 70. On the other hand a diagnosis already late in life leads to an even more sizeable immediate reduction in general health care spending (in absolute values).

The instantaneous reduction in general health care spending (difference between the dashed and dotted lines) can be traced back to the sudden and large increase in spending on cancer care that is triggered by a diagnosis (see the upper right panel of Figure 6). For all age groups, the initial expenditures are highest immediately following the diagnosis and then decline rapidly within around ten years after the diagnosis where the residual level can be be interpreted as regular screenings to avoid the reemergence of cancer. Strikingly, (initial) expenditures on cancer-care (dotted line) slightly increase with age from around 67000$ at young ages to 72000$ if the diagnosis occurs at age 60. As the time horizon quickly shortens when cancer is diagnosed closer to the end of life, the benefits of cancer care are less pronounced, and as a result the initial spending level declines steeply down to insignificant values. Expenditures for cancer care might appear rather high, but are in line with the average per-patient costs, that are reimbursable for breast cancer. Blumen et al. (2016) estimate the allowed costs to lie between roughly 60000$ and 134000$ within the first 12 months after a diagnosis (depending on the stage of cancer) and between 13000$ and 70000$ through the second year.
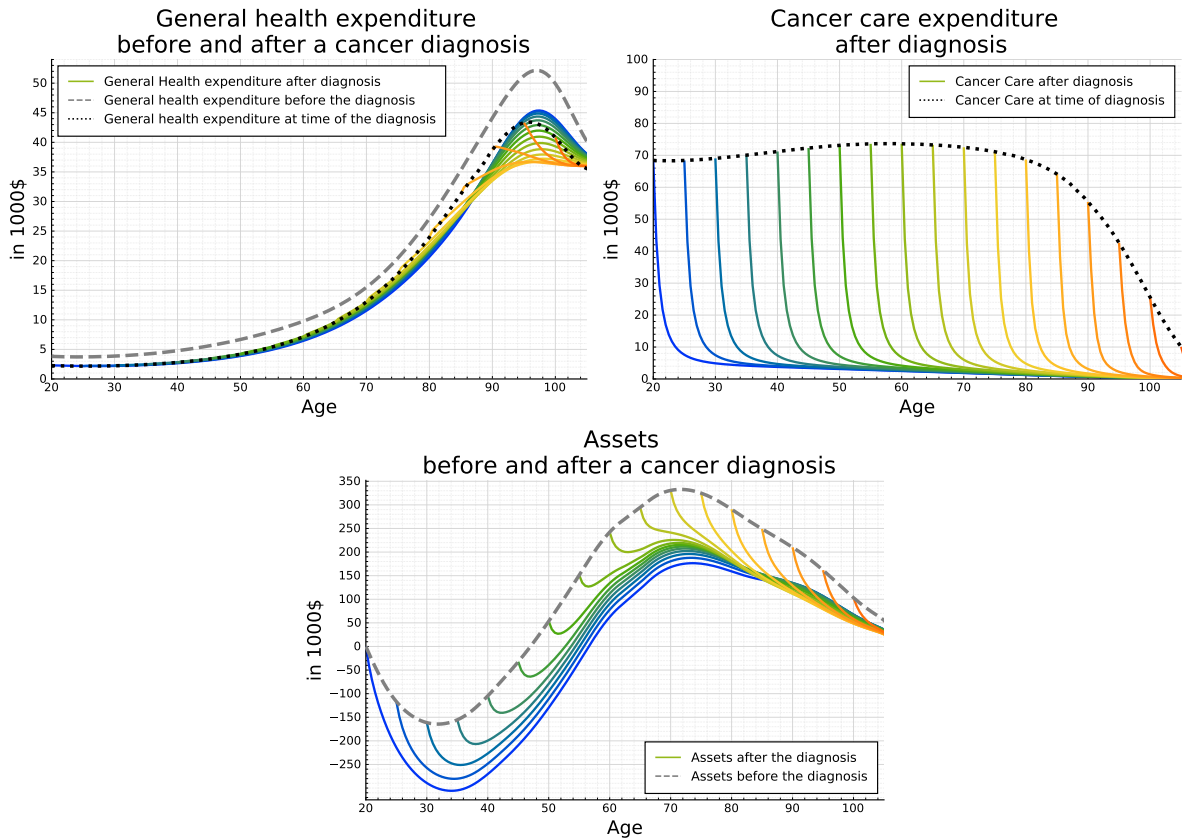

Figure 6: First row: General and cancer specific health expenditures before and after the diagnosis. Second row: Financial assets before and after the diagnosis.

The lower panel of Figure 6 shows the financial burden associated with the high expenditures for cancer care. Right after the diagnosis, assets decline steeply, where young individuals (holding negative assets already (dashed line)) go further into debt in order to finance cancer treatment, despite concomitantly reducing general health expenditure and consumption, as in Figure 3. We need to stress, however, that the structure of the annuity market plays a crucial role. As we assume that cancer patients are not identified by the market, they obtain the same annuity rate as a representative member of the

population. Despite their higher mortality risk individuals with a cancer diagnosis can thus still go into debt to finance cancer care, facing the same full interest rate as the average population. In contrast, adjustments in the annuity price to the individual health state and associated mortality would strongly compromise the individual's scope for undertaking the investments that are necessary for treating cancer effectively.

Health state profiles

In Figure 7 we see the outcomes related to general health care and chronic care spending, as reflected in the profiles of the health/survival state and the cancer stock. In the left panel we see the survival/health state profiles in the absence of a cancer diagnosis (dashed grey line) and those following a cancer diagnosis (solid grey-scaled/coloured lines) at various ages (again in 5-year steps). Most prominently we see a rapid decline in survival immediately after a diagnosis, reflecting high mortality risk during the first years after a diagnosis. While the decline in survival becomes less distinct after ten years (for most age groups), even for diagnoses early in life, the survival path never returns to the one of cancer-free individuals.

In the right panel we see that the (optimal) pattern of cancer care is effective in diminishing the cancer stock for all ages at diagnosis. This suggests that for our specification, cancer kills patients relatively early on and is otherwise transformed into a chronic disease, where a small residual stock is not eliminated but rather held in check by chronic care. This can be interpreted as a situation in which cancer can, indeed, be reduced to negligible levels (and, thus, be considered to be eliminated for all practical purposes) but where the risk of its resurgence requires a small amount of care in the form of regular screenings. The residual cancer stock also induces a mortality risk (depending on the age at diagnosis), which covers the chance of dying through a potential relapse. This also explains intuitively, why the survival profiles never return to the qualitative shape of cancer-free individuals. Meanwhile the impact of the residual cancer stock on consumption utility is insignificant in the long run. For our calibration, after ten years the remaining cancer stock decreases utility by less than one percent.
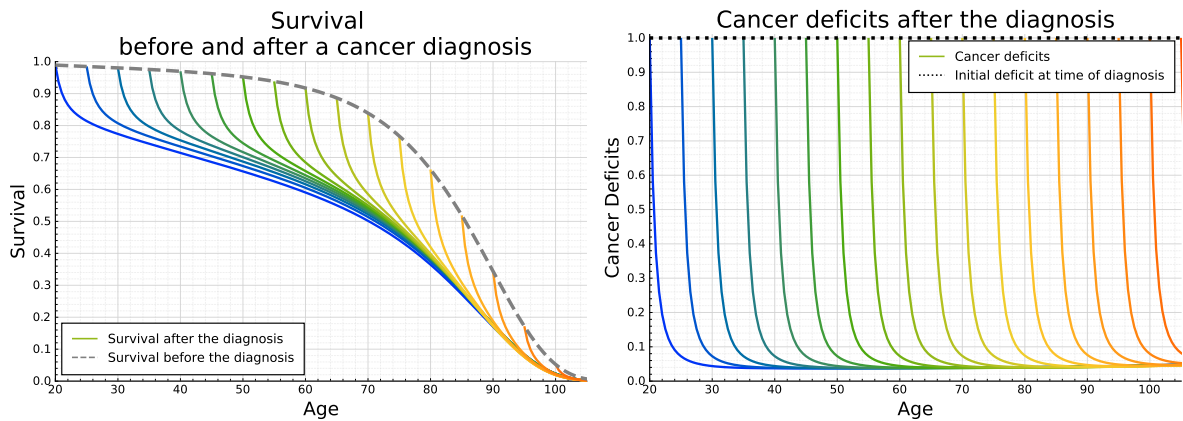

Figure 7: Health/survival state and cancer deficits before and after the diagnosis

Valuations of Health

Before we explore the Euler-type equations for general health and chronic care expenditures, we will make a small detour providing numerical illustrations of the valuations of the various components of health derived in Section 3.1. These profiles play a crucial role for the shape of health investment profiles as we already indicated in Section 3.2 discussing the first order optimality conditions.

In the upper left panel of Figure 8 we compare the different valuations. The solid black line shows the value of health (VOH) in the first, cancer-free stage ψ_H^1, which starts slightly above 16 million $ at age 20 and follows a concave-convex decline over the life-cycle. The ex-ante or average VOH, which takes account of the prevalence of cancer within each particular age-class (see definition 2), is illustrated by the dotted (blue/dark) line and differs only slightly from the first-stage VOH. The dash-dotted (green) line indicates the value of acute survival, which covers the willingness to pay for an increase in the continuation probability. Considering this value might look counter-intuitive at first, as in the case of a progressive disease such as cancer we do not have a mortality risk at the point of diagnosis. Following equation (26), however, the value of acute survival also covers the initial VOH after the shock if $P \equiv 1$. Hence, we can see that the VOH drops dramatically after a diagnosis before age 70, this difference being much less pronounced for diagnoses later in life.
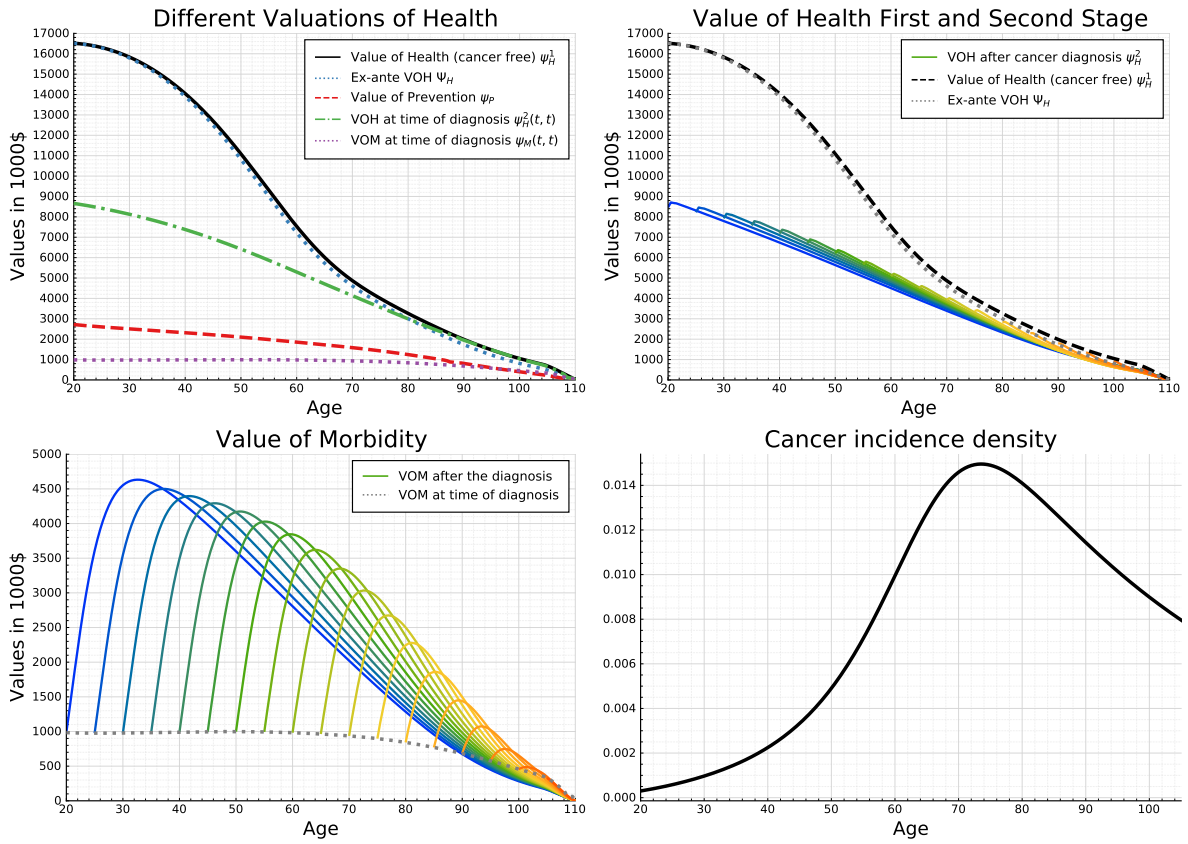

Figure 8: Valuations of Health

The upper right panel shows the development of the VOH in the second stage, i.e. the stage with cancer, for different ages at diagnosis as compared to the VOH in stage one. Apart from an age gradient there is only little variation in the second-stage VOH with respect to the age at diagnosis. Given the strong decline in the VOH at diagnosis for all ages up to 70, questions about the comparatively small difference between the ex-ante VOH and ψ_H^1 might arise. The answer originates in the low probabilities of getting diagnosed with cancer early in life, as can be seen in the lower right panel in Figure 8. This implies a low weight of ψ_H^2 in the calculation of the ex-ante VOH (see equation (33)). In contrast, the difference in the first and second stage VOH has become small for the age-groups older than 70, which exhibit significant prevalence of cancer. Thus the difference between the "cancer-free" VOH and the ex-ante VOH remains remarkably small over the whole life-cycle. This has an interesting policy implication:

Corollary 3 (Expected Value of Health)**.** *Even health shocks that are wide-spread (such as cancer) and have large impacts on the VOH from an individual perspective, exert little influence on the VOH from a population perspective (as represented by the ex-ante or average VOH) if their prevalence distribution is centred on high ages.*

Returning to the upper left panel, the dashed (red) line indicates the value of prevention, i.e. the willingness to pay for a reduction of the hazard rate. This value is significantly smaller, but still starts above 2.5 million \$ at age 20 and is still around one million \$ at the age of 85. Finally the dotted (purple/light) line represents the value of reducing morbidity (VOM) right at the time of the diagnosis. This value stays fairly constant around one million \$ up to age 70 and then declines slowly towards the end of the time-horizon. The lower left panel in Figure 8 shows the VOM after a cancer diagnosis for varying ages at diagnosis. Strikingly, the value of morbidity increases over the first 10-12 years to values more than fourfold the VOM at the time of diagnosis, followed by a continuous decline over the remaining life-course. These patterns can be understood when recalling from the FOC (37) that in the optimum the VOM equals the effective price of reducing the cancer stock by one unit, $\psi_M = \frac{p^2}{-f_{h_2}}$. As cancer care becomes less effective with a diminishing cancer stock, this implies an increasing effective price of reducing cancer (further). For an optimal allocation, this, in turn, implies that the VOM (i.e. the marginal rate of substitution between reductions in the cancer stock and consumption) must increase as well. Once the cancer stock has been reduced to its residual value after around ten years, the effective price of controlling it remains constant, and the VOM (and efforts towards controlling cancer) now declines in line with the reduction of the remaining life expectancy.

