Ca' Foscari University of Venice
Management Department

Master's Degree in
Accounting and Finance

Final Thesis

Economics and Health: The Behavioral Economics' Approach to Drug Addiction

Supervisor
Ch. Prof. Valeria Maggian

Graduand
Thomas Merotto

ID Number
853309

Academic Year 2018-2019

Contents

Introduction

How do individuals take decisions? For decades economics, psychology, sociology, political science, law and history's scholars have been trying to answer this question. In order to simplify things, neo-classical economics' theories and models embrace the assumption of perfect rationality of the individuals. Indeed, neo-classical economists believe that economic agents act according to well-defined preferences so that, in the canonical model of decision making under certainty, individuals choose the set of alternatives that better match their needs. However, these kinds of theories are quite far from reality and other possible economic assumptions (e.g., asymmetric information) are often not enough to justify possible irrational or not optimal decisions of individuals. Indeed, especially when dealing with complex decisions, people are not able to see the whole picture and they go for sub-optimal choices. Economic models based on perfect rationality may find some difficulties in explaining these kinds of economic behaviors. In particular, given the complexity of the medical knowledge and the possibility of myopic decisions of individuals, the health sector provides a fertile field for complex and difficult to handle issues. In spite of neo-classical economists recognizing the particular features of health economics, models based on perfect rationality are often not persuasive in explaining health related conditions such as obesity or addictions caused by irrational choices of individuals.

Nevertheless, Economics has evolved through times. Particularly, developments from cognitive and social psychology have been playing an important role in modifying the main economic thought. Starting from the 1950s, some economists questioned the assumption of perfect rationality. Psychology provided a vast set of new variables that could be easily mathematically implemented within the already existing neo-classical models. It is when economists started to recognize the role of psychology and tried to integrate its findings within economic theories that the branch, today known as behavioral economics, was born.

Moreover, more recent researches and findings from neuroscience have allowed to better investigate the role of brain in decision making processes. Neuroscience provides a scientific approach to psychology and has the potential to either confirm or neglect psychological theories. This, in turn, it may affect behavioral economics' theories. Starting from the late '90s, the economic branch called "neuroeconomics" began its development in order to integrate results from behavioral economics with the more recent findings from neurosciences.

It is important to note that neither behavioral economics nor neuroeconomics assert the incorrectness of the neo-classical theories. In fact, classical economic models can be used to potentially describe every situation from an economic point of view, and they are often a starting point for the development of behavioral economics theories. For instance, the Prospect Theory developed by Kahneman & Tversky (1979) derives from the Expected Utility Theory (e.g., Bernheim & Winston, 2013). In the same way, neuroeconomics does not state that behavioral economics is not useful but, on the contrary, it may represent its natural evolution and a further support for behavioral economics models when the psychologic assumptions at the base are confirmed by the neuroscience's research.

Hence, thanks to behavioral (and neuro) economics, economists have now a wider set of tools to better explain irrational and sub-optimal choices not being bounded to perfect rationality assumptions. The purpose of the thesis is to investigate and describe how individuals take decisions about an awkward but complex topic, namely, health. Particularly, the thesis faces the topic of addictions, especially with respect to those derived by the abuse of drugs. This work tries to identify justifications or explanations for possible irrational and dangerous decisions, yet always maintaining, when possible, an economic point of view. In fact, it almost impossible to deal with such topics without coming across medical and psychological implications. While medical involvement is minimized, given the nature of behavioral economics, very often it is somehow inevitable to bump into topics and theories that arise from psychology's literature. Furthermore, since neuroeconomics stems from neuroscience, which is a multidisciplinary discipline, some underlying insights about the brain need to be presented in order to fully understand the neuroeconomics theories at the base of drug abuse disorders. Moreover, analogies with other types of addictions such as alcoholism, smoking or compulsive gambling are often presented.

Having said that, the thesis is structured as follow: in the first chapter, an introduction of health economics and drug addiction is provided. Particularly, through the use of empirical evidence, it analyses why healthcare differs from other commodities and why health-related indicators are worth of attention from a macroeconomic perspective. Instead, the second part of the first chapter is dedicated to the drug addiction. First, some general concepts from psychology are provided, then the topic is analyzed by using economic tools. As we will see, the abuse of drugs is capable of generating relevant socio-economic costs both for its implications in health and criminality that require the intervention of the states and governments to be contained. Empirical data about the spread of drug abuse and drug-

related issues is also supplied. Furthermore, some insights about real-life policies against (and in support) of illegal drugs are described.

The second chapter deals with behavioral economics. First, the historical evolution from neo-classicism to behavioral economics is described. The theory of the rational customer at the base of the neo-classical approach is also summarized. Then, the main features at the base of behavioral economics are explained. The thesis continues by providing some real-life examples about how this newer approach to economics has been applied in order to correct some common health issues which involved wrong choices from individuals are also described. Finally, without going to much in depth, an introduction to neuroeconomics, particularly with respect to long-term decision-making topic, is presented.

The third chapter wants to further investigate those variables that are able to affect health decisional processes. For this purpose, some common theories and models from psychology and behavioral economics are described, often together with results from empirical researches aimed at analyzing their actual usefulness in changing and driving behaviors. Abstaining from being exhaustive given the vast literature, especially from psychology, of cognitive models and theories, the chapter deals with only with six of them. First, the health belief and transtheoretical models are presented, then the theories of fear appeal, nudging, framing effect, status quo bias are briefly described. As we will better see, the last two concepts, namely framing and status quo, both stem from the Prospect Theory developed by the psychologists Kahneman and Tversky (1979).

The fourth chapter somehow represents the core of the thesis. It describes and analyzes in detail the explanations from both behavioral and neuroeconomics about drug addiction. An early theory on addictions based on perfect rationality assumptions (i.e., the Model of Rational Addiction) is also discussed. Both neo-classical and behavioral economics make a large use of delay discounting processes to explain this kind of phenomenon. However, differences arise since classical theories adopt an exponential discounting rate while hyperbolic discounting rates are adopted in behavioral approaches. The main reason of this choice is that exponential discounting is not able to take into account the possibility of reversal in preference driven by impulsivity and loss of control, two kinds of behaviors very often exhibited by persons affected by a dependence. As it will be explained, these two psychological variables are at the base of behavioral economics theory on drug addiction. Furthermore, scientific insights and explanations from neuroeconomics are also reported in support to behavioral theories, also presenting data on other, but similar, kinds of addiction. Moreover, while presenting experiments on discounting rates, the chapter

supplies some insights about how economics can give its contribute in studying, preventing and treating the drug abuse.

Finally, the conclusions' section reports a summary of all fundamental results from the four chapters. Additionally, possible future directions in the development of the economic study of the drug dependence are supplied.

Chapter I: Health Economics and Drug Addiction

1.1 Health Economics: Definition and Features

One of the earliest definitions of health economist goes back up to 1958. According to this definition «Health economists are concerned with the organization of the market for health services and the net yield of investment in people for health. The "optimum" use of resources for the care of the sick and the promotion of health defines the special field of inquiry» (Mushkin, 1958, p.785). From this definition we can infer that health economics is a branch of economics whose purpose is to explain how health services are produced and provided and the drivers which lead personal' choices about health. Emphasis is put on the optimal allocation of resources. According to microeconomics theory, if no asymmetric information or externalities[1] are present, any competitive market would reach its optimal equilibrium where it is not possible to improve an individual's situation without making someone else worse off (Bernheim & Whinston, 2013). Healthcare can be supplied by both private and public agents; however, it represents a particular good since it provides challenges in terms of value, effectiveness, efficiency and ways of production and consumption that distinguish it from other commodities (e.g., food, clothing) especially if supplied by public organizations. Very often governments have an interest to intervene in public health in order to ensure a minimum level of health services to citizens and improve welfare. In addition to this, health indicators and diseases have proved to affect other economic variables such as growth. These effects will be investigated in the subchapters 1.2.1 and 1.2.2 of the thesis.

For many years, scholars did not treat medical-care market differently at all. Although a primitive literature was already present at the end of the '50s, the founder of Health Economics discipline is considered the American economist Kenneth J. Arrow. In his article "Uncertainty and the Welfare Economics of Medical Care", he identifies the main features of medical-care services. As asserted by Arrow (1963), the elements that characterize medical-care market with respect to classic goods and services are:

- The nature of demand: many casual factors in health cannot be foreseen in advance. For this reason, the demand of medical services is irregular and unpredictable. Excluding preventive services, individuals ask for healthcare when illness events occur. Some extreme events can also threat personal integrity with

---

[1] Asymmetric information: situation in which a party (typically the seller) owns a greater material knowledge than the other party (the buyer) involved in the same economic transaction.
Externality: cost or benefit caused by a third party who has no control over how that cost or benefit is produced (Bernheim & Whinston, 2013).

consequent risk of impairment or, in the worst case, death. Furthermore, there is also a potential risk for loss or reduction of the individuals' earning ability that can worsen if the cost of medical care is exceptionally high (criticality). These features are not unique of medical-care market: the unpredictable demand is shared with some legal services while food deprivation can threat personal integrity: However, in the last case, criticality can be avoided by governments ensuring a minimal income. Another element that can affect health services demand are externalities. They play an important role since some health status can be contagious. Individual decisions can threat the health situation of other persons. For example, we can think about a worker that decide to go to the office despite the flu. With this behavior he threats the health status of all his colleagues. Another example can be the recent phenomenon of no-vaxs (Bhattacharya *et al.*, 2014).

- Expected behaviour of physicians: medical care services belong to the category of commodities for which the product and the activity of production are identical. This means that the customer cannot try the product before purchasing it, so trust is involved in the process. Medical care sellers should have a "collectivity-orientation" that distinguishes them from the "profit-orientation" typical of business men. Several implications stem from collectivity-orientation: (1) competitive costs (i.e., advertising) are virtually eliminated among physicians; (2) self-interest is not involved as physician behaviour should be driven by a concern for individual's welfare; (3) medical treatments are usually required due to objective needs and should not be restricted by personal financial limitations despite the necessary allocation of resources required. Traditional human rights about an adequate medical care are important here; (4) heavy burden of asymmetric information. Physician is considered an expert in certifying the existence of injuries and illnesses for various legal and other purposes. When needed, his concern should be the correct convey of information rather than please his customer.

- Product uncertainty: most commodities offer the possibility of learning from personal experience or the one of others. In health case, this is generally not allowed. There is uncertainty about the quality of the product since recovery from disease is as unpredictable as its incidence. Severe cases further increase the degree of uncertainty, measured in term of utility variability. Asymmetric information plays a relevant role also in this case. Medical knowledge is complex, and the information owned by physicians about possible treatments and consequences is very much

greater than the one of patients. Both parties are aware of this inequality. Information asymmetries are almost always present in the seller-customer relationship, but it is difficult to identify another commodity where their presence is as big as in medical care case.

- Supply conditions: the supply of a commodity is linked to the expected return from its production. However, health represents an awkward theme and for this reason there are some departures from the classical commodity supply methods. In order to ensure a minimum level of quality, to enter in the profession licensing is required. Consequently, licensing restricts supply and increases the cost of medical care. A second aspect is linked to the high cost, especially in US, of medical education. These high costs are the result of a tight regulation on medical education which lead to a further limitation of physicians' supply.

- Pricing practices: the private medical sector applies an extensive price discrimination: hospitals charge rich individuals more than poor ones. The cost is almost zero for very low-income individuals. Arrows also states that there is opposition to prepayment and closed-panel practice from the side of hospitals. Closed-panel plan is a contractual arrangement that ties the patient to a specific group of physicians. However, Arrow provides no evidence of this point and is somehow questionable since a common health insurance offering is precisely that of binding the patient to hospitals or medical experts tied to the insurance groups.

The last point leads to two other problems linked to insurances, namely, Adverse Selection (selection of the worst) and Moral Hazard (opportunistic behavior). Both Adverse Selection and Moral Hazard occur when unequal information in consumer and producer can drive to undesirable output in the market. To explain these concepts, Akerlof (1970) uses the example of the so-called market of "Lemons". He considers the market of second-hand cars. Just like every market, it is composed by sellers and buyers. Sellers have full information about cars whereas buyers only know that inside the market are present good cars and bad cars (the "lemons"). Lemons have an intrinsic value (e.g., \$500) lower than good cars (e.g., \$1000). Not having further information, buyers assume the same likelihood for both buying a good car or a lemon. This implies that a rational buyer is prepared to pay $E(x) = \$1000 * P_{(GC)} + \$500 * P_{(L)} = \$750$, where P represent the probability of buying a good car or a lemon and it is 50% for both. Sellers of good cars, knowing that buyers are willing to pay only \$750, they are not going to sell the car and exit the market (Adverse Selection). Consequently, only lemons are still available, resulting in an

inefficient market (in a efficient market both good cars and lemons are present). The problem could be avoided if the sellers of bad cars would disclose more information, but they do not do that because they have an interest in hiding the real conditions of cars (Moral Hazard).

When dealing with insurance market, Adverse Selection inefficiency depends on the fact that the insurance company does not own full-information about the level of risk of its clients. Since the price of a policy should be equal to the expected value of the loss associated with the bad event (i.e. fair premium), it is fundamental for the insurance company to be able to identify the probability associated with the bad event. However, consumers have no incentive to reveal their personal data, on the contrary, they can decide to hide some relevant information in order to pay a lower price. If the insurer decides to raise the price of its policies what may happen is that clients with low likelihood of bumping into the event that leads to the refund will not subscribe the policy anymore. On the other hand, high-risk individuals will continue to subscribe it despite the higher price because risk averse. The result is that the company will have only high-risk clients requiring a higher refund in the case of event occurrence.

Moral Hazard generally concerns possible opportunistic behaviours of the insurance company in order to reduce the cost in case that the insured event happens. However, opportunistic behaviours can also stem from the insured client. It can be classified into two categories: *ex-ante*, which encompasses behaviours that make the insured event more likely and are often performed by the client (e.g., skipping the flu vaccine, driving fast, wrong diet); *ex-post*, when the behaviour changes after that the insured event happens and makes the recovery more expensive (e.g., expensive drug instead of cheaper remedies, unnecessary treatment) (Carande-Kulis *et al.*, 2007; Bhattacharya *et al.*, 2014).

Both Adverse Selection and Moral Hazard stem from the presence of asymmetric information in insurance markets that leads to a so-called market failure[2] where governments need to intervene. However, in a world where imperfect insurance and distortionary taxation are present, social insurance is a necessary part of redistribution because it permits to supplement the tax-transfer system. Additionally, in the healthcare sector and pensions area, private insurances show higher transaction costs than in the social insurance sector due to administrative costs that represent a further cause for the government intervention (Boadway *et al.*, 2006).

---

[2] Market failure: economic situation in which the optimal allocation of resources is not automatically reached through free market mechanisms, resulting in market inefficiencies (Gruber, J., 2013).

1.2 Health Impact on Macroeconomic Variables

*1.2.1 Life-expectancy and Economic Growth*

Health Economics and in general the medical-care market represents a complex object of study that should be approached with a different philosophy with respect to classical commodities. As asserted in the subchapter 1.1 of this thesis, the public utility, externalities, asymmetric information presence, and health insurances related issues represent *per se* justifications of the state intervention in health. An additional incentive for the state to interfere in this field is given by the way in which some health indicators may affect other macroeconomic variables. This section analysed the relationship between life-expectancy and growth.

The first thing we need to recognize is that healthcare is expensive. Developed countries spend each year a relevant percentage of their total GDP in health structures and research. Figure 1.1 and Figure 1.2 report data about health expenditure as percentage of GDP for European Union and United States respectively, in the period between 2005 and 2017. The growth (GDP) can be read on the left axis while the total expenditure on the right one. Both European Union and US show a similar trend for GDP which is steadily increasing, except in 2009, one year later the beginning of the financial crisis. If GDP growth is after all constant, health expenditure is more irregular and shows a fast increase between 2007 and 2009 in both EU and US. However, if in US health expenditure tends to remain constant or grow again in the years after 2009, in EU there is a downward tendency.

Another difference between US and EU is given by the percentage of expenditure with respect to GDP. Indeed, in 2017 US spent the 17.2% of its total GDP in health while the average in EU is just 8.60%. A possible explanation of this is the fact that, in EU case, we are dealing with averages that encompass countries sometimes very different among them. As was expectable, Western European countries tend to invest more in health with respect to Eastern countries. However, there is not any country in Europe that invest as much as US in health. The biggest European investor is represented by Switzerland that in 2017 spent in health the 12.7% of its total GDP, followed by France with 11.5%, and Germany with 11.3% (OECD, 2018). However, a bigger expenditure does not necessarily mean better health services, quite the contrary. Indeed, healthcare in US has always been a very much debated topic. The current American health system follows a model called "Fee for Service" where payments are based on the number of services provided. Each payment is unique so that services are billed and paid separately. This model well performs in

companies, but it has shown to be inefficient and ineffective if applied to healthcare, with a consequent raise of costs (Iglehart, 1992).

Figure 1.1: Health Expenditure in EU (2005-2017)
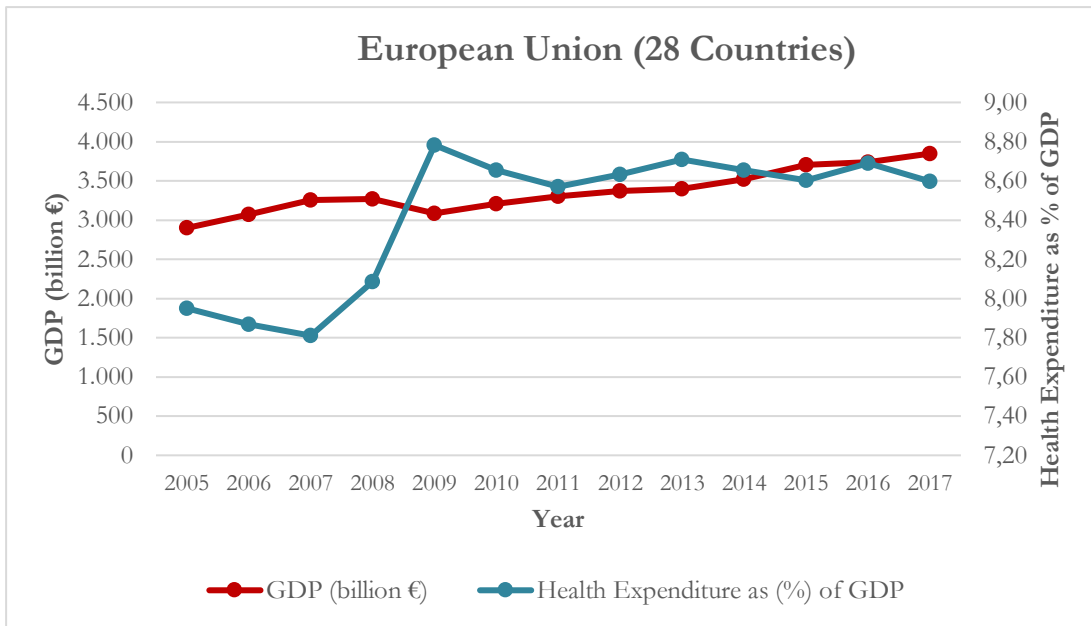
Source: OECD (2018)

Figure 1.2: Health Expenditure in US (2005-2017)
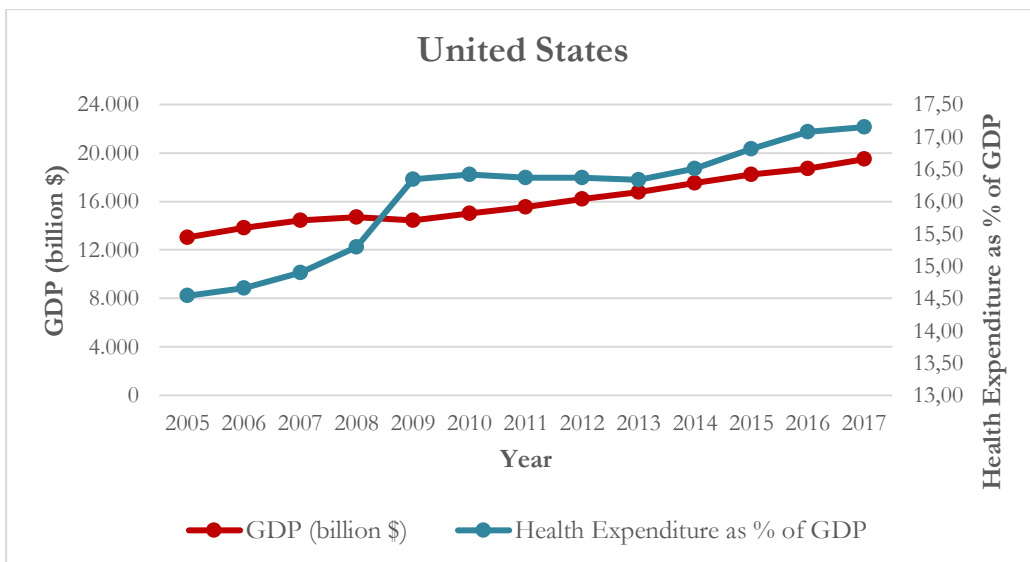
Source: OECD (2018)

Leaving aside the differences in health system among countries, we are interested in how different health indicators are able to affect growth. Generally, Economic growth (Y) depends on three main elements: Technical progress (A), Human and Physical Capitals (K) and, Labour (L).

$$Y = F(A, K, L)$$
$$Y = F\big(A(t), K(t), L(t)\big)$$
$$\frac{dY}{dt} = FA\frac{\partial A}{\partial t} + Fk\frac{\partial K}{\partial t} + FL\frac{\partial L}{\partial t}$$

Governments have an interest in investing in these inputs through growth-enhancing public expenditure which are usually funded by taxation (Baiardi *et al.*, 2018). While there is no doubt that medicine participates in an active way to the technical progress through the research and development of new drugs and treatments (whose extent mostly depends on the investments that pharmaceutical companies and states make in R&D), the impact on the two other variables may not be so immediate to get.

According to Swift (2010), health indicators play a role in determining both capital and labour. Longer life expectancy brings an increase in population, so to an increase in labour supply. However, possible troubles arise in the case that capital-to-labour ratios decrease due to population growth, namely, if labour grows less proportionally with respect to population. Additionally, Human capital is determined by the level of skills acquired by citizens and is influenced by some variables such as educational infrastructure, child nutrition and households' situation, including parents' physical health and cognitive attainment. Physical capital depends on savings whose amount is also affected by adult health.

A research conducted by Bhargava *et al.* (2001), based on data from the Penn World Table (PWT), and World Bank studies the effects that fertility rate, life expectancy and, population variables have on the GDP of countries. The probability of surviving the 60th birthday after reaching the age 15 years (ASR) has been used as a proxy of life-expectancy. Since data on ASR and fertility rate are gathered at irregular intervals, to reduce the impact of projections on the empirical results, they analysed panel data applying 5-year intervals. The analyses are based on 92 countries for PWD data and 73 for World Bank data.

Figure 1.3 shows the results of the regression analysis on PWT data. Two analyses have been conducted: under Specification 1, time varying variables are treated as exogenous while in Specification 2 lagged GDP is treated as a fully endogenous variable. However, according to the appropriate Chi-square statistic, the exogeneity null hypothesis for investment/GDP can be accepted.

Figure 1.3: Coefficients for real per capita GDP growth rates 1965–1990 (PWT)

Variable	Specification 1	Specification 2
Constant	0.274 (0.041)	0.407 (0.069)
Tropics	−0.011 (0.004)	−0.012 (0.004)
Openness	0.026 (0.005)	0.028 (0.005)
Logarithm of fertility rate lagged 5 years	−0.015 (0.006)	−0.028 (0.008)
Logarithm of investment/GDP ratio lagged 5 years	0.014 (0.002)	0.014 (0.002)
Logarithm of adult survival rate lagged 5 years	0.181 (0.075)	0.358 (0.114)
Interaction between lagged adult survival rate and GDP	−0.024 (0.011)	−0.048 (0.016)
Logarithm of GDP lagged 5 years	−0.026 (0.004)	−0.041 (0.008)
GDP at which partial derivative of GDP growth rate with respect to lagged adult survival rate is zero	2123	1714
Chi-square (20) test for exogeneity of lagged GDP	63.89	
Chi-square (20) test for exogeneity of investment/GDP ratio	18.28	
Chi-square (15) test for exogeneity of means of lagged investment/ GDP ratio, ASR, and interaction of ASR and GDP	15.22	
Number of countries	92	92
Number of time observations	5	5

Source: Bhargava *et al.* (2001)

The results of the regression indicate that:

- The percentage area in the tropics for a country is negative and statistically significant.

- Openness of the economy is positively associated with growth rates;

- Fertility rate is negatively associated with GDP growth and statistically significant. A possible explanation is that high fertility rates are common in developing countries. However, if the demand of resources for health care and education grows equi-proportionately with the population the same cannot be said for the labour force. Additionally, unwanted fertility in developing countries negatively affects households' resource allocation impacting the physical capital;

- Investment/GDP ratio is statistically significant and positive meaning economic growth is affected by investments in physical capital;

- The ASR and the interaction between ASR and GDP are significant predictors of economic growth. ASR impact is particularly relevant for low-income countries but approaches the zero for high levels of GDP per capita (in specification 1 when the GDP per capita was 2123 in terms of 1985 international dollars);

The same analysis was conducted using World Bank data as reported in Figure 1.4.

Figure 1.4: Coefficients for real per capita GDP growth rates 1965–1990 (WDI)

Variable	Specification 1	Specification 2
Constant	0.220 (0.033)	0.310 (0.048)
Tropics	−0.014 (0.004)	−0.016 (0.005)
Openness	0.041 (0.006)	0.047 (0.007)
Logarithm of fertility rate lagged 5 years	−0.023 (0.007)	−0.034 (0.008)
Logarithm of investment/GDP ratio lagged 5 years	0.009 (0.003)	0.008 (0.003)
Logarithm of adult survival rate lagged 5 years	0.192 (0.061)	0.333 (0.084)
Interaction between lagged adult survival rate and GDP	−0.029 (0.010)	−0.052 (0.014)
Logarithm of GDP lagged 5 years	−0.022 (0.004)	−0.034 (0.006)
GDP at which partial derivative of GDP growth rate with respect to lagged adult survival rate is zero	684	580
Chi-square (20) test for exogeneity of lagged GDP	62.41	
Chi-square (20) test for exogeneity of investment/GDP ratio	13.23	
Chi-square (15) test for exogeneity of means of lagged investment/ GDP ration, ASR, and interaction of ASR and GDP	3.04	
Number of countries	73	73
Number of time observations	5	5

Source: Bhargava *et al.* (2001)

The results of Figure 1.4 are coherent with those found in the analysis reported in Figure 1.3. The only difference is in the level of GDP for which the effect of ASR approaches the zero (in specifications 1 when the per capita GDP is 684 in 1985 international dollars, so at a lower level compared with PWT data).

The most important finding of Bhargava *et al.* (2001) is that a better health situation of the country (approximated by life-expectancy indicators) has a positive effect on growth but this is relevant only for low-income countries.

This result is denied by Swift (2010). Swift gathered data for 13 OECD high-income countries in order to determine if there is a long-term relationship between health and GDP, and if this relationship has remained constant over time. Again, life expectancy is used as a proxy of the health status of the nation. A limitation of this indicator is that it suffers for the disadvantage of not including improvements (i.e., better nutrition) that may improve productivity but have little effect on the length of life. The most important result of Swift is a two-way relationship: not only better health can drive economic growth through an increase total GDP as population increases and long-term gains in physical and human capital (higher productivity) but, in turn, growths in GDP can lead to an increase in life-expectancy. A virtuous cycle is generated. In average an 1% increase in life expectancy leads to a 6% increase in total GDP in the long run, and 5% increase in GDP per capita. A possible shortcoming of this research is the absence of other measures of health, in particular, morbidity and mortality that may help explaining in a better way the relationship between GDP and health. A further limit is that differences among countries are not taken into account.

1.2.2 Diseases and Unemployment Rate

In this section, we will go more in detail analysing how certain diseases and health status may affect another macroeconomic variable, namely, the unemployment rate. As we already saw in subchapter 1.2.1, labour is one of the variables that determines the economic growth of a country. What we need to keep in mind is that a raise on the unemployment rate is likely to affect the economic growth in a negative way. The purpose of this section is not to state that diseases are detrimental for employment (which is quite intuitive), but to examine the extent of their impact. Since the diseases and health statuses that may potentially affect the unemployment rate are too many to be described in detail, here we focus only on two common health situations: obesity and diabetes.

Overweight and obesity are even a more relevant problem, it is estimated that since the 1980s the number of persons with overweight and obesity related problems has more than tripled in many European countries. In average the 7% of national health budgets across the EU are invested on diseases linked to obesity each year (European Commission, 2014). This data encouraged Europe to intervene directly through the "EU Action Plan on Childhood Obesity 2014-2020", an initiative that aims at stopping the growth in overweight and obesity in children between 0-18 years.

Besides possible health complications that may arise due to obesity, Morris (2007) states that there is a two-way relationship between obesity and employment because:

- Obesity causes unemployment because it might lead to a debilitating condition that lowers productivity. Additionally, prejudices and stereotypes cannot be neglected because they can result in discrimination toward the obese;

- Unemployment causes obesity since the low income of an unemployed person may spur him in consuming more cheap fattening food;

In his work, Morris investigates the impact of obesity on employment in England. Figure 1.5 shows the features of the sample analysed in the study.

Figure 1.5: Employment by obesity categories

Obesity category	BMI (kg/m²)	Males (n=8,324)		Females (n=8,643)	
		% Sample	% Employed	% Sample	% Employed
Underweight	<20	3	66	7	55
Healthy	20–25	34	78	43	69
Overweight	25–30	46	81	31	66
Obese	>30	17	74	19	62

Source: Morris (2006)

The samples consider 8,324 males and 8,643 females. BMI represents the Body Mass Index. In this case, a healthy person is considered healthy if its BMI goes from 20 to 25.

Just the 34% of the males and 43% of female are in the healthy category. Overweight represents the biggest category for men (46% while for women 31%), while obese are respectively 17% of men and 19% of women. Figure 1.5 also shows the percentage of each category that is employed. In the male case, the biggest proportion is represented by the overweight category with 81% of persons employed while for women is the healthy one (69%).

Other than obesity, several possible explanatory variables are included, grouped into four categories: educational variables, selected health variables, home and family variables and, additional variables (i.e., age, ethnic group). The results of Morris' analyses on the samples for both univariate and bivariate analysis are reported in Figure 1.6. For simplicity, only the results of the variable "obesity" on employment are reported.

Figure 1.6: The impact of obesity on employment

The impact of obesity on employment: univariate probit

	Males			Females		
	Coef.	z	ME	Coef.	z	ME
Males						
Obese	−0.113	−2.1	−0.021	0.015	0.4	0.004
Observations		8,324			8,643	
Pseudo-R^2		0.363			0.218	

The impact of obesity on employment: bivariate probit

	Males			Females		
	Coef.	z	ME	Coef.	z	ME
Impact of obesity on employment						
Obese	−0.420	−1.6	−0.084	−0.696	−2.9	−0.213
Observations	8,324			8,643		
ρ	0.172			0.400		
Wald test $\rho=0$ [p value]	$\chi^2(1)=1.6$ [0.21]			$\chi^2(1)=7.1$ [<0.01]		
	Coef.	z		Coef.	z	
Impact of the instrument on obesity						
Prevalence of obesity	3.643	7.4		3.584	7.6	
Observations	8,324			8,643		
Wald test instrument=0 [p value]	$\chi^2(1)=54.41$ [<0.01]			$\chi^2(1)=58.09$ [<0.01]		

Source: Morris (2006)

The result of the univariate probit shows how obesity has a negative and significative impact on employment for males. Obese men have an employment probability of 0.021 lower than non-obese men. On the contrary, obesity seems to have a positive, although insignificant, effect on the women employment (0.004 higher likelihood).

