



HAL
open science

Individual behaviours when facing health risk and their aggregate impacts on the society

Ze Zhong Shang

► **To cite this version:**

Ze Zhong Shang. Individual behaviours when facing health risk and their aggregate impacts on the society. Economics and Finance. Le Mans Université, 2019. English. NNT : 2019LEMA2001 . tel-02383237

HAL Id: tel-02383237

<https://theses.hal.science/tel-02383237>

Submitted on 27 Nov 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

THESE DE DOCTORAT DE

LE MANS UNIVERSITE
COMUE UNIVERSITE BRETAGNE LOIRE

ECOLE DOCTORALE N° 597
Sciences Economiques et sciences De Gestion
Spécialité : Science économique

Par

« Ze Zhong SHANG »

« Individual behaviours when facing health risk and their aggregate impacts on the society »

Thèse présentée et soutenue à « LE MANS », le « 03/05/2019 »
Unité de recherche : Groupe d'Analyse des Itinéraires et des Niveaux Salariaux
Thèse N° :2019LEMA2001

Rapporteurs :

Ferhat MIHOUBI Professeur à l'Université Paris-Est
Emmanuel DUGUET Professeur à l'Université Paris-Est

Composition du Jury :

Président :	Thomas WEITZENBLUM	Professeur à l'Université du Mans
Examineurs :	Ferhat MIHOUBI	Professeur à l'Université Paris-Est
	Emmanuel DUGUET	Professeur à l'Université Paris-Est
	Pierre-Carl MICHAUD	Professeur à HEC Montréal
Dir. de thèse :	Xavier FAIRISE	Professeur à l'Université du Mans
Co-dir. de thèse :	Laurent DENIS	Professeur à l'Université du Mans
	François LANGOT	Professeur à l'Université du Mans

Titre : Les choix individuels en face du risque de santé et leur impact agrégé sur l'ensemble de l'économie

Mots clés : Santé; Risque; Assurance; Inégalité; Choix d'individu hétérogène; Economie

Résumé : La santé joue un rôle de plus en plus important dans l'économie dans les dernières années: d'une part, on observe une amélioration significative du niveau moyen de la durée de vie, d'autre part, il y a également une forte croissance sur les dépenses de santé. Dans cette thèse, on commence par présenter ces 2 faits stylisés mais on présente également que même s'il y a une amélioration du niveau moyen de la durée de vie, l'inégalité de la santé est toujours rapportée est en effet a tendance d'accroître, on présente aussi que les pays dont les dépenses de santé sont assez importantes en général n'ont pas un système de santé efficace. Par conséquent, cette amélioration de la santé humaine semble de bénéficier plus les gens qui peuvent payer le coût. Afin de trouver les raisons qui cause ce phénomène et proposer des solutions pratiques qui permet de résoudre ce problème, dans cette thèse, on prend 2 approches :

Title: Individual behaviors when facing health risk and their aggregate impacts on the society

Keywords: Health; Risk; Insurance; Inequality; Heterogeneous individual behaviors; Economy

Abstract: During the past 2 decades, health has become a more and more important role in our economy life: on the one hand, we observe a significant improvement of average lifespan across the globe, on the other hand, the health expenditures are also increasing enormously, which has become a shake to the public health system of many countries. In this dissertation, we begin with these 2 stylized facts but we also show that there is more to it: though the average level of health has been improved significantly, health related inequalities are still being reported and actually tend to increase, plus, we also show that countries with important health expenditures performs generally poorly in terms of health system efficiency, in short, this improvement of human health we are talking about seems to benefit more those who can pay the bills.

In order to figure out what causes this phenomenon and eventually propose practical solutions that help solve the problem, in this dissertation we take 2 approaches: first we start with the deterministic approach and also the more theoretical approach, in this approach we build our model based on the Grossman health capital model and we examine how people would behave when their health fluctuates, then we turn to the second approach, which is the stochastic approach and also the more practical approach, in this approach we use the Markov chain to simulate the real health risk and examine the behaviors of individuals of different social-economic status(SES) under this circumstance, we also aggregate these behaviors to see what impact could be generated on the whole society and we test reversely how these behavior would affected by public policies.

Abstract

Health, as one of the most important indicators of human life, is frequently measured by the famous self-assessed health status, indeed, health is rather a subjective sentiment than some objective measurements, uniqueness resides in its very nature. If we are to analyse health and its impact on human life, it is therefore essential to understand how individual reacts when their health varies. In this thesis I first focus on individual level to study what should be their decisions when their health varies, then I aggregate these decisions to see what impacts could be generated on the whole society. This dissertation is composed of 3 chapters, first I present some stylized facts and do some reviews on existed literatures, I also propose some questions that I aim to solve by this thesis, then I present 2 models that contribute to solve the problems that proposed at the end of the first chapter. Results of this dissertation show that first, a simplified Grossman health capital model with endogenous lifespan can be constructed, it's able to replicate a growing profile of health spending and a significant level of health spending at the end of life, the numerical solution it produces also fits roughly well the U.S. data, second, an augmented Aiyagari model with stochastic health shocks can reproduce pretty well some key macroeconomic moments of health economics, such a model can be therefore used to study the impacts of different public policies on the whole society: in this study, I compare the differences of public health policies between France and the U.S., the results show that low unemployment risk, larger costs of health services and higher risky behaviours contribute to explain the poorer health of the Americans than French.

Acknowledgements

First I want to present my profound gratitude to my three directors, Laurent Denis, Xavier Fairise and François Langot, they've been continuously teaching me with patience along these years, their understandings on economics, enthusiasm for research works, their kindness and tolerance of a sometimes stubborn and lazy student (like me) set me a model that I wish one day I can catch up with. I am sincerely thankful to have them as the directors of my thesis.

I would also like to thank my parents, for all the love they give me, when I decided to go to France for university studies 10 years ago, I was called crazy and nuts by my other family members, friends of my parents told them that I would never make it and it's all a waste of money and time, but still they took the risk and sent me aboard and give me encouragements continuously, without their help, I could never make all this happen.

My wife gives me invaluable help, she gives me accompaniment and comforts when I mostly need them. I always think I might spend the luck of my entire life just to meet her, I love her with all my heart.

My sincere thanks go to my colleagues in Laboratory GAINS, for being kind and welcome me warmly, it was fantastic to have the opportunity to work with so many outstanding people.

Last but not least, a very special gratitude goes out to the project PANORisk, my research would have been impossible without the help and the funding from it.

Contents

1 General Introduction	4
2 Stylized facts and review of literatures	9
2.1 Introduction	9
2.2 The 2 aspects of health inequalities	10
2.3 A simplified interpretation of the Grossman model	22
2.4 Justification of the lack of financial market in Laporte's model	33
2.5 Optimal health investment choice strategy under uncertainty	41
2.5.1 The risk of health	42
2.5.2 The risk of income	52
2.6 Conclusion	55
3 A Life-Cycle Model with health spending and endogenous lifetime	57
3.1 Introduction	57
3.2 A life-cycle model with health capital	60
3.2.1 Presentation of the model	60
3.2.2 The constant term in the utility function	67

3.3 Numerical investigations	74
3.3.1 Calibration strategy	74
3.3.2 Numerical results	75
3.3.3 Sensitivity analysis	78
3.4 Conclusion	83
4 A comparison between the US and France based on a general equilib-	
rium mode	85
4.1 Introduction	85
4.2 The theoretical model	89
4.3 Calibration and results	96
4.3.1 Calibration strategy	96
4.3.2 Results	97
4.4 Counterfactual experiments	99
4.5 Conclusion	104
5 General Conclusion	105
Bibliography	108
Appendix A One example of the DEA method	114
Appendix B The DEA methodology	122
B.1 The CCR mode	122
B.1.1 The multiplier form	122

B.1.2 The envelopment form	127
B.2 The BCC mode	131
B.2.1 The multiplier form	131
B.2.2 The envelopment form	138
Appendix C Input-oriented and output-oriented efficiency score	139
Appendix D The Health Gradient	144
Appendix E Equilibrium on financial markets	145

Chapter 1

General Introduction

In 2017, the share of health spending on GDP of the U.S. is about 17.2%, by 2020, this share is expected to reach 20%, of course, among all OECD countries, the U.S. ranks at the top when we compare the health spending, but even if we look at the bottom of this list, the countries that rank at the bottom, Mexico and Turkey, still has a significant level in terms of health spending-GDP ratio at around 5.4% and 4.1% respectively. The rapid growth of health investment could be related to multiple factors, on the one hand, progress in medical technology push up the price of health related goods and services, but the increase in incomes help ease the burden(Suen,2006)[48], on the other hand, increased life expectancy make households invest more in health while health spending at elder age is particularly expensive(Hall and Jones(2007)[26], Eric French, J. McCauley, M. Aragon, *et al.*(2017)[20]). In any cases, the direct consequence of the rapid growth of health investment is a shake to public health insurance program: in UK, the National Health Service(NHS) is reported to be in crisis with ever increasing budget, in the U.S., even with limited public health insurance, the financial pressure still make congressional leaders want to cut Medicare and Medicaid spending, in France, while being regarded as one of the best in the world, the system *l'assurance maladie* still face relentless financial

pressure.

Besides the rapid increasing health spending, another phenomenon is that countries with high health spending would perform poorly in terms of healthcare system efficiency and inequalities of health spending are still reported, so it's natural for us to ask why health spending becomes so expensive and why this 'efficiency paradox', in order to understand this phenomenon, we have to first understand individual behaviours concerning health, which are at the core of this phenomenon, that is, it is essential to first understand 2 questions that are lurking behind: first, how to define the value of health? Second, what influences health? These 2 questions link tightly to individual behaviours: the answer to the first question decides how an individual values his health in terms of utility, and the answer to the second question decides how an individual chooses to improve or to sacrifice his health in order to achieve the maximum of his utility. This dissertation gives our own perspective to these 2 questions, health is rather a subjective sentiment than some objective measurements, uniqueness resides in its very nature, personal characteristics (such as time preferences, risk tolerance, etc.), health status and financial constraints (such as assets and income) can alter one's decision when facing health shocks. If we are to understand the cause of the rapid growth of medical spending, it's better to start with the understanding of what makes a person want to invest in health and how would he do it, after figuring out the key factors in individual's decision making in health investment, it is therefore possible to aggregate these decisions made by people from different social-economic status and estimate the impact they would generate on the whole economy.

This dissertation is composed of 3 independent chapters, the first chapter lays the ground work for the 2 following chapters, we start with the situation we face nowadays, in the first part of this chapter, we use some empirical evidences to show the persistent trend of an increasing health spending and an increasing life expectancy during the last 3 decades among OECD countries, we also reveal that there are actually 2 aspects of this

increase: overall, for all countries, life expectancy has been largely improved and health spending has been growing quickly, but when we look at the details, health inequalities between different social-economic classes exist as always and seem to have been widening, as a result, we see that, gap in terms of life expectancy has been enlarged between the high-income class and the low-income class within countries, in terms of health spending, its growth benefits more the rich while aggravating the burdens on the poor. Afterwards, we begin to answer the 2 questions discussed above by reviewing literatures, the review of literatures is undertaken by 2 approach: first the deterministic approach then the stochastic approach, all these 2 approaches are based on the Grossman health capital model, the most important idea of this model is the notion of ‘health capital’, the principal ideas of this notion are: 1) Human health will deteriorate with time but can be ‘repaired’; 2) Unlike other corporate or incorporate assets, health can not be traded, thus the only option to insure health is via the channel of health spending, either before getting sick or after, Grossman chooses the former, we define ‘health insurance’ in a similar way in the latter chapters; 3) Health affects household’s utility by directly presenting in the utility function; 4) The household has to choose the allocation between medical spending, consumptions of other commodities and time devoted in either of these 2 actions in order to maximize his lifetime utilities; 5) As the household ages, the depreciation rate of health capital will increase so that at some point health capital would eventually fall below a certain level and the household dies. In the deterministic approach, we first present a simplified version of the Grossman health capital model, we justify its choices of simplifications and show the predictions it obtains, but do note that the discussion presented in this part are mostly theoretical, at the end of this part, we also discuss the flaws and potential problems that the Grossman model leaves because of its deterministic nature, in order to bring in the next part, in which we focus on the stochastic approach. In the stochastic approach, we first discuss the criticisms towards the deterministic nature of the Grossman model, then we use a simple example to show that a Grossman model

can be turned into a stochastic one and how the uncertainty would alter the individual's decisions.

In the following 2 chapters we discuss our perspectives to the 2 questions proposed at the beginning, in order to understand how an individual would value health and the interactions between his behaviours and his health, we first start with the deterministic perspective, in chapter 3 we discuss the possibility of building a Grossman model with endogenous lifespan, in this chapter, we study a life-cycle model of health stock and health spending. We use a simplified version of the Grossman model, where the trade-off occurs only between the allocation of health investment and consumption of other commodities. By introducing a non-negative health investment constraint, an explicit survivability threshold and a strong enough age-dependant health depreciation function, the model is able to derive optimality conditions with endogenous lifetime, the model is also numerically solved and we simulate the path followed by main variables over the agent's life-cycle. Calibrated on the US economy with realistic values of parameters, the model is able to replicate a growing profile of health spending and a significant share of health spending at the end of life roughly in accordance with the data. However, due to the low elasticity of income with respect to health stock and the lack of a financial market, we are not able to replicate a hump shaped income profile and a smoothed consumption profile.

Having confirmed that a model inheriting the spirit of the Grossman model is still able to replicate individual's health investment profile over his lifetime, we then move on to the stochastic approach, in chapter 4 we introduce the Markov chain to simulate a more realistic health shock and discuss the aggregate impact of individual's health decisions but also the impacts of different policies on individual's health decisions, using France and US as examples. A general equilibrium model *à la* Aiyagari (1994)[\[3\]](#) augmented by investments in health *à la* Grossman (1972)[\[23\]](#) is used to analyze the market allocations

of these 2 particular economies. The US is characterized by low unemployment risk, associated to low unemployment insurance, French is characterized by cheaper health goods and services and higher benefits of being good health. The turnover on the labour market in the US is therefore larger than in France. We show that expenditures in health are strongly related to the labour market turnover leading the American to perceived their employment spells as a more risky events, therefore reducing their incentives to invest in health, even when the turnover is principally associated to shorter periods of unemployment, in other words, the stability of jobs. This contributes to explain the poorer health of the Americans than French. In addition, counterfactual experiments show that expensive health related goods and services and risky behaviours are the other 2 main factors that explain the poorer performance of US healthcare system, indeed, when we swap the US-style health market by its French counterpart, we find out that there is a clear improvement in both the health and the welfare of the Americans.

Chapter 2

Stylized facts and review of literatures

2.1 Introduction

Without doubt, the average level of human health has been improved significantly: during the last 3 decades, overall life expectancy has been improved in most countries, in most developed countries like France, life expectancy at birth increases to 82.4 in 2015 from 75.7 in 1986, in less developed countries like China, life expectancy increases to 76 in 2015 from 68.4 in 1986(OECD,2018). But the word ‘average’ never seems to be accurate when we look at the details, indeed, average level of human health has been improved, but it doesn’t necessarily mean that health of every individual has been improved, in this chapter, by empirical data, we show that though the globally improved human health, health related inequalities are still reported and seem to have been widening. The word ‘inequality’ could be applied to many aspects, for example, inequality of mortality, inequality of life expectancy, or inequality of health spending, the last one seem to be mostly intriguing because of the rapid growth of health spending, it seems like that the improvement

of human health benefits more of those who can afford it. In order to understand this phenomenon, we suggest that first we need to understand individual behaviours when it concerns his own health, individual behaviours are at the core of health spending, if we can understand how individual of different social-economic status (SES) behave when facing health risks, we could then proceed to study the impacts and the consequences of this behaviours. In this chapter, we present 2 approaches: the deterministic approach and the stochastic approach, the first approach would provide some general ideas and theoretical basis of how an individual would react to a health fluctuation, the second approach would show us how to simulate the uncertain nature of health and sickness in real life when we want to apply our model to a more practical problem. This is the plan for this chapter, and also the plan we will follow in the whole dissertation, as the aim of this study is to understand individual behaviours and their aggregate impacts.

In this chapter we begin by reviewing existing literatures, the organisation of this chapter is as follows: in section 2 we show empirical evidences of health related inequalities, in section 3, section 4 and section 5 we review literatures, which is done through 2 approaches: in the deterministic approach, we choose to focus on reviewing the Grossman health capital model, in the stochastic approach, we choose to concentrate on showing how to turn the Grossman model into a stochastic model, finally, in section 6, we conclude.

2.2 The 2 aspects of health inequalities

Health is playing more and more an important role in nowadays economy life, specifically, during the last 3 decades, with the development of iatrology and medical technologies, two issues emerge in the domain of health economics: the first issue is that, despite life expectancy in general has increased across the globe, life expectancy inequalities within countries are still reported between people of different social-economic status (Blaise and

Lefebvre (2018)^[7]). At 30 years old, the life expectancy gap between low and high educated male in Europe is equal to 8 years (with a max in Estonia equals to 13.6 years, and a min in Sweden equals to 4.3 years). For the females, these gaps are lower but still significant: in average, gap in life expectancy between low and high educated female weights to 4.1 years (with a max in Estonia equals to 8.6 years and a min in Portugal equals to 2.3 years).

	Men		Women	
	Primary education level	Bachelor degree or higher	Primary education level	Bachelor degree or higher
Bulgaria	36.3	47.2	45.4	51.9
Czech Republic	35.7	48.3	49.8	52.4
Denmark	45.9	51.6	50.8	54.3
Estonia	34.9	48.5	45.9	54.5
Finland	46.1	51.5	52.4	55.8
France	46.8	53.1	54.1	57.1
Hungary	36.4	48.1	46.6	51.9
Italy	49	53.5	54.4	57.2
Norway	47.7	52.4	52.1	55.7
Poland	37.9	44.2	49.1	54.2
Portugal	47.5	52.6	54.1	56.4
Romania	38.1	46.3	48.1	51.2
Sweden	48.5	52.8	52	55.5
Average	42.4	50.4	50.4	54.5

Table 2.1: Europeans life expectancies at 30 years old between people of different education levels (Blaise and Lefebvre (2018)^[7])

In the U.S. it's the same thing, health gaps between different social-economic gradients still exists and seem to have been widened, people with high education level in general live longer than those with low education level: between 1990 and 2010, disparities in life expectancy caused by education levels increase across all gender and race groups (Sasson (2016)^[43]), the difference in life expectancy between the least educated white women and the most educated white women increases from 2.5 to 9.3 years, the difference between the most educated black women and the least educated black women also increases but a little bit smaller from 1.9 to 4.7 years, the life expectancy gap is even larger among men:

from 6.1 to 11.9 years for white men and from 6.9 to 8.6 years for black men between 1990 and 2010. An decrease of life expectancy is also observed among the least educated white people between 1990 to 2010 while all other groups have at least an increase in life expectancy during the same period. As generally we consider that a higher education level leads to a higher ability to get high revenue and thus a better social-economic status, we can also say that this result suggests a strong and positive relationship between social-economic status and life expectancy.

Education(yrs.)	Non-Hispanic White				Non-Hispanic Black			
	Women		Men		Women		Men	
	1990	2010	1990	2010	1990	2010	1990	2010
0-11	54	50.9	46	45.4	49.9	51.8	39.6	45.5
12	55.1	55.9	48.7	50.5	49.2	52.7	41.4	46.5
13-15	55.2	56.7	49.7	52.8	49.6	54.2	43.1	50.9
16+	56.5	60.2	52.1	57.3	51.8	56.5	46.5	54.1

Table 2.2: U.S. life expectancy at age 25 between 1990 and 2010(Sasson (2016)[\[43\]](#))

Income, as another indicator of social-economic status, is also proved to be positively related to longevity, in Europe, differences in life expectancy have been reported between people who have the highest revenue and people who have the lowest revenue: in France, for men in age cohort 30 to 34, the life expectancy gap between people who have the highest revenue and people who have the lowest revenue is roughly 20 years, in Sweden, for the same group of people, this gap is remarkably reduced but still exists at around 10 years(Blaise and Lefebvre (2018)[\[7\]](#)). In the U.S., gap in life expectancy caused by income has been widened since 2001, though life expectancy increases across all income quartiles, life expectancy grows 2.5 times faster for the wealthiest people than for the poorest people(Chetty, Stepner, Abraham *et al.* (2016)[\[11\]](#)). Consequently, the increase of life expectancy seems to benefit more people with high education level and high revenues. Last but not least, another interesting finding is that the marginal increase of life expectancy by revenue decreases with revenue, that is, a concave relationship between revenue and longevity(Blaise and Lefebvre (2018)[\[7\]](#)).

Annual income quartile	2001	2014	Annual increase(yrs.)
1st(\$17000)	76	77.2	0.08
2nd(\$47000)	80	81.8	0.12
3rd(\$87000)	82.2	84.9	0.18
4th(\$256000)	85	88	0.2

Table 2.3: Life expectancy at birth by income, U.S., 2001-2014(Chetty, Stepner, Abraham *et al.* (2016)[\[11\]](#))

Overall, during the last 3 decades, though this progress seems to stop after 2014¹, life expectancy increases substantially with in average an increment of 8 years(OECD,2018), but it clearly benefits more people in high social-economic gradient, as we see that health gap between high social-economic gradient and low social-economic gradient remains always large and actually tends to increase.

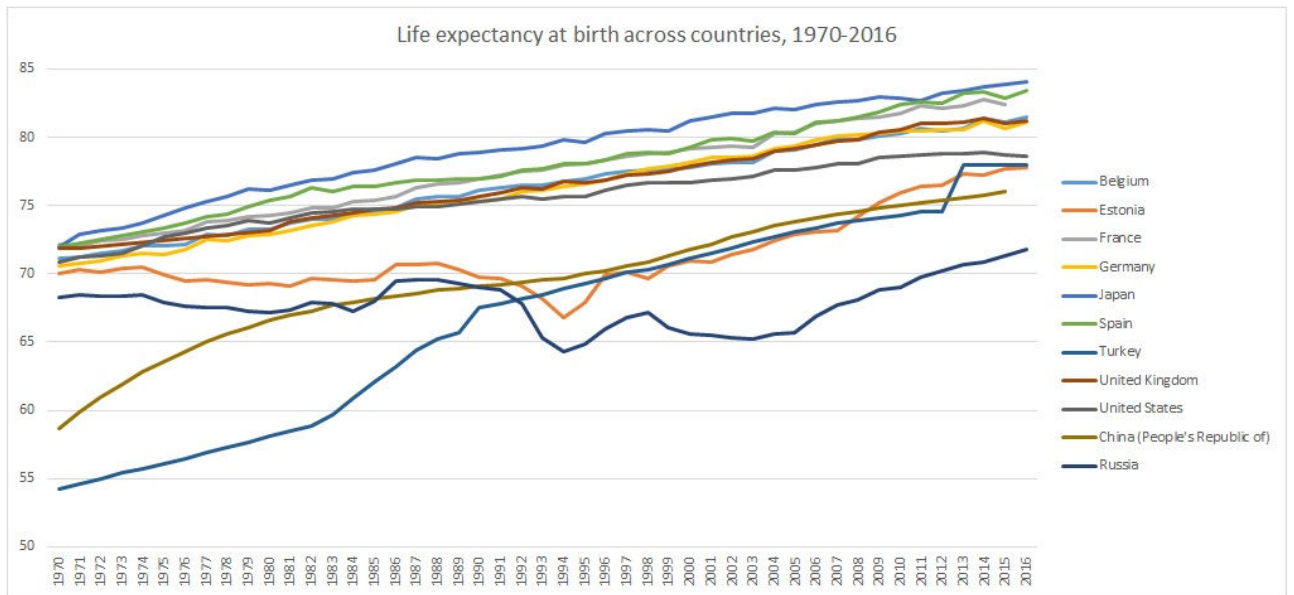


Figure 2.1: Life expectancy across countries,1970-2016(OECD,2018)

¹OECD data shows that the average life expectancy of all countries in its database reaches its peak in 2014 at 78.9 years, then in 2015 it drops to 78.8 years, furthermore, JY HO and AS Hendi (2018)[\[28\]](#) also reports a decline of life expectancy in most high income countries after 2014.

The second issue is the rapidly rising health expenditures. In 1990, the average health expenditure share of GDP across all OECD countries is about 6.5%, this number rises to 8.81% in 2017. Rapid growth of health spending among OECD countries have been reported, in the U.S. the health-GDP ratio rises most quickly from 8.2% in 1980 to 11.3% in 1990 and to 17.1% in 2013, in Japan this share rises relatively slow, from 6.2% in 1980 to 10.8% in 2013, in France, the growth of health expenditure is relatively not so quick(compared to the U.S.) but the number itself is still too large to ignore, with 6.7% in 1990 to 11.4% in 2013(OECD,2018). Squires and Anderson (2015)[46] compare the health expenditures of 13 high-income countries and find out that health spending generally grows quickly during the last 3 decades and by 2013, the health-GDP ratio in all 13 studied countries reach around at least 10%. Table 2.4 shows the updated data by OECD 2018 database and Squires and Anderson(2015)'s data of the 13 high-income countries, we see the differences are small and indeed by 2013, among the 13 countries, health spending represents at least roughly 10% of their GDP.

Country	Percentage of health cares on GDP				Squires and Anderson's Data
	1980	1990	2000	2013	2013
Australia	5.8	6.5	7.6	8.8	9.4
Canada	6.6	8.4	8.1	10.1	10.7
Denmark	8.4	8	8.1	10.2	11.1
France	6.7	8	9.5	11.4	11.6
Germany	8.1	8	9.8	10.9	11.2
Japan	6.2	5.8	7.2	10.8	10.2
Netherlands	6.6	6.9	7.1	10.9	11.1
New Zealand	5.8	6.7	7.5	9.4	11
Norway	5.4	7.1	7.7	8.9	9.4
Sweden	7.8	7.3	7.4	11.1	11.5
Switzerland	6.6	7.9	9.8	11.3	11.1
United Kingdom	5.1	5.1	6	9.8	8.8
United States	8.2	11.3	12.5	16.3	17.1

Table 2.4: Total Health Care Expenditures as a Percentage of Gross Domestic Product between 1980-2013, by OECD country(OECD,2018, Squires and Anderson (2015)[46])

But higher health spending doesn't necessarily bring better health system, Squires

and Anderson (2015)[46] uses the U.S. as an example to illustrate that though the U.S. spends more than twice the OECD median per capita, its health system still performs badly in terms of medical outcomes such as life expectancy or mortality rate as a result lethal diseases such as cancer and heart diseases. By Squires and Anderson (2015)'s data[46], if we compare the U.S. with France, which has a far better performance health system, we can see that despite of spending doubled health expenditures per capita(\$9086 vs \$4361) and 4 times more out-of-pocket health expenditures(\$1074 vs \$277), the Americans live in average 4 years less than the French people(78.8 vs 82.3) and have a similar mortality rate as a result of cancer but 3 times higher when the mortality rate is related to ischemic heart disease.

Country	Per Capita Health expenditures(US\$)	Per Capita Out-of-pocket Health expenditures(US\$)
Australia	4115	771
Canada	4569	623
Denmark	4847	625
France	4361	277
Germany	4920	649
Japan	3713	503
Netherlands	5131	270
New Zealand	3855	420
Norway	6170	855
Sweden	5153	726
Switzerland	6325	1630
United Kingdom	3364	321
United States	9086	1074
OECD median	3661	625

Table 2.5: Per Capita Health Care Expenditures, by OECD country:2013(Squires and Anderson (2015)[46])

Thus efficiency of health system needs to be checked to see if it's indeed in some high health spending countries a relative important part of health expenditures are spent for nothing. In this thesis we choose to focus on the U.S. and France, since the 2 countries have similar average life expectancies but health spending in the U.S. far exceeds that in France. To examine the efficiency, Data Envelopment Analysis(DEA) is frequently applied(Afonso

and Aubyn (2004)[\[2\]](#), Afonso and Kazemi (2016)[\[11\]](#), Cetin and Bahce (2016)[\[8\]](#), Önen and Sayin (2018)[\[39\]](#)), the DEA methodology is a linear programming method for evaluating the relative efficiencies of Decision Making Units(DMUs) who produce output(s) using similar inputs, the method is originally designed to evaluate the efficiency of an American federal program “Follow through”(Charnes, Cooper and Rhodes (1978)[\[9\]](#)), then it’s generalised to be applied in various sectors including health economics. In the domain of health, the mostly used outputs would be Life Expectancy(LE) and Infant Mortality Rate(IMR), as for the mostly used inputs, there are 2 types: either physical inputs such as number of beds, number of physicians and nurses, or monetary inputs such as per capita health expenditures and health spending-GDP ratio.

The DEA methodology has been widely used in the domain of health², Afonso and Kazemi (2017)[\[11\]](#) calculate DEA efficiency scores of healthcare system of 20 OECD countries between 2009 and 2013, using both CCR and BCC model, they use normalized health expenditures as the only input and LE and IMR as outputs. Here we only show the results that we are concerned with, complete results are presented in the Appendix [C](#). According to their results, under BCC, France and the U.S. have a similar input-oriented efficiency score, with the U.S. a little bit higher at 0.77 and France a little bit lower at 0.741, suggesting that in terms of input efficiency, both countries are considered as inefficient and they should have been able to reduce around 25% of their inputs to produce their outputs compare to the input efficiency of their own *pairs*, under CCR, France gets an even lower score at 0.694 while the input-oriented efficiency score for the U.S. doesn’t change much, which equals to 0.76, suggesting that the scale efficiency is larger for France than for the U.S.. Cetin and Bahce (2016)[\[8\]](#) obtain a similar conclusion, they use the classic CCR model and include not only financial factors such as per capita health expenditures but also physical factors such as number of doctors and nurses as inputs, according to their

²Detailed presentation of the DEA method can be seen in Appendix [A](#) and in Appendix [B](#).

results, France obtains an input-oriented efficiency score of 0.739, which is lower than that of the U.S. which is of 0.782, plus, these authors also report that France is of nature DRS and the U.S. is of nature IRS, suggesting that France has already passed their optimal input size and the U.S. hasn't yet reached their optimal input size. These 2 literatures both seem to suggest that France is worse than the U.S. in terms of input-oriented efficiency, which is a little contradictory to the data we presented before, since the Americans do spend much more on health than French people but have a worse or at most similar overall health status than French people. This contradiction could be caused by multiple reasons: 1) First of all, the efficiency score obtained by the DEA methodology may not be able to be directly used in comparison, since for each inefficient unit, their *pairs* could be different, in the article of Afonso and Kazemi (2017)[[10](#)], for France, its *pairs* are Switzerland and Luxembourg under BCC and Ireland under CCR, for the U.S., its *pairs* is Ireland both under BCC and under CCR, thus at least we would say that under BCC the input-oriented efficiency scores of these 2 countries may not be able to be directly compared. This is also the case in the article of Cetin and Bahce (2016)[[8](#)], in this article, France's *pairs* are Chile, Korea and Slovenia, for the U.S., its *pairs* are Chile, Iceland and Ireland; 2) The second reason could be that by definition, the DEA methodology measures only the efficiency, so a country with relatively poor health outcomes but using also relatively low inputs could also be considered as efficient, for example, according to Cetin and Bahce (2016)'s[[8](#)] study, Chile is considered as one of the efficient countries, but according to their database(OECD,2012), life expectancy of Chile is lower than that of France(78.7 vs 82.1) and infant mortality rate(per 1000 live birth) of Chile is more than double than that of France(7.4 vs 3.5), yet Chile is considered as efficient because of its relatively low input(7% of GDP) according to the DEA methodology but France is not. But between France and the U.S. this doesn't seem to be the case, so this contradiction we discussed here might be caused by another reason; 3) The third one could be that, in fact, since by definition the DEA methodology is designed to maximise the ratio between

the weighted output and the weighted input for a particular unit, so the DEA methodology does have the possibility to overemphasize an unit's strengths while ignoring its weakness by giving its weakness a low or even 0 weight, thus additional restrictions on weights might be proposed, but this is not the case in both Afonso and Kazemi (2017)'s [10] work and in Cetin and Bahce (2016)'s [8] work. Since in the article of Afonso and Kazemi (2017) [10], in terms of health, there is only 1 input thus it might not be a problem, but in the article of Cetin and Bahce (2016) [8] there are both financial input and physical input, hence it could be possible when evaluating the U.S., weights of physical factors might be overemphasized and weights of financial factors might be underemphasized; 4) The last reason could be a biased using of data, in fact, Afonso and Kazemi (2017) [10] aim to evaluate government's performance in various sectors, so the health expenditure-GDP ratio they use might not refer to the total health expenditures but only the public health expenditures, in this sense, France government do expenses more than the U.S. government since the U.S. still yet to have a full health care system. Therefore, the U.S. could be more efficient than France in terms of input-oriented efficiency because of this 'advantage' but also the second reason we mentioned above.

