ESCOLA NACIONAL DE SAÚDE PÚBLICA
UNIVERSIDADE NOVA DE LISBOA

Economic Analysis in Health Care Regulation

Three essays

Advisors:

Professor João Pereira and Professor Cesaltina Pires

Filipa de Aragão Figueiredo Barbosa Pombeiro

Dissertation Submitted for the Degree of Doctor of Philosophy in Public Health

Health Economics Specialty

Lisbon, 2011

Support for this research has been provided by Fundação para a Ciência e Tecnologia, Programa POCTI - Formar e Qualificar - Medida 1.1., grant Praxis XXI/BD/19954/99

ABSTRACT

It is the purpose of the present thesis to emphasize, through a series of examples, the need and value of appropriate pre-analysis of the impact of health care regulation. Specifically, the thesis presents three papers on the theme of regulation in different aspects of health care provision and financing. The first two consist of economic analyses of the impact of health care regulation and the third comprises the creation of an instrument for supporting economic analysis of health care regulation, namely in the field of evaluation of health care programs.

The first paper develops a model of health plan competition and pricing in order to understand the dynamics of health plan entry and exit in the presence of switching costs and alternative health premium payment systems. We build an explicit model of death spirals, in which profit-maximizing competing health plans find it optimal to adopt a pattern of increasing relative prices culminating in health plan exit. We find the steady-state numerical solution for the price sequence and the plan's optimal length of life through simulation and do some comparative statics. This allows us to show that using risk adjusted premiums and imposing price floors are effective at reducing death spirals and switching costs, while having employees pay a fixed share of the premium enhances death spirals and increases switching costs.
Price regulation of pharmaceuticals is one of the cost control measures adopted by the Portuguese government, as in many European countries. When such regulation decreases the products' real price over time, it may create an incentive for product turnover.

Using panel data for the period of 1997 through 2003 on drug packages sold in Portuguese pharmacies, the second paper addresses the question of whether price control policies create an incentive for product withdrawal. Our work builds the product survival literature by accounting for unobservable product characteristics and heterogeneity among consumers when constructing quality, price control and competition indexes. These indexes are then used as covariates in a Cox proportional hazard model. We find that, indeed, price control measures increase the probability of exit, and that such effect is not verified in OTC market where no such price regulation measures exist. We also find quality to have a significant positive impact on product survival.

In the third paper, we develop a microsimulation discrete events model (MSDEM) for cost-effectiveness analysis of Human Immunodeficiency Virus treatment, simulating individual paths from antiretroviral therapy (ART) initiation to death. Four driving forces determine the course of events: CD4+ cell count, viral load resistance and adherence. A novel feature of the model

with respect to the previous MSDEMs is that distributions of time to event depend on individuals' characteristics and past history. Time to event was modeled using parametric survival analysis. Events modeled include: viral suppression, regimen switch due virological failure, regimen switch due to other reasons, resistance development, hospitalization, AIDS events, and death. Disease progression is structured according to therapy lines and the model is parameterized with cohort Portuguese observational data.

An application of the model is presented comparing the cost-effectiveness ART initiation with two nucleoside analogue reverse transcriptase inhibitors (NRTI) plus one non-nucleoside reverse transcriptase inhibitor(NNRTI) to two NRTI plus boosted protease inhibitor (PI/r) in HIV-1 infected individuals. We find 2NRTI+NNRTI to be a dominant strategy. Results predicted by the model reproduce those of the data used for parameterization and are in line with those published in the literature.

Acknowledgments

Dedicated to my mother

I would like to begin by thanking my dissertation supervisors: Professor João Pereira and Professor Cesaltina Pires.

I deeply thank Professor João Pereira for believing in me, for understanding and supporting me when I needed; and I thank him just as much for being demanding and intransigent when required. Most of all, I thank him for his wise guidance in the research path that led to the present thesis.

I deeply thank Professor Cesaltina for teaching me the fun of doing research, and for embedding in me the importance of performing it rigorously. Mostly, I thank her for pushing me forward at the right moments.

In this journey, Professor Pedro Pita Barros and Professor Randall Ellis also had a decisive role. I will always be thankful to Pedro Pita Barros for introducing me to health economics and, most of all, for teaching me to look at problems through the eyes of an economist. I thank Professor Randall Ellis for taking the first steps with me, for guiding me when I most needed, and for watching me at a distance when I started to walk on my own.

I am deeply indebted to Jorge Félix for both his teachings and the working hours he dispensed and I thank Tânia Furtado for believing in the work done and supporting its continuity and improvement.

I would also like to thank Steven Dunn for bugging me enough to make me complete my work.

I thank two great friends in their area of expertise: my sister Maria for all the help in *"landscaping"* the final document and Samir Noorali for technological support over the years.

I also would like to thank Dr. José Vera, Dr. Ricardo Camacho, Dr. Eugénio Teófilo, Drª Fátima Candoso, Drª Ana Margarida Oliveira, Drª Manuela Vinagre, Drª Andreia Pinto Ferreira, and all the entities that made this project possible, by providing me with the data I needed. In this regard, I specially thank Professor Teresa Paixão, whose good will made it possible for me to access the Communicable Diseases and Epidemiological Surveillance Center database. She was an enthusiast of the project and of the future utilization of the database to develop an epidemiological model.

I thank my boyfriend, Jorge, for bearing with me and for supporting me, even though he neither understood what I was doing nor saw any point in it, besides the fact that it was important for me.

Last, but certainly not least, I thank Fundação Para a Ciência e Tecnologia for providing financial support.

Table of Contents

Table of Tables

Table of Figures

1. Introduction

The field of Health Economics developed as a branch of Economic Analysis due to the specificities of the health sector. These specificities include not only the right to health protection (established in Article 64 of the Portuguese Constitution) but also the presence of multiple market failures deriving from uncertainty, externalities and incomplete information.

The unique and aggravated combination of such specificities has justified public intervention in health. Public Intervention in healthcare systems assumes many roles, namely those of insurance provider, healthcare provider and regulator.

The exact degree of governmental intervention in health care differs between countries. Portugal, like many other European countries, has opted for a National Health Service type system while the United States has traditionally favored a private sector approach.

Regardless of the level of intervention desired by each society, few argue with the regulatory role played by the State in the health sector. As a regulator, the State issues norms and establishes the rules of the game which condition the activities of those in this sector. As a regulator, it is up to the State to create, through adequate public policies, an environment that promotes equity and efficiency.

In order to reach this objective, it is essential that health care policies be analyzed in the light of economic theory, which is where Health Economics plays a predominant role. Therefore, it is the purpose of the present thesis to attest the need and value of appropriate analysis of the impact of health care regulation. It is also important to emphasize that such analysis should be performed in the specific context of its application, since, in terms of social welfare, the outcome of the actions of the regulator will depend on the context in which such regulation is performed.

With this goal in mind, the value of context specific economic analysis of health regulation is, in the present thesis, highlighted through a series of examples of different health issues in distinct health care sub-sectors. Specifically, the thesis presents three papers on the theme of regulation in different aspects of health care provision and financing. The first two consist of economic analyses of the impact of health care regulation, and the third comprises the creation of an instrument for supporting economic analysis of health care regulation, namely in the field of evaluation of health care programs.

The three articles deal with different health sub-sectors: the health insurance market, the pharmaceutical market and the provision of treatment for HIV infection. The choice of distinct subject matter serves to emphasize the broadness of the theme in question and, thus, highlight the need for a correct pre-analysis of the impact of health care regulation that is not limited to one specific area but must, in fact, be applicable to all problems that regulation tries to minimize.

The themes selected reflect premises of major relevance for health policies. Obviously, many other topics of equal or greater importance exist, but the chosen examples are especially appropriate for the purpose of the thesis. The first two examples evaluate the impact of the same regulatory issue (price regulation) with evidence suggesting opposite effects in terms of the impact on welfare. The first article focuses on an example where price control reduces switching costs while the second provides evidence of increased switching costs due to price control regulation. As such, these two papers demonstrate the need and value of *context specific* economic evaluation of health regulation. The third article emerges naturally as a response to such need by providing an instrument that will precisely contribute to such context specific economic evaluation of health regulation.

The aim of the first paper is to understand the dynamics of health plan entry, exit and pricing in the presence of switching costs and biased selection. Moreover, we aim at understanding how these dynamics are affected by alternative price regulation strategies.

The existing insurance market literature has, generally, modeled equilibrium patterns of adverse selection without trying to model the dynamics of health plan entry and exit, while in the non-health literature, a variety of models examine dynamic pricing strategies with switching costs. Consequently, we build on the existing insurance market literature by developing a new model of health plan competition and pricing, specifically accounting for the dynamics of health plan entry and exit in the presence of switching costs. In order to understand the impact of regulation we simulate the equilibrium under alternative health premium payment systems.

We obviously make a number of simplifying assumptions in order to focus attention on pricing dynamics, plan switching and entry/exit decisions: we assume that plans are *ex-ante* identical; we choose to model entry and exit of health plans into a given employer, not in the overall market; we allow for only two types of enrollees; and we assume the switching costs distribution to be independent of health status. We use a discrete time model to capture the fact that health plan pricing decisions and commercial plan enrollment decisions are each made

only once a year. We find it to be important enough to justify the additional computational requirement resulting from the absence of an explicit terminal condition.

The theoretical model explains the economic rationale behind the insurance provider's strategy of offering heath plans with extremely aggressive prices in the first years, followed by sharp increases in subsequent years. This strategy leads to eventual market exit, once the prices, despite the switching costs, are no longer competitive.

This phenomenon named "*churning*" [1] has been widely referred to in the health policy literature [1, 2]. In an attempt to minimize its impact some States, in the United States of America have opted to administratively establish an upper limit to price increases. More often than not, such policy leads the insurance companies to anticipate product removal from the market, thereby generating elevated switching costs to consumers.

Among our most interesting findings is the possibility that, a first period price floor strategy, can potentially be welfare improving. By reducing competition, the imposition of a price floor, depriving health plans of aggressive entrance, substantially reduces average switching costs plus average premium, and softens death spirals at the same time, as it yields health plans positive intertemporal profits. We also find that in the presence of switching costs, cost sharing is welfare reducing while, on the other hand, risk adjustment yields lower total costs (switching costs plus premium costs) to consumers.

While true that, in Portugal, the health sector is mainly financed by public health insurance, given the growing share of employer-provided private health insurance, the analysis to be performed and the conclusions to be drawn in the first article are also of relevance in the national context, as they may guide future regulatory action in that sector.

The topic for the second article arises from the assertion that the phenomenon observed in the Unites States (U.S) health insurance market, where decreasing profits lead to the eventual product exit, could be the origin of the observed significant turnover rate of products in the pharmaceutical market in Portugal.

In the U.S. health care insurance plan market analyzed in the first article, there are no restrictions to the increase of prices upon entry into the market; nonetheless, the rise of costs

[1] In general terms, "churning" refers to the following phenomenon: The insurance companies competing in prices enter the market with very aggressive prices that only cover the costs while the initial exclusion clauses are valid, but that are too low to cover the costs two years later, once these clauses are no longer applicable. The result is an abrupt rise in prices, one or two years after the product enters the market.

leads to a decrease in profit throughout the life of the product. In the Portuguese pharmaceutical market, once the initial price is set, subsequent price increases are rarely allowed and reductions may be mandated. Posterior price variations are determined by the annual pharmaceutical price revision, announced by the Government. These annual pharmaceutical price revisions set the maximum allowed price increase as a function of the previous year's inflation rate, and are usually set below the inflation rate, which means that the real price decreases over time as the product ages. Although the decrease in the actual price does not go hand in hand with the decrease in profit throughout the life of a product, reality dictates that the decrease of the actual price can imply that the maximization of profits be obtained through the re-introduction of the product in the market.

Thus, one way to regain control over the price is to replace the existing product by a "renewed version" of the same. Although, the price of the new product will still be set under the general rules, which, in essence, take the average price of that of a similar product from four reference countries, a price negotiation takes place once again. The renewed version may be reflected in something as simple as new packaging.

The aim of the second paper is, therefore, to understand whether price regulation of the Portuguese pharmaceutical market creates an incentive for product turnover. The research question is addressed by empirically evaluating, in a reduced-form model, the determinants of product life-cycle in the pharmaceutical market.

In line with analysis performed in other markets, we construct indices of competition, differentiation, price control and quality which are used as explanatory variables in survival analysis. If, indeed, price regulation does create an incentive for product turnover, a significant coefficient on the price control index is expected.

Methodologically, we build on the existing product survival literature by introducing in this literature methods developed in the Industrial Organization literature. While proxies for unobserved product quality have previously been considered as a determinant of exit, we apply a more flexible non-parametric method to the estimation of the unobservable component then obtain consumer preferences for both observable and unobservable characteristics and use those to build an index of product quality. To our knowledge, the methodology followed had not previously been applied in the product survival literature.

While only a dynamic structural model of entry and exit would allow for a full understanding of the impact of any given health policy and the consequences of any changes in the regulatory

environment, we find the proposed methodology, which is considerably more straightforward, to be most adequate as a first approach to examine whether the issue deserves future research.

We find that price control measures do, in fact, influence the decision of product withdrawal and, consequently, have an impact on turnover rates in the pharmaceutical market. While novel product introductions are expected to have a positive welfare impact due to therapeutic improvements, replacement of existing products carries costs to society (switching costs, marketing costs and licensing costs) with little or no benefit to consumers.

The development of the first two articles highlights the importance of context specific economic analysis of regulation in health care. In fact, in the matter of health insurance, examined in the first article, we find price regulation (namely, regulation of entry price in the market) to be potentially welfare improving while, in the pharmaceutical market, considered in the second article, the analysis suggests that price regulation, in its current form, may be welfare reducing.

Considering the need to perform economic analysis on a case-by-case basis, the third article aims to contribute to the process of economic analysis, by creating a support instrument for health care regulation. The instrument consists of a framework of economic evaluation for a broad range of measures aimed at combating a specific disease. Economic evaluation allows for the resources to be given opportunity cost value, thereby contributing to the efficient allocation of resources.

The instrument focuses on the area of Human Immunodeficiency Virus (HIV) infection – one of the priority intervention areas, according to the strategic orientation defined by the government's National Health Plan for 2004-2010. In line with such strategic orientation, the Health Ministry delegated the National Coordination for HIV/AIDS Infection, of the Office of the High Commissioner for Health, the responsibility to create a National Plan of HIV/AIDS Prevention and Infection Control – 2007-2010. The instrument aims to be a support tool for resource allocation decisions in the context of health strategies defined in that Program or others defined by the government in the HIV/AIDS field.

The third paper thus aims at providing *a framework* for cost-effectiveness analysis of HIV treatment in Portugal. It does not aim to compare specific named treatments or programs but, instead, to provide a model, general enough to allow, amongst others, for the comparison of antiretroviral treatment options, of different cut-offs for treatment initiation, or of different adherence improvement strategies. This type of instrument, properly parameterized to the national context, positively contributes, therefore, to an adequate regulation of health care.

The *framework* contributes to the existing Portuguese literature, given that no other model, parameterized with Portuguese cohort data, is available for economic evaluation of a broad range of HIV treatment related decisions. It should be noted that, due to the requirement of economic evaluation studies in applications for drug reimbursement, several economic evaluation models of HIV treatment have, obviously, been developed or at least adapted in Portugal. These models were nevertheless built to compare specific antiretroviral drugs, parameterized with international data taken from the international literature. Since economic evaluation models provided to National Authority of Medicines and Health Products, IP (INFARMED), in the context of reimbursement applications, are not publicly available direct comparison is not possible.

A discrete event microsimulation model is designed to fit the disease progression from start of antiretroviral treatment (ART) to death. The model is parameterized to reflect characteristics of HIV infected individuals at ART initiation in Portugal, as well as the country's clinical practice. Given the unavailability of a national database complete enough for estimation of all parameters of interest, several databases are used.

The work also contributes to the existing literature, in that the methodology applied, namely, microsimulation discrete events model has not, to our knowledge, been previously applied to cost-effectiveness of HIV treatment. Discrete events simulation (DES) has been applied in a variety of settings. Of particular similarity to our model is the one developed by Barton *et al.* [3]- *The Birmingham Rheumatoid Arthritis Model* (BRAM). We expand on their work by modeling the parameters of the Weibull distributions from which time to event is sampled as a function of patient characteristics and history. Such modeling allows for a better fit to the observed data.

DES is selected for three reasons. First, microsimulation (individual simulation) models have been reported to perform better than a Markov model in isolating long-term implications of small, but important, differences in crucial input data [4, 5]. Second, in DES event occurrences determine time advances so random draws are never performed for patients to remain in the same state. Moreover, they do not limit event occurrences to cycle length[2], as occurs in fixed cycle length models. Third, DES model are particularly useful when interaction between individuals is of relevance, as is the case with infectious diseases. Although such interactions are

[2] The shorter the cycle, the more computationally demanding the model will become.

not modeled in the present framework, the framework was designed to allow for its future integration in an epidemiological model.

Given the fact that the model developed is a *framework* for analysis, we have *a priori* no empirical findings to report. Nonetheless, while a model is, by definition, a simplification of reality, its usefulness depends on its ability to reproduce the main aspects under consideration. Overall, the model predicts values close to those observed in the Centro Hospitalar de Cascais (CHC) sample, from which the model was (for the most part) parameterized. Individual parameter values predicted by the model (such as, CD4+ cell count growth, % reaching viral suppression, first line duration, etc.) are also close to those published in the literature, when it is possible to perform such comparisons. It should, however, be noted that validation also requires the ability to reproduce results from different samples and such a step has not yet been performed due to lack of available data.

Even though the model developed is a *framework* for analysis, we do include an application of the model consisting of a cost-effectiveness analysis of two Nucleoside Reverse Transcriptase Inhibitor (NRTI) plus one Non-nucleoside Reverse Transcriptase Inhibitor (NNRTI) versus two NRTI plus one boosted protease inhibitor (PI/r) as a first line antiretroviral therapy regimen [6] and the results obtained are in line with those obtained by equivalent analysis in other countries [7, 8]. All three analysis reinforce the British HIV Association 2008 [9] recommendations to select NNRTIs as the third agent of choice in treating naïve patients, unless impeditive clinical factors exist.