Bivariate probit confirms the negative and significant effect of obesity on male. Here the negative marginal effect is bigger than the previous one (now -0.084). The results are yet different for women. Now, obesity has a negative and significant impact, whit a marginal

effect of -0.213. ρ is positive indicating that there is a positive correlation between unexplained factors that affect obesity and unexplained factors that affect employment. In order to confirm that the hypothesis that $\rho = 0$ can be rejected, Morris used the Wald Test. In males' case, $\rho = 0$ cannot be rejected meaning that the endogeneity of obesity does not significantly affect the univariate probit estimates in males. In females' case, $\rho = 0$ is rejected. The univariate probit results are underestimated and biased the negative impact of obesity on employment.

The overall result of Morris' work is a negative impact of obesity on unemployment for both men and women. However, the further analyses on female employment indicates that there are other omitted variables that are correlated with obesity and employment (i.e., time preferences).

In a Morris' similar way, Tunceli *et al.* (2004) investigated the effect of diabetes on labour market outcomes. This study is somehow correlated with the previous since individuals affected by obesity problems are more likely to develop different forms of diabetes. Like obesity, diabetes is correlated to employment for three main reasons:

- Diabetes complications may prevent working entirely or lead to absenteeism;
- Productivity on work is reduced;
- Discriminations are possible especially due to hypoglycemia risk. Employers may restrict access to the jobs designated as safety sensitive. Further discriminations may stem from the lower productivity.

Tunceli' s study is based on secondary data gathered by the Health and Retirement Study in two waves (1992 and 1994) on a sample of 7,055 respondents between 51 and 61 years, where 490 of them reported to have diabetes problems in wave 1. Even here, different variables such as educational level, ethical group, age, civil status are considered. Most important is the inclusion of other health variables (i.e., BMI, number of other chronic conditions). The study analysed the impact of diabetes on three different labour outcomes, namely, employment, presence of work limitations and, number of work-loss days. The last two elements represent indicators of the work productivity. The main findings of the research are reported in Figure 1.7. Two models are reported: the first one is computed including just the sociodemographic and occupation variables, whereas the second model also takes into account BMI and the number of other chronic health conditions. Both results for men and women are reported.

Figure 1.7: Diabetes effect on employment and productivity

| | Employment | | Work productivity | | | |
| | | | Presence of work limitations | | No. of work-loss days | |
	I	II	III	IV	V	VI
Women						
Sample size	3,670	3,670	3,146	3,146	3,146	3,146
Variables						
Diabetes	−0.059 ± 0.026†	−0.044 ± 0.026	0.107 ± 0.030‡	0.060 ± 0.024†	3.107 ± 1.005‡	2.050 ± 1.017‡
BMI		0.014 ± 0.006†		−0.003 ± 0.005		0.197 ± 0.301
BMI²		−0.0002 ± 0.0001		0.0001 ± 0.0001		−0.001 ± 0.005
No. of other chronic conditions		−0.020 ± 0.006‡		0.046 ± 0.005‡		1.395 ± 0.271‡
Men						
Sample size	3,362	3,362	2,891	2,891	2,891	2,891
Variables						
Diabetes	−0.090 (0.025)‡	0.071 ± 0.025‡	0.104 ± 0.028‡	0.054 ± 0.023‡	1.662 ± 1.164	0.897 ± 1.180
BMI		0.011 ± 0.010		−0.012 ± 0.009		−0.534 ± 0.567
BMI²		−0.0001 ± 0.0001		0.0003 ± 0.0001		0.011 ± 0.010
No. of other chronic conditions	0.104 ± 0.028	−0.023 ± 0.006‡		0.050 ± 0.005‡		1.427 ± 0.359‡

Source: Tunceli *et al.* (2004)

Diabetes negatively affects employment. According to the first model, women with this disease have 5.9% probability less of working whereas men the 9% with respect to their counterparts without diabetes. The negative impact of diabetes is weaker if BMI and other chronic conditions are included in the model. The probabilities decrease to 4.4% for women and 7.1% for men. However, other chronic conditions contribute to lower the likelihood of employment (-2% for women and -2.3% for men).

About productivity, the effect of diabetes is smaller when other health variables are included. Women have 6% more of possibility to meet work limitations while men 5.4%. These results are significantly lower if compared with the first model (10.7% for women and 10.4% for men). Work-loss days is smaller for men than women. However, the results are relatively low for both. Women with diabetes have two more work-loss days per year with respect to the counterpart without diabetes while men almost one day.

Some limitations due to data quality are worth noting. First, data available within the HRS do not allow to be adjusted for the severity of diabetes (i.e., diabetes' complications are not considered). Second, information is self-reported and may be subject to error. The last limitation is represented by the fact that longitudinal data on employment status before a diagnosis of diabetes is not provided.

1.3 Drug Addiction

1.3.1 Addictions: A Common Ground

Before analysing in detail drug addiction, it can be useful knowing some general concepts from psychology. There is a vast literature of studies and experiments around addictions. This is not surprising if we think about how many kinds of dependencies exist: some are linked to the consume of a certain substances such as tobacco, alcohol and drug; others regard a specific behavior such as gambling or the compulsive buying disorder; some have sexual nature (i.e., pornography addition) or arise from other diseases (i.e. Bulimia nervosa). Technology related addictions also exist (i.e., internet addition disorder). Despite all these dependencies may be very different among them with respect to "object desired", socio-economic costs and impact on health, a common ground exists.

A general definition of addiction is provided by the Diagnostic and Statistical Manual of Mental Disorders (DMS) that reports: «The essential feature of this disorder is a cluster of cognitive, behavioral, and physiologic symptoms that indicate that the person has impaired control of psychoactive substance use and continues use of the substance despite adverse consequences» (Goodman, 1990, p.1404). The definition focuses on substances use. This because terms such as addiction, dependency and abuse have traditionally been associated to persons with substance disorders. However, it can also be applied to the other form of addictions since an impaired control due to cognitive, behavioral, and physiologic alterations and the recurrent behavior despite adverse consequences are usually shared by all types of dependencies.

An indirect definition of addiction states that, to be formally declared, a dependence should present the following characteristics:

- Recurrent incapacity to resist impulses toward a specified behavior;
- Sense of tension that tends to increase immediately before starting the behavior;
- Relief or pressure while engaging in the behavior;
- A sensation of lack of control while engaging in the behavior;
- At least five of the following attitudes: (1) recurrent preoccupation about the behavior or with preparatory activities to the behavior; (2) recurrent interest in the behavior to a greater extent or over a longer period than intended; (3) frequent efforts to stop, control, or reduce the behavior; (4) the individual spends a great deal of time in activities necessary for the behavior, engaging in the behavior or recovering from its effects; (5) the individual engages in the behavior frequently when he is expected to carry out occupational, domestic, academic, or

social obligations; (6) important social, occupational or recreational activities are reduced or given up due to the behavior; (7) continuation of the behavior despite the awareness of having a persistent or recurrent psychological, social, financial or physical problem that is caused or exacerbated by the behavior; (8) higher tolerance that leads the individual to increase the frequency or intensity of the behavior in order to achieve the desired effect or diminished effect with continued behavior of the same intensity; (9) irritability or restlessness if impossibility to engage in the behavior arises.

- Some symptoms of the disturbance have persisted or occurred repeatedly for one month, or a longer period of time;

### 1.3.2 Definition and Spread of Drug Addiction

Substances Use disorders are often included within the Mental Health disorders. Drug addiction can be defined as «a chronic, relapsing disorder in which compulsive drug-seeking and drug-taking behavior persists despite serious negative consequences. Addictive substances induce pleasant states (euphoria in the initiation phase) or relieve distress. Continued use induces adaptive changes in the central nervous system that lead to tolerance, physical dependence, sensitization, craving, and relapse» (Camì & Farré, 2003, p.975). The definition presents two adjectives: Chronic indicates a persisting, long-lasting condition; Relapsing indicates a morbus process that happens recurrently in the same individual, usually in an even more acute form. The second part of the definition declares that the use of substance leads to pleasant states and relief that may be identified as possible reasons that drive an individual to abuse of them and then list the main consequences in the long-term.

A distinction should be made between drug addition and drug abuse. If drug addition implies drug abuse the contrary is not necessarily true. It may happen that a person abuses of drugs one or more times without developing a dependence. In other words, the individual does not present the features described in subchapter 1.3.1 of this thesis. So, the term "drug abuse" is not enough to identify a dependence, for this reason we generally talk about "drug abuse disorders".

In 2016, the share of worldwide population with drug abuse disorders was the 0.85%, about 62 million people. Of this 62 million, two-third were men (40 million) and one-third women (20 million). Adults it their twenties have more likelihood to have drug disorders (IHME, 2017). All in all, these figures are relatively low if compared with other disorders or

diseases. For example, the worldwide average of smokers in 2016 was 21.9%. Always in the same year, adults who suffered of overweight and obesity problems were the 39% for both men and women (aged 18 and over) while children and adolescents (between 5 and 19 years) were the 18% (World Health Organization, 2018). It is then clear how drug problems involve a tiny share of global population. Figure 1.8 reports a snapshot of the drug disorders' situation in 2016.

**Figure 1.8**: Share of population with Drug Abuse Disorders (2016)
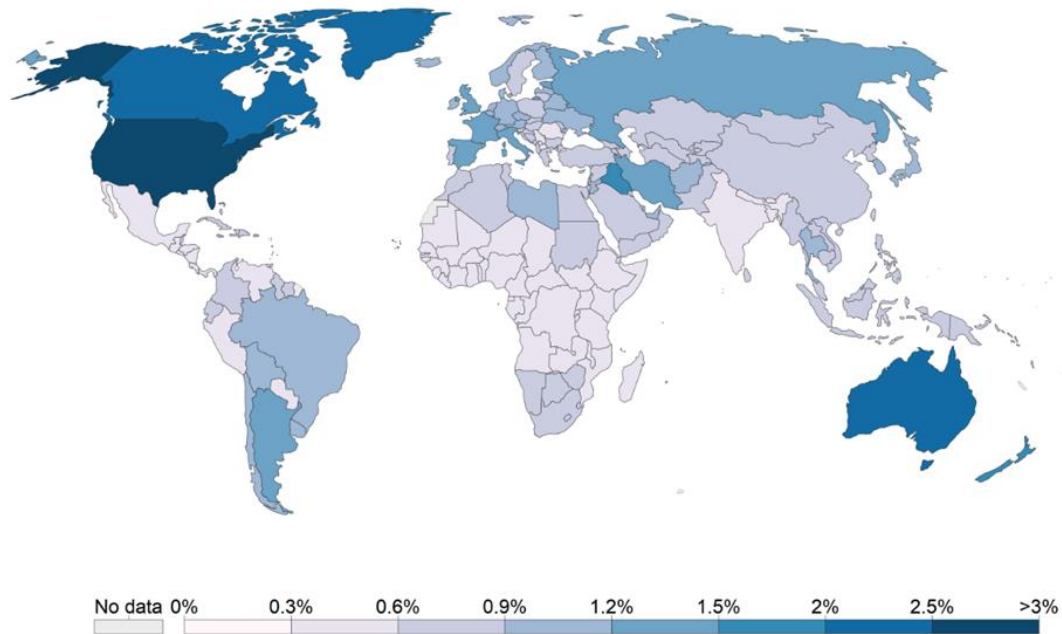
Source: IHME, 2017

As can be seen in the map, United States is the only country where drug addition problems involve more than the 3% of the population (3.31%). The second country for drug disorders is Australia (2.41%), followed by Canada (2.21%) and Iraq (1.57%). In Europe the situation is somehow less severe and there is not a big gap between Western Europe and Eastern Europe (1.16% against 1.15%). A similar situation is present in South America and Russia (1.25% for both) while drug disorders are less widespread in Africa and Asia. What is somehow surprising is that some countries traditionally linked to drug cartels such as Mexico and Colombia have lower rates than most developed countries (0.58% and 0.64% respectively).

Nevertheless, possible questions can arise about reliability of data from poor and developing countries. Very often data from developed and developing countries do not have the same quality and we may deal with estimates since the phenomenon is not clearly

observable or hidden (i.e., fake data) in some countries with a consequent lack of detailed and reliable data[3].

It could be interesting analysing historical data to identify if drug consumption has increased or decreased over time. Figure 1.9 shows how the number drug abuse disorders have changed over the last years.

Figure 1.9: Evolution of Drug Abuse Disorders (1990-2016)
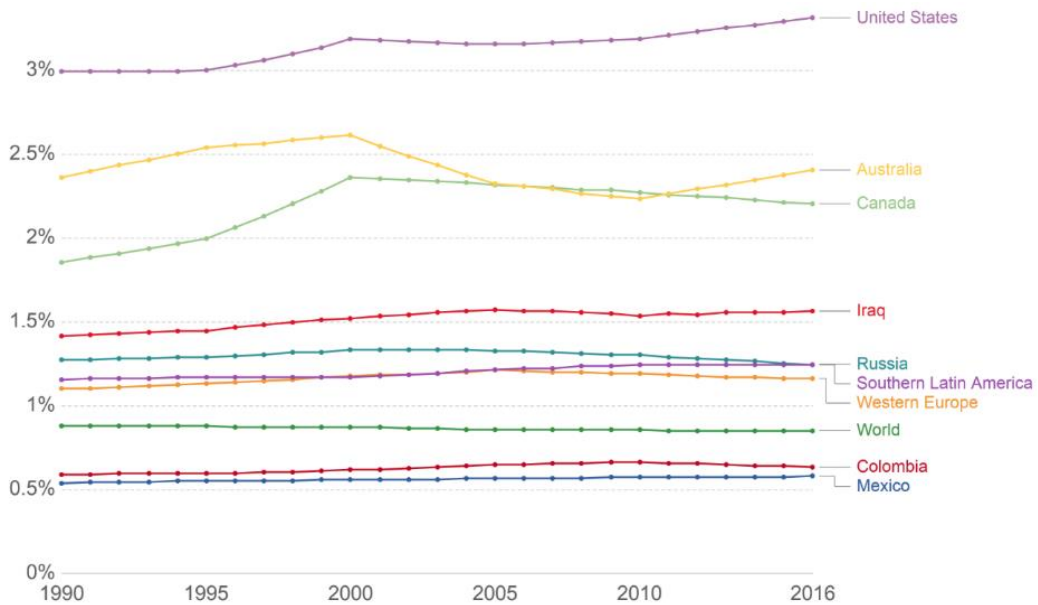
Source: IHME (2017)

From the graph we can seen how the situation at a worldwide level (the green line) remained stable over the period between 1990 and 2016. Europe, Russia, Iraq and South America also follow this trend. The biggest fluctuations are in those countries where there is the biggest presence of disorders due to drug. In particular, there is a tendency toward the growth in US while the situation is more complicated for Australia and Canada. Both shows a rapid growth till 2000 followed by a decrease that continues till 2016 for Canada but suddenly stop in 2010 for Australia. However, the stability of the world line implies that countries trends tend to offset each other.

---

[3] IHME (2017) uses epidemiological studies as source for global and national estimates. However, it recognizes how these are unequally distributed across disorders, countries, age groups and epidemiological parameters. Using these studies to provide a full pattern of drug disorders is challenging. For this reason, the use of specific software is required. If data is lacking for a given country, the software computes a proxy for each epidemiological parameter based on data available from surrounding countries.

1.3.3 The Economic and Social Costs of Drug Addition

Despite the fact that drug addiction involves a relatively tiny percentage of worldwide population (0.85% in 2016 as seen in subchapter 1.3.2), its impact cannot be neglected. When we hear about the consequences of drug abuse we probably think about medical and psychological implications of this disorder. However, especially in those countries where is more widespread, the socio-economic costs of drug abuse are relevant. These costs regard the medical resources necessary for the treatment, care, and rehabilitation of patients with drug related problems; the reduced or lost productivity due to drug consumption; the crime enforcement; and the pain and suffering of the drug abuser that may also affect its family and friends. Drug abuse imposes a burden on both the individual and whole society.

The following section is mainly based on the cost-of-illness studies of Rice *et al.* (1991) and Cartwright (2008). According to the first, costs can be classified into two categories: core costs, which are a direct result of the illness; and related costs, which are secondary and represent non-health effect of the disease. Both core and related costs can be further classified into direct (payments are actually made) and indirect costs (resources are lost).

Direct costs encompass those expenses that could be allocated differently in the absence of drug abuse, namely:

- Health expenditure for persons affected by drug addiction. This category comprises personal health care costs, including care in drug-related specialty, government institutions and short-term hospitalization. Expenditures for research, doctors and nurses training, program administration, and private insurance and, generally, all those support costs related to the treatment of drug abuse should be included.

- Costs of crime, so criminal justice system, public police protection costs, private legal defence, property destruction, drug traffic control, and social welfare administration.

Among indirect costs we find those elements which represent a reduction or loss of productivity, so:

- Morbidity cost: it is the value of reduced or lost productivity due to drug abuse. Sociological research highlights the difficulty of holding and finding a job and working productively in both workplaces and homes for drug abusers. It is computed as the number of individuals affected times the average income loss per individual due to drug abuse.

- Mortality cost: it represents the loss in productivity observed when an individual that would have continued to be productive for several years dies prematurely due to drug abuse. The cost (value for society) of all deaths is calculated as the product of the number of deaths and the expected value of an individual's future earnings. Different variables such as age, gender, life-expectancy are usually taken into account in the computation.

- Other costs: drug abuse can contribute to spread diseases such as HIV/AIDS (and other sexuality transmissible diseases), tuberculosis, and hepatitis B and C and lead to further medical complications such as infectious disease, respiratory and cardiovascular effects, cancers, prenatal effects and more. Costs of crime that are not included in direct costs namely victims of crime, incarceration and crime careers are included in indirect costs. Other personal issue that drug abuser causes to itself and the persons close to him in term of pain and suffer represent another kind of costs (even if difficult to quantify in monetary terms).

In figure 1.10 are reported the estimated costs of drug abuse in US in the years 1992 and 1998. Data has been adjusted for inflation and population change.

Figure 1.10: costs of drug abuse in United States (1992 and 1998)

| Variables	1992 (in millions of US$)	1998 (in millions of US$)
Health care expenditures	10,820	12,862
Specialty drug services	3,145	4,493
Medical and other	7,405	8,369
Productivity losses		
Earnings, premature death	14,575	16,611
Earnings, illness related to drugs	15,682	24,929
Earnings, crime and victims	39,164	56,925
Other impacts		
Criminal justice, private safety, and social welfare administration	21,912	32,083
Total costs	102,154	143,411

Source: Cartwright (2008)

The total costs of drug abuse in US between 1992 and 1998 increased from 102,154 million to 143,411 (+40.39%). This is coherent with the growth of this phenomenon in the country. Direct costs related to healthcare were 21,370 in 1992 and 25,724 in 1998 (+20.37%) while the indirect cost of productivity lost increased of the 41.84% (69,421 in 1992 and 98,465 in 1998). The direct cost of crime increased of the 46.42%. From the data

we can see how the biggest burden is represented by the losses in productivity. Cost of crime is also expensive whereas healthcare expenses represent the smallest share of the total expenditure due to drug abuse.

Since productivity losses are the biggest cost, further investigations can be interesting to perform. We start with respect to mortality cost. Figure 1.11 describes the death rates due to drug abuse disorders in 2016, measured per 100,000 individuals.

**Figure 1.11**: Death rates for drug use disorders (2016)
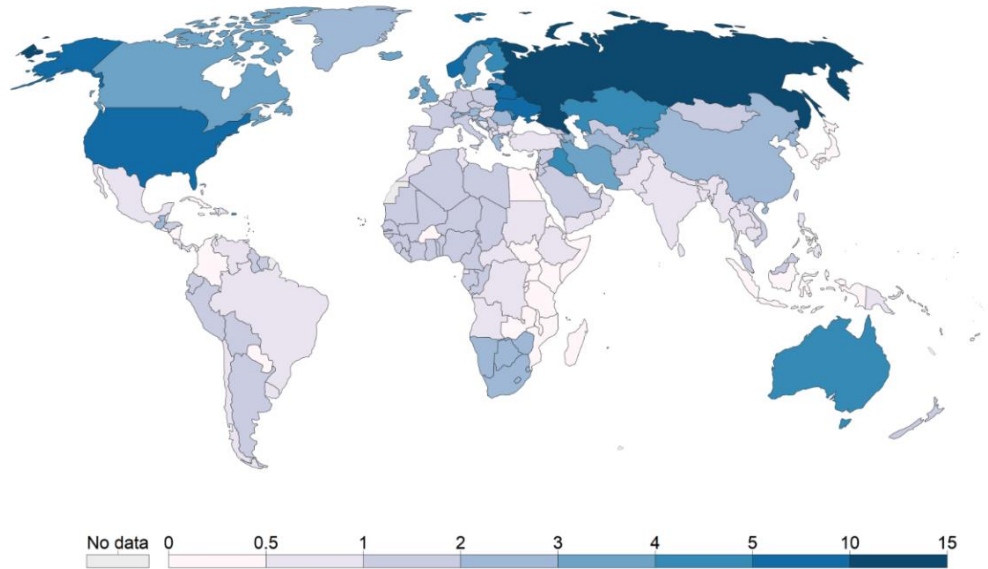
Source: IHME (2017)

According to the data, Russia is the first country for deaths due to drug abuse (10.27 deaths per 100,000 individuals) followed by United States (9.74 deaths). Russia seems to have a sort of negative effect since deaths are particularly high in adjacent countries (Ukraine, Belarus and Lithuania) and Kazakhstan. Also, Norway has a particularly high death rate (5.75) while the average rate in Western Europe is lower (2.11). Canada and Australia where drug abuse is particularly widespread have a death ratio of 3.88 and 4.05 respectively. Although in Central Africa and Asia drug disorders seemed not too may widespread (as seen in 1.3.2) they both show modest rates.

Opioids represent the main cause of death in United States with a death ratio of 10.34 with respect to 100,000 individuals in 2015. Of these, just 4.05 was caused by Heroin while the biggest share is represented by Pain Relievers (i.e., Morphine, Methadone). One possible explanation is that the latter can be prescribed to patients in order to treat several disorders and so they are more easily available. Cocaine (which is not an opioid) caused 2.11 death per 100,000 individuals in 2015, one-fifth with respect to general opioids and one half if we just consider Heroin. Looking at the historical evolution, the number of

deaths due to opioids overdose is rapidly increasing. While deaths due to cocaine increased till 2006 to decrease in the following years (till 2012). In 2012, the number of deaths due to Heroin overcame the deaths due to Cocaine. This data can be seen in Figure 1.12 which shows the evolution of drug overdose death rates in US with respect to Cocaine and opioids in the years between 1999 and 2015. Death rates are computed per 100,000 individuals.

Figure 1.12: Drug overdose death rates in US (1999-2015)
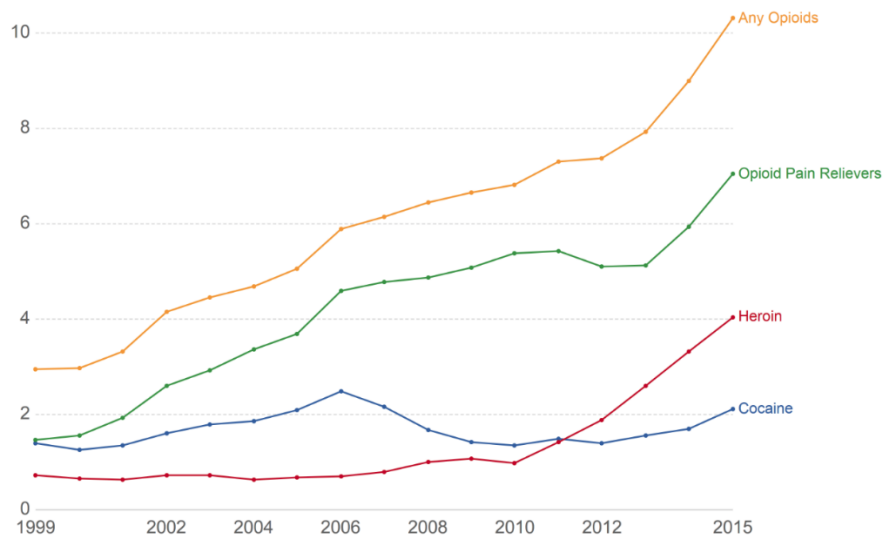
Source: IHME (2017)

Finally, Figure 1.13 shows the disease burden cause by drug abuse for each country measured in Disability-Adjusted Life Years (DALYs) in 2016. DALYs considers both death rates and years lived with disability or other health burden. Again, rates are computed per 100,000 individuals.

Figure 1.13: Drug abuse disorders DALYs (2016)
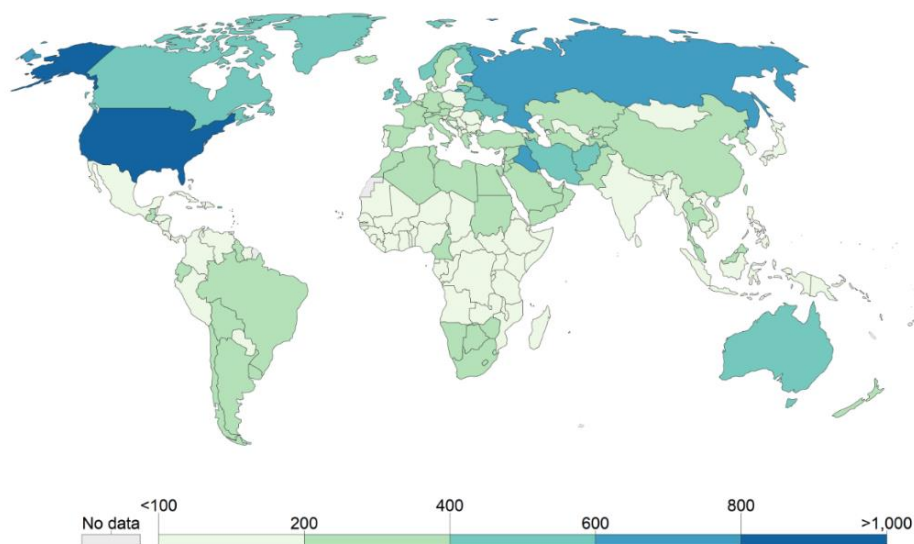
Source: IHME (2017)

As was expectable, US is the country with the highest DALYs (1,177.95 per 100,000 individuals), followed by Russia (750.79) and Iraq (710.43). Australia and Canada have a DALYs of 595.09 and 568.83 respectively. Western Europe presents a better situation with 290.13 as well as South America (260.43). DALYs are also modest in Northern and Southern Africa and Asia.

1.4 Reasons for the State Intervention against Drug Addiction

1.4.1 Externalities

An externality occurs when there is an interaction between two or more subjects. Generally, we talk about externality if an individual (A) carries out an action that has an effect/impact on a second individual (B) that it is not compensated. It is a not symmetric situation since A acts without considering the effect that its actions might have on B. According to the economic theory, an externality can be either positive or negative. Depending on the situation, we talk about marginal benefit or marginal cost for the individual.

Externalities represent a problem because implies a loss of efficiency. In a perfect world, the individual who provides a positive externality should be compensated by those who receive the benefit. On the contrary, who suffers the effect of a negative externality should receive a payment from the causer. If this does not happen, problems such as free-riding may arise. Economics typically identifies two possible solution to correct externalities:

- Coase Theorem: it is based on the assumption that each market is efficient. Bargaining will lead to a Pareto efficient result regardless of the initial allocation of property rights if trade in an externality is possible and there are low transaction costs;

- Pigouvian taxation: it represents a payment (or subsidy) whose purpose is to modify the behavior of the individual who generates the externality. It is usually applied in field such as health, environment and R&D to boost efficiency (Gruber, 2013).

As we said in subchapter 1.1 of this thesis, healthcare can be provided by both private and public institutions. In the second case, one of the problems of public goods[4] is the difficulty to find someone who pays for them. For this reason, they are often publicly

---

[4] A good is considered "public" when it is both non-excludable and non-rivalrous meaning that an individual can consume it without preventing others to do the same and from which no one is deprived. Public healthcare can be considered a public good because it presents these characteristics. Additionally, providing health services to citizens is in the government interest in order to increase the social welfare (Gruber, J., 2013).

funded through i.e. taxation. Some behaviours such as sedentarily life-style, proliferation of junk-food, smoking or drug abuse are examples of private or public actions that create externalities. Since another feature of public goods is the impossibility to exclude individuals from consuming them, these externalities influence the quantity and quality of work time and leisure, raise healthcare expenses for individuals and employers, health insurance premiums for consumers, and impact the overall financial performance of corporations. Financial disincentives, taxation and property rights legislation are used to "internalize" harmful externalities into the marketplace (Carande-Kulis *et al.*, 2007).

The usefulness of these instruments to deal with drug addiction is still relatively uncertain. However, some deductions can be made. The first is that, how we will better see in subchapter 1.4.3, drug market (except for those used for medical purposes) is in bigger part out from the state regulation. The question is how the government could apply a tax on i.e. cocaine when its market is out from the government control. Under this point of view, the control of illicit drugs market may represent an incentive to legalize them but, other than the social and health implications of the decision, a further problem may arise.

Addictive commodities (i.e. cigarettes, alcohol, drugs) are typically considered goods that do not follow the basic law of economics due to addictive behaviours such as irrationality of their consumers. The most significative implication is perhaps the down-ward sloping of the demand curve. This assumption implies that a change in price of the good would not affect the quantity demand making i.e. a tax on tobacco useless if applied with the purpose of reducing the tobacco consumption (Winston, 1980). A graphically representation of this situation is displayed in Figure 1.14.
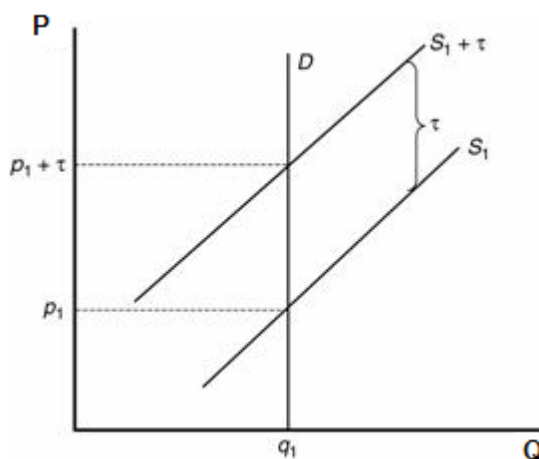
Figure 1.14: Tax effect with perfectly inelastic demand

Figure 1.14 shows a situation in which the demand for a product is perfectly inelastic. In this case, τ represents a tax on the commodity. The tax application leads to an upward

change of the supply curve. The new optimal equilibrium is found in a price higher than the original one but in which the quantity sold remains unchanged. Governments might take advantage from this situation imposing a very high tax burden on addictive commodities and incrementing their income (that can represent an additional point in favour to drug legalization). A further implication of this model is that, in case of perfectly inelastic demand, the whole burden of the tax is paid by the consumer.