Nevertheless, in terms of output-oriented efficiency, France does perform better than the U.S.: using the BCC model, Afonso and Kazemi (2017) [10] also calculate the output-oriented efficiency score for OECD countries, according to their results, France has an output-oriented efficiency score of 0.992 and its *peers* is Japan, the U.S. has an output-oriented score of 0.972 and its *peers* is also Japan, suggesting that though both France and the U.S. are still considered as inefficient, France healthcare system performs better than that of the U.S. in terms of medical outcomes. The 2 scores do seem close but note that according to Afonso and Kazemi (2017) [10], 0.972 makes the U.S. rank at the very bottom among all 20 studied OECD countries while 0.992 gives France a better ranking at 9th, Önen and Sayin (2018) [39] reaches the same conclusion as they calculate

the efficiency scores of OCED countries with their output-oriented DEA model, they only take physical factors such as number of physicians, number of nurses and number hospital beds as inputs and take LE and IMR as outputs, they also add additional restrictions on weights to alleviate the overemphasising and underemphasising problem we mentioned above: concerning the weights of outputs, for an output term y_{sk} , they introduce both an upper and a lower bound of the proportion of this particular weighted output $u_s y_{sk}$ to the overall weighted output of unit k

$$L_s \leq \frac{u_s y_{sk}}{\sum_{r=1}^S u_r y_{rk}} \leq U_s$$

in terms of inputs, they add similar restrictions on input weights v_i . They call this new form of DEA model(BCC+restrictions on weights) as ARG model, they then use both BCC and ARG model to evaluate the output-oriented efficiency score of 34 OECD countries both in 2008 and in 2012. In general output-oriented efficiency scores are lower in the ARG model than in the BCC model, which is normal since the ARG model introduces more restrictions on weights. Here we present only the latest results of their work, full results could be seen in Appendix [Q](#). By their calculation, in 2012, France has an efficiency score of 0.991 under BCC and 0.658 under ARG, which are both higher than the U.S. who only gets 0.966 as an efficiency score under BCC and 0.565 under ARG. So these results by Önen and Sayin (2018)[\[39\]](#) also suggest that France outperforms the U.S. in terms of medical outcome efficiency but do note that in their work the reference set for France and for the U.S. are different, for example, under ARG, in 2012, the reference set of France is composed of Slovenia and Iceland and the reference set of the U.S. is composed of Chile, Iceland and Slovenia.

Overall, though the DEA methodology could be questioned and the scores obtained by this method might not be able to be directly used in comparison, we are at least safe to say that health care system in France and in the U.S. are rather inefficient, both input-

oriented and output-oriented. France and the U.S. enlist certainly among the countries with highest health expenditures, but as we have discussed above, they certainly not enlist among countries with most efficient healthcare systems. This ‘efficiency paradox’ could lead to inequalities of health spending since the health related goods and services would be so expensive that only the rich could afford, indeed, inequalities of health spending have been reported in the U.S.: before ACA(Affordable Care Act)’s passage in 2010, 39% of Americans with below-average income admit that the sole reason that they don’t go see a doctor is the cost, yet even after ACA’s expansion of health insurance coverage, 27 million Americans still remain uninsured and most of them have annual incomes near or below the official poverty line(Dickman, Himmelstein and Woolhandler (2017)[16]). In addition, in the private insurance market, the premium has increased substantially, in 2016, the average premium for family coverage rise to \$18142, which has increased 58% since 2006(Claxton, Rae, Long, Damico and Sawyer (2016)[12]), more than 80% of employer-based insurance plans include an certain amount of money that a insured must pay before his insurer covers the rest, in 2016, this amount of money in average weights \$1478, 2.5 times higher than in 2006. The rising costs of having medical treatments leads predictably a decline use of health services by the poor, and doctors and hospitals would eventually fill their empty appointment slots and beds with people who can afford the costs(Dickman, Himmelstein and Woolhandler (2017)[16]). The situation is much better in France with a universal health insurance that aims at allowing every individual living in France to get access to the minimum health care, indeed, by French healthcare system, the poor benefits more than the rich: 90% of the costs of medical treatments are covered for the poorest, compared to less than 70% coverage for the rich, out-of-pocket health spending ranges from 4% for the poor to 13% for the rich(Nay *et al.* (2016)[38]). Yet inequalities still exists: based on the survey of ESPS, 15.8% of interviewees declare that they have to give up looking for medical treatments because of financial reasons(Dourgnon *et al.* (2012)[17]) and the wealthiest 50% people are reported to see specialists 2 times

more frequently than do the poorest 25%(Or, Jusot and Yilmaz (2012)[33]). Nevertheless, even excluding the U.S., inequalities on access to care seems to be a universal phenomenon, Devaux and Looper (2012)[45] report that in 19 major OECD countries, there exists a clear-cut on need-adjusted doctor visit frequency between high-income people and low-income people, this gap is even larger when consider need-adjusted specialist visit such as dentist visit or cancer screening, France ranks at the median according to this report and inequalities in healthcare utilisation has been clearly justified, for example, though France ranks at the top in overall need-adjusted doctor visit frequency, there is still a 10% gap between the utilisation frequency of high-income people and that of low-income people, when consider need-adjusted specialist visit, this gap is widened to 24%.

In general, we see a rapid increase of health expenditures across the globe, but this trend doesn't seem to make countries with high health expenditures necessarily efficient, furthermore, inequalities of health expenditures are still reported. With longer lifespan and widened health gap, we see a need of economists in the domain of health: it's thus necessary, and it's up to economists to model the economy of health, to understand the determinants when it comes to decisions concerning health and to figure out the differences between health policies in different countries and their impacts.

2 fundamental questions lie in the center of answering these needs: first of all, how to define the value of health? Secondly, what influences health? In the rest of this chapter, we try to respond to these questions by works that have been previously done: there are essentially 2 approaches of modelling individual behaviours when their health fluctuates, either by the deterministic approach or the stochastic approach, we will review them separately. We also dig more deeply to show the flaws of the original Grossman model: its potential permission of infinite lifespan and its deterministic nature that doesn't suit properly the uncertain nature of the health risk, these are also the reasons why we first start with the deterministic approach and then turn to the stochastic approach.

2.3 A simplified interpretation of the Grossman model

Up to today, The Grossman model [23] still holds an iconic position in terms of health economy, but it is such a complex model that even Grossman himself needs to divide the model into 2 branches³ in order to explain the individual's optimal stock of health capital and individual's optimal health investment decisions. In fact, in Grossman (1972)[23], in order to simplify the theoretical analysis and to contrast human health with other forms of human capital, Grossman uses only the pure investment model, however it's important to include both 2 benefits of health(consumption benefits and investment benefits) in the analysis of individual behaviours on health investment since leaving out either 1 of these 2 benefits will ultimately alter individual's decisions. In this section, in order to illustrate more clearly the trajectory of individual's optimal health stock and correspondingly the trajectory of individual's health decisions overtime, we would use a simplified version of the Grossman model by Laporte (2015)[34] which includes both benefit components as the full Grossman model.

Laporte (2015)[34] takes the continuous time optimal control approach to analyse the Grossman model, and phase diagrams are drawn to present the optimal trajectory of individual's stock of health and health investment. The model could be presented as below:

Utility function The utility function in Laporte (2015)[34] could be written as

$$\max \int_0^T U(C_t, H_t)e^{-\rho t} dt \quad (2.1)$$

³Grossman divides his model into 2 different branches: one is the pure investment model and the other is the pure consumption model, in the former branch Grossman assumes that only the monetary aspect of health matters, in the latter branch Grossman assumes that only the psychic aspect of health takes effect.

where C denotes the non health-related consumption and H denotes the stock of health capital, ρ denotes the subjective discount factor. In addition, Laporte assumes that $U_C > 0$, $U_{CC} < 0$, $U_H > 0$, $U_{HH} < 0$, so individual's utility function is concave on C and H , finally, Laporte assumes that $U_{CH} > 0$.

T is assumed to be the finite length of individual's lifespan, which is quite an unrealistic assumption since it suggests that individual would have full control on their lifespan, we shall return to this point later on.

Motion of health capital Health evolves essentially in the same way than in the original Grossman model, by the following equation

$$\dot{H} = G(I) - \delta H \quad (2.2)$$

where $G(I)$ is the production function of health capital, $G(I)$ is supposed to be instantaneous, which follows the instantaneous health production hypothesis in the original Grossman model, plus, Laporte also assumes that $G_I > 0$, $G_{II} \leq 0$ and $G(0) = 0$, thus the production function of health investment is supposed to have a positive marginal productivity. \dot{H} represents therefore the net investment in health, with $G(I)$ the gross investment and δ the health depreciation rate, as we can see, health depreciation rate is assumed to be constant, which actually leaves window for immortality, given that the technology of health investment production is sufficiently efficient, we shall discuss this point later.

The budget constraint Laporte assumes an instantaneous budget constraint instead of a lifetime budget constraint from the original Grossman model. His budget constraint could be written as

$$Y(0) + Y(H) = C_t + p_I I_t \quad (2.3)$$

where health investment I is defined as a commodity that can be directly purchased from the market at price p_I , the price of C is set to be 1. There is no production function of health investments as in the original Grossman model, plus, I is also supposed to be non-negative ($I \geq 0$). $Y(0)$ denotes the part of income that is independent of the individual's health and $Y(H)$ denotes that part of income that depends on the individual's health, in addition, Laporte assumes a diminishing effect of health on $Y(H)$, that is $Y_H > 0$ and $Y_{HH} < 0$.

There is no time constraint in Laporte's version, the reason of this simplification is that in this article, Laporte chooses to focus more on the inter-temporal issue, instead of picking up choices among different health investment strategies, in addition, health capital conceived in this article is defined in a more general sense as "... Health capital is best thought of as relating to the individual's ability to resist disease, and to perform what the health care literature refers to as activities of daily life... ", to this extent, health capital has no need to be tied with healthy days within a certain period of time, the individual receives higher utilities because of being able to live in good health rather than having more days being healthy, therefore in this version of the Grossman model, time is excluded as an input variable and there is also no need for the time constraint as in the original Grossman model.

The individual's problem The individual's problem could be written as

$$\max \int_0^T U(C_t, H_t) e^{-\rho t} dt \quad (2.4)$$

subject to

$$\begin{cases} Y(0) + Y(H) = C_t + p_I I_t \\ \dot{H} = G(I) - \delta H \\ I \geq 0 \end{cases}$$

The Lagrangian could be hence written as

$$\mathcal{L} = U(Y(0) + Y(H) - p_I I, H) + \phi[G(I) - \delta H] + \lambda I \quad (2.5)$$

where Eq 2.5 substitute C using the budget constraint. In this formulation, ϕ denotes the shadow price of health capital, thus the marginal benefit of 1 unit of H , λ is the Kuhn-Tucker expression, when $I > 0$, $\lambda = 0$, when $I = 0$, $\lambda > 0$.

The problem could be easily solve, Laporte writes the first of the Pontryagin necessary conditions as

$$\frac{\partial \mathcal{L}}{\partial I} = -p_I U_C + \phi G_I(I) \quad (2.6)$$

which gives out the interior equilibrium solution of I

$$p_I U_C = \phi G_I(I) \quad (2.7)$$

The LHS of Eq 2.7 shows the marginal benefit of 1 unit of I , for it's the product of the marginal product of health investment, G_I , times the shadow price of health capital, ϕ ; the RHS shows the marginal cost of I , for it's the product of reduction of consumption of 1 unit plus of I , p_I , times the marginal utility of consumption, U_C .

Now differentiate Eq 2.5 with respect to H , we can write

$$\frac{\partial \mathcal{L}}{\partial H} = U_H + U_C Y_H - \phi \delta = 0 \quad (2.8)$$

combining Eq 2.7 and Eq 2.8 we can obtain the equilibrium condition

$$U_H + U_C Y_H = p_I U_C \times \frac{\delta}{G_I} \quad (2.9)$$

⁴Since the price of I is p_I , 1 unit of I would therefore cost p_I , as the price of C is set to be 1, thus 1 unit increase in I would be matched by p_I units of C , in other words, the opportunity cost of I .

Eq 2.9 actually reveals a major problem of Laporte's model, which is that it still permits individuals to live forever, as we can see, the LHS shows the marginal benefits of every unit of health produced by health investment, yet for every unit of increase in H , δ units will be depreciated over time, and thus requires $\frac{\delta}{G_I}$ units of I invested to repair the loss, that's to say, at the optimum, Eq 2.9 suggests that the loss of health should always be fully repaired by health investment, and given that δ is hold constant, this might not be so implausible, in this sense, Laporte's model actually suggests that at the optimum the individual would simply live forever.

After justifying the simplifications he make in his version of the Grossman model, Laporte then sets out to show several predictions that could be obtained using his model, in other words, the predictions that should be obtained by the Grossman model.

Stationary loci of I and H The one-state variable form of his model allows Laporte to use the phase diagram to present the stationary loci of I and H and eventually the optimal trajectories of health capital and of health investments.

Thus the first step is to draw the stationary loci of I and H . By definition, the stationary locus of I is defined by $\dot{I} = 0$ and similarly the stationary locus of H is defined by $\dot{H} = 0$.

To be able to draw the stationary loci of I and H , Laporte further assumes that $G(I) = I$, which also suggests a CRS technique on health production, as in the original Grossman model. Therefore, Eq 2.2 becomes

$$\dot{H} = I - \delta H \tag{2.10}$$

thus the stationary locus of H becomes a straight line

$$I = \delta H \quad (2.11)$$

the phase arrows of H can be therefore found by $\frac{\partial \dot{H}}{\partial I} = 1 > 0$ and $\frac{\partial \dot{H}}{\partial H} = -\delta < 0$.

Concerning the stationary locus of I , first write the Hamiltonian as

$$\mathcal{H} = U(Y(0) + Y(H) - p_I I, H) + \phi(I - \delta H) \quad (2.12)$$

The Pontryagin necessary condition with respect to I would be

$$\frac{\partial \mathcal{H}}{\partial I} = -p_I U_C + \phi = 0 \quad (2.13)$$

The second of Pontryagin's necessary conditions is

$$\dot{\phi} = \rho \phi - \mathcal{H}_H \quad (2.14)$$

$$= \rho \phi - [U_C Y_H + U_H - \delta \phi] \quad (2.15)$$

Total differentiating Eq [2.13](#) with respect to time and substituting Eq [2.10](#) and Eq [2.14](#), we would get an equation of \dot{I}

$$\dot{I} = \frac{p_I [U_{CH} + U_{CC} Y_H] [I - \delta H] + [U_C Y_H + U_H] - [\rho + \delta] p_I U_C}{p_I^2 U_{CC}} \quad (2.16)$$

at the intersection point of stationary locus of I and H , i.e. $\dot{I} = \dot{H} = 0$, the above equation becomes

$$\dot{I} = \frac{[U_C Y_H + U_H] - [\rho + \delta] p_I U_C}{p_I^2 U_{CC}} \quad (2.17)$$

The stationary locus of I is defined by $\dot{I} = 0$, thus the slope of the stationary locus of I could be written as

$$\frac{\partial I}{\partial H} = \frac{[\rho + 2\delta][U_{CH} + U_{CC}Y_H] - [U_{CC}Y_H^2 + 2U_{CH}Y_H + U_C Y_{HH} + U_{HH}]}{[\rho + \delta]p_I^2 U_{CC}} \quad (2.18)$$

assuming that $U_{CH} + U_{CC}Y_H > 0$ and also assume the second term in square bracket is negative⁵, thus the slope of the stationary locus of I is negative, the phase arrows of I can be found by $\frac{\partial \dot{I}}{\partial I} > 0$ and $\frac{\partial \dot{I}}{\partial H} > 0$, hence the phase diagram of the problem could be drawn as Figure 2.2.

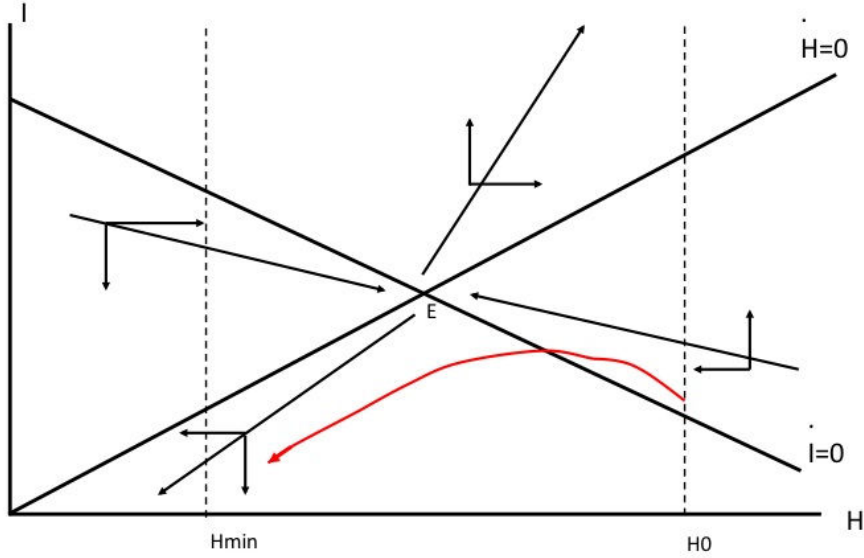


Figure 2.2: The phase diagram

Figure 2.2 shows the stationary locus of both I and H , $\dot{I} = 0$ and $\dot{H} = 0$, the intersection of these 2 locus(point E) is therefore referred as the equilibrium point. Every

⁵It means that the overall marginal utility of health stock is diminishing when both of its effect on income and its direct effect on utility are accounted in the same time.

trajectory in this phase diagram is by definition an optimal trajectory since the phase diagram is drawn using first order conditions, we see as well that there are only 2 trajectories that would converge to the equilibrium, referred as the stable branches, and also only 2 trajectories that would diverge directly from the equilibrium, referred as the unstable branches, other trajectories will first move toward the equilibrium but then diverge from it.

Point E is referred as the equilibrium point, but mostly in terms of macroeconomics, since on the microeconomic level, it usually takes forever for an individual to converge to the equilibrium, which no one has ever achieved. Thus the individual's optimal trajectory would not include this equilibrium, as what we show in Figure 2.2 by the curved trajectory. Since on the microeconomic level point E can not be reached in an finite amount of time, an specific endpoint condition has to be chosen in order to obtain a solution, Laporte chooses a fixed endpoint H_{min} as in the original Grossman model, meaning that once the individual's health capital falls to H_{min} , he dies.

The curved trajectory in Figure 2.2 hence shows a typical case of the Grossman model: given the initial health H_0 and the fixed length of life T (recall that for now T is still fixed and known in advance), the individual should choose a trajectory of (I, H) that would bring him from H_0 to H_{min} during T time periods, since he starts with a rather high H_0 , thus initially he choose a low I , which would make his H fall, then he chooses to increase his I , slowing down the rate of depreciation of his H , afterwards, when his I increases to a certain level, which is shown in Figure 2.2 by the point where the curved trajectory cuts the stationary loci of I , he chooses to let his I fall again, and this would eventually make his H decreases to H_{min} at T .

Effects of income Using the phase diagram, Laporte illustrates the effects of social factors(income or human capital)on the individual's optimal health stock and health in-

vestment decisions. In order to do so, he first further assumes that $Y(H)$ is dropped from the constraint budget, therefore individuals with different level of incomes $Y(0)$ will still have the same equation of motion for health.

Concerning the stationary locus of H , it should not be affected by a higher $Y(0)$, for it simply suggests that $I = \delta H$.

However, concerning the stationary locus of I , Eq [2.17](#) becomes

$$\dot{I} = \frac{U_H - [\rho + \delta]p_I U_C}{p_I^2 U_{CC}} \quad (2.19)$$

for $\dot{I} = 0$ it needs $U_H - [\rho + \delta]p_I U_C = 0$, hence we can eventually obtain

$$\frac{\partial I}{\partial Y(0)} = \frac{U_{HC} - [\rho + \delta]p_I U_{CC}}{p_I U_{HC} - p_I^2 U_{CC}[\rho + \delta]} = \frac{1}{p_I} > 0 \quad (2.20)$$

Eq [2.20](#) shows that even the individual's income has no effect on potential health investment and thus future health, an increase in $Y(0)$ will still pushes up I , in economic terms this can be explained by the fact that an increase of $Y(0)$ will increase also C for any given point of I , thus the opportunity cost of increasing I decreases for the marginal utility of consumption U_C becomes smaller.

Figure [2.3](#) shows the effect of an increase in $Y(0)$, the new stationary locus of I and the new optimal trajectory are represented by the dashed line. We can see clearly that higher $Y(0)$ shifts up the stationary locus of I , thus pushes up the equilibrium point but also make an individual with the same initial health capital invest more in health.

The optimal length of life Now let's discuss the optimal length of life. In Laporte's version, as in the original Grossman model, the individual's lifespan T is assumed to be known in advance and the health depreciation rate δ is considered as constant, given a CRS

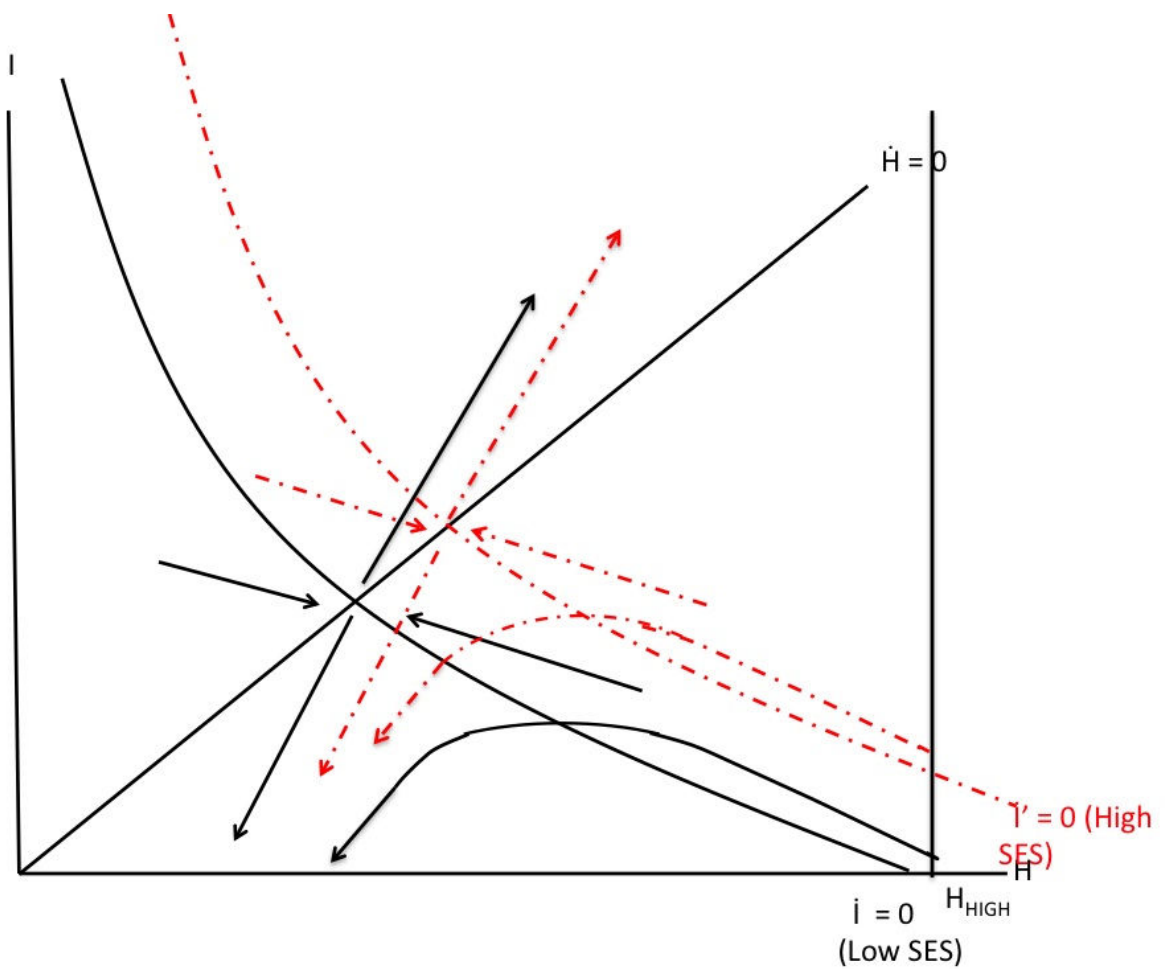


Figure 2.3: The effect of an higher income

health production function, these 2 assumptions would eventually make an individual live forever, given an enough sufficient health production technology, even if it wasn't this case, having full control on one's life is already beyond realism, Grossman [24] addresses to this issue, the solution he provides is to turn δ variable so that health capital will depreciate more and more severely as the individual ages, this seems to be a reasonable way to imitate the collapsing process of human health, but how δ varies with age needs to be carefully chosen to make sure that the speed of collapsing would be neither too quick nor too slow, we follow this path in chapter 3 and show how we set the equation of motion of δ so that the virtual individuals in our model have a finite and endogenous lifespan. Another criticism towards this method is that, as Laporte argues, even δ increases with age, it

still leaves window for the possibility of living forever, assuming a really enough sufficient health production technology, if it was true, it basically suggests that the existence of a somehow magical health treatment technology that would be able to repair human body instantly, no matter how high the health depreciation rate would be at the end of life, which, as we can see, turns to be actually another beyond-realism assumption.

But this method proposed above only ease the worries toward the constant δ issue, the other one, as Zweifel (2012)[\[49\]](#) argues, knowing the length of life is still pretty much unrealistic, yet because of the deterministic nature of the Grossman model, this seems to be unavoidable. Laporte argues that this assumption would rather be not so unreasonable in terms of empirical investigation, but admits that this problem needs to be dealt with since it really doesn't make any sense on the theoretical level. Laporte then propose 2 options to fix this problem:

- Either we assume a maximum length of life but the actual lifespan is endogenous, the optimal life length is chosen under condition that the Hamiltonian at this end of life equals to 0. In chapter 3, we choose this option, without stepping too early into the stochastic domain, since that's what we do in the fourth chapter. ;
- Either we take the stochastic approach, by making the health evolution random, as what we do in the fourth chapter. In the second to last section of this chapter, we shall present a simple example of this approach, as a prelude or an introduction to the fourth chapter, in this section, we shall see how the introduction of health risk would alter the agent's decisions.

2.4 Justification of the lack of financial market in Laporte's model

In this section we shall discuss the justifications for the simplifications Laporte made in his version of the Grossman model, since this model also serves as a basic model in chapter 3, these justifications also would provide a firm standpoint to which extent we build our own version of the Grossman model.

The previous section presents the basic layout of Laporte's simplified version of the Grossman model, as we can see, there are 2 major differences if we compare this version and the original one: first of all, there is no time constraint, as we have already discussed, Laporte chooses this simplification because he focus more on the intertemporal issue and the health capital is defined differently from the original Grossman model, we inherent this idea in chapter 3 for these very same reasons, in our version of the Grossman model, there is also no constraint on detailed time allocations, the second major difference is that there is no financial market, Laporte doesn't provide too much details on this simplification, thus we shall provide our own version to justify this choice since in chapter 3 our model also doesn't include a financial market.

Let us assume a similar setting of environment like Laporte, except for now there exists a financial market, the interest rate is denoted as r . Let us also suggest that the exogenous part of income $Y(0)$ is dropped out, thus the individual's income depends completely on his health capital defined by the equation below

$$Y(H) = \theta H^\alpha$$

where θ and α are 2 exogenous constants. Plus, α are supposed to be less than 1.

Production function of health $G(I)$ is supposed to be equal to I for simplicity, for

the same purpose we also suppose that the individual only lives 2 periods, with the initial health capital H_1 given, the individual's problem could be therefore written as

$$\max U(C_1, H_1) + \beta U(C_2, H_2)$$

subject to

$$\begin{cases} H_2 = I_1 + (1 - \delta)H_1 \\ \theta H_1^\alpha = C_1 + I_1 + S_1 \\ \theta H_2^\alpha + (1 + r)S_1 = C_2 \end{cases}$$

The Lagrangian is written as

$$\begin{aligned} \mathcal{L} = & U(C_1, H_1) + \beta U(C_2, H_2) + \phi_H [H_2 - I_1 - (1 - \delta)H_1] \\ & + \phi_{A,1} (C_1 + I_1 + S_1 - \theta H_1^\alpha) \\ & + \phi_{A,2} (C_2 - \theta H_2^\alpha - (1 + r)S_1) \end{aligned}$$

where ϕ_H and ϕ_A are the shadow price of health and assets.

Differentiate the Lagrangian with respect to C_1 , H_2 , I_1 and S_1 , we get

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial C_1} &= U_{C_1} + \phi_{A,1} = 0 \\ \frac{\partial \mathcal{L}}{\partial H_2} &= \beta U_{H_2} + \phi_H - \phi_{A,2} \theta \alpha H_2^{\alpha-1} = 0 \\ \frac{\partial \mathcal{L}}{\partial I_1} &= -\phi_H + \phi_{A,1} = 0 \\ \frac{\partial \mathcal{L}}{\partial S_1} &= \phi_{A,1} - \phi_{A,2} (1 + r) = 0 \end{aligned}$$

combining these first order conditions we get the equilibrium condition

$$U_{C_1} = \beta [U_{H_2} + \frac{U_{C_1}}{1+r} \theta \alpha H_2^{\alpha-1}] \quad (2.21)$$

Now let's apply the method used by Grossman in his original article: let us also divide

our example into 2 branches, first we assume the health capital has (almost) no impact on the income, which corresponds to the pure consumption branch of the Grossman model, then we assume that the health capital doesn't enter directly into the utility function, which corresponds to the pure investment branch of the Grossman model. In addition, in order to illustrate concretely the situation, we also assume that the utility function takes the form

$$U(C, H) = \mu \ln(C) + (1 - \mu) \ln(H) \quad (2.22)$$

To neutralize the effect of health capital on the income, we assume that $\alpha = 0$, in this case, Rq [2.21](#) becomes

$$U_{C_1} = \beta U_{H_2} \quad (2.23)$$

By the first order conditions that give us Eq [2.21](#) we still have $U_{C_1} = \beta(1 + r)U_{C_2}$, let us also substitute $U(C, H)$ by Eq [2.23](#), thus we can write

$$\begin{cases} \frac{\mu}{C_1} = \beta(1 + r)\frac{\mu}{C_2} \\ \frac{\mu}{C_1} = \beta\frac{1-\mu}{H_2} \end{cases} \quad (2.24)$$

Since $\alpha = 0$, the budget constraint is reduced to

$$C_1 + \frac{C_2}{1 + r} + I_1 = \frac{\theta}{1 + r} + \theta = W \quad (2.25)$$

Substituting Eq [2.24](#) into Eq [2.25](#) and the law of motion of health capital, finally we can obtain

$$C_1 = \frac{\mu}{\mu + \beta} [W + (1 - \delta)H_1] \quad (2.26)$$

and

$$C_2 = \beta(1 + r)C_1 = \frac{\mu(\beta(1 + r))}{\mu + \beta} [W + (1 - \delta)H_1] \quad (2.27)$$

and

$$I_1 = \frac{1-\mu}{\mu}\beta C_1 - (1-\delta)H_1 \quad (2.28)$$

Now let's proceed to the second scenario, in which the health stock doesn't enter directly into the utility function. In this case, Eq 2.21 becomes

$$U_{C_1} = \frac{U_{C_1}}{1+r}\theta\alpha H_2^{\alpha-1} \quad (2.29)$$

Thus the health stock in the future H_2 only depends on interest rate

$$H_2 = \frac{1+r}{\theta\alpha} \frac{1}{\alpha-1} \quad (2.30)$$

In other words, as long as Eq 2.30 holds, under the assumption that the health capital has no direct impact on the individual's utility, to the individual, health investment and saving become indifferent.