In conclusion, the thesis as a whole contributes to highlight the importance of economic analysis as a support instrument to health care policies. Although economic analysis, namely, economic evaluation, has been adopted and even become mandatory prior to political decisions in some health sub-sectors, such as in pharmaceutical reimbursements applications, the need and value of such analysis in other health sub-sectors has not yet been fully embraced by decision makers. In the last decade we have witnessed several examples of policy changes based on economic rationale (the separation of the financing and provider roles in the health sector is such an example), but by highlighting the value of information provided by economic analysis, and by reinforcing the need for context specific analysis, the present thesis holds as an argument in favor of its broader utilization and a consequently more efficient allocation of scarce resources.

Knowledge driven by research is built by adding small pieces of a puzzle, one after the other. Looking at what has been done and adding something to it. Throughout the dissertation, we carefully reference the work, developed by others, on which our work is based. The remaining, (unreferenced) developments are the authors' contribution. This dissertation adds three little pieces to the puzzle of health sector regulation through economic analysis. Each chapter adds one tiny piece to the complete picture and, in so doing, we believe that it will be of use to those who in the future will look at what has been done and build from there.

2. Death Spirals, Switching Costs, and Health Premium Payment Systems

Death Spirals, Switching Costs, and Health Premium Payment Systems

Filipa Aragão*

Escola Nacional de Saúde Pública, Universidade Nova de Lisboa

Randall Ellis+

Department of Economics, Boston University

ABSTRACT

This paper develops a model of health plan competition and pricing in order to understand the dynamics of health plan entry and exit, in the presence of switching costs and alternative health premium payment systems. We build an explicit model of death spirals, in which profit-maximizing competing health plans find it optimal to adopt a pattern of increasing relative prices culminating in health plan exit.

We find the steady-state numerical solution for the price sequence and the plan's optimal length of life through simulation and do some comparative statics. This allows us to show that using risk adjusted premiums and imposing price floors are effective at reducing death spirals and switching costs, while having employees pay a fixed share of the premium enhances death spirals and increases switching costs.

JEL Classification: I18, I11

Keywords: health insurance, switching costs, premiums payment systems, biased selection

*Support for this research has been provided by Fundação para a Ciência e Tecnologia, Programa POCTI - Formar e Qualificar - Medida 1.1., grant Praxis XXI/BD/19954/99

+Support for this research has been provided by the Agency for HealthCare Research and Quality (AHRQ), grant number R01-HS10620-01A

We are thankful to Ipek Gurol, Guillermo Caruana, Tom McGuire and Andy Weiss and seminar participants at Brown University, Boston University, Minnesota University and University United Methodist Church, Chapel Hill and Nova University of Lisbon for helpful discussion. The opinions expressed are the authors alone.

Abbreviations and notation

Abbreviations and notation	Description
hat notation	Competitor defined value
star notation	Optimal values
AP	Average premium
ASWC	Average switching cost
AVC_t	Average cost in period t
BU	Boston University
C_M	Annual costs of enrolling a sick person
C_N	Annual costs of enrolling a healthy person
d	Rate at which employees leave the firm per year
FC	Annual fixed costs
f_t	Probability of not switching from a plan aged t
F_t	Cumulative probability of not switching from a plan aged t
HPPC	Health plan participation constraint
IHS	Involuntary healthy switchers
IS	Involuntary switchers
ISS	Involuntary sick switchers
k	(1-s)
l	(1-d)
M_t	Number of sick individuals
NE	New employees
N_t	Number of healthy individuals
P_{min}	Lowest premium charged by any competitor
P_t	premium charged by a plan aged t
s	Rate at which healthy employees become sick per year
SW_t	Expected number of people switching out of a t year old plan
T	Health plan´s life span
THC	Total health costs of the firm's employees
VHS	Voluntary healthy switchers
VS	Voluntary switchers
VSS	Voluntary sick switchers
v_t	Cumulative discounted profits from period t onwards
w	Switching cost
W	Maximum switching cost
α	Risk adjustment factor
Δ	Variation
λ	Proportion of premium payed by the employee under cost sharing
ρ	Discount factor
Ψ	Total number of plans offered by the employer
$\prod_t$	Health plan profits in period t

2.1. Introduction

This paper is inspired in part by the history of health plan entry, exit and pricing in the Boston area. Both Boston University (BU) and Harvard University experienced rapid rates of premium escalation during the 1990's, culminating in numerous health plans being canceled by employers or plan providers. The Harvard University Experience is analyzed in Cutler and Reber [10], while the BU Experience is discussed further below.

We develop a model of health plan competition and pricing in order to understand the dynamics of health plan entry and exit, in the presence of switching costs and alternative health premium payment systems. Switching costs are defined as costs incurred by the enrollee of a health plan if and when he decides to change his health plan. These costs may result, among others, from required paper work, change in health care provider or information gathering. We build an explicit model of death spirals, defined as a pattern of increasing relative prices culminating in health plan exit, in which profit-maximizing competing health plans find it optimal to adopt such a pattern. Plans may be forced to enter such death spirals, because of the introduction of new plans that charge low premiums, thereby attracting the employees with lower switching costs.

In order to focus attention on pricing dynamics and plan switching, we construct a model in which all of the plans are *ex ante* identical. We fully appreciate that product differentiation is one factor in recent plan entry and exit, among various types of managed care plans. However, we believe that emphasis on these differences clouds the understanding of some of the basic forces driving pricing, entry and exit. Regardless of whether plans are identical or differentiated, entry of new plans creates an important asymmetry, changing the mixture of healthy and sick enrollees in each health plan. Plan entry and exit permits a recurring cycle of pricing, in which the costs of enrollees in existing plans are always increasing. Rather than product differentiation, our focus on identical plans emphasizes this process as a rationale for plan entry.

The existing insurance market literature has generally modeled equilibrium patterns of adverse selection without trying to model the dynamics of health plan entry and exit [11-16]. Rothschild and Stiglitz [11] analyze adverse selection equilibrium in multiple plan types, but without modeling the process of entry, exit, or changes in pricing over time. Neipp and Zeckhauser [17] introduce the notion of death spirals and examine the empirical patterns, but do not attempt to build an analytical model, and treat cost increases of existing plans as exogenous to pricing

decisions. Keeler, Carter and Newhouse [18] develop a model of consumer choice among differentiated health plans, and use a dynamic process to understand equilibrium enrollments, premiums and social costs, however they do not attempt to model health plan entry and exit. Vistnes, Cooper and Vistnes [19] use a two period game in which health plans first compete to be selected by employers and, if selected, compete to be chosen by consumers, but it does not explore the dynamics of pricing or exit *per se*. Ellis and Ma [20] consider a related issue by analyzing the impact of employer provided health insurance on job turnover. The authors focus on the firms' insurance offer decision, while we focus on the health plans' offer decision.

In the non-health literature, a variety of models examine dynamic pricing strategies with switching costs. Our model shares some features with that of Beggs and Klemperer [21][3], but their framework focuses on two infinitely lived firms, and does not support death spirals in which plan prices increase as market share shrinks, to the point where a firm decides to exit. Gale and Rosenthal [22] develop a model of firms whose quality is only imperfectly observed[4].The model in this paper differs from theirs in that it focus on biased selection[5] rather than adverse selection.

Switching costs introduce the potential for biased enrollment, if they differ across enrollees and are correlated with health costs [10, 16]. Strombom *et al. [23]*, examine switching among employees of the University of California system during the mid 1990's. They find that, under a fixed contribution scenario, consumers price elasticity ranges from -1.8 to -10 indicating considerable price sensitivity but also a significant heterogeneity among consumers[6]. This finding, together with the fact that new employees are, on average, much healthier than existing enrollees, lays the basis for potentially severe biased selection[7].

Grönqvist [24] empirically analyses the biased selection problem, arguing that this can explain the limited empirical evidence for adverse selection in insurance markets in the literature, and presents a model of insurance choice focusing on the decision of whether to purchase

[3] Beggs and Klemplerer examine the implications of switching costs in an infinite period market in which new consumers arrive each period that are not yet attached to a seller.

[4] In that model firms enter at a high quality and high price. Once they have established themselves to have this high quality and before exiting, firms switch to low quality, in order to take advantage of the greater profitability of charging high prices with low quality.

[5] Biased selection includes both favorable selection (which plans want) and adverse selection (which enrollees want individually).

[6] The analysis suggests that the healthiest enrollees may be four times as responsive to price differences as the sickest group of employees. This difference increases nine fold when they also allow for variation in age and employment tenure. Age and tenure are shown to significantly (and negatively) affect price sensitivity.

[7] Mean annual premiums in the Strombom *et al.* sample are approximately $3900, while employee annual spending on premiums averages just under $700 (18%). His demand model predicts that a $60 per year increase in the employee premium will result in a 50% reduction in the proportion of healthy individuals choosing a plan, versus only an 8.8% reduction for sick enrollees.

insurance, which is a complement to our analysis. Most recently, Cutler *et al.* [25] estimate the determinants of health plan switching (adverse selection, adverse retention, aging, demographics and the level of cost-sharing) and incorporate these in a simulation model to predict the long-term dynamics. Their analysis focuses on the determinants of switching between two different types of plans while we abstract from plan differences and focus instead on entry, exit and dynamic pricing behavior.

Several published empirical papers support the idea of death spiral and biased selection in the health insurance market [26-29]. Pauly *et al.* [30], on the other hand, argue that the phenomenon interpreted as death spiral may instead be an adjustment towards more preferred products which would have occurred even in the absence of adverse selection[8]. Our model provides a rational for death spirals even in the presence of identical plans.

Of direct relevance to our theoretical framework is the recently published article by Handle [31] who builds a structural model to empirically investigate consumer switching costs in the context of health insurance markets, accounting for potential adverse selection. The author estimates an average switching cost of $1,570 (75% of the average employee premium).

Fixed costs for health plans prevent plans from continuing to exist with minimal enrollments at very high prices, and are plausible given that certain transaction and administrative costs are independent of the number of plan enrollees. Biased selection and fixed costs are, thus, the driving forces of the health plan death spirals modeled here.

Death spirals are a consequence of how contracts are written between health plan providers and employers, not a characteristic of the plans themselves. Therefore, we model entry and exit of health plans to a given employer, not entry and exit of health plans overall. So, the same health plan can be a new entrant with one employer, while being a long time incumbent with other employers. Glazer and McGuire [32] have documented that health plans in the Boston area routinely charge different premiums for identical plans. Once plans are allowed to price discriminate between different employers, death spirals can occur, even when health plans are profitable.

Our model generates many insights. Plan deaths impose much higher switching costs than voluntary switches, since higher prices induce the lowest switching cost enrollees to switch

[8] Their conclusions are drawn from the fact that "implementing a significant risk adjustment had no discernable effect on adverse selection against the most generous indemnity insurance policy"

plans, while plan deaths force those with the highest switching costs to involuntarily change plans. We use the model to simulate how premium cost sharing schemes, risk-adjustment policies and the imposition of price floors, affect entry and exit, as well as health plan pricing and enrollee choice patterns.

Among our most interesting findings is the possibility that a first period price floor strategy can potentially be welfare improving. By reducing competition, the imposition of a price floor deprives health plans of aggressive loss leading entry. A price floor can potentially reduce the sum of switching costs plus premiums and soften death spirals at the same time, as it yields health plans positive intertemporal profits.

We also find that, in the presence of switching costs, cost sharing is welfare reducing, while, in contrast, risk adjustment yields lower total costs (switching costs plus premium costs) to consumers.

While the analysis has been inspired by the dynamics of entry and exit of health plans in the United States, lessons to be learned from this analysis are relevant to other countries, such as Portugal, where the employer-provided private insurance market is growing [33]. A proper health policy environment will minimize phenomena's such as "*churning"*[9]and death spirals observed in countries where health care is extensively financed through private health insurance.

The remainder of the paper is organized into four sections. In Section 2.2, a motivating example drawing upon empirical results from three markets in the United States is presented. Sections 2.3 and 2.4 describe the analytical model and discuss the attainable analytical results in a model with no closed-form solution. In Section 2.5 the four premium payment systems under consideration are presented in the context of the analytical model. In Section 2.6 we summarize the steady-state equilibrium solutions to our model under different premium payment systems and perform some sensitivity analysis. We conclude the paper with a discussion of the implications and limitations of our model.

[9] In general terms, "churning" refers to the following phenomenon: The insurance companies competing in prices enter the market with very aggressive prices that only cover the costs while the initial exclusion clauses are valid, but that are too low to cover the costs two years later, once these clauses are no longer applicable. The result is an abrupt rise in prices, one or two years after the product enters the market.

2.2. Motivating Evidence

Death spirals are not occurring in all employers or in all parts of the United States, but they do appear to be occurring under some contracts, and Boston University seems to be an example of this. Results in Cutler and Reber [10] and Yu, Ellis and Ash [34] suggest that the problems facing Boston University are similar to those experienced by other large employers in Massachusetts.

We present the evolution of total premiums for non-retired single employees at Boston University and compare it to two other markets: Minnesota and California premiums for state employees. A health plan is here defined as any separately, independently named and marketed insurance product for which a separate premium is charged. In many cases, the same insurance company offers multiple health plans.

2.2.1. Premium Patterns

Table 1 presents data on health plan entries (births) and exits (deaths) from the available portfolio of three employer markets (Boston University, Minnesota and California). The table reveals that the three employer markets have very different rates of entry and exit over time. Boston University has had the highest rates of entry and exit. California, despite having the largest number of health plans offered to its state employees, has the lowest rate of exit, and Minnesota lies in between the two.

Table 1: Key Descriptive Statistics

	Sample Period	Years	Plan Births	Plan Deaths	Average Number of Plans	Average Plan Births per Year	Average Plan Deaths per Year	Average Percentage of Plans Dying each Year
Boston University	1987-2001	14	11	9	6.10	0.79	0.64	10.5%
Minnesota	1984-1995	12	3	6	8.00	0.25	0.50	6.3%
California	1984-1995	12	6	2	22.70	0.50	0.17	0.7%

Changes in premium cost sharing occurring in each region during the sample periods are shown on the top of Figure 1 to Figure 4. The figures suggest that those changes may have a significant impact on premiums dispersion and exit rates.

2.2.1.1. *Boston University*

Boston University has offered fifteen different health plans over the 14 year period from 1987-2001. There were eleven new plans offered, and nine plan deaths. The overall number of plans

has remained relatively constant at five or six per year, implying a plan death rate of about 10.5 percent per year.

Figure 1 illustrates the three primary features that we model:

1. New plan entrants are generally priced lower than existing plans[10].
2. Existing plans tend to increase their premiums relative to the average
3. Higher premium plans have a greater probability of exiting than low premium plans.

Indeed, except in 1992[11], average premiums among exiting plans were always above average premium of those not exiting.

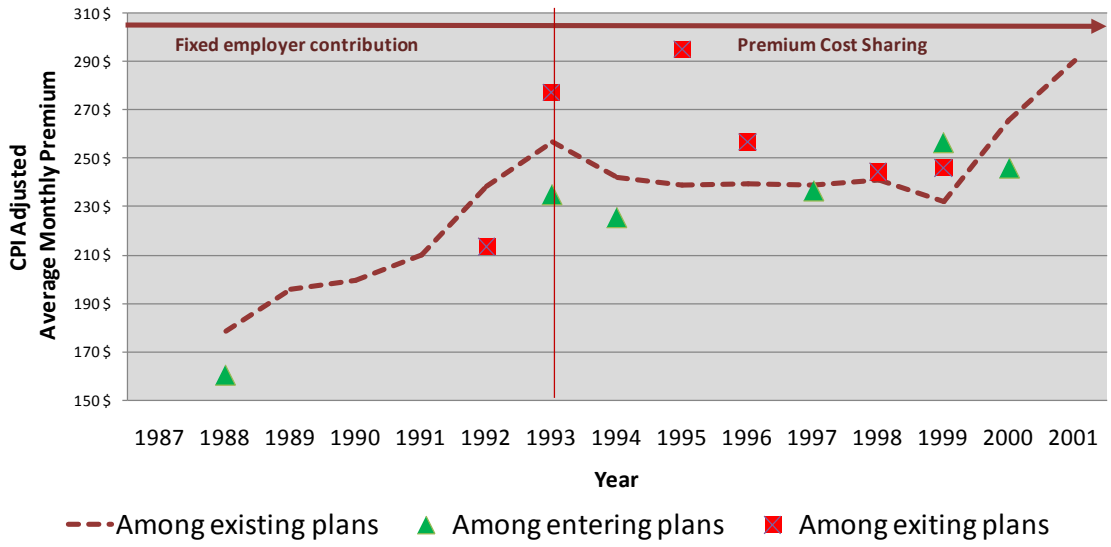


Figure 1: Boston University Average Single Premiums 1987-2001

From 1987 to 1992, BU made level premium contributions towards the six health plans that were offered at that time. As a result, employees faced the full cost burden of premium differences across plans. In 1993, BU changed its premium cost sharing so that, instead of level contributions, BU paid a fixed percentage of all premiums[12]. Greater dispersion in premiums across plans in their real costs in 2001 is evident in Figure 2.

[10] The only entrance above average occurred in 1999 with a plan which was offered for a single year.

[11] This was the average of two plan deaths, only one of which was priced below the average premium of exiting plans.

[12] BU also adopted a strategy of freezing enrollment in two health plans, the highest cost FFS plan (BCBS Comprehensive) and the highest cost HMO (Tufts). Except for those already enrolled in these plans, employees are not allowed to join them. BU also decided to enter the market in 2000 with its own plan, the BU Medical Center Preferred HMO. We do not try to model these alternative employer strategies.