Despite this, the more recent economic research has demonstrated that also the demand of addictive commodities responds to changes in prices and other factors contradicting the assumption of inelasticity of the demand curve. An example is Chaloupka (2018) which analysed the effect on tobacco consumption to cigarettes price variations. A fragment of Chaloupka's work is reported in Figure 1.15.

Figure 1.15: Tobacco consumption and Cigarette price in New Zealand (1990-2013)
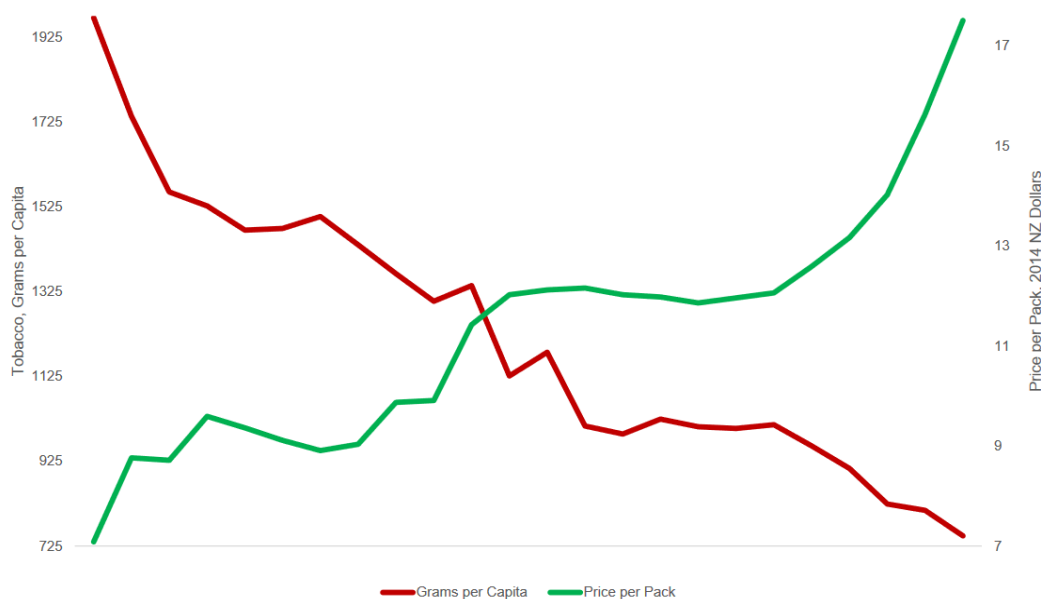
Source: Chaloupka (2018)

The graph clearly shows a correlation between cigarettes price and tobacco consumption. The consumption of tobacco can be read on the left axis whereas the price of cigarettes on the right one. We can see that as the price per pack increased over time the quantity of grams of tobacco proportionally decreased.

Chaloupka (2018) also provides some data about the relationship between monetary price and drugs use[5], in particular:

- Heroin: a 10% growth in price is reflected in a over 9% decrease in use;

- Cocaine: a 10% increase in price causes a reduction of about 3% in consumption;

- Marijuana: a 10% price increase lowers prevalence of youth marijuana use by 3%.

---

[5] Computations are based on the economic complementarity among licit and illicit substances.

Nevertheless, the way in which governments can control and affect these prices is still unclear because, as we will be seen in subchapters 1.4.3 and 1.5.1 of this thesis, states' attempts to control and reduce illegal drug markets have proved to have little or no effect.

1.4.2 Paternalism

A state/government is considered paternalistic when acts (or forbid to act) on behalf of an individual, against that individual's willingness or without its consent, with the explicit aim of doing good for (or avoiding harm to) that individual. In other words, paternalism regards an unwanted intervention, or the intentional withholding of something that is desired, by the targeted person or group (Cody, 2003).

Historically, the main justification of paternalism stemmed from the disbelief about the capability of certain categories of people (on the base of i.e. sex, sexuality, ethnicity, religion) to make decisions in their best interest and possibly controlling them. Today, this justification is no longer applied, and paternalism is mainly aimed at persons who are poor, at minorities, and at those with less power such as individuals with a chronic and critical physical or mental illness. A further reason arises from the fear that even rational people might not behave in their long-term self-interest in some situations. Example of paternalistic behaviours are the restrains of free-market, the compulsory education of the young, limits on purchasing and owning firearms.

In healthcare, paternalism is commonly identified in the conflict between the primary obligation of healthcare professional (i.e., physicians, nurses) to respect the principle of beneficence and the rights of persons who are receiving services in their practices in order to make autonomous decisions about their lives. Recommended interventions and treatments are justified by beneficence regarding the care of patients' health. Health-related disciplines are driven by scientific principles which provide guidelines. For this reason, authorities tend to believe that persons should do what medical personalities recommend in order to maintain their health. From this assumption some obligations such as compulsory vaccination, limits on legal drug availability and use arise. Then, a paternalistic state has an interest in preventing drug abuse: banning drugs can be considered a classic paternalistic policy with respect to this problem.

Generally, we can distinguish weak and strong forms of paternalism. In weak paternalism an agent intervenes for beneficence purposes to prevent nonvoluntary conduct, namely, to protect persons against their own nonautonomous action. It might be the case of persons who are in a state of severe depression, misinformed or under the control of addictions.

On the contrary, strong paternalism deals with interventions that wants to be beneficial for the individual, although it takes its risky choices and actions in an autonomous, voluntary and informed way.

Paternalism has been largely criticized, especially in its stronger forms because, if on one hand it prevents irrational individual from possible harmful choices, on the other hand it provides too big limitations to rational people's willingness. A possible solution to this problem is identified by behavioral economics in the so-called asymmetric paternalism (Camerer, 2003).

While we will better deal with behavioral economics in Chapter 2, we can define a regulation as asymmetric paternalistic when it provides large benefits for irrational individuals, while imposing little or no burden on those individuals who are fully rational. To understand asymmetric paternalism, we first need to consider how paternalistic policies are generally evaluated. A policy is considered beneficial if:

$$(p * B) - [(1 - p) * C] - I - \Delta\pi > 0$$

Individuals are divided into two categories: a fraction, p, who is boundedly rational and another fraction, $1 - p$, who is fully rational. A general paternalistic policy is developed in order to hinder, and possibly avoid, mistakes made by boundedly rational individuals but, imposing limitations, it may cause a cost on rational individuals. In the formula, the benefit for irrational agents is represented by B, whereas C is the net cost for rational agents. I denotes the cost of implementation of the policy while $\Delta\pi$ represents the possible policy effects on the firms' profits.

A policy is asymmetric paternalistic if it creates a large benefit for boundedly rational agents (so a large B) while imposing little or no cost for rational agents (small value of p or zero in case of pure asymmetric paternalism). Even possessing little information about the frequency of individuals' errors, as long as we have low p and since we consider that also the most rational agents sometimes may exhibit bounded rationality, we can assert that the policy is on net beneficial. However, we should not focus only on the value of p. A asymmetric paternalism policy is worth little if the cost of its implementation is very high (high value of I). Additionally, consumers' errors can be exploited by firms, which have an advantage from them. For this reason, an asymmetric paternalistic policy may be harmful for them. However, Camerer *et al.* (2003) claim that asymmetrically paternalistic policies enable irrational consumers to take better choices and this effect is reflected on a net increase of economic efficiency computed as the sum of consumer and producer's surplus.

But, how can asymmetric paternalism be applied to drug abuse disorders? To answer this question, we need to consider the role of cooling-off periods. When people are driven by emotions rather than rationality (the so-called "hot" states) they might take decisions that are costly or impossible to reverse. Behavioral economics suggests that these suboptimal decisions are taken due to an overestimation of how long these hot states will last. An extreme example may be an individual who commits suicide when depression is particularly intense.

To avoid hot and irrational decision-making, cooling-off periods force individuals to delay its decision for a certain period of time, allowing people to revaluate it when free of heat-of-the-moment impulses. If μ denotes the net benefit of carrying out an action now, a rational person would perform that action if $\mu > 0$. Suppose that bundled rational people also are present and experience a true net benefit of μ. However, because of evaluation errors, irrational people would perform an activity when $\mu + \epsilon > 0$ (where ε represents the value added by the evaluation error). Therefore, the individual may hurt itself when it takes a decision despite $\mu < 0$.

Let's consider now the case in which there is a cooling-off period before the action is undertaken. μ' represents the net benefit of carrying out the action after the delay. Irrational individuals might take advantage from the cooling-off period reversing a costly decision to perform the action, and hence benefit by as much as ε.

A possible problem of cooling-off periods is that they should be implemented after a careful analysis. This to avoid two potential costs, namely:

- People who carry out the activity regardless of the cooling-off period, whether rational or irrational, may face a loss of net benefit due to delay equal to $\mu - \mu'$.

- Individuals for who the net benefit is $\mu > 0$ before the cooling-off might decide to not perform the activity in a second moment ($\mu' < 0$).

Nevertheless, if implemented correctly, cooling-off periods represent a good example of asymmetric paternalism since it imposes a minimal cost on rational individuals while protecting those people that may act in the heat of the moment. Its implementation can occur in two ways: (1) forcing individuals to wait a certain period before undertaking the action; (2) enabling immediate decision but allowing the possibility to retract it during the cooling-off period. However, for some choices, i.e. suicide, this second option is not available.

Coming back to the drug abuse problem, scholar recognize that craving is the main driver for this kind of behavior and, just like depression, may lead to suboptimal choices. A

characteristic of craving is that it drastically impacts present decisions but has little effect on decisions involving future outcomes. For this reason, an addict might show a high willingness to pay to get drug immediately, but it would not be agree to pay the same amount in future. A possible asymmetric paternalistic policy in this case would be to dispense drug legally with a mandatory waiting period (instead of banning drugs). The forced waiting allows to protect the future self from the craving current self. Given the fact that perfectly rational users plan ahead, the delay imposes little cost and may be beneficial for drug users who have the possibility to make comparatively rational decisions for the future.

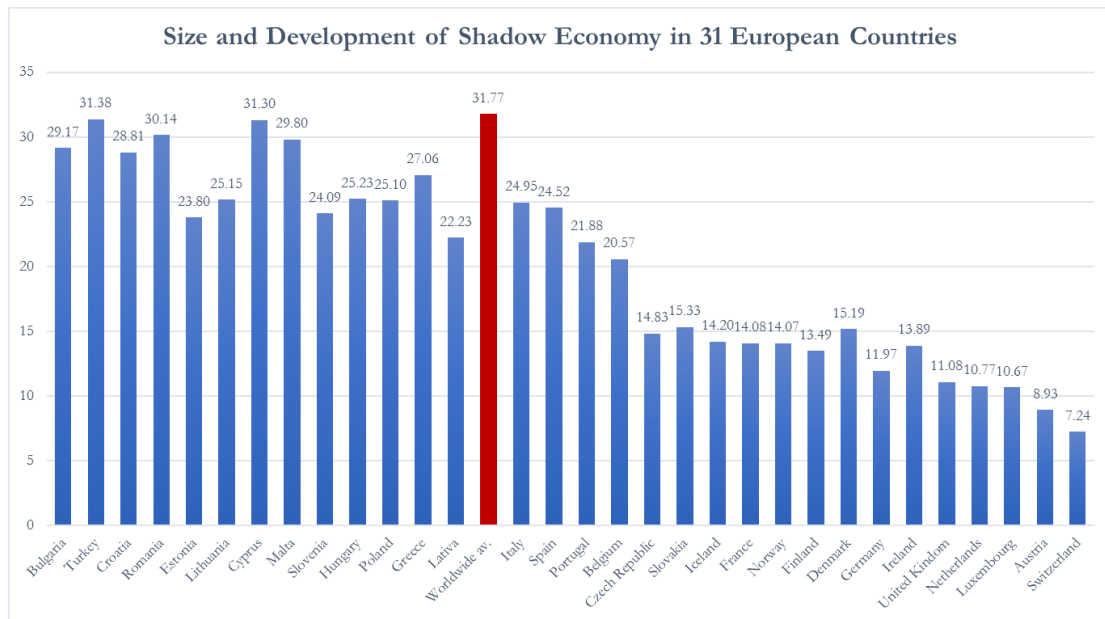### 1.4.3 Shadow Economy and Illegal Markets

The shadow economy represents a form of unofficial economy. Also known as hidden, informal, black economy, it can be defined as the economy that covers all those activities that are not encompassed in the national income accounts. Shadow economy does not necessarily mean criminal. Within shadow economy are also included those activities that would be *per se* legal but that becomes illegal in absence of some requisites (i.e., unreported employment). At the same time, criminal activities such as drug trade are considered (Fleming *et al.*, 2000).

According to Medina & Schneider (2018) shadow economy is by nature difficult to estimate because agents involved in illegal activities try to remain anonymous. However, efforts for its computation are justified by its political and economic relevance. First, because illegal activities participate in the formation of a country's GDP. Second, the size of shadow economy is the core input for the computation of tax evasion. Monetary (i.e., avoid paying taxes), regulatory (i.e., avoid governmental bureaucracy) and institutional (i.e., weak rule of law) are the main reasons of shadow economy. In addition, some drivers such as high unemployment rate, high tax rates, corruption, government instability and more are proved to boost shadow economy's size.

Figure 1.16 reports average data about 31 European countries over the period 2004-2015. Numbers are expressed as percentage of the total country's GDP. How can be seen, shadow economy represents a relevant portion of the economy in most countries especially in Eastern Europe. However, Western countries such as Belgium, Italy, Portugal and Spain exhibit high rates. Despite the higher rate in developing European countries, the overall European situation is below the worldwide average of 31. A so high number is worrying but should not surprise if we think that poorest states exhibit rates higher than 40.

Particularly critical is the African situation where countries such as Zimbabwe and Nigeria have rates of 60.64 and 56.67, respectively.

Figure 1.16: Average Shadow Economy as percentage of GDP in EU (2004-2015)

Source: Medina & Schneider (2018)

Incomes deriving by illegal markets are part of the shadow economy of a state. A market is considered illegal when a product, the exchange of it, or the way in which it is sold or produced goes against legal jurisdictions. Law plays an important role in defining which markets are illegal and they may differ among countries and over time. When a market is illegal the state does not intervene to protect neither consumers nor producers (i.e., ensuring a minimum level of quality or protecting property rights) and can prosecute the agents involved in it.

Nevertheless, not every criminal activity represents an illegal market, some requisites should be present. Demand and a supply for the product must be present just like in the legal case. This is not enough; other requisites are needed. From a demand perspective, a legal market for the product either does not exist or it exists but the good is provided at a lower price though the illegal market. The customer must be willing to support the moral cost and accept risks that may derive from the illegal purchase. The supply, as for the legal case, depends on the expected profitability for the supplier but also from its organizational and financial capabilities to produce/sell the product while avoiding the law (Beckert & Wehinger, 2012).

Furthermore, the illegality of the market can stem from four sources: (1) outlawing of the specific product; (2) exchange of a product that can be obtain legally; (3) exchange of a good that has been stolen or forged;(4) violation of regulatory stipulations. The trade of

illegal drugs is comprised in the first category. Other outlaw activities in this category can be child pornography or human trafficking. However, those drugs that are smuggled but also have medical purposes and can be obtain legally though i.e. medical prescription, are within the second category.

A question remains unsolved: why are some markets illegal and the state should intervene to block them? From the standard economic perspective, if a market provides utility to parties involved it should exist. Some justifications are found in values. Some markets may have both individually and socially harmful consequences: drugs have negative effects on their customers and go against the social values of the majority. Illegal markets are often characterized by high levels of asymmetric information that put one of the two parties in a situation of extreme vulnerability. Finally, empirical evidence has shown as the prohibition of some kinds of markets lead societies to the creation of social values economically relevant toward those goods considered harmful, obnoxious or repugnant.

Nevertheless, values are not enough. Economic consideration about the efficiency of market functioning and the protection of its participants should be considered. For example, enabling the trade of stolen goods not only is a violation of the property right but would provide incentives for stealing.

About the specific case of illegal drug markets, Bouchard (2007) identifies them as a particularly resilient kind of market. Resiliency is defined as the capability of market participants to resist external pressure aimed at obstructing the trade maintaining the existing levels of exchanges between consumers and sellers. Bouchard noticed that, despite the increasing pressure put by repressive policies in US in the period between mid-1980s and mid-1990s (drug control expenditures grew from $10 billion to $35 billion) with a consequent raise in drug dealers arrested, indicators showed that the markets for illegal drugs remain substantially unaffected. Indeed, the recruitment and replacement of dealers did not present any significant change; availability of most drugs remained stable; prices were constant or in decline even in the weeks following high profile police operations in specific markets.

Bouchard attributes this to the fact that law enforcement on illegal markets is not fully understood. There is lack of a good knowledge about the size of the market and the structure of distributional nets. Lack of conceptual tools that explain why law enforcement interventions are more susceptible to work or not also represents a problem. Additionally, three characteristics of drug market contribute to make it a particularly resilient:

- Low vulnerability to external shocks: vulnerability is the level of exposure to attacks. A natural protection in this sense in drug markets is represented by the small size of drug-dealing organizations. Small sizes enable to better resist to arrests and seizures other than reducing exposure to hazard, members' mistakes and likelihood of partners becoming police informants. A similar relationship exists also regarding drug shipment. Additionally, the high price of drugs allows to compensate the small size of most businesses. Dealers are satisfied even with a small number of customers since the high price for transaction keeps the activity profitable. Furthermore, the decentralized structure of drug markets helps minimizing the harmful effects of high-profile law enforcement operations. The effects of an external shock are contained when the removed player or organization plays a negligible role, or when the quantity of drug removed represents a small share if compared with the size of the market.

- High elasticity to external shocks: market's elasticity to attacks concerns the efficiency of the replacement process. It depends on two variables: the attractiveness of the market and the presence of barriers to entry. Drug markets are elastic because provide both incentives and low barrier for people who want to entry. Again, the core is represented by the decentralized structure and the high price of drugs. The high prices of the drug trade enable to increase the attractiveness for potential dealers while decentralization contributes to lower the financial costs necessary to enter in the drug market. Generally, firms want to become even bigger in order to exploit economies of scale or advantages in borrowing funds from banks. But, in the drug sector, small size represents an advantage for market entry since it implies no investment by dealers to enter in the trade. Furthermore, economies of scale and large-scale production are rare in drug economy. However, two barriers exist: (1) a certain level of skills is required (i.e., ability to avoid police detection and willingness to take risks); (2) social capital, intended as trustworthy relationships with partners and dealers since establishing reputation on drug market may require time.

- High capacity to adapt: adaptive capacity regards the ability of dealers of modifying work conditions making themselves less vulnerable. There are many forms of adaptation: dealers can substitute one product with another (e.g., from cocaine to opium, from heroine to crack) or can modify the methods of distribution (e.g., from fixed location to delivery) or displacing activities to

different locals. The important thing is that market continues to work and persist. Again, the small size of drug businesses represents an advantage making adaptation easier since they are less subject of the inertia of large organizations. The high flexibility of the market makes sound events such as arrests more as a "routine issue" that is quickly resolvable. Adaptation is enhanced by technology: mobile phones have almost completely eliminated street-selling methods allowing a better organization and to reach a larger share of the market while limiting risks.

1.5 Real-Life Policies on Drugs

1.5.1 Drug Prohibition Law

Illegal Drug trade is a global market that encompasses the cultivation, production, distribution and sale of drug. The major trafficking routes originate from Mexico, South America countries, Africa and Southern Asia where plants are cultivated and then spread throughout the world.

According to Levine (2003), despite the fact that each country around the world has its own laws and policies, a global drug prohibition system exists. Drug prohibition is born in the '20s in US as a subset of alcohol prohibition. In 1930, the US Congress divided drug prohibition from alcohol prohibition and created a new federal agency, the Federal Bureau of Narcotics. In 1948, the new United Nations (UN) made drug prohibition a priority matter. The UN single convention of 1961 (then modified in 1971 and 1988) established the current system of global drug prohibition. In the following 80 years, almost all countries endorsed drug prohibition laws making national drug prohibition one of the most widely accepted, reputable, legitimate government policy of the 20th century.

The success of these policies is partially identifiable in the huge pressure from the US government and the United Nations. A further reason is the increased power of police and military. Indeed, thanks to drug prohibition, government have used anti-drug squads to carry out surveillance operations and military raids that they would not have been able to justify since anybody could be in the drug business. In addition, politicians and governments could use drugs as scapegoat for many long-standing and recent issues such as corruption, fraud, physical violence, theft, robbery, rape, low productivity and more. In fact, drugs were described by governments, media and religions as extremely dangerous often with exaggerated and extreme terms. Last but not least, drugs prohibition proved to unite political opponents. Conservative, liberals, fascists, socialists, communists, populists all agreed on the use of state power for pursuing the common good. They also competed

in designing the more punitive anti-drug laws, building prisons, expanding antidrug police military forces, and funding government sponsored anti-drug messages.

However, a question may arise: are these drug prohibition policies useful? As we saw in subchapter 1.4.3 of this thesis, drug markets exhibit a particular high resilience to external shocks many times making law enforcements useless. Other than provide the features of resilience, Bouchard (2007) also provides a framework to evaluate the resilience of illegal drug markets. This framework is represented in Figure 1.17.

Figure 1.17: The Resilience of illegal Drug Markets
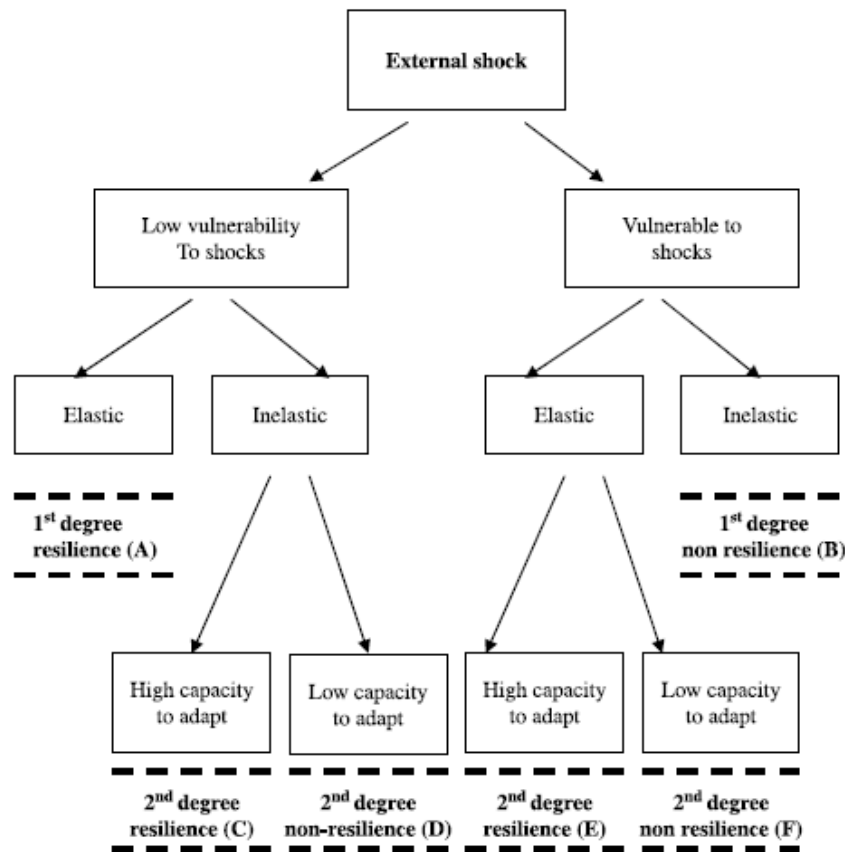
Source: Bouchard (2007)

Recalling the determinants of resilience: vulnerability, elasticity and adaptation can or cannot be correlated among them. Figure 1.17 describes the potential relationships between these three elements: some links lead to resilience some do not. The first case is the one in which both low vulnerability and high elasticity are present (A). We are talking about first-degree resilience because damages provided by law enforcement agencies are never so significant and adaptability is not necessary. These markets are the most challenging to drug law enforcement. On the contrary, in (B) we have vulnerability to shocks and inelasticity, so these markets are likely to collapse even before possible adaptation implementations. Situations (C) and (D) consider that if the market has a low vulnerability but is inelastic,

adaptation is necessary to get resilience. The same reasoning can be done for situations (E) and (F), when the market is vulnerable but has high elasticity. Some segments of Marijuana market are considered to be (A) type due to the very widespread decentralized structure. The historical evolution of heroin markets in US can be associated to (C). Originally, heroin market started from opium that was imported by Chinese immigrants in the 19th century. For many years, they were able to avoid attacks by hiding behind legal commercial companies. However, when opium became subject of major law enforcement laws, they could not continue to operate in the same way, and they needed to adapt also finding different and comparable products that could be used outside opium dens. Finally, markets (E) are identified in those dealers that are very vulnerable to arrest because they accept to sell also to strangers. Despite this inconvenient, the management of these markets has shown a good elasticity thank to its ability of finding replacement dealers quickly. However, when it becomes impossible to substitute the ones who have been arrested, instead of disappearing, dealers have shown adaptation abilities moving to another neighbourhood.

Returning to drug prohibition policies, Levine (2003) asserts that starting from the '80s, they have been facing some turning points and crises. For many years, global drug prohibition operated in at almost invisible way and was often taken for granted. Now that drug prohibition is easier to see, it has lost some of its ideological and political powers. One of the consequences of the increased visibility is a growing opposition by a relevant number of people such as physicians, public health officials, judges, lawyers, police, journalists, social workers and drug users toward the more extreme and punitive forms of drug prohibitions. They are considered to be too expensive (taking away resources for other health and political policies), often ethnically discriminatory and ineffective for reducing drug use. About this last point, global drug prohibition has often been accused to be unable to stop the cultivation and use of cannabis around the world.

The birth of harm reduction movements to respond to the spread of hepatitis and AIDS epidemics in the '80s can be considered an additional motivation of the drug prohibition crises. Even if they were accused by US conservatives to be "drug legalizers", outside the US other politicians, public health professionals and even the United Nations supported harm reduction programs as practical public health policy. Very often, changes in laws and policies have been required for the application of some practice such as syringe distribution and exchange, heroin prescriptions, methadone maintenance, cannabis application in medicine, or drug education for users. However, the use of these methods is justified by the fact that harm reduction programs provide a tolerant and pragmatic approach to drug

use where reasonable and competent specialists try to persuade those who use drugs to minimise the harms that their activities produce (e.g., reducing the spread of AIDS).

1.5.2 Cannabis Legalization

One of the most discussed and controversial topics about drugs regards the legalization of the so-called soft drugs, in particular, of the cannabis. Over the time, many country's legislations have moved toward an increasing tolerance of cannabis where, even if formally illegal, has been unenforced or decriminalized. Today, the use of cannabis for medical purposes has been approved and is legal in several countries.

Possible arguments arise when we talk about the legalization of cannabis for recreational use. The probably most well-known case is the Dutch one. The Netherlands *de facto* legalized cannabis for recreative use in 1976. Despite the formal legalization, some limits persist: cannabis can be purchased only for personal use and in small units and inside specific coffee shops where the consumption must take place. Outside coffee shops, the sale and possession are still substantially illegal (even if decriminalized) (MacCoun & Reuter, 1997).

The recent legalization of cannabis for recreative use in four US countries since 2012 (Alaska, Colorado, Oregon and Washington State) have brought the legalization topic back into the world. According to Hall & Lynskey (2016), the biggest concerns regard the possible consequences for the health of cannabis users since the bigger availability and the lower price may lead to adverse events. Indeed, the main driver through which legalization may lead to an increase in cannabis users is identified in the reduction in cannabis' price with respect to purchasing it through illegal markets. Three reasons for the lower price are identified:

- Legal cannabis' price does not include the black-market premium for the risks of arrest or drug market violence;
- The legal production is more efficient than the clandestine one. Growers can exploit the bigger scale of production reducing the unit price with a consequent benefit for the consumer;
- Even if taxation is applied, this is not enough to bring the price at the same level of black-markets one.

The first concern about Marijuana laws regards the possible effect on adolescents. It is reasonable to think that cannabis legalization would lead to an increased use by adolescents also due to the fact that legalization may pass the message that cannabis is not risky.

However, research seems to deny this risk. Comparing adolescent in countries where cannabis is legal and countries where is not, surveys did not find any significant difference. Differences in cannabis use were not found even before and after the legalization law (Hasin *et al.*, 2015).

Different results have been found for adults (over the 21 years). If there is substantially no difference in rates of initiation of cannabis use among adults between MML and non-MML states, it was found that the legalization of cannabis has led to a growth of its use in those countries where legal.

Research has found an increased percentage in cannabis-impaired drivers in fatal crashes in those countries where marijuana is legal with respect to the past. However, these results are somehow controversial since studies on the impact of cannabis on crashes were not so widespread and were mainly focused on alcohol before the implementation of the law.

Positive effects on health has also been found: a decrease in suicide among males between 20-30 years in those countries where Marijuana has been legalized and lower death rates due to opioids overdose.

Legalization can drive to other advantages: the government can now monitor with certain how much cannabis is sold and though taxation it represents a further income for the state. To simplify regulation, governments restrict the number of licensed producers and retailers creating a sort of cannabis oligopoly. Measures of THC and cannabidiol (CBD), two potentially dangerous elements contained in cannabis, can be required by regulatory authorities in order to concede and maintain the license.

Furthermore, legalization simplifies the study of health effects due to recreational use. Individuals who use marijuana have no longer reasons to lie or hide, making the identification of representative samples easier.

However, the creation of legal markets for cannabis recreational use is a relatively new phenomenon and it is maybe too soon to state in a certain way whether it has increased cannabis use and cannabis-related harm. Future evaluations will be needed to confirm the actual effects of legalization.

Chapter II: Behavioral Economics

### 2.1 From the Neo-Classical Theory to Behavioral Economics

The term "Classicism" was coined by Karl Marx in 1847 in order to mark the thought evolution from Mercantilism, which characterized the economic way of thinking in the period between the 16th and 18th centuries. It posits a clear separation: we pass from an economy where the value was based on metals (i.e., silver and gold) and focused on the maximization of the power of the empire, to an economy where its value was based on labour. So, from a macroaggregate (the welfare of the community) we move toward a microaggregate (productive labour). Adam Smith is probably the most well-known example of classical economist.

However, in the 1870s a further qualitative change in economists' approach happened. The aim was shifted from labour to the single individual, with the purpose of understanding his behaviour. This new school of thinking was named "Neo-Classicism" by Thorstein Veblen in 1900 (Colander, 2000).

The base of the Neo-Classical perspective is represented by the Utility theory. The value of a good is related to the (expected) utility that its purchase can provide to the individual. The neo-classical customer is identified in the Homo-Oeconomicus, a perfectly rational and self-concerned individual who is aware of all consumption options and has all the information necessary in order to take always the best decision (i.e., the one that maximize its utility).

According to the Neo-classical theory, individuals' decisions are affected by the budget constraint and preferences. Taking into account the price of goods, the budget constraint marks a limit to the customer's possibilities. If we have two commodities, X and Y, and their prices are respectively p_x and p_y while I is the individual's income, this limit is given by $P_x + P_y \leq I$. Instead, preferences are represented by the agent's utility function, where the so-called indifference curves represent its level curves. An indifference curve shows a set of consumption bundles among which the customer is indifferent. Due to perfectly rationality assumptions, the customer is always able to evaluate different consumption bundles and he makes choices by establishing a preference order. To identify the indifference curves (which are subjective), it is necessary to know the marginal rate of substitution (MRS) of the individual. MRS is the rate at which an individual is willing to give up some amount of one good in exchange for another one while maintaining the same level of utility. The implications of the neo-classical consumer theory so far can be summarized using a cartesian system, as shown in Figure 2.1.