But we still have $U_{C_1} = \beta(1+r)U_{C_2}$, thus we can write

$$\begin{cases} C_2 = \beta(1+r)C_1 \\ H_2 = \left(\frac{1+r}{\theta\alpha}\right)^{\frac{1}{\alpha-1}} \end{cases} \quad (2.31)$$

Substituting Eq 2.31 into the budget constraint and the law of motion of health, finally we can obtain

$$C_1 = \frac{1}{1+\beta} \left[\frac{1-\alpha}{\alpha} \left(\frac{1+r}{\theta\alpha}\right)^{\frac{1}{\alpha-1}} + \theta H_1^\alpha + (1-\delta)H_1 \right] \quad (2.32)$$

and

$$C_2 = \frac{\beta(1+r)}{1+\beta} \left[\frac{1-\alpha}{\alpha} \left(\frac{1+r}{\theta\alpha}\right)^{\frac{1}{\alpha-1}} + \theta H_1^\alpha + (1-\delta)H_1 \right] \quad (2.33)$$

and

$$I_1 = -(1 - \delta)H_1 + \left(\frac{1+r}{\theta\alpha}\right)^{\frac{1}{\alpha-1}} \quad (2.34)$$

Now let's see what would change if there doesn't exist a financial market. We assume that the setting of the model is exactly the same under this assumption than the previous one, except there is no interest rate. We shall use the same method to analyse the health investment and consumption decision, the individual's problem could be written as

$$\max U(C_1, H_1) + \beta U(C_2, H_2) \quad (2.35)$$

subject to

$$\begin{cases} H_2 = I_1 + (1 - \delta)H_1 \\ C_1 + I_1 = \theta H_1^\alpha \\ C_2 = \theta H_2^\alpha \end{cases}$$

If $\alpha = 0$, the Lagrangian would be written as

$$\begin{aligned} \mathcal{L} = & U(C_1, H_1) + \beta U(C_2, H_2) + \phi_H[H_2 - I_1 - (1 - \delta)H_1] \\ & + \phi_{A,1}(C_1 + I_1 - \theta) \\ & + \phi_{A,2}(C_2 - \theta) \end{aligned}$$

Differentiate the Lagrangian with respect to C_1 , H_2 and I_1 , we get

$$\beta U_{H_2} = U_{C_1} \quad (2.36)$$

which would give us

$$H_2 = \beta \frac{1 - \mu}{\mu} C_1 \quad (2.37)$$

Substituting Eq [2.37](#) into the budget constraint and the law of motion of health,

finally we can obtain

$$C_1 = \frac{\mu}{\mu + \beta - \mu\beta} [\theta + (1 - \delta)H_1] \quad (2.38)$$

and

$$C_2 = \theta \quad (2.39)$$

and

$$I_1 = \frac{1 - \mu}{\mu} \beta C_1 - (1 - \delta)H_1 \quad (2.40)$$

From Eq 2.38 and Eq 2.39, we can already sense the lack of consumption smoothing, indeed, when the financial market is excluded and under the assumption that health stock doesn't impact income, the individual invests in health only to raise his U_H , there is no substitution between the 2 periods in terms of consumption since C_2 is fixed no matter the health stock in the second period.

When $U_H = 0$, the Lagrangian of the individual's problem would be written as

$$\begin{aligned} \mathcal{L} = & U(C_1) + \beta U(C_2) + \phi_H [H_2 - I_1 - (1 - \delta)H_1] \\ & + \phi_{A,1} (C_1 - I_1 - \theta H_1^{\alpha-1}) \\ & + \phi_{A,2} (C_2 - \theta H_2^\alpha) \end{aligned}$$

Differentiate the Lagrangian with respect to C_1 , C_2 , I_1 and H_2 , we can obtain

$$U_{C_1} = \beta \theta \alpha H_2^{\alpha-1} U_{C_2} \quad (2.41)$$

which would give us

$$C_2 = \beta \theta \alpha H_2^{\alpha-1} C_1 \quad (2.42)$$

Eq 2.41 shows the substitution between C_1 and C_2 under the assumption that $U_H = 0$. Unlike the case when there is a financial market, here the substitution is taking place via

the channel of health: health investment in the first period pushes up the health stock, which in turn produces more income in the second period that would be completely transferred into consumption.

Substituting Eq [2.42](#) into the budget constraint and the law of motion of health, finally we can obtain

$$C_1 = \frac{\theta H_1^\alpha + (1 - \delta)H_1}{1 + \beta\alpha} \quad (2.43)$$

and

$$C_2 = \theta H_2^\alpha = \theta\beta^\alpha \alpha^\alpha \left(\frac{\theta H_1^\alpha + (1 - \delta)H_1}{1 + \beta\alpha} \right)^\alpha \quad (2.44)$$

and

$$I_1 = \beta\alpha C_1 - (1 - \delta)H_1 \quad (2.45)$$

The above paragraphs conclude the consumption and health investment decisions in each scenario, with or without the financial market. As we can see, when there is a financial market, the Euler condition $U_{C_1} = \beta(1+r)U_{C_2}$ always holds, thus the individual would always want to smooth his consumption as long as β tends to $\frac{1}{1+r}$, no matter how he invests in health. On the contrary, when there is no financial market, this substitution more or less disappears, under the pure consumption assumption, C_1 and C_2 are detached, under the pure investment assumption, the substitution effect between C_1 and C_2 is still there but it's taking place via the channel of health capital.

To see more clearly the change, let us take the numeric approach. Let us assume that $\theta = 1.4$, $\alpha = 0.35$, $\beta = 0.96$, $\mu = 0.9$, $\delta = 0.7$ and $H_1 = 0.6$, we first set $H_1 = 0.6$, and we calculate C_1 , C_2 , I_1 and S_1 in each scenario, afterwards, in order of comparison, we set $H_1 = 0.5$ and we recalculate the values of each variable in each scenario. The values of parameters are chosen such that the comparison of results could be more evident, other than this there is no particular reason that we made this choice. Table [2.6](#) summarizes

the value of each variable.

$H_1 = 0.6$		C_1	C_2	I_1	S_1
With a financial market	Pure Investment	1.5988	1.9966	0.3308	0.0765
	Pure Consumption	2.1417	2.6728	0.0484	0.2099
Without a financial market	Pure Investment	1.637	1.9469	0.37	0
	Pure Consumption	2.3313	2.4	0.0687	0
$H_1 = 0.5$		C_1	C_2	I_1	S_1
With a financial market	Pure Investment	1.5212	1.8984	0.3608	0.0011
	Pure Consumption	2.1272	2.6547	0.0769	0.1959
Without a financial market	Pure Investment	1.5217	1.8978	0.3613	0
	Pure Consumption	2.3042	2.4	0.0958	0

Table 2.6: Consumption and health investment decisions under each assumption, with or without the financial market

We can see that with or without a financial market, the consumption smoothing exists always, especially under the pure investment assumption, nevertheless, under the pure consumption assumption, since C_1 and C_2 are completely detached, we see there is nothing can be described as substitution between the 2 periods.

The key assumption that leads us to this result is that δ is assumed to be constant, which means that the health capital would deteriorate at a constant rate. Given the hypothesis that human body would not simply collapse in a single moment, it means that at each period the individual is promised to receive a certain amount of money as income, even it decreases with time, plus, the production function of health is assumed to be CRS, thus the health investment would effectively help slow down the health depreciation without a concern of its scale, consequently, this ‘permanent’ income flow that is promised to the individual would be rather stable than volatile⁶, in other words, when health still has impacts on the individual’s income and doesn’t just decrease way too fast, health capital functions similarly as the financial market: it provides a channel that allows that individual to substitute his consumptions between time periods, the individual could

⁶Of course the elasticity of income with respect to health α should also be relatively small, otherwise we would have a rather volatile income flow.

choose to consume less and invest in health in the present and receive an extra amount of health in the future, this extra amount of health would help ease the depreciation and thus help stabilize the health stock, which in turn would stabilize also the income flow, as we can see, even with our a little extreme example, the consumption flow is not as volatile as one would expect. In short, when there isn't a financial market, the health capital could play a similar role, thus the individual's consumption choice wouldn't be affected too much, of course, when the financial market and the health capital exist in the same time, the consumption flow would be even smoother, but as we can see, the absence of the financial market doesn't impact too much to the extent that it's unacceptable.

Therefore it justifies Laporte's choice of simplification, which is to exclude the financial market from his model, since he assumes a constant health depreciation rate and a CRS production function of health, this simplification shouldn't generate too much loss on the analysis of human behaviours concerning health, on the contrary, in chapter 3, though we use Laporte's model as the basis of our model, we assume an increasing health depreciation rate that depends on age, plus, in our model, the production function of health is still CRS, hence as we shall see in the result section of chapter 3, our model generates a decreasing income profile as well as a decreasing consumption profile, this is indeed one of the flaws of our model and this is also why we conclude that the existence of a financial market is necessary when health depreciation rate is increasing with age.

2.5 Optimal health investment choice strategy under uncertainty

In this section, we shall continue using the simple example of our own, but this time, a risk would be introduced, the risk would take 2 forms: first a risk of health, then a risk of income.

2.5.1 The risk of health

Let us assume that there is no change in the first period, but in the second period, the depreciation rate of health stock could take 2 values: either $\underline{\delta}$ with probability p or $\bar{\delta}$ with probability $(1 - p)$ and $\bar{\delta} > \underline{\delta}$. Therefore the health stock in the second period could also take 2 values: either \bar{H} or \underline{H} with

$$\begin{cases} \bar{H} = (1 - \underline{\delta})H_1 + I_1 \\ \underline{H} = (1 - \bar{\delta})H_1 + I_1 \end{cases} \quad (2.46)$$

The individual's problem Therefore the individual's problem would be written as

$$\max U(C_1, H_1) + \beta E[U(C_2, H_2)] \quad (2.47)$$

subject to

$$\begin{cases} \theta H_1^\alpha = C_1 + I_1 + S_1 \\ \bar{H}_2 = I_1 + (1 - \underline{\delta})H_1 \\ \theta \bar{H}_2^\alpha + (1 + r)S_1 = \bar{C}_2 \\ \underline{H}_2 = I_1 + (1 - \bar{\delta})H_1 \\ \theta \underline{H}_2^\alpha + (1 + r)S_1 = \underline{C}_2 \end{cases} \quad (2.48)$$

The Lagrangian could be written as

$$\begin{aligned}
\mathcal{L} = & U(C_1, H_1) + \beta[pU(\bar{C}_2, \bar{H}_2) + (1-p)U(\underline{C}_2, \underline{H}_2)] \\
& + \phi_{A,1}(\theta H_1^\alpha - C_1 - I_1 - S_1) \\
& + \phi_{\bar{A},2}(\theta \bar{H}_2^\alpha - \bar{C}_2 - (1+r)S_1) \\
& + \phi_{\underline{A},2}(\theta \underline{H}_2^\alpha - \underline{C}_2 - (1+r)S_1) \\
& + \phi_{\bar{H}}(\bar{H}_2 - I_1 - (1-\underline{\delta})H_1) \\
& + \phi_{\underline{H}}(\underline{H}_2 - I_1 - (1-\bar{\delta})H_1)
\end{aligned} \tag{2.49}$$

Differentiate the Lagrangian with respect to C_1 , S_1 , I_1 , \bar{C}_2 , \underline{C}_2 , \bar{H}_2 and \underline{H}_2 , we obtain

$$\begin{aligned}
\frac{\partial \mathcal{L}}{\partial C_1} &= U_{C_1} - \phi_{A,1} = 0 \\
\frac{\partial \mathcal{L}}{\partial S_1} &= -\phi_{A,1} + (1+r)(\phi_{\bar{A},2} + \phi_{\underline{A},2}) = 0 \\
\frac{\partial \mathcal{L}}{\partial I_1} &= -\phi_{A,1} - \phi_{\bar{H}} - \phi_{\underline{H}} = 0 \\
\frac{\partial \mathcal{L}}{\partial \bar{C}_2} &= \beta p(U_{\bar{C}_2}) - \phi_{\bar{A},2} = 0 \\
\frac{\partial \mathcal{L}}{\partial \underline{C}_2} &= \beta(1-p)(U_{\underline{C}_2}) - \phi_{\underline{A},2} = 0 \\
\frac{\partial \mathcal{L}}{\partial \bar{H}_2} &= \beta p(U_{\bar{H}_2}) + \phi_{\bar{A},2}\theta\alpha\bar{H}_2^{\alpha-1} + \phi_{\bar{H}} = 0 \\
\frac{\partial \mathcal{L}}{\partial \underline{H}_2} &= \beta(1-p)(U_{\underline{H}_2}) + \phi_{\underline{A},2}\theta\alpha\underline{H}_2^{\alpha-1} + \phi_{\underline{H}} = 0
\end{aligned}$$

Combing these first order conditions, we can write

$$\begin{cases} U_{C_1} = \beta E(U_{H_2}) + \phi_{\bar{A},2}\theta\alpha\bar{H}_2^{\alpha-1} + \phi_{\underline{A},2}\theta\alpha\underline{H}_2^{\alpha-1} \\ U_{C_1} = \beta(1+r)E(U_{C_2}) \end{cases} \tag{2.50}$$

Eq 2.50 shows that when an uncertainty is introduced, a precautionary action will take place. In a standard stochastic model, this action would function via the channel of the financial market, hence the famous ‘‘precautionary saving’’. In our example, since one’s health stock also affects one’s future income, thus the precautionary action would

function via 2 channels: either by precautionary savings or by health investment, the second equation of Eq 2.50 shows how the individual would smooth his consumptions between periods via the precautionary savings, which is pretty standard, the first equation of Eq 2.50 shows that how the individual could smooth his consumption by investing in health: the more the individual invests in health, the more income it would produce in the second period⁷($\phi_{\bar{A},2}\theta\alpha\bar{H}_2^{\alpha-1} + \phi_{\underline{A},2}\theta\alpha\underline{H}_2^{\alpha-1}$), thus the more he would be able to consume, plus, the higher the individual's health stock, also the higher his utilities($E(U_{H_2})$). The individual has to choose between the 2 channels in order to maximize his utilities.

Let us continue using the same form of the utility function as in the previous section, plus, let us also assume that $E(\delta) = p\underline{\delta} + (1-p)\bar{\delta} = \delta^d$, where δ^d indicates the depreciation rate of health in the deterministic environment.

Since we have 12 unknown variables: $\phi_{A,1}$, $\phi_{\bar{A},2}$, $\phi_{\underline{A},2}$, $\phi_{\bar{H}}$, $\phi_{\underline{H}}$, C_1 , S_1 , I_1 , \bar{C}_2 , \underline{C}_2 , \bar{H}_2

⁷Of course it also depends on the elasticity of income with respect to health, α .

and \underline{H}_2 and also 12 equations:

$$\left\{ \begin{array}{l} \theta H_1^\alpha = C_1 + I_1 + S_1 \\ \bar{H}_2 = I_1 + (1 - \underline{\delta})H_1 \\ \theta \bar{H}_2^\alpha + (1 + r)S_1 = \bar{C}_2 \\ \underline{H}_2 = I_1 + (1 - \bar{\delta})H_1 \\ \theta \underline{H}_2^\alpha + (1 + r)S_1 = \underline{C}_2 \\ \frac{\mu}{C_1} - \phi_{A,1} = 0 \\ -\phi_{A,1} + (1 + r)(\phi_{\bar{A},2} + \phi_{A,2}) = 0 \\ -\phi_{A,1} - \phi_{\bar{H}} - \phi_{\underline{H}} = 0 \\ \beta p \left(\frac{\mu}{C_2} \right) - \phi_{\bar{A},2} = 0 \\ \beta(1 - p) \left(\frac{\mu}{C_2} \right) - \phi_{A,2} = 0 \\ \beta p \left(\frac{1 - \mu}{\bar{H}_2} \right) + \phi_{\bar{A},2} \theta \alpha \bar{H}_2^{\alpha-1} + \phi_{\bar{H}} = 0 \\ \beta(1 - p) \left(\frac{1 - \mu}{\underline{H}_2} \right) + \phi_{A,2} \theta \alpha \underline{H}_2^{\alpha-1} + \phi_{\underline{H}} = 0 \end{array} \right. \quad (2.51)$$

thus it should be sufficient to find the solution to the problem, using a standard Newton algorithm.

In the previous section, we divided our example into 2 branches: either health status doesn't affect the individual's income, either health doesn't enter the individual's utility function. Let us begin with the former branch, in our example, it would mean that $\alpha = 0$, under this condition, Eq 2.50 becomes

$$\left\{ \begin{array}{l} U_{C_1} = \beta E(U_{H_2}) \\ U_{C_1} = \beta(1 + r)U_{C_2} \end{array} \right. \quad (2.52)$$

Note that when $\alpha = 0$, the individual would receive the same amount of income in both periods(θ), no matter his health status, thus the expected value symbol is excluded in the second equation of Eq 2.52, this also means that there will be no precautionary

savings because of the uncontrollable income fluctuation.

Take the Taylor expansion of RHS of the first equation of Eq 2.52 around $E(H_2)$, we can write

$$\begin{aligned} \beta E((1 - \mu)H_2^{-1}) &= \beta \{E[(1 - \mu)(E(H_2))^{-1}] - (1 - \mu)[(E(H_2))^{-2}]E[H_2 - E(H_2)] \\ &+ (1 - \mu)[(E(H_2))^{-3}]E[H_2 - E(H_2)]^2 + o(\sigma^2)\} \end{aligned} \quad (2.53)$$

where $o(\sigma^2)$ represents high order moments.

Since the term $E[H_2 - E(H_2)] = 0$, and let us omit the term $o(\sigma^2)$, then Eq 2.53 would become

$$\beta E((1 - \mu)H_2^{-1}) \approx \beta \{(1 - \mu)(E(H_2))^{-1} + (1 - \mu)[(E(H_2))^{-3}]E[H_2 - E(H_2)]^2\} \quad (2.54)$$

By definition, $E(H_2)$ denotes the expected value of health stock in the second period, however, it could also indicate, under certainty and with the same amount of health investment as in the stochastic environment, the amount of health capital that the individual could possess, if this is the case, then such amount of health $E(H_2)$ would also be matched by a set of consumption choice, let us call it $\{EC_1, EC_2\}$.

Now take the Taylor expansion of the LHS of the first equation of Eq 2.52 around EC_1 , we can write

$$U_{C_1} \approx U_{EC_1} + U_{EC_1 EC_1}(C_1 - EC_1) \quad (2.55)$$

Substituting Eq 2.55 and Eq 2.54 into the first equation of Eq 2.52 and rearrange, one can easily obtain

$$C_1 - EC_1 = \frac{\beta}{U_{EC_1 EC_1}} \times (1 - \mu)[(E(H_2))^{-3}]E[H_2 - E(H_2)]^2 \quad (2.56)$$

Since $U_{EC_1EC_1}$ is obviously less than 0, hence the RHS of Eq 2.56 would also be less than 0, meaning that $C_1 < EC_1$. That's to say, because of the introduction of a health risk, the individual would choose to consume less in the first period, plus, as what we have mentioned, in this scenario there would be no precautionary saving because of income fluctuations, thus this supplement amount of resources that the individual possesses would be completely invested into health. In addition, Eq 2.56 states that the difference between C_1 and EC_1 depends on the importance of the variance of health, which is directly affected by the variance of δ , thus we can deduce that the larger the variance of δ , the smaller C_1 .

The problem can only be solved numerically, since the investment aspect of health is excluded, the system 2.51 can be reduced to

$$\left\{ \begin{array}{l} \theta = C_1 + I_1 + S_1 \\ \bar{H}_2 = I_1 + (1 - \underline{\delta})H_1 \\ \theta + (1 + r)S_1 = C_2 \\ \underline{H}_2 = I_1 + (1 - \bar{\delta})H_1 \\ \frac{\mu}{C_1} - \phi_{A,1} = 0 \\ -\phi_{A,1} + (1 + r)\phi_{A,2} = 0 \\ -\phi_{A,1} - \phi_{\bar{H}} - \phi_{\underline{H}} = 0 \\ \beta p \left(\frac{\mu}{C_2} \right) - \phi_{A,2} = 0 \\ \beta p \left(\frac{1 - \mu}{\underline{H}_2} \right) + \phi_{\bar{H}} = 0 \\ \beta(1 - p) \left(\frac{1 - \mu}{\underline{H}_2} \right) + \phi_{\underline{H}} = 0 \end{array} \right. \quad (2.57)$$

Rearrange the system 2.57, we can reduce it to

$$\left\{ \begin{array}{l} (1 + \beta)C_1 + I_1 - \frac{\theta}{1+r} - \theta = 0 \\ \frac{\mu}{1-\mu} \left[\frac{p}{(1-\underline{\delta})H_1 + I_1} + \frac{1-p}{(1-\underline{\delta})H_1 + I_1} \right] - \beta C_1 = 0 \end{array} \right. \quad (2.58)$$

The system [2.58](#) is composed of 2 unknown variables (C_1 and I_1) and 2 equations, it should be fairly easy enough to compile the programme, using the standard Newton algorithm. For the purpose of comparison, we shall take the same parameter values as in the last section as benchmark calibrations. As for the value of $\bar{\delta}$, $\underline{\delta}$ and p , we choose $\bar{\delta} = 0.8$, $\underline{\delta} = 0.6$ and $p = 0.5$, once again there is no particular reason that we make this choice, these values are chosen only to make sure that the expected value of δ would be equal to its deterministic counterpart. The results can be seen in [Table 2.7](#)

	C_1	I_1	S_1
Under certainty, benchmark	2.1417	0.0484	0.2099
Under certainty, with same I_1 as under uncertainty	2.2734	0.0625	0.218
Uncertainty, $\bar{\delta} = 0.6$, $\underline{\delta} = 0.8$	2.1345	0.0625	0.203
Uncertainty, $\bar{\delta} = 0.5$, $\underline{\delta} = 0.9$	2.1166	0.0976	0.1858

Table 2.7: Consumption and health investment decisions under the pure consumption assumption, with the financial market

[Table 2.7](#) shows that when the uncertainty is introduced and under the pure consumption assumption, the individual will indeed choose to reduce his consumption in the first period, in order to invest more in health rather than saving, since in this situation the individual's income is not affected, note that if we are to have the same health investment under certainty than under uncertainty, θ needs to be higher than in the benchmark calibration, which means that the individual's health investment choice under uncertainty actually matches the individual's health investment choice under certainty but with a higher health production efficiency, showing the effect of the uncertainty. In addition, as we can see, the more dispersion between the "good" situation ($\underline{\delta}$) and the "bad" situation ($\bar{\delta}$), the more health investment, and the less the C_1 as well as S_1 , this result corresponds to our assumption and is actually logic: the riskier the uncertainty, the more incentive the individual would have to ensure themselves against it, and under the pure consumption branch, the only channel that remains is the health investment channel.

Now let's take a look at the other branch, in which health stock is excluded from the utility function, therefore Eq 2.50 would become

$$\begin{cases} U_{C_1} = +\phi_{\bar{A},2}\theta\alpha\bar{H}_2^{\alpha-1} + \phi_{\underline{A},2}\theta\alpha\underline{H}_2^{\alpha-1} \\ U_{C_1} = \beta(1+r)E(U_{C_2}) \end{cases} \quad (2.59)$$

In this scenario, the arbitrage that the individual has to make becomes more apparent: since health doesn't affect one's utility any more, the sole reason to invest in health is to compensate the uncontrollable loss of health via this investment, and thus ensure a more predictable income flow, or, the individual can achieve this goal via savings, the 2 channels of our example functions now to the same end. Thus we can deduce that in this scenario, the individual's consumption choice won't be affected too much with the introduction of the uncertainty, and the arbitrage of the choice between health investment and saving should depends on the attraction of these 2 channels(α and r).

Unfortunately, it's not likely that we can give out some analysis of this branch using the same fashion as what we did for the previous branch, thus numerically, we can only compare the results under uncertainty with those under certainty and not with their certainty equivalents.

The problem can also be solved numerically using the same method as for the pure consumption branch, since health is excluded from the utility function, the system 2.51

can be rewritten as

$$\left\{ \begin{array}{l} \theta H_1^\alpha = C_1 + I_1 + S_1 \\ \bar{H}_2 = I_1 + (1 - \underline{\delta})H_1 \\ \theta \bar{H}_2^\alpha + (1 + r)S_1 = \bar{C}_2 \\ \underline{H}_2 = I_1 + (1 - \bar{\delta})H_1 \\ \theta \underline{H}_2^\alpha + (1 + r)S_1 = \underline{C}_2 \\ \frac{\mu}{C_1} - \phi_{A,1} = 0 \\ -\phi_{A,1} + (1 + r)(\phi_{\bar{A},2} + \phi_{\underline{A},2}) = 0 \\ -\phi_{A,1} - \phi_{\bar{H}} - \phi_{\underline{H}} = 0 \\ \beta p \left(\frac{\mu}{C_2} \right) - \phi_{\bar{A},2} = 0 \\ \beta(1 - p) \left(\frac{\mu}{C_2} \right) - \phi_{\underline{A},2} = 0 \\ \phi_{\bar{A},2} \theta \alpha \bar{H}_2^{\alpha-1} + \phi_{\bar{H}} = 0 \\ \phi_{\underline{A},2} \theta \alpha \underline{H}_2^{\alpha-1} + \phi_{\underline{H}} = 0 \end{array} \right. \quad (2.60)$$

Rearrange the system [2.60](#), it can be reduced to

$$\left\{ \begin{array}{l} \theta H_1^\alpha - C_1 - I_1 - S_1 = 0 \\ \frac{1}{C_1} - \beta(1 + r) \left[\frac{1}{\bar{C}_2} + \frac{1-p}{\underline{C}_2} \right] = 0 \\ \frac{1}{C_1} - \frac{\beta p}{C_2} \theta \alpha \bar{H}_2^{\alpha-1} - \frac{\beta(1-p)}{\underline{C}_2} \theta \alpha \underline{H}_2^{\alpha-1} = 0 \end{array} \right. \quad (2.61)$$

where $\bar{H}_2 = I_1 + (1 - \underline{\delta})H_1$, $\underline{H}_2 = I_1 + (1 - \bar{\delta})H_1$, $\bar{C}_2 = \theta \bar{H}_2^\alpha + (1 + r)S_1$ and $\underline{C}_2 = \theta \underline{H}_2^\alpha + (1 + r)S_1$.

The system [2.61](#) is composed of 3 equations and 3 unknown variables (C_1 , I_1 and S_1), thus the same method can be applied to solve the problem. We shall always use the same parameter values and we first calculate the values of unknown variables with $\alpha = 0.35$, then we raise it to 0.38 and decrease it to 0.25 and recalculate the values of these variables in each scenario. As for the interest rate r , we do the same thing: we first calculate the values of unknown variables with $r = 0.3$, then we raise it to 0.4 and decrease it to 0.25

and recalculate the values of these variables in each scenario.

The results are shown in Table 2.8 and Table 2.9.

	C_1	I_1	S_1
Under certainty, benchmark	1.5988	0.3308	0.0765
Under certainty, $\alpha = 0.25$	1.7155	0.1767	0.2201
Uncertainty, $\alpha = 0.25$	1.7125	0.1874	0.2124
Uncertainty, $\alpha = 0.35$	1.5975	0.3388	0.0708
Uncertainty, $\alpha = 0.38$	1.5680	0.392	0.0165

Table 2.8: Consumption and health investment decisions under the pure investment assumption with change of α , with the financial market

	C_1	I_1	S_1
Under certainty, benchmark	1.5988	0.3308	0.0765
Under certainty, $r = 0.25$	1.6299	0.3625	0.0147
Uncertainty, $r = 0.25$	1.6276	0.3702	0.0092
Uncertainty, $r = 0.3$	1.5975	0.3388	0.0708
Uncertainty, $r = 0.4$	1.5451	0.2845	0.1775

Table 2.9: Consumption and health investment decisions under the pure investment assumption with change of r , with the financial market

Table 2.8 and Table 2.9 show that when the uncertainty is introduced, the individual will choose to either save more or invest more in health. The arbitrage actually depends on the importance of the income elasticity α and the interest rate r : if α is not large enough, then health investment would not produce enough resources in the second period, thus the individual will choose to save more, on the contrary, if α is large enough that sufficient income could be produced via the channel of health, then the individual would choose to invest more in health, even it means to cut part of his consumption. The same trend could be observed when r changes: the higher the interest rate, the more important the savings and the less important the health investment, on the contrary, when r is low, the individual would choose to invest more in health and save less.

2.5.2 The risk of income

Now let us examine what would happen when the uncertainty is brought upon income and not on health. By this assumption, δ would return constant and α should be equal to 0, plus, for there is a risk of income, we assume that in the second period, the individual's income θ_2 would take 2 values: either $\bar{\theta}$ with probability p or $\underline{\theta}$ with probability $(1-p)$, we also assume that $E(\theta_2) = \theta$, in the meantime, we continue assuming that the individual's income in the first period remains unchanged with $\theta_1 = \theta$.

The individual's problem Therefore the individual's problem would be written as

$$\max U(C_1, H_1) + \beta[pU(\bar{C}_2, H_2) + (1-p)U(\underline{C}_2, H_2)] \quad (2.62)$$

subject to

$$\begin{cases} \theta_1 = C_1 + I_1 + S_1 \\ H_2 = (1 - \delta)H_1 + I_1 \\ \bar{\theta} + (1 + r)S_1 = \bar{C}_2 \\ \underline{\theta} + (1 + r)S_1 = \underline{C}_2 \end{cases} \quad (2.63)$$

The Lagrangian could be written as

$$\begin{aligned} \mathcal{L} = & U(C_1, H_1) + \beta[pU(\bar{C}_2, H_2) + (1-p)U(\underline{C}_2, H_2)] \\ & + \phi_{A,1}(\theta - C_1 - I_1 - S_1) \\ & + \phi_{\bar{A},2}(\bar{\theta} - \bar{C}_2 + (1+r)S_1) \\ & + \phi_{\underline{A},2}(\underline{\theta} - \underline{C}_2 + (1+r)S_1) \\ & + \phi_H(H_2 - I_1 - (1-\delta)H_1) \end{aligned} \quad (2.64)$$

Differentiate the Lagrangian with respect to C_1 , S_1 , I_1 , \bar{C}_2 , \underline{C}_2 and H_2 , we obtain

$$\begin{aligned}
\frac{\partial \mathcal{L}}{\partial C_1} &= U_{C_1} - \phi_{A,1} = 0 \\
\frac{\partial \mathcal{L}}{\partial S_1} &= -\phi_{A,1} + (1+r)(\phi_{\bar{A},2} + \phi_{\underline{A},2}) = 0 \\
\frac{\partial \mathcal{L}}{\partial I_1} &= -\phi_{A,1} - \phi_H = 0 \\
\frac{\partial \mathcal{L}}{\partial \bar{C}_2} &= \beta p(U_{\bar{C}_2}) - \phi_{\bar{A},2} = 0 \\
\frac{\partial \mathcal{L}}{\partial \underline{C}_2} &= \beta(1-p)(U_{\underline{C}_2}) - \phi_{\underline{A},2} = 0 \\
\frac{\partial \mathcal{L}}{\partial H_2} &= \beta U_{H_2} + \phi_H = 0
\end{aligned}$$

By these first order conditions we could obtain

$$\begin{cases} U_{C_1} = \beta U_{H_2} \\ U_{C_1} = \beta(1+r)E(U_{C_2}) \end{cases} \quad (2.65)$$

System [2.65](#) states the classic intertemporal substitution between periods, as what the first equation of the system shows, since here the uncertainty is only on the income, then there should be no precautionary actions in terms of health investment, in the meantime, the individual should prepare some precautionary savings against the odds of income.