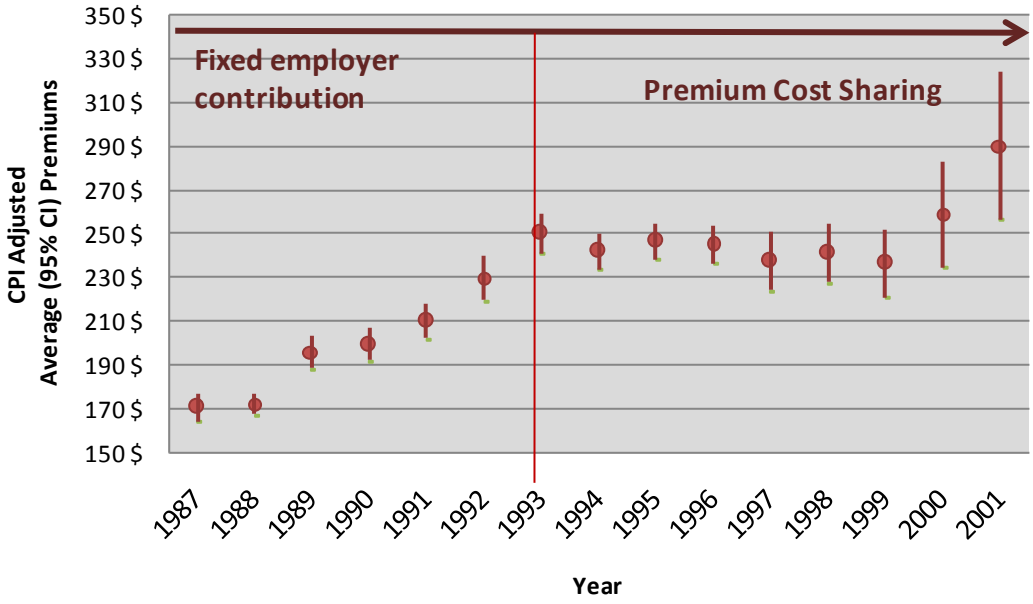


Figure 2: Boston University Average (95% CI) Single Premiums

2.2.1.2. *Minnesota*

Minnesota offered twelve different health plans to its state employees over the sample period ranging from 1984 to 1995. There were only three plan births versus six plan deaths[13] implying a plan death rate of 6.3 percent per year, and suggesting some evidence of death spirals. However, the most evident plan deaths occurring above the average premium happened in 1989. This was a year after Minnesota changed its health premium payment system, moving from a price floor system to a level premium contribution. Under the new system, employer contributions are calculated as equal to the premium of the lowest cost plan. Figure 3 shows premiums tracking together very closely from 1984-1988, at which point an enormous change in the dispersion of premiums occurred, quickly resulting in three plan deaths (marked with stars). This increased dispersion is plausible due to Minnesota switching from a price floor regimen to a fixed employer contribution.

[13] Several of which reflect mergers rather than true plan exits.

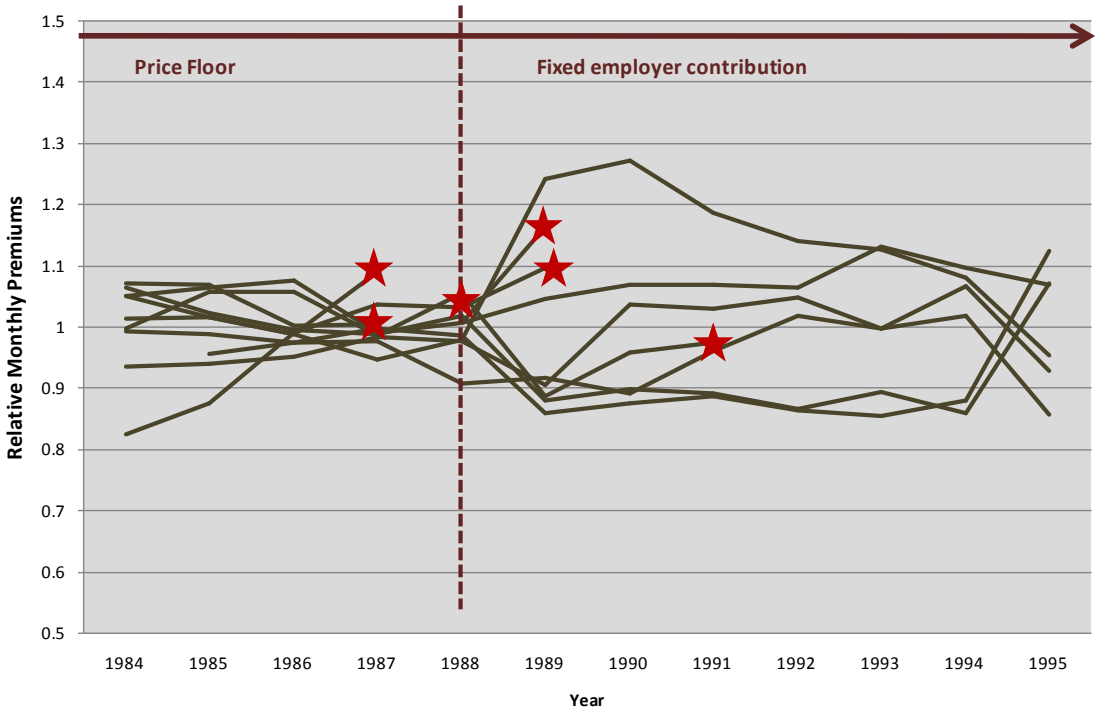


Figure 3: Minnesota Relative Single Premiums 1984-95

2.2.1.3. *California*

California offered 26 plans over the twelve years for which we have premium data. There were six plan births and two plan deaths, implying an annual plan death rate of only 0.7 percent. There is no evidence of death spirals in the California sample: the (only) two plan deaths in our sample period were at average rather than high premiums.

California had consistently maintained a system where the state's contribution is not tied to the lowest cost premium. California premiums contributions are such that employees pay no premium for all but the top four of its twenty three plans in 1996. This fact may be linked to the striking pattern shown in Figure 4: the high number of premiums clustered at the price floor imposed by the state. Moreover, even plans charging premiums higher than the price floor seem to survive for many years rather than dying, as was the case in Massachusetts and Minnesota.

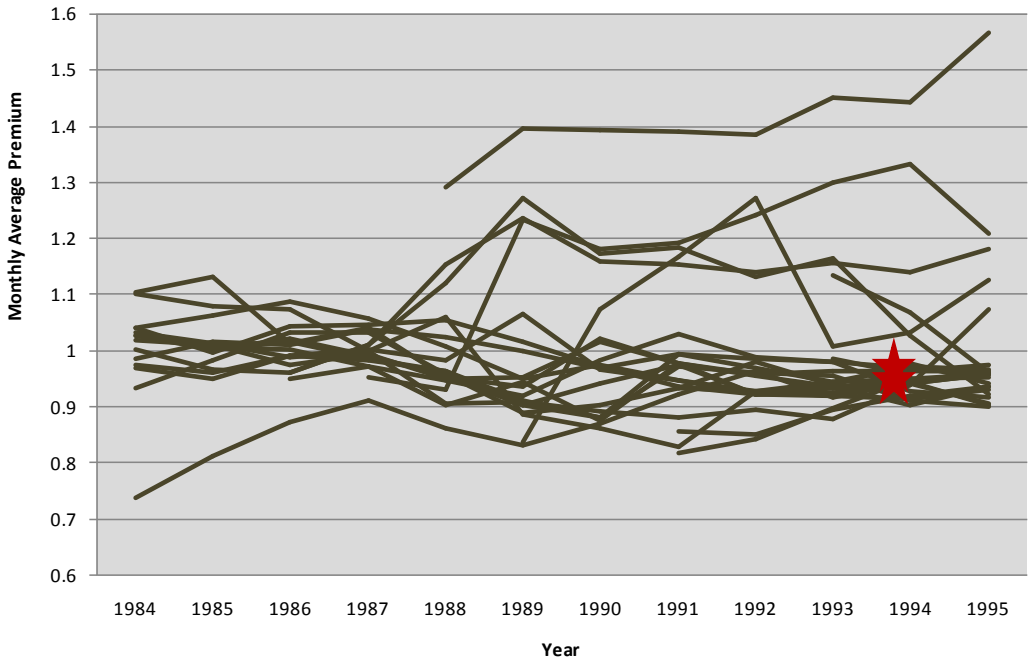


Figure 4: California Relative Single Premiums 1984-95

With no intention of generalizing based on only three markets, the premium patterns above suggest four patterns. First, death spirals do occur under some circumstances. Second, proportional premium cost sharing is associated with increases in the dispersion of premiums across plans. Third, maintaining a price floor is associated with lower rates of plan exit and slower rates of premium growth. Fourth, years in which the premium payment system is changed causes enormous turmoil in health plan pricing, entry and exit.

2.3. Model

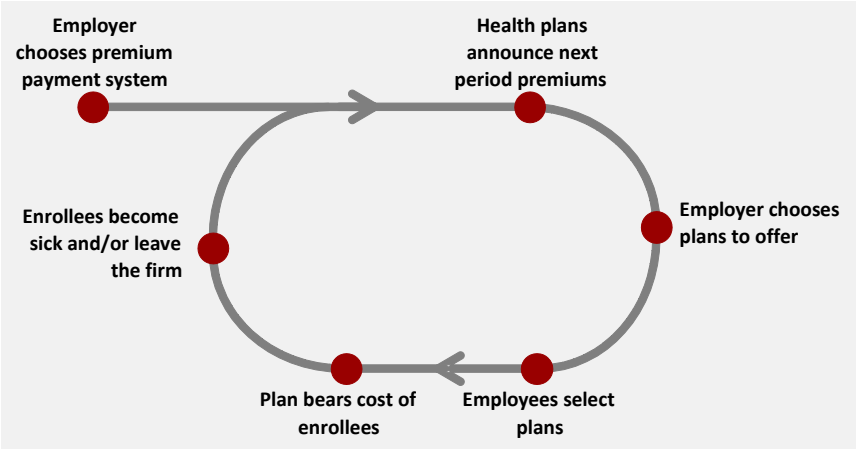


Figure 5: Timing

We model the process as a repeated game, as represented by Figure 5. At time 0, the employer chooses a premium payment system, which is to say a mechanism for making payments to health plans and charging premiums to employees. This decision is made only once, at the beginning of the process, and we model how plans and consumers respond to this premium payment system in equilibrium, without trying to model how agents would react to changes in the premium payment system.

In each period thereafter, there are five stages. During the first stage both potential new entrants and incumbent health plans simultaneously announce their premiums. Existing plans, which are unable to find a price yielding a nonnegative expected sum of future profits, exit.

In the next stage, the employer chooses which plans to offer. Since all plans are, *a priori*, alike, the main reason for the employer to offer a new plan is to enable price competition among plans and keep premiums down. Among several potential new plans, the employer chooses one new entrant each period - the entrant announcing the lowest first period premium. If several plans offer the same lowest price, the employer randomly picks one.

In the third stage of the process, each employee selects one of the offered health plans. Since all plans offer identical benefits, the only information employees use when choosing plans is the premium they will have to pay for each plan and the switching costs they will have to incur, should they switch plans.

Once employees have made their choice, the next step is for plans to bear the costs of treating health care needs according to health status of their enrollees. The fifth, and final stage, occurs after costs are borne by the plans but before the prices are chosen for the next period. Chance determines whether a healthy consumer becomes sick and/or departs the firm, and whether a sick enrollee leaves the employer. The five stages are repeated in each subsequent period.

2.3.1. Employers

The employers are assumed to care about the sum of total premium payments plus total switching costs. Even though employers do not directly bear the burden of switching costs, they should care about the value of the health plan benefits for their employees. The employer is assumed not to care about switching costs borne by employees leaving the firm.

The employer benefits from competition among potential entrants, but cannot force any plan to commit to a price beyond the current period. If plans are *ex ante* identical, the employer's best strategy is to select only one new entrant each period - the health plan announcing the

lowest first period premium. This will avoid duplication of fixed costs and induce Bertrand competition among potential new entrants[14].

2.3.2. Employees

There are only two types of employees: low cost (healthy) and high cost (sick). By sick employees, we mean chronically ill patients. All other employees are considered healthy. All new employees are assumed to be healthy and we normalize the number of new employees to be one. Each employee has single coverage: family contracts are not modeled. Healthy employees become sick at the rate s per year, while sick employees never recover to become healthy again. All employees leave the firm at the rate d per year, for reasons that are independent of their health status[15]. Employees leave the firm for reasons that are exogenous and independent of plan prices and enrollee switching costs.

From the consumer's point of view, all plans offer the same coverage and give the consumer the same utility, except that consumers bear a switching cost when changing from their current plan to another. Consequently, when switching, consumers always choose the health plan charging the lowest premium. If, the optimal price sequence is increasing in time, the health plan charging the lowest premium will always be the youngest plan.

A more general model would consider heterogeneity in switching costs. For the time being, we assume that each time an employee, either healthy or sick, abandons a health plan and joins a new one, she is assigned a switching cost w, where w is uniformly distributed over the interval $[0, W]$.

The structure we have just described implies that we can take expectations across all individuals in a health plan and represent the expected number of individuals in the plan at time t as the vector $\{M_t, N_t\}$, where N_t is the number of healthy individuals, and M_t is the number of sick individuals.

In the absence of any plan switching, the stochastic process can be represented by the following stationary transition matrix A:

[14] Not offering a new plan each year is never optimal since existing firms will raise premiums even more. We do not model the optimal length of a period.

[15] We have also experimented with departure rates that are correlated with health states. This added complexity with few new insights. Ellis and Ma (2011) show that the higher job turnover of young employees more than offsets the predictable turnover of the older retirees.

$$\begin{bmatrix} N_t \\ M_t \end{bmatrix} = A \begin{bmatrix} N_{t-1} \\ M_{t-1} \end{bmatrix}$$

where, $A = \begin{bmatrix} (1-s)(1-d) & 0 \\ s(1-d) & (1-d) \end{bmatrix}$. With plan switching, the transition matrix is no longer stationary. In this case, the transition matrix, A_t, will be:

$$A_t = f_t \, A$$

where f_t is the probability of a current enrollee not switching from a plan of age t. The complete specification of this probability depends on the premium payment system as described in Section 2.5. The expected number of people switching out of a t year old health plan, SW_t, can be written as:

$$SW_t = [A - A_{t-1}] \begin{bmatrix} N_{t-1} \\ M_{t-1} \end{bmatrix}$$

Equation 1: Expected number of people switching out of a *t* year old plan

2.3.3. Health Plans

Every year, many potential new entrant health plans compete *à lá Bertrand* to gain access to the market. Since only one new health plan is selected each year, as long as there are at least two potential new entrants, first period premiums are bid down to the point where the sum of discounted profits over the life span of the plan is exactly zero; unless a first period price floor is set by the employer.

Insurance companies are assumed not to be allowed by employers to charge different premiums to different individuals in the same plan. However, because of switching costs, plans are not perfect substitutes and, consequently, each plan faces a downward sloping demand curve. Insurance companies are thus price setters, and each year they choose the premium to charge all enrollees from a given employer.

Health plan profits in period t ($\prod_t$) can be written as:

$$\prod_t = (P_t - C_N)N_t + (P_t - C_M)M_t - FC$$

where P_t is the premium charged by a plan aged t, FC are the fixed costs, and C_N and C_M are the one year costs of enrolling a healthy person and a sick person, respectively. Cumulative discounted profits from period t onwards, V_t, is given by:

$$V_t = \sum_{i=t}^{T} \rho^{i-t} \prod_i$$

where ρ is the discount factor and T is the health plan´s optimally chosen life span.

At any age t, if V_t is negative, it is optimal for the plan to exit. Consequently, in order for the plan to remain active at age t, the following health plan participation constraint (HPPC) has to be satisfied:

$$HPPC: V_t \geq 0$$

2.4. Results

Given the structure of our model, with one unit of healthy new employees arriving each period, the proportion s of existing healthy employees becoming sick each period, and the proportion d of existing healthy employees departing each period, then the steady-state total number of healthy enrollees, (N), that will be choosing among all of the health plans is:

$$N = \frac{1}{(1-kl)}$$

Equation 2: Steady-state total number of healthy enrollees

where, $k = (1-s)$ and $l = (1-d)$. Similarly, since a proportion s of healthy workers are becoming sick each period, and a proportion d of both newly sick and previously sick are departing the firm, then the steady-state total number of sick employees in the firm at any moment, (M), must be:

$$M = \frac{l}{(1-l)} - \frac{kl}{(1-kl)}$$

Equation 3: Steady-state total number of sick employees

Regardless of how workers distribute themselves among the total number of plans offered by the employer (Ψ), total health costs (THC) of the firm's employees will be:

$$Total\ Health\ Costs = \frac{1}{(1-kl)} * C_N + \left[\frac{l}{(1-l)} - \frac{kl}{(1-kl)}\right] * C_M + \Psi * FC$$

2.4.1. The Social Optimum

It is straightforward to see that the social optimum would have only one plan (and hence plan lifetime, T , will be infinite), and it would charge a constant premium over time. Having a constant premium guarantees that consumers have no incentive to switch plans, thereby minimizing switching costs. From Equation 1 it can be seen that switching costs would be zero. Having a single health plan also minimizes the fixed costs. Hence we have:

$$P^{FB} = \frac{\frac{1}{(1-kl)} * C_N + \left[\frac{l}{(1-l)} - \frac{kl}{(1-kl)}\right] * C_M + FC}{\frac{1}{1-l}}$$

$$T^{FB} = \infty$$

2.4.2. Competitive Solution

As described in Section 2.3.3, in the absence of a price floor, competition among potential entrants reduces overall discounted profits to zero. The first period price (P_1) is thus implicitly defined by:

$$(P_1 - C_N)N_1 + (P_1 - C_M)M_1 - FC + \sum_{t=2}^{T} \rho^{i-t} \prod_t = 0$$

Equation 4: Lifetime profit

In order to solve Equation 4, three preliminary steps are required: The first is the specification of the optimal sequence of prices and profits, given T; the second is the definition of the vector $\{N_1, M_1\}$, also for a given life span, T; The third, and final step, is the determination of the optimal life span. We turn to these issues next.

2.4.3. The Optimal Price Sequence

The optimal price sequence is the solution to the following optimization problem:

$$\underset{\{P_1,\ldots,P_T\}}{Max} \quad V_t$$
$$s.t. \begin{cases} HPPC \\ P_t, N_t, M_t \geq 0 \end{cases}$$

We assume Bertrand pricing behavior. Let the hat notation define variables set by competitors and let $\widehat{\mathrm{P_{Min}}}$ be the lowest premium charged by any competitor. As shown in Appendix I, the number of healthy and sick enrollee at any moment in time is a function of the parameters $\widehat{\mathrm{P_{Min}}}$, $\mathrm{P_t}$, and $\mathrm{P_1}$, and of $\{\mathrm{N_1}, \mathrm{M_1}\}$. $\{\mathrm{N_t}\}$ and $\{\mathrm{M_t}\}$, are independent of any other past premiums charged by the plan.