**Figure 2.1**: Customer Optimal Decision under Perfectly Rationality
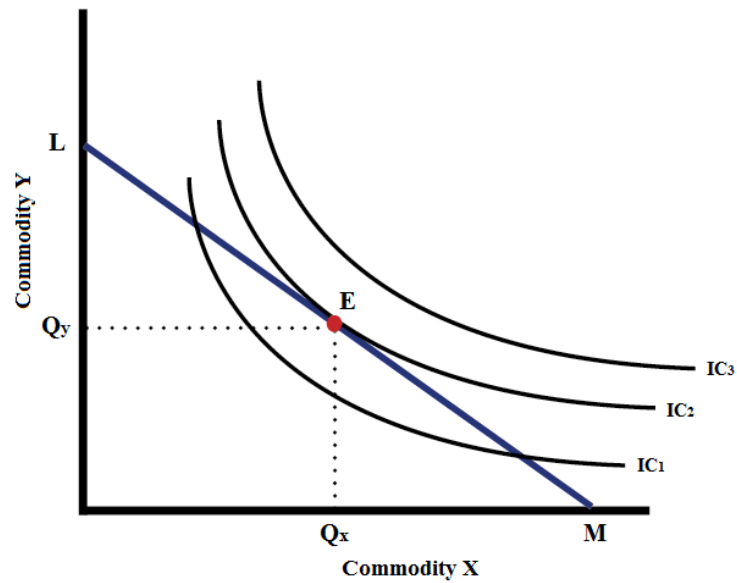
According to the graph, the line ML represents the budget constrain where $M = \dfrac{I}{px}$ and $L = \dfrac{I}{py}$. Bundles beyond the budget line would provide a bigger utility (because they lie on a higher indifference curve) but the customer cannot afford them. On the contrary, bundles within the budget line could be purchased but they provide a lower utility (they lie on a lower indifference curve). The individual would choose the bundle that maximize its utility, namely, the bundle where the marginal rate of substitution equals the price ratio $\dfrac{px}{py}$. In other words, the optimal bundle (E) is found in the point where the indifference curve is exactly tangent with respect to the budget line.

If we modify the price of e.g. the good X, we can determine how the quantity of X varies with respect to Y. Generally, an increase in price of X leads to a bigger consumption of Y and vice versa. If we perform this process for each price, we are able to identify the demand curve of the commodity X for the individual. Since each demand curve is subjective (individuals have different preferences), this process should be performed for all persons and commodities. Individuals' demands are then summed together in order to compute the market demand of the good that will be then combined with the market supply curve (which depends on the expected return of its production) to find the quantity and the price of equilibrium (Bernheim & Winston, 2013).

The theory just described represents the base of the neo-classical approach. A graphical representation of the well-known demand-supply model is reported in Figure 2.2.

Figure 2.2: Demand-Supply Model
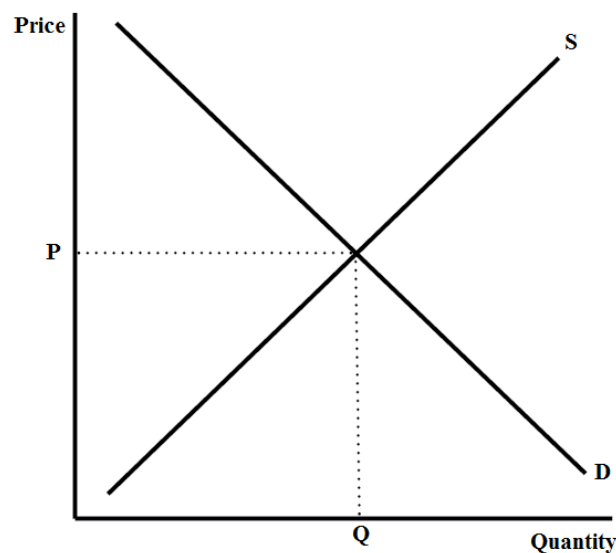
According to the law of supply and demand, the supply curve (S) is positive since an increase in price generates a growth in the quantity supplied of a good or service. On the contrary, the demand curve (D) is negative: an increment of price leads to a lower quantity purchased. The match between the demand and the supply curves represent the equilibrium of the market, namely, the optimal level of price and quantity.

According to Camerer & Loewenstein (2003), the neoclassical approach based on utility maximization, efficiency, and equilibrium is still applied nowadays because it provides economists with a theoretical framework that can be used to describe almost any form of economic behavior. However, they often encompass strong assumptions (i.e., perfect rationality ones) that are difficult to observe in the real world. This is where behavioral economics comes into play. The purpose of behavioral economics is to increase the explanatory power of economics models by adding more realistic psychological elements. The merge between economics and psychology enhance the realism and enable to make better predictions of a phenomena, generate theoretical insights and suggest better policies.

Most of the ideas behind behavioral economics are not new. The same Adam Smith, in its book called "The Theory of the Moral Sentiment", provided some psychological insights applied to economics and that are developed by the current behavioral economics (e.g., loss aversion). However, when economics was recognized as a distinct field of study, psychology did not exist as a discipline. Furthermore, divergent opinions arose since psychologists tend to view human beings as fallible and sometimes even self-destructive; on the other hand, economists tend to view the individual as an efficient maximiser of its self-interest who makes mistakes only when imperfect information about the consequences of its actions are present. Also for this reason, for many years economists refused

45

psychology, which began emerging in the 20th century, because considered not very scientific and incompatible with economics. Theorists such as Arrow and Debreu worked at formalizing economics into mathematics terms. Instead, psychology was a verbal construct, difficult to translate into mathematics. But when expected utility and discounted utility models started to show their limits and became subject to critiques, economists developed the awareness that these anomalies could not be permanently neglected.

Developments in psychology represented a promising direction for new theories, especially around the 1970s when cognitive psychology started supporting the idea of the brain as an information-processing device substituting the behaviourist theory that asserted the brain operating through a stimulus-response mechanism. The information-processing idea shed a light on neglected topics like memory, problem solving and decision making, and better suit to the neoclassical conception of utility maximization than the behaviourist approach. Cognitive models can be organized into two different classifications: prescriptive and analytical.

Prescriptive models concern frameworks to describe how behavior is structured. They are useful to identify which variables should be modified to attract a certain response. The most widespread prescriptive model is probably the Theory of Planned Behavior (TPB) of Ajzen (1991) which is graphically reported in Figure 2.3.

**Figure 2.3**: Theory of Planned Behavior
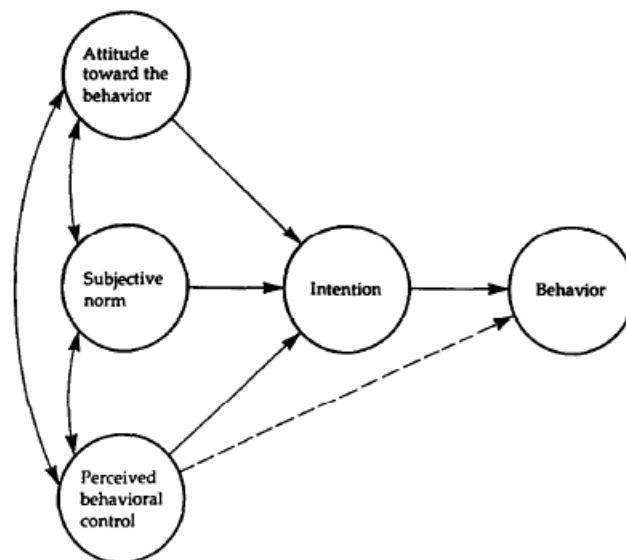
Source: Ajzen (1991)

The theory represents an evolution of the Theory of Reasoned Action always developed by Ajzen in 1975. According to TPB, intention represents the closest antecedent of the behavior. However, an intention-behavior gap is acknowledged. Three elements lead to the formation of an intention: (1) the attitude toward the behavior, so the general

46

predisposition to act that can be either positive or negative and it represents the belief toward an outcome; (2) the subjective norm, namely the moral perception of the individual that is influenced by the expectations that others have toward a certain behavior (i.e., social pressure). Again, subjective norm can be positive or negative; (3) the perceived behavioral control is the perception of how much difficult or easy an activity is to perform. It depends on contextual and subjective factors and it is identified as the main cause of the intention-behavior gap in the case that both attitude and subjective norm are positive. The model is very general and do not refer to any specific behavior, for this reason it has been largely applied, especially in social psychology.

On the other hand, analytical models provide frameworks of key elements that explain the behavior of the individual. One of the most widely diffuse in economics, especially in marketing, is the Consumer Decision Model developed by Blackwell in 1968 (Bray, 2008). The model has been subject of numerous revisions. One of the latest publications depicted the model as reported in Figure 2.4.

Figure 2.4: Consumer Decision Model
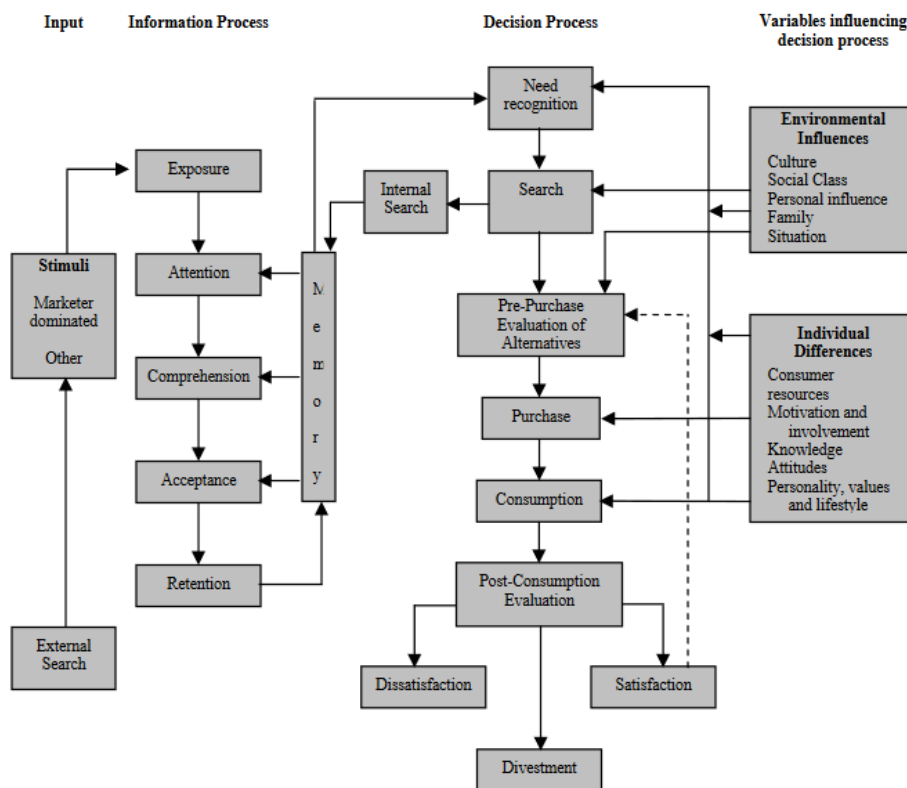
Source: Bray (2008)

Consumer decision model is structured around seven sequential points: need recognition, search of information, alternative evaluations, purchase, consumption, post-consumption evaluation and divestment. All these decisions are affected by marketer and non-marketed

stimuli which are received and processed by the individual in conjunction with memories of past experiences. External variables also affect the final decision both in the form of environmental influences and individual differences.

Leaving aside the cognitive theories on decision-making, we need to say that two influential contributes for the success of behavioral economics were published by the psychologists Amos Tversky and Daniel Kahneman. The first in 1974, arguing that heuristic short-cuts created probability judgments which differ from statistical principles. The second, in their paper "Prospect theory: decision making under risk" where they propose an axiomatic theory based on psychological principle to explain violations in expected utility.

A further milestone that contributes to the establishment of behavioral economics was the conference at the University of Chicago in the 1985 where an extraordinary range of social scientists presented papers. A sort of match between behaviouralists and traditionalists took place. Behaviourists presented violations of the rational choice model while traditionalists tried to explain why these findings do not matter. Despite the success of behaviouralists in 1985, the recognition of behavioral economics in wider academic circles was still far. Just in the 1997, ten years later, a special issue in the Quarterly Journal of Economics was dedicated to behavioral economics (Illiashenko, 2017).

Today, behavioral economics is a growing branch of economics. One proof of this is the American economist Richard Thaler, the father of the so-called Nudge Theory and one of the pioneers of behavioral economics that, in 2017, received the Nobel prize for its contributions in this filed. Thaler challenged the assumptions of perfectly rationality encompassed in the traditional economic theory adding realistic psychological hypothesis (e.g., lack of self-control, limited rationality, social preferences) and highlining how these affect individuals' decisions and market's results. This is not the first time that the Nobel prize for economics is related to a behavioral economics' topic. Already in 2002, Daniel Kahneman won the prize for its capability of integrating psychology into economics (The Nobel Prize, 2017)[6].

Possible concerns about the future remain because of the fear that psychological evidence is too fragmented to provide coherent alternatives to rationality. However, how it will be described in section 2.2, behavioral economics theories have been able to explain things that the traditional economic approach has not. For instance, anomalies in stock market

---

[6] The Nobel Prize (2017), "Richard H. Thaler Facts", Available at: https://www.nobelprize.org/prizes/economic-sciences/2017/thaler/facts/. Last visit: 07 April 2019

pricing have been explained by the Prospect Theory as well as low-probability decisions related to gambling whereas hyperbolic discounting can be used to explain addictions and procrastination (Camerer, 1999).

2.2 How Psychology Affects Economics

As already mentioned, behavioral economics' purpose is to improve the realism of economics models by adding psychological insights, reunifying psychology and economics in the process. Reunification should guide to better guesses about economic behavior and better policy prescription without necessarily rejecting the neo-classical approach.

Economic theories should be evaluated according to three criteria: congruence with reality, generality and tractability. Behavioral economics theories should respect these standards too. As stated, the enhanced realism is one of the main objects of behavioral economics theories while generality is maintained since usually only one or two variables representing the individual behaviours are added to the standard model. However, the inclusion of these variables might make the model less tractable. The increased difficulty in handling behavioral economics models is offset by the possibility of obtaining more precise results assuming more rationality and strategic and dynamics interactions.

The methods used in behavioral economics are the same of the other economic areas with a strong reliance on evidence from experiments, especially in the initial phase, because experimental control is useful for marking the differences between behavioral explanations and the standard ones (Camerer & Loewenstein, 2003).

A first attempt for the reunification between economics and psychology was carried out by Herbert Simons that, already in 1955, developed theories based on bounded rationality assumptions. Nevertheless, these theories were difficult to translate into economic and, especially, into mathematical terms. Just when cognitive psychologist started studying economic decision-making processes in 1970s, findings in this area provided principles and construct that could be included into economic models in terms familiar to economists. Behavioral economics suggests mathematical alternatives including psychological foundation to rationality assumptions (Sent, 2004).

Before analysing how psychology affects rational economic theories, some key behavioral concepts should be explained:

- Bounded rationality: this assumption was developed by Simons (1955) in contrast with the standard economic theory in which persons are full informed when taking decisions. Behavioral economics assumes that when individuals are dealing

with a decision, especially if a complex one, they limit themselves to a sub-optimal set of possible options due to e.g., lack of information, cognitive limitation of the mind or time restrictions. Therefore, many choices might not be considered, including the one that could represent the best option for the individual.

• Reference dependence and Loss aversion concepts were developed by Kahneman & Tversky (1979). Classical theory assumes that preferences are not affected by the current asset situation of the individual. Reference dependence states that individuals do not evaluate solely the expected outcome value when taking decisions, but they also compare this outcome value to some reference point (typically the current endowment) basing on past experiences, expectations and circumstances of other people. Loss aversion occurs because, when people are dealing with a decision, they prefer avoiding losses, even small ones, to a greater extent than acquiring equivalent gains. This implies a change in the shape of the indifference curve as shown by Figure 2.5.

Figure 2.5: Indifference Curve in Loss Aversion Case
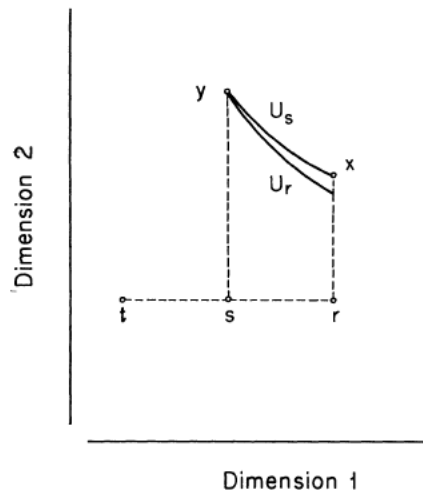
Source: Kahneman & Tversky (1991)

The graph shows two indifference curves evaluated with respect to the commodity y. In the Us curve, y is evaluated from s to r (implying a decreasing amount of y) while in the Ur case, y is evaluated from r to s (positing a growing amount). Loss aversion implies that the indifference curve through y is flatter when y is evaluated from s (Us) than when is evaluated from r. This because, if an individual is loss-averse, he would accept a decrease in commodity y only if compensate with a bigger amount of commodity x.

Both reference dependence and loss aversion (together with the flaming and reflection effects) are at the base of the Prospect Theory of Kahneman & Tversky and possible drivers of the so-called Status quo bias.

- Preference for immediacy: big decisions such as savings, labour supply, educational investments, health and diet, and drug use, have different costs and benefits according to the specific point in time in which they are taken. A central issue in economic is to compute the trade off between performing an action now or at a certain point in the future. The standard assumption states that people weight their future utilities by an exponentially-declining discount factor $\delta^t = \frac{1}{1+r}$ where $0 < \delta^t < 1$ and r is the discount rate. Behavioral economics asserts the importance of the so-called Immediacy Effect, according to which discounting is dramatic when individuals delay an action that would otherwise be immediate. The main driver in this kind of effect has been identified in people's self-control. Empirical tests have shown how a simple hyperbolic time discounting factor of $d(t) = \frac{1}{1+rt}$ can better describe this effect. This implies that people will make good decision when planning in advance (all costs and benefits are in the future) but short-sighted decisions when some costs and benefits occur immediately. But a time-inconsistency is created. Indeed, in the exponential discount factor case, an individual would take the same decision prospectively as he would when the time for a decision finally arrives, if the same choice and information are present. This does not happen in the hyperbolic case since, when the future arrives, the individual may behave against its earlier purposes, preferring immediate consumption rather than long-run well-being.

To solve the problem of inconsistency, Laibson (1997) proposes the "quasi-hyperbolic" time discounting function. The hyperbolic function is approximated by two parameters, β and δ (where β represents the immediacy effect), in which the weight on current utility is 1 and the weight on period-t instantaneous utility is $\beta\delta^t$ for $t > 0$. Therefore, we have the basic exponential time discounting plus an immediacy effect. This way, a person would discount delays in consumption equally at all moments over time except in the current one, providing one simple but powerful model for measuring the taste for immediate gratification.

- Fairness and social preferences: the assumption that individuals act to maximize their wealth caring about just self-interest has been widely used in economics to

simplify models. However, many times individuals can use their wealth to reward those who helped them, punish who damaged them or to make outcomes fairer. Thus, people are driven by the social utility of their action. Understanding these social preferences is required since this new assumption implies a change in the demand for goods and in the utility function of individuals (Camerer & Loewenstein, 2003).

- Market salience: regards the specific case of how taxes are presented. According to the standard economic assumption, persons respond to the substance of a cost irrespective of its presentation. However, tax presentation affects market decisions and economic activity. If persons do not alter their behavior after the tax is applied there is no loss of efficiency (i.e., deadweight loss). Lower salience taxes help to reduce efficiency losses (Gamage & Shanske, 2011).

Except for the market salience, Camerer (1999) explains how these behavioral economic assumptions have been integrated into some classic economic theories. A table which summarized its work is reported in Figure 2.6.

Figure 2.6: Rational and Behavioral Principles

Rational principle	Behavioral principle	Psychological foundation
Expected utility $\sum_i P_i u(X_i)$	Prospect theory $\sum_i \pi(P_i) u(X_i - r)$	Psychophysics, adaptation: loss-aversion, reflection, mental accounting, nonlinear $\pi(P_i)$
Equilibrium (mutual best response)	Learning, evolution	Generalized reinforcement, replication by fitness
Discounted utility $\sum_t \delta^t u(x_t)$	Hyperbolic discounting $u(X_0) + \sum_{t=1} \beta \delta^t u(X_t)$	Preference for immediacy (temptation)
Own-payoff maximixation $u_{X_2}^1(X_1, X_2) = 0$	Social utility $u_{X_2}^1(X_1, X_2) \neq 0$	"Spend" money on other people (reciprocate, dislike inequality)

Source: Camerer (1999)

The expected utility theory asserts that people evaluate risky investments by weighting the utility of an outcome (Xi) by its probability (Pi), denoted as $\sum_i Pi * u(Xi)$ where u is the function that measures the value of an outcome. Expected Utility assumes that the outcome is integrated into individuals' overall wealth, and if two persons have a common likelihood of a common outcome, that outcome is counteracted when deciding between the two. Expected utility is at the base of theories such as asset pricing, corporate structure, educational investments and purchase of insurance. The behavioral alternative is represented by the Prospect theory. People adapt to past experiences and weight probabilities in a non-linear way. Adaptation entails that utility is determined by gains and losses from some reference point r, and not by the overall wealth. Additionally, rather than aggregating it all together, persons "mentally account" for money in separate categories.

Loss aversion implications are also added to the model. Studies by Kahneman & Tversky (1991) have shown how losses are disliked about twice with respect to gains equal in size. People look for risk when the investments involve only losses (when the best they can reach is zero loss) whereas they avoid risk when investments may lead to gains. The big difference between hypothetical buying and selling prices for nontraded goods (e.g., environmental damages) can be explained by loss-aversion. Evidence has demonstrated how individuals tend to overestimate low probabilities. This fact helps to explain why people invest money in low-probability events (e.g., lottery tickets, slot machines) and to insure against low-probability catastrophes that are not easily justified by expected utility.

Equilibrium is a common concept in economics: the market is in equilibrium when the demand meets the supply; in a strategic game, the equilibrium is reached when all agents choose the optimal alternative (i.e. Nash Equilibrium). According to the Game Theory[7], we have a Nash Equilibrium when players have no reasons to change their chosen strategy after considering the expected opponent's choice (Bernheim & Winston, 2013). A classic example of strategic game is the so-called "Prisoner's Dilemma": two members of a criminal gang (A and B) are arrested and imprisoned in two different rooms without the possibility of communicating. The police officers explain them that they have proofs only to accuse them of a minor crime, but they promise the freedom for the criminal that confess the other crimes. However, the admission must be relevant (i.e., only one criminal confesses). Consequently, the two criminals need to take a choice: (1) confessing, betraying this way the partner or (2) being silence. The payoffs are the following:

- If both A and B confess, the admission is not relevant and both criminals will be incarcerated for e.g., nine years;
- If A confesses while B is silence, B will be incarcerated for, e.g. ten years while A is free (and vice versa);
- If they both are silence, they will be accused for the minor crime and incarcerated for one year.

The prisoner's dilemma somehow represents a deviation from perfect rationality principle since it demonstrates how two rational individuals do not always cooperate, even if it is in their best interest to do so. Indeed, when both players have a dominant strategy, namely a choice that is the best regardless of the other player's possible choice, non-optimal

---

[7] Branch of mathematics and economics that deals with the analysis of strategies between two or more players who are in competition among them and whose outcome critically depends on the choices of the other participants (Bernheim & Winston, 2013).

equilibriums are possible. The numerical representation of the payoffs reported in Figure 2.7 may be useful to understand these concepts.

Figure 2.7: Payoffs in Prisoner's Dilemma

According to Figure 2.7, the numbers on the left side represent the A's payoffs (expressed in years of detention). We can see how, in the case that A thinks that B will not confess (the left column), he has an incentive to betray him since his payoff would be bigger than in the silence case (i.e., 0 against -1). Again, if A expects that B will talk, he has an incentive to confess too in order to make B's admission irrelevant and avoiding this way one year of prison (i.e., -9 against -10). Thus, A has a dominant strategy since "confess" leads always to a better outcome independently from the expected behavior of B. The same reasoning can be done under the B's perspective. As a consequence, the equilibrium is represented by (-9;-9) which is not efficient and, for this reason, it cannot be considered a Nash equilibrium. Indeed, both individuals would gain from collaboration because, if both A and B do not confess, they would be incarcerated only for one year instead of nine.

Going beyond the Game Theory, theories on learning from others, population evolution and rules of individual learning derived from experimental observation suggest parsimonious principles of equilibration. Generally, people learn by "reinforcing" strategies that performed well or would have performed well if they had been chosen.

Laboratory experiments contribute to demonstrate how individuals do not care only about the self-utility maximization when taking a decision, but they might act driven by altruism or revenge desires. One of these experiments is reported by Camerer and Thaler (1995). In the so-called "ultimatum games", a player can offer a portion of e.g., $10 to another player while keeping the rest for himself. The responder can either accept or refuse the offer leaving both them with nothing. If a player has a wealth-maximizing purpose, he should accept anything hence the first player has a convenience to offer a little portion. What is surprising from an economic point of view is that in most countries, players

routinely offer about 4$ to 10$ whereas offers of $2 or less are refused half. Mathematical theories of social utility integrate these findings by assuming that people like to reciprocate and do not like allocations where they earn different amounts than others.

However, this may not be the only explanation. Additional studies on Ultimatum Game have been carried out. Sociologists use Ultimatum Game to illustrate the concept of honour and the human unwillingness to accept injustices. Moreover, it has been used to analyse how psychiatric and personality disorders such as depression or anxiety can affect decision-making during social interactions (e.g., Grecucci *et al.*, 2013). A research conducted by Paz *et al.* (2017) makes use of the Ultimatum game to study the effect of self-esteem in social interactions. Self-esteem is defined ad the degree to which an individual likes, values, and accepts himself and it represents a relevant trait of the human personality since it affects the individual's health and quality of life (e.g., low self-esteem persons are more likely to develop disorders such as depression or substance dependence) other than social interactions. The experiment carried out by Paz *et al.* (2017) did not find any relevant difference in the rejection rate between low and high self-esteem individuals. Nevertheless, it was able to demonstrate how women with a low self-esteem tend to react in a more aggressive way, showing strong anger reactions, to unfair offers than individuals with a higher self-esteem while no relevant results were found for men. The study suggests that low self-esteem women are more reactive to social stressors, yet the necessity of more works about self-esteem and sex on emotional reactions is acknowledged.

We here abstain from further explanations about discounted utility. The topic will be again covered in the subchapter 2.4 from a neuroeconomics' point of view and in Chapter IV since delay discounting processes are at the base of behavioral economics' explanations of drug addiction.

2.3 Behavioral Economics and Health

In 1.1, we discussed how the medical-care market presents different features and issues with respect to a classical commodity market. Frank (2004) proposes how some behavioral economics intuitions could be applied to health economics, in particular, with respect to the behavior of physicians.

We said that the behavior of physicians departures from the classic profit maximization purpose. Therefore, physician market experiences an excess of demand. A first behavioral economics explanation could be that physician deliberately set prices below the market

equilibrium in order to reach a "target income" because maximizing behaviours are seen as socially undesirable.

Further deductions can be made analysing the physician's profit maximization problem proposed by McGuire (2000). Physicians are highly-educated and skilled professionals and they are often in position of clinical uncertainty. Evidence has also demonstrated how they are habitudinal when making medical choices and are slow to adapt to new technologies and practices that would improve the quality of care. A patient is assumed to benefit from the receipt of medical care (m) according to the function $B = B(m; e; u)$, where e represents the physician effort and considers the asymmetric information between patient and medic. It is assumed to be not observable and non-contractible. u represents the uncertainty (or random error) of the medical decision. The patient is assumed to take a Bayesian decision[8] since it can only infer the physician effort observing m and the outcome of the treatment (B0). The inference can be seen as the probability that a high level of effort was taken, $L_1 = (B_0, m)$ while the likelihood that the appropriate effort was not taken is $L_0 = 1 - L_1$.

The physician's profit maximization problem in presence of asymmetric information is given by $\pi = n(L_1)[R + (p - c)m]$, where n is the number of patient visited, and the term within parenthesis is a general statement about revenues and costs that takes into consideration the continuum of payment arrangements commonly applied to pay doctors.

The model assumes that the patient chooses according to its expectation about the doctor's effort. Behavioral economics can be useful to investigate the ways these expectations are developed and their implications. Results from surveys show how persons tend to rely on family, friends and their usual physician when taking this kind of decision. Psychology uses the term "availability heuristic" to describe the reliance on more vivid or memorable evidence to construct the expectation about the likelihood that a physician will perform an adequate effort. However, reliance on family and friends' advice might produce distortions in perceived quality because healthcare consumers tend to be more aware of certain attributes of care (e.g., clarity in explaining clinical issues, attentive) and less of others (e.g., technical quality of care). Health care sector offers also a favourable context for the "law of small numbers" bias: the reports of family and friends are usually based on

---

[8] Bayesian inference is a statistical inference method in which probabilities are not interpreted as frequencies or proportions but rather as levels of trust in the accomplishment of a certain event. Since not all information is available, the patient can only try to estimate the probability of a high (or low) physician effort basing on the elements he can observe, namely, the outcome of the treatment and the receipt (Frank, 2004).

a small number of encounters with a physician. An individual might overestimate the extent to which information from a relatively small number of events represents the experience of the "population" from which those events are drawn.

Arrow (1963) also stressed the concept of "trust" in the patient-doctor relationship. Trust is a non-market institution that aims to decrease the inefficiency by reducing the disutility due to the uncertainty and the anxiety of illness, as well as the agency problems caused by the informative asymmetry between medic and patient. Studies have demonstrated how the trust between patient and doctor remains very high also in case of poor outcome of the treatment since it is generally imputed to events that go beyond the physician's control (Hall, 2001). Since the outcome of the treatment represents an element for the evaluation of the physician's effort, the role of trust cannot be neglected.

A last point embedded in the behavior of physicians regards how these take their choice among a complex set of alternatives to treat a patient or disease. The implications are similar to the ones seen for the patient choosing a doctor. The Bayesian physician is assumed to combine new knowledge about technologies and treatments with prior information, giving each case equal weight to form the posterior distribution. The result is given by a weighted average of past observations from medical training and subsequent practice. Again, availability heuristic implications are possible when the personal advice of colleagues and peers about how apply a specific medical technique plays an important role.

Another behavioral construct that may help to explain why other opinions receive such high weight is "regret". Physicians commonly have to choose among competing treatments for caring a particular condition and they are often largely responsible for the consequences of their complex choices. Many of them, are taken under a relevant time pressure (i.e., the average medical visit lasts 15 minutes). Therefore, the potential regret is very high, and physicians have an incentive in trying to lower the responsibility costs due to medical decision-making. Physicians are influenced by the professional norm, which provides a guideline for their choice. In case of poor outcome, the physician that practiced according to norms can assert that he had followed standard operating practices whereas they could more easily judged as guilty if they departed from local practice.

However, behavioral economics implications on health economics have not been formally acknowledged yet (Frank, 2004). Nevertheless, behavioral economics has been recognised able to explain some issues and anomalies related to health. Before analysing in detail five of them, some behavioral economics' concepts need to be briefly discussed in addition to those described in section 2.2.