The situation is completely contrary to the pure consumption branch we have mentioned in the previous subsection when the risk of health is introduced, but they all follow the same logic: to see this, take the Tylor expansion of the RHS of the second equation of system [2.65](#) around EC_2 , where EC_2 denotes the certainty equivalent of C_2 , the equation could be written as

$$\beta(1+r)E(U_{C_2}) = \beta(1+r)\left\{U_{EC_2} + \frac{1}{2}U_{EC_2 EC_2 EC_2}E[C_2 - EC_2]^2\right\} \quad (2.66)$$

Since we have defined EC_2 , we could also define EC_1 as its counterpart of the first period in the deterministic environment, now take the Tylor expansion of the LHS of the

second equation of system [2.65](#) around EC_1 , we can write

$$U_{C_1} = U_{EC_1} + U_{EC_1EC_1}(C_1 - EC_1) \quad (2.67)$$

Substituting Eq [2.66](#) and Eq [2.67](#) into the second equation of the system [2.65](#), we can find

$$C_1 - EC_1 = \frac{1}{2} \frac{U_{EC_2EC_2EC_2}}{U_{EC_1EC_1}} E[C_2 - EC_2]^2 \quad (2.68)$$

Since $U_{EC_2EC_2EC_2}$ is greater than 0 and $U_{EC_1EC_1}$ is less than 0, then we have $C_1 < EC_1$, showing that when the individual's income becomes unpredictable, the individual would reduce his consumption in order to save more, thus the famous “precautionary savings”. In addition, the difference between C_1 and EC_1 depends on the variance of C_2 , which, in our example, is directly affected by the variance of θ_2 .

The problem can also be solved numerically, using the very same Newton algorithm: with the same form of utility function, the problem could be reduced to

$$\begin{cases} \frac{1}{C_1} - \beta(1+r) \left[\frac{p}{\theta+(1+r)S_1} + \frac{1-p}{\underline{\theta}+(1+r)S_1} \right] = 0 \\ \frac{1-\mu}{\mu} \beta C_1 - C_1 + \theta + (1-\delta)H_1 - S_1 = 0 \end{cases} \quad (2.69)$$

There are 2 unknown variables (C_1 and S_1) and 2 equations, thus it's sufficient to find a solution. The results are shown in the Table [2.10](#).

	C_1	I_1	S_1
Under certainty, benchmark	2.1417	0.0484	0.2099
Under certainty, with same S_1 as under uncertainty	2.1493	0.0493	0.2103
Uncertainty, $\theta = 2.45$, $\underline{\theta} = 2.35$	2.1413	0.0484	0.2103
Uncertainty, $\theta = 2.5$, $\underline{\theta} = 2.3$	2.1403	0.0483	0.2114

Table 2.10: Consumption and health investment decisions with change of θ_2

From Table [2.10](#) we can see that under uncertainty of income, the individual would

choose to consume less and save more, plus, as θ_2 becomes more and more disperse, the individual would choose to consume even less in the first period, in order to save even more rather than investing more in health, for the only uncertainty in this situation is the odds of income. Besides, if we are to have the same amount of savings under certainty than under uncertainty, it also requires a higher θ , which is the same thing as what we have seen in the pure consumption assumption when there is a risk of health.

The above fictive example shows one simple way to introduce the risk, though its simplicity and its naivety, the example provides a first impression of how the uncertainty would alter the individual's decisions. In chapter 4, we shall take this stochastic approach and build a structural model upon the aiyagari's model, in which a risk of health is added, as well as a risk of income⁸.

2.6 Conclusion

In this chapter we answer the 2 questions proposed at the beginning of the thesis by reviewing existing literatures, the review of literatures is undertaken by 2 approaches: both the deterministic approach and the stochastic approach, in the deterministic approach we show that the Grossman health capital model turns to be still very handful in terms of estimating individual life-cycle profiles when there are fluctuations of health, but criticisms argue that the lack of uncertain nature of health makes the Grossman health capital model unreliable, nevertheless, as what we have shown in the second approach, the stochastic approach, this problem can be eventually 'fixed' by introducing a stochastic process into the model and we also show that how the uncertainty would alter the individual's decisions.

In latter chapters, we will continue following this path and show our contributions to

⁸But it doesn't depend on the agent's health status

each approach. In addition, we also seek to answer the following 2 more precise questions by the following chapters:

- We argue that the human capital model leaves window for immortality, so is it possible to close this window by adding an finite time horizon? That is, a lifetime health decision model within a reasonable life length.
- Characterizing shocks(for example, earning shocks or health shocks) by a Markov chain, we can construct models that replicate the structure of a certain economy, but is it possible to have a unified framework that can be applied to multiple countries ? If yes, can then we use this model to not only identify the differences across countries in terms of health, but also explore the key elements of public policies that can help improve the situation ?

Chapter 3

A Life-Cycle Model with health spending and endogenous lifetime

3.1 Introduction

Health spending holds a growing place in economic debate. At the macroeconomic level, health spending represents a growing share of the GDP. At the microeconomic level, health spending is a key variable in agent choices over the life-cycle. Moreover, it tends to represent a growing share in agents' spending over his life cycle. The issue of health spending and its growing trend may stem from several factors: ageing, scientific innovations or new medical treatments, etc.. Ageing leads to an increase in demand and scientific innovations lead to an increase in supply (new treatments, quality improvement,...).

In this article, we are interested in studying health spending decisions over the life-cycle. What are the main features of health spending over the life-cycle? French, McCauley, Aragon *et al.* (2017)[\[20\]](#) provide a lot of data concerning end-of-life medical spending in a large set of countries. They focus particularly on the last 3 years (and the last year) of life. In most countries, medical spending made at the last year of life

represents 8 to 11% of the aggregate medical spending (8.5% in the US). Concerning the last 3 years of life, the medical spending made during this period represents between 16.7 and 24.5% of the aggregate medical spending made over the life-cycle (16.7% in the US). Even if we consider the heterogeneity between countries, it is clear that medical spending in the last 3 years of life represents a significant share of the aggregate medical spending. Using US data, Halliday, He, Ning, and Zhang (2017)[27] and Fonseca *et al.* (2008)[18] report an increasing profile of medical spending over the life-cycle, medical spending being approximatively zero around 20-25 years old. Jung and Tran (2013)[31] reach similar conclusions. They also estimate the evolution of the share of health spending. At 20-25 years old, health spending is zero while at 85, it roughly represents more than 50% of income. The profile is increasing and convex. Finally, we underline that at the aggregate level, medical spending represent around 17% of the GDP in the US.

From a theoretical point of view, Grossman (1972)[23] develops a framework including a modeling of health spending. Grossman[23]'s model is a life-cycle one and he considers the health level may be represented by a stock variable, that is the health stock or health capital. The health stock has a positive impact on the agent's welfare. It depreciates at a rate possibly increasing with age. The health stock can also be improved or maintained by continuously health investment. Finally, the income of the agent is an increasing function of his health. If the health stock becomes less than a threshold, the agent dies, lifespan is an endogenous variable. The health capital concept is in the continuity of the human capital concept introduced among others by Ben-Porath (1967)[5]. The agent has also access to a financial market on which a risk-free asset is exchanged. Grossman[23]'s model includes several other features we will not detail.

Grossman[23]'s model is at the base of a vast literature. Halliday, He, Ning, and Zhang (2017)[27] and Fonseca *et al.* (2008)[18] use life-cycle models including an health stock variable, an endogenous health spending and a financial asset. They evaluate the

ability of their models to quantitatively replicate some stylized facts. Fonseca *et al.* (2008)[18] are interested by retirement choices in an uncertain environment. They simulate the path of some key variables and discuss their adequacy with their empirical counterpart. Halliday, He, Ning, and Zhang (2017)[27] focus on the growing profile of health spending over the life-cycle. Their model, with endogenous labour supply, is calibrated, numerically solved and simulated. They discuss the ability of their model to replicate the profile of some variables such as consumption or medical spending.

Some authors theoretically study the Grossman[23]'s model. In order to obtain analytical results and to isolate the key mechanism, they use a simplified version (often in continuous time) with only a health stock variable and without a financial asset. Laporte (2015)[34] derives some results concerning the dynamic of the model (uses of phases diagrams). Strulik (2015)[47] obtains a closed-form solution and performs exercises of comparative dynamics. Finally, Laporte and Ferguson (2017)[35] introduce uncertainty into the model.

In the contributions of Laporte (2015)[34], Laporte and Ferguson (2017)[35] and Strulik (2015)[47], the depreciation rate of health stock is constant over time, which allows the agent to potentially live forever. Indeed, in the Grossman's model, the agent stays alive as long as the health stock is above the threshold. If the health stock depreciation rate is constant, the agent may have the possibility to indefinitely maintain its health stock above the threshold. In this article, we consider a more realistic assumption and suppose that the health stock depreciation rate increases as the agent ages. Lifespan is now an endogenous variable. In order to avoid utility being negative, we use an utility function with a strong and positive enough constant, which is in line with that of Rosen (1988)[42] and Hall and Jones (2007)[26]. This model, written in discrete time, is solved numerically. A realistic calibration with a life span of 85 years shows that our model is able to replicate a growing path of the health spending over the life-cycle and a significant share of health

spending especially at the end of life. We also perform a sensitivity analysis, which allows us to evaluate how the numerical results are modified if the value of some key parameters are changed.

The paper is organized as follows. In section 2 we present the model and derive the optimality conditions, in section 3, we discuss the calibration strategy and the numerical results.

3.2 A life-cycle model with health capital

3.2.1 Presentation of the model

Our model is based on Grossman (1972)[\[23\]](#). We use a simplified version similar to the one developed by Strulik (2015)[\[47\]](#), Laporte (2015)[\[34\]](#) and Laporte and Ferguson (2017)[\[35\]](#) while keeping the main ingredients of Grossman’s model. Agent’s health is characterized by its health capital level and agent’s utility depends on consumption and health. The income of the agent is an increasing function of the health capital. The agent must allocate its income between consumption and investment in health capital. Finally, we assume that health capital depreciates at a rate increasing with age. The agent ceases to live if the health capital fall below a threshold.

The intertemporal utility function The representative agent maximizes its intertemporal discounted utility:

$$\sum_{t=0}^T \beta^t \left(\frac{(C_t^\gamma H_t^{1-\gamma})^{1-\sigma}}{1-\sigma} + b \right) \quad (3.1)$$

with C_t the consumption and H_t the health capital. T is the agent lifespan which is an endogenous variable. More precisely, we consider $t = 0$ corresponds to the beginning of the active life (that is 25 years) and T is the lifespan in adulthood (that is $T+25$ years).

$\beta \in]0, 1[$ is the discount factor, $\sigma > 1$ is the intertemporal elasticity of substitution and $\gamma \in]0, 1[$ is the weight of consumption in utility. It is also possible to interpret $C_t^\gamma H_t^{1-\gamma}$ as an aggregate good produced with private consumption and health capital through a constant return to scales technology.

b is a constant positive term. The key role of this parameter is widely discussed by Rosen (1988)[42] and Hall and Jones (2007)[26]. In particular, the value of b must be sufficiently high to ensure a positive flow of utility. We'll discuss this point in Section 3.2.2.

The law of motion of health capital Health capital evolution is described by the following equation:

$$H_t = (1 - \delta_t)H_{t-1} + AI_t \quad (3.2)$$

I_t represents the investment in health capital. Health capital is produced through a linear technology, A being the efficiency of this technology. Thus, at each date, the health capital is increased by an amount AI_t .

Health capital depreciates at a rate δ_t increasing with the age of the agent. Following Fonseca *et al.* (2008)[18], we guess that δ_t could be written as follows:

$$\delta_t = \delta_1 \exp(\delta_2 t)$$

with $\delta_1 > 0$ and $\delta_2 > 0$. δ_1 represents the health depreciation at the beginning of the adulthood.

Death occurs when health capital falls below the threshold $\underline{H} > 0$. The agent is still alive at period t if $H_t \geq \underline{H}$.

The aggregate resource constraint There is a unique good produced by the household which is consumed and invested. The good is produced by the household through a decreasing return to scales technology. Health capital H_t is the unique input. The household's production can also be interpreted as an income in real terms. The agent can improve its health by reducing consumption and increasing investment in health capital.

At each period t , one has:

$$\theta H_t^\alpha - I_t - C_t = 0 \quad (3.3)$$

where $\theta > 0$ denotes the efficiency of the production technology. Having assuming a decreasing return to scales technology, one has $\alpha \in]0, 1[$.

Finally, we impose the following positivity condition on investment, $I_t \geq 0$, $\forall t = 0, 1, \dots, T$. This condition is natural since health capital can not be “uninstalled” and sold.

Household's programme The initial health capital of the agent is $H_{t-1} > \underline{H}$. The agent chooses the optimal path of consumption and investment maximizing his lifetime utility subjected to resource constraints. First, let us assume a given lifespan T . The household's program is therefore written as:

$$\max \sum_{t=0}^T \beta^t \left(\frac{(C_t^\gamma H_t^{1-\gamma})^{1-\sigma}}{1-\sigma} + b \right) \quad (3.4)$$

subject to

$$\left\{ \begin{array}{ll} -H_t + (1 - \delta_t)H_{t-1} + AI_t = 0 & (q_t) \\ \theta H_t^\alpha - C_t - I_t = 0 & (\mu_t) \\ I_t \geq 0 & (\eta_t) \\ H_t - \underline{H} \geq 0 & (\xi_t) \end{array} \right. \quad (3.5)$$

The optimality conditions The Lagrangian associated to the above program is written as:

$$\begin{aligned}
\mathcal{L} &= \sum_{t=0}^T \beta^t \left(\frac{(C_t^\gamma H_t^{1-\gamma})^{1-\sigma}}{1-\sigma} + b \right) \\
&+ \sum_{t=0}^T \beta^t q_t (-H_t + (1-\delta_t)H_{t-1} + AI_t) \\
&+ \sum_{t=0}^T \beta^t \mu_t (\theta H_t^\alpha - I_t - C_t) + \sum_{t=0}^T \beta^t \eta_t I_t + \sum_{t=0}^T \beta^t \xi_t (H_t - \underline{H}) \\
&= \sum_{t=0}^T \beta^t \left(\frac{(C_t^\gamma H_t^{1-\gamma})^{1-\sigma}}{1-\sigma} + b \right) \\
&+ \sum_{t=0}^T \beta^t q_t (-H_t + AI_t) + \sum_{t=-1}^{T-1} \beta^{t+1} q_{t+1} (1-\delta_{t+1}) H_t \\
&+ \sum_{t=0}^T \beta^t \mu_t (\theta H_t^\alpha - I_t - C_t) + \sum_{t=0}^T \beta^t \eta_t I_t + \sum_{t=0}^T \beta^t \xi_t (H_t - \underline{H})
\end{aligned}$$

The following optimality conditions are obtained:

$$\begin{aligned}
\frac{\partial \mathcal{L}}{\partial C_t} &= \beta^t \gamma \frac{(C_t^\gamma H_t^{1-\gamma})^{1-\sigma}}{C_t} - \beta^t \mu_t = 0, \quad t = 0, \dots, T \\
\frac{\partial \mathcal{L}}{\partial I_t} &= \beta^t q_t A - \beta^t \mu_t + \beta^t \eta_t = 0, \quad t = 0, \dots, T \\
\frac{\partial \mathcal{L}}{\partial H_t} &= \beta^t \left[(1-\gamma) \frac{(C_t^\gamma H_t^{1-\gamma})^{1-\sigma}}{H_t} + \mu_t \theta \alpha H_t^{\alpha-1} + \xi_t - q_t \right] + \beta^{t+1} q_{t+1} (1-\delta_{t+1}) = 0, \\
&t = 0, \dots, T-1 \\
\frac{\partial \mathcal{L}}{\partial H_T} &= \beta^T \left[(1-\gamma) \frac{(C_T^\gamma H_T^{1-\gamma})^{1-\sigma}}{H_T} + \mu_T \theta \alpha H_T^{\alpha-1} + \xi_T - q_T \right] = 0 \\
&t = T
\end{aligned}$$

Otherwise, one has the following exclusion conditions:

$$\begin{aligned}
\eta_t I_t &= 0, t = 0, 1, \dots, T, \text{ with } \eta_t \geq 0 \\
\xi_t (H_t - \underline{H}) &= 0, t = 0, 1, \dots, T-1, \text{ with } \xi_t \geq 0
\end{aligned}$$

The system of equations that characterizes the solution The initial health capital

H_{-1} is supposed to be given. The unknown variables are:

$$C_t, \quad t = 0, \dots, T$$

$$H_t, \quad t = 0, \dots, T$$

$$I_t, \quad t = 0, \dots, T$$

$$q_t, \quad t = 0, \dots, T$$

$$\mu_t, \quad t = 0, \dots, T$$

$$\eta_t, \quad t = 0, \dots, T$$

$$\xi_t, \quad t = 0, \dots, T$$

The system to be solved is defined by the following set of equations:

$$\gamma \frac{(C_t^\gamma H_t^{1-\gamma})^{1-\sigma}}{C_t} - \mu_t = 0, \quad t = 0, \dots, T$$

$$q_t A - \mu_t + \eta_t = 0, \quad t = 0, \dots, T$$

$$(1 - \gamma) \frac{(C_t^\gamma H_t^{1-\gamma})^{1-\sigma}}{H_t} + \mu_t \theta \alpha H_t^{\alpha-1} + \xi_t - q_t + \beta q_{t+1} (1 - \delta_{t+1}) = 0, \quad t = 0, \dots, T - 1$$

$$(1 - \gamma) \frac{(C_T^\gamma H_T^{1-\gamma})^{1-\sigma}}{H_T} + \mu_T \theta \alpha H_T^{\alpha-1} + \xi_T - q_T = 0,$$

$$\eta_t I_t = 0, \quad t = 0, \dots, T \text{ with } \eta_t \geq 0$$

$$\xi_t (H_t - \underline{H}) = 0, \quad t = 0, \dots, T \text{ with } \xi_t \geq 0$$

$$-H_t + (1 - \delta_t) H_{t-1} + A I_t = 0, \quad t = 0, \dots, T$$

$$\theta H_t^\alpha - I_t - C_t = 0, \quad t = 0, \dots, T$$

There are $(T + 1) \times 7$ equations and $(T + 1) \times 7$ unknown variables. The system of equations is thus “well-posed” and we can undertake to solve it numerically.

Guess about the characterizations of the solution As previously mentioned, the system of equations can not be solved analytically. The system is solved numerically using a standard Newton algorithm. The problem includes two positivity constraints and we have to determine when these constraints are binding or not. This is done numerically.

If the path followed by the variables can only be determined numerically, it is however possible to provide some intuitions about the optimal path of the variables. The optimal solution may reasonably take two forms (T being taken as given):

1. Let suppose that $H_t > \underline{H}$, $\forall t = 0, \dots, T$. If the health capital is strictly greater than \underline{H} at the last period of life, investment may be equal to zero during the last periods.

To sum up:

$$I_t > 0 \text{ and } \eta_t = 0, \forall t = 0, \dots, \bar{T} - 1$$

$$I_t = 0 \text{ and } \eta_t \geq 0, \forall t = \bar{T}, \dots, T$$

$$H_t \geq \underline{H} \text{ and } \xi_t = 0, \forall t = 0, \dots, T$$

This case may occurs if the initial health capital H_{-1} is sufficiently high or the health capital production sufficiently efficient.

2. Let suppose that $H_t > \underline{H}$, $\forall t = 0, \dots, T - 1$ and $H_T = \underline{H}$.

$$I_t > 0 \text{ and } \eta_t = 0, \forall t = 0, \dots, T$$

$$H_t > \underline{H} \text{ and } \xi_t = 0, \forall t = 0, \dots, T - 1$$

$$H_T = \underline{H} \text{ and } \xi_T \geq 0$$

In this case, we guess the investment constraint is never bound. Indeed, investment will necessarily be strictly positive during the last periods of life to ensure that

health capital is not below the threshold \underline{H} . Intuitively, this case may occur if the initial health capital is low.

Optimal lifetime We previously considered that the lifetime T was given. In our setting, T must be chosen in such a way it maximizes the agent's intertemporal utility. The optimal value of T may be determined in a simple way. For all the possible values of T , solve (numerically) the above problem and compute the value of the intertemporal utility, that is:

$$\sum_{t=0}^T \beta^t \left(\frac{(C_t^\gamma H_t^{1-\gamma})^{1-\sigma}}{1-\sigma} + b \right)$$

Finally, choose the value T^* that maximizes the intertemporal utility.

But is there a more sophisticated method to determine the optimal value T^* ? In a continuous time framework, the question is well documented and optimality conditions can easily be found to determine an optimal lifespan. Yet, in a discrete time framework, to our knowledge, there is a lack of references and the derivation and the writing of the optimality conditions is not immediate¹.

Here we propose an optimality condition obtained through an adaptation of the continuous time case.

We define the following Hamiltonian:

$$\mathcal{H}(T) = \left(\frac{(C_T^\gamma H_T^{1-\gamma})^{1-\sigma}}{1-\sigma} + b \right) + q_T(-\delta_T H_{T-1} + I_T)$$

In continuous time, the optimal time satisfies² $\mathcal{H}(T^*) = 0$. This condition has to be adapted if time is discrete because problems of “indivisibility” occurs.

¹In a continuous time problem, the optimality condition giving the optimal time is obtained by deriving the intertemporal utility with respect to T and applying the envelop theorem.

²This condition can be found in any standard optimal control book.

Now, assume that for $T = T^+$, one has:

$$\begin{aligned}\mathcal{H}(T^+) &= \left(\frac{(C_{T^+}^\gamma H_{T^+}^{1-\gamma})^{1-\sigma}}{1-\sigma} + b \right) + q_{T^+}(-\delta_{T^+} H_{T^+-1} + I_{T^+}) > 0 \\ \mathcal{H}(T^+ + 1) &= \left(\frac{(C_{T^++1}^\gamma H_{T^++1}^{1-\gamma})^{1-\sigma}}{1-\sigma} + b \right) + q_{T^++1}(-\delta_{T^++1} H_{T^+} + I_{T^++1}) < 0\end{aligned}$$

We can conclude that:

$$\begin{aligned}T^* = T^+ &\quad \text{if} \quad |\mathcal{H}(T^+)| < |\mathcal{H}(T^+ + 1)| \\ T^* = T^+ + 1 &\quad \text{if} \quad |\mathcal{H}(T^+)| > |\mathcal{H}(T^+ + 1)|\end{aligned}$$

We previously discussed the characteristics of the optimal path for a given value of T . What can be said about the optimal path of health capital if the life time is also optimal (T^*)? The agent is alive as long as $H_t \geq \underline{H}$. If at date T , one has $H_T > \underline{H}$ and $(1 - \delta_T)H_T \geq \underline{H}$, the agent can live one more period without investing in health capital. Clearly, T is thus not an optimal life time. We conclude at the optimal time T^* , one has $H_{T^*} = \underline{H}$.

The above optimality condition is not necessary to solve numerically the problem. However, it is useful to provide some economic interpretations concerning the determinants of the optimal lifetime.

3.2.2 The constant term in the utility function

Here we discuss the meaning of the term b in the utility function. Rosen (1988)[\[12\]](#) is interested by the question of the value of life and develops theoretical frameworks allowing to highlight its determinants. Two classes of models are considered, models with a survival rate and deterministic life-cycle models. Note that this two classes of

models are closely related and have very similar results. In the line of Rosen (1988)[\[42\]](#), we use a deterministic life-cycle model in which the lifetime is endogenously determined. In models like ours, the constant term b of the utility function plays an important and “non-standard” role because it has implications on the decisions of the agents. This point is discussed in detail by Rosen (1988)[\[42\]](#) and Hall and Jones (2007)[\[26\]](#). In standard microeconomic theory, it is always possible to consider an increasing transformation of a utility function. This transformation does not impact the marginal utility ratios and thus the optimal decisions. Our model works very differently than standard ones. Indeed, in our framework, as in Rosen[\[42\]](#)’s, the choice problem is a very particular one. The agent has to determine the optimal consumption and health investment plans. These decisions only depend on the marginal utility and the constant b does not impact the optimality conditions. Our model is also a life-cycle one and the agent has to determine his optimal lifetime T^* . The associated optimality condition (see subsection [3.2.1](#)) significantly differs from the other optimality conditions and the level of utility now matters. It is thus easy to understand that the value of b will significantly alter agent decisions (through the value of the optimal lifetime). Furthermore, the optimization problem determining the optimal lifetime has a solution if b is positive and sufficiently large.

The life-cycle utility We follow Rosen (1988)[\[42\]](#) who explains how this utility function may be set and how the constant b may be understood.

The lifetime is T . The instantaneous utility of the agent (during his lifetime) is:

$$u(C_t, H_t) = \frac{(C_t^\gamma H_t^{1-\gamma})^{1-\sigma}}{1-\sigma}$$

The agent utility over its life-cycle is given by:

$$\sum_{t=0}^T \beta^t u(C_t, H_t)$$

However, the above expression does not provide a complete evaluation of the agent utility over his life-cycle. Indeed, we suppose that the agent incurs a disutility of being death, the instantaneous disutility being $-b$. Evaluated at $t = 0$, the disutility of being death at $t > T$ is given by $\beta^t(-b)$. To sum up, the life-cycle utility of the agent is written as:

$$\sum_{t=0}^T \beta^t u(C_t, H_t) + \sum_{t=T+1}^{\infty} \beta^t (-b)$$

Rearranging the above expression, one gets:

$$\sum_{t=0}^T \beta^t u(C_t, H_t) + \sum_{t=T+1}^{\infty} \beta^t (-b) = \sum_{t=0}^T \beta^t (u(C_t, H_t) + b) - \frac{b}{1-\beta}$$

The term $\frac{b}{1-\beta}$ being independent of T may be dropped. We thus obtain the intertemporal utility function of subsection [3.2.1](#), that is:

$$\sum_{t=0}^T \beta^t (u(C_t, H_t) + b)$$

Optimal longevity The optimality condition giving the optimal lifetime T^* is given in Subsection [3.2.1](#). To simplify the analysis, assume that T^* is such that one exactly has $\mathcal{H}(T^*) = 0$. Note that it is the standard optimality condition obtained in the continuous time equivalent model. At subsection [3.2.1](#), we just use the discrete time counterpart.

$$\mathcal{H}(T^*) = \frac{(C_{T^*}^\gamma H_{T^*}^{1-\gamma})^{1-\sigma}}{1-\sigma} + b + q_{T^*}(-\delta_{T^*} H_{T^*-1} + I_{T^*}) = 0$$

Evaluated at $t = 0$, the above condition can be rewritten as follows:

$$\beta^{T^*} \mathcal{H}(T^*) = \beta^{T^*} [u(C_{T^*}, H_{T^*}) + b] + \beta^{T^*} q_{T^*}(-\delta_{T^*} H_{T^*-1} + I_{T^*}) = 0$$

or

$$\begin{aligned}
\beta^{T^*} \mathcal{H}(T^*) &= \beta^{T^*} [u(C_{T^*}, H_{T^*}) + b] + \beta^{T^*} q_{T^*} (H_{T^*} - H_{T^*-1}) \\
&= \beta^{T^*} [u(C_{T^*}, H_{T^*}) + b] + \beta^{T^*} u_C(C_{T^*}, H_{T^*}) \frac{1}{A} (H_{T^*} - H_{T^*-1}) \quad (3.6) \\
&= 0
\end{aligned}$$

It is possible to show that $\beta^{T^*} \mathcal{H}(T^*)$ is the net gain of being alive one more period³. If the agent stays alive one period more, his gain in utility is $\beta^{T^*} [u(C_{T^*}, H_{T^*}) + b]$, the sum of the utility derived from consumption and health capital and of the utility of being alive (b). However, in order to ensure the supplementary consumption (and maintain the health capital above or equal to the threshold \underline{H}) induced by the increase of the lifespan, the agent must reduce its consumption over his life-cycle. The cost, in terms of utility, is given by $\beta^{T^*} u_C(C_{T^*}, H_{T^*}) \frac{1}{A} (H_{T^*} - H_{T^*-1}) = \beta^{T^*} q_{T^*} (H_{T^*} - H_{T^*-1})$. We underline that $H_{T^*} - H_{T^*-1}$ is necessarily negative. At the end of his life, the health capital of the agent is decreasing, so that $\beta^{T^*} q_{T^*} (H_{T^*} - H_{T^*-1}) < 0$. At the optimum, the net gain of being alive one more period is equal to zero.

This optimality condition is written in terms of utility level, and thus, the constant term b plays a key role in the tradeoffs made by the agent. To underline the role played by b assume that $b = 0$. With our specification of the utility function and assuming $\sigma > 1$ (see Hall and Jones (2007)^[26]), the condition ^(3.7) becomes:

$$\frac{1}{1 - \sigma} + \frac{\gamma}{C_{T^*}} \frac{1}{A} (H_{T^*} - H_{T^*-1}) = 0$$

All the terms of the above expression are negative. It follows the optimality condition that

³In the continuous time model, this can be shown by differentiating the agent welfare with respect to T along an optimal path.

gives the optimal lifespan works only if $b > 0$ is sufficiently large to ensure a strictly positive utility. In other words, the parameter b is crucial to obtain a well-posed optimization problem.

The role of parameter b Here we discuss the role of parameter b and how it determines the agent's choices.

As previously mentioned, b is the utility of being alive. An increase in b will provide incentives to increase lifetime for a given amount of resources. Given the resources constraints, the agent should then reduce his consumptions. We consider a simplified version of Rosen (1988)[\[42\]](#)'s model. The agent is endowed with wealth W and the interest rate is equal to the time preference rate. We thus use β to compute the sum of discounted consumption flows in the budget constraint. We also impose that consumption is constant over time. One has:

$$\sum_{t=0}^T \beta^t \left[\frac{C^{1-\sigma}}{1-\sigma} + b \right] = \frac{1 - \beta^{T+1}}{1 - \beta} \left[\frac{C^{1-\sigma}}{1-\sigma} + b \right]$$

$$W = \sum_{t=0}^T \beta^t C = \frac{1 - \beta^{T+1}}{1 - \beta} C$$

Determine now the optimality conditions with respect to C and T , that is:

$$\frac{1 - \beta^{T+1}}{1 - \beta} C^{-\sigma} - \lambda \frac{1 - \beta^{T+1}}{1 - \beta} = 0$$

$$-\frac{\log(\beta)\beta^{T+1}}{1 - \beta} \left[\frac{C^{1-\sigma}}{1-\sigma} + b \right] + \lambda \frac{\log(\beta)\beta^{T+1}}{1 - \beta} C = 0$$

λ being the Lagrange multiplier.

The optimal consumption is given by:

$$C = \left(\frac{\sigma - 1}{\sigma} \right)^{\frac{1}{1-\sigma}} b^{\frac{1}{1-\sigma}}$$

Note that C exists only if $b > 0$ and $\sigma > 1$. Let's differentiate C with respect to b :

$$\frac{\partial C}{\partial b} = \left(\frac{\sigma - 1}{\sigma} \right)^{\frac{1}{1-\sigma}} \frac{1}{1-\sigma} b^{\frac{1}{1-\sigma}-1} < 0$$

The optimal value of T is given by:

$$W = \frac{1 - \beta^{T+1}}{1 - \beta} \left(\frac{\sigma - 1}{\sigma} \right)^{\frac{1}{1-\sigma}} b^{\frac{1}{1-\sigma}}$$

It is easy to show that $\frac{\partial T}{\partial b} > 0$. This example shows clearly the role of parameter b and what are the tradeoffs made by the agent. An increase in b correspond to an increase in the utility of being alive. If b increases, the agent chooses to live longer and, knowing that its wealth is given, reduces his instantaneous consumption.

To finish, we discuss what happens if the agent can invest in health capital. We use the model with a survival rate of Hall and Jones (2005), which allows to have analytical results. The agent's problem writes:

$$\max p(H)U(C)$$

$$\text{s.t. } Y - C - H \geq 0$$

$p(H) = H^\omega$, with $\omega < 1$ is the survival rate and Y is the agent's income.

We deduce the following optimality conditions:

$$\begin{aligned} p(H)U'(C) - \lambda &= 0 \\ p'(H)U(C) - \lambda &= 0 \end{aligned}$$

Note that the level of utility must be positive, otherwise the second optimality condition is not consistent.

Let's define the elasticities $\xi_H = H \frac{p'(H)}{p(H)}$ and $\xi_C = C \frac{U'(C)}{U(C)}$. From the optimality conditions, we deduce:

$$\frac{C}{H} = \frac{\xi_C}{\xi_H} = \frac{\xi_C}{\omega}$$

Assuming that $\frac{\partial C}{\partial Y} > 0$, we determine the effect of an increase in Y on the ratio $\frac{C}{H}$. One gets:

$$\frac{\partial}{\partial Y} \left(\frac{C}{H} \right) = \frac{1}{\omega} \frac{\partial \xi_C}{\partial C} \frac{\partial C}{\partial Y}$$

The sign of the above expression depends on the sign of the derivative of the elasticity ξ_C . If this elasticity is decreasing in consumption, then, the ratio $\frac{C}{H}$ is decreasing in Y . It follows that the share of health expenditures increases as the income Y increases.