Claim:

$$X_t \equiv \begin{bmatrix} N_t \\ M_t \end{bmatrix} = \begin{bmatrix} F_t l^{t-1} k^{t-1} N_1 \\ F_t l^{t-1} \{M_1 + [1 - k^{t-1}] N_1\} \end{bmatrix}$$

Equation 5: Number of healthy and sick employees in a given plan at time t

where the cumulative probability of not switching up to period t is $F_t \equiv \prod_{i=1}^{t} f_i = \frac{W - P_t + \widehat{P_{Min}}}{W}$.

Proof: See Appendix I.

This result implies that the premium charged by the health plan at any moment t will have no impact on future profits. Consequently, the optimization problem can be solved one period at a time through T- 1 problems of the form:

$$Max_{\{P_t\}} \quad (P_t - C_N) N_t + (P_t - C_M) M_t - FC$$
$$s.t. \begin{cases} \Pi_t \geq 0 \\ P_t, N_t, M_t \geq 0 \end{cases}$$

Problem 1: Maximization problem for period t

This problem is solved by assuming that all restrictions are met, and then verifying if they are indeed met. By solving Problem 1, we obtain the equation defining $P_t, \forall\, t = 2, .., T$.

Claim:

$$P_t = \frac{\left(W + \widehat{P_{Min}}\right)}{2} + \frac{AVC_t}{2}$$

where $AVC_t \equiv \frac{\{C_N k^{t-1} N_1 + C_M [M_1 + (1-k^{t-1})N_1]\}}{N_1 + M_1}$ and $\widehat{P_{Min}}$ is the lowest premium charged by any competitor.

Equation 6: Optimal price at $t > 1$

Proof: See Appendix I.

Moreover, the expression for the optimal price depends on time, exclusively through average costs. Consequently, as long as $\mathrm{AVC_t}$ is increasing in time, so is $\mathrm{P_t}$. As shown in Appendix I, it is easily verified that: $\mathrm{AVC_t} > AVC_{\mathrm{t}-1,}, \forall_{\mathrm{t}}$, as long as $\mathrm{C_M} > \mathrm{C_N}$.

Claim: The optimal price sequence $\{P_2, \dots, P_T\}$ is strictly increasing in t.

Proof: From Equation 6, it is clear that P_t is increasing in AVC_t , the average cost at time t:

$$\frac{\partial P_t}{\partial AVC_t} = \frac{1}{2} > 0$$

The intuition behind the above claim is the following: in our model, two forces cause health plan average costs, and hence health plan premium, to increase as a plan ages. First, new enrollees are healthier than average and perfectly price elastic: they always join the lowest premium plan, lowering the new plan costs relative to existing plans. Secondly, even in the absence of plan switching, the healthy get sicker, raising plan costs over time[16].

The fact that the optimal sequence of prices is increasing in t has an important consequence to our model. In Section 2.3.2 we argued that switchers would always select the health plan with the lowest premium. If prices are increasing with time, this implies that all employees deciding to switch will select the youngest plan ($\widehat{P_{Min}} = P_1$). Equation 6 may, then, be rewritten as:

$$P_t = \frac{W + \widehat{P_1}}{2} + \frac{AVC_t}{2}$$

Equation 7: Equilibrium optimal price at time $t > 1$

[16] Cutler and Reber (1998) highlight the importance of the third factor, which they term "adverse retention." The classic Rothschild and Stiglitz (1976) frameworks focus on the adverse selection resulting from the fact that healthier employees are more likely to switch plans. This fact would occur in our framework if heterogeneity in switching costs was included.

Consequently, changes in pricing by incumbents older than t do not affect the pricing decision of a firm of age t. After establishing its market share in the first period, and given the premium charged by the youngest plan, P_1, the health plan tries to extract the highest possible rent from consumers over which it has market power due to switching costs, *i.e.* its enrollees. Equation 7 is analogous to the well-known monopolist optimal price with $(W + P_1)$ as the consumer's reservation price. Indeed, once we have solved Equation 5 for P_t, it becomes obvious that $(W + P_1)$ may be interpreted as the consumer's reservation price.

This notation for the reservation price also creates a useful criterion of the firms exit decision. Note that the average cost is a deterministic function of the parameters C_N, C_M, k and the initial enrollments N_1 and M_1. The firms exit decision will be when $P_{Max} \equiv W + \widehat{P_1} < AVC_t + AFC_t$. This can be simplified to yield the following exit condition:

$$Exit\ when\ W + \hat{P}^1 < \frac{C_N k^{t-1} N_1 + C_M[M_1 + (1 - k^{t-1})N_1]}{N_1 + M_1} + \frac{FC}{l^{t-1}(N_1 + M_1)}$$

Equation 8: Exit condition

This equation shows that there are two reasons why firms exit: One is because of the increase in the average variable cost of their enrollees as the proportion of healthy enrollees declines relative to the sick (aging); The other is due to the increase in average fixed costs as total plan enrollment declines. This equation makes it clear that in this model, firm exit is exogenous to the pricing strategy of any firm other than the new entrant, which sets $\widehat{P_1}$.

From Equation 8, it can be seen that firm lifetime will decrease as FC increases, plan departure rate increases ($l \equiv 1 - d$ decreases), rate of becoming sick increases, and the maximum support for switching costs increases.

As shown in the Appendix I, substituting Equation 7 back in the objective function of Problem 1, we obtain an expression for the optimal period t profit.

$$\prod_t = \frac{l}{4W}(N_1 + M_1)[W + P_1 - AVC_t]^2 - FC$$

Equation 9: Profit in period t

Claim: Profits, after the first period, are decreasing in t.

Proof: In the optimal profit expression (Equation 9), we see that profit at time t depends on t exclusively through AVC_t and, as it can easily be verified, $AVC_t > AVC_{t-1}, \forall_t$.

Substituting Equation 9 back in Equation 4 yields:

$$P_1 = AVC_1 + \frac{\sum_{t=1}^{T} \rho^{t-1} FC}{N_1 + M_1} - \frac{1}{4W} \sum_{t=2}^{T} \rho^{t-1} l^{t-1} [(W + \widehat{P_1}) - AVC_t]^2$$

Equation 10: Optimal first period price

Proof: See Appendix I.

Equation 10 reveals that P_1, as a function of N_1 and M_1, is actually quite simple. P_1 is the average marginal cost in period one, plus the sum of discounted fixed costs per initial enrollee, less the sum of future profits each person will, on average, generate. Thus, for a given $\{N_1, M_1\}$ vector, the more profitable each person is in the future (because new plans enter less aggressively), the more willing the plan is to incur losses in its first period. As a result, high first period prices are not sustainable in equilibrium unless a price floor is set.

2.4.4. The Optimal Vector of First Period Enrollees

Our next step is to find expressions defining the vector of first period enrollees, $\{N_1, M_1\}$. Since, all switchers and all new employees select the youngest plan, N_1 will be the sum of new employees (NE), plus involuntary healthy switchers (IHS), plus voluntary health switchers (VHS). IHS are employees who, due to their high switching costs, chose never to switch out of the plan that just died, but instead were forced to switch when the plan exited. VHS are enrollees who choose to switch while their health plan was still available. They have, of course, lower switching costs than IHS.

$$N_1 = NE + IHS + VHS = 1 + kl\, N_t + \sum_{t=1}^{T-1} k^t l^t \widehat{F_t} \left(1 - \widehat{f_{t+1}}\right) N_1$$

Equation 11: Number of healthy enrollees in year 1

where,

$$\widehat{f_{t+1}} = \frac{W - \widehat{P_t} + P_1}{W - \widehat{P_{t-1}} + \widehat{P_1}} \text{ and } \widehat{F_t} = \frac{W - \widehat{P_t} + \widehat{P_1}}{W}.$$

When a plan exits, all its former enrollees are forced to switch, thus the probability of not switching, $\widehat{f_{T+1}}$, is zero. Including this, Equation 11 simplifies to:

$$N_1 = 1 + \sum_{t=1}^{T} k^t l^t \widehat{F_t}\left(1 - \widehat{f_{t+1}}\right)\widehat{N_1}$$

Equation 12: N_1

The derivation of M_1 follows approximately the same reasoning as that of N_1, except that, because all new entrants are assumed healthy, the new plan's market share of sick enrollees is simply the sum of involuntary sick switchers (ISS) and voluntary sick switchers(VSS) it is able to "steal" from existing plans:

$$M_1 = ISS + VSS = M_t l + slN_t + \sum_{i=1}^{T-1}\left(\widehat{M_i} l + sl\widehat{N_i}\right)(1 - f_{i+1})$$

Equation 13: Number of sick enrollees in year 1

$$\Leftrightarrow M_1 = \sum_{i=1}^{T}(\widehat{M_i} l + sl\widehat{N_i})(1 - \widehat{f_{i+1}})$$

Equation 14: M_1

2.4.5. The Determination of the Optimal Life Span

Results derived until now have all been conditional on T, the optimal lifespan of a plan. The procedure for the determination of the optimal life span is explained in more detail in Appendix I, but we turn next to a brief description of it.

At the beginning of each period, the plan will determine the price that maximizes its profits, according to the HPPC, $V_t \geq 0$, if no price yields non-negative profits, the plan will exit. The determination of the steady-state optimal life span, T, requires, nevertheless, one additional condition: that it would not have been feasible for the health plan to select the optimal price sequence of any other life span. Any $\tilde{T} > T$ would not be feasible because the HPPC would not have been satisfied for all t, and $\tilde{T} < T$ would not be feasible because P_1 is decreasing in T. Bertrand competition among potential entrants implies that the lowest possible P_1 yielding zero intertemporal profits will be selected.

2.4.6. The Steady-State Competitive Equilibrium

Because one plan is added to the menu of choices each year and one plan dies, the optimal life span, T, is also the number of plans offered by the employer at any moment in time. Moreover, by symmetry, the steady-state equilibrium $N_1 = \widehat{N_1} = N_1^*$, $M_1 = \widehat{M_1} = M_1^*$ and $P_t = \widehat{P_t} = P_t^*, \forall_t$, where the star indicates equilibrium values.

Imposing the equilibrium conditions on Equation 12, Equation 14, Equation 10 and Equation 7 and then substituting this last one on the first three, we obtain the system of equations defining the first period variables of interest, from which all others may be obtained. This system does not have a closed-form solution. Instead, we use it to simulate the model as described in the next section.

$$\begin{cases} \mathrm{P}_1^* = AVC_1^* + \dfrac{\sum_{t=1}^{T} \rho^{t-1} FC}{N_1^* + M_1^*} - \dfrac{1}{4w} \displaystyle\sum_{t=2}^{T} \rho^{t-1} l^{t-1} [(W + \mathrm{P}_1^*) - AVC_t^*]^2 \\ N_1^* = \dfrac{2W}{(1-kl)\sum_{i=1}^{T} k^{t-1} l^{t-1} (W + P_1^* - AVC_t^*)} \\ M_1^* = \dfrac{N_1^* \sum_{i=1}^{T} [l^t (1-k^t) - l^{t-1}(1-k^{t-1})]\,(W + P_1^* - AVC_t^*)}{(1-l)\sum_{i=1}^{T} l^t (W + P_1^* - AVC_t^*)} \\ (P_1^* - C_N) N_1^* + (P_1^* - C_M) M_1^* - FC + \\ + \displaystyle\sum_{t=2}^{T} \rho^{t-1} \left\{ \frac{l}{4W} (N_1^* + M_1^*)[(W + P_1^*) - AVC_t^*]^2 - FC \right\} = 0 \end{cases}$$

Equation 15: Equilibrium conditions

In the price floor scenario, the system is precisely the same, except for the third equation: the equation defining P_1. We assume the price floor is such that first period profits are zero. Consequently,

$$P_1^* = AC_1 + \frac{FC}{N_1^* + M_1^*}$$

2.5. Premium Payment Systems

We consider four possible premium payment systems: Base Case, Cost Sharing, Risk Adjustment and Price Floor.

2.5.1. Base Case

Our base case assumption is that employers make level contributions toward all the health plans offerings, so that employees bear the full cost differential of any premium charges made by health plans. For simplicity, we assume that employees pay the entire premium in this case, and ignore the well-known problem of tax exempt health care spending. In the absence of any discounting, this assumption also implies that total lifetime premiums exactly equals total treatment costs, hence comparisons of pricing paths under this assumption focus on differences in total switching costs.

Since under this system employees pay the full premium, a given employee will choose to switch out of a plan in period t if $P_t > \widehat{P_{Min}} + w$, where P_t is the premium charged by a plan of vintage t and $\widehat{P_{Min}}$ is the lowest premium charged by any competitor. The switching cost, w, is uniformly distributed over the interval $[0, W]$, hence the probability that a current enrollee will not switch from a plan aged t, denoted by f_t, is:

$$f_t = \begin{cases} \dfrac{W - P_t + \widehat{P_{Min}}}{W - P_{t-1} + \widehat{P_{Min}}} \; if \;\; W \geq P_t - \widehat{P_{Min}} \;\wedge P_t \geq P_{t-1} \\ 0 \qquad\qquad\qquad if\; W < P_t - \widehat{P_{Min}} \;\wedge P_t \geq P_{t-1} \\ 1 \qquad\qquad\qquad\qquad\qquad if\; P_t < P_{t-1} \end{cases}$$

Equation 16: Probability of not switching at time t

Under our model assumptions, it turns out that in equilibrium conditions $W \geq P_t - \widehat{P_{Min}}$ and $P_t \geq P_{t-1}$ are both satisfied. As described in the Section 2.3.3, existing plans are unable to charge a price that yields strictly positive profits in the forthcoming periods exit in stage 2. This implies that the first condition defined, $W \geq P_t - \widehat{P_{Min}}$, will always be met for an active plan because if it were not, all enrollees would switch and fixed costs would imply negative profits. The second condition, $P_t \geq P_{t-1}$, means that prices are non-decreasing over time. As we show in Section 2.4.2 this condition is also met in equilibrium.

Intuitively one would expect the probability of switching to be decreasing in W. For a given price sequence, this is easily confirmed by taking the partial derivative of $(1 - f_t)$ with respect to W. However, W will have an impact on the optimal price sequence and possibly on the optimal life span. Without an analytical solution to our model, the total effect has to be simulated, and this is done in Section 2.6.

Switching costs depend on the prices, the number of healthy and sick people in each vintage plan, and especially on how many people are in a plan when it dies. The expected switching cost of enrollees leaving a plan at age t is:

$$\frac{P_t - \widehat{P_{Min}} + P_{t-1} - \widehat{P_{Min}}}{2}[A - A_t]\begin{bmatrix} N_{t-1} \\ M_{t-1} \end{bmatrix}$$

Equation 17: Expected switching cost at *t*

2.5.2. Cost Sharing

The second premium payment system we consider is proportional premium cost sharing. In this system, employees pay only the proportion λ, with $0 < \lambda < 1$, of the full premium P_t, while the employer pays the remaining $(1-\lambda)P_t$. Since employees pay only a proportion of the total plan premium, their willingness to switch health plans will depend on $\lambda\left(P_t - \widehat{P_{Min}}\right)$ rather than $\left(P_t - \widehat{P_{Min}}\right)$, as occurs in the base case. With cost sharing, the probability of a current enrollee not switching from a plan aged t, denoted by f_t^{CS}, is defined as:

$$f_t^{CS} = \frac{W - \lambda\left(P_t - \widehat{P_{Min}}\right)}{W - \lambda\left(P_{t-1} - \widehat{P_{Min}}\right)} = \frac{\frac{W}{\lambda} - P_t + \widehat{P_{Min}}}{\frac{W}{\lambda} - P_{t-1} + \widehat{P_{Min}}}$$

Equation 18: Probability of not switching under CS

The probability described in Equation 18 is identical to f_t, the probability of not switching in our base case, after replacing W with $W^{CS} = \frac{W}{\lambda}$. Hence, with regard to the probability of switching, premium cost sharing is equivalent to increasing switching costs.

2.5.3. Risk Adjustment

The third premium payment system considered involves risk adjustment. Risk adjustment is a supply side payment system that uses signals predictive of enrollee health costs to, partially or fully, compensate health plans for predictable differences in the expected cost of their enrollees. Risk adjustment does not directly affect demand responsiveness or switching, however, it does change the relative profitability of healthy and sick consumers. We implement imperfect risk adjustment in our model by assuming that the employer is able to observe a signal that is imperfectly correlated with enrollee costs. They use this signal to increase payments to the sick and reduce them to the healthy. Hence, under imperfect risk adjustment, health plan profits at time t can be written as:

$$\prod_t^{RA} = (P_t - C_N + \gamma_N)N_t + (P_t - C_M + \gamma_M)M_t$$

Taking expectations, the effect of risk adjustment is to reduce the profitability of healthy enrollees and reduce losses on sick enrollees. We simulate this effect by assuming that $\gamma_N = \alpha(C_N - \bar{C})$ and $\gamma_M = \alpha(C_M - \bar{C})$. Since $C_N < \bar{C} < C_M$, then $\gamma_N < 0 < \gamma_M$. That is, risk

adjustment raises the cost of enrolling healthy people and lowers the cost of enrolling sick people. We model the effect of risk adjustment by replacing C_N with $C_N^{RA} = (1-\alpha)\, C_N + \alpha \bar{C}$, and replacing C_M with $C_M^{RA} = (1-\alpha)\, C_M + \alpha \bar{C}$.

Reflecting the fact that existing risk adjustment formulas are imperfect, in our simulations we assume $\alpha = 0\,.1$; that is, risk adjustment moves costs only ten percent of the way towards fully compensating health plans for enrolling sicker than average enrollees[17].

Risk adjustment affects only the supply side of the market, and does not affect demand responsiveness or plan switching, other than through its effect on pricing behavior. Consequently, the probability of a current enrollee not switching from a plan of vintage t is precisely the same as in the base case scenario.

2.5.4. Price Floor

The fourth and final premium payment system we consider is a price floor on first period premiums. By price floor we mean a minimum premium such that, if a health plan charges a lower premium, employees do not reap any of the reduction. Price floors arise naturally when employers choose a base plan for calculating their employee contribution which is not the lowest premium health plan. While price floors are generally considered bad by economists, in our context they serve the useful purpose of thwarting the ability of the entrant plan to attract the healthiest employees.

In our simulations, the first period price floor value is the price yielding zero profits in the first period. This would be feasible if employers have some, but not full, power to negotiate first period prices. If employers know the true costs of first period plan entrants, for instance, then they might be able to enforce zero profits in the first period, even if they cannot enforce zero profits in subsequent periods.