- Endowment effect and Status quo bias: evidence demonstrated how, when people come into possession of something, they develop a sense of ownership toward it and tend to over evaluate their goods. This way, people hold onto it even when not convenient and avoid moving toward other alternatives when advisable. A similar behavior is represented by the Status quo bias which refers to a preference for the current status of affair that can drive to bad decisions (Kahneman & Tversky, 1991).

- Framing effect: concerns the way in which the probability of an event is described. Studies suggest that people perceive probabilities and situations in different ways according to how the information is communicated, affecting the decision-making process (Kahneman & Tversky, 1979).

- Decision fatigue: very often, economic requires that individuals need to take decisions under the same circumstances each time. However, after making decision after decision, the brain gets lazy and stop working in an efficient way, making the decision maker more susceptible to bad decisions (Rice, 2013).

- Too much choice: consumers think that they are good at making decisions, in particular, the ones who are not, in fact, skilled. However, experiments have demonstrated how people are apparently overwhelmed when there is a vast amount of possibilities among which to make choice. More choice tends to lead people toward low-yield, less risky investments, and to adopt conservative strategies that are generally not recommended (Rice, 2013).

- Nudge: the term stems from the libertarian paternalism idea, which asserts that persons should be given choices, but the final decision making should be guided by experts who nudge persons toward the choice that is in its best interest (Thaler & Sunstain, 2008).

The complexity of most health care related decisions implicates that people often do not act in their best interest. Failing to enrol in health insurance or engaging in dangerous behaviours without consideration for future implications (e.g., smoking, consuming junk food) are some examples. It follows that these behaviours are in contrast and difficult to justify through classical economic tools. Under perfect rationality assumptions, little or no explanations are provided for these bad decisions.

Behavioral economics has only recently been applied to health decisions. However, it provides a vast set of tools that are useful to understand these anomalies under an economic point of view. When dealing with health care, people do not always gather all

necessary information, nor use them properly. Very often, there is a lack of self-control that spur people to put too much value in the present enjoyment, take myopic decisions and repeat past mistakes to the detriment of future well-being. Rice (2013) analysed five health topics where people tend to misbehave, providing both behavioral economics justifications and solutions.

The first topic regards the donation of organs. In a world dominated by traditional economic rules, the only driver in the decision of donating or not would be the psychic benefit the person derives (in this case, the fact of helping someone). However, most countries have an undersupply of donors. Johnson and Goldstein (2004) provide evidence about how the willingness rate is almost 100% in Austria but just 16% in the near Germany; 86% in Sweden but just 4% in Denmark. The most controversial situation is in the United States where we have a 79% in countries such as Alaska and 1% in Montana. The cause of these differences has been identified in how the decision is framed. In countries with very high rates, an individual is assumed to be a donor unless it explicitly indicates the unwillingness to participate. On the contrary, in countries with low rates there is an opt-in system and only by explicitly giving the permission an individual is considered a donor. Since the biggest source of donors is represented by victims of traffic incidents, a compromise could be found in the requirement of explicitly indicate if a person wants or not be a donor at the time of driver's license renewal. Such policy has been applied in United Kingdom. Here, behavioral economists' thinking splits. One may argue that an opt-out system represents a government overstep and that an individual that does not explicitly perform the effort of opting-out actually wants to be a donor. But the opposite reasoning can be done: if persons do not explicitly choose to be donors, maybe they still would want to be after a fatal accident but they might not fully understand this option when they are alive or they prefer not to think about this option because too painful to consider.

The second subject is the enrolment in government-sponsored health insurance programs. Data shows how about five million of children in US that could be insured through the Children's Health Insurance Program (CHIP) are not because parents have not signed up. This phenomenon is almost intractable using traditional economic tools. However, behavioral economics offers some insights stating that low enrolment rates are not due lack of interest on the part of parents but should be found in possible difficulties in understanding the costs and benefits of alternatives, the big amount of choice, a preference toward the present over the future, a devaluation of the risk faced, and how persons frame the issues in their own minds due to enrolment procedures, along with administrative

roadblocks imposed by state agencies. Automatic enrolment or, at least, no interpersonal interviews could represent possible solutions to the problem.

A similar problem regards the prescription drug insurance plan. Since 2006 in US, by enrolling in private health insurance plans, people can obtain new benefits for the prescription of drugs (e.g., 75% of premium cost covered by the federal government). However, a problem arises due to the sheer of choices available: people need to choose among 30 different drug plans with the consequence that almost nobody selects the most cost-effective plan for itself. Despite individuals have the opportunity to change plan each year, few persons do it, showing a preference for the status quo. Typical economic solution would be providing people with more information, but this approach has proved to be ineffective. Behavioral economics suggests rather to apply the opposite method, namely, to provide just the most salient information and making it more comprehensible for individual (e.g., using symbols rather than numbers), in addition to a reduction of the number of options.

Behavioral economics also deals with the reduction of tobacco smoking. Traditional economic approaches such as taxation on cigarettes or providing information about detrimental effects on health have proved to be moderately effective in lowering tobacco use. Nevertheless, smoking rates remain very high (21.9% worldwide in 2016 as seen in the subchapter 1.3.2 of this thesis). A behavioral economics intervention would be signing a contract to stop smoking and providing reward to those who do it. Another initiative could be encouraging the use of electronic cigarettes. The behavioral assumption is that is easier to get people engaged in related behaviours rather than eliminate them completely. Electronic cigarettes provide a dose of nicotine just like the normal ones but the bad effects on health are limited with respect to traditional smoking.

Finally, the last topic is obesity. Traditional economic theory asserts that the increase in the obesity rate is due to a fall in the price of (junk) food with respect to other goods and services. Therefore, people may have substituted food for other commodities, resulting in weight gain. Even if coherent with the neo-classical view, this explanation seems unlikely if applied to reality. Behavioral economics sees obesity as an irrational choice of individuals, not aimed at maximizing the utility and influenced by food-producing companies which try to change persons' tastes toward high-calorie and fat foods. Thus, marketing strategies are drivers of obesity through direct advertising and product placement (i.e., positioning snacks and sweets near cashier, marketing can exploit decision fatigue: when customers reach it, they have already made a big amount of decisions on what to buy, and they suffer from ego

depletion). Behavioral economics suggests remedies aimed at contrasting and offsetting the pervasive power of marketers. As we will better see in Chapter III, this can be done through a direct application in the real world of the so-called Nudge Theory. Experiments have demonstrated that also monetary incentives can be useful to affect behaviours: immediate cash can be implemented for affecting difficult to change habits.

2.4 How the Study of the Brain Can Inform Economics: the Neuroeconomics

Behavioral economics is now becoming even more a reference point in economics due to its applications in each topic such as finance, game theory, law, public finance and macroeconomics. The biggest contributors to behavioral economics are represented by the cognitive and social psychology, namely, those branches of psychology that study how behavioral decisions (both individual and within groups) are taken. However, other cognitive sciences have been developing in the recent years. Among them, important findings surely come from neuroscience.

Neuroscience is an interdisciplinary discipline focused on the study of the nervous system. It merges knowledge from different fields of study such as biology, medicine, chemistry, physic, statistics, linguistics and, of course, psychology to infer details about how the brain works. Neuroscience has been able to reshape what is believed about psychology, which in turn had an impact on behavioral economics. When this occurred, in the late '90s, the new field called neuroeconomics was born (Loewenstein *et al.*, 2008).

Given its interdisciplinary and complex nature, we refrain from going to much in depth about neuroeconomics. Explanations about the brain and the nervous system functioning would be required to fully understand most of the topics. We here provide just some basic insight about decision-making and theoretical frameworks provided by neuroscience and neuroeconomics mainly basing on the works of Camerer *et al.* (2005) and Loewenstein *et al.* (2008).

In a certain sense, we can define neuroeconomics as the natural evolution of behavioral economics since it has further unified the once disparate fields of economics and psychology. The value added that neuroscience gives to economics can be viewed according to an incremental or a radical approach. In the incremental approach, neuroscience suggests variables that should be added to conventional decision-making models or that should replace those assumptions that have never been well demonstrated empirically; on the other hand, the radical approach adopts a "what if" method, asking to

itself how economics could have developed differently if it had been informed from the beginning by the findings now available from neuroscience.

When we dealt with behavioral economics in subchapter 2.1, we stated that the idea of the brain seen an information-processing device posit by cognitive psychology represented, in a sense, the turning point that allowed economics and psychology to reconcile. One of the fundamental of neuroscience is that brain is not a homogeneous processor, but a merger of diverse specialized processes that are integrated in different ways when the brain challenges different kinds of problems. A distinction between automatic and controlled processes was proposed by Schneider and Shiffrin (1977):

- Controlled processes are serial (i.e., they follow a step-by-step logic or computations coherently with the cognitive psychology's theory), are invoked deliberately by the individual when it faces a challenge or surprise and are often related with a subjective feeling of effort. People can typically provide a good introspective explanation of controlled processes meaning that, for instance, if a person is asked to solve a math problem, he/she can easily recall the steps and considerations that have led to the solution.

- Automatic processes present exactly the opposite features with respect to controlled processes: they operate in parallel, are not accessible to consciousness, and, generally, they do not require any effort. Parallelism allows rapid response, multitasking, and gives the brain a notable power when it need to perform certain types of tasks, such as visual identification. Additionally, not being accessible to consciousness, people often cannot provide any explanation about how a certain automatic judgement or decision have been taken.

Neuroscience also deals with in which area of the brain automatic and controlled processes happen. Since the study of the brain goes beyond our purposes, we here limit to say that those regions that bear automatic processes are concentrated in the back (occipital), top (parietal), and side (temporal) parts of the brain whereas controlled processes happen mainly in the front (orbital and prefrontal) parts. Figure 2.8 can be used to better understand and identify these areas.

Figure 2.8: economically relevant areas of the human brain
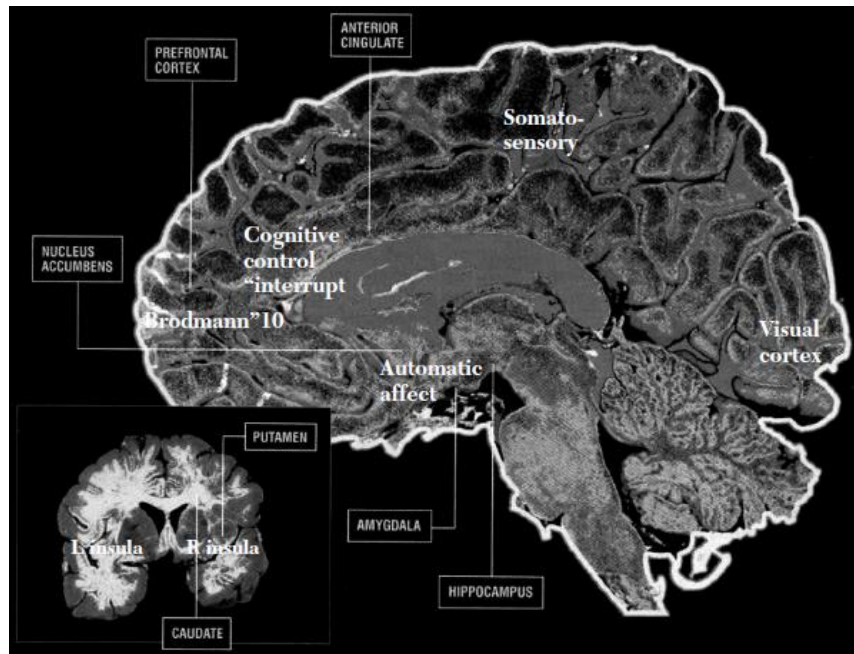
Source: Camerer *et al.* (2005)

A further distinction is represented by affective and cognitive processes. Affective processes are those processes that answer to "go/no-go" issues, namely, that motivate approach or avoidance behaviours. In contrast, cognitive processes, are those that address true/false questions. To better explain these two concepts, we can assert that cognitive processes cannot by themselves produce an action: to influence behaviours, cognitive system needs to operate though the affective system. In this case, affect encompasses emotions (e.g., anger, fear, anxiety, jealousy) but also needs such as hunger, thirst and sexual desire, and motivational states (e.g., physical pain, discomfort, drug craving).

The combination of these two dimensions (i.e., controlled-automatic processes and affective-cognitive processes) results in the formation of four quadrants, labelled from I to IV as shown in Figure 2.9. Quadrant I (cognitive controlled processes) represents those processes that are activated under the individual's willingness and, therefore, introspective access is possible. Actions in quadrant I may require a big cognitive effort to be activated; quadrant II (affective controlled processes) is the rarest in pure form. Camerer *et al.* (2005) recognizes the difficulty to observe this kind of processes and provides the example of actors who imagine previous emotional experiences so as to actually experience those emotions during a performance; quadrant III (cognitive automatic processes) encompasses all those action that the brain has learn from past cognitive experiences (e.g., the skier that suddenly change direction to avoid an obstacle in the track); quadrant IV (affective automatic processes) is contraposed to the first since the processes that characterize it

occur in an automatic way and without a cognitive-rational base. They can affect in a preponderant way decisions, perceptions and behaviours. Quadrant IV is the one that makes an individual jump when somebody on its back screams "Booo!".

Figure 2.9: The Two Dimensions of Neural Functioning

	Cognitive	Affective
Controlled Processes ▪ serial ▪ effortful ▪ evoked deliberately ▪ good introspective access	I	II
Automatic Processes ▪ parallel ▪ effortless ▪ reflexive ▪ no introspective access	III	IV

Source: Camerer *et al.* (2005)

Neuroscience heavily rely on technology for its studies and it includes methods such as brain imaging, single-neuron measurement and electrical brain stimulation[9] in order to observe the brain area of interest. Neuroscience findings can be possibly applied to each theory developed within the framework of behavioral economics (i.e., expected utility, discounted utility, equilibrium and own-payoff maximization), as we have seen in the subchapter 2.2 of this thesis. Here, we discuss just the neuroeconomics implications on the intertemporal choice that will be useful in Chapter 3 to deepen the economic explanation of drug addiction.

The main shortcoming of the Discounted Utility model typically used in economics to explain intertemporal choice is the assumption of exponential discounting. This posits that a certain time delay drives to the same discounting regardless of when it takes place. Behavioral economics stresses the concept of self-control and propose a quasi-hyperbolic discounting model to address traditional economics limits as explained in subchapter 2.2 of this thesis. Neuroeconomics takes a further step. The studies on brain have some important implications on interpersonal choice, that is seen as the result of both affective and cognitive processes.

---

[9] In brain imaging images are taken (through electro-encephalogram or magnetic resonance) while subject is performing different tasks. Images are then compared to provide a picture of regions of the brain that are differentially activated. Single-neuron measurement and electric brain stimulation are restricted to animals. The first involves tiny electrodes inserted into brain; the latter involves brief electric brain stimulations as a reward for e.g., rats. Evidence has shown how many abused drugs lower the limit at which animals will lever-press for electric brain stimulation (Camerer *et al.*, 2005)

The affective system is designed to assure the survival of the individual. In most animals, emotions motivate those behaviours that have a short-term purpose, such as eating. Thus, the affective system, that humans share with most animals, leads to choices that are mainly myopic. Despite some animals can display long-term behaviours such as storing the food for the winter, humans are the only among animals who carries out immediate sacrifices for desired future consequences. The capability of considering future implications when making a decision seems to be a result of our prefrontal cortex, a part of the brain that is uniquely human (Figure 2.8). Indeed, patients with damages on the prefrontal region tends to behave more myopically. Thus, intertemporal choices can be viewed as the result of impulsive and affective (more deliberative) processes and far-sighted (more cognitive) processes driven by the prefrontal cortex.

The notion of hyperbolic time discounting implies that people behave impulsively when the right combination of incentives is available. Even if neuroeconomics does not reject the hyperbolic time discounting model, it asserts that this is not always the case. Instead, neuroeconomics proposes the idea that impulsivity in decisions depends on the competition between affective and cognitive systems and determine those factors that strengthen or weaken one or the other.

Prefrontal cortex is that part of the brain that deals with the subjective feeling of effort. Any factor that requires an effort of the prefrontal cortex, so of the controlled and cognitive system, should decrease the influence of this system and, therefore, the individual control on its own behavior. An example in how cognitive system can be manipulated is provided by Shiv and Fedorikhin (1999). They divided a group of people into two sub-groups. A group were asked to memorize a two-digit number, the other to memorize a seven-digit one. Subjects were then asked to walk to another room where a table presented a choice between a caloric piece of cake and a fruit salad. The experiment shown how the 59% of the individuals who had had to memorize a the seven-digit number before entering in the second room chose the piece of cake whereas only the 37% individuals of the other group made this choice. The results of the experiment imply that the effort required to memorize seven-digit numbers takes deliberative resources off from self-control, driving those subjects who need to remember more to choose cake.

Furthermore, a prior exercise of self-control tends to reduce the capability and, therefore, the propensity to perform self-control in the present. Studies demonstrated how subjects on diet, once the diet is over, are spurred to consume more fat food than previous and also give up earlier when faced with an intellectual problem they couldn't solve.

Finally, the activation of affective states should accentuate temporal myopia. Research has shown, for instance, that addicts show a higher discount rates, not only for drugs but also for money, when individuals are craving drugs than when they are not. However, exceptions of this rule are possible when there is an interaction between cognitive and affective systems. Decisions in delay often involve a mixture of cognition and affect. Cognitive awareness of the delayed positive consequences is required but it is not enough; emotions play an important role in the choice of delaying decisions since people care about the delayed benefits of their decisions only to the extent that contemplating such benefits evokes immediate affect.

But, how can these neuroscience findings be useful to develop an economic model on time preferences? First, neuroscience research provides a way to "unpack" the time preference concept. Ability to think about future pros and cons is important, which is probably the reason why time preferences are correlated with measured intelligence. Second, a key point is represented by the fact that people are more likely to perform myopic actions when affected by powerful emotions and drivers. This suggests that research should focus on which kind of situations get individuals more prone to act irrationally. Third, individuals have different willpower, namely, the availability of the scarce internal resource that allows individuals to inhibit emotional driven behaviours.

A neuroeconomics model that encompasses both cognitive and affective implication can be useful to explain not only impulsivity, but also why some persons display opposite self-control issues with respect to those types typically examined (e.g., workaholics who do not want to take a break). These behaviours can be explained by the human propensity to experience emotions, such as fear and anxiety, when thinking about the future. Indeed, one of the instruments used by our brains to impose self-control when affective drivers would otherwise drive toward myopic self-destructive behaviours is to create a "deliberative affect" through directed thought and imagery.

This kind of framework may also explain why people act inconsistently when their behavior is analysed through the lens of discounted utility. The capacity to forecast future consequences may not be strongly correlated with the extent to which different experiences produce emotional reactions, and these in turn may not be necessarily correlated with an individual's willpower. Loewenstein *et al.* (2001) have found close to zero correlations between large decisions that have an intertemporal component, but high correlations between behaviours that seemed to draw on the same dimension of intertemporal choice (e.g., which required the inhibition of specific feelings such as anger).

Chapter III: Psychology and Behavioural Economics' Theories on Health

The purpose of this chapter is to present some frameworks and theories both from psychology and behavioral economics literature. We will discuss two cognitive frameworks, the first is prescriptive and is called Health Belief Model; the latter, the Transtheoretical Model, is analytical (section 3.2). Then, we will discuss other theories that can be applied to health decisions. Some of them, such as the framing effect and the nudging, have already be cited in subchapter 2.3 of this thesis. Instead, the topic of the drug abuse from a behavioral economic point of view will be deeply covered in Chapter IV.

3.1 Health Belief Model

According to Granz *et al.* (2008), one of the most widespread models used to explain and predict health-related behaviour is the Health Belief Model (or HBM). In a similar way to the Theory of the Planned Behavior of Ajzen (1991), also HBM provides some key variables that should be modified to move toward a certain behavioral outcome. The framework was developed by the social psychologists Rosenstock, Hochbaum, Kegeles and Leventhal at the U.S. Public Health Service to find a justification to the failure of people to participate in programs to detect and prevent diseases (e.g., tuberculosis) in the early 1950s. Before analysing the model in detail, a visual representation of this framework can be useful to simply its explanation.

Figure 3.1: Health Belief Model
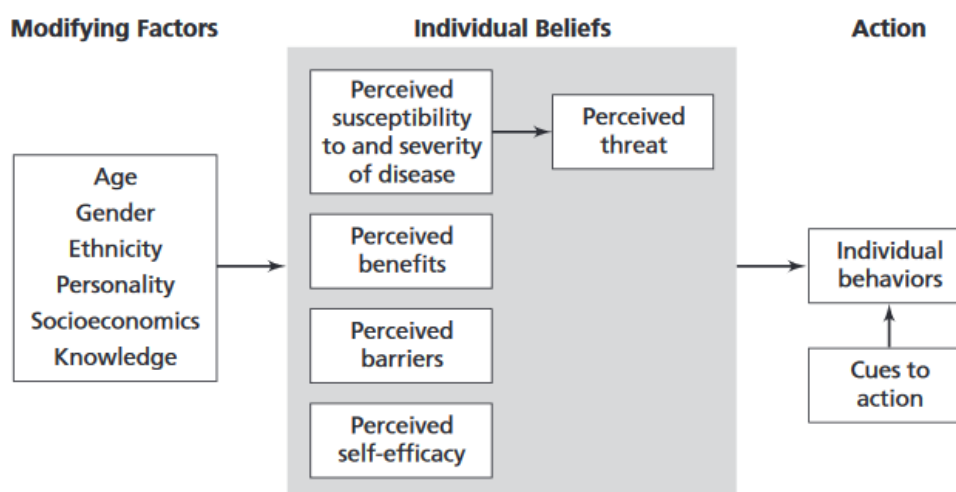
Source: Glanz *et al.* (2008)

Figure 3.1 represents one of the last versions of the model. The model is organized into three categories: modifying factors, individual beliefs and action. On the left side, we find

modifying factors which include demographic and sociopsychological variables that might influence perceptions and, therefore, indirectly affect health-related behaviours. For instance, education is believed to affect how individual beliefs are perceived by individuals.

In the middle, there are individual beliefs, namely, (1) perceived susceptibility represents the belief about the probability of contracting a disease or a condition; (2) perceived severity are perceptions about the seriousness of getting an illness or of leaving it untreated. It includes evaluations of both possible medical (e.g., death, disability, pain) and social consequences (e.g., work, family life, relationships). The combination of susceptibility and severity forms the perceived threat; (3) perceived benefits is the belief about the efficacy of the advised medical action to reduce the risk or the seriousness of the disease. Also, non-health-related perceptions (e.g., financial savings of quitting smoking) can affect the final behavior. Therefore, individuals who have positive beliefs about susceptibility and severity would not accept any suggested health action unless they also perceive that the action is potentially beneficial by reducing the threat; (4) perceived barriers are the potential negative aspects of a particular health action that act as impediments to undertaking recommended behaviours (e.g., high costs, side effects, unpleasantness); (5) perceived self-efficacy is the conviction that an individual can successfully perform the behavior required to generate the outcomes.

On the right side of Figure 3.1, there is the action part. The term "cues" refers to strategies to activate "readiness". Readiness to perform an action (perceived susceptibility and perceived benefits) can only be enhanced by other factors, particularly by cues to induce actions (e.g., bodily and environmental events).

As displayed in Figure 3.1, arrows indicate the relationships among the components. Modifying factors might influence health perceptions. Individual beliefs about health represent the major construct of the model and are included within the central box. Both modifying factors and cues to action can affect these perceptions. The combination of individual beliefs leads to the final behavior/decision. Although the Health Belief Model provides those constructs that drive to the outcome behaviours, relationships among these variables are not well defined. This ambiguity has allowed different applications and approaches of the HBM interpretation. Many studies have tried to set each of the major dimensions as independent, others have tried multiplicative approaches. One of these studies was performed by Janz & Becker (1984). They performed a critical review of different HBM studies carried out between 1974 and 1984. Aggregate results suggest that perceived barriers are the most decisive predictor of behaviours. Moreover, even though

both perceived susceptibility and perceived benefits are overall important, perceived susceptibility plays a more crucial role in preventive health behavior than for sick-role behavior. The reverse is true for perceived benefits. Finally, perceived severity seems the less powerful driver. Nevertheless, severity has proved to be strongly related to sick-role behaviours. Analytical approaches to better identifying these relationships would be needed to boost the precision of the HBM in predicting health-related decisions.

3.2 Transtheoretical Model

The transtheoretical model considers the temporal dimension and the stages of changes in order to integrate processes and principles of change from psychotherapy and behavior change. The model has been largely used to study topics such as smoking cessation, weight management, depression prevention and stress management. According to the framework, health behavior change goes through six stages of change, namely: precontemplation, contemplation, preparation, action, maintenance, and termination as shown in Figure 3.2.

Figure 3.2: Transtheoretical Model's Stages
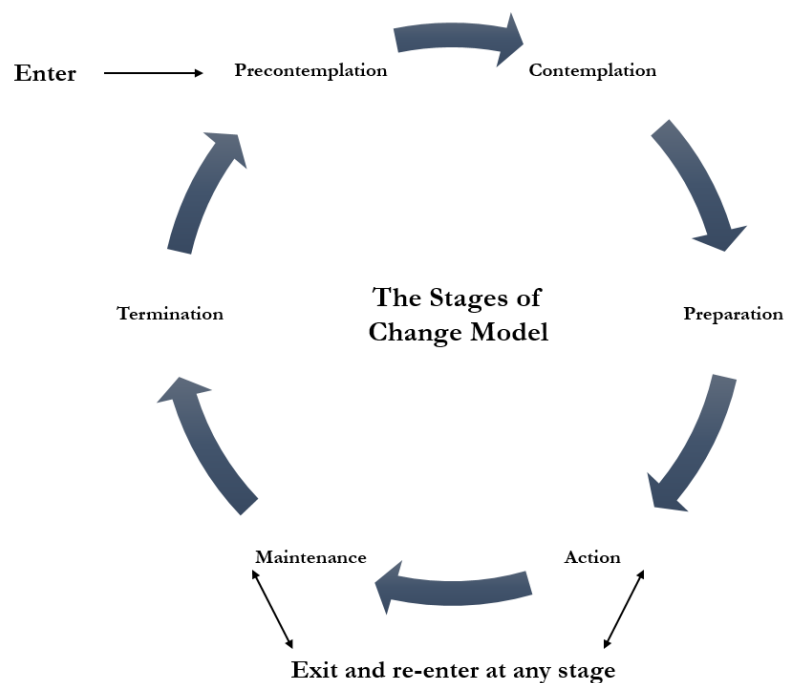
Source: Prochaska & Velicer (1997)

In the first stage, precontemplation, people are not willing to take any action in the foreseeable future (the next 6 months) because they may be under-informed about the consequences of their behavior or they are demoralized due to past attempts to change a dangerous habit. In this stage, people tend to avoid talking, thinking or reading about their risky behaviour. The second stage, contemplation, is the stage in which people are willing to change in the short-term (6 months). They are aware of both pros and cons of changing.

69

The trade-off between costs and benefits of changing can produce ambivalence that may keep people stuck in contemplation for longer periods of time leading to chronic contemplation or procrastination; in determination (or preparation) people are intending to act in the immediate future (in the next month). They typically have already developed a plan of action, such as joining a health education class or talking to their physician; during the action stage, people have made specific changes in their life styles in the previous six months. Not all changes of behavior are relevant according to the transtheoretical model; the criterion is that the change must be sufficient to reduce the risk for diseases. In the maintenance stage, people work to prevent a possible relapse. However, they do not apply change processes as frequently as do people in action since they are now less tempted to give up and increasingly more confident that they can continue in their processes. Finally, in the termination stage, people have zero temptation to return to the old habit and 100% self-efficacy regardless possible averse events such as depression, anxiety, boredom or rage.

The model also recognizes the role of different processes of change, namely, covert and overt activities that people use to move from one stage to the next. In an empirical study, Prochaska & Velicer (1997) have identified the most relevant ten. (1) Consciousness Raising concerns increased awareness about causes, consequences and treatments of a dangerous behavior. Media campaigns, education and bibliography can be used to increase individuals' awareness; (2) Dramatic Relief initially produces increased emotional experiences followed by reduced affect if appropriate action can be taken. (3) Self-reevaluation how people image themselves with and without a particular unhealthy habit, e.g. one's image as a couch potato and an active person. (4) Environmental Reevaluation encompasses both affective and cognitive evaluations of how the presence or absence of a personal habit can have an impact on the one's social environment (e.g., passive smoking). (5) Self-liberation represents both the belief that one can change and the commitment to act on that belief. (6) Social Liberation is the increase in social opportunities or alternatives that derives from stopping a certain behavior. The burden is relevant especially for people who are relatively deprived or oppressed. (7) Counterconditioning is the adoption of healthier behaviours that can be used as substitute for problem habits. (8) Stimulus Control regards removing cues for unhealthy habits and adding prompts for healthier alternatives. (9) Contingency Management concerns the use of incentives or punishments for taking steps toward a particular direction. Evidences highlights how self-changers rely on rewards much more than punishments. (10) Helping Relationships means providing caring, trust, openness, and acceptance as well as support for the healthy behavior change.

Additionally, Prochaska & Velicer (1997) stresses the concepts of decisional balance, self-efficacy and temptation. The first regards the way in which people weight pros and cons of changing; the second represents the self-confidence that people have about coping a specific risky situation that may lead to relapse; finally, temptation is the intensity of urges to engage in a dangerous behavior when difficult situations occur.

3.3 Fear Appeal Theory

Fear appeal is a term used in psychology, sociology and marketing that generally indicates a strategy that uses pictures and sentences for motivating people to carry out a certain action, buy a certain product or keep a determined behavior by arousing fear. There is a wide range of frameworks that take into account fear appeal elements (e.g., Fear-as-acquired-drive model, Nonmonotonic model), the same Health Belief Model discussed in subchapter 3.1 of this thesis considered them when dealing with perceived susceptibility and severity. Without getting too much into details, fear appeal theories and models have been largely used to address many public health related issues such as smoking cessation, reduction of alcohol usage while driving, condom usage to reduce sexual communicable diseases, and more.

Three key independent variables drive the theory: fear, perceived threat and perceived efficacy. Fear is identified as a negative emotion, accompanied by a high level of arousal; Perceived threat is composed by two element: (1) perceived susceptibility of the threat, hence how much an individual feels itself exposed to a certain threat and (2) perceived severity of the threat, namely the expected negative impact of the threat. Fear and threat are reciprocally related: the higher the perceived threat, the higher the fear experienced. Also perceived efficacy is composed by two dimensions: (1) perceived self-efficacy, the belief of an individual to be able to perform recommended responses and (2) perceived response efficacy, the belief of an individual about the efficacy of the response in averting the threat (Rogers, 1983).

A study conducted by Witte & Allen (2000) uses meta-analysis[10] to identify the impact of fear appeal in affecting intentions and behaviours. The first analysis investigates how people perceive fear appeal messages in term of fear, susceptibility, severity, self-efficacy and response. The results of this analysis are reported in Figure 3.3.

---

[10] Meta-analysis is a quantitative technique that synthetizes the outcomes of a particular group of studies in order to produce an average result for different variables across the literature (Lanzini, 2018).

Figure 3.3: Effects of Fear Appeal Message Features on Perceptions

	Fear	Severity	Susceptibility	Response Efficacy	Self-Efficacy
k	51	33	29	24	17
N	12,735	5,531	4,731	4,739	2,731
r	.297*	.439*	.301*	.358*	.361*
95% confidence interval	±.089	±.109	±.077	±.113	±.117
χ^2	276.31*	200.31*	140.73*	263.35*	99.72*

*$p < .05$.