Suppose now that $U(C) = \frac{C^{1-\sigma}}{1-\sigma} + b$. One has:

$$\frac{\partial \xi_C}{\partial C} = \frac{(1-\sigma)bC^{1-\sigma}}{\left(\frac{C^{1-\sigma}}{1-\sigma} + b\right)^2}$$

Knowing that $\sigma > 1$, the above derivative is negative. It is also easy to check that if $b = 0$, $\xi_C = 1 - \sigma$. It follows that $\frac{C}{H} = \frac{1-\sigma}{\omega} < 0$ and $\frac{\partial \xi_C}{\partial C} = 0$. Once again, we can underline that the importance of parameter b which must be large enough to ensure the existence of a

solution to the agent's problem⁴.

To conclude this paragraph about parameter b . Assuming that $\sigma > 1$, we saw, through several examples, that b must be positive and sufficiently large to ensure the consistency of a life-cycle model with endogenous life-time (or endogenous survival rate). Furthermore, $b > 0$ ensures that the consumption-elasticity of utility is decreasing. It follows that the share of health expenditures in income increases as income increases. This is consistent with the data providing evidences that the share of income devoted to health expenditures increases as income increases.

3.3 Numerical investigations

3.3.1 Calibration strategy

The model is calibrated in order to match US data. The discount factor β is set to 0.96, which correspond to an annual interest rate of 4%. The risk aversion coefficient take the traditional value of 1.5. As previously discussed, σ must be greater than 1. Parameter α is the elasticity of income with respect to the health stock. We choose $\alpha = 0.15$, which is of the same order as the value of 0.146 estimated by Fonseca *et al.* (2008)^[18]. Parameter γ (resp. $1 - \gamma$) corresponds to the weight of private consumption (resp. health stock) in utility. In a static model, this parameter, together with α , determines the share of consumption (and health spending) in income⁵. In USA, this share is roughly of 17% and in France, it is of 11%. Given the retained value of α , a value of $\gamma = 0.95$ would be

⁴If $b = 0$, the problem has a solution only if $\sigma \in]0, 1[$ and the ratio is constant. The health spending share is then independent of income.

⁵Consider the simplified model:

$$\begin{aligned} & \max_{C, I, H} \frac{(C^\gamma H^{1-\gamma})^{1-\sigma}}{1-\sigma} \\ & \text{s.t.} \begin{cases} \theta H^\alpha - C - I = 0 & (\lambda) \\ AI - H = 0 & (\mu) \end{cases} \end{aligned}$$

necessary to obtain (in the static model) a health spending share of 19%. There is no clear evidences in the literature concerning the numerical value of this parameter. For instance, Fonseca *et al.* (2008)[18] set $\gamma = 1$ whereas Halliday, He, Ning and Zhang (2017)[27] chose $\gamma = 0.64$. Parameter γ is set at 0.98, which allows to obtain a health spending share close to 13 – 14%. The parameters A and θ are normalized to 1.

Fonseca *et al.* (2008)[18] assume the health stock is between 0 and 100, in other words they choose a 0-100 scale. They also report that the health stock decreases from 73.3 at age 26 to 39.7 at age 85. We calibrate the model in order to satisfied the same proportions and we impose $\frac{H}{H_{-1}} = \frac{39.7}{73.3}$. The level of these variables is set in order to have an health investment equal to 0 during the first periods (as it is the case in Halliday, He, Ning and Zhang (2017)[27] and Fonseca *et al.* (2008)[18]). One has $H_{-1} = 4.1441$ and $H = 2.2140$. Finally, following Fonseca *et al.* (2008)[18], the parameters describing the depreciation of health capital take the values $\delta_1 = 0.035$ and $\delta_2 = 0.025$. We impose $b = 2.24$, so that the agent lifetime is of around 85 years.

The benchmark calibration is summarized in Table 3.1.

β	b	σ	γ	α	θ	A	δ_1	δ_2	H_{-1}	H
0.96	2.24	1.5	0.98	0.15	1	1	0.035	0.025	4.1441	2.2140

Table 3.1: Benchmark calibration

3.3.2 Numerical results

The system of equations characterizing the equilibrium is solved numerically by the Gauss-Newton algorithm.

Is is easy to determine that:

$$\frac{I}{C+I} = 1 - \gamma + \gamma\alpha$$

This benchmark calibration gives an optimal T of 59, which corresponds to a lifespan of 85 years. The share of health spending ($\frac{I}{C+I}$), that is its average value calculated over the life-cycle, is of 13.91%. French *et al.* (2017) report that the medical spending of the last three years represents 16.70% of the US aggregate medical spending. In the other countries they study, this value is around 20%. As we consider a representative agent over his life-cycle, we calculate the share of the last 3 years health spending in total health spending. We find a value of about 10.5%. This value is less than the one reported by French *et al.* (2017). However, health spending during the last 3 years of life have a significant level.

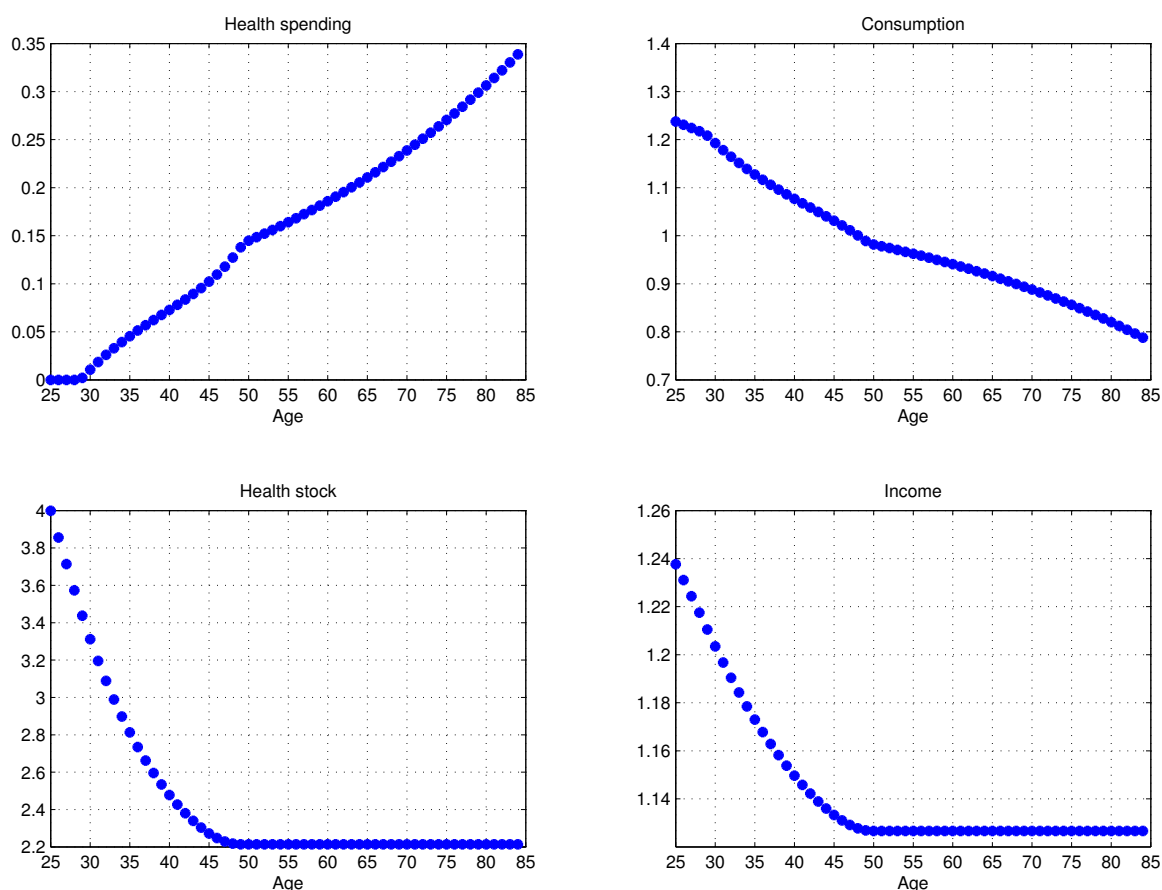


Figure 3.1: Evolution of main variables

Figure 3.1 presents the evolution of health spending, consumption, health stock and income over the life-cycle. The health spending is zero around the age of 25. Thereafter,

the profile of health expenditures is increasing, which is roughly consistent with the data (see Halliday, He, Ning and Zhang (2017)[\[27\]](#) and Fonseca *et al.* (2008)[\[18\]](#)). The health stock is decreasing. However, around the age of 50, the health stock attains its critical threshold \underline{H} . Health spending allows to maintain health stock at this critical level until the end of life⁶. The income follows the same trend as the health stock. Finally, consumption is decreasing. This is obviously due to the decreasing of income and the increasing of health spending. However, in this class of models, agents wish to smooth their consumption over their life-cycle. The wish to smooth consumption tends to limit the decreasing trend of consumption.

Our model is able to reproduce an increasing profile of health spending consistent with the data. On the other hand, our model does not reproduce the standard hump shaped profile of income reported in life-cycle literature. Likewise, consumption profile is less smooth than in a standard life-cycle model. In our framework, there is no financial market. The agent cannot transfer income over time and consumption cannot be smoothed. Concerning the income, our model looks like life-cycle models with human capital and endogenous wage (see Ben-Porath (1967)[\[5\]](#)). This class of model is able to generate an hump shaped income profile. In our model, the health stock variable plays a similar role to human capital. We point out two differences. Firstly, we guess the depreciation rate increases with age. Secondly, the elasticity of income with respect to health capital is, in our benchmark calibration, equal to 0.15. This value is too small to generate an hump shaped profile of income. A value of around 0.8 would be necessary to obtain it. Income is less sensitive to health than to education.

⁶Following Halliday, He, Ning and Zhang (2017)[\[27\]](#) and Fonseca *et al.* (2008)[\[18\]](#), the health index is decreasing over the life-cycle. Our results are consistent with their data, except that in our model (and with our benchmark calibration, the health stock is staying at its critical threshold for about 35 years.

3.3.3 Sensitivity analysis

In this subsection, we examine how the numerical results are modified if we change some key parameters.

The elasticity σ We simulated the model for two other value of σ , that is $\sigma \in \{1.5; 3; 5\}$ (Figure 3.2). Recall that $\frac{1}{\sigma}$ is the intertemporal elasticity of substitution. This variable ex-

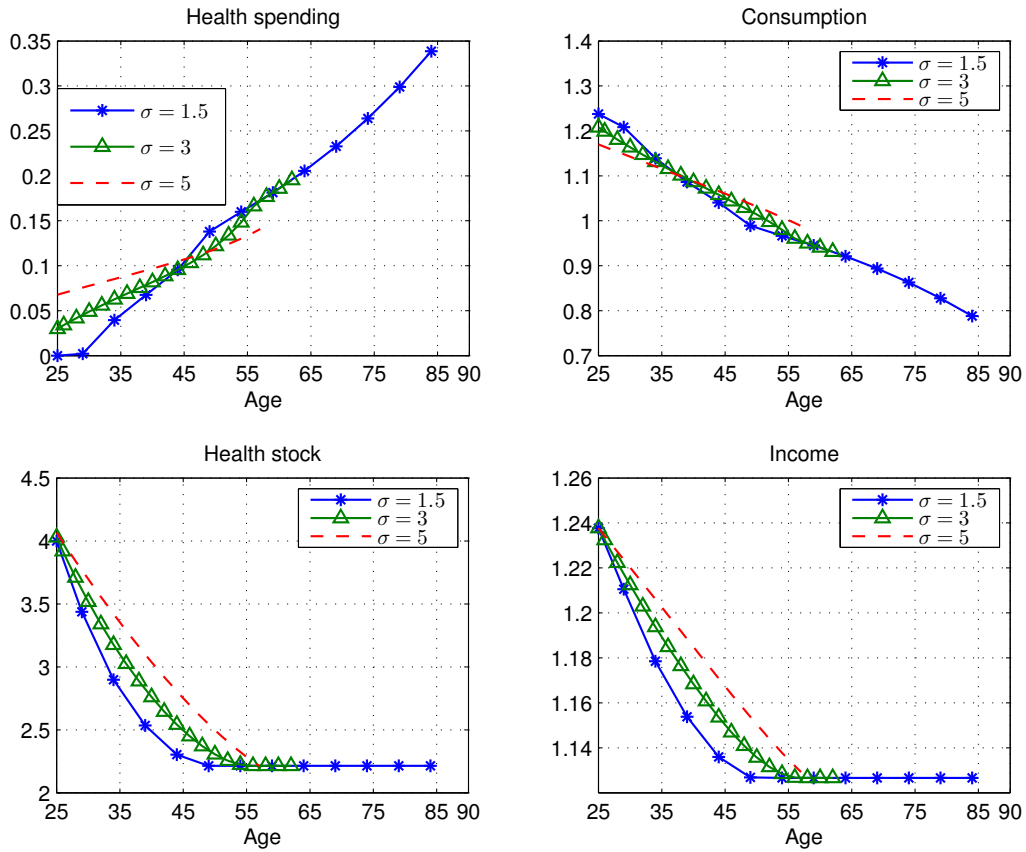


Figure 3.2: Elasticity σ

presses the will of the agent to substitute intertemporally consumption (or more precisely, the “composite good” $C^\gamma H^{1-\gamma}$). Thus, increasing σ means a reduction of the willingness of the agent to make intertemporal substitutions. In other word, the agent prefers a smoother consumption profile. As we want to evaluate the impact of a perturbation of the elasticity σ on the trade-off made by the agent, the numerator of the utility function

$1 - \sigma$ is always parametrized using the benchmark value of 1.5. Otherwise, a perturbation of σ would have effects similar to a change in b .

σ	1.5	3	5
Lifespan	85	63	58
Health spending share	13.91%	8.96%	8.56%

Table 3.2: Elasticity σ

Increasing parameter σ leads to a decrease in the lifespan and in the health spending share (Table 3.2). This last point is obviously related to the lifespan decrease. A lower lifespan combined with a lower health stock depreciation mechanically leads to a reduction in health spending.

Figure 3.2 presents the paths followed by health spending, consumption, health stock and income for the three values of σ . The increase in σ significantly alters the tradeoffs made by the agent. The path followed by consumption is smoother as σ increases. A reduction in the lifespan follows.

The parameter b We consider two alternative values for parameter b , that is $b \in \{2; 2.24; 2.5\}$ (Figure 3.3). As previously discussed, this parameter plays a key role in the determination of the lifespan. The results shown in Table 3.3 point that the lifespan is increasing relative to b . This quantitative exercise points out that a small perturbation of parameter b significantly changes the lifespan of the agent. We also observe an increase in the health spending share. This is obviously related to the lifespan increase and the increase in the health stock depreciation rate. To live longer, the agent must increase his health spending share as he gets holder. It follows an increase in the health spending share.

Figure 3.3 presents the paths followed by health spending, consumption, health stock and income for the three values of b . Surprisingly, for each variables, the paths seem to be superimposed. In other words, the path of a given value of b approximatively extends

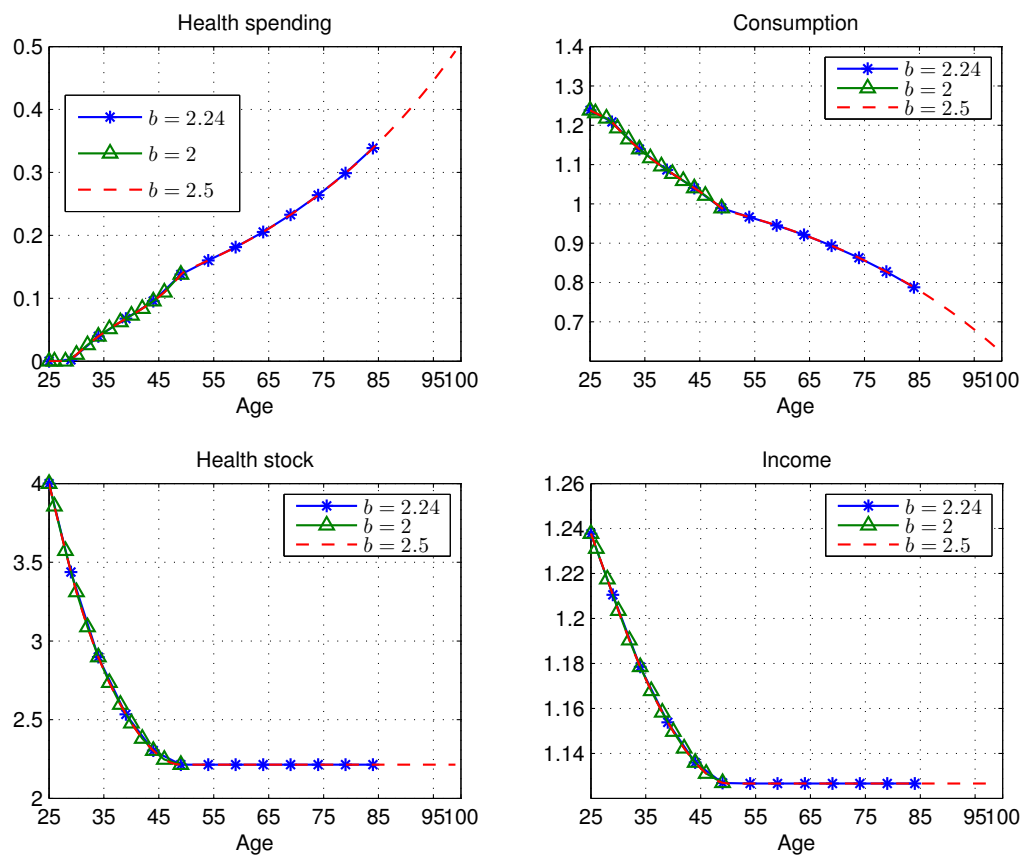


Figure 3.3: Parameter b

b	2	2.24	2.5
Lifespan	50	85	100
Health spending share	4.99%	13.91%	18.52%

Table 3.3: Parameter b

the one obtained with a smaller value. An increase in b corresponds to an increase in the utility of being alive. Consequently, the agent is willing to accept a reduction in consumption to finance health spending in order to stay alive.

The parameter γ We suppose that $\gamma \in \{0.9; \mathbf{0.98}; 1\}$. Recall this parameter has a direct impact on the behaviour of the agent in regards to health spending.

The more interesting case is $\gamma = 0.9$. Under this value of γ , the agent chooses to devote a greater share of his expenses to health spending (Table 3.4). The health spending is then strictly positive from the age of 25 years (Figure 3.4). Recall that with higher values of γ , there is no health spending around 25 years. The paths followed by the variables is significantly different from those of the benchmark calibration. During the first years, consumption is lower with a smoother profile. It follows that the agent invests in health from 25 years and health stock and income are improved.

γ	0.9	0.98	1
Lifespan	93	85	83
Health spending share	17.92%	13.91%	13.26%

Table 3.4: Parameter γ

The parameter α We suppose that $\alpha \in \{0.12; \mathbf{0.15}; 0.18\}$. This parameter corresponds to the elasticity of income with respect to health stock. An increase in α obviously means an improvement of the technology linking health stock and income.

Results of Figure 3.5 show that as α increases, the income path significantly moves upwards. Consumption and health spending do the same. However, the lifespan and the health spending share increase moderately as α increases. Thus, an increase in α essentially modify the level of consumption and health spending.

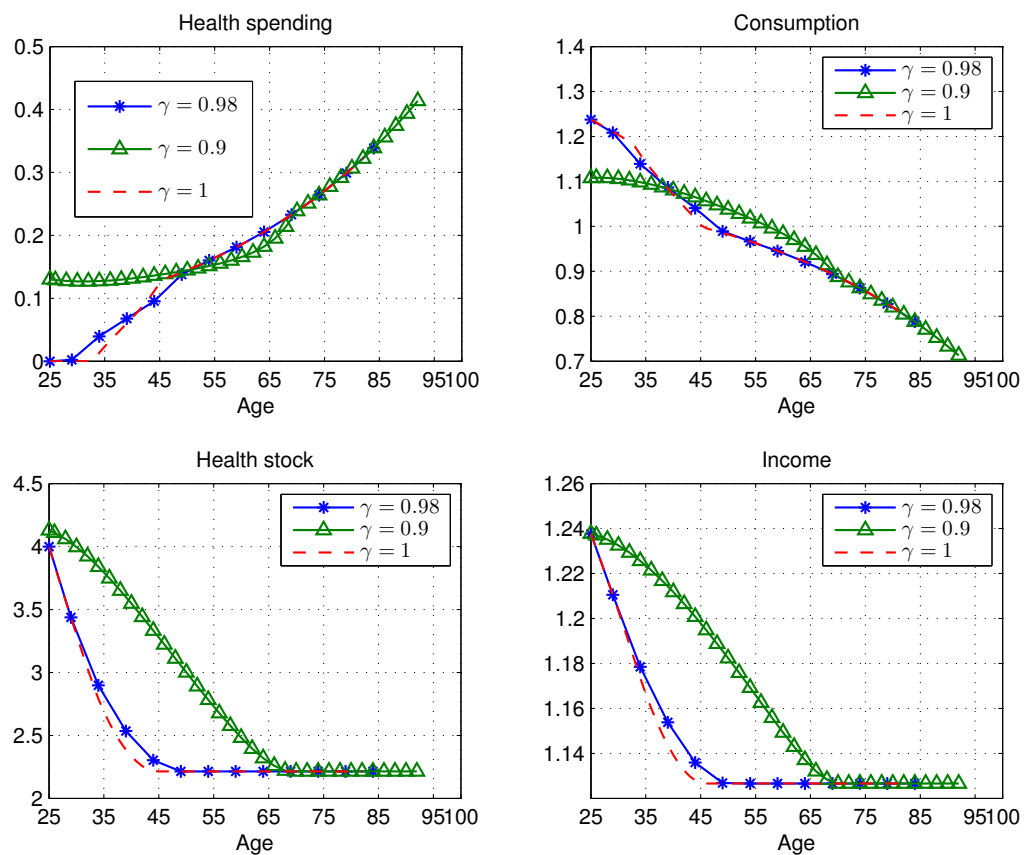


Figure 3.4: Parameter γ

α	0.12	0.15	0.18
Lifespan	82	85	88
Health spending share	13.27%	13.91%	14.96%

Table 3.5: Parameter α

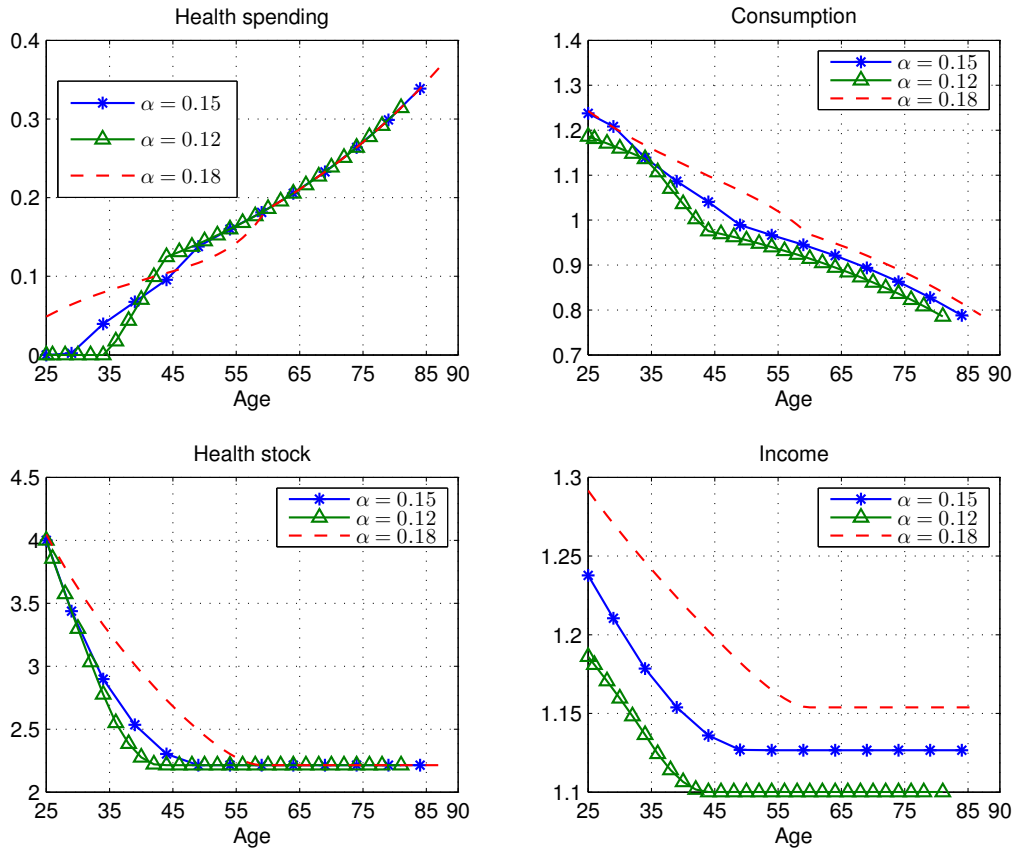


Figure 3.5: Parameter α

3.4 Conclusion

In this article, we study a simplified version of the Grossman[23]’s model. We mainly follow Laporte (2015)[34] and Strulik (2015)[47]. The agent can allocate his income between consumption and health spending. The health of the agent is characterized by his health stock or health capital. The depreciation of the health stock may be compensated by health spending. The agent stays alive as long as the health stock is above a threshold. One key assumption concerns the depreciation rate of the health stock. Contrarily to Laporte (2015)[34] and Strulik (2015)[47], we suppose the depreciation rate increases with age. Finally, we point out that there is no financial market.

We derive the optimality conditions that allow to characterize the solution. The optimal control problem presents some difficulties. There are two occasionally binding

constraints (health spending is positive or zero and the health stock is greater than or equal to the critical threshold) and the depreciation rate of health stock increases with age. However, the optimality conditions are obtained with an endogenous lifespan, thus the problem of potential immortality is avoided.

We are able to characterize numerically the solution of the model. The model is calibrated on the US economy. The model is able to reproduce an increasing profile of health spending over the life-cycle and a significant share of health spending especially at the end of life. The profile of some other important variables do not replicate their empirical counterpart. Likewise, the health stock attains the critical threshold before the age of 50 and stays at this level until the age of 85. This is due to the assumptions of this simplified model. For instance, there is no financial market. The agent can not transfer income through time and cannot smooth its consumption path. Further research should include a financial market or an hump-shaped profil of income over the life-cycle.

Chapter 4

A comparison between the US and France based on a general equilibrium model

4.1 Introduction

During recent years, the health inequalities have become more and more a severe issue around the globe. This rise of health inequalities does not only take place in poorer countries¹, but also happens in developed countries: in Britain, the mortality rate of the richest people is about 15 per 1000 new-births, however, for the poorest people, the rate rises up to 42 per 1000 new-births (WHO (2008)[13]). In Europe, for middle-aged individuals who are in the lowest socioeconomic class, the spillover effect of their mortality risk varies from 25% up to 50%(Mackenbach, Meerding and Kunst (2007)[37]). The prevalence of long period disability for European males who are older than 80 is

¹For example, in India, among the most rich people, the mortality rate of child who are younger than 5 years old is around 50 per 1000 new-births, however, among the most poor people, this rate rises sharply to around 150 per 1000 new-births.

about 58.8% for those who only have limited education, versus 40.2% for those who receive university education (Huisman, Kunst and Mackenbach (2003) [30]).

In this paper, we deal with these health inequalities from a macroeconomic perspective. Indeed, there exists a large literature on the interactions between socioeconomic position of the agents and their health status. But, the analysis of the health insurance system and its implications on the agents wellbeing can be enlightened by general equilibrium approaches. Indeed, the share of the health expenditures is large enough to change the allocation of saving and thus the capital available for the production. On the other end, the risks on the labour market also provide incentives for selfinsurance, which can crowd-out the incentives for health insurance. In order to examine contrasting experiences, we choose to apply our model to two countries: the US where the unemployment insurance and health insurance programs are limited and France where they are generous. Our objective will be then to evaluate the impact of these public insurance system on the health expenditures and on health inequalities. Given our general equilibrium approach, we can also measure their impacts on the efficiency of the allocation.

Despite its generous social security system, France must deal with health inequality problem. Even if the health status has been improved for a long time, the gap between different socioeconomic classes has yet enlarged: between 1968 and 1996, the death rate for people who don't have a diploma rests unchanged while for those who have an advanced diploma, it has continuously been improving. This trend continues to our days as the health inequality continues increasing, because the gap between employees and unemployed people increases sharply (Louise Potvin (2010) [41]). More precisely, a 35 years old unemployed male expects to live 28.5 years more while a 35 years-old employee can expect to live 40 years more. Furthermore, the survey ESPS [47] also reports that 15.8% of its interviewees declare that they give up spending on health for financial reasons. These health inequalities occur despite the increase of the weight of Consumption of Health and

Medical Goods (CHMG) in GDP (from 11.1% in 2006 to 12% in 2014). The percentage of unhealthy people is slightly higher in America than in France. But the disparities in health is way larger in America: if we compare the top socioeconomic class with the bottom socioeconomic class, after adjustments, in America, the percentage of healthy people in the former class is about 1.5 times higher than in the latter one, while in France, it's only 1.2 times higher (see Fonseca, Langot, Michaud and Sopraseuth (2018)[\[19\]](#)). The CDC Health Disparities and Inequalities Report of United States (2013)[\[40\]](#) reports also that the percentage of poor health agents among workers is about 8.2%, but it's basically doubled among agents who experience less than one year unemployed duration but even tripped among agents who experience longer than one year unemployed duration. The paradox is that the the share of health expenditures in GDP is larger in the US than in France, and that this share also increases rapidly, from 12.5% in 2000 to 16.9% in 2016. Hence, this trend doesn't seem to favor poor people. According to the report Health Insurance Coverage in the United States(Barnett and Berchick (2017)[\[6\]](#)), though the percentage of uninsured people drops to 8.8% in 2016, the percentage of buying additional private health insurance is still about 67.5% among people who have health insurance plans.

The objective of this paper is to identify, using a structural model, the main sources of these health inequalities by taking advantage of the two contrasting experiences of France and the US. For this purpose, we develop an original extension of the Aiyagari's model[\[3\]](#) where besides the labor income risk, an uncertainty on the health status is introduced. Agents decide to invest or not in health. In order to construct and integrate this new risk into our model, we follow Titus J.Galama (2008)[\[22\]](#) who argues that the "optimal health stock" of Grossman (1972)[\[23\]](#) is only an health threshold.² Thus, only individuals whose health stocks are below Grossman's health threshold would invest in

²In the initial model of Galama, the initial health capital is different from the Grossman's "optimal" health stock. It is then shown that situations with "excessive" initial health are preferable.

health. We inherent this idea so that in our model only individuals in bad health have the need of investing in health. Nevertheless, we take into account that investment in health only gives a better chance to restore health: it is a risky project. For simplicity and as in De Nardi, French and Jones (2016)^[14] or in Fonseca, Langot, Michaud and Sopraseuth (2018)^[19], we assume that the benefice of being healthy is directly added in the utility function. Finally, we do not introduce bonus in wage when individual are in the good health state, because Eric French (2005)^[21] find that the impact of this channel is negligible.³ For simplicity, we also assume that the health investment is a fixed cost and thus the generosity of the health insurance system is simply revealed by the gap between the values of these fix costs among countries: the decision to invest in health is then a discrete choice, as in search models of the labor market where multiple offers are available (see e.g. Algan, Hairault, Langot and Chéron (2001)^[10]). Finally, empirical evidences by Florence Jusot^[32] show that the health inequality between different social and economic classes is not only caused by social health gradient or dangerous habits (alcohol addictive, heavy smoker,etc.), but also by several barriers such as information barriers, or culture barriers, and most importantly, financial barriers. Hence, we not only set health investment constant but also we set it really costly for all agents so that not everyone can afford it. Consequently, agents will choose to invest in health mainly based on their assets accumulation. The complete history of the agents on the labour market determines this wealth distribution and thus the heterogeneous financial capacities to invest in health.