2.6. Simulation Results

In order to simulate the model, we selected values for the parameters and used Mathematica 4.0® to find the steady-state equilibrium in each of the four premium payment systems considered. Comparative statics is performed by linearizing the system and then, applying the

[17] We are using conventional, not optimal risk adjustment here, as Glazer and McGuire (2000) define it. That is, we assume that the employer does not use its knowledge of the demand model of how consumers are selecting health plans to risk adjust payments.

implicit function theorem, again use Mathematica 4.0®, to understand the impact of each parameter on the equilibrium. This analysis is, obviously, performed locally around the equilibrium found, which itself depends on the values selected for the parameters. We conclude this section with some sensitivity analysis meant to verify the degree of dependence of our results on the parameter values chosen.

2.6.1. Starting Values

Table 2: Simulation Parameters

	Sample Period	Years
Number of New Entrants Each Period	**NE**	1
Proportion of Healthy Who Become Sick	**s**	0.03
Proportion of Employees Departing the Firm	**d**	0.1
Annual Medical Cost per Healthy Enrollee	**CN**	1,000
Annual Medical Cost per Sick Enrollee	**CM**	2,000
Maximum Switching Cost	**W**	1,000
Plan Fixed Cost per Year	**FC**	100
Discount Factor	**ρ**	1

Table 2 displays the assumed values of our base case simulation model. Healthy enrollees are assumed to cost $C_N = \$1{,}000$ per year, with sick enrollees costing twice as much $C_M = \$2{,}000$. The maximum price differential that will result in all enrollees switching is $W = \$2{,}000$. Employees, both healthy and sick, are assumed to leave the firm at the rate $d = 0.1$. The rate at which healthy people become sick was selected to insure that there were a total of eight units of healthy employees and two units of sick employees. From Equation 2 and Equation 3, we obtained $k = \frac{7}{8l}$, which is to say that healthy people become sick at a rate, s, of approximately three percent. Health plans are assumed to have fixed costs of $FC = \$100$ per plan/year. We assume no discounting for our baseline. These six parameters $(C_N, C_M, W, s, d, FC, \rho)$ fully characterize our model for a given premium payment system.

With eight units of healthy employees and two units of sick employees, the average treatment costs will be:

$$\frac{8 * 1000 + 2 * 2000}{10} = \$1200$$

Once fixed costs are added, we obtain the social optimum premium that one (single) plan could charge and achieve zero profits:

$$P^{SO} = 1200 + \frac{100}{10} = 1210$$

2.6.2. Competitive Equilibrium Simulation Results

In this section we provide an overview of the simulation results obtained from dozens of reasonable combinations of the six parameters $\{s, l, W, C_N, C_M, CF\}$. In almost all the sets of parameters chosen, the model has one (unique) solution yielding real, positive values for all variables. The exception occurs for low values of W (the highest possible switching cost) in which case no solution exists. We believe that this is due to the fact that for low values of W, the model reaches close to the Bertrand paradox, where the first order conditions no longer hold.

We discuss the simulation results by first commenting on the determination of the optimal life span (Section 2.6.3) and then on the equilibrium values of our base case (Section 2.6.4). Once the main features of our model are presented, the comparison of equilibria under the four systems (Section 2.6.5) should run smoothly.

2.6.3. Determination of the optimal life span

In all but the price floor scenario, the determination of the optimal life span of the plan follows the following steps (exemplified in Table 3):

1. Assume the life span is $\tilde{T} = 2$
2. Find the steady-state optimal stream of prices and profits given that $\tilde{T} = 2$
3. If at time $\tilde{T}$, profits are positive, assume the life span is $\tilde{T} = 3$
4. Find the steady-state optimal stream of prices and profits given that $\tilde{T} = 3$
5. If at time $\tilde{T}$, profits are positive, assume the life span is $\tilde{T} = 4$

This loop goes on until the optimal price strategy yields negative profits at time $\tilde{T}$. When it occurs, it means the optimal life span, T, is $T = \tilde{T} - 1$

Table 3: Determination of T

Life Span		3	4	5	6	7	8	9	10	11
Average Premium		$1,397	$1,422	$1,439	$1,451	$1,462	$1,472	$1,481	$1,490	$1,499
Average Switching Cost		$240	$202	$176	$157	$142	$130	$120	$111	$104
Average Premium plus Average Switching Cost		$1,637	$1,624	$1,614	$1,608	$1,604	$1,602	$1,601	$1,602	$1,603
Inv. Healthy Switchers		1.24	0.89	0.67	0.52	0.41	0.33	0.27	0.22	0.18
Inv. Sick Switchers		0.42	0.32	0.26	0.22	0.18	0.16	0.14	0.12	0.11
Vol. Healthy Switchers		2.67	2.48	2.34	2.23	2.14	2.06	1.98	1.92	1.86
Vol. Sick Switchers		0.71	0.64	0.58	0.53	0.49	0.46	0.43	0.40	0.38
% Inv. Switchers		16.6%	12.1%	9.3%	7.4%	6.0%	4.9%	4.1%	3.4%	2.9%
% Vol. Switchers		33.8%	31.2%	29.2%	27.6%	26.3%	25.1%	24.1%	23.2%	22.4%
Cost Inv. Healthy Switchers		$1,007	$735	$560	$439	$350	$283	$231	$191	$158
Cost Inv. Sick Switchers		$341	$268	$220	$185	$158	$136	$118	$104	$91
Cost Vol. Healthy Switchers		$830	$806	$782	$758	$736	$714	$693	$672	$652
Cost Vol. Sick Switchers		$221	$207	$195	$183	$173	$163	$154	$146	$139
Premium	1	$986	$941	$912	$893	$880	$873	$868	$867	$867
	2	$1,598	$1,572	$1,554	$1,541	$1,532	$1,525	$1,520	$1,516	$1,514
	3	$1,609	$1,583	$1,565	$1,552	$1,543	$1,536	$1,531	$1,528	$1,526
	4		$1,594	$1,576	$1,563	$1,554	$1,547	$1,542	$1,539	$1,537
	5			$1,587	$1,574	$1,565	$1,558	$1,553	$1,550	$1,548
	6				$1,584	$1,575	$1,569	$1,564	$1,561	$1,559
	7					$1,585	$1,579	$1,574	$1,571	$1,569
	8						$1,589	$1,584	$1,581	$1,579
	9							$1,594	$1,591	$1,589
	10								$1,600	$1,599
	11									$1,608
N	1	4.91	4.37	4.01	3.75	3.55	3.39	3.25	3.14	3.04
	2	1.67	1.41	1.26	1.15	1.08	1.03	0.99	0.96	0.94
	3	1.42	1.20	1.06	0.98	0.92	0.87	0.84	0.81	0.80
	4		1.02	0.90	0.83	0.78	0.74	0.71	0.69	0.67
	5			0.76	0.70	0.66	0.62	0.60	0.58	0.57
	6				0.59	0.55	0.53	0.51	0.49	0.48
	7					0.47	0.45	0.43	0.42	0.41
	8						0.38	0.36	0.35	0.34
	9							0.31	0.30	0.29
	10								0.25	0.25
	11									0.21
M	1	1.13	0.96	0.84	0.75	0.68	0.61	0.56	0.52	0.48
	2	0.44	0.36	0.31	0.27	0.24	0.22	0.21	0.19	0.18
	3	0.43	0.35	0.30	0.26	0.24	0.22	0.20	0.19	0.18
	4		0.33	0.29	0.25	0.23	0.21	0.20	0.18	0.18
	5			0.27	0.24	0.22	0.20	0.19	0.18	0.17
	6				0.23	0.21	0.19	0.18	0.17	0.16
	7					0.19	0.18	0.17	0.16	0.15
	8						0.17	0.16	0.15	0.14
	9							0.14	0.14	0.13
	10								0.13	0.12
	11									0.11
Profit	1	-$1,316	-$1,376	-$1,367	-$1,331	-$1,281	-$1,225	-$1,166	-$1,107	-$1,049
	2	$719	$554	$459	$400	$362	$335	$317	$304	$296
	3	$596	$453	$372	$322	$289	$266	$251	$240	$233
	4		$369	$299	$256	$227	$208	$195	$186	$180
	5			$237	$200	$176	$159	$148	$141	$136
	6				$153	$132	$118	$109	$102	$98
	7					$95	$83	$75	$70	$67
	8						$54	$47	$43	$40
	9							$24	$20	$17
	10								$1	**-$1**
	11									**-$17**

In the absence of a first period price floor, as shown in Table 3 and Figure 6, the longer health plans intend to live, the more aggressive they will enter the market, *i.e.* P_1 is decreasing in T. This occurs because health plans are willing to "pay" more (in the form of a first period price even lower than marginal cost) for potential consumers, if they expect to extract future rents over a longer life span.

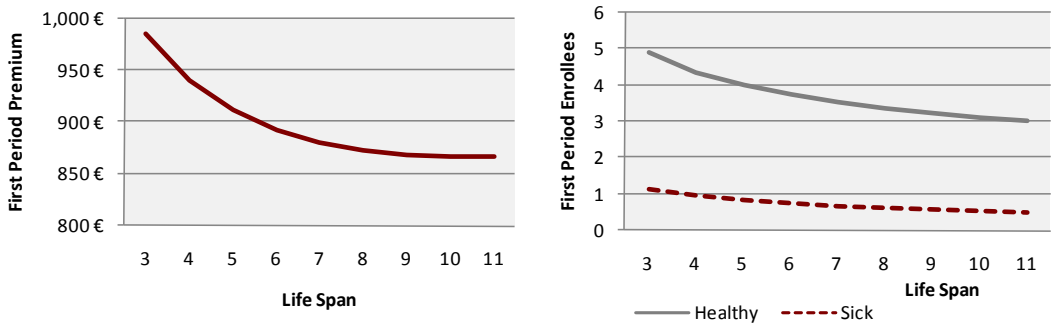


Figure 6: First Period Premium, Number of Healthy and Number of Sick Enrollees for Different Life Spans

As P_1 decreases in T, one might expect to find higher market shares and bigger losses in the first period as T increases. However, that is not the case (Figure 6). As Table 3 shows, the number of (healthy and sick) enrollees in the first year of the plan's life is decreasing in T, and so is the loss incurred in that period. In our model, the life span equals the number of plans in any given period, meaning that longer life spans correspond to a higher degree of competition among existing plans. Although the loss per enrollee in the first period is higher, with more plans in the menu of choices and a constant number of employees, each plan will have less enrollees and lower losses in the first period.

Table 3 and Figure 7 also show that the average premium is increasing in T. For all T, prices increase sharply from period one to period two and only slightly after that, so the increase in average price, as T increases, simply reflects the fact that P_1 weighs less as plans live longer. In fact, the average premium, excluding the first period price is slightly decreasing in T. When plans live longer, less switching will occur especially among involuntary switchers, who are also the ones with the higher switching cost. Table 3 shows that, in our base case, the percentage of switchers decreases from 50% to 27% as the life span increases from three to ten, and that the largest relative reduction occurred among involuntary switchers. The decrease in average

switching cost is also mainly driven by the significant reduction (78%) in switching costs of involuntary switchers.

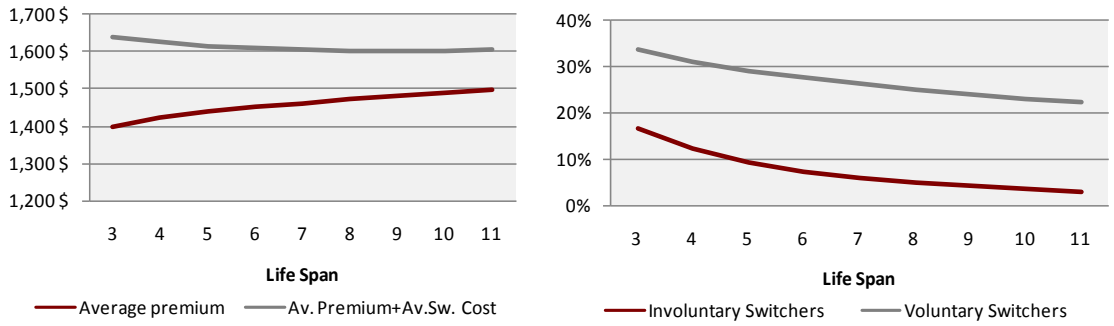


Figure 7: Average Premium, Average Total Cost to Consumers and Percentage of Switchers, for Different Life Spans

While switching costs of voluntary switchers are taken into account by health plans that have to set premiums sufficiently low to compensate for switching costs, switching costs of involuntary switchers do not affect demand. It is true that, once forced to switch, involuntary switchers will select the plan offering the lowest premium but ***all*** involuntary switchers will select the youngest plan, even if its premium is just slightly below the cheapest of the remaining premiums.

With average premiums increasing and average switching costs decreasing as plans live longer, it is not clear what will happen to the total cost borne by consumers (average premium plus average switching cost). Our simulations show that the total cost has a minimum at some $\tilde{T} < T$, where T is the steady-state equilibrium life span (Table 3).

2.6.4. Base Case Competitive Equilibrium Simulation Results

Table 4 presents the steady-state competitive equilibria of our model. There are four blocks (premium, profit, healthy and sick) with four columns each. Each column shows the simulation results for one premium payment system. Our base case competitive equilibrium results, discussed in this Section, are presented in column "BC" in each block. Comparison of the results for the different premium payment systems is performed in 2.6.5.

Table 4: Equilibrium Solution under Different Premium Payment Systems

	Premium				Profit				Nt				Mt			
Year	BC	CS	RA	PF	BC	CS	RA	PF	BC	CS	RA	PF	BC	CS	RA	PF
1	$867	$823	$867	$1,155	-$1,107	-$1,276	-$1,123	$0	3.14	3.16	3.13	2.23	0.519	0.523	0.515	0.289
2	$1,516	$1,544	$1,518	$1,647	$304	$332	$299	$484	0.96	0.95	0.96	0.99	0.191	0.189	0.189	0.160
3	$1,528	$1,556	$1,528	$1,659	$240	$265	$238	$401	0.81	0.81	0.81	0.85	0.190	0.188	0.188	0.165
4	$1,539	$1,567	$1,538	$1,671	$186	$209	$186	$330	0.69	0.68	0.69	0.72	0.185	0.183	0.184	0.166
5	$1,550	$1,578	$1,548	$1,682	$141	$161	$143	$269	0.58	0.58	0.58	0.62	0.177	0.177	0.177	0.163
6	$1,561	$1,589	$1,558	$1,693	$102	$121	$105	$217	0.49	0.49	0.50	0.53	0.168	0.168	0.169	0.159
7	$1,571	$1,599	$1,567	$1,704	$70	$86	$74	$172	0.42	0.42	0.42	0.45	0.158	0.159	0.160	0.152
8	$1,581	$1,609	$1,576	$1,714	$43	$57	$47	$134	0.35	0.35	0.36	0.39	0.148	0.148	0.149	0.145
9	$1,591	$1,619	$1,585	$1,724	$20	$33	$24	$101	0.30	0.30	0.30	0.33	0.137	0.138	0.139	0.137
10	$1,600	$1,628	$1,593	$1,734	$1	$12	$5	$73	0.25	0.25	0.26	0.28	0.126	0.128	0.129	0.128
11				$1,743				$48				0.24				0.120
12				$1,753				$28				0.21				0.111
13				$1,762				$10				0.18				0.103

Note: BC: Base Case; CS: Cost Sharing; RA: Risk Adjustment; PF: Price Floor; Nt: Number of healthy employees at time *t*; Mt: Number of sick employees at time *t*

For any given life span, the optimal price sequence is always increasing and the correspondent sequence of profits is strictly decreasing, which confirms our analytical results presented earlier. More specifically, prices are below average cost in the first period, increase abruptly from the first to the second period and only slightly after that. Health plans have losses in the first period in order to establish their market share. Then, at vintage two, plans have their best opportunity to recover losses incurred in the previous period, because that is when the price differential with respect to the youngest plan is lowest and fewest enrollees have become sick.

As plans become older, healthy enrollees become sick raising plan costs and prices over time. With higher costs, not only are plans incapable of attracting (healthy) incoming employees but they are also unable to hold their enrollees with lower switching costs. It is thus predictable, and Table 4 confirms it, that the number of healthy enrollees (N_t) decreases as the plan grows older.

As to the evolution of the number of sick enrollees with time, the results are not so obvious. On one hand, two forces drive the number of sick enrollees down. First, employees leave the firm at a rate d; Second, as the price differential between their current premium and the premium charged by the youngest plan increases, more and more sick enrollees will find it profitable to switch. On the other hand, healthy enrollees become sick, thus increasing the number of sick enrollees and it is not clear which force will *a priori* prevail. However, in all our simulations (even with s much higher than d) the price spiral was steep enough to imply a decrease in the number of sick employees over time.

Having summarized the main features, common to all simulations carried out and before we proceed to the comparison of the steady-state equilibria of the four premium payment system considered, it is interesting to do some comparative statics. As described in Section 2.4, with no closed-form solution to our model, comparative statics is performed locally. Table 5 shows the

impact of a 0.1 increase in k and l, a 10% increase in W,C_N and C_N on the equilibrium value of our dependent variables, the price sequence and first period market shares, for our base case. In performing such analysis, we have kept constant the life span. This implies that we are possibly comparing optimal solutions to non-optimal ones (accounting for the change in the optimal life span is performed in Section 2.7). Because the model is well-behaved, interpretation of the results presented in Table 5 shares much in common to the discussion presented in Section 2.7. Consequently, in order to avoid repetition we refer the reader to that Section.

Table 5: Analysis of the Impact of the Parameters on the Equilibrium Values

	Δk = - 0,1	Δl = - 0,1	ΔW = +10%	ΔCN = +10% & ΔCM = - 10%
Ntotal	4.61	4.46	7.87	7.87
Mtotal	5.39	0.54	2.13	2.13
ΔN1	1.16	2.12	0.00	0.00
ΔM1	-1.60	1.10	0.00	0.00
ΔP1	-$576	-$10	-$0.44	$0.60
ΔP2	-$536	$81	$0.28	$0.63
ΔP3	-$570	$79	$0.28	$0.61
ΔP4	-$603	$76	$0.28	$0.58
ΔP5	-$634	$74	$0.28	$0.56
ΔP6	-$663	$72	$0.28	$0.54
ΔP7	-$689	$70	$0.28	$0.52
ΔP8	-$714	$68	$0.28	$0.50
ΔP9	-$738	$66	$0.28	$0.48
ΔP10	-$760	$64	$0.28	$0.46

2.6.5. Competitive Equilibrium Simulation Results Under Different Premium Payment Systems

Table 4 presents the steady-state competitive equilibria of our model, under each of the four premium payment systems considered, while Figure 8 graphs the correspondent price sequences. The first column refers to our base case, in which consumers pay the full premium; the second describes the equilibrium with costs sharing, assuming that employees pay ninety percent of the premium; column three provides the steady-state values under risk adjustment; and finally, the last column shows the results in the presence of a first period price floor yielding zero profit in the first period.