In the next step we discuss the VOH before a diagnosis (left panel) and the value of prevention (VOP) (right panel) and their respective decompositions for an optimal allocation within our numerical setting in Figure 9. The solid (black) lines equal the sum of the respective sub-parts and correspond to the curves in the upper left panel in Figure 8. For the VOH in the left panel the conventional VOL (dashed red line) adds the main share over the whole life-cycle. This part thereby decreases nearly linearly with age and accounts for most of the VOH after the age of 70. Up to that age the benefit of health, as measured by S_1, through better prevention of cancer (dotted blue line) contributes significantly and explains more than a third of the total cancer-free VOH at age 20. The remaining component of the VOH captures the effect of first-stage health, S_1, on second-stage health, S_2. At values below 1 million \$ this part is relatively small compared to the others, but still far from insignificant in absolute terms.

Corollary 4 (Value of Health before diagnosis)**.** *A potential cancer diagnosis significantly adds to the value of health for relatively young ages if health has a preventive effect on the hazard rate.*

In the right panel, we decompose the value of prevention into its three components. The direct (dashed red line) and indirect (dotted blue line) benefits of postponing a cancer diagnosis get partially offset by the utility which would have been generated after a diagnosis (dash-dotted green line). As the absolute value of the losses are less than even the direct gains alone, we obtain an overall positive value of prevention over the life-cycle.
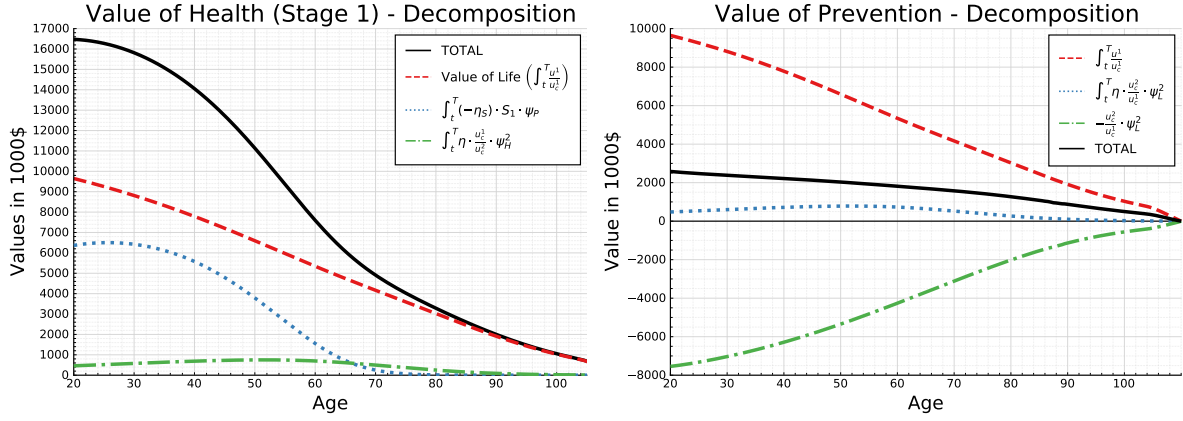
Figure 9: Valuations of Health

5.4 Euler equations for health care investments

Using the first-order optimality conditions for the different types of health expenditures presented in Proposition 2, which hold for each separate point in time, we can derive Euler type equations for general health care (before and after the diagnosis), chronic care and also preventive care[33]. The full derivations are again relegated to appendix 7.5.

The Euler equations for all types of health investments share a similar pattern as they consist of three main terms. (a) Intuitively the dynamics of a specific type of health care are connected to the dynamics of the valuation of health, which enters the respective FOC. E.g. for life-cycle pattern of general health care investments, the value of health is the decisive valuation, while for chronic care, the value of morbity is a determining factor. (b) Changes in the efficiency of health investments over time due the impacts of other variables or external factors like age or technology can contribute significantly towards the qualitative pattern over the life-course. (c) The price development of different health care types imposes further incentives to either postpone investments to later ages or conduct them earlier in life.

5.4.1 General health expenditure

The Euler equations for general health expenditure before and after a cancer diagnosis in equations (44) and (45) directly illustrate these three main contributing factors. Note that all terms are finally scaled with $\frac{-\mu^i_{b_i}}{\mu^i_{b_i b_i} b_i}$ respectively, which is the equivalent to the inverse of the intertemporal elasticity of substitution in the consumption Euler equation.

$$\frac{\dot{b}_1}{b_1} = \frac{-\mu^1_{b_1}}{\mu^1_{b_1 b_1} b_1} \Bigg[-\frac{\mu^1_{b_1 S_1}}{\mu^1_{b_1}} \mu^1 S_1 + \frac{\mu^1_{b_1 t}}{\mu^1_{b_1}} - \frac{\dot{p}^b}{p^b} +$$

$$\underbrace{+ r + \bar{\mu} + \mu^1_{S_1} S_1 - \frac{u^1/u^1_{c_1}}{\psi^1_H} + \eta_{S_1} S_1 \frac{\psi_P}{\psi^1_H} - \eta P \frac{u^2_{c_2}}{u^1_{c_1}} \left\{ \frac{(\psi^2_H - \psi^1_H)}{\psi^1_H} + \frac{P_{S_1} S_1}{P} \frac{\psi^2_{life}}{\psi^1_H} + (-B_{S_1}) S_1 \frac{\psi_M}{\psi^1_H} \right\}}_{= \frac{\psi^1_H(t)}{\psi^1_H(t)}} \Bigg] \quad (44)$$

$$\frac{\dot{b_2}}{b_2} = \frac{-\mu_{b_2}^2}{\mu_{b_2 b_2}^2 b_2}\left[-\frac{\mu_{b_2 S_2}^2}{\mu_{b_2}^2}\mu^2 S_2 + \frac{\mu_{b_2 t}^2}{\mu_{b_2}^2} - \frac{\dot{p^b}}{p^b} + \underbrace{r + \bar{\mu} + \mu_{S_2}^2 S_2 - \frac{u^2/u_{c_2}^2}{\psi_H^2}}_{=\frac{\psi_H^2(t,s)}{\psi_H^2(t,s)}}\right] \tag{45}$$

We will now discuss the economic interpretations of the three driving forces (a), (b), and (c) and their respective decompositions in detail.

(a) Value of Health: The rate of change of the VOH, $\dot{\psi}_H^1(t)/\psi_H^1(t)$ resp. $\dot{\psi}_H^2(t,s)/\psi_H^2(t,s)$, plays a crucial part in the Euler equations. Using the results of Proposition 1 these rates can be decomposed into several parts. Parts (i) and (ii) discussed below show up in both stages and the interpretations are relevant for general health expenditures in both stages. Meanwhile the effects discussed in (iii)-(vi) are only present before the cancer diagnosis.

  (i) The discount rate of the VOH $(r + \bar{\mu} + \mu_{S_i}^i S_i \gtreqless 0)$ incentivizes later health investments for high market returns $(r + \bar{\mu})$. This effect is partially offset as the individuals take the effect of higher health on the mortality rate into account $(\mu_{S_i}^i S_i \leq 0)$.

  (ii) $-\frac{u^i/u_{c_i}^i}{\psi_H^i} \leq 0$ captures the depreciation of the value of life (relative to the VOH) and motivates health investments earlier in life.

  (iii) $\eta_{S_1} S_1 \frac{\psi_P}{\psi_H^1} \leq 0$ implies advancement of health investments towards younger ages, since S_1 has an impact on the hazard rate and decreases the probability of a cancer diagnosis. The extent of this effect depends on the value of prevention relative to the value of health.

  (iv) $-\eta P \frac{u_{c_2}^2}{u_{c_1}^1}\frac{(\psi_H^2 - \psi_H^1)}{\psi_H^1} \gtreqless 0$ infers that if the value of health is smaller in the second stage compared to the first stage (as we have seen in Figure 8), health investments are less attractive in the present (in case of a shock, health is instantaneously valued less) and consequently delayed to later stages in life. This factor becomes more pronounced at ages at which a cancer diagnosis is more likely (increased η).

  (v) $-\eta P \frac{u_{c_2}^2}{u_{c_1}^1}\frac{P_{S_1} S_1}{P}\frac{\psi_{life}^2}{\psi_H^1} \leq 0$ accounts for the positive effect of health S_1 on the continuation probability and results in another force shifting health investment to younger ages. *This aspect is equal to zero in our scenario, as health has no impact on the continuation probability*

  (vi) $-\eta P \frac{u_{c_2}^2}{u_{c_1}^1}(-B_{S_1})S_1\frac{\psi_M}{\psi_H^1} \leq 0$ lastly represents that as initial deficits are lower, if an individual is in good health at time of the diagnosis, individuals have another incentive to keep their health at high levels over the life-cycle. This results in health investments being advanced towards younger ages. *This aspect is equal to zero in our scenario, as health has no impact on the initial deficits.*

(b) Effectiveness: Changes in the effectiveness of health investments over the life-cycle also play a decisive role for the shape of the profile. This aspect contains two separate parts:

  (i) $-\frac{\mu_{b_i S_i}^i}{\mu_{b_i}^i}\mu^i S_i$ covers that if health investments tend to have a higher impact for already depleted health $(\mu_{b_i S_i}^i > 0)$, people are more likely to defer until their health is depleted to increase their health expenditures. *This aspect is equal to zero in our scenario, as health has no impact on the base mortality rate.*

  (ii) $\frac{\mu_{b_i t}^i}{\mu_{b_i}^i}$ shows that people have an incentive to postpone their health investment to later ages, if marginal effectiveness of health expenditures increase with age $(\mu_{b_i t}^i < 0)$. This could also include a set-up, in which individuals expect health technologies to improve during their lifetime.

(c) Price: Expected increases of the prices of health care over time push health investments towards earlier stages in life (and vice versa) as the term $-\frac{\dot{p}^b}{p^b}$ indicates. *This aspect is equal to zero in our scenario, as prices are assumed to be constant.*

Similarly to the consumption profiles we want to illustrate these theoretical driving forces using our numerical calibration. As already indicated above, some of the terms are equal to zero due to the assumptions made for the scenario of a cancer diagnosis and are therefore omitted in the presentation in the figures below. Like in Figure 4 and 5, we present the terms in the Euler equations aggregated up to timepoint t as their sum corresponds to the difference of log-expenditures at point t and age $t_0 = 20$. Figure 10 shows the numerical evaluations of the different (non-zero) parts and their aggregated sum for general health expenditure of cancer-free individuals.
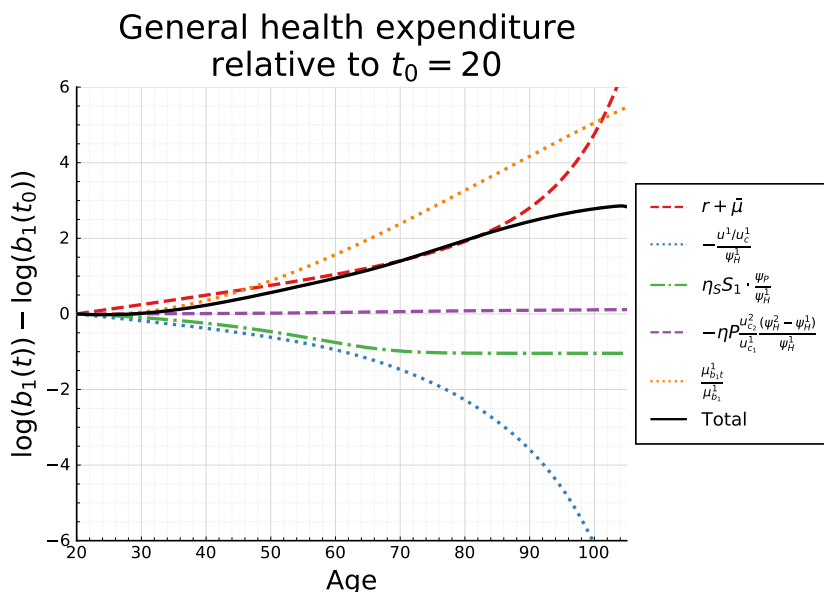

Figure 10: Illustrations of the impacts of the different parts of the Euler equations for general health expenditure before a cancer diagnosis

We can easily identify the two main driving forces behind the overall increasing health expenditures over the life-cycle. First, intuitively the sum of market and annuity interest alone would impose a strong delay in health expenditure (dashed red/dark grey line) especially towards the end of life, where the annuity rate strongly increases. The second strong incentive to postpone health investments is the increasing efficiency of general health care for older ages (dotted orange/light grey line), which by itself would imply nearly linearly increasing health expenditures after the age of 45. Consequently the effect of interest is more pronounced in the early ages and for ages 100+, while the impact of the increasing effectiveness is stronger between ages around 45 to 100.

However the aggregated effect of these two terms is dampened by two other factors. While the advancing effect of the preventive aspect of higher health is only present up to age 70 (dash-dotted green/grey line), the depreciation of the value of life (dotted blue/grey line) becomes increasingly strong over the life-course. This leads to the health expenditures in total becoming stagnant at the end of the time horizon (the solid black line becomes flat). Finally we see that the last term containing the impact of a change in the VOH through diagnosis (dashed purple/grey line) has no significant impact over the whole life-course.[34]

---

[34]The insignificant impact can be explained similarly to the small difference between the expected and the cancer-free VOH.