Combining all these factors mentioned above, a general prediction of our model is that first of all, being healthy permits agents to accumulate wealth but being wealthy doesn't necessarily lead to good health, secondly, since health investment is discrete, there should be a clear boundary that separates agents who choose to invest in health and

³In order to show this result, French (2005)^[21] estimates a life-time dynamic structural model of labor supply, retirement, and saving.

those who choose not to. Starting from a calibration where the health expenditures are larger in the US than in France but where the percentage of individuals in good health is smaller, counterfactual experiments show that larger costs of health care services as well as higher risky behavior are the main channels that explain these observed data. It is shown that these determinants of the health market equilibrium (the total cost of the health service and the subjective value of the health for an individual) dominate the lower risk of unemployment on the US labor market. Nevertheless, the low unemployment insurance also contribute to the poor health of the American. Hence, our results suggest that the health insurance system must take into account the status of the individual on the labor market in order to reduce the health inequalities.

The organization of this paper is as follows: in section 2 we present the theoretical model and the algorithm, in section 3 we discuss the method we use to calibrate different parameters, especially the ones related to the health risk. Then we present the results, alongside with their interpretations and intuitions.

4.2 The theoretical model

Individual agents The economy is populated by a continuum of infinitely lived households and the total population is normalized to 1. At the beginning of the current period, a household earns an amount of wealth denoted by a . The household is characterized by an employment status and an health status. Let denote by j the employment status. The household may be employed ($j = e$) or unemployed ($j = u$). The health status is denoted by h . The household may be in good health ($h = g$) or in bad health ($h = b$). To sum up, at the present time, an agent is characterized by a vector of three state variables (a, j, h) .

Job and health status evolve following Markov processes. The process describing the evolution of the job status is exogenous and is characterized by the following transition

matrix:

$$\begin{pmatrix} \pi_{e,e} & \pi_{e,u} \\ \pi_{u,e} & \pi_{u,u} \end{pmatrix}$$

where $\pi_{j,j'}$ is the probability of transition from state j and to state j' (for $j = e, u$ and $j' = e, u$). One obviously has $\pi_{j,e} + \pi_{j,u} = 1$, with $j = e, u$. The process describing the evolution of the health status will be described latter.

The instantaneous utility function of an individual agent includes two additive terms. The first term is the standard CRRA utility function depending on the individual consumption level. The second term corresponds to an additional utility of being in good health. One has:

$$u(c) = \frac{c^{1-\mu} - 1}{1 - \mu} + \mathbf{1}(h = g) \times \bar{u} \quad (4.1)$$

where μ is the risk aversion coefficient and \bar{u} denotes the additional amount of utility of being in good health. Finally, $\mathbf{1}()$ is an indicator function satisfying $\mathbf{1}(h = g) = 1$.

We now describe the budget constraint. At each period, an agent receives an income $y(j)$ depending on his job status. One has $y(e) = (1 - \tau)w$ and $y(u) = b$, with w the wage rate and b the unemployment benefit. The working time of an employed agent is normalized to 1, in other words, labor supply is inelastic. τ denotes the tax rate or payroll tax financing the unemployment insurance.

An agent will take the following decisions. He must choose his consumption c and his asset level a' , and he must decide whether or not to do a fixed health spending d , d represents the total cost of health spending. This is a discrete choice, the decision variable is denoted by m . The variable m takes the value of 1 if the agent invests in health and 0 if he chooses to do nothing. If the agent is in bad health, the health investment will increase the probability of transition to good health. We will discuss this point latter.

Capital is accumulated by the households. One unit of capital is rented by the representative firm at the rate R . Furthermore, capital depreciates at rate δ .

We denote by a the present wealth of the agent and by a' his wealth next period. The state variables characterizing the present situation of the agent are (a, j, h) .

The budget constraint of a state (a, j, h) agent writes:

$$a' = y(j) + (1 + R - \delta)a - c - \mathbf{1}(m = 1) \times d$$

We also assume that agents face a liquidity constraint in such a way that $a' \geq 0$.

Representative firm Let denote by K the aggregate capital and N the aggregate employment. There is a representative firm producing a single good, both consumed and invested, using a constant return-to-scale technology represented by a Cobb-Douglas production function.

$$Y = F(K, N) = AK^\alpha N^{1-\alpha}$$

with $A > 0$ and $\alpha \in]0, 1[$. Assume that markets are perfectly competitive, the firm is price-taker and chooses the amount of factors that maximizes its profits. One has the following optimality conditions:

$$\frac{\partial F(K, N)}{\partial K} = \alpha AK^{\alpha-1} N^{1-\alpha} = R \quad (4.2)$$

$$\frac{\partial F(K, N)}{\partial N} = (1 - \alpha)AK^\alpha N^{-\alpha} = w \quad (4.3)$$

Aggregate variables N denotes the aggregate employment and $U = 1 - N$ is the number of unemployed workers. Assume ρ is the replacement rate, one has $b = \rho w$. For a

given replacement rate, we impose the payroll tax τ adjusts to ensure the balance of the insurance system, that is:

$$\tau wN = b(1 - N)$$

which reduces to:

$$\tau = \rho \frac{1 - N}{N} \quad (4.4)$$

As we previously underlined, agents may decide to do or not the health spending d . The superscripts 1 and 0 respectively point out agents doing and not doing the health spending. One has:

$$\begin{aligned} N &= N^1 + N^0 \\ U &= U^1 + U^0 \end{aligned}$$

Let K' denotes the next period aggregate capital and C the aggregate consumption. Let $\lambda(a, j, h)$ be the density function giving the number of individuals at state (a, j, h) . One has:

$$\begin{aligned}
K &= \sum_j \sum_h \int \lambda(a, j, h) a da \\
K' &= \sum_j \sum_h \int \lambda(a, j, h) a'(a, j, h) da \\
N &= \sum_h \int \lambda(a, e, h) a da \\
U &= \sum_h \int \lambda(a, u, h) a da \\
C &= \sum_j \sum_h \int \lambda(a, j, h) c(a, j, h) da
\end{aligned}$$

The aggregation of the individual budget constraints gives:

$$K' = (1 - \tau)wN + bU + (1 + R - \delta)K - (N^1 + U^1)d$$

Using the balanced budget constraint of the insurance fund and the firm's optimality conditions, one gets:

$$K' = (1 - \delta)K + Y - (N^1 + U^1)d$$

Health status evolution The health status follows a Markov process, however, the transition probabilities are contingent to the care decisions. One has:

$$\begin{pmatrix}
\mathbb{1}(m = 1)\pi_{g,g}^1 + (1 - \mathbb{1}(m = 1))\pi_{g,g}^0 & \mathbb{1}(m = 1)\pi_{g,b}^1 + (1 - \mathbb{1}(m = 1))\pi_{g,b}^0 \\
\mathbb{1}(m = 1)\pi_{b,g}^1 + (1 - \mathbb{1}(m = 1))\pi_{b,g}^0 & \mathbb{1}(m = 1)\pi_{b,b}^1 + (1 - \mathbb{1}(m = 1))\pi_{b,b}^0
\end{pmatrix}$$

where $\pi_{h,h'}^m$ denotes the probability of transition if the care decision is $m = 0, 1$. Additional assumptions are necessary concerning the probabilities of transition. If the agent is in good health, health expenditures do not matter. One obviously has $\pi_{g,g}^0 = \pi_{g,g}^1$ and $\pi_{g,b}^0 = \pi_{g,b}^1$.

If the current status is bad health, the transition probabilities depend on the care decision. It is natural to admit that $\pi_{b,g}^1 > \pi_{b,g}^0$ and $\pi_{b,b}^1 < \pi_{b,b}^0$. The medical expense increases the probability to recover good health.

Household program The problem of an individual agent at state (a, j, h) can be written recursively. The agent has to choose its consumption level $c(a, j, h)$, its asset level $a'(a, j, h)$ and to take its care decision $m(a, j, h)$. Given the prices w and R , the household decisions are solution of the following Bellman equation:

$$V(a, j, h) = \max_{a', c, m} \left\{ \frac{c^{1-\mu} - 1}{1-\mu} + \mathbb{1}(h = g)\bar{u} + \beta \sum_{j'} \sum_{h'} \pi_{j,j'} \pi_{h,h'}^m V(a', j', h') \right\} \quad (4.5)$$

$$\begin{cases} -a' + y(j) + (1 + R - \delta)a - c - \mathbb{1}(m = 1)d \geq 0 \\ a' \geq 0 \end{cases}$$

$\beta < 1$ is the discount factor.

Stationary equilibrium We only focus on the stationary equilibrium. The model is solved recursively. We suppose the replacement rate ρ is given and the payroll tax τ is adjusted each period in order to balance the insurance fund. To begin, we consider a level aggregate capital K and a level of aggregate employment⁴ N . Using Eq (4.2), (4.3) and (4.4), we deduce the rental rate of capital, the wage rate w and the unemployment benefit b , and thus the income process $y(j)$. Solving the household program (4.5), one gets the policy rules $c = c(a, j, h)$, $a' = a'(a, j, h)$ and $m = m(a, j, h)$. Using these policy rules, it is possible to compute the density function $\lambda(a, j, h)$ giving the number of agents at state (a, j, h) . Using these density function, a new value of the aggregate capital and of the

⁴It is easy to show that the stationary level of employment is $N = \frac{1 - \pi_{u,u}}{1 - \pi_{e,e} + 1 - \pi_{u,u}}$.

aggregate employment are computed, that is:

$$\begin{aligned}
K &= \sum_j \sum_h \int \lambda(a, j, h) da \\
N &= \sum_h \int \lambda(a, e, h) da
\end{aligned}$$

The process is continued until convergence.

Note that one also has:

$$\begin{aligned}
N^1 &= \sum_h \int \mathbf{1}[1 = m(a, e, h)] \lambda(a, e, h) da \\
N^0 &= \sum_h \int (1 - \mathbf{1}[1 = m(a, e, h)]) \lambda(a, e, h) da \\
U^1 &= \sum_h \int \mathbf{1}[1 = m(a, u, h)] \lambda(a, u, h) da \\
U^0 &= \sum_h \int (1 - \mathbf{1}[1 = m(a, u, h)]) \lambda(a, u, h) da
\end{aligned}$$

The model is solved using standard numerical methods. The solution method is based on the discretization of the state space variables. We thus suppose the state variable a takes its value in a grid $\mathcal{A} = \{0, a_1, \dots, a_n\}$. One key point is the determination of the density $\lambda(a, j, h)$. In the case of a discretized state space (variable a), $\lambda(a, j, h)$ is given by the following expression:

$$\begin{aligned}
\lambda(a', j', h') &= \sum_{j \in \{e, u\}} \sum_{h \in \{g, b\}} \sum_{a | a' = a'(a, j, h)} \pi_{j, j'} \left\{ \pi_{h, h'}^1 \mathbf{1}[1 = m(a, j, h)] \right. \\
&\quad \left. + \pi_{h, h'}^0 (1 - \mathbf{1}[1 = m(a, j, h)]) \right\} \lambda(a, j, h)
\end{aligned}$$

4.3 Calibration and results

The model is calibrated to match the first order moments estimated using historical data. Under this calibration, we propose to measure the contributions of both health and labor market structures in the explanation on the contrasting equilibria that summarizes the French and the US stories.

4.3.1 Calibration strategy

The vector of the model's parameter is $\Phi = \{\Phi_1, \Phi_2\}$ with $dim(\Phi) = 21$. All parameters calibrated using external information are:

$$\Phi_1 = \{\beta, \mu, \delta, \alpha, \{\pi_{e,u}^j, \pi_{u,e}^j, \rho^j\}_{j=US,FR}\} \quad dim(\Phi_1) = 10$$

The discount factor β is calibrated to match a monthly discount factor consistent with an annual interest rate of 4%. We set μ that governs the relative risk aversion equals to 2. The depreciation rate of capital (δ) is set such that 8% per year of the used capital is depreciated. The parameter of the Cobb Douglas production function is equal to 0.3. The worker flows $\{\pi_{eu}^j, \pi_{ue}^j\}_{j=US,FR}$ are provided by Shimer (2005)[\[45\]](#) for the US and by Barnichon, Hairault and Sopraseuth (2015)[\[25\]](#) for France. These worker flows lead to a unemployment rate equal to 7.22% in the US and 10.26% in France. Finally, we estimate the replacement rates $\{\rho^j\}_{j=US,FR}$ as the gross unemployment replacement rate in OECD database and follow the results reported by CESifo Groupe (Unemployment Benefit Replacement Rates, 1961-2011, 2012).

	β	μ	δ	α	$\pi_{e,u}$	$\pi_{u,e}$	ρ
France	0.9975	2	8%	0.3	0.016	0.14	0.32
US					0.035	0.45	0.15

Table 4.1: Parameters for France and the US

For the other parameters, we need some restrictions in order to identify them using a set of moments computed using French and US data. We take as a reference the GDP of the US economy, and thus normalize a^{US} to unity. Hence, 11 parameters are calibrated:

$$\Phi_2 = \left\{ a^{FR}, \left\{ \bar{u}^j, d^j, \pi_{g,g}^j, \pi_{b,b}^{j,1}, \pi_{b,b}^{j,0} \right\}_{j=US,FR} \right\} \quad \dim(\Phi_2) = 11$$

The calibrated parameters are the solution to $\min_{\Phi_2} \|\Psi^{theo}(\Phi_2) - \Psi\|$, where the numerical solution for $\Psi^{theo}(\cdot)$ is provided by the algorithm based on value function iterations.⁵ The 11 free parameters are the elements of Φ_2 , whereas the 11 first-order moments provided by the data are:

$$\Psi = \left\{ \frac{Y^{FR}}{Y^{US}}, \left\{ \frac{M^j}{Y^j}, p^j(H=0), \frac{p^j(H=1|I \in Q2)}{p^j(H=1|I \in Q1)}, \frac{p^j(H=1|I \in Q3)}{p^j(H=1|I \in Q1)}, \frac{p^j(H=1|I \in Q4)}{p^j(H=1|I \in Q1)} \right\}_{j=US,FR} \right\}$$

with $\dim(\Psi) = 11$.

4.3.2 Results

The solution for the calibrated parameters are reported in Table 4.2. First of all, good health is more valued in France than in the US. This result is in accordance with the less risky behaviors of the French people, in particular in terms of obesity. This is already underlined by Fonseca, Langot, Michand and Sopraseuth (2017)[19]. Concerning the efficiency of the health sector, the transition matrices provide the probability gaps to be cured conditionally to spend money in health. If the agent don't buy health services, the risks are characterized by $\pi_{g,g}^{FR} > \pi_{g,g}^{US}$ and $\pi_{b,b}^{FR,0} < \pi_{b,b}^{US,0}$. This suggests that for agents

⁵Value function iteration algorithm applies the contraction mapping theorem. The algorithm is guaranteed to converge by this theorem. First, start by making an initial guess for the value function at each capital point (an initial guess of zero at each point for example). Second, compute the first iteration of the value function by considering the future value as your initial guess. This yields a new value: the sum of the current payoff and the discounted (expected) future payoff. Use this value as the future value in the next iteration to produce a new value, etc... The algorithm stops when the gap between the value functions of two successive iterations becomes negligible. See Ljungqvist and Sargent (2000)[36] for more details on this method.

who are already in good health, it's better for them to be in France than in the US, but if they are in bad health the probability to cure is lower in France. In average, without health expenditures, the probabilities to be in good health are 58.46% and 48.72% respectively in France and in the US. If the agent spend money in health services, we have $\pi_{b,b}^{FR,1} > \pi_{b,b}^{US,1}$, showing that the probability to cure is higher in the US. This suggests that the efficiency of the health sector is higher in the US than in France. Remark that for these agents spending money in health services, the expectations to be in good health are 80% and 79.44% respectively in France and in the US. Our results also show that the costs of health services are larger in the US than in France: we have $d^{US} > d^{FR}$. This is consistent with previous study of Fonseca, Langot, Michand and Sopraseuth (2017)[\[19\]](#). Remark that this gap in the health service costs are also due an higher Out-of-pocket (*OOP*) in the US than in France: in OECD data, we have $OOP^{US} = 13\%$ and $OOP^{FR} = 7\%$.

	a	\bar{u}	d	$\pi_{g,g}$	$\pi_{b,b}^1$	$\pi_{b,b}^0$
France	0.0803	2.295	0.01482	0.801	0.19	0.72
US	0.104	2.1	0.0347	0.78	0.15	0.791

Table 4.2: Calibrated parameters for France and the US

The fit of the model is reported in the Table [4.3](#). We find that almost all moments we pick match the data, all except one (the percentage of healthy agents in the third revenue quartile versus the ones in the first revenue quartile), but it's because the discrete nature of the decision that agents make. In a discrete choice model, there exists a wealth threshold for each type of agent (employed and unemployed workers) above which the optimal decision is to consume health services because they are in bad health. Given that this expenditure in health services is a fix cost, all the agents above the threshold have access to the same health services and thus have the same chance to become healthy agents. Hence, in the model, the population is divided into 2 groups: those who have a low health risk and those who have a high health risk.⁶ The equilibrium wealth distributions

⁶See the appendix [4](#) for more detailed statistics on the health gradient.

	France		US	
	model	data	model	data
Y^{FR}/Y^{US}	0.6769	0.666	1	1
M^j/Y^j	9.82%	10%	14.59%	14.6%
$p^j(H = 0)$	26.48%	28%	30.6 %	30%
$p^j(H = 1 I \in Q2)/p^j(H = 1 I \in Q1)$	1.1012	1.06	1.1728	1.196
$p^j(H = 1 I \in Q3)/p^j(H = 1 I \in Q1)$	1.2673	1.17	1.4631	1.337
$p^j(H = 1 I \in Q4)/p^j(H = 1 I \in Q1)$	1.2688	1.245	1.4741	1.475

Table 4.3: Simulation results for France and the US

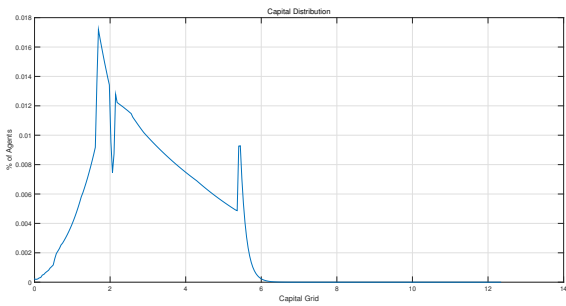
displayed in Fig 4.1 show that there exists 2 groups of agents in these economies⁷: those we have access to the health services, and those who have not. Hence, the model generates two health groups. The first one has a low risk to be in bad health and in this case the probability to cure is high. Hence, whether these agents are in good or bad health, with a job or not, are mainly characterized by a amount of assets larger than the threshold above which they pay the access to the health services. This feature of the equilibrium wealth distribution also explains why the percentage of healthy agents in the third revenue quartile does not match the data: in the top two revenue quartiles, all agents are rich enough to choose to invest in health. Thus the percentage of healthy agents in the third revenue quartile is rather close to the one in the last revenue quartile but relatively far from the data showing that our two-states model is too simple to account for this heterogeneity.

4.4 Counterfactual experiments

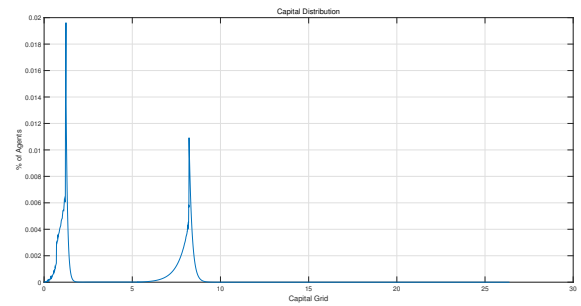
Given these characteristics of the benchmark equilibria, we can now proceed to counterfactual simulations in order to evaluate the relative contributions of each country-specific calibration. To this end, we apply six different scenarios starting from the US case:

- (1) The cost of health investment d is the same as the cost in France

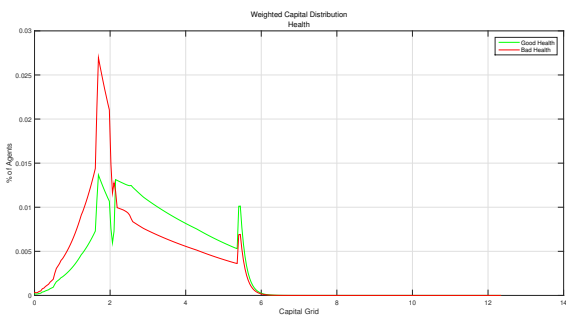
⁷See the appendix B for the general equilibrium on financial market.



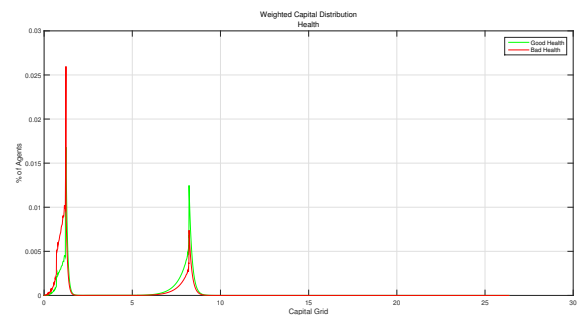
(a) France



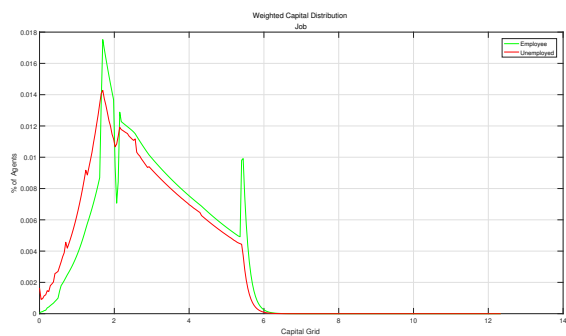
(b) The US



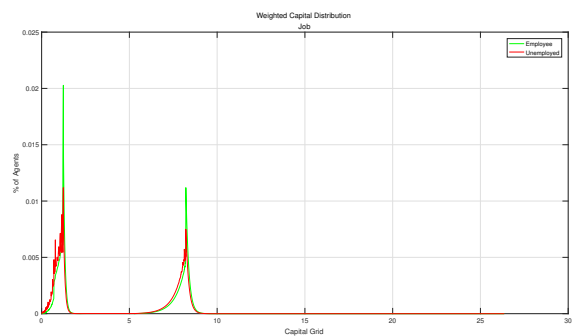
(c) By health status - France



(d) By health status - the US



(e) By employment status - France



(f) By employment status - the US

Figure 4.1: Capital Distribution

- (2) The benefit of being healthy \bar{u} is the same as the benefit in France
- (3) The health risk (the transition matrix Π_H) is the same as in France
- (4) The health risk Π_H , the d and the \bar{u} are all the same as in France
- (5) The employment risk (the transition matrix Π_e) is the same as in France
- (6) The replacement rate ρ is the same as the replacement rate in France

The results are in the Table [4.4](#).

Scenario	(1) d^{FR}	(2) \bar{u}^{FR}	(3) Π_H^{FR}	(4) d^{FR} & \bar{u}^{FR} & Π_H^{FR}	(5) Π_e^{FR}	(6) ρ^{FR}	(7) US bench.
$p(H = 1)$	79.44%	79.27%	60.12%	80.28%	75.5 %	77.5%	69.1 %
$p(Buy = 1)$	99.99%	99.36%	6.84 %	99.98%	86.55%	92.96%	64.26%
r	0.2395%	0.2458%	0.2449%	0.2373%	0.2296%	0.2474%	0.2459%
Welfare*	4.48%	2.92%	2.20%	7.22%	-3.65%	0.42%	-7.3545
Cons. losses	-4.3%	-2.9%	-2.15%	-6.8%	3.79%	0.4%	8.11%

* Variations of the welfare relative to the benchmark: $((\text{Welfare}(7) - \text{Welfare}(x)) / \text{Welfare}(7)) \times 100$

Table 4.4: Counterfactual experiments - US

There are two risks in the economy. The first is the job loss leading to low labor incomes and the second is the bad health leading to low welfare. Larger income risks lead the households to save more in order to smooth their consumption. At general equilibrium, this precautionary saving can reduce the interest rate. But it is not always the case because the low employment rate reduces the marginal return of the capital. Obviously these income risks (governed by the matrix Π_e) can also be reduced by the unemployment benefits (ρ). If France is characterized by an higher unemployment risk, this country is also more generous than the US for the unemployed workers.

Income risks (5). If the French labor market risks are "exported" in the US (column (6)), there will be more precautionary saving in the US: the US workers must insure

themselves against this higher risk. Hence, there is more capital supply in the economy. But, at the same time, the level of employment declines (the unemployment rate is multiplied by two), leading to a decline in the demand of capital. This last effect dominates: the interest rate increases and the aggregate production declines. Hence, the higher unemployment risk is not overcompensated by the increase of the precautionary saving of the household. There is less wealth in average in this economy than in the benchmark calibration of the US economy. The surprising result comes from the share of individuals in good health which increases. This is explained by the longer employment spells when the labor market transitions of the French economy replace those of the US economy. In this counterfactual experiment, these long employment spells allow the individuals to reach the wealth threshold from which the agents choose to buy health services. But, this cannot compensate the losses in consumption induced by the decline in production: the welfare is reduced with respect to the benchmark.

Unemployment benefits (6). An increase of the unemployment benefits reduces the incentives to save and thus the capital amount in the economy because the unemployment risk being more generously insured by inter-agent transfers. This increases the interest rate because the capital demand is not affected by the change in unemployment benefits. This lower level in capital can reduce the welfare. But, this is not the case because the consumption of health services increases. Indeed, less saving is needed to insurance labor market risks, then more resources are available for health expenditures. This is the case and thus more agents are in good health because a larger part of them choose to buy health services. Thus, the average welfare increases.

Cost of the health services (1). The decrease of d leads a larger set of agents to invest in health: the risk to be in bad health declines. Numerical results show that the percentage of healthy agents rise up to the limit. At the same time, more resources are

available to insure agents against the unemployment risk. Hence, there is more capital in the economy and thus more production. This allows household to reach a higher welfare.

Welfare benefits induced by good health status (2). In a society where the good health is more valued (the risky behaviors as obesity more penalized), the incentives to buy health services increase. Hence, more agents spend in health services and as a result more people are in good health. The counterpart of these choices is a reduction of the saving leading to a rise in the interest rate. The impact on the welfare is a priori undetermined: the direct impact of the rise value of the good health can be compensated by the decline in the aggregate capital. Numerical results show that the former effect dominates.

Efficiency of the health sector (3). When the French matrix of the health risks is implemented in the US economy, the average chance to be in bad health declines. This largely reduces the incentive to spend money in health services (only 6.84% of people buy health services). Given the high costs of the health services, agents prefer to play with the "health lottery" where the medical sector do not take part. Hence, more resources can be devoted to the saving (the insurance of the income risks) leading the capital stock to increase (the interest rate declines). At the end, the agents welfare is improved

Cost, welfare benefits and efficiency of the health sector (4). This last scenario shows that even if the risk to be in bad health is reduced to value for the French economy, a lower cost of the health services is crucial to provide incentives to spend in health services. Moreover, the larger utility value of the good health status magnifies this incentive. Hence, the combination of these three changes in the US economy would allow American people to have a higher welfare than the French citizens.

4.5 Conclusion

This paper shows how a general equilibrium model of Aiyagari (1994)[\[3\]](#) augmented by health care choices in the spirit of the Grossman (1972)[\[23\]](#) can match the US and France differences, namely the facts that the US the health expenditures are larger than in France but the percentage of individuals in good health is smaller. Using counterfactual experiments, it is shown that larger costs of health care services as well as higher risky behavior are the main channels that explain these results. More precisely, the welfare gain induced by the a reduction of cost of health service in the US such that it would be equal to the French one, is equivalent to a permanent increase in consumption of 4.3%. A reduction of the risky behaviors with respect to health, such that they would be equal to the French ones, will induce a rise by 2.9% permanent increase in consumption. The risks associated to the labor market interact with the health status of the individuals. We show that the large turnover observed in the US induces low investment in health, even if these turnovers are associated to short periods of unemployment, and low unemployment rate. Indeed, what matters for the health decision seems to be the stability of the jobs. These costs in term of health linked to turnover are partially compensated by the higher employment level allowing to reach a higher production level at the general equilibrium. Finally, we show that an increase of the unemployment benefits can improve significantly the health of the American. This rise in the distortion reduces the saving and thus the production, but this general equilibrium feedback does not compensate the impact of a more generous unemployment benefit system on health status. Nevertheless, the determinants of the health market equilibrium (the cost of the health care services and the risky behaviors) seem to dominate those of labor market.

In future researches, more detailed risks on the labor market must be introduced. It will be also necessary to distinguish between health care expenditures that have direct impact on the life expectancy and those that necessary during periods of dependency.

Chapter 5

General Conclusion

In this dissertation we study the impacts of health on human life, as mentioned in the introduction, we start with the deterministic approach, in this approach, we focus on individual's heterogeneous behaviours when facing 'health fluctuations', then we set out to the stochastic approach, in this approach, we use the Markov chain to simulate the uncertain nature of health, we then examine how individuals of different SES would react to this 'real' health risk, finally, we aggregate these behaviours to see the impacts on the whole economy. Concretely, for the first approach, we construct a simplified Grossman health capital model with finite and endogenous life horizon, as for the second approach, we construct an augmented Aiyagari model with stochastic health shocks using the Markov chain.

In this section, we aim to examine how the questions proposed at the end of the first chapter are resolved:

Question 1 the possibility of immortality has long been one of the major criticisms to the Grossman health capital model, so is it possible to have a Grossman health capital model with a reasonable life horizon?

Chapter 2 answers this question, we prove that, by adding explicitly a non negative health spending constraint(i.e. $I_t \geq 0$), a survivability condition(i.e. $H_t - \underline{H} \geq 0$) and a age-dependant health depreciation function, a simplified Grossman model following Laporte(2015)[34] and Strulik(2015)[47] can thereafter derive the optimality conditions that allow to characterize the solution to household's problem with endogenous lifespan, thus the possibility of immortality is avoided. In addition, using this model, a numerical solution can be obtained: we calibrate the model to fit the U.S. economy and we show that such a model can generate a decreasing health capital profile, an increasing health spending profile and a significant level of health spending especially at the end of life, but a non-smooth consumption profile is also obtained, which is mainly because of the lack of the financial market.

Question 2 Can there be a single framework that can replicate the key macroeconomic health economic moments of multiple countries?

Chapter 3 answer this question by constructing an augmented Aiyagari model, the original Aiyagari model includes a stochastic income shock, we extend this model by introducing a stochastic health shock. The model is proved to be able to fit pretty well the key macroeconomic moments of French and the U.S., and there is a clear separation in the demographic distribution between individuals who choose to invest in health and individuals who choose not to. Plus, with a series of counterfactual experiments, we show that though the income risk and the health risk are set independently, the former still interacts with the latter. Empirically, the Americans have a poorer health than French, we show that the low unemployment risk(shorter period of unemployment)in the U.S. contributes to explain this phenomenon, for a lower unemployment risk leads to a less incentive to invest in health. Indeed, what matters for the health decision seems to be the stability of the jobs. These costs in terms of health linked to turnover are partially

compensated by the higher employment level allowing to reach a higher production level at the general equilibrium.

Concerning the impacts of public policies, we show that a French-style health service market, which means lower health expenditures and a higher satisfaction of being healthy (higher utility bonus for being healthy) can improve both the health and the welfare of the American. Indeed, larger costs of health care services as well as higher risky behaviours in the U.S. are the main channels that explain the poorer health of the Americans if we compare the 2 countries. More precisely, the welfare gain induced by the a reduction of cost of health service in the U.S. such that it would be equal to the French one, is equivalent to a permanent increase in consumption of 4.3%. A reduction of the risky behaviours with respect to health, such that they would be equal to the French ones, will induce a rise by 2.9% permanent increase in consumption.

In future researches, we should always continue the study in both approaches: in the deterministic approach, introduction of financial market should be considered, plus, a real health shock should be integrated into the model to see its impact on the individual level; in the stochastic approach, a non-binary health investment choice should be included to make a more realistic assumption, in addition, we should continue testing the credibility of the model with more countries, furthermore, other nations like China or Japan could also be considered to see the differences of health decisions under alien cultures.