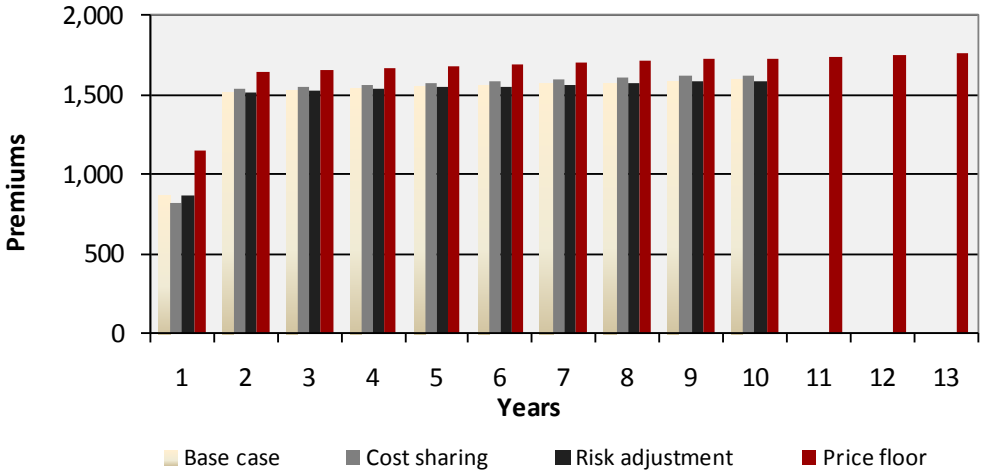


Figure 8: Optimal Price Sequences under Different Payment Systems

Column two of each block in Table 4 provides the steady-state equilibrium values of our variables of interest, in the context of premium cost sharing. To make the results comparable to the risk adjustment scenario, we chose to model the very modest change of imposing only ten percent premium cost sharing. That is, whereas in our base case model we have the consumer paying one hundred percent of the difference between the lowest cost plan and all other plans, here we have the consumer paying ninety percent of this difference. The employer is sharing the cost of the premiums by paying ten percent of the difference. As discussed above, ten percent premium cost sharing is equivalent in our framework to having switching costs that are approximately ten percent higher, and hence, we conducted the simulation by assuming $W = \$1{,}100$.

In our model, switching costs are the reason for plans to have market power. Higher switching costs mean that health plans are able to extract higher rents from their enrollees. The results in column two of each block of Table 4 reflect precisely that. If plans have more market power, they charge higher premiums. This ability to charge higher premiums in the future enhances first period competition and lowers the first period premium, which in turn attracts more healthy employees and implies bigger losses. The difference in premiums with respect to the first, $P_t - P_1$, is higher in all periods than in the base case. This causes much more switching than in the base case and, as a result, the cost of voluntary switchers is substantially higher with cost sharing (Table 6). The consequences of cost sharing, just described, prevailed in all simulations performed[18] indicating that, in the presence of switching costs, cost sharing as a

[18] In simulations performed with a lower proportion of healthy to sick, cost sharing did increase the optimal life span of the plan, but still at higher average premium plus average switching cost.

premium payment system enhances death spirals and total costs borne by consumers, with no additional increase in plans' profits.

As discussed in our model section, risk adjustment in our framework is equivalent to considering moving both C_N and C_M towards the population average ($C = 1{,}200$). We chose to reflect the empirical observation that risk adjustment appears to only move payments about ten percent of the distance toward perfectly reflecting costs. Consequently, the simulation results presented in column three of each block of Table 4 were obtained assuming the cost of healthy and sick people as $C_N = 0.9 * \$1{,}000 + 0.1 * \$1{,}200 = \$1{,}020$ and $C_M = 0.9 * \$2{,}000 + 0.1 * \$1{,}200 = \$1{,}980$, respectively.

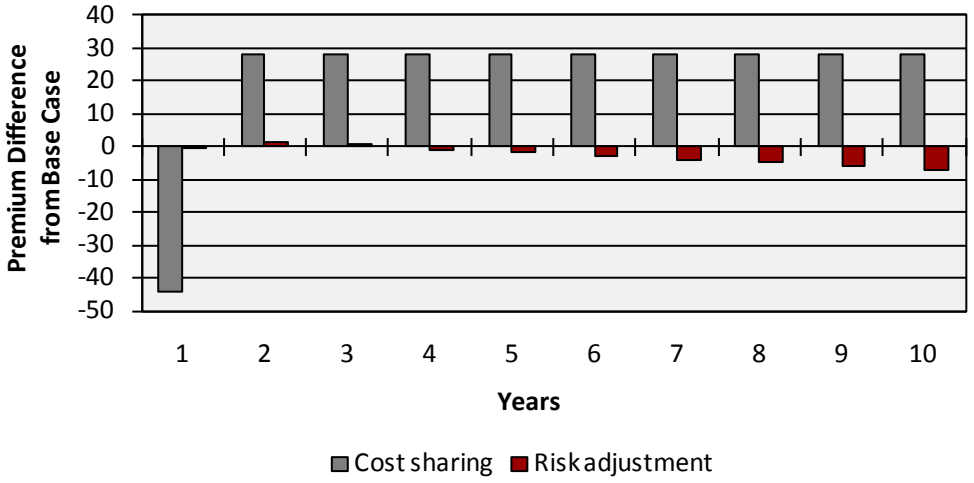


Figure 9: Premium Difference Cost Sharing and Risk Adjustment

Figure 9 compares price paths under risk adjustment and cost sharing with our base case. This figure shows that risk adjustment helps health plans by improving their resistance to death spirals. Health plans charge a first period premium only slightly lower than in our base case but prices increase less after the first period, implying a lower average premium for the same life span. With risk adjustment, there is less switching and the average switching cost is lower, as shown in Figure 8 and Table 6.

Although the results presented do not reflect a significant impact of risk adjustment, it should be said that, qualitatively, the effects described were present in all simulations performed. Moreover, the magnitude of these effects increases as the proportion of sick employees in the total population increases. Indeed, with a higher prevalence of illness among employees, risk adjustment allows plans to live longer than in the base case.

Even when risk adjustment increases the optimal life span, the sum of average premium plus average switching cost is lower than in our base case, while intertemporal profits are equally null (Table 6). Thus, our simulations indicate that risk adjustment is welfare improving.

The last premium payment system considered in our simulations is a price floor. Our discussion in Section 2.2 seems to indicate that death spirals are less frequent when a price floor is established by the employer. As our model exemplifies, health plans enter the market with an extremely aggressive first period price, in the prospect of making use of their future market power, over their enrollees, to extract rents. In the following periods, faced with new health plans entering equally aggressively, the health plans' best strategy is to "*take advantage while they can*", *i.e.* rapidly increase premiums to make up for first period losses, while able to retain some enrollees. The abrupt premium increase culminates in death spirals with huge switching costs to consumers. The rationale behind the price floor premium payment system is to curtail first period competition, in the hope of eliminating the need to drastically increase premiums in the following periods.

The final column in each block of Table 4 presents the steady-state equilibrium values under a price floor premium payment system. The increase in prices is substantially lower than in all other systems considered. Not only do premiums, from the first to the second period, increase substantially less in the second period but plans also choose to live three additional years. Consequently, the proportion of switchers is almost half that of the base case and average switching costs decreases from $111 to $56 (Table 6).

Table 6: Comparing Switchers, Switching Costs, Premiums and Profits for Different Payment Systems

	Base Case	Cost Sharing	Risk Adjustment	Price Floor
Life Span	10	10	10	13
Inv. Healthy Switchers	0.22	0.22	0.22	0.15
Inv. Sick Switchers	0.12	0.12	0.12	0.10
Vol. Healthy Switchers	1.92	1.94	1.90	1.07
Vol. Sick Switchers	0.40	0.40	0.39	0.19
% Inv. Switchers	3.4%	3.4%	3.5%	2.5%
% Vol. Switchers	23.2%	23.4%	23.0%	12.7%
Cost Inv. Healthy Switchers	$191	$201	$194	$124
Cost Inv. Sick Switchers	$104	$109	$106	$78
Cost Vol. Healthy Switchers	$672	$748	$663	$297
Cost Vol. Sick Switchers	$146	$161	$143	$58
Average Switching Cost (ASWC)	**$111**	**$122**	**$110**	**$56**
Average Premium (AP)	**$1,490**	**$1,511**	**$1,488**	**$1,665**
ASWC+AP	**$1,601**	**$1,633**	**$1,598**	**$1,720**
Sum of Profits	**$0**	**$0**	**$0**	**$2,265**

In Table 6, it is possible to compare average switching costs and average premiums for the four premium payment systems considered. Indeed, average premiums plus average switching costs are $118 per person/year higher with a price floor than in our base case, but profits are also substantially higher. While in all other scenarios the sum of discounted profits over the life span of the plan was zero, in the price floor case health plans have an overall profit of $2.265, *i.e.* a profit of $226.5 per employee. This means that there is margin for Pareto movements, when comparing our base case (or, actually, any of the other three premium payment systems) with the price floor scenario.

Figure 10 illustrates the evolution of the ratio of healthy to sick enrollees over time, in each of the four scenarios. The horizontal line at four identifies the population average ratio for healthy and sick employees. Points above that line reflect a favorable selection of enrollees, while points below indicate adverse selection. Under all four of the scenarios, new entrants attract favorable selections. In all but the price floor scenario, health plans have a favorable selection for only three periods. Both cost sharing and risk adjustment yield a more favorable (or less adverse) ratio than the base case in all periods, but the magnitude is bigger in the second (the difference is decreasing in time). Although risk adjustment somewhat slows down switching, it is not enough to maintain a favorable selection for more periods than cost sharing or the base case. The price floor scenario has the slowest rate of deterioration, with enrollments being above the average for four periods. After that, the ratio of healthy to sick falls below the population average but remains less adverse than in the other three cases.

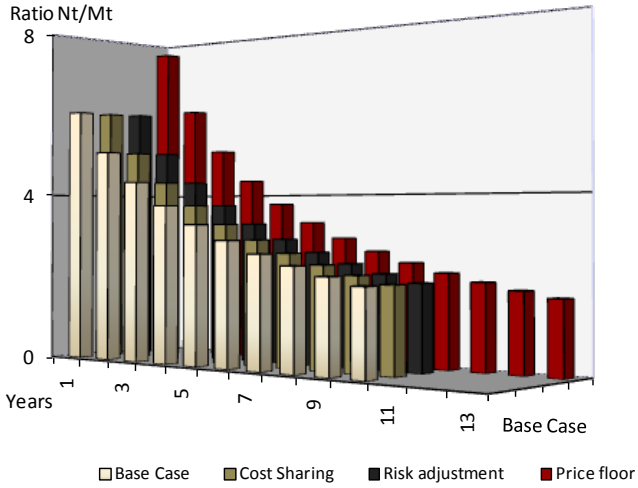


Figure 10: Ratio of Healthy to Sick for Four Premium Payment Systems

2.7. Sensitivity Analysis

In Table 7 to Table 12, we provide the results of the sensitivity analysis. In each table, we allow one parameter to vary and keep the remaining parameters at their baseline values. In each table, we consider two alternative values for the parameter under analysis: one below the baseline value and one above it. For comparison, all tables include the baseline itself.

Different parameter values result not only in different equilibrium values for $\{N_t\}$, $\{M_t\}$, $\{P_t\}$ and $\{\Pi_t\}$ but also in distinct optimal life spans. In Table 7, we provide an example of how the optimal life span is obtained in each sensitivity analysis table. This is done by solving the model for a sequence of possible life spans (three to seven years are shown in that table). Obviously, not all life spans considered are optimal life spans, in the sense that a profitable deviation might exist, but each column shows the optimal price sequence, given the life span. For example, as shown in Table 7, for a fixed cost of \$500, a life span of 6 years yields a negative profit in the last year $(-69\$)$, which indicates that the optimal lifespan is 5 years.

Table 7: Equilibrium Results under Different Fixed Costs Values

Life Span		3			4			5			6			7		
FC	Period	0	100	500	0	100	500	0	100	500	0	100	500	0	100	500
	1	$956	$986	$1,110	$901	$941	$1,111	$862	$912	$1,131	$833	$893	$1,164	$810	$880	$1,207
	2	$1,583	$1,598	$1,659	$1,552	$1,572	$1,656	$1,530	$1,554	$1,662	$1,512	$1,541	$1,674	$1,498	$1,532	$1,692
	3	$1,594	$1,609	$1,670	$1,563	$1,583	$1,667	$1,541	$1,565	$1,673	$1,523	$1,552	$1,686	$1,509	$1,543	$1,703
P	4				$1,574	$1,594	$1,678	$1,552	$1,576	$1,684	$1,534	$1,563	$1,697	$1,520	$1,554	$1,714
	5							$1,562	$1,587	$1,695	$1,545	$1,574	$1,707	$1,531	$1,565	$1,725
	6										$1,555	$1,584	$1,718	$1,541	$1,575	$1,736
	7													$1,551	$1,585	$1,746
	1	4.99	4.91	4.62	4.49	4.37	3.94	4.17	4.01	3.45	3.94	3.75	3.07	3.78	3.55	2.76
	2	1.63	1.67	1.82	1.37	1.41	1.57	1.21	1.26	1.42	1.11	1.15	1.32	1.03	1.08	1.25
	3	1.38	1.42	1.56	1.16	1.20	1.34	1.02	1.06	1.21	0.93	0.98	1.13	0.87	0.92	1.07
N	4				0.98	1.02	1.15	0.87	0.90	1.03	0.79	0.83	0.96	0.73	0.78	0.91
	5							0.73	0.76	0.88	0.67	0.70	0.82	0.62	0.66	0.78
	6										0.56	0.59	0.70	0.52	0.55	0.67
	7													0.44	0.47	0.57
	1	1.15	1.13	1.05	0.99	0.96	0.85	0.88	0.84	0.70	0.80	0.75	0.59	0.73	0.68	0.50
	2	0.43	0.44	0.48	0.35	0.36	0.39	0.30	0.31	0.34	0.26	0.27	0.30	0.24	0.24	0.27
	3	0.42	0.43	0.47	0.34	0.35	0.38	0.29	0.30	0.33	0.25	0.26	0.29	0.23	0.24	0.26
M	4				0.32	0.33	0.37	0.28	0.29	0.32	0.24	0.25	0.29	0.22	0.23	0.26
	5							0.26	0.27	0.31	0.23	0.24	0.27	0.21	0.22	0.25
	6										0.22	0.23	0.26	0.20	0.21	0.24
	7													0.18	0.19	0.23
	1	-$1,421	-$1,316	-$928	-$1,532	-$1,376	-$818	-$1,575	-$1,367	-$658	-$1,589	-$1,331	-$487	-$1,589	-$1,281	-$320
	2	$769	$719	$539	$600	$554	$395	$502	$459	$324	$439	$400	$291	$396	$362	$279
	3	$652	$596	$390	$506	$453	$267	$422	$372	$207	$368	$322	$179	$331	$289	$171
Profit	4				$427	$369	$157	$354	$299	$106	$308	$256	$84	$277	$227	$77
	5							$298	$237	$20	$258	$200	$1	$231	$176	-$3
	6										$216	$153	-$69	$193	$132	-$72
	7													$161	$95	-$131
%IS		16%	17%	18%	12%	12%	14%	9%	9%	11%	7%	7%	9%	6%	6%	7%
%VS		35%	34%	29%	33%	31%	24%	32%	29%	21%	30%	28%	18%	30%	26%	15%
AVSWC		$245	$240	$222	$209	$202	$175	$185	$176	$142	$169	$157	$116	$156	$142	$96

Note: FC=fixed cost, N=Healthy enrollees, M=Sick enrollees, IS=Involuntary switchers, VS=voluntary switcher, AVSWC=average switching cost

The optimal life span is highly sensitive to the value of fixed costs (Table 7). If no fixed costs are present, plans tend to live "forever" with minimal enrollments. But with fixed costs as low as one hundred (ten percent of treatment cost of a healthy enrollee), plans die after ten periods.

Overall, the model is "well-behaved" in that if an increase in a given parameter yields a decrease in one independent variable, it does so regardless of the parameters' chosen value or of how much it increases. To get a feel for how to interpret these tables, we describe the results from Table 8 in detail, and then focus on only the interesting features of the remaining tables.

Table 8: Equilibrium Results under Different Probabilities of Becoming Sick

s	0.01				0.03				0.1			
Ntotal	9				8				5			
Mtotal	1				2				5			
Life Span	T*=11				T*=10				T*=7			
	Prem.	Profit	Nt/ Ntotal	Mt/ MTotal	Prem.	Profit	Nt/ Ntotal	Mt/ MTotal	Prem.	Profit	Nt/ Ntotal	Mt/ MTotal
t=1	$750	-$1,146	358	238	$876	-$1,107	398	268	$1,182	-$1,095	508	388
t=2	$1,408	$263	118	88	$1,516	$304	128	108	$1,827	$401	148	148
t=3	$1,412	$218	98	88	$1,528	$240	108	108	$1,857	$280	108	138
t=4	$1,417	$179	88	88	$1,539	$186	98	98	$1,883	$189	88	128
t=5	$1,421	$144	78	88	$1,550	$141	78	98	$1,906	$122	68	118
t=6	$1,426	$114	68	88	$1,561	$102	68	98	$1,926	$71	48	108
t=7	$1,430	$87	68	78	$1,571	$70	58	88	$1,945	$33	38	88
t=8	$1,435	$64	58	78	$1,581	$43	48	88				
t=9	$1,439	$43	48	78	$1,591	$20	48	78				
t=10	$1,444	$25	48	78	$1,600	$1	38	78				
t=11	$1,448	$10	38	68								
	Voluntary		Involuntary		Voluntary		Involuntary		Voluntary		Involuntary	
Switchers	21%		3%		23%		3%		29%		5%	
Sw. Costs	$718		$277		$818		$295		$1,076		$443	
Average Sw. Costs	$100				$111				$152			

Table 8 provides the solution to our model, assuming $\{l, W, C_N, C_M, CF, \rho\} = \{0.9, 1000, 1000, 2000, 100, 1\}$ and considering three possible values for s, $\{0.01, 0.03, 0.1\}$. Each of the three blocks in Table 8 provides the solution to the model under one value of s. The second block, where $s = 0.03$, corresponds to our baseline. Each block has four columns providing the results for the four equilibrium series: $\{N_t\}$, $\{M_t\}$, $\{P_t\}$ and $\{\Pi_t\}$. Because varying s and d alters the (steady-state) total number of healthy (N) and sick employees (M) in the firm, Table 8 and Table 9 include two extra rows with the corresponding N and M. The third column, where $s = 0.1$, is of especial interest because it yields an equal number of healthy and sick employees in the firm, which allows us to isolate the effect of biased arrival *per se*.