In Figure 11 we continue with the analogous decomposition for general health expenditure after a cancer diagnosis. In the left panel we illustrate the decomposition for an early diagnosed individual at age 30, while in the right panel the diagnosis occurs later in life at age 70. Although we already identified in Figure 6 that the age-profile for health expenditures after a cancer diagnosis follows a qualitatively similar pattern to the one of cancer free patients, this decomposition still helps us to identify the smaller differences. First of all, there are only three terms remaining for health expenditures after the diagnosis
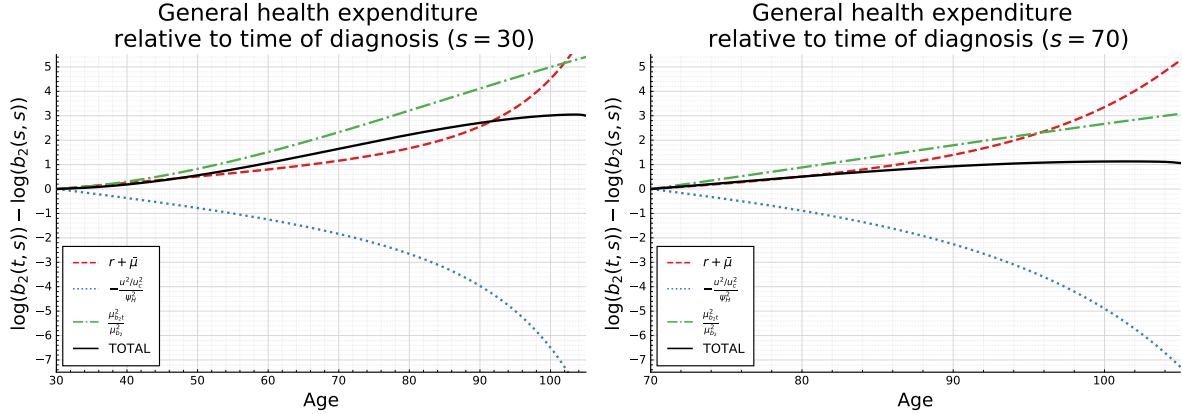

Figure 11: Illustrations of the impacts of the different parts of the Euler equations

compared to the previous five terms. However the three with the strongest magnitude are still present. Again the interest and annuity rate (dashed (red) line) incentiveses individuals to defer health care, however now the postponing effect of increasing effectiveness of health care (dash-dotted (green) line) is the most distinct reason behind overall increasing health expenditures (solid black line). The sole dampening effect of the otherwise strongly increasing profile results from the depreciation of the value of health (dotted (blue) line). The absence of an additional advancing force (like the preventive aspect of health before the diagnosis) at least partially explains, why the total effect is actually slightly higher for ages after 60 in case of an early diagnosis compared to the cancer free profile.[35]

Corollary 5 (General Health Expenditure). *The main driving forces behind increasing health expenditures over the life-course are the market and annuity interest rate and the increasing effectiveness of health care in older ages. In the presence of a potential cancer diagnosis the dampening effect of the depreciation of the value of life is enhanced by an additional term, which covers the preventive effect of health and motivates advancement of general health expenditure. This term is only present before the cancer diagnosis and only shapes the behaviour of cancer-free individuals.*

5.4.2 Cancer specific chronic care expenditure

As the last part of our numerical analysis we focus on cancer care after the diagnosis, present the corresponding Euler equation and again evaluate the decomposition along the optimal numeric solution. The dynamics of chronic care can differ quantitatively between the different shock scenarios, however the qualitative shaping forces are the same independent of the age at diagnosis. Similar to the general health expenditures all terms in equation (46) are scaled with $\frac{-f_{h_2}}{f_{h_2 h_2} \overline{h}_2}$, the equivalent to the inverse of

---

[35]Furthermore we would like to stress that these figures evidently show how an analytical derivation and assessment of dominating terms in the Euler equations would be fairly impossible. The numerical evaluation nevertheless give great insight into to driving behavioural forces.

the intertemporal elasticity of substitution in the consumption Euler equation.

$$\frac{\dot{h_2}}{h_2} = \frac{-f_{h_2}}{f_{h_2 h_2} h_2} \left[\underbrace{r + \bar{\mu} - f_E + \frac{w_E^2}{\psi_M} + \frac{u_E^2/u_{c2}^2}{\psi_M} - \mu_E^2 \frac{\psi_H^2}{\psi_M} + \frac{f_{h_2 t}}{f_{h_2}} + \frac{f_{h_2 E} f}{f_{h_2}} - \frac{\dot{p^2}}{p^2}}_{=\frac{\dot{\psi}_M(t,s)}{\psi_M(t,s)}} \right] \tag{46}$$

Furthermore we also obtain the same three main parts in the dynamics: (a) value of morbidity dynamics, (b) chronic care efficiency aspects and (c) the price developments.

(a) ***Value of Morbidity:*** As we have seen in Figure 8, the value of morbidity follows a surprising profile after the diagnosis. Consequently decomposing its rate of change $\frac{\dot{\psi}_M(t,s)}{\psi_M(t,s)}$ can enhance our understanding extensively:

   (i) The discount rate of the value of morbidity $(r + \bar{\mu} - f_E)$ again leads to chronic care postponement for high market rates $(r + \bar{\mu})$. If deficits accumulate slower (faster) if they are already on a high level, i.e. $f_E < 0$ $(f_E > 0)$ individuals further delay (accelerate) their chronic care expenditures.

   (ii) $\frac{w_E^2}{\psi_M} \leq 0$ captures the impact of the deficit stock on the wage, what also incites individuals to conduct chronic care earlier in life or closer to the time of the diagnosis. *This term is equal to zero in our scenario, as deficits are assumed to have no impact on working income.*

   (iii) $\frac{u_E^2/u_{c2}^2}{\psi_M} \leq 0$ is another reason for individuals to invest into chronic care earlier, as the deficits have a negative impact on the utility. The extent of the shift of chronic care towards younger ages depends on the ratio of marginal utility gains through deficit reduction and consumption.

   (iv) $-\mu_E^2 \frac{\psi_H^2}{\psi_M} \leq 0$ implies advancements of chronic care as increasing deficits impose an additional mortality risk. The strength of this factor depends on the value of health relative to the value of morbidity.

(b) ***Effectiveness:*** Changes in the effectiveness of chronic care over the life-course also play a decisive role for the shape of the patterns:

   (i) $\frac{f_{h_2 E} f}{f_{h_2}} \gtrless 0$ covers the effect of the cancer stock on the marginal impact of chronic care. The sign of this term depends on several factors, e.g. first the sign of the mixed derivative $f_{h_2 E}$ might be ambiguous. Depending on the specific health shock, chronic care might be more effective if a certain deficit stock has been accumulated $(f_{h_2 E} < 0)$. On the other hand it is also possible, that an higher accumulated deficit stock makes it harder to eliminate the remaining deficits as the treatment is effective $(f_{h_2 E} \geq 0)$. Second it is crucial whether deficits are accumulated or reduced at every point in time as it defines the sign of f.

   (ii) Again we might expect $\frac{f_{h_2 t}}{f_{h_2}} \gtrless 0$ to lead to a delay in chronic care investments, since it is reasonable for chronic care to be more effective for older ages. Still we cannot make a statement about the sign of this term in general. *This term is equal to zero in our scenario, as the progression of cancer is assumed to be independent of age.*

(c) ***Price:*** Expected increases of the price of cancer care over time can push cancer care closer to the diagnosis (and vice versa) as the term $-\frac{\dot{p^2}}{p^2}$ indicates. *This term is equal to zero in our scenario, as prices are assumed to be constant.*

In Figure 12 we present the numerical profiles aggregated from the time of diagnosis analogue to the previous analyses. First of all we find, that the qualitative patterns of each decompositional term are

similar regardless the age at diagnosis, however the absolute values and consequently the composite effect can be different.
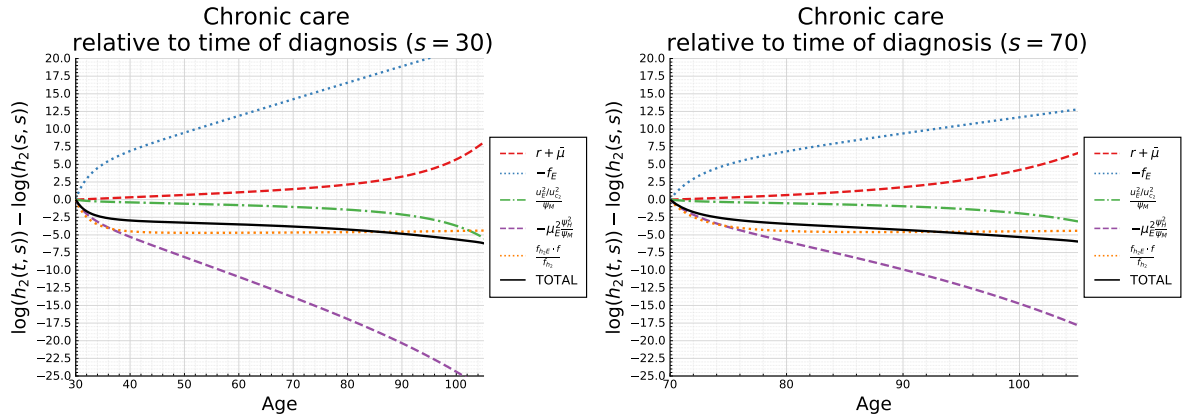

Figure 12: Illustrations of the impacts of the different parts of the Euler equations for cancer specific care after a diagnosis.

Similarly to all types of health expenditures presented so far, the total interest rate $r + \overline{\mu}$ alone would lead to deferral of chronic care especially late in life (dashed red/dark grey line). Strikingly, as the marginal impact of deficits on the efficiency of chronic care is (for the optimal chronic care profile) stronger than the cancer growth rate δ_0, the term $-f_E$ implies a delay in cancer care, especially in the first years after the diagnosis (dotted blue/dark grey line). Consequently through this channel we would expect a chronic care profile even more concentrated right after the diagnosis, if the efficiency of chronic care was independent or even increasing with a decreasing cancer stock.

On the other hand the negative impacts of the cancer stock are an incentive for the advancement of cancer treatments. Notably the term concerning the additional mortality risk (dashed violet/grey line) implies a continuously decreasing chronic care pattern after the diagnosis. This effect is strengthened by the term containing the impact of cancer deficits on utility, which, however, is comparatively small and only increases in magnitude at very late ages (dash-dotted green/light grey line).[36] As a last part the dotted orange/light grey line shows, that the decreasing efficiency of cancer treatments for decreasing cancer stock leads to investments closer to the diagnosis, but has no significant impact on the qualitative shape of the pattern after the first ten years. Adding all terms up, we obtain the total profile (solid black line), which highlights the dominance of negative terms over the positive ones. The strongest decrease in the cancer expenditure profile occurs right after the diagnosis, which is followed by a period of relatively flat resp. slightly diminishing expenditures.

Corollary 6 (Cancer treatment). *A spike in cancer care right after the diagnosis is mainly driven by the high additional mortality risk and decreasing efficiency of cancer care for a decreasing cancer stock. However the combined impact is attenuated as the direct effect of the cancer stock on cancer progression (under the optimal cancer care regime) acts in an opposing way. Compared to the other types of health care, the total interest rate significantly affects the cancer care profile only towards older ages 70+.*

5.4.3 Preventive care

Lastly we want to analyse the theoretical results we can derive for preventive care, which we unfortunately cannot underline with our numerical solution as for the other types of health care. However the

---

[36]Note that this reassures, that changes in our educated guesses for the parameters of the utility function likely do not have significant impact on the overall structure of the optimal solution.

interpretations of the analytical results still hold true in general. In equation (47) we present the Euler type equation for preventive care.

$$\frac{\dot{h_1}}{h_1} = \frac{-\eta_{h_1}}{\eta_{h_1 h_1} h_1} \left[\underbrace{ r + \bar{\mu} + \eta P \frac{u^2_{c_2}}{u^1_{c_1}} - \frac{u^1/u^1_{c_1}}{\psi_P} + \left(\rho - \frac{\dot{P}}{P} - \frac{\frac{d}{dt} V^*(t)}{V^*(t)} \right) \frac{u^2_{c_2}}{u^1_{c_1}} \frac{P \psi^2_{life}}{\psi_P} }_{= \frac{\dot{\psi_P}}{\psi_P}} - \frac{\eta_{h_1 S_1} \mu^1 S_1}{\eta_{h_1}} + \frac{\eta_{h_1 t}}{\eta_{h_1}} - \frac{\dot{p^1}}{p^1} \right]$$

(47)

Analogous to the other health care measures the Euler equation consists of the equivalent of the inverse of the intertemporal elasticity of substitution as a scaling factor and the three main impact factors: (a) value of prevention dynamics, (b) preventive care efficiency aspects, and (c) the price developments. While the latter two are straight forward as for the other types of health care, the decomposition of the former can be significantly more involved. This follows intuitively from preventive care (resp. the value of prevention) generating costs in the cancer-free present, while its impacts and benefits are on the probabilistic side.

(a) Value of Prevention: The rate of change of the value of prevention $\frac{\dot{\psi_P}(t)}{\psi_P(t)}$ can be split up in the following parts:

(i) The discount rate of the value of prevention ($r + \bar{\mu} + \eta P \frac{u^2_{c_2}}{u^1_{c_1}} > 0$) again leads to chronic care postponement for high market rates ($r + \bar{\mu}$). Additionally the desire for precautionary savings $\eta P \frac{u^2_{c_2}}{u^1_{c_1}}$ (which we also discussed in Section 5.2 regarding the consumption profiles) further delays preventive care.

(ii) $-\frac{u^1/u^1_{c_1}}{\psi_P} < 0$ captures the the value of life lost, when being diagnosed with cancer, relative to the value of prevention and motivates advancements of preventive care to earlier ages.

(iii) $\left(\rho - \frac{\dot{P}}{P} - \frac{\frac{d}{dt} V^*(t)}{V^*(t)} \right) \frac{u^2_{c_2}}{u^1_{c_1}} \frac{P \psi^2_{life}}{\psi_P} \gtrless 0$ is the most involved part. It contains the comparison between the value of life in the second stage and the value of prevention. The sign of this term depends on whether the utility discount rate (ρ) is higher then the sum of the rate of change in the continuation probability and the optimal objective value in the second stage ($\frac{\dot{P}}{P} + \frac{\frac{d}{dt} V^*(t)}{V^*(t)}$). However it is more than plausible that $\frac{\frac{d}{dt} V^*(t)}{V^*(t)} < 0$ in general, since the length of the remaining time horizon after a health shock shrinks as t increases. Furthermore as $\frac{\dot{P}}{P} \leq 0$ is more likely than not for most diseases, the total effect is probably positive in many cases (without general applicability) and implies deferral of preventive investments towards older ages, where the cancer diagnosis (or general health shock) has a more significant impact on the individual.

(b) Effectiveness: Changes in the effectiveness of preventive care over the life-cycle also play a decisive role for the shape of the profile:

(i) $-\frac{\eta_{h_1 S_1} \mu^1 S_1}{\eta_{h_1}}$ indicates, that depending on whether preventive care is more ($\eta_{h_1 S_1} < 0$) or less ($\eta_{h_1 S_1} > 0$) effective while in better health, individuals are either motivated to conduct preventive care earlier or later in life.

(ii) $\frac{\eta_{h_1 t}}{\eta_{h_1}}$ shows that the expectation of higher effectiveness of preventive care in the future ($\eta_{h_1 t} < 0$) leads to postponement of preventive care to older ages.

(c) Price: Expected increases of the price of preventive care over time push investments into preventive actions towards earlier stages in life (and vice versa) as the term $-\frac{\dot{p^1}}{p^1}$ indicates.