Source: Witte & Allen (2000)

According to Figure 3.3, k represents the number of studies analysed, N the participants. The table indicates that the stronger the fear appeal is, the greater is its impact on fear aroused, severity and susceptibility perceived. In the same way, there is a greater impact on response efficacy and self-efficacy. The heterogeneity[11] found in additional analyses is due to the fact that individual investigators largely vary in their fear appeal manipulations. Tests indicate a significant correlation between the year of the study and the size of the manipulation. Therefore, more recent studies should indicate larger manipulation effects than older studies ($r = 0.13, p < 0.05$).

Figure 3.4 displays the second part of the study, which is focused on how the fear appeal dimensions can alter attitudes, intentions and behaviours.

Figure 3.4: Effects of Message Feature on Attitudes, Intentions, and Behaviours

	Fear	Severity	Susceptibility	Response Efficacy	Self-Efficacy
Attitudes					
k	34	14	11	11	8
N	7,514	2,195	1,606	1,849	1,348
r	.147	.149	.119	.139	.116
95% confidence interval	.022	.041	.048	.044	.052
χ^2	79.99*	49.23*	50.86*	40.72*	44.85*
Intention					
k	43	26	27	24	21
N	9,686	3,923	3,948	4,348	3,873
r	.126	.142	.171	.166	.173
95% confidence interval	.020	.030	.030	.029	.030
χ^2	212.95*	81.80*	105.08*	99.29*	82.10*
Behavior					
k	28	16	11	12	11
N	4,666	2,528	1,797	1,608	1,475
r	.158	.132	.138	.129	.126
95% confidence interval	.028	.038	.045	.048	.050
χ^2	142.69*	34.28*	22.23*	16.74	20.46

*$p < .05$.

Source: Witte & Allen (2000)

---

[11] Heterogeneity in meta-analysis indicates the percentage of variation across studies' outcomes (Lanzini, 2018).

The table shows how all the features of the message (fear, severity, susceptibility, response efficacy and self-efficacy) result in a higher level of attitude, intentions and behavior change. The stronger the fear appeal, severity, susceptibility, response efficacy and self-efficacy in the message the stronger the attitudes, intentions, and behaviours toward the recommended response. Response efficacy and self-efficacy display homogeneous effects for behavior change whereas all other effects are heterogeneous. Heterogeneity indicates that we should interpret the average correlation carefully because a moderator variable affecting acceptance of a message may exist.

3.4 Framing Effect

The Framing Effect was first analysed by Kahneman & Tversky (1979) while developing the Prospect Theory. The Prospect theory suggests an utility function where people avoid risks (i.e., they are risk adverse) when considering gains but they prefer risks (i.e., they are risk seeking) when considering losses. As a consequence, the final decision about taking or not a risk also depends on whether the option is expressed (i.e., framed) in terms of gains or losses. Hence, the same problem can be framed positively, in terms of gain, or negatively, affecting this way the final decision. The classic problem reported by the two psychologists is the case of an epidemic that is expected to kill 600 persons and two options are offered to the decision-makers. When the decision is framed positively, the two options are expressed in terms of number of lives saved. On the contrary, when they are framed negatively, the options are expressed in terms of number of lives lost. People may prefer a certain option when expressed in terms of lives saved but reject the same option in favour of the riskier one when expressed in terms of lives lost.

The prospect theory also deals with a similar phenomenon called Reflection effect. The difference between reflection and framing is that, while the first implies preferences for a gain or a loss which are symmetric (e.g., winning or losing $1000), the latter concerns preferences for a decision that is framed to seem as either a gain or a loss.

Framing implications should be considered when developing health-related messages. In particular, we need to distinguish between different consequence framing and same consequence framing:

- Different consequence framing: if positively framed, the message emphasizes the positive consequences of taking a certain action (e.g., if a cancerous growth is detected, 19 of 20 growths are the less deadly nonmelanoma cancer), whereas if negatively framed, the message emphasized the negative consequences of taking

that decision (e.g., if a cancerous growth is detected, 1 of 20 growths are the more deadly melanoma cancer).

- Same consequence framing: instead of referring to the positive and negative consequences of performing the same action, the positive and negative frames describe the same consequences in terms of either carrying out or not carrying out a behavior. In this case, a positive framed message could be e.g., if you have a skin cancer examination, a cancerous growth can be detected before it becomes life threatening. On the contrary, the negative framed message is e.g, if you do not have a skin cancer examination, a cancerous growth cannot be detected before it becomes life threatening.

While same consequence framing can be used for every situation, different consequence framing can be applied only if the decision lead to specific negative and positive consequences (e.g., life vs death).

Studies on surgical preferences have largely used different consequence manipulation where the benefit or the negative consequences of surgery were described. On the contrary, Meyerowitz & Chaiken (1987) used same consequences framing in which both the positive and negative message emphasized the importance of performing breast self-examination. The way in which framing manipulations are applied are relevant because they may lead to an explanation of apparently contradictory findings. Indeed, positive framing increases the preference for the surgical option in the first kind of studies, however, negative framing has shown to be more useful in motivating breast self-examination.

Generally, for health-related behaviours, negative framing can lead the individual to feel worried about its present health status. Then, the individual may perform a more risk-seeking behaviour, as suggested by the prospect theory. On the other side, positive framing, emphasizing on gains and benefits, provides reassuring information that may lead the individual to be more risk averse in its choices. Rothman *et al.* (1992) offer the example of using sunscreen to prevent skin cancer. Preventive behavior helps to avoid bumping into risky events and, according to the prospect theory, the individual will choose the more risk averse option when a positively framed information is presented.

In their study, Garcia-Retamero & Galesic (2010) tried to find a way for reducing the effect of the frame effect in health-related decisions. The research was based on a computerized questionnaire carried out by 987 German and American participants. Numerical skills have shown to play a relevant role in the study of framing. Indeed, low numeracy participants perceive a surgical procedure as less risky when the risk is expressed

in terms of surviving rather than dying whereas there is no difference in perception among high numeracy individuals. Again, communicating relevant health information through visual aids make almost disappear framing in participant with low numerical skills. However, not all visual aids are effective in the same way: pie charts and bar graphs are more useful in reducing framing than icon arrays (i.e., icon arrays require a more precise and quantitative processing to be interpreted whereas pie charts and bar graphs are more "immediate"). Finally, Garcia-Retamero & Galesic (2010) also state that if the purpose is to persuade people rather than improve their informed decision making (e.g., cessation of smoking), positive framing is more useful in case of disease prevention or treatment selection while negative framing is more effective for disease detection.

3.5 Status Quo Bias

Like the Framing Effect, also the Status Quo Bias is considered in the Prospect theory of Kahneman & Tversky (1979). Most real decisions encompass a possible status quo alternative, which means doing nothing or maintaining the previous position. According to Samuelson & Zeckhouser (1988), the effect can be seen as a consequence of:

- Rational decision making: under many points of view, status quo bias is perfectly consistent with the neo-classical theory. An individual who takes the same decision more times may be justified by the fact that he is facing independent and identical decision settings (i.e., preferences and alternatives are always the same or at least similar). Therefore, rationality would impose to repeat always the same choice. Furthermore, transaction costs might explain the bias when decisions are not independent (i.e., when the initial choice affects preferences or choices in the following decisions). In this situation, transaction costs lead to status quo when the cost of switching is bigger than the efficiency improvement associated with a better alternative. Additionally, uncertainty can be a further driver toward the same decision when the utility provided by e.g. alternative products is unknown and a customer needs to try them in order to find out it. Optimal decision-making imposes a cut-off strategy, namely, a person repeats the same choice if the utility it supplies is sufficiently high, otherwise he tries something new.

- Cognitive misperceptions: recalling for a moment what have already been discussed in subchapters 2.2 and 3.4, Kahneman & Tversky (1979) have demonstrated how individuals tend to weight losses more heavily than gains when making decisions. In particular, people are loss averse with respect to gain but are

risk seeking when dealing with losses. Let's consider the situation in which an individual must choose between the current situation and a new one. If assumptions about loss aversion are true, taking the current situation as reference point, he would weight potential losses from switching heavier than potential gains. Then, the individual will decide not to change due to the fear of lose what he currently own (i.e., endowment effect). However, status quo also largely depends on how the situation is framed (i.e., in terms of losses or gains). Nevertheless, status quo bias has been shown to be present in situation where there are not explicit gain/loss framing effects (e.g., budget problem, car colour choice, job choice).

A second cognitive misperception is called Anchoring. It takes place when a person considers a certain value of an unknown quantity before assessing such quantity. The value he has considered or that has been shown to him before, strongly determines the assessment he is going to make, which will be always relatively close to that previous value, which is labelled as the anchor.

- Psychological commitment: psychologic and economic research has demonstrated how individuals' decisions are affected by sunk costs[12]. When sequential decisions need to be taken, status quo choices may be justified by the individual's reluctance to "cut his losses" or by the wish to not go against previous commitments to a (perhaps failing) course of action by making subsequent commitments. Thus, the presence of sunk costs (or other resource invested and non-recoverable) contributes to status quo bias in decision making. The greater the investment in the status quo alternative is, the less likely is a deviation from the previous taken decisions.

Regret avoidance is another factor contributing to psychological commitment. People try to avoid consequences of decisions which have proved to be wrong even if those decisions appeared to be correct given the information available at the time. Evidence from Kahneman & Tversky (1982) argues that individuals feel a stronger sense of regret for bad outcomes when these are the result of new actions taken than from inaction. Therefore, avoiding decision regrets is a

---

[12] Sunk costs represent the costs in which an individual has already incurred and cannot recover them (Bernheim & Winston, 2013).

possible justification of status quo bias. Individual may also try to avoid cognitive dissonance[13], searching for consistency when dealing with consequent decisions.

Therefore, Status quo bias may represent a problem when dealing with consecutive decisions. How this cognitive bias is applied to the health sector? In subchapter 2.3 we discussed how individuals tend to have a preference for the current situation when they need to choose their drug insurance plan. In particular, the presence of a big number of options seems to represent a further issue that might further worsen the status quo bias. Indeed, this kind of cognitive preference arises when dealing with health insurance choices. Further evidence of this is provided by Samuelson & Zeckhouser (1988) and displayed in Figure 3.5.

Figure 3.5: Transfer among health plans

	1984/1985	1985/1986	1986/1987
Total Transfers	770	385	330
as % of all enrollees	8.1%	3.8%	3.6%
Net Transfers by Plan			
BCBS	−575	−93	−127
HUGHP	−27	+12	+10
HCHP	−2	−24	+16
MGHP	+16	+4	+8
Bay St	+60	+57	+34
Tufts	+9	+9	+14
Lahey	+7	+10	−4
BC Low	+52	+23	+49
Total Net	0	0	0

Source: Samuelson & Zeckhouser (1988)

The experiments involved 9,185 employees from the Harvard University enrolled in eight different health plans. In 1980, only four plans were available namely Blue Cross/Blue Shield (BCBS), Harvard University Group Health Plan (HUGHP), Harvard Community Health Plan (HCHP), and Multigroup Health Plan (MGHP). The Lahey plan became available in 1982, Bay State and Tufts in 1984 and BC Low in 1985. At the beginning of the decade, the status quo was firmly represented by the BCBS plan which embedded the 62% of all enrollers, followed by HUGHP (31%) and HCHP (6%). However, even when more plan became available, we can see how individuals tend to maintain the previous choices. In the period between 1984 and 1985, only the 8.1% of the employees change health plan and this figure is even smaller in the following two periods (3.8% and 3.6%). The comparison of health plan choices supplies strong evidence of status quo bias. Old enrolees continue in

---

[13] Cognitive Dissonance is the mental discomfort that arises when individuals' beliefs go against their behaviours and/or new information that is presented to them (Festinger, 1957).

selecting the current plan (i.e., BCBS) while enrolments in the newer plans (and also in HUGHP and HCHP plans) are less frequent (even if they better match needs of most enrollees). The very low rate of transfer among plans is further evidence of status quo inertia. Nonetheless, a flaw in this study is present since little or no bias is evident in transfers between BCBS plans.

3.6 Nudge Theory

Thaler & Sunstein (2008, p.6) define nudge as «any aspect of the choice architecture that alters people's behaviour in a predictable way without forbidding any options or significantly changing their economic incentives». The definition stem from both behavioral economics and social psychology to justify why people act in ways that go against the economic assumption of perfect rationality. Indeed, very often people value their health but persist in behaviours that may undermine it. The gap between values and behavior can be explained by using a dual process model where human decisions are shaped by both a cognitive and an affective system. The cognitive system is a reflective and goal-oriented and it is controlled by values and intentions. Hence, it needs cognitive capacity or thinking space. The latter is automatic and requires little or no cognitive engagement. The affective system is mainly driven by immediate feelings and triggered by the surrounding environment. The purpose of nudging is that of leading individual to right decisions despite the affective system (Marteau *et al.*, 2011)

The definition also stems from the Libertarian Paternalism concept. We have already discussed the behavioral economics' theory of the asymmetric paternalism in section 1.4.2 which was based on the purpose of providing large benefits for irrational individuals, while imposing little or no burden on those individuals who are fully rational (Camerer, 2003). Despite for many economists the term libertarian in association with paternalism represent an oxymoron, libertarian paternalism wants to improve the directions of people's choices allowing both private and public intervention while maintaining the freedom of choice (Thaler & Sunstein, 2003).

However, the original definition of nudging excludes legislation, regulation, and interventions that alter economic incentives. In fact, one of the most appealing features of nudging is that it provides simple and low-cost solutions that do not necessarily require legislation or government intervention to be applied and are suitable to a wide array of problems arising from human behaviour. Figure 3.6 provides some examples of nudging applications to common health-related issues. The table also reports the classical regulation

provided by institutions. For example, making salad and vegetables the default side dish rather than chips in restaurants can be seen as an alternative to ban industrially produced trans fatty acids. In the same way, removing cigarettes, lighters and ashtrays from the view of smokers might substitute policies aimed at increasing the price of cigarettes.

Figure 3.6: Nudging and Regulation Actions

	Nudging	Regulating
Smoking	Make non-smoking more visible through mass media campaigns communicating that the majority do not smoke and the majority of smokers want to stop	Ban smoking in public places
	Reduce cues for smoking by keeping cigarettes, lighters, and ashtrays out of sight	Increase price of cigarettes
Alcohol	Serve drinks in smaller glasses	Regulate pricing through duty or minimum pricing per unit
	Make lower alcohol consumption more visible through highlighting in mass media campaigns that the majority do not drink to excess	Raise the minimum age for purchase of alcohol
Diet	Designate sections of supermarket trolleys for fruit and vegetables	Restrict food advertising in media directed at children
	Make salad rather than chips the default side order	Ban industrially produced trans fatty acids
Physical activity	Make stairs, not lifts, more prominent and attractive in public buildings	Increase duty on petrol year on year (fuel price escalator)
	Make cycling more visible as a means of transport, eg, through city bike hire schemes	Enforce car drop-off exclusion zones around schools

Source: Marteau *et al.* (2011)

Evidence seems to confirm nudging usefulness. A meta-analysis carried out by Arno & Thomas (2016) demonstrated how alterations of the choice architecture, namely, changes in the environment (e.g., olfactory or social), perception (e.g., emotional priming), availability of food (e.g., portion size), or knowledge-based changed (e.g., labelling) can increase in average the 15.3% in healthy dietary or nutritional choice. Again, providing information on the healthy behaviour of others (i.e., social norm feedback) is particularly useful to limit the consumption of alcohol among students.

Nevertheless, a dark side of nudging exists. As any intervention, nudging has the capacity to generate harm other than benefits, arising from perverse response to nudges. For instance, proposing healthier dishes as the default, or put too much emphasis on the healthy feature of foods can lead to Halo Effect[14] resulting in an excess consumption since people may underestimate the energy content. In a study carried out by Chernev (2011), people were asked to estimate the number of calories presented in a hamburger. Individuals were separated in two groups: one group needed to estimate the calories when the hamburger was served alone, the other group had to perform the estimation when the

---

[14] Halo Effect: cognitive bias that occurs when one trait of a person or thing is used to make an overall assessment of that person or thing. It enhances rapid decisions, even if biased ones. A simple example can be that of assuming an individual as intelligent at first glance just because good-looking (Marteau *et al.*, 2011)

hamburger was served together with some healthy food (e.g., three celery sticks). Results show how the average estimation for the first group was 697 calories whereas for the second group was 642. This is just an example of why evaluations about possible paradoxes or unexpected effects should be considered when developing a nudging strategy.

Moreover, nudging has been criticized by some scholars. Critiques mainly regard the short-termism of nudge strategies that does not enable a long-term behavior change. Goodwin (2012) also states that in UK, nudge goes against the coalition government's ambitions to encourage ideas such as empowerment, freedom and fairness. Additionally, he argues that it is not a good strategy to solve the big problems of the society such as climate change or public health. Instead, the government should adopt more deliberative models of democratic and public engagement, which are proved to encourage people to act more collectively and involve with issues more deeply, providing this way the opportunity for more substantive citizen empowerment. Goodwin (2012) also criticizes the paternalistic aspect of nudge, reprimanding that, although it tries to leave a sense of freedom to individuals, the fact that it exploits the imperfections of human judgement makes it something quite manipulative.

In conclusion, we can assert that nudging has encouraged policymakers to adopt alternative methods (i.e., altering environments) that differ from regulation to change people's behaviour. These developments are to be welcomed since they provide simple and low-cost alternatives applicable from both private and public institutions. However, there is lack of evidence in support of the effectiveness of nudge strategies as a means to improve population health and reduce health inequalities, especially in the long-term.

Chapter IV: Toward an Economic Explanation of Drug Abuse

4.1 A Neo-Classical Attempt: The Model of Rational Addiction

It is not correct to assert that classical economic theories do not deal with issues such as drug abuse. When behavioral economics theories were not fully recognized yet, an early economic attempt to explain addictions were performed. The most-known neo-classical theory about addictions takes the name of Model of Rational Addiction. The assumptions supporting the model are pretty far from the more recent ones of behavioral and neuroeconomics even if the two approaches share the role of future implications in decision-making. Indeed, the model of rational action tries to justify irrational choices such as drug abuse without denying classical economic assumptions (e.g. perfect rationality).

The model, developed by Becker & Murphy (1988), can be possibly applied to any kind of addiction other than drugs (e.g., alcoholism, smoking, work, technologies). Addictions are defined in a non-psychological way, yet as a causal effect of past consumption on current consumption. Authors recognize how they represent the antithesis of rational behaviours. In fact, "rational" is intended with the meaning that people maximize utility consistently over time (i.e., they have a consistent plan to maximize utility in the long-term), and a commodity may lead to addiction if increases in past consumption raise current consumption.

The starting assumption of the model is represented by the fact that the individual's utility in any moment depends on two good, c (the addictive good) and y. However, these goods defer for the feature that the current utility depends on the measure of the past consumption of c (through a "learning-by-doing"[15] process) but not from that of y, so that $u_t = u(y_t; c_t; S_t)$ where u is assumed to be concave and S represents the consumption capital. The investment function assumed for the present is given by $S_t = c_t - \delta S_t - hD_t$, where S is the rate of change over time in S, c is the past consumption (or investment in "learning"), δ is the instantaneous depreciation rate that measures the exogenous rate of disappearance of the physical and mental effects of past consumption of c, whereas D_t represents expenditures on endogenous depreciation or appreciation. Finally, assuming a life's length equal to T and a constant rate of time preference, σ, the utility function of an individual is given by:

$$U_0 = \int_0^T e^{-\sigma t} u(y_t; c_t; S_t) dt$$

---

[15] In economics, "Learning-by-doing" refers to a process through which improvements in efficiency are achieved through practice, self-perfection and innovations (Becker & Murphy, 1988).

The maximization problem of a rational individual is solved by imposing a budget constraint on its expenditure. Let's consider A_0 the initial value of assets and r the interest rate (which is assumed constant over time), if earnings at time t are a concave function of the stock of consumption capital at t, w_S, and assuming perfect capital markets, the budget equation is given by:

$$\int_0^T e^{-rt}\left(y_t + p_t^c c_t + p_t^d D_t\right)dt \le A_0 + \int_0^T e^{-rt} w(S_t)dt$$

Where y has a constant price over time. Thus, an individual in order to maximize its utility, it needs to take into account this budget constrain and the already mentioned investment equation $S_t = c_t - \delta S_t - hD_t$. The result obtained from the optimal solution, $V(A_0; S_0; w, p)$, represents the maximum reachable utility given by the initial assets A_0, the initial stock of capital S_0, the earnings function w_S, and the price structure p_t. Finally, the optimal path of y_t and c_t are computed through the first-order conditions from the equation:

$$a_t = \int_t^T e^{-(\sigma+\delta)(\tau-t)}u_S\, d\tau + \mu \int_t^T e^{-(\sigma+\delta)(\tau-t)}w_S\, d\tau$$

Where $\mu = \frac{\partial V}{\partial A_0}$. The function a_t represents the discounted utility of the individual. Monetary costs and benefits of the additional commodity, c, are computed through the implications on future stocks measuring this way the shadow price[16] of an additional unit of stock. The consumption of a good is considered dangerous if $u_S; w_S < 0$ and, consequently, it would have an adverse effect on future utility and earnings. On the contrary, a commodity is beneficial when its consumption leads to $u_S; w_S > 0$ with positive effects on future utility and earnings.

The discounted utility formula is the first relevant result of the model of rational addiction. Becker & Murphy (1988) also dealt with the dynamics of the model. However, we herein refrain from doing so in order to avoid going to much in depth in the mathematics underlying them. We limit to present the second relevant result of this model, namely, the optimal path of the capital stock, identifiable by the following formula:

$$S_t = de^{\lambda_1 t} + S^*$$

---

[16] Shadow price is the price assumed for a good or service for which there is not a market price (Becker & Murphy, 1988).

Where $\lambda_1 = \frac{-\sigma - \sqrt{\sigma^2 + 4B}}{2}$, S^* is the steady state value of S and $d = S_0 - S^*$. Instead, B is equal to $B = \delta(\sigma + \delta) + \frac{\alpha_{ss}}{\alpha_{cc}} + (\sigma + 2\delta)\frac{\alpha_{cs}}{\alpha_{cc}}$ while α_s and α_{ss}'s values depend on the coefficients of the utility and earnings functions of the individuals.

Finally, through further developments and studies of the model, Becker & Murphy were able to arrive to the following conclusions:

- The steady-state consumption of addictive commodities is unstable when the degree of addiction is strong, namely, when there is a strong complementarity between past and current consumption;

- Not all e.g., smokers or drug users are addicts. There are some factors such as the level of income, the price and temporally stressful events that affect the probability of becoming addicted. Furthermore, individuals who discount future in a heavy way are more likely to develop a dependence;

- Change in prices of addictive goods might have a modest short-run impact on their consumption (coherently with what asserted in section 1.4.1). However, their demand in the long-run tend to more elastic with respect to other commodities;

- When an individual is particularly devoted to a substance or behavior (e.g., cocaine, alcohol, gambling), the only way to break it is through Cold Turkey (i.e., by abrupt interruption of the consumption). The necessity for cold turkeys indicates the presence of a weak will or other form of less-than-rational behaviours. Furthermore, cold turkey is consistent with the rational behavior concept since rational individuals end strong addictions only with quick or discontinuous reductions in consumption.

The economic model of rational addiction, and generally all neo-classical models, assumes that delayed costs and benefits are discounted exponentially over time so that, for each unit of time delayed to delivery, the value of the benefit (or cost) decreases by a fixed proportion. Exponential discounting rate is usually identified by the formula:

$$v_d = Ve^{-kd}$$

Where v_d is the present value of a delayed reward, V is the actual value of the delayed reward, k is an empirically derived constant that measures the tendency to prefer smaller but immediate rewards to larger but delayed and d is the delay duration.

However, this way people's preferences remain always the same over time and they will never reverse regardless of the value of k as can be seen in Figure 4.1. This is the big

limitation of economic models in explaining addictions: assuming an exponential discounting factor they do not take into account the possibility of preference reversals that may be driven by factors such as loss of control and impulsive behaviours which are very frequent among individuals, especially if affected by an addiction.

**Figure 4.1**: Present value of delayed rewards using exponential discounting factor
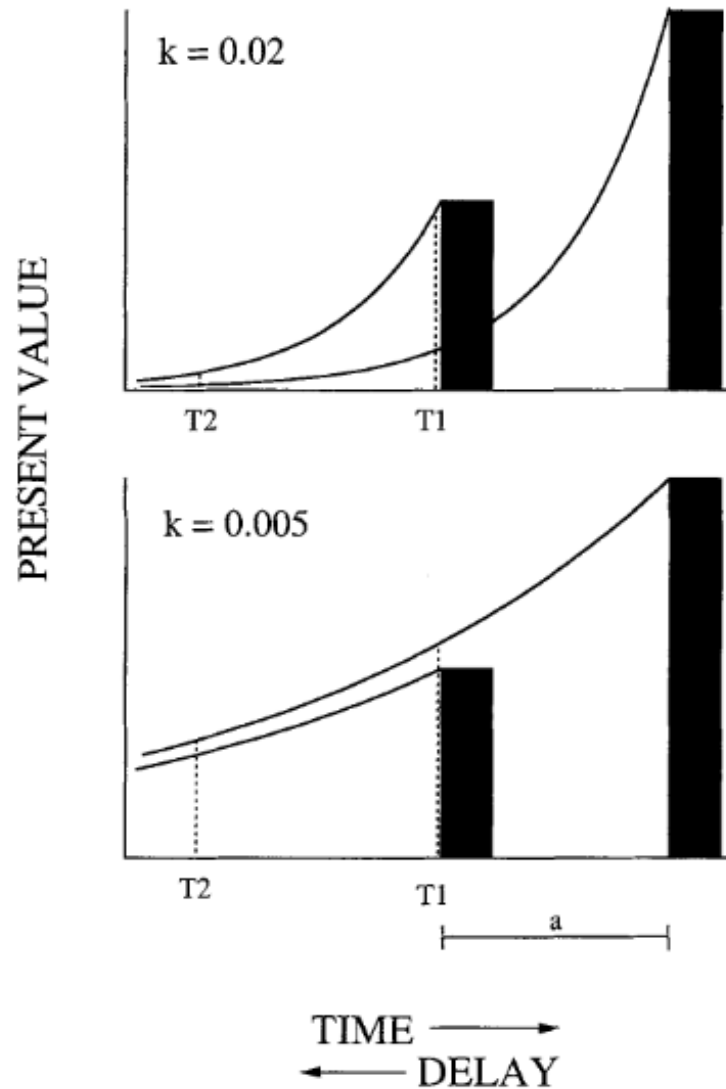
Source: Madden *et al.* (1999)

Figure 4.1 shows how an individual may prefer smaller but more immediate rewards (e.g. $100) than bigger but deferred rewards (e.g. $200) according to the value of k (the higher, the bigger is the preference for more immediate reward). However, the decision taken by the individual at time T_2 is consistent over time and it will never reverse. As we will better see in 4.2.1, this is assumption is unrealistic since this is not always the case.

Nevertheless, neo-classical economists note how exponential discounting can predict preference reversal if two different discount rates are applied to the small but more immediate reward and to the bigger, delayed reward as can be seen in Figure 4.2.

Figure 4.2: Preference reversal under exponential discounting
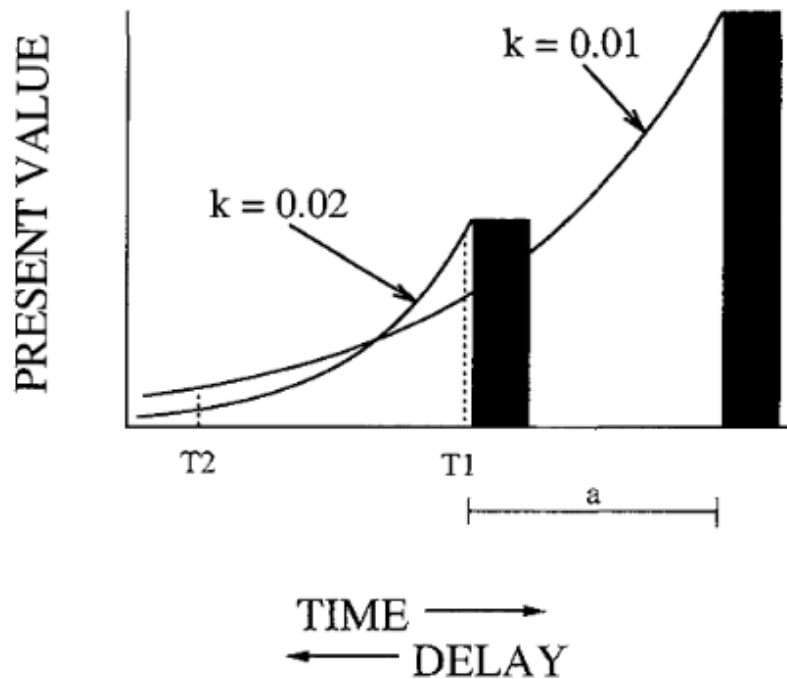
Source: Madden *et al.* (1999)

According to the Figure 4.2, the smaller, more immediate reward is discounted at a rate ($k = 0.02$) that is the double with respect to the one of the delayed reward ($k = 0.01$). This cause preference reversal since in T_2 the individual prefers to wait for the bigger reward but, as the time goes by, it will start to crave the smaller reward because closer in time. This is exactly the situation described in T_1.

Nevertheless, doubts about exponential discounting still persist since it has never been empirically demonstrated by behavioral researches conducted on both human and non-human subjects. On the contrary, researches are in favour of hyperbolic discounting which has proved to better fit to both animals and humans as we will better seen in section 4.2.2 (Ainslie & Haslam, 1992).

4.2 A Behavioral Economics Explanation of Drug Addiction

4.2.1 Impulsivity, Loss of Control and Hyperbolic Discounting

The shortcomings of the Model of Rational Addiction have been overtaken by behavioral economics. Indeed, recognizing the relevance of some psychological elements in affecting individuals' decision-making, behavioral economics is able to provide assumptions and tools that better suit to a topic such as drug abuse. However, like in the neoclassical attempt, also the behavioral approach recognizes the role of delay discounting factors.

One of the purposes of behavioral economics is the study of the allocation of behavior within a system of constrain emphasizing how some conditions can influence the consumption of goods, including drug dependence. According to Bickel & Marsh (2001), two concepts are fundamental in understanding drug abuse from an economic perspective:

- The elasticity of the demand, i.e. the proportional change in the quantity consumed due to a proportional change in price of a commodity. Despite the perfect inelasticity hypothesis of addictive commodities has been rejected by empirical evidence (section 1.4.1), it is also true that the same studies have demonstrated that the assumption was partially true since e.g., cigarettes among smokers and heroin among heroin-addicts are less sensitive to price (i.e., they have a more inelastic demand) than other commodities (Jacobs & Bickel, 1999);

- The second concept to understand is how delayed reinforcers are discounted by individuals. Specifically, discounting of delayed reinforcers explains how the value of a delayed reinforce (that can be, for example, a monetary incentive) is discounted (i.e., how it losses value or it is considered to be worth less) with respect to the value of an immediate reinforcer. The preference of the individual about how much it prefers a reinforcer now rather than in future determines the size of the discount rate for that individual.