Bibliography

- [1] António Afonso and Mina Kazemi. Assessing public spending efficiency in 20 oecd countries. In *Inequality and Finance in Macrodynamics*, pages 7–42. Springer, 2017.
- [2] António Afonso and Miguel St Aubyn. Non-parametric approaches to education and health expenditure efficiency in oecd countries. 2004.
- [3] S Rao Aiyagari. Uninsured idiosyncratic risk and aggregate saving. *The Quarterly Journal of Economics*, pages 659–684, 1994.
- [4] Rajiv D Banker, Abraham Charnes, and William Wager Cooper. Some models for estimating technical and scale inefficiencies in data envelopment analysis. *Management science*, 30(9):1078–1092, 1984.
- [5] Y. Ben-Porath. The production of human capital and the life cycle of earnings. *The Journal of Political Economy*, 75(4):352–365, 1967.
- [6] Edward Berchick and Edward R Barnett. Health insurance coverage in the united states: 2016. *Washington, DC <https://www.census.gov/library/publications/2017/demo/p60-260.html>*, 2017.
- [7] Blaise and Lefebvre. La mortalité selon le revenu:estimation à partir de données européennes. *Revue Française d'Économie*, 2018.

- [8] Volkan Recai Cetin and Serdal Bahce. Measuring the efficiency of health systems of oecd countries by data envelopment analysis. *Applied Economics*, 48(37):3497–3507, 2016.
- [9] Abraham Charnes, William W Cooper, and Edwardo Rhodes. Measuring the efficiency of decision making units. *European journal of operational research*, 2(6):429–444, 1978.
- [10] Arnaud Chéron and François Langot. Real wage dynamics in the labor market search model. Technical report, Mimeo, CEPREMAP, 2001.
- [11] Raj Chetty, Michael Stepner, Sarah Abraham, Shelby Lin, Benjamin Scuderi, Nicholas Turner, Augustin Bergeron, and David Cutler. The association between income and life expectancy in the united states, 2001-2014. *Jama*, 315(16):1750–1766, 2016.
- [12] Long M Damico A Sawyer B Claxton G, Rae M. 2016 employer health benefits survey. Technical report, Kaiser Family Foundation, 2016.
- [13] Organisation Mondiale de la Santé. Comblen le fossé en une génération. *Instaurer l'équité en santé en agissant sur les déterminants sociaux de santé*, 2008.
- [14] Mariacristina De Nardi, Eric French, and John Bailey Jones. Medicaid insurance in old age. *American Economic Review*, 106(11):3480–3520, 2016.
- [15] Marion Devaux and Michael De Looper. Income-related inequalities in health service utilisation in 19 oecd countries, 2008-2009. 2012.
- [16] Samuel L Dickman, David U Himmelstein, and Steffie Woolhandler. Inequality and the health-care system in the usa. *The Lancet*, 389(10077):1431–1441, 2017.

- [17] Paul Dourgnon, Florence Jusot, and Romain Fantin. Payer nuit gravement à la santé: une étude de l'impact du renoncement financier aux soins sur l'état de santé. *Économie publique/Public economics*, (28-29):123–147, 2012.
- [18] R. Fonseca, P-C. Michaud, T. Galama, and A. Kapteyn. Retirement and the demand for health. Mimeo, 2008.
- [19] Raquel Fonseca, Francois Langot, Pierre-Carl Michaud, and Thepthida Sopraseuth. Paying the price: Accounting for health status and expenditures across country. Technical report, Mimeo, CEPREMAP, 2018.
- [20] E.B. French, J. McCauley, M. Aragon, et al. End-of-life medical spending in last twelve months of life is lower than previously reported. *Health Affairs*, 36(7):1211–1217, 2017.
- [21] Eric French. The effects of health, wealth, and wages on labour supply and retirement behaviour. *The Review of Economic Studies*, 72(2):395–427, 2005.
- [22] Titus J Galama, Arie Kapteyn, Raquel Fonseca, and Pierre-Carl Michaud. Grossman's health threshold and retirement. 2008.
- [23] M. Grossman. On the concept of health capital and the demand for health. *The Journal of Political Economy*, 80(2):223–255, 1972.
- [24] Michael Grossman. The human capital model. In *Handbook of health economics*, volume 1, pages 347–408. Elsevier, 2000.
- [25] Jean-Olivier Hairault, Thomas Le Barbanchon, and Thepthida Sopraseuth. The cyclicity of the separation and job finding rates in france. *European Economic Review*, 76:60–84, 2015.
- [26] R.E. Hall and C.I. Jones. The value of life and the rise in health spending. *The Quarterly Journal of Economics*, 122(1):39–72, 2007.

- [27] T.J. Halliday, H. He, L. Ning, and H. Zhang. Health investment over the life-cycle. *Macroeconomic Dynamics*, 2017. doi:10.1017/S1365100516001152.
- [28] Jessica Y Ho and Arun S Hendi. Recent trends in life expectancy across high income countries: retrospective observational study. *bmj*, 362:k2562, 2018.
- [29] Rok Hren et al. Theoretical shortcomings of the grossman model. *Bulletin: Economics, Organisation and Informatics in Healthcare*, 28(1):63–75, 2012.
- [30] Martijn Huisman, Anton E Kunst, and Johan P Mackenbach. Socioeconomic inequalities in morbidity among the elderly; a european overview. *Social science & medicine*, 57(5):861–873, 2003.
- [31] J. Jung and T. Tran. Medical consumption over the life-cycle: Fact from u.s. medical expenditures panel survey. Working Paper 2010-09, Towson University, 2013.
- [32] Florence Jusot. Les inégalités sociales de santé: quel constat? quels leviers d’action? Technical report, Note pour le club parlementaire Hippocrate, Collège des Economistes de la santé, 2010.
- [33] Florence Jusot, Zeynep Or, and Nicolas Sirven. Variations in preventive care utilisation in europe. *European Journal of Ageing*, 9(1):15–25, 2012.
- [34] A. Laporte. Should the grossman model of investment in health capital retain its iconic status. Working Paper 2014-04, Canadian Center for Health Economics, 2015.
- [35] A. Laporte and B. Ferguson. Making the grossman model stochastic: Investment in health as a stochastic control problem. Working Paper 170009, Canadian Center for Health Economics, 2017.
- [36] Lars Ljungqvist and Thomas J Sargent. *Recursive macroeconomic theory*. MIT press, 2004.

- [37] JP Mackenbach, Willem Jan Meerdink, and Anton Kunst. *Economic implications of socio-economic inequalities in health in the European Union*. European Commission, 2007.
- [38] Olivier Nay, Sophie Béjean, Daniel Benamouzig, Henri Bergeron, Patrick Castel, and Bruno Ventelou. Achieving universal health coverage in france: policy reforms and the challenge of inequalities. *The Lancet*, 387(10034):2236–2249, 2016.
- [39] Zehra Önen and Serpil Sayın. Evaluating healthcare system efficiency of oecd countries: A dea-based study. In *Operations research applications in health care management*, pages 141–158. Springer, 2018.
- [40] Healthy People. Conclusion and future directions: Cdc health disparities and inequalities report united states, 2013. *CDC Health Disparities and Inequalities Report United States, 2013*, 62(3):184, 2013.
- [41] Louise Potvin, Marie-José Moquet, Catherine Jones, and Michael Marmot. *Réduire les inégalités sociales en santé*. INPES éditions, 2010.
- [42] S. Rosen. The value of changes in life expectancy. *Journal of Risk and Uncertainty*, 1(3):285–304, 1988.
- [43] Isaac Sasson. Trends in life expectancy and lifespan variation by educational attainment: United states, 1990–2010. *Demography*, 53(2):269–293, 2016.
- [44] Ardeshir Sepehri. A critique of grossmans canonical model of health capital. *International Journal of Health Services*, 45(4):762–778, 2015.
- [45] Robert Shimer. Reassessing the ins and outs of unemployment. *Review of Economic Dynamics*, 15(2):127–148, 2012.

- [46] David Squires and Chloe Anderson. Us health care from a global perspective: spending, use of services, prices, and health in 13 countries. *The Commonwealth Fund*, 15:1–16, 2015.
- [47] H. Strulik. A closed-form solution for the health capital model. *Journal of Demographic Economics*, 3(81):301–316, 2015.
- [48] Richard MH Suen et al. Technological advance and the growth in health care spending. *Economie D'Avant Garde, Research Report*, (13), 2006.
- [49] Peter Zweifel. The grossman model after 40 years, 2012.

Appendix A

One example of the DEA method

At the core of the DEA methodology is the concept ‘relative efficiency’, and since it’s ‘relative’, the basic idea of this concept is to compare performances of DMUs: first of all, among all units that are to be studied, the DEA methodology will identify the units that can be considered as 100% efficient, these units would be thereafter called efficient units and all other units would be called as inefficient units, secondly, the DEA methodology will identify for each inefficient unit a reference set of efficient unit(s) that are the closest to its production function, this reference set of efficient unit(s) are called *pairs* or *peers* to the inefficient unit, for each inefficient unit, *pairs* could be different, finally, the DEA methodology compare the performance of each inefficient unit to its *pairs* and calculate its relative efficiency score.

The DEA method measures unit’s efficiency by 2 orientations: either input-oriented or output-oriented. In general, input-oriented DEA method measures how much inputs one unit could save without reducing its outputs, correspondingly, output-oriented DEA method measures how much more outputs one unit can producing without using more inputs.

Let us use an example to illustrate how the DEA methodology functions in practice:

suppose that there are 5 general clinics, each clinic has its own number of doctors that would cure different number of patients, their monthly performances are summarised in the table below

Clinic	Doctors	cured patients($\times 100$)
A	2	1
B	3	4
C	5	5
D	4	3
E	6	7

Now let us apply the DEA methodology to compare the performances of these 5 clinics, first of all efficient unit(s) needs to be identified, and in order to identify the efficient unit(s), the DEA methodology uses the ‘efficiency frontier’, note that in general there are 2 types of ‘efficiency frontier’: a constant return to scale(CRS) efficiency frontier and a variable return to scale(VRS) efficiency frontier, this distinction also divides the DEA model into 2 general types: a CRS DEA model(Charnes, Cooper and Rhodes (1978)[9]) and a VRS DEA model(Banker, Charnes and Cooper (1984)[4]). Let us draw the performances of the 5 clinics of our example and also the 2 types of ‘efficiency frontier’ and show the differences under each circumstance.

Figure A-1 shows the efficiency frontier, which is a straight line that goes through origin and the point B , therefore in this case only clinic B is located on the frontier and thus considered as an efficient unit, all other clinics are considered as inefficient units. To calculate the relative efficiency, according to the DEA methodology, there are in general 2 ways, either input-oriented or output-oriented: for an inefficient unit, we compare its performance to the performance of its *pairs*, in this case, since only clinic B is considered as efficient unit, thus we compare the performances of all other clinics to clinic B , the efficiency of clinic B is $\frac{4}{3}$, which means per unit of inputs used by clinic B would produce $\frac{4}{3}$ units of outputs¹, or reversely, per unit of outputs that clinic B produces needs $\frac{3}{4}$ units

¹In our scenario, it means each doctor in clinic B would cure $\frac{4}{3} \times 100$ patients per month.

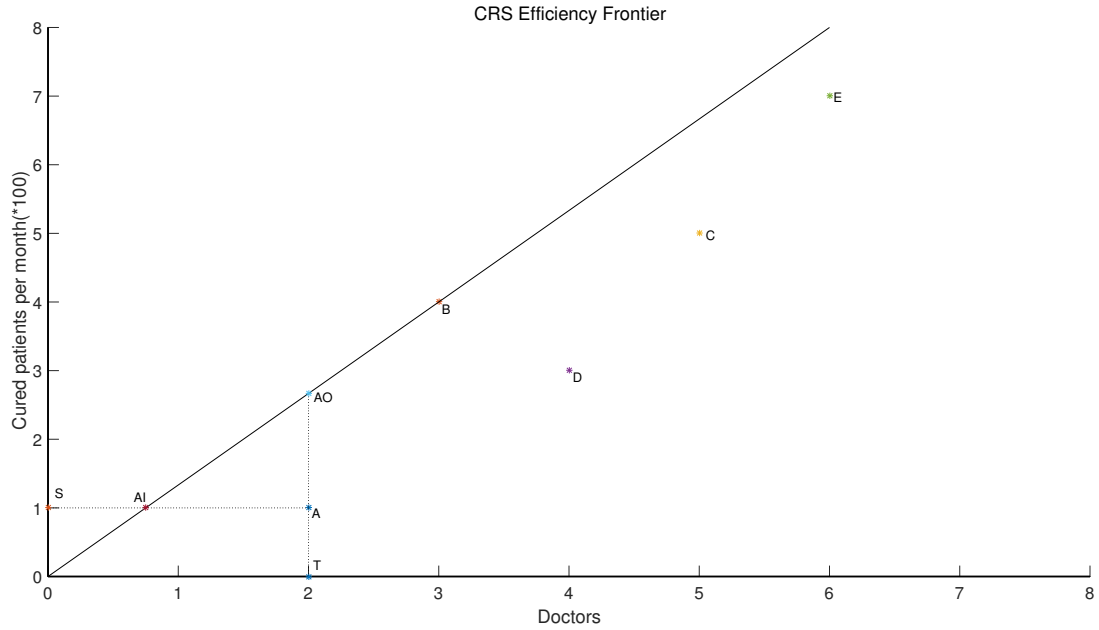


Figure A.1: CRS DEA efficiency

of inputs², therefore to compare the performance of an inefficient unit, for example, clinic A, to the performance of clinic B, either we compare their output efficiency by holding the same level of input, either we compare their input efficiency by holding the same level of output. If we take the former approach, then we calculate how much should clinic A produce with the same level of inputs(2) but with clinic B's input efficiency($\frac{3}{4}$), that's to say, for clinic A, with its 2 doctors, it should have been able to cure $\frac{8}{3} \times 100$ patients per month but in fact it only cures 1×100 patients, thus its output efficiency equals to $\frac{1}{\frac{8}{3}} = \frac{3}{8}$, it is shown in Figure A as point AO, to obtain clinic A's relative efficiency, we calculate $\frac{TA}{TAO} = \frac{3}{8}$, this is what is called the output-oriented efficiency, if we take the latter approach, then we calculate how much would clinic A expense with the same level of outputs(1) but with clinic B's output efficiency($\frac{4}{3}$), that's to say, for clinic A, with its current per month cured patients 1×100 , it should have used only $\frac{3}{4}$ doctors but in fact it uses 2, thus its input-efficiency equals to $\frac{1}{\frac{3}{4}} = \frac{4}{3}$, it is also shown in Figure A as point AI, to obtain clinic A's relative efficiency, we calculate $\frac{SAI}{SA} = \frac{3}{8}$, this is called the

²In our scenario, it means every 100 patients need $\frac{3}{4}$ doctor to cure.

input-oriented efficiency.

As what this example shows, input-oriented efficiency shows how much inputs could be reduced without affecting the outputs, for example, clinic A has an input-oriented efficiency of $\frac{3}{8}$, it means to be efficient, clinic A should reduce its number of doctors to $2 \times \frac{3}{8} = \frac{3}{4}$ while keeping curing 100 patients per month, correspondingly, the output-oriented efficiency measures how much outputs should be increased without using more inputs, clinic A has also an output-oriented efficiency of $\frac{3}{8}$, it means to be efficient, clinic A should increase its monthly cured patients to $\frac{1}{\frac{3}{8}} = \frac{8}{3}(\times 100)$ while keeping its 2 doctors.

Note that under CRS, input-oriented efficiency score is always equal to output-oriented efficiency score, but this is not the case under VRS efficiency frontier. Figure [A.2](#) shows the efficiency frontier under VRS of our example, we can see that under VRS, point B is still located on the frontier, but also point A and point E are on the frontier too, that's to say, under VRS, there are more efficient units. To calculate the relative efficiency score of an inefficient unit, for example clinic D, using the methods we show above, we can obtain the input-oriented efficiency score of clinic D by $\frac{UDI}{UD} = \frac{1.8}{4} = \frac{2}{3}$ and the output-oriented efficiency score by $\frac{VD}{VDO} = \frac{3}{5}$. Note that in this case the input-oriented score is different from the output-oriented score, which is logic since from Figure [A.2](#) we can see that when we calculate the input-oriented efficiency score of clinic D, we compare D's input efficiency with the performances of A and B, and when we calculate the output-oriented efficiency score of clinic D, we actually compare D's output efficiency with B and E, since the reference set or *pairs* are different and since AB and BE are 2 different lines with 2 different slopes, it's only normal that we would obtain 2 different scores.

Now let us combine these 2 efficiency frontiers in order to illustrate the scale efficiency. In fact, a CRS DEA model is mostly suitable under circumstances that all units are operating with their optimal size, this obviously refers to the perfect market situation which is rarely true, correspondingly, a VRS DEA model is often used when all units

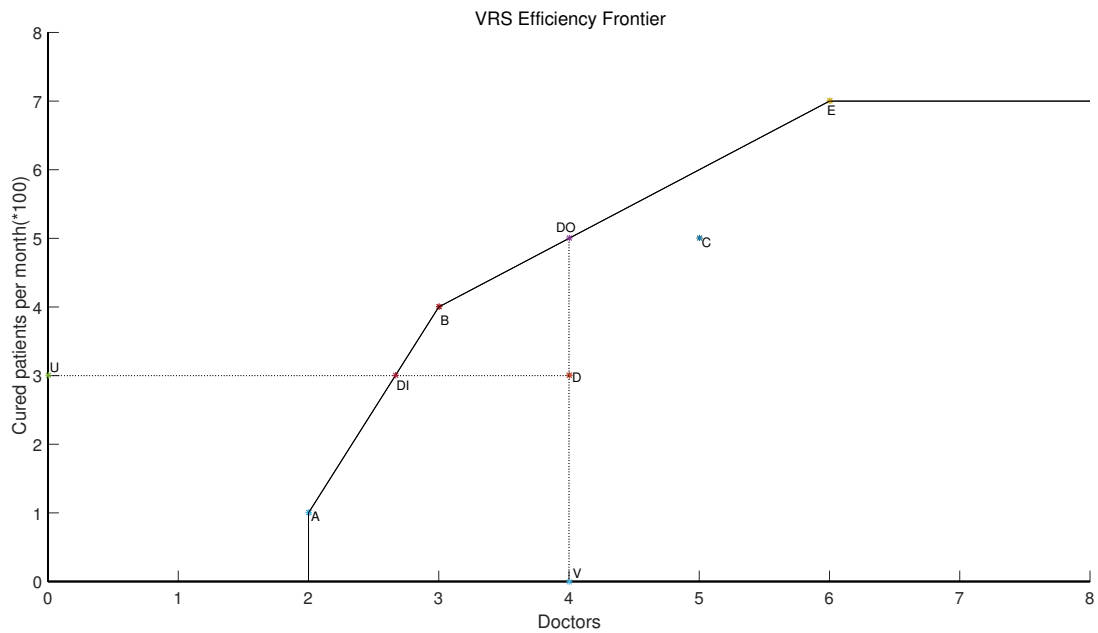


Figure A.2: VRS DEA efficiency

are operating without their optimal size. Obviously, a VRS DEA model is more realistic than a CRS DEA model but it doesn't mean that a CRS DEA model is meaningless, in fact, the DEA efficiency under CRS 'includes' that under VRS in the sense that the former measures the general performance of a DMU and the latter measures only the pure technical efficiency (e.g. bad performance because of deficient management), therefore the difference between these 2 is called scale efficiency since it is considered to evaluate DMU's inefficiency because of its size. Figure [A.3](#) shows the 2 types of efficiency frontier in one figure, we can see that clinic B is not only CRS but also VRS efficient, clinic A and E are VRS but not CRS efficient, clinic D and C are neither VRS nor CRS efficient. The margin between CRS frontier and VRS frontier could be explained by the scale inefficiency, that to say the inefficiency caused by unit's size, in this and the next 2 paragraphs we only show the inefficiency caused by unit's input size. as we discuss above, CRS is used under assumption that all units operate with their optimal size and VRS is used under assumption that all units operate with non-optimal size, thus for an unit that is VRS efficient to become CRS efficient, for example clinic A, it needs to modify its size, in

Figure A.3 we can see that, to become CRS efficient, clinic A needs to move to point *ACI*. For an unit who is neither VRS nor CRS efficient, to become CRS efficient, it needs to be first of all VRS efficient then CRS efficient, in our example, clinic D is neither VRS nor CRS efficient, from Figure A.3 we can see that first clinic D needs to move to point *DVI* to become VRS efficient, then to point *DCI* to become CRS efficient, the scale efficiency of clinic D could be therefore calculated by $\frac{UDCI}{UDVI} = \frac{9}{\frac{4}{3}} = 0.844$, which means in order to become CRS efficient, clinic D should still reduce 15.6% of its inputs even when become VRS efficient.

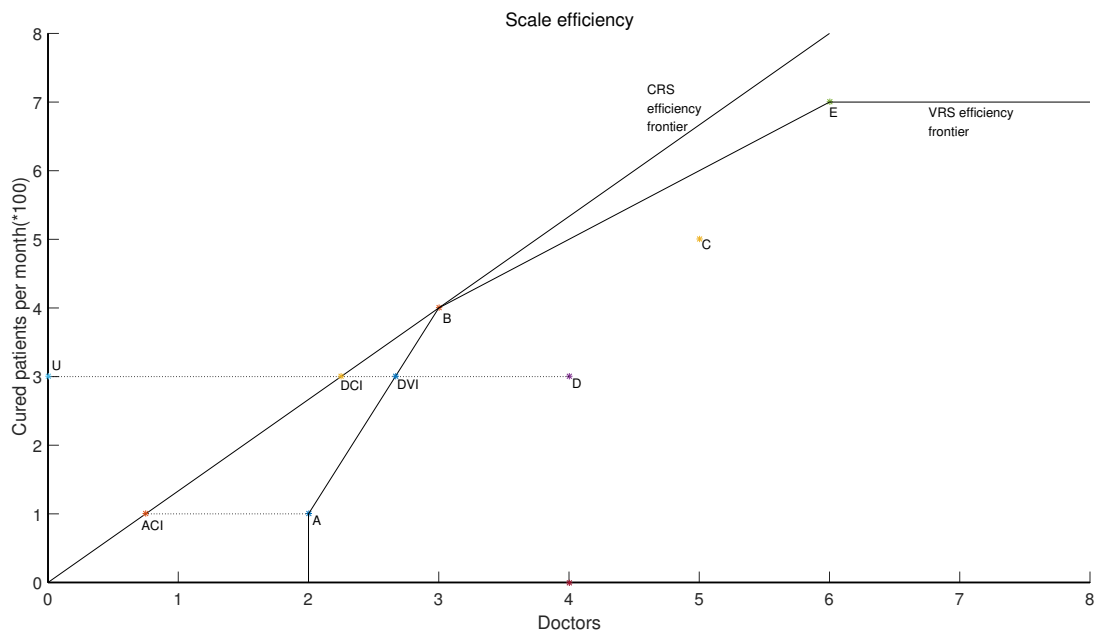


Figure A.3: Scale efficiency

From Figure A.3 we could see that the margin between VRS and CRS efficiency frontier could be measured by the scale efficiency and thus for units that are not CRS efficient they need to modify their size in order to become CRS efficient, thus another question could be asked that is how to modify their size. In this paragraph we still focus on input-oriented efficiency, Figure A.4 shows different natures of return to scale of different units, point *B* is located on both CRS and VRS efficiency frontier, thus for those who are below point *B* (point *A* and point *DVI*), we can see that the input efficiency at point

A is lower than that at point $DVI(0.5 < 1.12)$, thus we say that for all units that are in this area, their production would increase with their size, indicating that they haven't yet reach their optimal size, hence for units that are in this area, they need to increase their production, this is called increasing return to scale(IRS); for units who are above point B , like point CVI and point E , we see that their input efficiency decreases with their size($1.25 > 1.17$), thus for all units that are in this area, their production would decrease with their size, indicating that they have already passed their optimal size, hence they need to decrease their production, this is called decreasing return to scale(DRS).

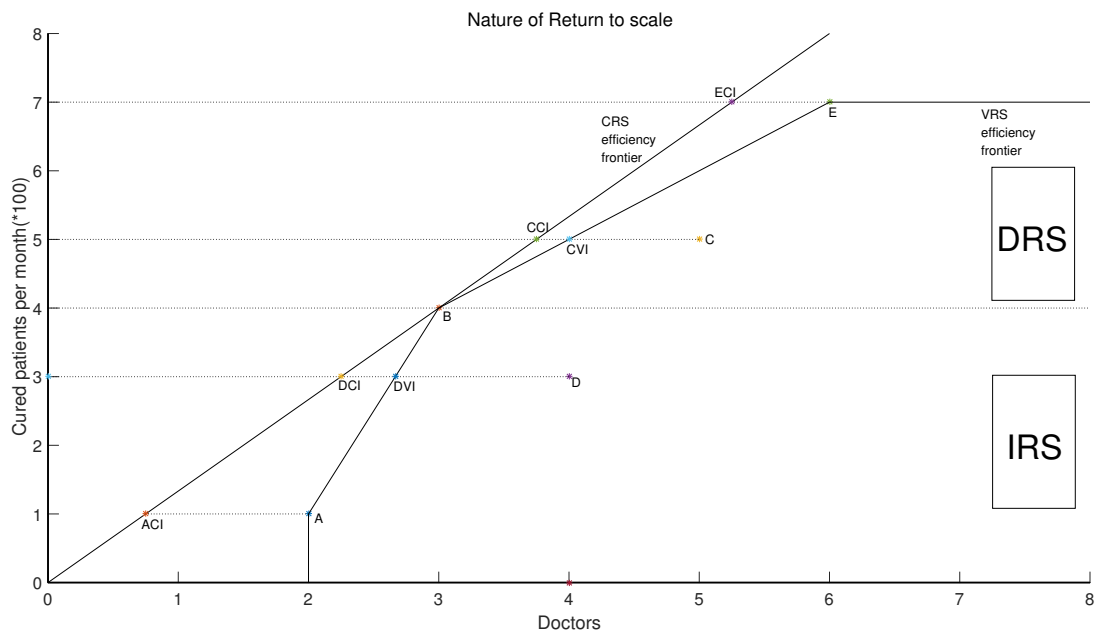


Figure A.4: Nature of Return to scale

Last point we want to show with our example is that for an inefficient unit how it could choose its *pairs*, as we mentioned above, for an inefficient unit, its *pairs* is the reference set of efficient unit(s) who are the closest to its production function, we are going to use a figure to simply illustrate the strategy, since in our example only under VRS there are multiple efficient units, thus we are going to show the choose strategy under VRS efficiency frontier.

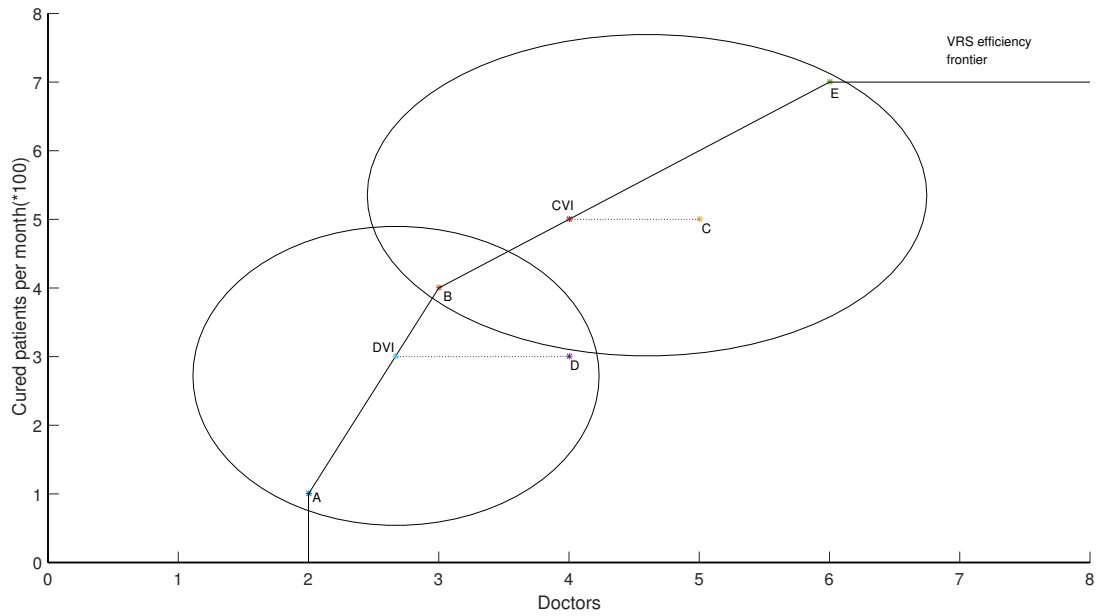


Figure A.5: Choose strategy

Figure [A.5](#) shows under VRS how to choose its *pairs* for clinic D and clinic C in terms of input-oriented efficiency, we can see that for clinic D, the closest efficient units are A and B, thus we choose A and B as its *pairs*, for clinic C, the closest efficient units to its position are B and E, hence we choose B and E as its *pairs*.

Appendix B

The DEA methodology

The precedent fictive example gives a general idea of how the DEA methodology functions, now let us present the formal mathematical formulation of the DEA methodology: in general, there are 2 types of DEA models, a CRS DEA model and a VRS DEA model, the CRS version of DEA model is also frequently called as a CCR model since it was firstly introduced by Charnes, Cooper and Rhodes in 1978[9], the VRS version is usually called as a BCC model since it was presented by Banker, Charnes and Cooper in 1984[4].

B.1 The CCR model

B.1.1 The multiplier form

We will begin with the CCR model, suppose that we have n units and for each unit $k = 1, 2, \dots, n$, it produces S outputs y_{rk} with M inputs x_{ik} , for each input we assign a weight v_i and for each output we also assign a weight u_r , denote h_k the efficiency score for unit k , the objection of DEA is to find the optimal weight solution (u^*, v^*) such that

the ratio of the weighted output over the weighted input is maximized for DMU_k

$$h_k = \max \frac{\sum_{r=1}^S u_r y_{rk}}{\sum_{i=1}^M v_i x_{ik}} \quad (\text{B.1})$$

subject to

$$\begin{cases} \frac{u_1 y_{1j} + u_2 y_{2j} + \dots + u_S y_{Sj}}{v_1 x_{1j} + v_2 x_{2j} + \dots + v_M x_{Mj}} \leq 1, & j = 1, 2, \dots, n \\ v_1, \dots, v_M \geq 0 \\ u_1, \dots, u_S \geq 0 \end{cases}$$

the problem of this formulation is that first of all the solution of this problem seems to be ambiguous, and second of all, even if was to exist an optimal solution (u^*, v^*) , then because of its fractional programming form, $(\alpha u^*, \alpha v^*)$ would also be optimal for $\alpha > 0$, thus the number of solutions would be infinite. To avoid this problem, Charnes, Cooper and Rhodes propose to transform the model from the fractional programming form to the linear programming form, the transformation is taken place by imposing that for unit k , either $\sum_{r=1}^S u_r y_{rk} = 1$ or $\sum_{i=1}^M v_i x_{ik} = 1$, the choice depends whether it's to measure the input-oriented efficiency or the output-oriented efficiency.