Corollary 7 (Preventive Care). *The Euler equation for preventive care is the most involved within all types of health care. Beside the price, effectiveness and interest rate impacts similar to other types of health care, the demand for precautionary savings potentially delays preventive efforts. While the depreciation of the value of life incentiveses earlier investments in preventive care, a counter acting deferral effect towards later ages where the shock has more significant impact is likely to exist.*

6 Conclusions

We have constructed a life-cycle model in which individuals respond to the risk of a singular, life-changing shock to their health (e.g. heart attack, stroke, cancer, diabetes, disabling accident) by investing besides their general health expenditures in a range of distinct forms of health care: preventive care to reduce, directly or indirectly, the arrival rate of the shock; acute care to lower instantaneous survival and the extent of morbidity/disability at the point of the shock; and chronic care to lower mortality and morbidity in the follow-up of the shock. We solve the complex underlying stochastic optimal control problem with a random time horizon by applying an innovative transformation into an age-structured control model. This enables us to derive (i) intuitive expressions for the first-order conditions for the choice of health care based on their respective (monetary) valuations; and (ii) Euler equations that reflect both age and duration of disease after the onset of the shock and forward-looking behaviour before the arrival of the shock. The first-order conditions first of all affirm the intuitive notion, individuals choose their optimal level of all distinct types of health care, so that the marginal costs of any additional effort match the monetary valuation of the respective aspect of health. On the other hand they also provide a basis for empirical derivations of the valuations of different facets of health following a revealed preference argument.

Calibrated to US data, our model illustrating the risk of onset of cancer provides a very good fit between (i) general and cancer-specific health care expenditure, (ii) age-specific survival and age-specific prevalence of cancer; (iii) age- and duration-specific cancer mortality rates, as well as (iv) average cancer and non-cancer specific mortality. Under the assumption that assets are exclusively held as annuities the return of which is based on the average mortality rate, relatively young individuals respond to a cancer diagnosis by adjusting their consumption in the light of an increased mortality risk, from a gradually increasing pattern to a U-shaped pattern (after a brief initial adjustment phase). To finance high temporary cancer-specific care they reallocate resources from consumption and health expenditures putting both respective profile on a lower level compared to cancer free individuals. Given the possibility to do so in a perfect annuity market, they also run up substantial debt in order to finance cancer-specific health care. With the duration of the disease, cancer-specific health care is reduced, reflecting an intuitive tendency of such care to evolve from urgent life-protecting care into less intensive chronic care, conditional on the individual's survival. A hike in consumption early-on for individuals diagnosed late in life coincides with a larger drop in general health expenditures and a smaller investment in cancer-specific care, showing that a cancer diagnosis affects individuals differently at different points within the life-cycle. However consumption of individuals diagnosed during old age subsequently steeply reduces to levels below those that would be realized for healthy individuals.

We calculate various expressions for the value of health. Here, the value of health before a diagnosis of cancer starts from a level of around 16.3 Mio $ at age 20 and then declines steadily. Our analysis shows that up to around age 70 a cancer diagnosis lowers the value of health dramatically. While both the cancer-free value of health and the value of health at the point of diagnosis decline with age, the decline is much more pronounced for the cancer-free value of health. From around age 85 onwards, the

cancer-free value of health no longer differs much from the value at the onset of diagnosis, reflecting the much diminished remaining life time in either case. Contrasting the cancer-free value of health with the ex-ante value of health, which involves a weighting with the prevalence function, we find little difference between the two. This may come as a surprise given the strong drop in the value of health for individuals diagnosed at young ages. Notably, however, for these age groups a cancer is an unlikely diagnosis, implying a low weight, whereas for higher ages, the two values of health have converged.

Considering the components of the cancer-free value of health, we find that while the largest part falls on the value of survival (tantamount to the conventional value of life), the preventive value of good health (i.e. the value of prevention weighted with the health-related reduction in the incidence of cancer) makes up for a significant proportion of the total value up to age 65. Standing at slightly above 6 Mio $ up to the early thirties, it compares to a value of survival around 9.5 Mio $ and makes up more than a third of the total value of health at age 20. From age 50 onwards, however, the preventive value of good health diminishes quickly and vanishes almost entirely for the highest ages.

Finally, turning to the Euler equations, we see that under our assumption that the rate of time preference equals the interest rate, two offsetting forces emerge as drivers of consumption in the healthy state: On the one hand, given that the individual remains healthy, the average mortality component (including expected mortality due to cancer) exceeds (to increasing extent) the non-cancer mortality as a source of risk, allowing the individual to increase consumption over the remaining life-course. On the other hand, a desire for consumption smoothing leads to the advancement of consumption into the healthy stage, which subsequently translates into a tendency to reduce consumption of the life-course. As it turns out, the former effect is slightly stronger for our calibration, leading to an increase in consumption with age for individuals who remain cancer-free. After the onset of cancer, again two main forces determine the consumption choice: Assuming that the marginal utility of consumption is reduced through cancer, there is a tendency for individuals to defer consumption into the future when (conditional on survival) they expect to experience a higher marginal utility from consumption. As it turns out, for the most part this tendency tends to be offset by the desire of individuals to consume instantaneously given the high mortality risk, following the onset of cancer.

The demand for health care generally develops under the presence of two forces: On the one hand, the return to annuities tends to imply an increase, whereas the writing off of life-years from the respective value of health tends to imply a decline. For general health investments, these effects tend to be moderated by age-related changes to their effectiveness. While for cancer-free individuals the shock-preventive aspect of health implies an additional incentive to advance general health investment, this factor is absent for individuals after a diagnosis. Overall, in our calibration there is a moderate increase in general health investments over the life-course both in the absence and, from a much lower level, in the presence of cancer. While the former three forces also tend to be at play for cancer-specific health care, there are two additional effects: On the one hand, individuals have an incentive to delay the consumption of cancer care, given it is more effective for a disease that has progressed already; on the other hand, the mortality risk that is increasing with the progression of the disease provides a strong incentive to advance cancer care; and this is, indeed, the dominating effect.

While in this paper we have developed a rich framework for the study of large, singular health shocks, there is considerable scope for extensions and applications. First, we are planning to provide analysis for other types of health shocks, in particular cardio-vascular events that involve an instantaneous threat to survival and, thus, warrant the consumption of acute health care. Second, especially in the case of cancer, there is an important issue about the lag between onset and diagnosis, which we plan to address in future work. Third, there is large scope for employing the model to study the role of the annuity

market and the role of health and disability insurance. Finally, we are considering extensions of the model to involve multiple shocks and differences in severity of the health shock as a second dimension of uncertainty.

References

Aldy, Joseph E and W Kip Viscusi (2008). "Adjusting the value of a statistical life for age and cohort effects". In: *The Review of Economics and Statistics* 90(3), pp. 573–581.

Bíró, Anikó (2013). "Subjective mortality hazard shocks and the adjustment of consumption expenditures". In: *Journal of Population Economics* 26(4), pp. 1379–1408.

Blumen, Helen, Kathryn Fitch, and Vincent Polkus (2016). "Comparison of treatment costs for breast cancer, by tumor stage and type of service". In: *American health & drug benefits* 9(1), p. 23.

Boukas, El-Kébir, Alain Haurie, and Philippe Michel (1990). "An optimal control problem with a random stopping time". In: *Journal of optimization theory and applications* 64(3), pp. 471–480.

Capatina, Elena (2015). "Life-cycle effects of health risk". In: *Journal of Monetary Economics* 74, pp. 67–88.

Cole, Harold L, Soojin Kim, and Dirk Krueger (2019). "Analysing the Effects of Insuring Health Risks: On the Trade-off between Short-Run Insurance Benefits versus Long-Run Incentive Costs". In: *The Review of Economic Studies* 86(3), pp. 1123–1169.

Dalgaard, Carl-Johan and Holger Strulik (2014). "Optimal aging and death: understanding the Preston curve". In: *Journal of the European Economic Association* 12(3), pp. 672–701.

De Nardi, Mariacristina, Eric French, and John B Jones (2010). "Why do the elderly save? The role of medical expenses". In: *Journal of Political Economy* 118(1), pp. 39–75.

Eeckhoudt, Louis, Christian Gollier, and Nicolas Treich (2005). "Optimal consumption and the timing of the resolution of uncertainty". In: *European Economic Review* 49(3), pp. 761–773.

Eeckhoudt, Louis and Harris Schlesinger (2008). "Changes in risk and the demand for saving". In: *Journal of Monetary Economics* 55(7), pp. 1329–1336.

Ehrlich, Isaac (2000). "Uncertain lifetime, life protection, and the value of life saving". In: *Journal of Health Economics* 19(3), pp. 341–367.

Feichtinger, Gustav, Gernot Tragler, and Vladimir M. Veliov (2003). "Optimality conditions for age-structured control systems". In: *Journal of Mathematical Analysis and Applications* 288(1), pp. 47–68.

Finkelstein, Amy, Erzo FP Luttmer, and Matthew J Notowidigdo (2013). "What good is wealth without health? The effect of health on the marginal utility of consumption". In: *Journal of the European Economic Association* 11(suppl_1), pp. 221–258.

Fonseca, Raquel, Pierre-Carl Michaud, Arie Kapteyn, and Titus J Galama (2013). "Accounting for the rise of health spending and longevity". In:

Frankovic, Ivan, Michael Kuhn, and Stefan Wrzaczek (2020). "On the anatomy of medical progress within an overlapping generations economy". In: *De Economist*, pp. 1–43.

Freiberger, M. and M. Kuhn (2020). *Modelling health processes through survival*. Mimeo.

French, Eric (2005). "The effects of health, wealth, and wages on labour supply and retirement behaviour". In: *The Review of Economic Studies* 72(2), pp. 395–427.

French, Eric and John Bailey Jones (2011). "The effects of health insurance and self-insurance on retirement behavior". In: *Econometrica* 79(3), pp. 693–732.

Goodwin, James S, Jonathan M Samet, Charles R Key, Charles Humble, Daniel Kutvirt, and Curtis Hunt (1986). "Stage at diagnosis of cancer varies with the age of the patient". In: *Journal of the American Geriatrics Society* 34(1), pp. 20–26.

Grossman, Michael (1972). "On the concept of health capital and the demand for health". In: *Journal of Political economy* 80(2), pp. 223–255.

Hall, Robert E and Charles I Jones (2007). "The value of life and the rise in health spending". In: *The Quarterly Journal of Economics* 122(1), pp. 39–72.

Hansen, Gary D and Selahattin İmrohoroğlu (2008). "Consumption over the life cycle: The role of annuities". In: *Review of economic Dynamics* 11(3), pp. 566–583.

Hassager, Christian, Ken Nagao, and David Hildick-Smith (2018). "Out-of-hospital cardiac arrest: in-hospital intervention strategies". In: *The Lancet* 391(10124), pp. 989–998.

Heuser, Louis, John S Spratt, and Hiram C Polk Jr (1979). "Growth rates of primary breast cancers". In: *Cancer* 43(5), pp. 1888–1894.

Hugonnier, Julien, Florian Pelgrin, and Pascal St-Amour (2013). "Health and (other) asset holdings". In: *Review of Economic Studies* 80(2), pp. 663–710.

Institute for Health Metrics and Evaluation (IHME) (2016). *Tracking personal health care spending in the US*. URL: http://vizhub.healthdata.org/dex.

Jung, Juergen and Chung Tran (2016). "Market inefficiency, insurance mandate and welfare: US health care reform 2010". In: *Review of Economic Dynamics* 20, pp. 132–159.

Khwaja, Ahmed, Frank Sloan, and Sukyung Chung (2006). "Learning about individual risk and the decision to smoke". In: *International Journal of Industrial Organization* 24(4), pp. 683–699.

Kopecky, Karen A and Tatyana Koreshkova (2014). "The impact of medical and nursing home expenses on savings". In: *American Economic Journal: Macroeconomics* 6(3), pp. 29–72.

Kuhn, Michael, Stefan Wrzaczek, and Jim Oeppen (2010). "Recognizing progeny in the value of life". In: *Economics Letters* 107(1), pp. 17–21.

Kuhn, Michael, Stefan Wrzaczek, Alexia Prskawetz, and Gustav Feichtinger (2011). "Externalities in a life cycle model with endogenous survival". In: *Journal of Mathematical Economics* 47(4-5), pp. 627–641.

Kuhn, Michael, Stefan Wrzaczek, Alexia Prskawetz, and Gustav Feichtinger (2015). "Optimal choice of health and retirement in a life-cycle model". In: *Journal of Economic Theory* 158, pp. 186–212.

Laporte, Audrey and Brian S Ferguson (2007). "Investment in health when health is stochastic". In: *Journal of Population Economics* 20(2), pp. 423–444.

Lee, Ronald Demos and Andrew Mason (2011). *Population aging and the generational economy: A global perspective*. Edward Elgar Publishing.

Mandelblatt, Jeanne, Howard Andrews, Jon Kerner, Ann Zauber, and William Burnett (1991). "Determinants of late stage diagnosis of breast and cervical cancer: the impact of age, race, social class, and hospital type". In: *American journal of public health* 81(5), pp. 646–649.

Marti, Joachim and Michael R Richards (2017). "Smoking response to health and medical spending changes and the role of insurance". In: *Health economics* 26(3), pp. 305–320.

Murphy, Kevin M and Robert H Topel (2006). "The value of health and longevity". In: *Journal of political Economy* 114(5), pp. 871–904.

Palumbo, Michael G (1999). "Uncertain medical expenses and precautionary saving near the end of the life cycle". In: *The Review of Economic Studies* 66(2), pp. 395–421.

Picone, Gabriel, Martın Uribe, and R Mark Wilson (1998). "The effect of uncertainty on the demand for medical care, health capital and wealth". In: *Journal of health economics* 17(2), pp. 171–185.

Reichling, Felix and Kent Smetters (2015). "Optimal annuitization with stochastic mortality and correlated medical costs". In: *American Economic Review* 105(11), pp. 3273–3320.

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

Rowen, Donna and John Brazier (2011). "Health utility measurement". In: *The oxford handbook of health economics.*

Satariano, William A, Steven H Belle, and G Marie Swanson (1986). "The severity of breast cancer at diagnosis: a comparison of age and extent of disease in black and white women". In: *American Journal of Public Health* 76(7), pp. 779–782.

Schuenemann, Johannes, Holger Strulik, and Timo Trimborn (2017). "Going from bad to worse: Adaptation to poor health health spending, longevity, and the value of life". In: *Journal of Economic Behavior & Organization* 140, pp. 130–146.

Shepard, Donald S and Richard J Zeckhauser (1984). "Survival versus consumption". In: *Management Science* 30(4), pp. 423–439.