Behavioral economics identifies two main behaviours often exhibited by drug-dependent individuals that are able to deeply affect the discount rate. The first is impulsivity that, under a discounting perspective, can be defined as the selection of a smaller more immediate incentive rather than a larger more delayed reward (i.e., the contrary of self-control). Indeed, as stated in subchapter 1.3.1 of this thesis, individuals who suffer the effect of a dependence very often abstain from or reduce occupational and recreational activities in order to consume drugs. Thus, drug addicts tend to prefer a relatively brief, but immediately available, pleasure derived from drug intoxication or they give up to temporarily withdrawal symptoms rather than wait for a variety of prosocial, but often delayed, rewards. In a similar way, intravenous drug users may choose to risk the conveyance of possible diseases sharing needles rather than postponing the drug use until they have time to disinfect the needle or obtain some clean ones. Concerns may regard if drug addicts display the same impulsivity for all choices or if their decisions vary with the kind of reinforcer considered. The scientific literature suggests that impulsivity can be linked to various psychological disorders other than addictions, such as delinquent behaviours, depression, suicidality, aggression and excessive spending. Moreover, many

types of psychological disorders tend to co-occur such as drug and alcohol use, enhancing the possibility of risky sexual behaviours, gambling, depression and personality disorders (e.g. Vitaro *et al.*, 1998).

The second behavior is "loss of control". Drug abusers may have a preference for the larger, delayed rewards. However, when the time to take a decision comes, the individual still chooses the smaller and more immediate reward. Therefore, displaying a reversal in preference. The main difference between loss of control and impulsivity is that the latter does not necessarily imply reversal in preferences since the individual may actually prefer the immediate reward rather than the delayed one. Instead, loss of control can be useful to explain, for instance, why many drug addicts voluntarily participate to outpatient drug-treatment programs, but they still continue to abuse of drugs. At the same way, drug abusers might show a strong preference for employment or relationships with family and friends over drug use, yet a short time later they may relapse and give up to drugs instead of going to work or spending time with their family and friends.

As argued in the end of section 4.1, economic models generally use an exponential discounting factor. However, studies have demonstrated how delay discounting is hyperbolic (Ainslie & Haslam, 1992). In a hyperbolic discounting regime, the devaluation of deferred rewards is directly proportional to their delay meaning that, for each unit of time of delay, the reward's present value decreases by an increasing smaller proportion. The discounting rate can be computed as:

$$v_d = \frac{V}{(1 + kd)}$$

Where, again, v_d is the present value of a delayed reward, V is the actual value of the delayed reward, k is the empirically derived constant that measures the tendency to prefer smaller but immediate rewards to larger, but deferred ones (i.e., the larger k is, the more impulsive, irrational or present-focused the behavior) and d is the delay duration. It may be interesting to note how, in subchapter 2.2, we stated that in order to overcome the time-inconsistency issue created by hyperbolic discounting, behavioral economics also proposes a "quasi-hyperbolic" time discounting function (Laibson, 1997). However, that right there is the feature of hyperbolic discounting that can be particularly useful in explaining the preference reversal due to losses of control often exhibited by drug abusers.

Figure 4.3 shows the present value of two rewards that have been discounted over time through a hyperbolic function. According to the figure, the first reward is smaller (i.e. $100) and it is available after a shorten period of time with respect to the second bigger reward

(i.e. $200). Thus, the two rewards are separated by a hypothetical delay of two years. The two hyperbolic curves display the discounted value (i.e. present value) of the two rewards as a function of time to their availability. The following situation is then possible: an individual who needs to take a choice in T1 recognizes the greater value of $200 and prefers it even if the reward is delayed of two years with respect to the $100 reward (i.e., self-controlled choice). However, what happens in T2 is that, as the time goes by, the $100 becomes even more appealing because more immediately available. Therefore, the present value of the rewards reverses and the $100 has a greater value at the individuals' eyes (i.e. impulsive choice) despite the objective value of the rewards ($100 and $200) and the time between their availability (two years) do not change.

Figure 4.3: Present value of delayed rewards using hyperbolic discounting factor
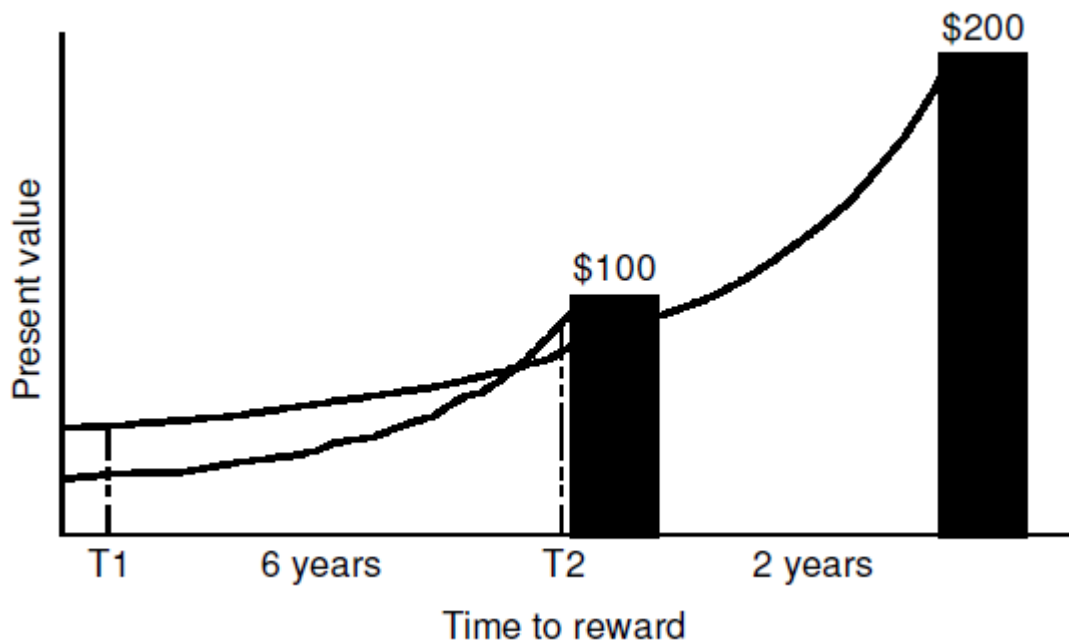
Source: Bickel & Marsh (2001)

Hyperbolic discounting suggests that when an individual has to take a decision about events which are distant in time, the initial choice could be made according to self-controlled, rationality and consistently with the objective value of the rewards (that for a drug abuser could be translate in e.g., "I want to work, be with my family without using drugs"). However, as the smaller sooner reward becomes available, impulsivity and irrationality may lead to preference reversals, leading to an inconsistent choice with respect to both the objective value of the rewards and the initial expressed preference.

4.2.2 Empirical Evidence on Hyperbolic Discounting

Hyperbolic discounting is able to account for both loss of control and impulsive behaviours. A vast literature of studies which examines the delay discounting in human subjects and non-human subject exists. Researches carried out on humans typically employ procedures very similar to the psychophysical ones. Generally, experiments on delay discounting are performed presenting to subjects a choice between a larger but delayed reward (e.g., $1000 deferred of 1 year) and an immediate reward whose amount is subjectively decided by participants to the amount that makes the two rewards to be equal worth (Green *et al.*, 1994). This point of equivalence represents the indifference point for that particular delay interval for a certain individual. An indifference curve can be drawn computing indifference points for a variety of delays, then an indifference curve may be plotted. Indifference curves are important since they allow to empirically determine the shape of the function, and to the rate that discounts delayed rewards to be empirically derived.

Empirical studies have demonstrated that indifference curves are hyperbolic rather than exponential. Furthermore, experiments on both humans and animals (using real and hypothetical money for human and food and water for animals as rewards) have found that the hyperbolic equation is able to model the discounting function accurately, accounting for more than 85% of the variance (Richards *et al.*, 1999).

A relevant empirical research is the one carried out by Madden *et al.* (1997). The study compares the delay discounting rate of opioid-dependent patients in treatment with the ones of community volunteers. The experiment was about choosing between a (hypothetical) monetary incentive immediately available or following by a delay. According to the procedure previously described, delayed reward was $1000 while the immediately available reward was adjusted according to subjective preferences for seven different delays (from 1 week to 25 years). Using the same procedures, the opioid-dependent patients also had to perform a second series of choices involving immediate or delayed heroin[17]. The experiment showed how hyperbolic discounting equation was able to capture from the 80% to 99% of the variance for both opioid-dependent and non-dependent individuals. Additionally, it demonstrated how the discounting rate for monetary rewards was much higher for drug addicts individual than for other participants as can be seen in Figure 4.4.

The graph on the left explains the monetary choice data for opioid-dependent and control participants, whereas the right ones shows opioid-dependent participants' monetary

---

[17] The amount of heroin was determined using street values so that it represented the quantity affordable with $1000 (Madden *et al.*, 1997).

and heroin choice data. Indifference points represent the present value of the delayed rewards namely, the value of the delayed rewards in immediate-reward terms. This way it is possible to compare heroin and monetary rewards on a common axis. The vertical axis shows the value of the delayed reward in percentage.

Figure 4.4: Median indifference points for opioid-dependents and control group
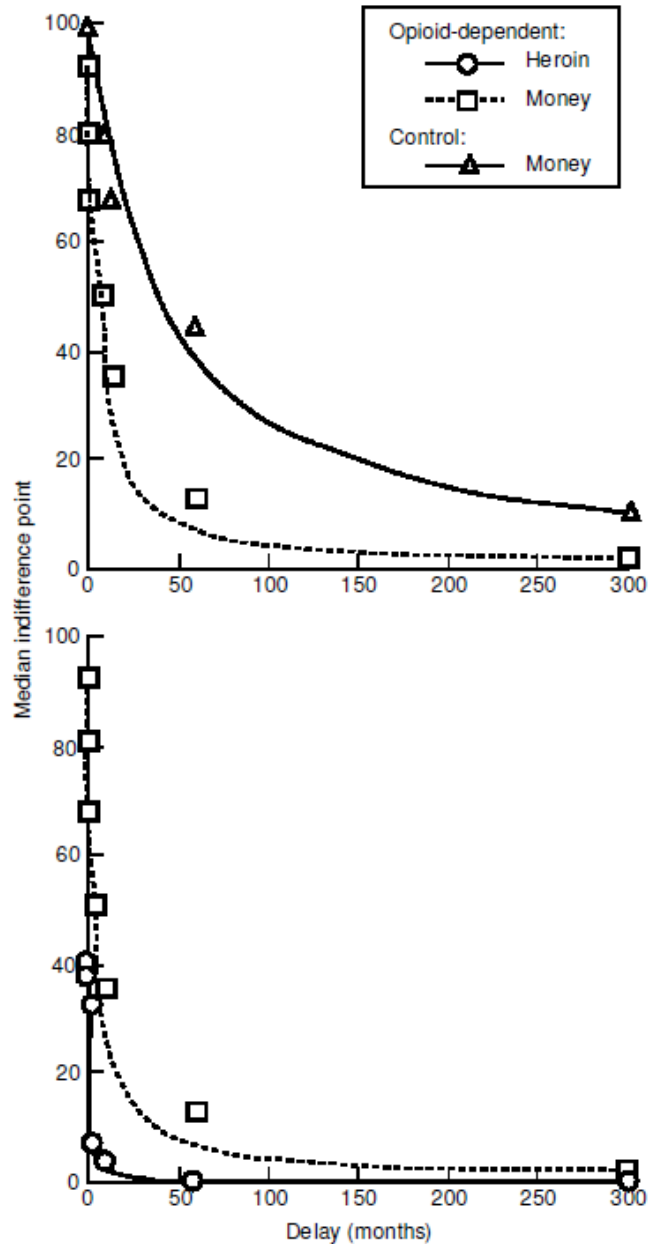
Source: Madden *et al.* (1997)

The graph on top displays how the rate of discounting monetary rewards was significantly greater for the opioid dependent. For instance, a delay of one year is enough to reduce the subjective value of $1000 by 60% for opioid-dependents whereas a delay of 5 years for the control group is necessary to reach the same value. Additionally, the graph below describes how opioid-dependents discount heroin in a way significantly greater than

90

money. Heroin losses the 60% of its value with just a week of delay. These findings lead to the conclusion that the magnitude of discounting is not invariant, but it is affected by the type of reward provided. Moreover, these results demonstrate that the procedures used to estimate delay discounting could be modified to use drugs as a reward. Possible concerns about this study regard the use of hypothetical rewards. Indeed, further analyses conducted by Kirby (1997) demonstrate how hypothetical rewards are expected to provide a proxy of discounting, nevertheless the discount rate might be underestimated.

Additional studies about hyperbolic discounting have been carried out also with respect to other kinds of addiction rather than drugs. Bickel & Madden (1999) compared the discounting of hypothetical monetary rewards among smokers, ex-smokers and never-smokers. Coherently with what found by Madden *et al.* (1997), smokers have been proved to be more impulsive discounting the value of cigarettes more steeply than delayed monetary rewards suggesting that, just like heroin in the previous study, cigarettes present a rapid loss of subjective value when delayed. Furthermore, smokers discount monetary rewards more aggressively than the other two groups. No relevant difference has been found between never- and ex-smokers who showed similar discounting rates.

In a similar way, Petry & Casarella (1999) examine the discounting rate between problem gambling substance abusers and other voluntary participants. The study leads to the conclusion that delay discounting rates of substance abusers with problem gambling are over three times larger than those of non-problem gambling substance-abusing participants. Hence, the abuse of substance and the compulsive gambling may interact additively in affecting discount rates.

Finally, a research conducted by Vuchinich & Simpson (1998) on college students was aimed at examining how heavy and light drinkers discounted hypothetical monetary rewards. Again, the experiment demonstrates how heavy drinkers devalued delayed rewards more than light drinkers. The study also highlights how through hyperbolic equation it is possible to account for significantly more variance than through exponential function developed via economic approaches.

In conclusion, results from empirical studies seem to confirm both the mathematical and behavioral assumptions described in section 4.2.1. Indeed, they were able to prove how hyperbolic discounting is more appropriate than exponential discounting to describe real-life situations. Furthermore, impulsivity and loss of control are proved by the experiments previously reported and these seem true not only for individuals with drug abuse disorders but for all those subjects who suffer of other kinds of disorders (e.g., smoking, gambling).

Nevertheless, an exception exists. A study carried out by Johnson *et al.* (2010) has demonstrated how marijuana dependence is the only commonly abused drug dependence disorder that is proved not to lead to an increased discount rate.

4.2.3 Some Insights from Neuroeconomics

Behavioral economics research mainly focuses on the stimulus input and behavioral output often neglecting the way in which information is processed and decisions are made (i.e., the intermediary steps). Understanding the role of the brain in decision making has made filling this gap possible. Neuroeconomics supplies an important scientific approach that can be applied to a vast set of issues affecting the society, including drug addiction. In particular, through the use of brain-imaging techniques such as functional magnetic resonance imaging (fMRI)[18] or positron emission tomography (PET), neuroeconomics has been able to identify those neural systems related to temporal discounting (Bickel *et al.*, 2007).

Important insights are provided by Bechara (2005). According to his studies, the competing interaction between two competing neural systems are responsible of suboptimal decision making such as drug related decisions. A graphical representation of these two systems is reported in Figure 4.5.

Figure 4.5: Competing Systems in Drug Related Decisions
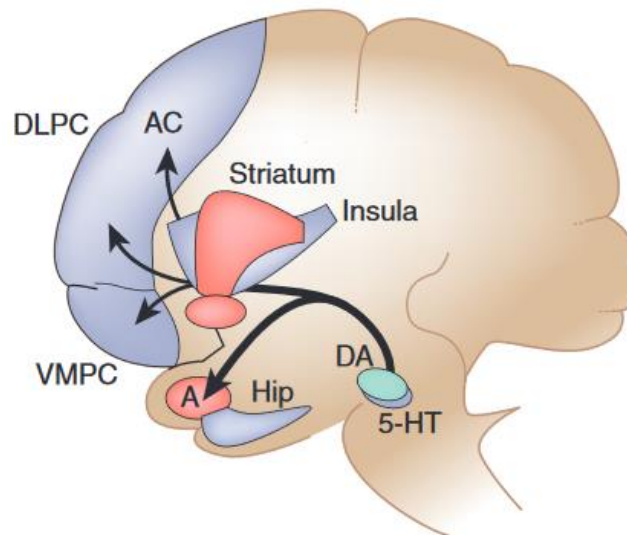
Source: Bechara (2005)

---

[18] Functional magnetic resonance imaging (fMRI) has established as the preferred brain imaging technique because it does not require the injection of any substance, so it is non-invasive, and it provides greater spatial and temporal resolution than PET. fMRI is based on the theory that brain activation requires increased oxygen metabolism, which in turn leads to an increased blood flow to toward activated regions. However, not all the oxygen delivered by the blood is used by activated regions causing oxyhaemoglobin concentration that can be detected by fMRI (Bechara, 2005).

The first system is labelled as "impulsive" and it is reported in red in Figure 4.5. This region of the brain is considered to be essential for reacting to stimuli, initiating a physiological response and decoding emotions. In particular, research seems to confirm that a part of the brain called Amygdala (indicated by the letter "A" in the figure) is an important set of nuclei in indicating the positive or negative valence of immediate outcomes. Indeed, studies on patients with lesions to amygdala show an inability of these to behave in their long-term interest. In a similar way, analyses conducted on comparable areas of rats' brain demonstrate that lesions lead to more impulsive behaviours.

The second system, labelled as "reflective" (or "executive") by Bechara (2005) and underlined in blue in Figure 4.5, includes the prefrontal cortex (PFC). This area of the brain is generally involved in executive functions regarding long-term achievements such as working toward a defined goal, forecasts and expectations about outcomes or consequences of current activities, and social control. PFC is also the likely location where information from various sources (e.g., stimulus, memory of previous events) converges to make decisions about the behavior to keep in order to achieve current goals. Persons with ventromedial prefrontal cortex (VMPC in Figure 4.4) damages, consistently repeat the same mistakes because they cannot take into account negative feedbacks from previous experiences. Since VMPC is the region of brain that stores and signals value of future consequences, possible damages to this area lead to behaviours that rely only on input from other brain's regions (e.g., the amygdala).

According to these findings, neuroeconomics proposes an explanation of drug addiction that confirms behavioral economics' results about how individuals with drug abuse disorders tend to temporally discount rewards more heavily than non-dependents people. The hyperactive impulsive system (i.e., amygdala) might lower the relative influence of the reflective system (i.e., PFC). Thus, drug addiction arises when there is too much emphasis on immediate consequences, namely, when the impulsive system overwhelms the reflective system. Then, the two systems are separated but interacting.

Experiments supports this hypothesis. For instance, McClure *et al.* (2004) used fMRI to scan participants while carrying out temporal discounting tasks. The analyses displayed a greater activity in limbic areas (the red ones according to Figure 4.5) for decisions that encompasses an immediate reward. On the contrary, greater activity was present in the frontal and parietal regions when participants chose the delayed reward.

The topic can also be seen from another point of view. Indeed, as already discussed, individuals who have a decreased activity in the reflective system may take fewer choices in

favour of delayed rewards. Many drugs produce neurological effects that lead to a decrease in regions of the cortex associated with the reflective system. Moreover, evidence has shown how teenagers and children affected by attention deficit hyperactivity disorder (ADHD) or individuals with antisocial personality disorder (APD) also tend to discount delayed awards heavily. Prefrontal cortex's dysfunctions are involved in these and other disorders, particularly, in those involving young people such as ADHD. It is interesting to note how the prefrontal cortex undergo developmental changes well into the 20s. Due to the continue maturation of the prefrontal cortex, teenagers appear to be more vulnerable to abuse of substances such as drugs. These insights may lead to the conclusion that impulsive behaviours are caused by the non-maturity of the reflective system (rather than the impulsive system). Consequently, dysfunctions of the prefrontal cortex are exacerbated in adolescents due to immaturity as well as the brain changes related to the pharmacological effects caused by the consumption of addictive drugs (Bickel *et al.*, 2007).

4.2.4 A Potential Behavioural Market of Addiction

So far, we have discussed how temporal discounting processes are at the base of both behavioral and neuroeconomics explanation of drug addiction. What we may ask to ourselves is whether these economics findings can be useful in preventing drug dependence or in treating addicted patients. Namely, we want to know if temporal discounting process can represent a possible behavioral market of addiction. A behavioral marker is «more than a risk factor or a mere correlate of disease progression if it also reveals facets of the disorder's mechanism, tracks treatment outcomes, and suggests novel avenues for treatment development» (Duka *et al.*, 2011). To be an useful behavioral market, temporal discounting should be able to distinguish whether individuals currently have a drug dependence disorder. Results from experiments reported in sections 4.2.2 and 4.2.3 repeatedly demonstrate how individuals affected by a drug addiction have a comparatively higher average discount rate with respect to those individuals who do not. The purpose of this section is to analyse temporal discounting with respect to the clinical course of addiction, the treatment outcomes and biological components. However, what we need to keep in mind is that these represent a relatively new fields of research, for this reason, further studies will be necessary in future to confirm or reject findings.

Bickel *et al.* (2014) reviews the economic literature about drug addiction in order to support the thesis that temporal discounting can be an appropriate behavioral market. The first topic analysed is the usefulness of temporal discounting processes in predicting the

entrance in drug use. The difficulty in performing this kind of research is very high due to the fact that, to be reliable, the relationship between delay discounting and drug use needs to be measured by assessing delay discounting rates before any drug use and after that individuals eventually used or abused a drug. While researches on animals such as rats are more common, only one study examines this kind of relationship in humans. In their study, Audrain-McGovern *et al.* (2009) found a positive relationship between high discount rates and subsequent initiation or increased use of cigarette and drugs. To 947 people were asked to respond to a questionnaire in three occasions between the high school and till two years after the diploma. Students with higher discount rates were proved to consume a larger number of cigarettes post high school, but no relationship was found with later marijuana use. Similarly, the presence of high discount rates in childhood significantly predicted cocaine and crack use in adulthood.

The second topic analysed are the relationships between discounting processes and the amount of drug use. Several studies (e.g., Vuchinich and Simpson, 1998) have shown how discount rates are correlated with the quantity of drug consumed, making it a possible indicator of addiction severity. Indeed, the number of cigarettes smoked, the amount of cocaine used, the number of years spent in abusing of heroin and the volume of alcohol consumed are all demonstrated to be positively correlated with one's rate of discounting.

Instead, contrasting results have been found between temporal discounting and comorbidities. Discounting has been shown to increase with the number of addictive behaviours (e.g., tobacco use and alcoholism), yet this is not always the case. For example, this relationship is not confirmed between individuals who smoke and have other substance abuse disorders (Businelle *et al.*, 2010). Additional research may be useful to confirm or neglect the relationship between discounting processes and comorbidities.

Withdrawal symptoms have shown to be relevant in affecting discounting rates. Indeed, most drugs produce withdrawal syndromes when the amount consumed is substantially or suddenly reduced after a period of heavy use. Despite the duration and the symptoms vary according to the drug and the dependence severity, during the withdrawal phase there is a drastically increased likelihood to relapse. Also in this case, few studies analyse the relationship between withdrawal symptoms and discounting rate. However, some have found that discount rate is elevated during the withdrawal period such as Giordano *et al.* (2002) who computed opioid-dependent discount rates on six delayed discounting procedures. Participants were tested either in a deprived (when their last dose was 5 days prior) or sated (when receiving 5 times their usual maintenance dose). Further studies on

animals, have allowed to achieve results still unexplored in the human literature assessing the amount of time necessary after drug administration before discounting behavior returns to pre-drug levels. For example, an experiment carried out on pigeons have shown how discounting behaviours return to baseline levels after five weeks of chronic morphine administration ceased (Eppolito *et al.*, 2013).

The capability of discount rates to distinguish between current and ex-drug addicts is less clear. If it is true that heroin and amphetamine users show higher discount rates than ex- and never-users subjects, the same cannot be said for cocaine users. Indeed, individuals who stopped cocaine consumption 30 days before the assessment show the same discount rates of current users (Heil *et al.*, 2006). Possible insights are deductible from these findings. Maybe discount rates may need some time to return to pre-drug use levels. Another possibility is that those users who successfully initiate and maintain abstinence had lower discount rates throughout their time in dependence, and for this reason, discount rates do not decrease after stopping the consume. Unfortunately, there is a lack of longitudinal studies in this field.

Another important employment for delay discounting serving as a behavioral marker is how discount rate could be correlated with recovery success. Indeed, temporal discounting may represent a predictor of later abstinence or future drug consumption among those individuals who were engaged in drug use, at least initially. Most studies found that measures of temporal discounting are related to measures of abstinence, cessation, or relapse. However, the number of researches available is limited and further studies are necessary to determine if any particular measure, treatment, or drug-dependent population is more or less likely to exhibit this relationship.

Temporal discounting can be used as an indicator of the effectiveness of treatments and interventions. Change in the rate of discounting is noteworthy since empirical evidence has shown how measures of temporal discounting tend to remain stable over varying timeframes ranging from 1 week to 1 year. Therefore, in the absence of an intervention and treatment discounting can be assumed to be stable. Yi *et al.* (2008) investigated the effects of five days of a contingency management procedure on the temporal discounting of smokers. Temporal discounting rates for money and cigarettes were computed both before and after implementation of certain intervention. The results, which are showed in Figure 4.6, display how discounting rates are significantly decreased for both cigarettes and money among the smokers under treatment, whereas the control group showed no significant change for either commodity.

Figure 4.6: Discounting rates before and after treatment
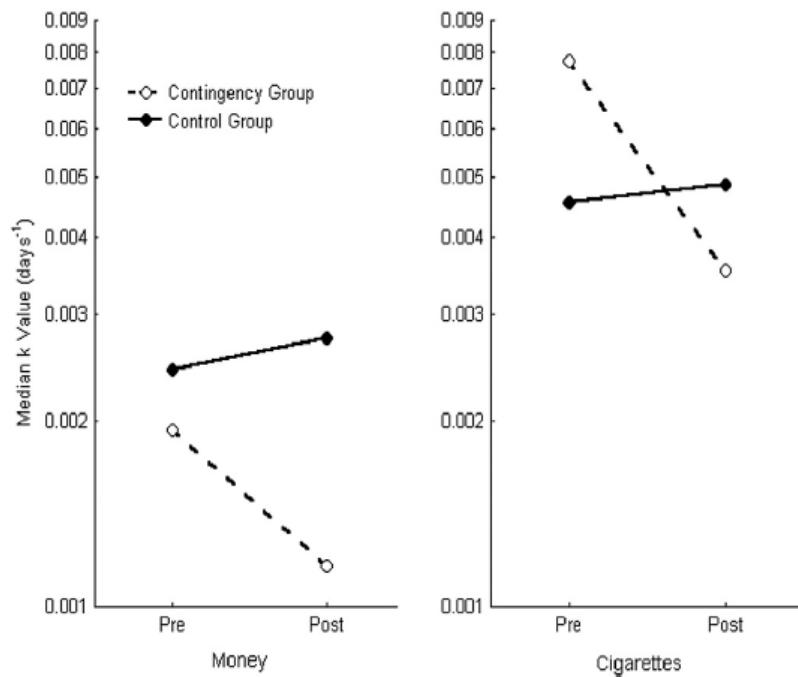
A further study conducted by Black & Rosen (2011) on cocaine users examined the effects of a money-management intervention or control intervention on discounting processes. The treatment was based on a multi-component package addressing substance abuse in the context of prospective money management concerns. Specifically, participants need to plan monthly budgets while taking into account long-term goals reachable through shorter spending plans. A temporal discounting questionnaire was filled four times across the 32 weeks of the experiment. Results describe how the money management intervention was useful since participants undergone to it then discounted future monetary rewards less and had more abstinence than the control participants by the end of the measurement period. Additionally, regardless of group assignment, those participants who discounted future rewards more heavily over time also increased the cocaine use.

The last point regards whether biological components underlie the empirical relationship between discounting and addiction. A behavioral marker should be able to clarify a disorder's mechanism of action. For this reason, if temporal discounting is accepted as a behavioral marker of addiction, extension of the biological underpinnings of temporal discounting to the understanding of addiction might advise innovative neurobehavioral indices of the disorder. These mechanisms may need as diagnostic assessment tools, predictors of illness or outcome, or they may even guide the implementation of individualized therapy. The main idea is that substance-using individuals have inefficient

recruitment of cognitive systems if compared to non-users (Hoffman *et al.*, 2008). However, research in neurobiological and genetic supporting discounting and addiction are not very widespread yet. Future investigations are needed to better identify those brain regions and networks that are differentially active in delayed discounting in individuals with and without drug dependence in order to characterize the relationship between temporal discounting and addiction.

Conclusions

The fundamental results of the thesis are now summed up. In the first chapter, we began with the definition of health economics and, in particular, we saw how health-related goods and services differ from the classical commodities for the nature of the demand, product uncertainty, expected behaviour of physicians, supply conditions and pricing practices (Arrow, 1963). Furthermore, we discussed the role of asymmetric information in health economics and how it might lead to moral hazard and adverse selection phenomena when dealing with health insurances. However, in addition to these market failures, governments have an interest in intervening in health economics due to the impact that some heath variables can have on macroeconomics variables. In particular, life-expectancy improvements have proved to boost growth, yet contradictory evidence is presented. Indeed, according to Bhargava *et al.* (2001), a better life-expectancy has a relevant impact only in poor and developing countries while it provides no or lower benefits in richer countries. Nevertheless, the study carried out by Swift (2010) on 13 rich OECD countries seems to deny results by Bhargava et al. (2001), showing a positive correlation between life-expectancy and GDP also for rich countries. In a similar way, we investigated the negative impact of some health status (i.e., obesity and diabetes) on the employment rate. Given the results from these researches, the big investments of states in health should not surprise. Indeed, United States spent about the 17% of its total GDP in health in 2017, while the European average in the same year is about 8% (OECD, 2018). As we briefly discussed, this big difference in health investments is mainly due to the inefficiencies of the American health care system rather than e.g., bigger investments in R&D.

The focus was then shifted toward the drug addiction issue. We herein refrain from repeating the psychological assumptions at the base of addictions. What we saw is that, despite drug addiction is less widespread with respect to other health status (e.g., smoking, obesity), it leads to relevant socio-economic costs in term of mortality and morbidity (e.g., loss of efficiency) other than for health, crime and interpersonal related costs. Since these costs are paid by the whole society, they can be seen as a negative externality. In this section, empirical evidence denies the economic assumption of perfect inelasticity demand for addictive commodities: increases in price of cigarettes or cocaine can actually lead to a lower consumption (Chaloupka, 2018). However, a bigger inelasticity of these kind of goods with respect to other commodities is acknowledged. Further reasons for the state intervention in limiting the drug abuse have been identified in paternalistic motivations and in the illegal markets. About the last point, we argue how governments have difficulties in

developing effective policies due to the drug market's features. Finally, some insights about drugs' regulation have been presented with respect to drug prohibition laws and the controversial topic of cannabis legalization.