Output-oriented efficiency To measure the output-oriented efficiency, we would want to minimise the input expensed per unit output produced, thus, with unit's current quantity of output, maximum output efficiency could be achieved. The linear programming form of the model is written as

$$\min \sum_{i=1}^M v_i x_{ik} \quad (\text{B.2})$$

subject to

$$\begin{cases} \sum_{i=1}^M v_i x_{ij} - \sum_{r=1}^S u_r y_{rj} \geq 0 & j = 1, 2, \dots, n \\ \sum_{r=1}^S u_r y_{rk} = 1 \\ v_1, \dots, v_M \geq 0 \\ u_1, \dots, u_S \geq 0 \end{cases}$$

this form of the CCR model shows how unit k should minimize its inputs per unit produced¹, exactly as what we did in our example when we were to calculate the output efficiency of clinic A under CRS, let us show how this form and our example are connected: in our example, there is only one input and one output and we have only 5 units, thus the linear programming form of our example could be written as

$$\min vx_k$$

subject to

$$\left\{ \begin{array}{l} vx_j - uy_j \geq 0 \quad j = 1, 2, \dots, 5 \\ uy_k = 1 \\ v \geq 0 \\ u \geq 0 \end{array} \right.$$

Substitute $uy_k = 1$ into $vx_j - uy_j \geq 0 \quad j = 1, 2, \dots, 5$, we can have all the possible values of v

$$\left\{ \begin{array}{l} v \geq \frac{y_1}{x_1} \times \frac{1}{y_k} \\ v \geq \frac{y_2}{x_2} \times \frac{1}{y_k} \\ v \geq \frac{y_3}{x_3} \times \frac{1}{y_k} \\ v \geq \frac{y_4}{x_4} \times \frac{1}{y_k} \\ v \geq \frac{y_5}{x_5} \times \frac{1}{y_k} \end{array} \right.$$

we want to use this model to calculate the input-oriented efficiency of clinic A, note that clinic A has 2 doctors that can cure 1($\times 100$)patients per month, therefore we can write

$$\min v \times 2$$

¹In fact, it's actually close to the notion of 'marginal cost' in economics.

subject to

$$\begin{cases} vx_j - uy_j \geq 0 & j = 1, 2, \dots, 5 \\ u \times 1 = 1 \\ v \geq 0 \\ u \geq 0 \end{cases}$$

since we know exactly $\frac{y_j}{x_j}$ for $j = 1, 2, \dots, 5$ and also y_k , we can thus calculate all the possible values of $V = [\frac{1}{2}, \frac{4}{3}, 1, \frac{3}{4}, \frac{7}{6}]$, then we need to test all these values to see if they can satisfy that $vx_j - uy_j \geq 0$ for $j = 1, 2, \dots, 5$, after testing, it turns out that this constraint can be satisfied only when $v = \frac{4}{3}$, this also corresponds to what shows by Figure [A.1](#) that under CRS only clinic B is considered as efficient. Therefore we have v and since we know already u , we can calculate clinic A's output-oriented efficiency by $\frac{1}{\frac{4}{3} \times 2} = \frac{3}{8}$, which is equal to the result we showed above.

This form is also called as the 'multiplier form', and since it's a form of linear programming, it should also have its dual, which we present in the next section.

This is how we measure unit's output-oriented efficiency using the DEA methodology, it's called as the output-oriented CCR model because of its envelopment form. We now continue to show how to measure unit's input-oriented efficiency, which is rather straightforward since it follows the very same idea as the output-oriented efficiency, the only difference is that when we are to measure unit's input-oriented efficiency, we hold unit's weighted input constant.

Input-oriented efficiency To measure unit k 's input-oriented efficiency, we wish to maximise unit's output per unit input expensed, the multiplier form of the CCR model could be written as

$$\max \sum_{r=1}^S u_r y_{rk} \tag{B.3}$$

subject to

$$\left\{ \begin{array}{l} \sum_{i=1}^M v_i x_{ij} - \sum_{r=1}^S u_r y_{rj} \geq 0 \quad j = 1, 2, \dots, n \\ \sum_{i=1}^M v_i x_{ik} = 1 \\ v_1, \dots, v_M \geq 0 \\ u_1, \dots, u_S \geq 0 \end{array} \right.$$

To see how this form could be used in practice to measure the input efficiency, let us use again our fictive example: first rewrite the model with 5 units and only 1 input and 1 output, we get

$$\max u y_k$$

subject to

$$\left\{ \begin{array}{l} v x_j - u y_j \geq 0 \quad j = 1, 2, \dots, 5 \\ v x_k = 1 \\ v \geq 0 \\ u \geq 0 \end{array} \right.$$

Suppose that we want to measure the input efficiency of clinic A, note that thus $x_k = 2$ and $y_k = 1$, therefore

$$\max u \times 1$$

subject to

$$\left\{ \begin{array}{l} v x_j - u y_j \geq 0 \quad j = 1, 2, \dots, 5 \\ v \times 2 = 1 \\ v \geq 0 \\ u \geq 0 \end{array} \right.$$

and we can get all the possible values of u by

$$\left\{ \begin{array}{l} u \leq \frac{x_1}{y_1} \times \frac{1}{2} \\ u \leq \frac{x_2}{y_2} \times \frac{1}{2} \\ u \leq \frac{x_3}{y_3} \times \frac{1}{2} \\ u \leq \frac{x_4}{y_4} \times \frac{1}{2} \\ u \leq \frac{x_5}{y_5} \times \frac{1}{2} \end{array} \right.$$

since we know $\frac{x_j}{y_j}$ for $j = 1, 2, \dots, 5$, hence we can obtain all the possible values of $U = [1, \frac{3}{8}, \frac{1}{2}, \frac{2}{3}, \frac{3}{7}]$, obviously only $u = \frac{3}{8}$ can satisfy the constraints, thus the optimal solution of u is $u^* = \frac{3}{8}$. Therefore we can calculate clinic A's input-oriented efficiency score by $\frac{\frac{3}{8} \times 1}{1} = \frac{3}{8}$, which equals to what we have obtained before.

The multiplier form of the model shows how to maximise k 's outputs per unit of input expended. Its dual, the envelopment form, we shall also present it in the next subsection.

B.1.2 The envelopment form

Output-oriented efficiency If output-oriented, the dual of the multiplier form could be written as

$$\max \phi_k \tag{B.4}$$

subject to

$$\left\{ \begin{array}{l} \phi_k y_{rk} - \sum_{j=1}^n \lambda_j y_{rj} \leq 0 \quad r = 1, \dots, S \\ x_{ik} - \sum_{j=1}^n \lambda_j x_{ij} \geq 0 \quad i = 1, \dots, M \\ \lambda_j \geq 0 \quad \forall j = 1, \dots, n \end{array} \right.$$

where we want to maximise ϕ_k but in fact is $\frac{1}{\phi_k}$ measures the output-oriented efficiency of unit k , λ_j represents the weights of unit j . This form is also called as the 'envelopment form' and it's actually preferred in terms of compilation because it only has $M + S$

constraints instead of $n + 1$ constraint as in the ‘multiplier form’.

Up to now, what we present seems to be not connected with the notion *pairs*, so next we are going to show how this ‘envelopment form’ could be connected with this notion. In order to simplify the interpretation, let us suppose that for now there is only 1 input, therefore the ‘envelopment form’ could be written as

$$\max \phi_k \tag{B.5}$$

subject to

$$\begin{cases} \phi_k y_k - \sum_{j=1}^n \lambda_j y_j \leq 0 \\ x_{ik} - \sum_{j=1}^n \lambda_j x_{ij} \geq 0 \quad i = 1, \dots, M \\ \lambda_j \geq 0 \quad \forall j = 1, \dots, n \end{cases}$$

now make the scaling change of variable $\lambda_j = \frac{\lambda'_j}{y_j}$ and $\phi_k = \frac{\phi'_k}{y_k}$, then drop the primes and also the constant $\frac{1}{y_k}$, the model could be written as

$$\max \phi_k = \sum_{j=1}^n \lambda_j \tag{B.6}$$

subject to

$$\begin{cases} \frac{x_{ik}}{y_k} - \sum_{j=1}^n \frac{x_{ij}}{y_j} \lambda_j \geq 0 \quad i = 1, \dots, M \\ \lambda_j \geq 0 \quad \forall j = 1, \dots, n \end{cases}$$

thus we can see the ‘envelopment form’ is transformed into a form in which we seek to maximise the sum of weights of all units. Charnes, Cooper and Rhodes(1978)[\[9\]](#) propose this transformation in order to make contact with the work of M.J.Farrel, whose model could be presented as

$$\max \phi_k = \sum_{j=1}^n \lambda_j \tag{B.7}$$

subject to

$$\begin{cases} \sum_{j=1}^n P_j \lambda_j \leq P_k \\ \lambda_j \geq 0 \quad \forall j = 1, \dots, n \end{cases}$$

where $P_j = \sum_i \frac{x_{ij}}{y_j}$ and $P_k = \sum_i \frac{x_{ik}}{y_k}$

This is the example illustrated by Charnes, Cooper and Rhodes(1978)[9], in order to explain its meaning, Charnes, Cooper and Rhodes(1978)[9] use the following figure

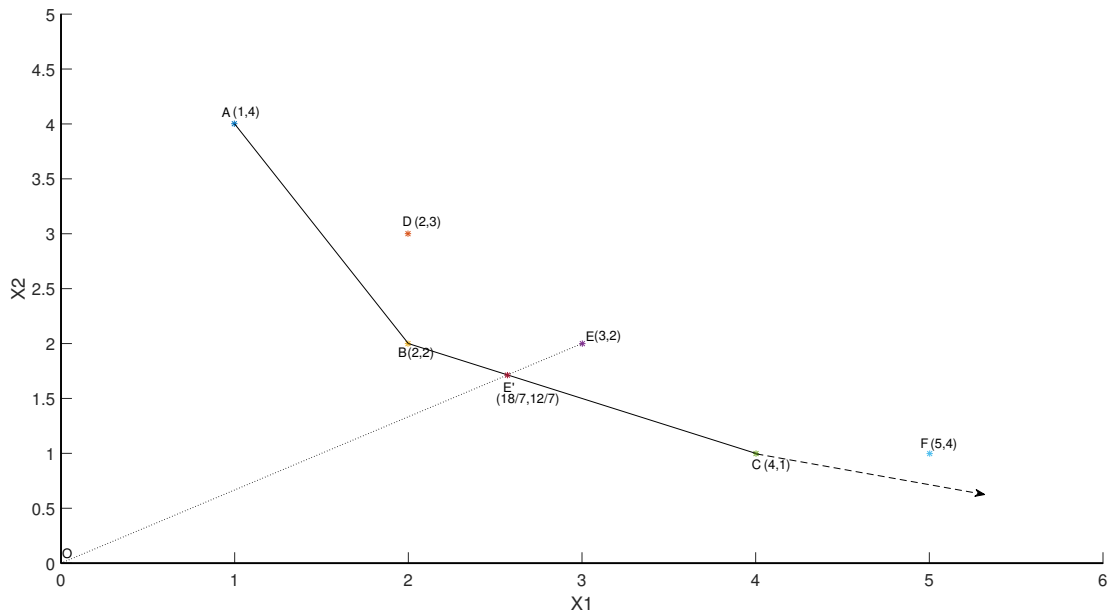


Figure B.1: Envelopment form

Figure B.1 shows a simple example in which there are 6 units, for each unit it has 2 inputs(x_1 and x_2) and 1 output(y), Figure B.1 shows the input efficiency of each unit per unit of y produced, for example, unit E needs 4 units of x_1 and 1 unit of x_2 to produce 1 unit of y , the efficient frontier is composed of line AB and BC and obviously unit E is not on the frontier, in addition, we can also see that for unit E, the efficient units that are closest to it are unit B and unit C, thus for unit E, its *pairs* should be composed of B and E. The output-oriented efficiency score of unit E could be therefore obtained by the performance of unit B and of unit E with any adjacent extreme point method(such

as the simplex or dual methods),

$$\frac{5}{6}B + \frac{1}{3}C = E \quad (\text{B.8})$$

i.e. the weight of unit B is $\lambda_B = \frac{5}{6}$ and the weight of unit E is $\lambda_E = \frac{1}{3}$, according to Eq [B.6](#), $\phi_E = \lambda_B + \lambda_E = \frac{5}{6} + \frac{1}{3} = \frac{7}{6}$, since unit k 's output-oriented efficiency score equals to $\frac{1}{\phi_k}$, thus for unit E, its output-efficiency score equals to $\frac{6}{7}$, indicating that unit E should have been able to produce 1 unit of output with only $\frac{6}{7}$ of the amount of x_1 and x_2 it expensed.

Point E' is the projection of E on the efficiency frontier, it shows the amount of x_1 and x_2 unit E should have expensed to produce 1 unit of output in order to be considered as efficient, to obtain E' , we can either calculate the intersection between line BC and line OE , or we can convexify Eq [B.8](#), which we do by dividing Eq [B.8](#) by ϕ_E , to get

$$\frac{5}{7}B + \frac{2}{7}C = E' = \frac{6}{7}E \quad (\text{B.9})$$

thus the output-efficiency score of unit E is indeed $\frac{6}{7}$.

We can apply the same method to calculate the output-efficiency score of all inefficient units, for example, unit D, in fact, by compare unit D's input efficiency to the input efficiencies of its *pairs* A and B, we can obtain its input-oriented efficiency score which actually also equals to $\frac{6}{7}$, but this doesn't necessarily mean that unit D and unit E have the same input-oriented efficiency, because though they do have the same score, their *pairs* are different: for unit E, its reference set is composed of B and C, yet for unit D, its reference set is composed of A and B.

Input-oriented efficiency If input-oriented, the envelopment form could be written as

$$\min \theta_k \tag{B.10}$$

subject to

$$\begin{cases} y_k - \sum_{j=1}^n \lambda_j y_j \leq 0 \\ \theta_k x_{ik} - \sum_{j=1}^n \lambda_j x_{ij} \geq 0 \quad i = 1, \dots, M \\ \lambda_j \geq 0 \quad \forall j = 1, \dots, n \end{cases}$$

where θ_k measure unit k 's input-oriented efficiency score.

B.2 The BCC model

B.2.1 The multiplier form

Now we proceed to the VRS version or the BCC model, as what we have shown above, the biggest difference between the CCR model and the BCC model is that in a BCC model the efficiency frontier is no longer a straight line, the difference between a CCR model and a BCC model is measured by the scale efficiency. The setting is pretty much the same as in the CCR model, we still suppose n units that each of which produces S outputs y_{rk} with M inputs x_{ik} and each output y_{rk} and each input x_{ik} are assigned a weight u_r and v_i respectively, note that in the BCC model, an unit's input efficiency would be different from its output efficiency, thus we will present the problem separately.

Output-oriented efficiency In the BCC model, if it was to measure unit k 's output efficiency score, the problem could be written as

$$h_k = \max \frac{\sum_{r=1}^S u_r y_{rk}}{\sum_{i=1}^M v_i x_{ik} - c_k} \tag{B.11}$$

subject to

$$\left\{ \begin{array}{l} \frac{\sum_{r=1}^S u_r y_{rj}}{\sum_{i=1}^M v_i x_{ij} - c_k} \leq 1, \quad j = 1, 2, \dots, n \\ v_1, \dots, v_M \geq 0 \\ u_1, \dots, u_S \geq 0 \end{array} \right.$$

where c_k is free in sign, we will explain its role later.

The objective is always to maximise the ratio between the weighted output and the weighted input. Once again, the same problem of fractional programming form emerges, in order to avoid this problem, we need to transform this fractional programming form into a linear programming form, which could be written as

$$\min \sum_{i=1}^M v_i x_{ik} - c_k \tag{B.12}$$

subject to

$$\left\{ \begin{array}{l} \sum_{i=1}^M v_i x_{ij} - \sum_{r=1}^S u_r y_{rj} - c_k \geq 0 \quad j = 1, 2, \dots, n \\ \sum_{r=1}^S u_r y_{rk} = 1 \\ v_1, \dots, v_M \geq 0 \\ u_1, \dots, u_S \geq 0 \end{array} \right.$$

this is essentially the same thing as what we have shown in the CCR model, except the term c_k , now let us use our fictive example to show how to use this form to calculate the output-oriented efficiency score of unit k . Since in our example there is only 1 input and 1 output, thus the model above could be rewritten as

$$\min vx_k - c_k$$

subject to

$$\begin{cases} vx_j - uy_j - c_k \geq 0 & j = 1, 2, \dots, 5 \\ uy_k = 1 \\ v \geq 0 \\ u \geq 0 \end{cases}$$

substitute $uy_k = 1$ into $vx_j - uy_j - c_k \geq 0 \quad j = 1, 2, \dots, 5$ we get

$$\begin{cases} vx_1 - c_k \geq \frac{y_1}{y_k} \\ vx_2 - c_k \geq \frac{y_2}{y_k} \\ vx_3 - c_k \geq \frac{y_3}{y_k} \\ vx_4 - c_k \geq \frac{y_4}{y_k} \\ vx_5 - c_k \geq \frac{y_5}{y_k} \end{cases}$$

Suppose that we want to calculate the output efficiency score of clinic D, recall that clinic D has 4 doctors and cure 3($\times 100$)patients per month, thus $x_k = 4$ and $y_k = 3$, thus the problem could be written as

$$\min v \times 4 - c_k$$

subject to

$$\begin{cases} vx_j - uy_j - c_k \geq 0 & j = 1, 2, \dots, 5 \\ u \times 3 = 1 \\ v \geq 0 \\ u \geq 0 \end{cases}$$

since we know $\frac{y_j}{y_k}$ for $j = 1, 2, \dots, 5$ and also x_j for $j = 1, 2, \dots, 5$, thus we can obtain all

the possible values of v and c_k by

$$\left\{ \begin{array}{l} 2v - c_k \geq \frac{1}{3} \\ 3v - c_k \geq \frac{4}{3} \\ 5v - c_k \geq \frac{5}{3} \\ 4v - c_k \geq \frac{3}{3} \\ 6v - c_k \geq \frac{7}{3} \end{array} \right.$$

Resolving these inequations, we can have all the possible values of v and c_k that satisfy the constraints, which are summarized in the table below

		$4v - c_k$
$v = \frac{2}{3}$	$c_k = \frac{2}{3}$	2
$v = \frac{1}{3}$	$c_k = -\frac{1}{3}$	$\frac{5}{3}$
$v = \frac{1}{2}$	$c_k = -\frac{1}{6}$	$\frac{11}{6}$
$v = \frac{1}{6}$	$c_k = -\frac{4}{3}$	2
$v = \frac{4}{9}$	$c_k = 0$	$\frac{16}{9}$
$v = 1$	$c_k = -\frac{5}{3}$	$\frac{7}{3}$

It's obvious that the optimal solution for c_k and v are $v^* = \frac{1}{3}$ and $c_k^* = -\frac{1}{3}$. Thus we can calculate the output efficiency score of clinic D by $h = \frac{1}{\frac{1}{3}} = \frac{3}{5}$, which equals to the result we have obtained before.

Note that the optimal solution gives us $c_k = -\frac{1}{3}$. In this linear programming form c_k indicates the nature of return of scale of unit k , the rules are:

- if $c_k > 0$, then unit k is of nature IRS;
- if $c_k = 0$, then unit k is of nature CRS;
- if $c_k < 0$, then unit k is of nature DRS

in our case c_k is negative, therefore clinic D is DRS in terms of output efficiency.

Like in the CCR model, in the BCC model, this linear programming form is also called as the multiplier form, and its dual is also called as the envelopment form, which we present in detail in the next section.

Input-oriented efficiency Similar to the output-oriented efficiency, in the BCC model, the input-oriented efficiency problem could be presented as

$$h_k = \max \frac{\sum_{r=1}^S u_r y_{rk} + c_k}{\sum_{i=1}^M v_i x_{ik}} \quad (\text{B.13})$$

subject to

$$\begin{cases} \frac{\sum_{r=1}^S u_r y_{rj} + c_k}{\sum_{i=1}^M v_i x_{ij}} \leq 1, & j = 1, 2, \dots, n \\ v_1, \dots, v_M \geq 0 \\ u_1, \dots, u_S \geq 0 \end{cases}$$

where c_k is always free in sign.

The multiplier form of this problem could be written as

$$\max \sum_{r=1}^S u_r y_{rk} + c_k \quad (\text{B.14})$$

subject to

$$\begin{cases} \sum_{r=1}^S u_r y_{rj} - \sum_{i=1}^M v_i x_{ij} + c_k \leq 0 & j = 1, 2, \dots, n \\ \sum_{i=1}^M v_i x_{ik} = 1 \\ v_1, \dots, v_M \geq 0 \\ u_1, \dots, u_S \geq 0 \end{cases}$$

Now let us see how can we use this form in our example: with 1 input and 1 output, the model could be rewritten as

$$\max u y_k + c_k$$

subject to

$$\begin{cases} uy_j - vx_j + c_k \leq 0 & j = 1, 2, \dots, 5 \\ uy_k = 1 \\ v \geq 0 \\ u \geq 0 \end{cases}$$

substitute $uy_k = 1$ into $uy_j - vx_j + c_k \leq 0 \quad j = 1, 2, \dots, 5$ we get

$$\begin{cases} uy_1 + c_k \leq \frac{x_1}{x_k} \\ uy_2 + c_k \leq \frac{x_2}{x_k} \\ uy_3 + c_k \leq \frac{x_3}{x_k} \\ uy_4 + c_k \leq \frac{x_4}{x_k} \\ uy_5 + c_k \leq \frac{x_5}{x_k} \end{cases}$$

We want to measure the input efficiency of clinic D, hence $x_k = 4$ and $y_k = 3$, we can write

$$\max u \times 3 + c_k$$

subject to

$$\begin{cases} uy_j - vx_j + c_k \leq 0 & j = 1, 2, \dots, 5 \\ v \times 4 = 1 \\ v \geq 0 \\ u \geq 0 \end{cases}$$

since we know $\frac{x_j}{x_k}$ for $j = 1, 2, \dots, 5$ and also y_j for $j = 1, 2, \dots, 5$, thus we can obtain all

the possible values of u and c_k by

$$\left\{ \begin{array}{l} u + c_k \leq \frac{2}{4} \\ 4u + c_k \leq \frac{3}{4} \\ 5u + c_k \leq \frac{5}{4} \\ 3u + c_k \leq \frac{4}{4} \\ 7u + c_k \leq \frac{7}{4} \end{array} \right.$$

Resolving these inequations, we can have all the possible values of u and c_k that satisfy the constraints, which are summarized in the table below

		$3u + c_k$
$u = \frac{1}{8}$	$c_k = \frac{1}{4}$	$\frac{5}{8}$
$u = \frac{1}{4}$	$c_k = -\frac{1}{4}$	$\frac{1}{2}$
$u = \frac{1}{6}$	$c_k = \frac{1}{12}$	$\frac{7}{12}$
$u = \frac{1}{2}$	$c_k = -2$	$-\frac{1}{2}$
$u = \frac{3}{16}$	$c_k = 0$	$\frac{9}{16}$
$u = \frac{1}{12}$	$c_k = \frac{5}{12}$	$\frac{2}{3}$

Obviously the optimal solution for u and c_k are $u^* = \frac{1}{12}$ and $c_k^* = \frac{5}{12}$. Thus the input-oriented efficiency of clinic D is $h = \frac{3 \times \frac{1}{12} + \frac{5}{12}}{1} = \frac{2}{3}$, which equals to the results we have obtained before.

Here we can also detect the nature of unit k 's return to scale by judging the sign of c_k , the rules are exactly the same as in the output efficiency case. In our example clinic D has a $c_k^* > 0$, thus it's of nature IRS in terms of input-oriented efficiency.

The dual of this form, the envelopment form, we present it in detail in the next section.

B.2.2 The envelopment form

Output-oriented efficiency In the BCC model, if output-oriented, the envelopment form could be written as

$$\max \phi_k \tag{B.15}$$

subject to

$$\begin{cases} \phi_k y_k - \sum_{j=1}^n \lambda_j y_j \leq 0 \\ x_{ik} - \sum_{j=1}^n \lambda_j x_{ij} \geq 0 \quad i = 1, \dots, M \\ \sum_{j=1}^n \lambda_j = 1 \\ \lambda_j \geq 0 \quad \forall j = 1, \dots, n \end{cases}$$

where we want to maximise ϕ_k but is $\frac{1}{\phi_k}$ that measures the output-oriented efficiency. We can also see that there is a new constraint $\sum_{j=1}^n \lambda_j = 1$, this is the convexity constraint.

Input-oriented efficiency If input-oriented, the envelopment form could be written as

$$\min \theta_k \tag{B.16}$$

subject to

$$\begin{cases} y_k - \sum_{j=1}^n \lambda_j y_j \leq 0 \\ \theta_k x_{ik} - \sum_{j=1}^n \lambda_j x_{ij} \geq 0 \quad i = 1, \dots, M \\ \sum_{j=1}^n \lambda_j = 1 \\ \lambda_j \geq 0 \quad \forall j = 1, \dots, n \end{cases}$$

where θ_k measures the input-oriented efficiency score of unit k , note that there is also the convexity constraint $\sum_{j=1}^n \lambda_j = 1$.

Appendix C

Input-oriented and output-oriented efficiency score

Country	CRS		Input oriented			Output oriented		
			VRS	Peers	Rank	VRS	Peers	Rank
Australia	AUT	0.76	0.76	IRL	16	0.986	JPN	14
Belgium	BEL	0.764	0.767	IRL	15	0.982	JPN	17
Canada	CAN	0.823	0.827	LUX/IRL	10	0.988	CHE/JPN	11
Denmark	DNK	0.679	0.684	IRL	20	0.979	JPN	19
Finland	FIN	0.954	0.956	IRL	6	0.994	CHE/JPN	7
France	FRA	0.694	0.741	CHE/LUX	17	0.992	JPN	9
Germany	DEU	0.71	0.711	IRL	18	0.985	JPN	16
Greece	GRC	0.952	0.954	IRL	7	0.994	LUX/CHE	7
Ireland	IRL	1	1	IRL	1	1	IRL	1
Italy	ITA	0.856	0.932	LUX/CHE	8	0.996	JPN/CHE	6
Japan	JPN	0.782	1	JPN	1	1	JPN	1
Luxembourg	LUX	0.968	1	LUX	1	1	LUX	1
Netherlands	NLD	0.682	0.69	LUX/IRL	19	0.987	JPN	13
Norway	NOR	0.789	0.802	LUX/IRL	13	0.988	CHE/JPN	11
Portugal	PRT	0.866	0.873	IRL	9	0.982	CHE/JPN	17
Spain	ESP	0.929	0.993	LUX/CHE	5	0.999	CHE/LUX	5
Sweden	SWE	0.757	0.805	LUX/CHE	12	0.991	JPN	10
Switzerland	CHE	0.884	1	CHE	1	1	CHE	1
United Kingdom	GBR	0.806	0.807	IRL	11	0.986	JPN/CHE	14
United states	USA	0.76	0.77	IRL	14	0.972	JPN	20

Table C.1: Input-oriented and output-oriented efficiency scores of OECD countries, Afonso and Kazemi (2017)[\[11\]](#)

Country	Efficiency score	Peers	Return to scale	Referencing frequency
Australia	0.8095	Chile; Iceland; Slovenia	Decreasing	0
Austria	0.5212	Chile; Estonia; Iceland	Decreasing	0
Belgium	0.8138	Chile; Korea; Slovenia	Decreasing	0
Canada	0.9835	Chile; Iceland; Ireland	Decreasing	0
Chile	1		Constant	17
Czech Republic	0.8311	Estonia Korea; Slovenia; Turkey		0
Denmark	0.7891	Chile; Iceland; Sweden	Increasing	0
Estonia	1		Constant	5
Finland	0.8242	Iceland; Korea; Slovenia	Increasing	0
France	0.7393	Chile; Korea; Slovenia	Decreasing	0
Germany	0.6099	Chile; Korea; Slovenia	Decreasing	0
Greece	0.8083	Estonia; Israel; Turkey	Decreasing	0
Hungary	0.7533	Estonia; Poland; Turkey	Increasing	0
Iceland	1		Constant	12
Ireland	1		Constant	4
Israel	1		Constant	3

(Continued)

Country	Efficiency score	Peers	Return to scale	Referencing frequency
Italy	0.8843	Chile; Iceland; Israel	Decreasing	0
Japan	1		Constant	0
Korea	1		Constant	7
Luxembourg	0.8008	Chile; Korea; Slovenia	Decreasing	0
Mexico	1		Constant	0
Netherlands	0.8354	Chile; Ireland; Slovenia	Decreasing	0
New Zealand	0.9031	Chile; Iceland; Sweden	Decreasing	0
Norway	0.8859	Chile; Iceland; Sweden	Increasing	0
Poland	1		Constant	1
Portugal	0.9329	Chile; Iceland; Israel	Increasing	0
Slovakia	0.7089	Estonia; Korea; Slovenia; Turkey	Increasing	0
Slovenia	1		Constant	11
Spain	0.9188	Chile; Iceland; Sweden	Decreasing	0
Sweden	1		Constant	5
Switzerland	0.6954	Chile; Ireland; Slovenia	Decreasing	0
Turkey	1		Constant	4
UK	0.9161	Chile; Iceland; Sweden	Decreasing	0
USA	0.7823	Chile; Iceland; Ireland;	Increasing	0

Table C.2: CCR efficiency scores of OECD countries, Cetin and Bahce (2016)[\[8\]](#)

Country	BCC	ARG	Country	BCC	ARG
Australia	0.996	0.776	Japan	1	0.992
Austria	0.98	0.769	South Korea	1	0.994
Belgium	0.972	0.699	Luxembourg	1	1
Canada	1	0.801	Mexico	1	1
Chile	1	1	Netherlands	0.981	0.771
Czech Republic	0.945	0.872	New Zealand	1	0.973
Denmark	0.965	0.756	Norway	0.992	0.929
Estonia	0.903	0.682	Poland	0.936	0.727
Finland	0.980	0.886	Portugal	0.974	0.955
France	0.987	0.771	Slovak Republic	0.916	0.627
Germany	0.975	0.727	Slovenia	0.983	0.975
Greece	1	1	Spain	1	1
Hungary	0.903	0.652	Sweden	1	1
Iceland	0.997	0.869	Switzerland	1	0.693
Ireland	0.978	0.728	Turkey	0.940	0.632
Israel	0.995	0.910	UK	0.982	0.786
Italy	1	0.934	USA	0.964	0.698

Table C.3: Output-oriented efficiency scores of OECD countries, 2008, Önen and Sayin (2018)[\[39\]](#)

Country	BCC	ARG	Country	BCC	ARG
Australia	0.995	0.673	Japan	1	1
Austria	0.978	0.730	South Korea	1	1
Belgium	0.976	0.611	Luxembourg	0.990	0.777
Canada	1	0.676	Mexico	1	1
Chile	1	1	Netherlands	0.981	0.5651
Czech Republic	0.945	0.794	New Zealand	0.993	0.575
Denmark	0.969	0.572	Norway	0.982	0.657
Estonia	0.928	0.756	Poland	0.950	0.730
Finland	0.976	0.751	Portugal	0.975	0.864
France	0.991	0.658	Slovak Republic	0.925	0.591
Germany	0.975	0.598	Slovenia	1	1
Greece	1	1	Spain	1	1
Hungary	0.913	0.623	Sweden	1	0.832
Iceland	1	1	Switzerland	0.997	0.523
Ireland	0.991	0.732	Turkey	1	1
Israel	1	0.970	UK	0.992	0.722
Italy	0.997	0.915	USA	0.966	0.565

Table C.4: Output-oriented efficiency scores of OECD countries, 2012, Önen and Sayin (2018)[\[39\]](#)

Appendix D

The Health Gradient

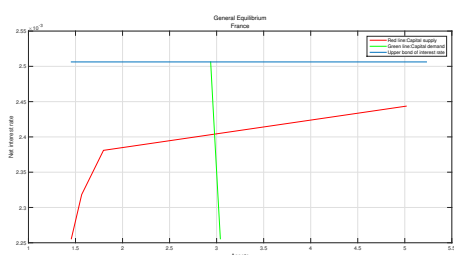
	France		US	
	model	data	model	data
$p^j(H = 1 I \in Q1)$	63.34%	65 %	54.31%	55.95%
$p^j(H = 1 I \in Q2)$	69.59%	68.9 %	63.69%	66.9 %
$p^j(H = 1 I \in Q3)$	80.28%	76.05 %	79.46%	74.8 %
$p^j(H = 1 I \in Q4)$	80.45%	80.925%	80.06%	82.5 %

Table D.1: Percentage of agents in good health by income quartiles - France and US

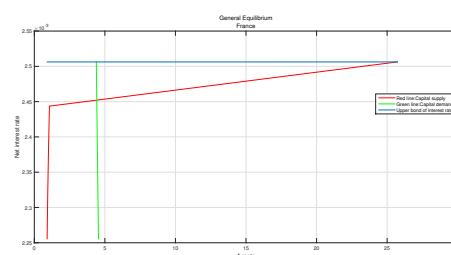
Appendix E

Equilibrium on financial markets

For each simulation we check if the equilibrium interest rate is below the time preference rate. The two panels of the Fig E.1 represent the general equilibrium point for the 2 benchmark simulations. In either case, the general equilibrium interest rate is well below the time preference rate.



(a) General Equilibrium for France



(b) Capital Distribution for USA

Figure E.1: General Equilibrium