Smith, James E and Ralph L Keeney (2005). "Your money or your life: A prescriptive model for health, safety, and consumption decisions". In: *Management Science* 51(9), pp. 1309–1325.

Smith, V Kerry, Donald H Taylor Jr, Frank A Sloan, F Reed Johnson, and William H Desvousges (2001). "Do smokers respond to health shocks?" In: *Review of Economics and Statistics* 83(4), pp. 675–687.

Surveillance Research Program, National Cancer Institute (Sept. 2020). *SEER*Explorer: An interactive website for SEER cancer statistics [Internet].* URL: `https://seer.cancer.gov/explorer/`.

Talkington, Anne and Rick Durrett (2015). "Estimating tumor growth rates in vivo". In: *Bulletin of mathematical biology* 77(10), pp. 1934–1954.

United Nations (2013). *National transfer accounts manual: Measuring and analysing the generational economy.* UN. URL: `https://www.ntaccounts.org/web/nta/show/Browse%20database`.

University of California, Berkeley (USA) and Max Planck Institute for Demographic Research (Germany) (2020). *Human Mortality Database.* Available at www.mortality.org or www.humanmortality.de.

Veliov, Vladimir M. (2003). "Newton's method for problems of optimal control of heterogeneous systems". In: *Optimization Methods and Software* 18(6), pp. 689–703.

Wrzaczek, Stefan, Michael Kuhn, and Ivan Frankovic (2020). "Using Age Structure for a Multi-stage Optimal Control Model with Random Switching Time". In: *Journal of Optimization Theory and Applications* 184(3), pp. 1065–1082.

Yaari, Menahem E (1965). "Uncertain lifetime, life insurance, and the theory of the consumer". In: *The Review of Economic Studies* 32(2), pp. 137–150.

Yancik, Rosemary, Margaret N Wesley, Lynn AG Ries, Richard J Havlik, Brenda K Edwards, and Jerome W Yates (2001). "Effect of age and comorbidity in postmenopausal breast cancer patients aged 55 years and older". In: *Jama* 285(7), pp. 885–892.

Yogo, Motohiro (2016). "Portfolio choice in retirement: Health risk and the demand for annuities, housing, and risky assets". In: *Journal of Monetary Economics* 80, pp. 17–34.

7 Appendix

7.1 The model as a vintage optimal control model

The transformation from the stochastic formulation defined by equations (12)–(20) into a vintage structured optimal control problem follows the technique developed by Wrzaczek et al. (2020). We will present the two crucial steps here, for the proof and detailed derivations we will however refer to the work of the original authors.

In the first step we transform the stochastic formulation into an equivalent deterministic version. The dynamics of the state variables thereby do not change, but we need to introduce the new state variable $Z_1(t)$, which is already presented in Section 2 and its dynamics are described in equation (21). Following the description in Wrzaczek et al. (2020) we obtain the new deterministic objective function shown below.

$$\max_{\substack{c_1(t),h_1(t),b_1(t)\geq 0 \\ c_2(t,s),h_2(t,s),b_2(t,s)\geq 0 \\ d(t)\geq 0}} \int_0^T e^{-\rho t}\Bigg[Z_1(t)S_1(t)u^1(c_1(t)) +$$

$$+\ \underbrace{Z_1(t)\eta(t,S_1(t),h_1(t))P(S_1(t),d(t))\cdot\left(\int_t^T e^{-\rho(\tau-t)}S_2(\tau,t)u^2(c_2(\tau,t),E(\tau,t))d\tau\right)}_{:=\tilde{V}(t)}\Bigg]\,dt$$

The integrand consists of two parts. The first part is the probability of surviving in good health $Z_1(t)$ multiplied with the expected utility derived from consumption at this age $S_1(t)u^1(c_1(t))$. The second part captures the "probability" of suffering the health shock at age t ($\to Z_1(t)\eta(t,S_1(t),h_1(t))$) times the probability of surviving the shock ($\to P(S_1(t),d(t))$) times the expected aggregated utility over the remaining life-time after health shock at age t. To abbreviate some equations below we define this term as $\tilde{V}(t)$ (as it directly corresponds to the objective value of the second stage).

For the next step in the transformation process we need to introduce the variable $Q(t)$ as the expected utility at age t with the expectation taken over the distribution of all possible prior shocks $s \leq t$, i.e.

$$Q(t) = \int_0^t Z_2(s)S_2(t,s)u^2(c_2(t,s),E(t,s))ds \tag{48}$$

with $Z_2(s)$ being defined as in equation (22). Wrzaczek et al. (2020) then show, that the equation (49) for the aggregated utility over all possible second stage scenarios holds,

$$\int_0^T e^{-\rho t}\tilde{V}(t)dt = \int_0^T e^{-\rho t}Q(t)dt \tag{49}$$

This is illustrated by a diagram similar to a Lexis-diagram (see figure 13). The left respectively right panel in figure 13 illustrate the different integrations techniques on the left respectively right side of equation (49). In both panels (corresponding to the resp. side in equation (49)) the utility is aggregated over the triangle below the 45°-line. In the left panel the integration first takes place along each shock scenario from the age at the health shock until the end of the maximum life-span. In a second step all scenarios weighted with their occurrence probability of $Z_1(t)\eta(t,S_1(t),h_1(t))$ get aggregated. On the right side, for each age t, first the utility generated in each scenario, where the shock has occurred at an age $s < t$, is aggregated and in the second step the integration over all ages takes place.
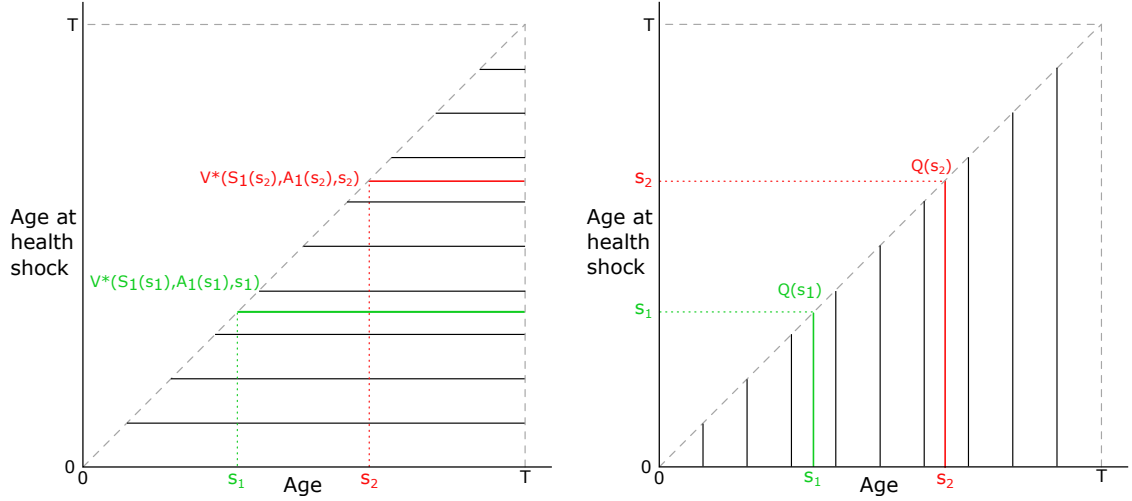
Figure 13: Graphical illustration of the change in the order of integration

In a final step we need to introduce another artificial dimension to the variable $Z_2(s)$ as follows:

$$\frac{dZ_2(t,s)}{dt} = 0, \quad t \geq s,$$

$$Z_2(s,s) = Z_1(s)\eta(s, S_1(s), h_1(s))P(S_1(s), d(s)), \quad \forall s \geq 0.$$

This is a pure technical point (i.e. it eliminates the time lag) and ensures that the optimisation problem fits the problem framework of Feichtinger et al. (2003). Finally we can summarise the whole problem formulation in equations (50)–(66).

$$\max_{\substack{c_1(t),h_1(t),b_1(t)\geq 0 \\ c_2(t,s),h_2(t,s),b_2(t,s)\geq 0 \\ d(t)\geq 0}} \int_0^T e^{-\rho t} \left[Z_1(t)S_1(t)u^1(c_1(t)) + Q(t) \right] \, dt \tag{50}$$

s.t.
$$\dot{S}_1(t) = -\mu^1(t, S_1(t), b_1(t))S_1(t) = -\mu^b(t, S_1(t), b_1(t))S_1(t) \tag{51}$$

$$S_1(0) = 1 \, , \tag{52}$$

$$\dot{A}_1(t) = (r(t) + \bar{\mu}(t))A_1(t) + w^1(t) - c_1(t) - p^b(t)b_1(t) - p^1(t)h_1(t) \tag{53}$$

$$A_1(0) = 0 \, , \; A_1(T) = 0 \tag{54}$$

$$\dot{Z}_1(t) = -\eta(t, S_1(t), h_1(t))Z_1(t) \tag{55}$$

$$Z_1(0) = 1 \, , \tag{56}$$

$$\frac{dS_2(t,s)}{dt} = -\mu^2\big(t, s, S_2(t,s), b_2(t,s), E(t,s)\big)S_1(t) \tag{57}$$

$$= -\left[\mu^b\left(t, s, S_2(t,s), b_2(t,s)\right) + \mu^m(t, s, E(t,s)) \right] S_2(t,s), \qquad t \geq s \tag{58}$$

$$S_2(s,s) = S_1(s) \, , \quad \forall s \geq 0 \tag{59}$$

$$\frac{dA_2(t,s)}{dt} = (r(t) + \bar{\mu}(t))A_2(t,s) + w^2(t, s, E(t,s)) - c_2(t,s) - p^b(t)b_2(t,s) - p^2(t)h_2(t,s) \tag{60}$$

$$A_2(s,s) = A_1(s) - p^d(s)d(s), \quad , \quad A_2(T,s) = 0, \quad \forall s \geq 0 \tag{61}$$

$$\frac{dZ_2(t,s)}{dt} = 0 \tag{62}$$

$$Z_2(s,s) = Z_1(s)\eta(s, S_1(s), h_1(s))P(S_1(s), d(s)), \quad \forall s \geq 0 \tag{63}$$

$$\frac{dE(t,s)}{dt} = f(t,s,E(t,s),h_2(t,s)) \tag{64}$$

$$E(s,s) = B(S_1(s),d(s)), \quad \forall s \geq 0 \tag{65}$$

$$Q(t) = \int_0^t Z_2(t,s)S_2(t,s)u^2(c_2(t,s),E(t,s)) \quad ds \tag{66}$$

7.2 Necessary optimality conditions

Using the vintage structured model formulation (50)–(66) we can derive the necessary optimality conditions following the theoretical work of Feichtinger et al. (2003). These conditions consist of (i) a system of differential equations (67)–(71) for first stage co-state variables, (ii) a system of partial differential equations (72)–(78) and (iii) a set of first order optimality conditions (accounting for boundary solutions for the controls) (79)–(92). Note there are not endpoint condition for $\lambda_A(t)$ and $\xi_A(t,s)$ as we included the terminal conditions $A_1(T) = 0$ and $A_2(T,s) = 0 \,\forall s \geq 0$ in the problem formulation.

$$\dot{\lambda_S} = (\rho + \mu^1 + \mu_{S_1}^1 S_1)\lambda_S + Z_1\eta_{S_1}\left[\lambda_Z(t) - P\xi_Z(t,t)\right] - \xi_S(t,t) - \xi_E(t,t)B_{S_1} - $$
$$ - Z_1\left[u^1 + \eta P_{S_1}\xi_Z(t,t)\right] \tag{67}$$

$$\lambda_S(T) = 0 \tag{68}$$

$$\dot{\lambda_A} = (\rho - r - \bar{\mu})\lambda_A - \xi_A(t,t) \tag{69}$$

$$\dot{\lambda_Z} = (\rho + \eta)\lambda_Z - \eta P\xi_Z(t,t) - S_1 u^1 \tag{70}$$

$$\lambda_Z(T) = 0 \tag{71}$$

$$\frac{d\xi_S(t,s)}{dt} = (\rho + \mu^2 + \mu_{S_2}^2 S_2)\xi_S - Z_2 u^2 \tag{72}$$

$$\xi_S(T,T) = 0 \tag{73}$$

$$\frac{d\xi_A(t,s)}{dt} = (\rho - r - \bar{\mu})\xi_A \tag{74}$$

$$\frac{d\xi_Z(t,s)}{dt} = \rho\xi_Z - S_2 u^2 \tag{75}$$

$$\xi_Z(T,T) = 0 \tag{76}$$

$$\frac{d\xi_E(t,s)}{dt} = (\rho - f_E)\xi_E + \mu_E^2 S_2\xi_S - w_E^2\xi_A - S_2 Z_2 u_E^2 \tag{77}$$

$$\xi_E(T,T) = 0 \tag{78}$$

$$0 \geq Z_1 S_1 u_{c_1}^1 - \lambda_A \tag{79}$$

$$0 = \left(Z_1 S_1 u_{c_1}^1 - \lambda_A\right) \cdot c_1 \tag{80}$$

$$0 \geq -Z_1\eta_{h_1}\left(\lambda_Z - P(S_1,d)\xi_Z\right) - \lambda_A p^1 \tag{81}$$

$$0 = \left(-Z_1\eta_{h_1}\left(\lambda_Z - P(S_1,d)\xi_Z\right) - \lambda_A p^1\right) \cdot h_1 \tag{82}$$

$$0 \geq -\mu_{b_1}^1 S_1\lambda_S - \lambda_A p^b \tag{83}$$

$$0 = \left(-\mu_{b_1}^1 S_1\lambda_S - \lambda_A p^b\right) \cdot b_1 \tag{84}$$

$$0 \geq \xi_E(t,t)B_d + Z_1\eta\xi_Z(t,t)P_d - \xi_A(t,t)p^d \tag{85}$$

$$0 = \left(\xi_E(t,t)B_d + Z_1\eta\xi_Z(t,t)P_d - \xi_A(t,t)p^d\right) \cdot d \tag{86}$$

$$0 \geq Z_2 S_2 u_{c_2}^2 - \xi_A(t,s) \tag{87}$$

$$0 = \left(Z_2 S_2 u_{c_2}^2 - \xi_A(t,s)\right) \cdot c_2 \tag{88}$$

$$0 \geq \xi_E(t,s)f_{h_2} - \xi_A(t,s)p^2 \tag{89}$$

$$0 = \left(\xi_E(t,s)f_{h_2} - \xi_A(t,s)p^2\right) \cdot h_2 \tag{90}$$

$$0 \geq -\mu_{b_2}^2 S_2 \xi_S - \xi_A(t,s) p^b \tag{91}$$

$$0 = \left(-\mu_{b_2}^2 S_2 \xi_S - \xi_A(t,s) p^b\right) \cdot b_2 \tag{92}$$

7.3 Proof of proposition 1 on the valuations of health

Following the calculations of Rosen (1988) we derive the following terms for the valuations:

$$\psi_H^1(t) = \frac{S_1 \lambda_S}{\lambda_A} \tag{93}$$

$$\psi_H^2(t,s) = \frac{S_2 \xi_S}{\xi_A} \tag{94}$$

$$\psi_P(t) = \frac{Z_1(\lambda_Z - P\xi_Z(t,t))}{\lambda_A} \tag{95}$$

$$\psi_{AS}(t) = \frac{Z_1(t)\eta(S_1(t), h_1(t))\xi_Z(t,t)}{\xi_A(t,t)} \tag{96}$$

$$\psi_M(t,s) = \frac{-\xi_E(t,s)}{\xi_A(t,s)} \tag{97}$$

Since we assume an interior solution for the consumption profiles, the first order conditions for consumption (79) and (87) have to be fulfilled in strict form and we can derive the following equations.

$$\lambda_A(t) = Z_1(t)S_1(t)u_{c_1}(c_1(t))$$

$$\xi_A(t) = Z_2(t,s)S_2(t,s)u_{c_2}(c_2(t,s), E(t,s))$$

$$\forall t \in [0,T] \text{ and } s \in [0,t]$$

Furthermore we can easily derive the following relationship, which will be useful in the calculations below.

$$\lambda_A(t) = \frac{\xi_A(t,t)}{\eta(t)P(t)} \frac{u_{c_1}^1(t)}{u_{c_2}^2(t,t)}. \tag{98}$$

In a next step we summaries the explicit solutions for the costate ODEs and PDEs.