In the second chapter, we quickly review the evolution of the economic thought in the last centuries. A summarized version of the neo-classical utility theory of the customer under perfect rationality is reported. We asserted how, when expected utility and discounted utility models started to show their limits and became subject to critiques, economists began to acknowledge the possible involvement of psychology within economics. Simons (1955) was one of the first to hypothesize bounded rationality assumptions within the economics model, despite already Adam Smith in the XVIII century provided some insights about it. However, the real development of the field of study that today is known as Behavioral Economics was possible only starting from the 1970s thanks to cognitive psychology progresses. The Planned behavior theory (Ajzen, 1991) and the Consumer decision model are explained as examples of respectively prescriptive and analytical cognitive models. Furthermore, we saw how the concepts of bounded rationality, loss aversion, preference for immediacy and fairness and social preferences introduced in economics through the integration of psychological variables have played a fundamental role in modifying the neo-classical theories of expected utility, equilibrium, discounted utility and own-payoff maximization making them more realistic (Camerer, 1999). The implementation of behavioral economics in health economics has been discussed. Even if behavioral economics implications on health economics have not been formally acknowledged yet (Frank, 2004), the tools it supplies have seen to be useful in many health-related issues namely organs donation, health (and drug) insurance choices, smoking and obesity. Finally, an introduction to neuroeconomics and some general concepts about brain functioning with respect to long-term decisions are provided.

The third chapter begins with the explanation of the Health Belief and Transtheoretical Models. The first is prescriptive and provides some relevant variables where a policy-maker should intervene in order to lead individuals toward the one that it considers to be the best choice. On the contrary, the second is analytical and try to study the stages of changes an individual goes through while he/she tries to modify e.g. a dangerous habit also identifying possible weak or sensitive spots. Descriptions and evidence for fear appeal, framing effect, status quo bias and nudging are provided. Without further remarking the individual results for all of them, we can assert that they generally proved to be effective in driving people's behaviours and choices. However, there are possible concerns regard nudging that, if not

correctly implemented, it might lead to adverse effects. Moreover, since nudging stems from paternalism, some scholars largely criticized it because manipulative (Goodwin, 2012).

Finally, in chapter four we dealt with the drug addiction issue. The Model of Rational Addiction developed by Becker & Murphy (1988) under perfect rationality assumptions has been described. Even if it has enabled to lead to important economic results, this model (and generally, all neo-classical models) has been criticized because it makes use of exponential discounting which has never been empirically demonstrated in real-life situations. Furthermore, exponential discounting does not take into account the possible reversal in preferences that involves long-term decision-making (even if preference reversal is possible through exponential discounting applying a different discount rate for immediate and more delayed rewards). What behavioral economics wants to do is trying to overcome these limits, first acknowledging the role of impulsivity and loss of control, two behaviours very common among drug abusers, and secondly applying hyperbolic discounting rather than exponential. Through the use of hyperbolic discounting, behavioral economics is able to provide a more realistic economic explanation of why people abuse of drug (and generally, of all the other substances), also explaining preference reversal which is seen as a consequence of both impulsivity and loss of control. Particularly, the latter is seen as the main driver since an individual may actually want to stop using drug. However, as the time goes by, he/she is subject to temptations and the perspective of an more immediate consume of drug might be seen as more appealing than waiting for the end of abstinence symptoms and get this way other social but delayed benefits (e.g., relationships with family and friends, employment, sobriety).

The chapter continues providing empirical evidence about hyperbolic discounting. What most studies are agree with is that drug abusers, and more generally all individuals affected by a dependence (e.g., smoking, alcoholism) tend to discount delayed money and other rewards more heavily than individuals without a dependence, showing a certain impulsivity and preference for smaller, but closer in time, rewards. Moreover, suggesting that addiction is the result of the interaction between two neural systems and through empirical evidence, neuroeconomics seems to confirm results from behavioral economics' researches.

Possible future directions in the study of discounting processes are proposed by Bickel & Marsh (2001). One of this regards the possible implication of loss aversion elements in discounting processes. Particularly, the so-called "Sign Effect" which refers to the fact that rewards are discounted at a higher discount rate than comparably valued losses. This means that, for example, the value of a \$10 reward delivered in 1 week is heavier discounted than

the value of a $10 loss delayed of the same period of time. However, some evidence (Petry *et al.*, 1998) suggests that alcoholics and opioid dependents are insensitive or react less to small immediate losses. If this would be further confirmed, some social and behavioral problems stemming from forms of drug dependence may be better understood (e.g. family, occupational, legal and health issues). These problems may result from those individuals who prefer not to solve problems immediately, but they wait till they become much larger at a later time.

Another potential direction in research can be the relationship between elasticity of demand and temporal discounting. Since drug consumption has shown to be less sensitive to the price than other commodities, possible questions about a possible linkage between inelasticity of demand and extreme discounting may arise. If a relationship between these two behavioral economic factors should be demonstrated, then more elaborated models on drug addiction might be able to meaningfully distinguish recreational drug users from the drug dependents basing on the individual's elasticity of demand for the drug and discounting of the drug.

Finally, the last section of the chapter dealt with the possibility of delay discounting processes to be a behavioral market for addiction. Indeed, temporal discounting can be used to predict the entrance in drug use, assess the quantity of drug consumed by individuals, analyse withdrawal symptoms, indicate the effectiveness of treatments and more. Despite the necessity of further studies has been largely stressed throughout the section 4.2.4, if additional studies were found to in favour of temporal discounting, then it could represent an useful tool that might be applied to examine the efficacy of novel treatments and could be the basis for personalizing prevention or treatment's procedures (Bickel *et al.*, 2014).

In conclusion, integrating elements from psychology and scientific disciplines, the relatively new fields of behavioral and neuroeconomics working together have the potential to lead toward new empirical findings about addictions. The application of these approaches to addiction are promising, and the more scientific and sceptical approaches adopted by neuroeconomics have the potential to confirm or reject theories assumed by behavioral economics. Thanks to these new two approaches, economics can now provide an additional contribute in fighting drug abuse, other than helping in the development of new prevention and treatment programs that may drive to fully solution of this serious public health issue.

Bibliography

Ainslie, G.; Haslam, N. (1992), "Irrationality, impulsiveness, and selfishness as discount reversals", Choice Over Time, Russell Sage Foundation, Pages 57–92.

Ajzen, I. (1991), "The Theory of Planned Behavior", Organizational Behavior and Human Decision Processes, Volume 50, Issue 2, Pages 179-211, DOI: https://doi.org/10.1016/0749-5978(91)90020-T.

Akerlof, G. A. (1970), "The Market for "Lemons": Quality Uncertainty and the Market Mechanism" The Quarterly Journal of Economics, Volume 84, No. 3, Pages 488-500, URL:http://links.jstor.org/sici?sici=00335533%28197008%2984%3A3%3C488%3ATM F%22QU%3E2.0.CO%3B2-6.

Arno, A.; Thomas, S. (2016), "The efficacy of nudge theory strategies in influencing adult dietary behaviour: a systematic review and meta-analysis", BMC Public Health, Volume 16, Pages 676-694, DOI: https://doi.org/10.1186/s12889-016-3272-x.

Arrow, K. J. (1963), "Uncertainty and the Welfare Economics of Medical Care", The American Economic Review, Volume 53, Issue 5, Pages 941-973.

Audrain-McGovern, J.; Rodriguez, D.; Epstein, L. H.; Cuevas, J.; Rodgers, K.; Wileyto, E.P. (2009), "Does delay discounting play an etiological role in smoking or is it a consequence of smoking?", Drug and Alcohol Dependence, Volume 103, Issue 3, Pages 99-106, DOI: https://doi.org/10.1016/j.drugalcdep.2008.12.019.

Baiardi, D.; Profeta, P.; Puglisi, R.; Scabrosetti, S. (2018), "Tax policy and economic growth: does it really matter?", International Tax and Public Finance, Volume 26, Issue 2, Pages 282-316, DOI: https://doi.org/10.1007/s10797-018-9494-3.

Bhargava, A.; Jamison, D. T.; Lau, L. J.; Murray, C., J., L. (2001), "Modeling the Effects of Health on Economic Growth", Journal of Health Economics (20), Pages 423–440, DOI: https://doi.org/10.1016/S0167-6296(01)00073-X.

Bhattacharya, J.; Hyde, T.; Tu, P. (2014), "Health Economics", Palgrave Macmillian, Mannheim Institute of Public Health, Heidelberg University, Germany.

Bechara, A. (2005), "Decision making, impulse control and loss of willpower to resist drugs: a neurocognitive perspective", Nature Neuroscience, Volume 8, Number 11, Pages 1458–1463, DOI: https://doi.org/10.1038/nn1584.

Becker, G. S.; Murphy, K. M. (1988), "A Theory of Rational Addiction", The Journal of Political Economy, Volume 96, Number 4, Pages 675-700, DOI: https://doi.org/10.1086/261558.

Beckert, J.; Wehinger, F. (2012), "In the shadow: illegal markets and economic sociology", Socio-Economic Review, Volume 11, Issue 1, Pages 5–30, DOI: https://doi.org/10.1093/ser/mws020.

Bernheim, B. D. & Whinston, M. (2013), *Microeconomics*, McGraw-Hill Education.

Bickel, W. K.; Madden, G. J. (1999), "A comparison of measures of relative reinforcing efficacy and behavioral economics: Cigarettes and money in smokers", Behavioural Pharmacology, Volume 10, Issue 6-7, Pages 627-637, DOI: http://dx.doi.org/10.1097/00008877-199908001-00023.

Bickel, W. K.; Marsh, L. A. (2001), "Toward a behavioral economic understanding of drug dependence: delay discounting processes", Addiction, Volume 96, Issue 1, Pages 73-86, DOI: https://doi.org/10.1046/j.1360-0443.2001.961736.x.

Bickel, W. K.; Miller, M. L.; Yi, R.; Kowal, B. P.; Lindquist, D. M.; Pitcock, J. A. (2007), "Behavioral and neuroeconomics of drug addiction: Competing neural systems and temporal discounting processes", Drug and Alcohol Dependence, Volume 90, Issue 1, Pages 85-91, DOI: https://doi.org/10.1016/j.drugalcdep.2006.09.016.

Bickel, W. K.; Koffarnus, M. N.; Moody, L.; Wilson, A. G. (2014), "The behavioral- and neuro-economic process of temporal discounting: A candidate behavioral marker of addiction", Neuropharmacology, Volume 76, Part B, Pages 518-527, DOI: https://doi.org/10.1016/j.neuropharm.2013.06.013.

Black, A. C.; Rosen, M. I. (2011), "A money management-based substance use treatment increases valuation of future rewards", Addictive Behaviors, Volume 36, Issues 1-2, Pages 125-128, DOI: https://doi.org/10.1016/j.addbeh.2010.08.014.

Boadway, R.; Leite-Monteiro, M.; Marchand, M.; Pestieau, P. (2006), "Social Insurance and Redistribution with Moral Hazard and Adverse Selection", The Scandinavian Journal of Economics, 108(2), Pages 279–298, DOI: 10.1111/j.1467-9442.2006.00446.x.

Bouchard, M. (2007), "On the Resilience of Illegal Drug Markets", Global Crime, Volume 8, Issue 4, Pages 325-344, DOI: 10.1080/17440570701739702.

Bray, J. (2008), "Consumer Behaviour Theory: Approaches and Models", Discussion Paper. Bournemouth University. Bournemouth.

Businelle, M. S.; McVay, M. A.; Kendzor, D.; Copeland, A. (2010), "A comparison of delay discounting among smokers, substance abusers, and non-dependent controls", Drug and Alcohol Dependence, Volume 112, Issue 3, Pages 247-250, DOI: https://doi.org/10.1016/j.drugalcdep.2010.06.010.

Camerer, C. F.; Thaler, R. H. (1995), "Anomalies: Ultimatums, Dictators and Manners", Journal of Economic Perspectives, Volume 9, no. 2, Pages 209-219, DOI: 10.1257/jep.9.2.209.

Camerer, C. F. (1999), "Behavioral economics: Reunifying psychology and economics", PNAS, Volume 19, Pages 10575-10577, DOI: https://doi.org/10.1073/pnas.96.19.10575

Camerer, C. F.; Issacharoff, S.; Loewenstein, G., O'Donoghue, T., Rabin, M. (2003), "Regulation for Conservatives: Behavioral Economics and the Case for "Asymmetric Paternalism"", University of Pennsylvania Law Review, Volume 151, No. 3, Pages 1211-1254, DOI: 10.2307/3312889.

Camerer, C. F.; Loewenstein, G. (2003), "Behavioral Economics: Past, Present, Future", Advances in behavioral economics, Pages 1-61, URL: http://resolver.caltech.edu/CaltechAUTHORS:20110204-152338626.

Camerer, C. F.; Loewenstein, G.; Prelec, D. (2005), "Neuroeconomics: How Neuroscience Can Inform Economics", Journal of Economic Literature, Volume 43, no. 1, Pages 9-64, DOI: 10.1257/0022051053737843.

Camì, J.; Farré, M. (2003), "Drug Addiction", The New England Journal of Medicine, Pages 975-986, DOI: 10.1056/NEJMra023160.

Carande-Kulis, V. G.; Getzen, T. E.; Thacker, S. B. (2007), "Public Goods and Externalities: A Research Agenda for Public Health Economics", Journal of Public Health Management and Practice, Volume 13, Issues 2, Pages: 227-232, https://journals.lww.com/jphmp/Fulltext/2007/03000/Public_Goods_and_Externalities__A_Research_Agenda.24.aspx.

Cartwright, W. S. (2008), "Economic costs of drug abuse: Financial, cost of illness, and services", Journal of Substance Abuse Treatment, Volume 34, Issue 2, Pages 224-233, DOI: https://doi.org/10.1016/j.jsat.2007.04.003.

Chaloupka, F. J. (2018), "The Economics of Addiction", Annual Meetings of the Pakistan Society of Development Economists, https://tobacconomics.org/wp-content/uploads/2018/12/Chaloupka_Addiction_13-Dec-2018.pdf.

Chernev, A. (2010), "The Dieter's Paradox", Journal of Consumer Psychology, Volume 21, Issue 2, Pages 178-183, DOI: https://doi.org/10.1016/j.jcps.2010.08.002.

Cody, W. K. (2003), "Paternalism in Nursing and Healthcare: Central Issues and Their Relation to Theory", Nursing Science Quarterly, Volume 16, No. 4, Pages 288-296, DOI: 10.1177/0894318403257170.

Colander, D. (2000), "The Death of Neoclassical Economics" Journal of the History of Economic Thought, Volume 22, Issue 2, Pages 127-143, DOI: https://doi.org/10.1080/10427710050025330.

Duka, T.; Crombag, H. S.; Stephens, D. N. (2011), "Experimental medicine in drug addiction: towards behavioral, cognitive and neurobiological biomarkers", Journal of Psychopharmacology, Volume 25, Issues 9, Pages 1235-1255, DOI: https://doi.org/10.1177/0269881110388324.

Eppolito, A. K.; France, C. P.; Gerak, L. R. (2013), "Effects of acute and chronic morphine on delay discounting in pigeons", Journal of Experimental Analysis of Behavior, Volume 99, Issue 3, Pages 277-289, DOI: https://doi.org/10.1002/jeab.25.

European Commission (2014), "EU Action Plan on Childhood Obesity 2014-2020", https://ec.europa.eu/health/sites/health/files/nutrition_physical_activity/docs/childhoodobesity_actionplan_2014_2020_en.pdf.

Festinger, L. (1957), "A Theory of Cognitive Dissonance", Stanford University Press.

Fleming, M. H.; Roman, J.; Farrell, G. (2000), "The Shadow Economy", Journal of International Affairs, Volume 53, No. 2, Pages 387-409, DOI: https://www.jstor.org/stable/24357758.

Frank, R. G. (2004), "Behavioral Economics and Health Economics", The National Bureau of Economic Research, Working Paper No. 10881, DOI: 10.3386/w10881.

Gamage, D.; Shanske, D. (2011), "Three Essays on Tax Salience: Market Salience and Political Salience", Articles by Maurer Faculty, 2416, Pages 19-98, URL: http://www.repository.law.indiana.edu/facpub/2416.

Garcia-Retamero, R.; Galesic, M. (2010), "How to Reduce the Effect of Framing on Messages About Health", Journal of General Internal Medicine, Volume 25, Issue 12, pp 1323–1329, DOI: https://doi.org/10.1007/s11606-010-1484-9.

Giordano, L.; Bickel, W. K.; Loewenstein, G.; Jacobs, E. A.; Marsch, L.; Badger, G. J. (2002), "Mild opioid deprivation increases the degree that opioid-dependent outpatients discount delayed heroin and money", Psychopharmacology, Volume 163, Issue 2, Pages 174-182, DOI: https://doi.org/10.1007/s00213-002-1159-2.

Goodman, A. (1990), "Addiction: definition and implications", British Journal of Addiction, Volume 85, Pages 1403-1408, DOI: https://doi.org/10.1111/j.1360-0443.1990.tb01620.x.

Goodwin, T. (2012), "Why We Should Reject 'Nudge'", Politics, Volume 32, Issue 2, Pages 85-92, DOI: https://doi.org/10.1111/j.1467-9256.2012.01430.x.

Granz, K.; Rimer, B. K; Viswanath, K. (2008), "Health Behavior and Health Education: Theory, Research and Practice", 4th Edition, Jossey-Bass.

Grecucci, A.; Giorgetta, C.; Brambilla, P.; Zuanon, S.; Perini, L.; Balestrieri, M.; Bonini, N.; Sanfey, A.G. (2013), "Anxious ultimatums: How anxiety disorders affect socioeconomic behaviour", Cognition and Emotion, Volume 27, Issue 2, Pages 230-244, DOI: https://doi.org/10.1080/02699931.2012.698982.

Green, L.; Fry, A. F.; Myerson, J. (1994), "Discounting of Delayed Rewards: A Life-Span Comparison", Psychological Science, Volume 5, Issue 1, Pages 33-36, DOI: https://doi.org/10.1111/j.1467-9280.1994.tb00610.x.

Gruber, J. (2013), "Public Finance and Public Policy: Fourth Edition", Worth Publishers.

Iglehart, J. K. (1992), "The American Health Care System — Managed Care", Health Policy Report, DOI: 10.1056/NEJM199209033271029.

Hall, M. A. (2001), "Arrow on Trust", Journal of Health Politics, Policy and Law, Volume 26, Number 5, Pages 1131-1144, DOI: https://muse.jhu.edu.

Hall, W.; Lynskey, M. (2016), "Evaluating the public health impacts of legalizing recreational cannabis use in the United States", Addiction, Volume 111, Issue 10, Pages 1764-1774, DOI: https://doi.org/10.1111/add.13428.

Hasin D. S.; Wall M.; Keyes K. M.; Cerdá M.; Schulenberg J.; O'Malley P. M.; *et al.* (2015), "Medical marijuana laws and adolescent marijuana use in the USA from 1991 to 2014: results from annual, repeated cross-sectional surveys", Lancet Psychiatry, Volume 2, Issue 7, Pages 604-608, DOI: https://doi.org/10.1016/S2215-0366(15)00217-5.

Heil, S. H.; Johnson, M. W.; Higgins, S. T.; Bickel, W. K. (2006), "Delay discounting in currently using and currently abstinent cocaine-dependent outpatients and non-drug-using matched controls", Addictive Behaviors, Volume 31, Issue 7, Pages 1290-1294, DOI: https://doi.org/10.1016/j.addbeh.2005.09.005.

Hoffman, W. F.; Schwartz, D. L.; Huckans, M. S.; McFarland, B. H.; Meiri, G.; Stevens, A. A.; Mitchell, S. H. (2008), "Cortical activation during delay discounting in abstinent methamphetamine dependent individuals", Psychopharmacology, Volume 201, Pages 183-193, DOI: https://doi.org/10.1007/s00213-008-1261-1.

IHME (2017), *Global Burden of Diseases*. Available at: http://www.healthdata.org/gbd.

Illiashenko, P. (2017), "Misbehaving: The Making of Behavioral Economics", International Journal of Applied Behavioral Economics, Volume 6, Issue 1.

Jacobs, E. A.; Bickel, W. K. (1999), "Modeling drug consumption in the clinic via simulation procedures: demand for heroin and cigarettes in opioid-dependent

outpatients", Experimental and Clinical Psychopharmacology, Volume 7, Issue 4, Pages 412-426, DOI: http://dx.doi.org/10.1037/1064-1297.7.4.412.

Janz, N. K.; Becker, M. H. (1984), "The Health Belief Model: A Decade Later" Health Education Quarterly, Volume 11, Issue 1, Pages 1-47, DOI: https://doi.org/10.1177/109019818401100101.

Johnson, E. J.; Goldstein, D. G. (2004), "Defaults and Donation Decisions", Transplantation, Volume 78, Issue 12, Pages 1713-1716, DOI: 10.1097/01.TP.0000149788.10382.B2.

Johnson, M. W.; Bickel, W. K.; Baker, F.; Moore, B. A.; Badger, G. J.; Budney, A. J. (2010), "Delay discounting in current and former marijuana-dependent individuals", Experimental and Clinical Psychopharmacology, Volume 18, Issue 1, Pages 99-107, DOI: http://dx.doi.org/10.1037/a0018333.

Kahneman, D.; Tversky, A. (1979), "Prospect Theory: An Analysis of Decisions under Risk", Econometrica, Volume 47, Pages 263-291.

Kahneman, D.; Tversky, A. (1982), "The Psychology of Preference," Scientific American, Volume 246, Pages 160-173.

Kahneman, D.; Tversky, A. (1991), "Loss Aversion in Riskless Choice: A Reference-Dependent Model", The Quarterly Journal of Economics, Volume 106, Issue 4, Pages 1039–1061, DOI: https://doi.org/10.2307/2937956.

Kirby, K. N. (1997), "Bidding on the future: Evidence against normative discounting of delayed rewards", Journal of Experimental Psychology: General, Volume 126, Issue 1, Pages 54-70, DOI: http://dx.doi.org/10.1037/0096-3445.126.1.54.

Lanzini, P. (2018), "Responsible Citizens and Sustainable Consumer Behavior", Routledge.

Laibson, D. (1997), "Golden Eggs and Hyperbolic Discounting", The Quarterly Journal of Economics, Volume 112, Issue 2, Pages 443–478, DOI: https://doi.org/10.1162/003355397555253.

Levine, H. G. (2003), "Global drug prohibition: its uses and crises", International Journal of Drug Policy, Volume 14, Issue 2, Pages 145-153, DOI: https://doi.org/10.1016/S0955-3959(03)00003-3.

Loewenstein, G. F.; Weber, E. U.; Hsee, C. K.; Welch, N. (2001), "Risk as feelings", Psychological Bulletin, Volume 127, No 2, Pages 267-286, DOI: 10.1037//0033-2909.127.2.267.

Lowenstein, G. F.; Rick, S.; Cohen, J. D. (2008), "Neuroeconomics", Annual Review of Psychology, Volume 59, Pages 647-672, DOI: https://doi.org/10.1146/annurev.psych.59.103006.093710.

MacCoun, R.; Reuter, P. (1997), "Interpreting Dutch Cannabis Policy: Reasoning by Analogy in the Legalization Debate", Science, Volume 278, Issue 5335, Pages 47-52, DOI: 10.1126/science.278.5335.47.

Madden, G. J.; Petry, N. M.; Badger, G. J.; Bickel, W. K. (1997), "Impulsive and self-control choices in opioid-dependent patients and non-drug-using control patients: Drug and monetary rewards", Experimental and Clinical Psychopharmacology, Volume 5, Issue 3, Pages 256-262, DOI: http://dx.doi.org/10.1037/1064-1297.5.3.256.

Madden, G. J.; Bickel, W. K.; Jacobs, E. A. (1999), "Discounting of Delayed Rewards in Opioid-Dependent Outpatients: Exponential or Hyperbolic Discounting Functions?", Experimental and Clinical Psychopharmacology, Volume 7, Issues 3, Pages 284-293, DOI: http://dx.doi.org/10.1037/1064-1297.7.3.284.

Malenka, D. J.; Baron, J. A.; Johansen, S.; Wahrenbergen, J. W.; Ross, J. M. (1993), "The Framing Effect of Relative and Absolute Risk", Journal of General Internal Medicine, Volume 8, Issue 10, pages 543–548, DOI: https://doi.org/10.1007/BF02599636.

Marteau, T. M.; Ogilvie, D.; Roland, M.; Suhrcke, M.; Kelly, M. P. (2011), "Judging nudging: can nudging improve population health?", BMJ, Volume 342, Pages 263-265, DOI: https://doi.org/10.1136/bmj.d228.

McClure, S. M.; York, M. K.; Montague, P. R. (2004), "The Neural Substrates of Reward Processing in Humans: The Modern Role of fMRI", The Neuroscientist, Volume 10, Issue 3, Pages 260-268, DOI: https://doi.org/10.1177/1073858404263526.

McGuire, T.G. (2000), "The Economics of Physician Behavior", Handbook of Health Economics Amsterdam: North Holland.

Medina, L.; Schneider, F. (2018), "Shadow Economies Around the World: What Did We Learn Over the Last 20 Years?", IMF Working Paper No. 18/17, Available at: https://ssrn.com/abstract=3124402.

Meyerowitz, B. E.; Chaiken, S. (1987), "The effect of message framing on breast self-examination attitudes, intentions, and behavior", Journal of Personality and Social Psychology, Volume 52, Issue 3, Pages 500-510, DOI: http://dx.doi.org/10.1037/0022-3514.52.3.500.

Morris, S. (2006), "The impact of obesity on employment", Labour Economics (14), Pages 413–433, DOI: https://doi.org/10.1016/j.labeco.2006.02.008.

Mushkin, S. J. (1958), "Toward a Definition of Health Economics", Public Health Report, 73(9), Pages 785-794, PMC:1951624.

OECD (2018), *Health Statistics. Health Expenditure and Financing*. Available at: http://www.oecd.org/els/health-systems/health-data.htm.

Paz, V.; Nicolaisen-Sobesky, E.; Collado, E.; Horta, S.; Rey, C.; Rivero, M.; Berriolo, P.; Díaz, M.; Otón, M.; Pérez, A.; Fernández-Theoduloz, G.; Cabana, A.; Gradin, V. B. (2017), "Effect of self-esteem on social interactions during the Ultimatum Game", Psychiatry Research, Volume 252, Pages 247-255, DOI: http://dx.doi.org/10.1016/j.psychres.2016.12.063.

Petry, N. M; Bickel, W. K. (1998), "Polydrug abuse in heroin addicts: a behavioral economic analysis", Addictions, Volume 93, Issue 3, Pages 321-335, DOI: https://doi.org/10.1046/j.1360-0443.1998.9333212.x.

Petry, N. M.; Casarella, T. (1999), "Excessive discounting of delayed rewards in substance abusers with gambling problems", Drug and Alcohol Dependence, Volume 56, Issue 1, Pages 25-32, DOI: https://doi.org/10.1016/S0376-8716(99)00010-1.

Prochaska, J. O.; Velicer, W. F. (1997), "The Transtheoretical Model of Health Behavior Change", American Journal of Health Promotion, Volume 12, Issue 1, Pages 38-48, DOI: https://doi.org/10.4278/0890-1171-12.1.38.

Rice, D. P.; Kelman, S.; Miller, L. S. (1991), "Economic Costs of Drug Abuse", Economic Costs, Cost-Effectiveness, Financing, and Community-Based Drug Treatment, Pages 10-32.

Rice, T. (2013), "The Behavioral Economics of Health and Health Care", Annual Review of Public Health, Volume 34, Pages 431-447, DOI: https://doi.org/10.1146/annurev-publhealth-031912-114353.

Richards, J. B.; Zhang, L.; Mitchell, S. H.; De Wit, H. (1999), "Delay or probability discounting in a model of impulsive behavior: effect of alcohol", Journal of Experimental Analysis of Behavior, Volume 71, Issue 2, Pages 121-143, DOI: https://doi.org/10.1901/jeab.1999.71-121.

Rogers, R. W. (1983), "Cognitive and psychological processes in fear appeals and attitude change: A revised theory of protection motivation", Social psychophysiology: A sourcebook, Pages 153-176.

Rothman, A. J.; Salovey, P.; Antone, C.; Keough, K.; Martin, C. D. (1992), "The Influence of Message Framing on Intentions to Perform Health Behaviors", Journal of

Experimental Social Psychology, Volume 29, Issue 5, Pages 408-433, DOI: https://doi.org/10.1006/jesp.1993.1019.

Samuelson, W.; Zeckhauser, R. (1988), "Status quo bias in decision making", Journal of Risk and Uncertainty, Volume 1, Issue 1, Pages 7-59, DOI: https://doi.org/10.1007/BF00055564.

Schneider, W.; Shiffrin, R. M. (1977), "Controlled and automatic human information processing: II. Perceptual learning, automatic attending and a general theory", Psychological Review, Volume 84, Pages 127-190, DOI: http://dx.doi.org/10.1037/0033-295X.84.2.127.

Sent, E. M. (2004), "Behavioral Economics: How Psychology Made Its (Limited) Way Back Into Economics", History of Political Economy, Volume 36, issue 4, Pages 735-760, URL: https://EconPapers.repec.org/RePEc:hop:hopeec:v:36:y:2004:i:4:p:735-760.

Shiv, B.; Fedorikhin, A. (1999), "Heart and Mind in Conflict: the Interplay of Affect and Cognition in Consumer Decision Making", Journal of Consumer Research, Volume 26, No. 3, Pages 278-292, DOI: 10.1086/209563.

Simons, H. A. (1955), "A Behavioral Model of Rational Choice", The Quarterly Journal of Economics, Volume 69, Issue 1, Pages 99–118, DOI: https://doi.org/10.2307/1884852.

Swift, R. (2010), "The relationship between health and GDP in OECD countries in the very long run", Health Economics, Volume 20, Issue 3, Pages 306-322, DOI: 10.1002/hec.1590.

Thaler, R. H.; Sunstein, C. R. (2003), "Libertarian Paternalism", American Economic Review, Volume 93, no. 2, Pages 175-179, DOI: 10.1257/000282803321947001.

Thaler, R. H.; Sunstein, C. R. (2008), "Nudge: improving decisions about health, wealth, and happiness", Yale University Press.

Tuncelli, K.; Bradley, C. J.; Nerenz, D.; Williams, L., K.; Pladevall, M.; Lafata, J., E. (2004), "The Impact of Diabetes on Employment and Work Productivity", Diabetes Care, Volume 28 (11), Pages 2662-2667, DOI: https://doi.org/10.2337/diacare.28.11.2662.

Vitaro, F.; Ferland, F.; Jacques, C.; Ladouceur, R. (1998), "Gambling, substance use, and impulsivity during adolescence", Volume 12, Issue 3, Pages 185-194, DOI: http://dx.doi.org/10.1037/0893-164X.12.3.185.

Vuchinich, R. E.; Simpson, C. A. (1998), "Hyperbolic temporal discounting in social drinkers and problem drinkers", Experimental and Clinical Psychopharmacology, Volume 6, Issue 3, Pages 292-305, DOI: http://dx.doi.org/10.1037/1064-1297.6.3.292.

Winston, G. C. (1980), "Addiction and backsliding: A theory of compulsive consumption", Journal of Economic Behavior & Organization, Volume 1, Issue 4, Pages 295-324, DOI: https://doi.org/10.1016/0167-2681(80)90009-8.

Witte, K.; Allen, M. (2000), "A Meta-Analysis of Fear Appeals: Implications for Effective Public Health Campaigns", Health Education & Behavior, Volume 27, Issue 5, Pages 591-615, DOI: 10.1177/109019810002700506.

World Health Organization (2018), *Global Health Observatory (GHO) data*, available at: https://www.who.int/gho/en/.

Yi, R.; Johnson, M. W.; Giordano, L. A.; Landes, R. D.; Badger, G. J.; Bickel, W. K. (2008), "The effects of reduced cigarette smoking on discounting future rewards: an initial evaluation", The Psychological Record, Volume 58, Issue 2, Pages 163-174, DOI: https://doi.org/10.1007/BF03395609.