For a given life span, an increase in s is associated with steeper spirals (defined as a higher difference between the price in the first period and prices in subsequent periods)[19] and higher percentage of voluntary switchers. Nevertheless, because varying s, changes the ratio of healthy to sick employees, an increase in s changes the optimal life span and results in a mixed effect on death spirals steepness. Note that, for the optimal life span associated with each value of s (as shown in Table 8), the percentage increase in price from the first to the second period is decreasing in s while the percent increase in price in subsequent periods is increasing in s. The first effect results from the fact that a market where individuals become sick at a lower rate is a more attractive one and plans are thus more willing to forgo profits in the first period. The second effect is a consequence of the fact that with a higher proportion of sick enrollees average cost increases faster. The overall effect results in a lower percentage of voluntary

[19] Results not shown.

switching for more favorable healthy to sick ratio (lower s). Because a lower probability of becoming sick allows plans to live longer (the optimal life span is decreasing in s) the percentage of involuntary switchers also decreases and, as a result, average switching costs ($AVSWC$) are lower for low values of s.

Table 9: Equilibrium Results under Different Probabilities of Leaving the Firm

d	0.05				0.1				0.15			
Ntotal	13				8				6			
Mtotal	7				2				1			
Life Span	T*=15				T*=10				T*=7			
	Prem.	Profit	Nt/ Ntotal	Mt/ MTotal	Prem.	Profit	Nt/ Ntotal	Mt/ MTotal	Prem.	Profit	Nt/ Ntotal	Mt/ MTotal
t=1	$892	-$2,255	328	228	$867	-$1,107	398	268	$883	-$730	468	318
t=2	$1,589	$406	98	78	$1,516	$304	128	108	$1,502	$264	158	138
t=3	$1,599	$350	88	78	$1,528	$240	108	98	$1,514	$190	128	138
t=4	$1,608	$300	78	78	$1,539	$186	98	98	$1,526	$131	98	128
t=5	$1,618	$255	68	68	$1,550	$141	78	98	$1,538	$84	78	118
t=6	$1,627	$215	68	68	$1,561	$102	68	88	$1,549	$46	68	108
t=7	$1,636	$180	58	68	$1,571	$70	58	88	$1,560	$16	58	98
t=8	$1,644	$148	58	68	$1,581	$43	48	78				
t=9	$1,653	$120	48	58	$1,591	$20	48	78				
t=10	$1,661	$95	48	58	$1,600	$1	38	68				
t=11	$1,669	$73	38	58								
t=12	$1,677	$53	38	58								
t=13	$1,684	$36	38	48								
t=14	$1,691	$20	28	48								
t=15	$1,698	$6	28	48								
	Voluntary		Involuntary		Voluntary		Involuntary		Voluntary		Involuntary	
Switchers	21%		3%		23%		3%		25%		5%	
Sw. Costs	$1,658		$469		$818		$295		$540		$256	
Average Sw. Costs	$106				$111				$119			

In Table 9, we see that plans enter more aggressively (the entering price is lower) as d increases, which might seem unlikely. Why would plans be more willing to incur losses if they are less likely to be able to retain their enrollees? The reason being that the ratio of healthy to sick also changes; in particular the ratio is increasing with d. At $d = 0.15$ the ratio is six, making this market a quite attractive one. This "healthier market" implies that, even at a lower first period price, losses incurred in the first period are lower than losses incurred when entering at a higher price with a less attractive ration of healthy to sick. Obviously, plans live longer when d is low.

With respect to Table 11 and Table 12, it is worth mentioning that as health costs of healthy and sick become closer, spirals become less steep ($P_T - P_1$ is smaller). There is also less switching and lower average switching cost.

Table 10: Equilibrium Results under Different maximum Switching Costs Values

W	$400				$1,000				$1,500			
Life Span	T*=5				T*=10				T*=12			
	Prem.	Profit	Nt/ Ntotal	Mt/ MTotal	Prem.	Profit	Nt/ Ntotal	Mt/ MTotal	Prem.	Profit	Nt/ Ntotal	Mt/ MTotal
t=1	$1,116	-$371	498	418	$867	-$1,107	398	268	$639	-$1,827	388	238
t=2	$1,356	$171	178	178	$1,516	$304	128	108	$1,647	$410	118	88
t=3	$1,367	$111	148	168	$1,528	$240	108	98	$1,659	$338	98	88
t=4	$1,378	$63	118	148	$1,539	$186	98	98	$1,670	$275	88	88
t=5	$1,388	$25	98	138	$1,550	$141	78	98	$1,682	$222	78	88
t=6					$1,561	$102	68	88	$1,692	$176	68	88
t=7					$1,571	$70	58	88	$1,703	$137	58	78
t=8					$1,581	$43	48	78	$1,713	$104	48	78
t=9					$1,591	$20	48	78	$1,723	$75	48	68
t=10					$1,600	$1	38	68	$1,733	$50	38	68
t=11									$1,742	$29	38	68
t=12									$1,751	$11	28	58
	Voluntary		Involuntary		Voluntary		Involuntary		Voluntary		Involuntary	
Switchers	28%		9%		23%		3%		23%		3%	
Sw. Costs	$378		$298		$818		$295		$1,208		$335	
Average Sw. Costs	$68				$111				$154			

Table 10 shows the solution under three possible values for the maximum switching cost (W). As clearly shown in the table, price escalation is highly associated with the presence of switching costs. As expected, for higher values of W, plans are more willing to invest in the first period because higher switching costs will allow them to compensate for initial losses, through higher future prices. Moreover, plans will not only charge higher prices but also survive longer, which, again, creates an incentive for aggressive entrance.

Table 11: Equilibrium Results under Different Cost of Healthy Enrollees

CN	$100				$1,000				$1,500			
Life Span	T*=8				T*=10				T*=11			
	Prem.	Profit	Nt/ Ntotal	Mt/ MTotal	Prem.	Profit	Nt/ Ntotal	Mt/ MTotal	Prem.	Profit	Nt/ Ntotal	Mt/ MTotal
t=1	$154	-$1,106	448	328	$867	-$1,107	398	268	$1,268	-$1,131	378	238
t=2	$798	$371	148	128	$1,516	$304	128	108	$1,923	$268	118	88
t=3	$819	$274	118	118	$1,528	$240	108	98	$1,929	$220	108	98
t=4	$840	$196	98	118	$1,539	$186	98	98	$1,935	$178	88	98
t=5	$861	$133	78	108	$155	$141	78	98	$1,940	$142	78	88
t=6	$881	$82	68	98	$1,561	$102	68	88	$1,946	$111	68	88
t=7	$900	$41	58	88	$1,571	$70	58	88	$1,951	$84	58	88
t=8	$919	$9	48	78	$1,581	$43	48	78	$1,956	$60	58	78
t=9					$1,591	$20	48	78	$1,961	$40	48	78
t=10					$1,600	$1	38	68	$1,966	$22	38	78
t=11									$1,970	$6	38	68
	Voluntary		Involuntary		Voluntary		Involuntary		Voluntary		Involuntary	
Switchers	27%		4%		23%		3%		21%		3%	
Sw. Costs	$994		$369		$818		$295		$728		$274	
Average Sw. Costs	$136				$111				$100			

Table 12: Equilibrium Results under Different Cost of Sick Enrollees

CM	$1,100				$2,000				$4,000			
Life Span	T*=12				T*=10				T*=7			
	Prem.	Profit	Nt/ Ntotal	Mt/ MTotal	Prem.	Profit	Nt/ Ntotal	Mt/ MTotal	Prem.	Profit	Nt/ Ntotal	Mt/ MTotal
t=1	$691	-$1,149	368	78	$867	-$1,107	398	268	$1,305	-$921	468	368
t=2	$1,353	$236	118	88	$1,516	$304	128	108	$1,934	$450	158	148
t=3	$1,354	$201	98	88	$1,528	$240	108	98	$1,968	$309	128	138
t=4	$1,356	$169	88	88	$1,539	$186	98	98	$2,000	$200	98	118
t=5	$1,357	$140	78	88	$1,550	$141	78	98	$2,032	$116	78	108
t=6	$1,358	$115	68	88	$1,561	$102	68	88	$1,064	$53	68	98
t=7	$1,359	$92	58	88	$1,571	$70	58	88	$1,094	$5	48	78
t=8	$1,360	$72	58	78	$1,581	$43	48	78				
t=9	$1,361	$54	48	78	$1,591	$20	48	78				
t=10	$1,362	$38	48	78	$1,600	$1	38	68				
t=11	$1,363	$23	38	68								
t=12	$1,364	$10	38	68								
	Voluntary		Involuntary		Voluntary		Involuntary		Voluntary		Involuntary	
Switchers	20%		3%		23%		3%		30%		4%	
Sw. Costs	$656		$253		$818		$295		$1,113		$400	
Average Sw. Costs	$91				$111				$100			

Table 11 and Table 12 provide the results of the sensitivity analysis to the cost of healthy and sick enrollees. Higher costs of healthy enrollees, while maintaining the cost of sick enrollees at its base value, reduces the cost disadvantage between incumbent and new entrants, thereby allowing plans to live longer. By the same token, reducing the cost of sick enrollees allows plans to live longer.

As expected, for the same life span, if discounting is included (not shown) and the future is worth less, plans are less willing to incur losses in the first period and will set a higher first period price. This will, in turn, generate lower first period market shares, lower profits and lower switching.

2.8. Discussion

This paper has examined the implications of a premium payment system in a model in which identical health plans attract different mixes of healthy and sick enrollees, according to how long a health plan has been offered by an employer. We first reviewed some empirical evidence that highlights that some insurance organizations, namely the Blue Cross Blue Shield plans, have been regularly exiting and entering with different health plan options, as our theoretical model would predict.

The experience at Boston University, Minnesota and California also differs in that the highest rates of exit and entry and greatest dispersion of pricing have occurred with partial premium cost sharing, and the lowest rates have occurred with premium price floors. Clearly, there are a

great many other changes taking place in these markets for health plans, however our model provides a rationale for this empirical evidence and our simulations confirm it.

Our model shows that, in the presence of switching costs, cost sharing is potentially harmful because it increases plans' market power.

In our simulations, we show that, by contradicting the forces driving prices up, the increase in costs due a deterioration of the mix of enrollees, risk adjustment creates the right incentives for plans not to follow a death spiral pattern. Slowing down death spirals, risk adjustment also decreases switching costs, thus increasing consumers' welfare with no impact on profits.

Price floors were found to be extremely effective at reducing switching costs. They are also found to increase average premiums and profits but, at least in our simulations, the increase in profits is higher than the increase in total costs to consumers, implying a margin for welfare improvement. We are perfectly aware of the fact that our simple model may not capture many relevant aspects and that simulations are simulations. However, the possibility for price floors to be welfare improving exists, and it is our hope that this paper will stimulate further research on the subject.

Such research will shed light on the full policy implications of our findings. The "*churning*" phenomenon, widely referred to in the health policy literature [1, 2], has led some States in the United States of America to administratively establish an upper limit to price increases in health plan premiums. More often than not, such policy has led the insurance companies to anticipate product removal from the market, thereby generating elevated switching costs to consumers. If, as suggested by our model, price floors are found to be more effective than current approaches aimed at minimizing the impact of *"churning"* and death spirals, than such option should be tested and implemented.

What is more, lessons to be learned from such analysis are relevant to other countries, such as Portugal, where the employer-provided private insurance market is growing [33]. Because death spirals are a consequence of how contracts are written between health plan providers and employers, awareness of both the consequences and the mechanisms to minimize such consequences, will be useful in setting the appropriate grounds on which the employer-provided private insurance market is to grow.

The model we use is limited in a number of ways that could be relaxed and studied in future research. We mention a few that seem particularly important to us. The first, which we have

mentioned previously, is that it assumes an equal distribution of switching costs for healthy and sick enrollees. One would expect that a consumer who visits his physician quite often to be less willing to change provider than someone who almost never makes use of his insurance plan. Second, we focus on the simple case in which there are only two types of consumers, healthy and sick, whereas in the real world there are a continuum of types, with different demand responsiveness and treatment costs. Health care spending is stochastic and, for this reason, consumers may need to make choices prior to knowing their health costs needs. We focus on the case in which plans are all identical, and the only switching cost is a fixed cost for changing plans. An alternative framework, that might be interesting to contemplate, is one that is closer to a matching model, in which consumers derive utility each period that depends on the quality of their match with a health plan. Finally, there is no moral hazard problem in our paper, and plans do not explicitly adopt strategies to affect health plan choices other than through pricing. If service distortion and/or explicit dumping are permitted, it would clearly change the optimal pricing, entry, and exit decisions.

3. Price Regulation and Product Survival in the Portuguese Pharmaceutical Market

Price Regulation and Product Survival

The Portuguese Pharmaceutical Market

Filipa Aragão*

ABSTRACT

As in many European countries, price regulation of pharmaceuticals is one of the cost control measures adopted by the Portuguese government. When such regulation decreases the products' real price over time it may create an incentive for product turnover.

Using panel data for the period of 1997 through 2003 on drug packages sold in Portuguese pharmacies, this empirical analysis addresses the question of whether price control policies create an incentive for product withdrawal.

Our work builds the product survival literature by accounting for unobservable product characteristics and heterogeneity among consumers when constructing quality, price control and competition indexes. These indexes are then used as covariates in a Cox proportional hazard model.

We find that indeed price control measures increase the probability of exit and that such effect is not verified in OTC market where no such price regulation measures exist. We also find quality to have a significant positive impact on product survival.

JEL Classification: I18, I11, C41

Keywords: Regulation, demand, product differentiation, discrete choice, survival analysis, competition, pharmaceuticals

*Support for this research has been provided by Fundação para a Ciência e Tecnologia, Programa POCTI - Formar e Qualificar - Medida 1.1., grant Praxis XXI/BD/19954/99

We are thankful to Céu Mateus, Pedro Pita Barros and seminar/conference participants at Escola Nacional de Saúde Pública, IX Conferência Nacional de Saúde Pública and XIII Jornadas de Classificação e Análise de Dados for helpful discussion. The opinions expressed are the author's alone.

Abbreviations

Abbreviations	Description
AIM	Autorização de Introdução no Mercado Maketing authorization
BB	Banjari and Benkard
BLP	Berry, Levinsohn and Pakes
INFARMED	Autoridade Nacional do Medicamento e Produtos de Saúde, I.P. National Authority of Medicines and Health Products, IP
NHS	National Health Service
ICD	International Common Denomination
DGAE	Direcção-Geral das Actividades Económicas Directorate General for Economic Activities
NROTC	Over-the-counter, non-reimbursable drugs
PD	Prescription drugs
ROTC	Reimbursable OTC
OTC	Over-the-counter drugs
IO	Industrial Organization
ATC	Anatomical Therapeutic Chemical
ATC3	Level 3 Anatomical Therapeutic Chemical
OECD	Organization for Economic Co-operation and Development
PH	Proportional Hazard
LR	likelihood ratio

3.1. Introduction

In Portugal, health expenditure represented, in 2008, 10.8% of the Gross Domestic Product and pharmaceuticals and other medical non-durable accounted for 21% of the total expenditure on health [35]. In an attempt to reduce the National Health Service (NHS) deficit, several cost containment policies have been implemented over the years.

Oliveira and Pinto [36] and more recently Barros and Nunes [37] have assessed the impact of such measures on pharmaceutical expenditure. The overall conclusion in both papers is that policy measures aimed at controlling pharmaceutical expenditure have, in general, been unsuccessful.

As in many European countries, price regulation of pharmaceuticals is one of the cost control measures adopted by the Portuguese government [38]. Commercialization of pharmaceutical products in Portugal requires a license ("*Autorização de Introdução no Mercado*" (AIM)). This license provides a number that uniquely identifies the product and all its characteristics (packaging, dosing, International Common Denomination (ICD), dosage, therapeutic class and the population to whom it is recommended).

The owner of such license then proposes a price to the *Direcção-Geral das Actividades Económicas* (DGAE)[20] [39]. While the price of over-the-counter, non-reimbursable (NROTC), drugs may be freely set by the firm at the resale level [40], in the case of non-generic prescription drugs (PD) or reimbursable OTC (ROTC), the Government sets a maximum price, based on the average price of the same product in Spain, France, Italy and Greece[21] [41, 42]. Comparison is based on the pharmaceutical form, dosage and packaging format, and these price setting rules apply to both new products and changes in pharmaceutical form and dosage.[22] If the firm applies for reimbursement, the price should be revised by National Authority of Medicines and Health Products, IP (INFARMED), based on both clinical efficacy and the cost-effectiveness evaluation results.

[20] NROTC drug prices although freely set are to be "approved" by the Direcção-Geral Comércio e Concorrência (Portaria 713/2000 de 5 Setembro, DR I Série B, 4693)

[21] This rule was introduced by Decreto-Lei n.º 65/2007 de 14 de Março de 2007. DR I Série n.º 52, de 14 de Março de 2007 - Ministério da Saúde. Before that, the rule was to take the maximum price of 3 reference countries.

[22] Special rules apply for generic drugs. In that case, the price has to be 20% or 35% below the reference product or reference homogenous group price. Also, different sets of rules apply for drugs sold in ambulatory and those sold exclusively through the hospital.