$\xi_\mathbf{Z}$ - dynamics

From equation (75) together with the terminal condition $\xi_Z(T,s) = 0$ it directly follows that

$$\xi_Z(t,s) = \int_t^T e^{-\rho(\tau - t)} S_2(\tau, s) u^2(\tau, s) d\tau.$$

$\xi_\mathbf{S}$ - dynamics

From equation (72) together with the terminal condition $\xi_S(T,s) = 0$ it directly follows that

$$\begin{aligned}
\xi_S(t,s) &= \int_t^T e^{-\int_t^\tau [\rho + \mu^2 + \mu_{S_2}^2 S_2] d\tau'} Z_2(\tau, s) u^2(\tau, s) d\tau \\
&= \int_t^T e^{-\rho(\tau - t)} Z_1(s)\eta(s) P(s) u^2(\tau, s) \underbrace{e^{-\int_t^\tau \mu^2 d\tau'}}_{\frac{S_2(\tau, s)}{S_2(t, s)}} e^{-\int_t^\tau \mu_{S_2}^2 S_2(\tau', s) d\tau'} d\tau \\
&= \frac{Z_1(s)\eta(s)P(s)}{S_2(t,s)} \int_t^T e^{-\rho(\tau - t)} S_2(\tau, s) u^2(\tau, s) \underbrace{e^{-\int_t^\tau \mu_{S_2}^2 S_2(\tau', s) d\tau'}}_{\geq 1} d\tau
\end{aligned}$$

$$\geq \frac{Z_1(s)\eta(s)P(s)}{S_2(t,s)}\xi_Z(t,s).$$

ξ_A - dynamics

As there is no terminal condition for ξ_A we can only derive the following equation from equation (74) which will still be crucial further down below.

$$\xi_A(t,s) = \xi_A(s,s) \cdot e^{\int_s^t (\rho - r - \bar{\mu})ds'} = \xi_A(T,s) \cdot e^{-\int_t^T (\rho - r - \bar{\mu})ds'}$$

ξ_E - dynamics

The dynamics for $\xi_E(t,s)$ described in equation (77) together with the terminal condition $\xi_E(T,s) = 0$ yield

$$\xi_E(t,s) = \int_t^T e^{-\int_t^\tau [\rho - f_E(\tau',s)]d\tau'} \left[w_E^2(\tau,s)\xi_A(\tau,s) + Z_2(\tau,s)S_2(\tau,s)u_E^2(\tau,s) - \mu_E^2(\tau,s)S_2(\tau,s)\xi_S(\tau,s) \right] d\tau$$

λ_Z - dynamics

Similar to the ξ_Z dynamics from equation (70) together with the terminal condition $\lambda_Z(T) = 0$ it directly follows that

$$\lambda_Z(t) = \int_t^T e^{-\int_t^\tau (\rho + \eta)ds'} \left[\eta(\tau)P(\tau)\xi_Z(\tau,\tau) + S_1(\tau)u^1(\tau) \right] d\tau$$

λ_A - dynamics

For our calculation of the λ_A-dynamics we first use equation (98) to modify the $\dot{\lambda}_A$-equation what results in

$$\dot{\lambda}_A = (\rho - r - \bar{\mu})\lambda_A - \xi_A(t,t),$$
$$\dot{\lambda}_A = (\rho - r - \bar{\mu})\lambda_A - \eta(t)P(t)\frac{u_{c_2}^2(t,t)}{u_{c_1}^1(t)}\lambda_A(t),$$
$$\dot{\lambda}_A = \left(\rho - r - \bar{\mu} - \eta(t)P(t)\frac{u_{c_2}^2(t,t)}{u_{c_1}^1(t)} \right)\lambda_A(t).$$

Now similar to the ξ_A-dynamics, without a terminal condition, we can still derive the equation below:

$$\lambda_A(s) = \exp\left(\int_t^s \rho - r - \bar{\mu}(s') - \eta(s')P(s')\frac{u_{c_2}^2(s',s')}{u_{c_1}^1(s')}ds' \right)\lambda_A(t).$$

λ_S - dynamics

Finally from equation (67) together with the terminal condition $\lambda_S(T) = 0$ we can derive the following representation of $\lambda_S(t)$.

$$\lambda_S(t) = \int_t^T e^{-\int_t^\tau (\rho + \mu^1(\tau') + \mu_{S_1}^1 S_1)d\tau'} \Big[Z_1 u^1 + \xi_E(\tau,\tau)B_{S_1} + \xi_S(\tau,\tau) -$$
$$- Z_1 \eta_{S_1} [\lambda_Z(\tau) - P(\tau)\xi_Z(\tau,\tau)] + \eta P_{S_1}(\tau)Z_1(\tau)\xi_Z(\tau,\tau) \Big] d\tau$$

Coming back to the different valuation of health, we are going to start with the simpler ones of the second stage and proceed to the more involved ones of the first stage.

7.3.1 Value of health in the second stage $\psi_H^2(t,s)$

$$
\psi_H^2(t,s) = \frac{S_2(t,s)\xi_S(t,s)}{\xi_A(t,s)} = \frac{S_2(t,s)\frac{Z_2(t,s)}{S_2(t,s)}\int_t^T e^{-\rho(\tau-t)}S_2(\tau,s)u^2(\tau,s)e^{-\int_t^\tau \mu_{S_2}^2 S_2(\tau',s)d\tau'}d\tau}{\xi_A(t,s)}
$$

$$
= \int_t^T \left(e^{-\rho(\tau-t)}Z_2(\tau,s)S_2(\tau,s)u^2(\tau,s)e^{-\int_t^\tau \mu_{S_2}^2 S_2(\tau',s)d\tau'}\frac{1}{\xi_A(\tau,s)}e^{\int_t^\tau \rho-r-\bar\mu d\tau'}\right)d\tau
$$

$$
= \int_t^T e^{-\int_t^\tau r+\bar\mu+\mu_{S_2}^2 S_2(\tau',s)d\tau'}\frac{Z_2(\tau,s)S_2(\tau,s)u^2(c_2(\tau,s),E(\tau,s))}{Z_2(\tau,s)S_2(\tau,s)u_{c_2}^2(c_2(\tau,s),E(\tau,s))}d\tau
$$

$$
= \int_t^T e^{-\int_t^\tau r+\bar\mu+\mu_{S_2}^2 S_2(\tau',s)d\tau'}\frac{u^2(c_2(\tau,s),E(\tau,s))}{u_{c_2}^2(c_2(\tau,s),E(\tau,s))}d\tau
$$

7.3.2 Value of morbidity $\psi_M(t,s)$

$$
\psi_M(t,s) = \frac{-\xi_E(t,s)}{\xi_A(t,s)} = \frac{\int_t^T\left(e^{-\int_t^\tau \rho-f_E d\tau'}\left[-w_E^2(\tau,s)\xi_A(\tau,s)-Z_2(\tau,s)u_E^2(\tau,s)+\mu_E^2(\tau,s)S_2(\tau,s)\xi_S(\tau,s)\right]\right)d\tau}{\xi_A(t,s)}
$$

$$
= \int_t^T e^{-\int_t^\tau \rho-f_E d\tau'}\left[-w_E^2(\tau,s)-\frac{Z_2(\tau,s)S_2(\tau,s)u_E^2(\tau,s)}{\xi_A(\tau,s)}+\mu_E^2(\tau,s)\frac{S_2(\tau,s)\xi_S(\tau,s)}{\xi_A(\tau,s)}\right]e^{\int_t^\tau \rho-r-\bar\mu d\tau'}d\tau
$$

$$
= \int_t^T e^{-\int_t^\tau r+\bar\mu-f_E d\tau'}\left[-w_E^2(\tau,s)-\frac{u_E^2(\tau,s)}{u_{c_2}^2(\tau,s)}+\mu_E^2(\tau,s)\psi_H^2\right]d\tau
$$

7.3.3 Value of acute survival $\psi_{AS}(t)$

$$
\psi_{AS}(t) = \frac{Z_1(t)\eta(t)\xi_Z(t,t)}{\xi_A(t,t)} = \frac{Z_1(t)\eta(t)\int_t^T e^{-\int_t^\tau \rho d\tau'}S_2(\tau,t)u^2(\tau,t)d\tau}{\xi_A(t,t)}
$$

$$
= Z_1(t)\eta(t)\int_t^T e^{-\int_t^\tau \rho d\tau'}e^{\int_t^\tau \rho-r-\bar\mu d\tau'}\frac{S_2(\tau,t)u^2(\tau,t)}{\xi_A(\tau,t)}d\tau
$$

$$
= Z_1(t)\eta(t)\int_t^T e^{-\int_t^\tau (r+\bar\mu)d\tau'}\frac{S_2(\tau,t)u^2(\tau,t)}{Z_2(\tau,t)S_2(\tau,t)u_{c_2}^2(\tau,t)}d\tau
$$

$$
= Z_1(t)\eta(t)\int_t^T e^{-\int_t^\tau (r+\bar\mu)d\tau'}\frac{u^2(\tau,t)}{Z_1(t)\eta(t)P(t)u_{c_2}^2(\tau,t)}d\tau
$$

$$
= \frac{1}{P(t)}\int_t^T\left[e^{-\int_t^\tau (r+\bar\mu)d\tau'}\frac{u^2(c_2(\tau,t),E(\tau,t))}{u_{c_2}^2(c_2(\tau,t),E(\tau,t))}\right]d\tau
$$

$$
=: \frac{1}{P(t)}\psi_{life}^2(t,t)
$$

7.3.4 Value of prevention ψ_P

$$
\psi_P(t) = \frac{Z_1(t)(\lambda_Z(t)-P\xi_Z(t,t))}{\lambda_A(t)} = \frac{Z_1(t)\lambda_Z(t)}{\lambda_A(t)}-\frac{Z_1(t)P\xi_Z(t,t)}{\lambda_A(t)} = \Pi_1 - \Pi_2
$$

$$
\Pi_1 = \frac{Z_1(t)\lambda_Z(t)}{\lambda_A(t)}
$$

$$
= \frac{1}{\lambda_A(t)}\int_t^T e^{-\int_t^\tau \rho ds}\left[Z_1(\tau)\eta(\tau)P(\tau)\xi_Z(\tau,\tau)+S_1(\tau)Z_1(\tau)u^1(\tau)\right]d\tau
$$

$$
= \int_t^T e^{-\int_t^\tau \rho ds}e^{\int_t^\tau\left(\rho-r-\bar\mu-\eta P\frac{u_{c_2}^2}{u_{c_1}^1}\right)ds}\left[\frac{Z_1(\tau)\eta(\tau)P(\tau)\xi_Z(\tau,\tau)}{\lambda_A(\tau)}+\frac{S_1(\tau)Z_1(\tau)u^1(\tau)}{\lambda_A(\tau)}\right]d\tau
$$

$$= \int_t^T e^{-\int_t^\tau \left(r+\bar\mu+\eta P \frac{u_{c_2}^2}{u_{c_1}^1}\right)ds} \left[\frac{Z_1(\tau)\eta(\tau)P(\tau)\xi_Z(\tau,\tau)}{\frac{\xi_A(\tau,\tau)}{\eta(\tau)P(\tau)}\frac{u_{c_1}^1(\tau)}{u_{c_2}^2(\tau,\tau)}} + \frac{S_1(\tau)Z_1(\tau)u^1(\tau)}{S_1(\tau)Z_1(\tau)u_{c_1}^1(\tau)} \right] d\tau$$

$$= \int_t^T e^{-\int_t^\tau \left(r+\bar\mu+\eta P \frac{u_{c_2}^2}{u_{c_1}^1}\right)ds} \left[\frac{u^1(\tau)}{u_{c_1}^1(\tau)} + \eta(\tau)P(\tau)\frac{u_{c_2}^2(\tau,\tau)}{u_{c_1}^1(\tau)} P(\tau)\frac{Z_1(\tau)\eta(\tau)\xi_Z(\tau,\tau)}{\xi_A(\tau,\tau)} \right] d\tau$$

$$= \int_t^T e^{-\int_t^\tau \left(r+\bar\mu+\eta P \frac{u_{c_2}^2}{u_{c_1}^1}\right)ds} \left[\frac{u^1(\tau)}{u_{c_1}^1(\tau)} + \eta(\tau)P(\tau)\frac{u_{c_2}^2(\tau,\tau)}{u_{c_1}^1(\tau)} \psi_{life}^2(\tau,\tau) \right] d\tau$$

$$\Pi_2 = \frac{Z_1(t)P(t)\xi_Z(t,t)}{\lambda_A(t)} = \frac{Z_1(t)P(t)\xi_Z(t,t)}{\frac{\xi_A(t,t)}{\eta(t)P(t)}\frac{u_{c_1}^1(t)}{u_{c_2}^2(t,t)}}$$

$$= \frac{u_{c_2}^2(t,t)}{u_{c_1}^1(t)}P(t)\frac{Z_1(t)\eta(t)\xi_Z(t,t)}{\xi_A(t,t)}P(t) = \frac{u_{c_2}^2(t,t)}{u_{c_1}^1(t)}P(t)\psi_{life}^2(t,t)$$

$$\psi_P(t) = \int_t^T e^{-\int_t^\tau \left(r+\bar\mu+\eta P \frac{u_{c_2}^2}{u_{c_1}^1}\right)ds} \left[\frac{u^1(\tau)}{u_{c_1}^1(\tau)} + \eta(\tau)P(\tau)\frac{u_{c_2}^2(\tau,\tau)}{u_{c_1}^1(\tau)} \psi_{life}^2(\tau,\tau) \right] d\tau - \frac{u_{c_2}^2(t,t)}{u_{c_1}^1(t)}P(t)\psi_{life}^2(t,t)$$

7.3.5 Value of health in the first stage $\psi_H^1(t)$

$$\psi_H^1(t) = \frac{S_1(t)\lambda_S(t)}{\lambda_A(t)} = \frac{S_1(t)}{\lambda_A(t)} \int_t^T e^{-\int_t^\tau (\rho+\mu^1(\tau')+\mu_{S_1}^1 S_1)d\tau'} \Big[Z_1 u^1 - Z_1\eta_{S_1}\left[\lambda_Z(\tau)-P(\tau)\xi_Z(\tau,\tau)\right] +$$

$$+ \xi_E(\tau,\tau)B_{S_1} + \xi_S(\tau,\tau) + \eta P_{S_1}(\tau)Z_1(\tau)\xi_Z(\tau,\tau) \Big] d\tau$$

$$= \int_t^T e^{-\int_t^\tau (\rho+\mu^1+\mu_{S_1}^1 S_1)d\tau'} e^{\int_t^\tau \mu^1 d\tau'} e^{\int_t^\tau \left(\rho-r-\bar\mu-\eta P \frac{u_{c_2}^2}{u_{c_1}^1}\right)d\tau'} \left[\frac{S_1(\tau)}{\lambda_A(\tau)}Z_1 u^1 - \right.$$

$$\left. - \frac{S_1(\tau)}{\lambda_A(\tau)}\eta_{S_1}Z_1\left[\lambda_Z(\tau)-P(\tau)\xi_Z(\tau,\tau)\right] + \frac{S_1(\tau)}{\lambda_A(\tau)}\xi_E(\tau,\tau)B_{S_1} + \frac{S_1(\tau)}{\lambda_A(\tau)}\xi_S(\tau,\tau) + \frac{S_1(\tau)}{\lambda_A(\tau)}\eta P_{S_1}(\tau)Z_1(\tau)\xi_Z(\tau,\tau) \right] d\tau$$

$$= \int_t^T e^{-\int_t^\tau \left(r+\bar\mu+\mu_{S_1}^1 S_1+\eta P \frac{u_{c_2}^2}{u_{c_1}^1}\right)d\tau'} \left[\frac{u^1(\tau)}{u_{c_1}^1(\tau)} - \eta_{S_1}S_1\frac{Z_1\left[\lambda_Z(\tau)-P(\tau)\xi_Z(\tau,\tau)\right]}{\lambda_A(\tau)} + \right.$$

$$\left. + \eta P\frac{u_{c_2}^2}{u_{c_1}^1} \left(B_{S_1}S_1(\tau)\frac{\xi_E(\tau,\tau)}{\xi_A(\tau,\tau)} + \frac{S_2(\tau,\tau)\xi_S(\tau,\tau)}{\xi_A(\tau,\tau)} + P_{S_1}(\tau)S_1(\tau)\frac{Z_1(\tau)\eta(\tau)\xi_Z(\tau,\tau)}{\xi_A(\tau,\tau)} \right) \right] d\tau$$

$$= \int_t^T e^{-\int_t^\tau \left(r+\bar\mu+\mu_{S_1}^1 S_1+\eta P \frac{u_{c_2}^2}{u_{c_1}^1}\right)d\tau'} \left[\frac{u^1(\tau)}{u_{c_1}^1(\tau)} - \eta_{S_1}S_1\psi_P(\tau) + \right.$$

$$\left. + \eta P\frac{u_{c_2}^2}{u_{c_1}^1}\left\{(-B_{S_1})S_1(\tau)\psi_M(\tau,\tau) + \psi_H^2 + P_{S_1}(\tau)S_1(\tau)\psi_{AS}(\tau)\right\} \right] d\tau$$

$$= \int_t^T e^{-\int_t^\tau \left(r+\bar\mu+\mu_{S_1}^1 S_1+\eta P \frac{u_{c_2}^2}{u_{c_1}^1}\right)d\tau'} \left[\frac{u^1(\tau)}{u_{c_1}^1(\tau)} + (-\eta_{S_1})S_1\psi_P(\tau) + \right.$$

$$\left. + \eta P\frac{u_{c_2}^2}{u_{c_1}^1}\left\{\psi_H^2 + \frac{P_{S_1}(\tau)S_1(\tau)}{P(\tau)}\psi_{life}^2 + (-B_{S_1})S_1(\tau)\psi_M(\tau,\tau)\right\} \right] d\tau$$

7.4 Proof of proposition 2 on the first order optimality conditions

Under the assumption of interior solutions the FOCs (79), (81), (83), (85), (87), (89) and (91) all have to hold in strict form. Simple rearrangements of these equations and substituting the different valuations of health in equations (93)-(97) result in the equations presented in proposition 2.

7.5 Derivation of the Euler equations

In this section we present the derivation of all the Euler equations presented in the numerical analysis.

7.5.1 c_1 - Dynamics

$$\lambda_A(t) = Z_1(t)S_1(t)u^1_{c_1}(c_1(t))$$

$$\frac{\dot{\lambda}_A(t)}{\lambda_A(t)} = \frac{\dot{Z}_1(t)}{Z_1(t)} + \frac{\dot{S}_1(t)}{S_1(t)} + \frac{\frac{d}{dt}\left[u^1_{c_1}(c_1(t))\right]}{u^1_{c_1}(c_1(t))}$$

$$\rho - r - \bar{\mu} - \eta P \frac{u^2_{c_2}}{u^1_{c_1}} = -\eta - \mu^1 + \frac{u^1_{c_1 c_1} c_1}{u^1_{c_1}} \frac{\dot{c}_1}{c_1}$$

$$\frac{\dot{c}_1}{c_1} = \frac{u^1_{c_1}}{-u^1_{c_1 c_1} c_1}\left[r - \rho + \bar{\mu} - \mu^1 - \eta + \eta P \frac{u^2_{c_2}}{u^1_{c_1}}\right]$$

7.5.2 c_2 - Dynamics

$$\xi_A(t,s) = Z_2(t,s)S_2(t,s)u^2_{c_2}(c_2(t,s), E(t,s))$$

$$\frac{\dot{\xi}_A(t,s)}{\xi_A(t,s)} = \frac{\dot{Z}_2(t,s)}{Z_2(t,s)} + \frac{\dot{S}_2(t,s)}{S_2(t,s)} + \frac{\frac{d}{dt}\left[u^2_{c_2}(c_2(t,s), E(t,s))\right]}{u^2_{c_2}(c_2(t,s), E(t,s))}$$

$$\rho - r - \bar{\mu} = 0 - \mu^2 + \frac{u^2_{c_2 c_2} c_2}{u^2_{c_2}} \frac{\dot{c}_2}{c_2} + \frac{u^2_{c_2 E}}{u^2_{c_2}} f$$

$$\frac{\dot{c}_2}{c_2} = \frac{u^2_{c_2}}{-u^2_{c_2 c_2} c_2}\left[r - \rho + \bar{\mu} - \mu^2 + \frac{u^2_{c_2 E}}{u^2_{c_2}} f\right]$$

7.5.3 b_2 - Dynamics

$$p^b(t) = \left[-\mu^2_{b_2}(t)\right] \cdot \psi^2_H(t)$$

$$\frac{\dot{p}^b(t)}{p^b(t)} = \frac{\frac{d}{dt}\left[-\mu^2_{b_2}(t)\right]}{-\mu^2_{b_2}(t)} + \frac{\dot{\psi}^2_H(t)}{\psi^2_H(t)}$$

$$\frac{\dot{p}^b}{p^b} = \frac{\mu^2_{b_2 b_2} b_2}{\mu^2_{b_2}} \frac{\dot{b}_2}{b_2} + \frac{\mu^2_{b_2 S_2} \dot{S}_2}{\mu^2_{b_2}} + \frac{\mu^2_{b_2 t}}{\mu^2_{b_2}} + \frac{\dot{\psi}^2_H(t)}{\psi^2_H(t)}$$

$$\frac{\dot{b}_2}{b_2} = \frac{-\mu^2_{b_2}}{\mu^2_{b_2 b_2} b_2}\left[\frac{\dot{\psi}^2_H(t)}{\psi^2_H(t)} - \frac{\dot{p}^b}{p^b} - \frac{\mu^2_{b_2 S_2}}{\mu^2_{b_2}} \mu^2 S_2 + \frac{\mu^2_{b_2 t}}{\mu^2_{b_2}}\right]$$

$$\dot{\psi}^2_H = \left(r + \bar{\mu} + \mu^2_{S_2} S_2\right)\psi^2_H - \frac{u^2}{u^2_{c_2}}$$

$$\frac{\dot{b}_2}{b_2} = \frac{-\mu^2_{b_2}}{\mu^2_{b_2 b_2} b_2}\left[r + \bar{\mu} + \mu^2_{S_2} S_2 - \frac{u^2/u^2_{c_2}}{\psi^2_H} - \frac{\dot{p}^b}{p^b} - \frac{\mu^2_{b_2 S_2}}{\mu^2_{b_2}} \mu^2 S_2 + \frac{\mu^2_{b_2 t}}{\mu^2_{b_2}}\right]$$

7.5.4 b_1 - Dynamics

$$p^b(t) = \left[-\mu^1_{b_1}(t)\right] \cdot \psi^1_H(t)$$

$$\frac{\dot{p^b}(t)}{p^b(t)} = \frac{\frac{d}{dt}\left[-\mu_{b_1}^1(t)\right]}{-\mu_{b_1}^1(t)} + \frac{\dot{\psi_H^1}(t)}{\psi_H^1(t)}$$

$$\frac{\dot{p^b}}{p^b} = \frac{\mu_{b_1b_1}^1 b_1}{\mu_{b_1}^1}\frac{\dot{b_1}}{b_1} + \frac{\mu_{b_1S_1}^1 \dot{S_1}}{\mu_{b_1}^1} + \frac{\mu_{b_1t}^1}{\mu_{b_1}^1} + \frac{\dot{\psi_H^1}(t)}{\psi_H^1(t)}$$

$$\frac{\dot{b_1}}{b_1} = \frac{-\mu_{b_1}^1}{\mu_{b_1b_1}^1 b_1}\left[\frac{\dot{\psi_H^1}(t)}{\psi_H^1(t)} - \frac{\dot{p^b}}{p^b} - \frac{\mu_{b_1S_1}^1}{\mu_{b_1}^1}\mu^1 S_1 + \frac{\mu_{b_1t}^1}{\mu_{b_1}^1}\right]$$

$$\dot{\psi_H^1} = \left(r + \bar{\mu} + \mu_{S_1}^1 S_1 + \eta P\frac{u_{c_2}^2}{u_{c_1}^1}\right)\psi_H^1 - \frac{u^1}{u_{c_1}^1} + \eta_{S_1}S_1\psi_P - \eta P\frac{u_{c_2}^2}{u_{c_1}^1}\left\{\psi_H^2 + \frac{P_{S_1}S_1}{P}\psi_{life}^2 + (-B_{S_1})S_1\psi_M\right\}$$

$$\frac{\dot{\psi_H^1}}{\psi_H^1} = r + \bar{\mu} + \mu_{S_1}^1 S_1 - \frac{u^1/u_{c_1}^1}{\psi_H^1} + \eta_{S_1}S_1\frac{\psi_P}{\psi_H^1} - \eta P\frac{u_{c_2}^2}{u_{c_1}^1}\left\{\frac{(\psi_H^2 - \psi_H^1)}{\psi_H^1} + \frac{P_{S_1}S_1}{P}\frac{\psi_{life}^2}{\psi_H^1} + (-B_{S_1})S_1\frac{\psi_M}{\psi_H^1}\right\}$$

$$\frac{\dot{b_1}}{b_1} = \frac{-\mu_{b_1}^1}{\mu_{b_1b_1}^1 b_1}\left[r + \bar{\mu} + \mu_{S_1}^1 S_1 - \frac{u^1/u_{c_1}^1}{\psi_H^1} + \eta_{S_1}S_1\frac{\psi_P}{\psi_H^1} - \right.$$

$$\left. - \eta P\frac{u_{c_2}^2}{u_{c_1}^1}\left\{\frac{(\psi_H^2 - \psi_H^1)}{\psi_H^1} + \frac{P_{S_1}S_1}{P}\frac{\psi_{life}^2}{\psi_H^1} + (-B_{S_1})S_1\frac{\psi_M}{\psi_H^1}\right\} - \frac{\dot{p^b}}{p^b} - \frac{\mu_{b_1S_1}^1}{\mu_{b_1}^1}\mu^1 S_1 + \frac{\mu_{b_1t}^1}{\mu_{b_1}^1}\right]$$

7.5.5 h_2 - Dynamics

$$p^2(t) = [-f_{h_2}(t,s)]\cdot\psi_M(t,s)$$

$$\frac{\dot{p^2}(t)}{p^2(t)} = \frac{\frac{d}{dt}[-f_{h_2}(t,s)]}{[-f_{h_2}(t,s)]} + \frac{\dot{\psi_M}(t,s)}{\psi_M(t,s)}$$

$$\frac{\dot{p^2}}{p^2} = \frac{f_{h_2h_2}h_2}{f_{h_2}}\frac{\dot{h_2}}{h_2} + \frac{f_{h_2t}}{f_{h_2}} + \frac{f_{h_2E}\dot{E}}{f_{h_2}} + \frac{\dot{\psi_M}(t,s)}{\psi_M(t,s)}$$

$$\frac{\dot{h_2}}{h_2} = \frac{-f_{h_2}}{f_{h_2h_2}h_2}\left[\frac{\dot{\psi_M}(t,s)}{\psi_M(t,s)} + \frac{f_{h_2t}}{f_{h_2}} + \frac{f_{h_2E}f}{f_{h_2}} - \frac{\dot{p^2}}{p^2}\right]$$

$$\dot{\psi_M} = (r + \bar{\mu} - f_E)\psi_M + w_E^2 + \frac{u_E^2}{u_{c_2}^2} - \mu_E^2\psi_H^2$$

$$\frac{\dot{h_2}}{h_2} = \frac{-f_{h_2}}{f_{h_2h_2}h_2}\left[r + \bar{\mu} - f_E + \frac{w_E^2}{\psi_M} + \frac{u_E^2/u_{c_2}^2}{\psi_M} - \mu_E^2\frac{\psi_H^2}{\psi_M} + \frac{f_{h_2t}}{f_{h_2}} + \frac{f_{h_2E}f}{f_{h_2}} - \frac{\dot{p^2}}{p^2}\right]$$

7.5.6 h_1 - Dynamics

$$p^1(t) = [-\eta_{h_1}(t)]\cdot\psi_P(t)$$

$$\frac{\dot{p^1}(t)}{p^1(t)} = \frac{\frac{d}{dt}[-\eta_{h_1}(t)]}{[-\eta_{h_1}(t)]} + \frac{\dot{\psi_P}(t)}{\psi_P(t)}$$

$$\frac{\dot{p^1}}{p^1} = \frac{\eta_{h_1h_1}h_1}{\eta_{h_1}}\frac{\dot{h_1}}{h_1} + \frac{\eta_{h_1S_1}\dot{S_1}}{\eta_{h_1}} + \frac{\eta_{h_1t}}{\eta_{h_1}} + \frac{\dot{\psi_P}}{\psi_P}$$

$$\frac{\dot{h_1}}{h_1} = \frac{-\eta_{h_1}}{\eta_{h_1h_1}h_1}\left[\frac{\dot{\psi_P}}{\psi_P} - \frac{\eta_{h_1S_1}\mu^1 S_1}{\eta_{h_1}} + \frac{\eta_{h_1t}}{\eta_{h_1}} - \frac{\dot{p^1}}{p^1}\right]$$

$$\psi_P(t) = \int_t^T e^{-\int_t^\tau \left(r+\bar{\mu}+\eta P \frac{u_{c_2}^2}{u_{c_1}^1}\right)ds} \left[\frac{u^1(\tau)}{u_{c_1}^1(\tau)} + \eta(\tau)P(\tau)\frac{u_{c_2}^2(\tau,\tau)}{u_{c_1}^1(\tau)}\psi_{life}^2(\tau,\tau)\right] d\tau - \frac{Z_1(t)P(t)\xi_Z(t,t)}{\lambda_A(t)}$$

$$= \Pi_1 - \Pi_2$$

$$\dot{\Pi}_1 = \left(r+\bar{\mu}+\eta P \frac{u_{c_2}^2}{u_{c_1}^1}\right)\Pi_1 - \frac{u^1}{u_{c_1}^1} - \eta P \frac{u_{c_2}^2}{u_{c_1}^1}\psi_{life}^2$$

$$\dot{\Pi}_2 = \left(\frac{\dot{Z}_1}{Z_1} + \frac{\dot{P}}{P} + \frac{\frac{d}{dt}\xi_Z(t,t)}{\xi_Z} - \frac{\dot{\lambda}_A}{\lambda_A}\right)\Pi_2$$

$$= \left(-\eta + \frac{\dot{P}}{P} + \frac{\dot{\xi}_Z}{\xi_Z} + \frac{\frac{\partial}{\partial s}\xi_Z}{\xi_Z} - \left(\rho - r - \bar{\mu} - \eta P\frac{u_{c_2}^2}{u_{c_1}^1}\right)\right)\Pi_2$$

$$\frac{\dot{\xi}_Z}{\xi_Z} = \frac{\rho\xi_Z - S_2 u^2}{\xi_Z} = \rho - \frac{S_2 u^2}{\xi_Z} = \rho - S_2 u^2 \frac{Z_1\eta}{\xi_A\psi_{AS}} = \rho - \frac{Z_1\eta S_2 u^2}{Z_2 S_2 u_{c_2}^2 \psi_{AS}} = \rho - \frac{Z_1\eta S_2 u^2}{Z_1\eta P S_2 u_{c_2}^2 \psi_{AS}} = \rho - \frac{u^2}{u_{c_2}^2\psi_{life}^2}$$

$$\dot{\Pi}_2 = \left(r+\bar{\mu}-\eta+\eta P\frac{u_{c_2}^2}{u_{c_1}^1} + \frac{\dot{P}}{P} + \frac{\frac{\partial}{\partial s}\xi_Z}{\xi_Z}\right)\Pi_2 - \frac{u^2}{u_{c_2}^2\psi_{life}^2}\Pi_2$$

$$= \left(r+\bar{\mu}-\eta+\eta P\frac{u_{c_2}^2}{u_{c_1}^1} + \frac{\dot{P}}{P} + \frac{\frac{\partial}{\partial s}\xi_Z}{\xi_Z}\right)\Pi_2 - \frac{u^2}{u_{c_1}^1}P$$

$$\dot{\psi}_P(t) = \dot{\Pi}_1 - \dot{\Pi}_2 = \left(r+\bar{\mu}+\eta P\frac{u_{c_2}^2}{u_{c_1}^1}\right)\Pi_1 - \frac{u^1}{u_{c_1}^1} - \eta P\frac{u_{c_2}^2}{u_{c_1}^1}\psi_{life}^2 -$$

$$- \left(\left(r+\bar{\mu}-\eta+\eta P\frac{u_{c_2}^2}{u_{c_1}^1} + \frac{\dot{P}}{P} + \frac{\frac{\partial}{\partial s}\xi_Z}{\xi_Z}\right)\Pi_2 - \frac{u^2}{u_{c_1}^1}P\right)$$

$$= \left(r+\bar{\mu}+\eta P\frac{u_{c_2}^2}{u_{c_1}^1}\right)\psi_P - \frac{(u^1 - Pu^2)}{u_{c_1}^1} - \left(\frac{\dot{P}}{P} + \frac{\frac{\partial}{\partial s}\xi_Z}{\xi_Z}\right)\frac{u_{c_2}^2}{u_{c_1}^1}P\psi_{life}^2$$

Since it also holds that

$$\xi_Z(t,t) = V^*(t,S_1(t),A_1(t),Z_1(t)) =: V^*(t)$$

we can alternatively derive

$$\dot{\psi}_P(t) = \left(r+\bar{\mu}+\eta P\frac{u_{c_2}^2}{u_{c_1}^1}\right)\psi_P - \frac{u^1}{u_{c_1}^1} + \left(\rho - \frac{\dot{P}}{P} - \frac{\frac{d}{dt}V^*(t)}{V^*(t)}\right)\frac{u_{c_2}^2}{u_{c_1}^1}P\psi_{life}^2$$

$$\frac{\dot{\psi}_P}{\psi_P} = r+\bar{\mu}+\eta P\frac{u_{c_2}^2}{u_{c_1}^1} - \frac{(u^1 - Pu^2)/u_{c_1}^1}{\psi_P} - \left(\frac{\dot{P}}{P} + \frac{\frac{\partial}{\partial s}\xi_Z}{\xi_Z}\right)\frac{u_{c_2}^2}{u_{c_1}^1}\frac{P\psi_{life}^2}{\psi_P}$$

$$= r + \bar{\mu} + \eta P \frac{u_{c_2}^2}{u_{c_1}^1} - \frac{u^1/u_{c_1}^1}{\psi_P} + \left(\rho - \frac{\dot{P}}{P} - \frac{\frac{d}{dt}V^*(t)}{V^*(t)} \right) \frac{u_{c_2}^2}{u_{c_1}^1} \frac{P \psi_{life}^2}{\psi_P}$$

$$\frac{\dot{h_1}}{h_1} = \frac{-\eta_{h_1}}{\eta_{h_1 h_1} h_1} \left[r + \bar{\mu} + \eta P \frac{u_{c_2}^2}{u_{c_1}^1} - \frac{(u^1 - Pu^2)/u_{c_1}^1}{\psi_P} - \left(\frac{\dot{P}}{P} + \frac{\frac{\partial}{\partial s} \xi_Z}{\xi_Z} \right) \frac{u_{c_2}^2}{u_{c_1}^1} \frac{P \psi_{life}^2}{\psi_P} + \frac{\eta_{h_1 S_1} \dot{S}_1}{\eta_{h_1}} + \frac{\eta_{h_1 t}}{\eta_{h_1}} - \frac{\dot{p^1}}{p^1} \right]$$

$$= \frac{-\eta_{h_1}}{\eta_{h_1 h_1} h_1} \left[r + \bar{\mu} + \eta P \frac{u_{c_2}^2}{u_{c_1}^1} - \frac{u^1/u_{c_1}^1}{\psi_P} + \left(\rho - \frac{\dot{P}}{P} - \frac{\frac{d}{dt}V^*(t)}{V^*(t)} \right) \frac{u_{c_2}^2}{u_{c_1}^1} \frac{P \psi_{life}^2}{\psi_P} + \frac{\eta_{h_1 S_1} \dot{S}_1}{\eta_{h_1}} + \frac{\eta_{h_1 t}}{\eta_{h_1}} - \frac{\dot{p^1}}{p^1} \right]$$

7.6 Estimation and calibration strategies

Estimation of cancer incidence rate $\eta(S_1)$

For the estimation of the cancer incidence rate we adjust the total age-specific mortality rate obtained from the human mortality database (University of California, Berkeley (USA) and Max Planck Institute for Demographic Research (Germany) (2020)) for the average age-specific cancer mortality provided in the SEER-database (Surveillance Research Program, National Cancer Institute (2020)). Using this adjusted mortality profile, we can derive the empirical equivalent of the S_1-profile in our model. Using the age-specific cancer incidence rate (again from the SEER database), the functional form

$$\eta(t, S) = \frac{\beta_0}{1 + \beta_1 (\frac{1-S_1}{S_1})^{\beta_2}}$$

turns out to provide an appropriate fit. Consequently we use a general least square fitting function for non-linear functions to obtain estimates for β_0, β_1, and β_2.

Estimation of cancer specific mortality rate $\mu^m(s, E)$

The estimation procedure the parameters of the cancer specific mortality rate is more involved. Using the duration dependent cancer specific survival data for different ages at diagnosis (4 broad age groups), we first derive the corresponding empirical information about the cancer mortality rate. The remaining procedure is less straight forward. Compared to the estimation of the cancer incidence rate, we do not have an empirical counterpart for the development of the cancer stock E. First note, that we can derive the empirical duration specific mortality only between full years. However especially within the first two years after a diagnosis, there can be significant differences in cancer mortality risk and expenditure between the beginning and the end of the first year. Hence the derived empirical mortality rates are in a continuous time setting more appropriate estimates for mortality at 6 months, 1.5 years, 2.5 years, etc. Furthermore our specified cancer mortality function

$$\mu^m(t, s, E) = \psi_0 \cdot E \cdot \exp\left\{ \psi_1 \cdot \left(\frac{s}{T} \right)^{\psi_2} \right\}$$

and the normalisation of the initial cancer stock $E(s, s) = 1.0$ imply that we can potentially estimate the impact of age at diagnosis using mortality data right at the point of diagnosis. Hence we use the cancer specific mortality data for 6 months, 1.5 years, 2.5 years to estimate a non-linear function

$$M^j(t - s) = \exp\left\{ -a_1 + a_2 \cdot e^{-a_3(t-s)} \right\}$$

for each age-group $j \in \{1, 2, 3, 4\}$ separately (which provides excellent fit) and use it to derive extrapolated estimations for the mortality rates right after the diagnosis, i.e. $M^j(0)$.

However the data shows that the impact of the age at diagnosis can change for increasing duration of cancer. Therefore estimating ψ_1 and ψ_2 only from data $M^j(0)$, might lead to skewed results, which are not able to replicate the mortality risk of cancer also in the long-run. Hence we evaluated $M^j(t)$ at $t = 0, 1, \ldots, 10$ to obtain a relevant data-set $\mathcal{M}$. While we still do not know the development of E ex-ante, we can estimate ψ_0, ψ_1, and ψ_2 as follows:

(i) We assume a highly general mortality function (with $d = t - s$ describing the duration of the diagnosed cancer disease) in the form of

$$\widetilde{\mu^m}(d, s) = \nu_d \cdot \exp\left\{ \zeta_1 \cdot \left(\frac{s}{T} \right)^{\zeta_2} \right\}$$

(ii) Now we can use the dataset $\mathcal{M}$ to estimate the 13 parameters $\nu_0, \nu_1 \ldots \nu_{10}$, ζ_1, and ζ_2 using a generalised least-square fit. However to obtain the corresponding mortality rates for the four age-groups, we need to calculate the averaged values within each age-group. Thereby we need to account for the varying incidence of cancer depending on age and use the empirical equivalents to S_1, Z_1, and Z_2, which we can all derive from the estimates of the cancer incidence rate.

(iii) Finally we can use the estimates for ζ_1 and ζ_2 as estimates for ψ_1 and ψ_1 in μ^m and the estimated value for ν_0 provides an approximation for ψ_0.

Calibration of the base mortality function $\mu^b(t, b)$

We use a very general form of the mortality function, so we are able to match the health expenditure profile and the survival profile at the same time, i.e. we assume the functional form presented in Section 4.

$$\mu^b(t, S_1, b_1) = \mu^b(t, b_1) = g(t) b_1^{\varepsilon(t)} \qquad\qquad g(t) = \exp\left\{ \gamma_0 + \gamma_1 t + \gamma_2 t^2 + \gamma_3 t^3 \right\}$$
$$\varepsilon(t) = \alpha_0 + \alpha_1 \cdot t \quad (< 0)$$

As the data delivers better information about the health expenditure rather than the health investments/measures we will directly connect the mortality rate to the health expenditures and as a result also set the price for health investments $p^b = 1$. To find fitting values for γ_i and α_i we use an iterative procedure:

(i) We start with a reasonable estimate for the value of health profile in the first stage.

(ii) Now consider the first-order optimality condition

$$(-\mu^1_{b_1}(t, b_1(t)))\psi^1_H(t) = p^b(= 1)$$

and the fact, that for our functional specification it holds that

$$\mu^1_{b_1}(t, b_1(t)) = \varepsilon(t)\frac{\mu^1(t, b_1(t))}{b_1(t)}$$

(iii) Consequently we can rewrite the optimality condition as:

$$\varepsilon(t) = \frac{b_1(t)}{-\mu^1(t, b_1(t))\psi^1_H(t)}$$

Inserting the general expenditure data $b_1^{Data}(t)$ obtained from the NTA database, the calculated value of health profile and the non-cancer mortality rates from the data, we get an estimated profile for $\varepsilon(t)$, which can then be used to estimate the parameters α_1 and α_2.

(iv) Having found the parameters for $\varepsilon(t)$, we can then simply use the definition of the mortality function to find the parameters $\gamma_i, i \in \{1, 2, 3, 4\}$ by using

$$\ln(g(t)) = \ln\left(\frac{\mu^1(t)}{b(t)^{\varepsilon(t)}}\right)$$

and inserting the mortality and health expenditure data together with the estimate for $\varepsilon(t)$ on the right side of the equation.

(v) As a next step we undertake several steps in our general optimisation algorithm to get closer to the optimal solution.

(vi) After a certain number of steps we stop and calculate the a-priori value of health (average value of health in the first and second stage) for the current solution of controls, states and costates and return to step (ii).

As the optimal solution converges during the process, also the calibration for the parameters γ_i and α_i should converge for proper guesses for the initial choices.