Once the initial price is set, subsequent price increases are rarely allowed and decreases may be mandated [43, 44]. The initial price will remain unchanged for 3 years[23] [42], after which, periodic price variations will be determined by the annual pharmaceutical prices' revision, determined by the Government in office. These annual pharmaceutical prices' revision set the maximum allowed price increase[24] and are often set below the inflation rate [45, 46], which means that the real price decreases over time, as the product ages.

Although both the formula for initial price determination and the formula for annual pharmaceutical price revisions have changed over the years, one essential feature remains: once the initial price is set, the pharmaceutical company loses control over subsequent price variations, that is, all subsequent price increases are exogenous to the firm[25].

One way to regain control over the price is to replace the existing product by a "renewed version" of the same. Although, the new product's price will still be set under the general rules, a price negotiation again takes place. This renewed version may reflect something as simple as a new packaging.

The aim of the present analysis is to take a first step towards understanding whether price regulation of the Portuguese pharmaceutical market creates an incentive for product turnover, defined as the substitution of existing products by slightly different versions of the same product. Our approach is to evaluate whether price control measures have an impact on product survival after controlling for quality and competition.

There is evidence in the literature to suggest that price regulation creates an incentive for product turnover when such regulation decreases the products' real price over time. Danzon and Chao [47] use data from seven countries to examine price competition between generic competitors (different manufacturers of the same compound) and therapeutic substitutes (similar compounds), under different regulatory regimes. The authors find that "*generic competition is ineffective and may be counterproductive in countries with strict price or reimbursement regulation*". In interpreting their findings the authors argue that one plausible explanation is that in regulatory regimens (which decrease the products' real price over time),

[23] The law has recently changed to 1 year (Decreto-Lei 48/A 2010, de 13 de Maio de 2010, DR I Série, n.º 93 de 13 de Maio de 2010 – Ministério da Saúde)

[24] Currently, the review is based on the comparison with the average of the prices in force in the reference countries on 1st January each year (Decreto-Lei 48/A 2010, de 13 de Maio de 2010, DR I Série, n.º 93 de 13 de Maio de 2010 – Ministério da Saúde).

[25] Exceptional price revisions are contemplated in the Law. These may occur due to public interest or by initiative of the Marketing Authorization owner. Any such change must be approved by DGAE, INFARMED I.P, Ministry of Health and Ministry of Economy.

generic equivalents are often me-too products introduced by the manufacturers as a strategy to obtain higher prices.

This issue is relevant because introducing new products, especially in such regulated markets, is expensive and time consuming, both for applying firms and for licensing authorities. Moreover, both the prescribing physician and the consuming patients incur in non-neglectable switching costs. These costs may well exceed potential savings obtained through price control.

Oliveira and Pinto [36] argue that the growth in pharmaceutical expenditure is related, among others, to a shift towards more expensive drugs. Are these, "more expensive drugs", new chemical entities (recovering research and development cost while protected by patent legislation) or "me-too" products"?. In Danzon, Wang &Wang [48], Portugal appears second to last in terms of the number of launches of new chemical entities among major markets in the 90s. Nevertheless, as shown in Table 13, Portugal has a product turnover rate of about 7% a year[26] and as noted by Cardoso [49] the proliferation of me-too drugs was one of the main reasons for authorities to allow, in 2001[27], the transformation of branded drugs into generic drugs.

Using panel data for the period of 1997 through 2003 on drug packages sold in Portuguese pharmacies, this empirical analysis addresses the question of whether price control policies create an incentive for product withdrawal.

Methodologically, our work builds on the product survival literature [50-52] by first estimate unobservable characteristics and preference parameters for both observable and unobservable characteristics using one method commonly applied in the empirical IO literature, namely that developed by Bajari end Benkard (BB) [53], and then using those preference estimates to build an index of product quality. While proxies for unobserved product quality have previously been considered as a determinant of exit in the product survival literature, to our knowledge, the methodology followed had not previously been applied in the product survival literature.

Our approach is as follows: First, we apply the methodology developed by Bajari and Benkard (BB) [53] to estimate structural preferences parameters for both observable and unobservable characteristics. Secondly, we use these preferences parameters' estimates to construct quality, competition and market structure indexes. Thirdly, we perform a (reduced-form) survival analysis for product exit using the constructed indexes, along with a proxy for price control, as

[26] The two-digit figures for new products in 2002 and 2003 are due regulatory changes and consequent massive entrance of generics.

[27] This regulatory change only became effective in October 2003.

covariates. Finally, we compare the determinants of product withdrawal in the prescription and OTC markets.

The structure of the paper is the following: In Section 3.2, we briefly review the two strains of literature in which our work is based. In Section 3.3 we describe the key features of the Portuguese pharmaceutical market and describe our data. In Section 3.4, we describe the hedonic approach to demand estimation used in this analysis and present our preference parameters' estimates. In Section 3.5, we describe our modeling approach to estimate the determinants of product survival. This includes presenting both the covariates (constructed indexes of quality, differentiation, price control, age and market structure) and the estimation framework. In Section 3.6 provides our results and in the last section, we discuss and conclude.

3.2. Literature Review

Our work relates to two distinct strains of the economic literature: the product survival literature and the differentiated products empirical literature of Industrial Organization (IO). In essence, our work contributes to the literature by incorporating developments of the second into the first. As such, we will begin by reviewing the product survival literature, thereby contextualizing the benefits of the developments in the IO literature and then proceed to a brief review of empirical IO literature of relevance for the present work.

Product survival literature

The idea that price regulation of pharmaceutical products has an impact on the firm's entry decisions is not new [38, 47, 48, 54-56] and the general conclusion has been that price regulation contributes to launch delay novel products while apparently encouraging the introduction of me-too products.

In 2000, Danzon and Chao [54] use 1992 data from 7 countries to evaluate the impact of different regulatory environments on competition. Evidence provided in that paper, namely, that of ineffective competition among same compound products in more regulated countries supports (as noted by the authors themselves) our hypothesis of me-too introductions as a strategy to overcome price regulation.

Following that work, Danzon and several colleagues have produced a series of papers on the impact of regulation on launch delay and innovation at the country level [38, 48]. While that

line of research focuses on the delayed introduction of novel products resulting from price regulation, the present analysis focuses on a complementary aspect, the impact of price regulation on the probability of withdrawal after controlling for quality and competition indicators.

A few studies have analyzed survival at the product level. Determinants of survival which have been studied include, among others, product characteristics, market structure and firm characteristics. In 1995, Stavins [50] investigated product quality as a determinant of the likelihood of exit of models in the US personal computer market (without accounting for product duration). Later, Greenstein and Wade [57] estimated the relationship between the number of competitors and the survival of products in the US mainframe computer market. And Asplund and Sandin [58] studied the relationship between market share and the survival of products in the Swedish beer market. Neither of these two studies examined quality as a determinant of the product life cycle.

In 2005, Figueiredo and Kyle [51], improve on the previous work by including three proxies for quality, one of which was the residual from the hedonic regression on product characteristics, and complement the survival analysis (exit decision) by also modeling the determinants of product entry. In the same year Ruebeck [59] models the determinants of product exit in a vertically differentiated market (applied to the computer hard disk). The work by Requena-Silvente and Walker [52], in the UK automobile market, innovates by including indexes of quality, intra-firm and inter-firm competition constructed from the parameter estimates obtained from a hedonic regression (similar to those used by Stavins [50]).

Our work builds on this vein of literature, by applying a more flexible methodology to estimate unobservable product characteristics, estimating its associated preference parameter and incorporating this information when constructing the indexes. Some of the indexes we use are also different from those of Requena-Silvente and Walker since our main focus is the impact of price regulation on product withdrawal.

Most recently, Cardoso [49] analyzed survival of pharmaceutical products in three countries, including Portugal. Part of her analysis refers to approximately the same period as ours and utilizes a sample of one of the databases we also use in our analysis (the INFARMED database). Her analysis focused on the impact of regulatory changes on product withdrawal (including both market withdrawal and transformation in to generic products) while ours focuses on the impact of price regulation on product withdrawal. Our analysis differs from Cardoso's in that we

include price and quantity information (not available to her) to construct quality, competition and price control indexes which are included are regressors in the survival analysis.

Empirical IO literature

Estimation of demand in differentiated products markets has seen major developments in the empirical IO literature in the last two decades. Within such developments, unobservable characteristics and consumer heterogeneity have been shown to be of significant relevance in markets of differentiated products [53, 60, 61].

In 1994, Berry [62] operationalized the empirical framework for demand estimation of differentiated products with unobservable characteristics and heterogeneous consumers. One year later, Berry, Levisohn and Pakes (BLP) [60] applied that methodology and estimate a structural model of demand (and supply) in the automobile industry. This line of research has had many followers and consequent advances over the years [63, 64]. As shown by these authors omitting unobservable characteristics from the demand system will generate biased estimates of the willingness to pay for product attributes[28].

In 2005, Banjari and Benkard (BB) [53] presented an alternative approach - the hedonic approach - to demand estimation with unobservable characteristics and heterogeneous consumers. Building[29] on the pioneer work by Rosen [65], the authors develop an easily implemented, although data demanding, two-step method which overcomes (i) some of the limitations identified in the BLP approach [61, 63] and (ii) some of the criticisms to Rosen work [53]. The BB approach has since been applied in several settings [66, 67].

Simply put, in the hedonic approach consumer demand is deterministic (unlike the BLP-like random utility models). This method has advantages and disadvantages with respect to the random utility BLP approach. The BB approach avoids the need for available valid instruments, does not require an independent and identically distributed (*i.i.d*) random error in the utility function, and is computationally simpler (see BB for a formal discussion). However, it is more demanding on the data and may require a stronger assumption of independence among observable and unobservable characteristics. In practice, the two approaches may lead to different results, although more research is needed (see [61, 64] for a discussion on this issue).

[28] If unobservable characteristics are correlated with prices, as expected, price elasticity estimates will be biased downward.

[29] BB relax three assumptions in Rosen's work, namely, perfect competition, continuum of products and perfect observability of characteristics.

Although accounting for unobservable characteristics has become the rule rather than the exception in empirical IO, to our knowledge, we are the first to explicitly incorporate preference parameters for both unobservable and unobservable characteristics in the survival analysis literature[30]. In doing so, we closely follow the BB approach to estimate structural preferences parameters for both observable and unobservable characteristics.

3.3. The Database

The analysis performed is based on the Intercontinental Medical Systems (IMS) database for the period 1997-2003. This panel data provides information, at the product level, for all outpatient drug sales in Portugal, with revenues above 300 Euros. The data set includes information on brand name, Anatomical Therapeutic Chemical (ATC), manufacturer, mode of administration, pharmaceutical form, strength, package size, prescription requirements, value added tax (VAT) class, reimbursement rate, consumer price per pack and first marketing date.

This data set was matched to the INFARMED website information [68] in order to identify the International Common Denomination (ICD), the marketing authorization license (AIM) owner, AIM number, AIM date, details on interior packaging, prescription requirements, licensing status and reference price. The licensing date refers to the licensing date of the first package. The reference price is that of 2004, the time at which the website was consulted.

It should be noted that the licensing date, quite often, differs significantly from the first "positive" sales date, as reported by IMS. The average difference is 4 years between the marketing date and the licensing date, but with a wide variation from -46 years to 52 years. Although a posterior marketing date may be due to entry deterrence strategies, a marketing date before the licensing date (5% of the sample) implies that, in legal terms, (and for pricing concerns) the product has been replaced. Given the aim of the present analysis, the most recent of the two dates will be assumed as the entering date. In those cases where the AIM date was not available, the date reported by IMS was assumed.

The INFARMED website provides information on licensed products, independently of whether these products have actually been placed on the market. Likely due to economies of scale in the licensing process, the number of variants of the same product, licensed but not introduced in

[30] Note that while the residual of the hedonic regression has been used as a proxy of unobserved quality, in our work we use flexible non-parametric methods based on the techniques developed by Fan and Gijbels (1996) to estimate the hedonic unobservable, then estimate its associated preference parameter (as with any other observable characteristic).

the market, is significant. This phenomenon, certainly of interest in a structural approach to entering strategies, is beyond the scope of the present analysis and this information was, therefore, discarded.

On request, INFARMED also kindly provided information on all AIM requests and revocation requests in the period of analysis. The AIM number was used to link this information to the remaining data.

Due to price control, price variations are determined jointly by the Ministry of Health and the Ministry of Economics. Any price variations beyond that, are to be reported (and approved) by *Direcção-Geral das Actividades Económicas (DGAE*). Information on all requests for price changes in the period of analysis were kindly provided by that governmental department and used to validate the prices provided by IMS. As expected, a lag was often found between the new price year of approval and the year in which the drug was sold at the new price. In the analysis, the IMS price was assumed, since this reflects the price faced by consumers (and AIM owners).

The relevant price from the patients' perspective is the co-payment they will be charged. If price sensitivity were driven solely by consumer co-payments, we might expect small quantity elasticity's with respect to manufacturer prices. However, considering that demand decisions also reflect physicians' incentives through detailing and pharmacists' through more attractive margins, the full price was used in the analysis. The price per pack was used in the analysis to reflect the possible waste associated with package size.

Table 13: Product Turnover in the Portuguese Pharmaceutical Market

Year	# AIM owners	#ICDs	Products	Entries	Exits	Generics Volume mk share	OTC Volume mk share
1997	217	1,273	3,843		5%	0.07%	23%
1998	222	1,268	3,921	7%	8%	0.07%	21%
1999	217	1,266	3,947	8%	7%	0.14%	20%
2000	216	1,261	3,947	7%	6%	0.20%	19%
2001	218	1,255	4,079	9%	6%	0.42%	18%
2002	221	1,225	4,687	18%	9%	1.35%	17%
2003	229	1,199	5,378	21%		3.32%	18%
1997-2003	268	1,491	6,899	44.3%	22%		

As shown in Table 13, the data consists of 6,899 distinct products sold, at some point, during the 7 year period of analysis. A product is, in the present context, defined by the complete set of characteristics (active ingredient, dosage, strength, packaging format and size, administration mode, generic or not, prescription drug or OTC) and AIM number. This means that if a product

is withdrawn from the market and reintroduced with the exact same characteristics but a new licensing process (and thus a new AIM number), it is assumed as a new product. By the same token, if a product is transformed into generic, it is considered a new product. The only exception concerns the reimbursement rate, which in case of change from one year to the other, it is not considered to be a new product.

These 6,899 products are built on 1,491 different ICD and registered to 268 different pharmaceutical firms. An ICD is defined as an active ingredient or a combination of active ingredients. Thus, for example, acetylsalicylic acid is different from acetylsalicylic+ascorbic acid.

We are aware that multiple links between pharmaceutical companies exist and that, as a consequence, these 268 entities possibly reflect a much smaller number of firms. Nevertheless, in the absence of such information, these firms were assumed to act independently.

Although generic products had legally been introduced in the Portuguese pharmaceutical market for almost a decade, the impact of such regulation in terms of actual sales/market shares was so diminute that new regulation, creating incentives for generic substitution, was published in the 2000-2003 period [69]. The increase in the percentage of new products in the 2002-2003 period reflects precisely that legislation. This increase in the number of new products result not only from the generics introduced, but also from the "response" of the branded market to that new environment, as shown in Table 14. As noted by Cardoso [49] the possibility of transforming branded products in generic ones only became effective in the end of 2003, consequently our data does not show too much of an abnormal increase in the number of exits in the period of analysis.

Table 14: Entry and Exit, Branded Versus Generics

Year	Brand Market				Generic Market			
	Total	%	Entries	Exits	Total	%	Entries	Exits
1997	3,826			5%	17			12%
1998	3,901	99%	7%	8%	20	1%	25%	10%
1999	3,901	99%	8%	7%	46	1%	61%	4%
2000	3,881	98%	6%	6%	66	2%	33%	3%
2001	3,955	97%	8%	6%	124	3%	48%	6%
2002	4,459	95%	17%	9%	228	5%	49%	6%
2003	4,811	89%	16%		567	11%	62%	
Average	4,105	96%	10%	7%	153	4%	46%	6%

Due to the already mentioned divergence between the licensing year and the year of first reported sales, we assume a product enters the market in year X if it was not sold in year X-1

and it exited the market if it was not sold the following year. Consequently, we do not have information on which products were launched in 1997, so it is not possible to estimate the percentage of entries in that year. The same occurs with respect to exits in 2003[31]. Overall, 44% of the 6899 products in the data were introduced after 1997 and 22% exited between 1997 and 2003. There was, in the analysis period, a turnover rate of about 7% per year in this market.

The generics market share, in terms of packages sold, increased significantly during the period of analysis, reflecting the incentives created, but was still less than 4% at the end of the study period. The OTC market share remained fairly constant over the period of analysis, with an average of 19%.

As shown in Table 15, the turnover rate (excluding the abnormal 2002-2003 years in the prescription market) was about the same in the OTC and the prescription drug market. The average age of the products in the market was higher ($p < 0.001$) in the OTC market.

Table 15: Entry and Exit, Prescription Versus OTC

Year	Prescription Drugs					OTC Drugs				
	Total	%	Entries	Exits	Average Age	Total	%	Entries	Exits	Average Age
1997	3,156	82%		5%	11.5	687	18%		6%	16.8
1998	3,248	83%	8%	7%	11.7	673	17%	4%	9%	17.7
1999	3,296	84%	9%	7%	11.8	651	16%	6%	7%	18.2
2000	3,317	84%	7%	6%	12.2	630	16%	4%	6%	18.7
2001	3,427	84%	9%	6%	12.2	652	16%	9%	7%	18.3
2002	4,056	87%	20%	9%	11.1	631	13%	4%	8%	18.5
2003	4,757	88%	22%		9.8	621	12%	6%		18.2
Average	3,608	85%	13%	7%	11.5	649	15%	6%	7%	18.1

Note: OTC= Over the Counter Drugs

Due to control measures, price increases determined by the Government are often below inflation rate, implying a decline in real prices over the product's life cycle. Our sample is left censored with respect to price, as we do not have information on launch prices for products launched before 1997, nor do we known about subsequent negotiated price changes. It is, therefore, not possible to have an estimate of the real price of each product. We may, nevertheless, evaluate the loss, in real terms, in the 7 years of analysis, as shown in Table 16. Real price estimation were based on the inflation rates provided by UNECE Statistical database [70].

[31] The IMS data does actually provided such information with a coded sign next to each registry. Because our data was obtained by scanning paper versions of the data set, often signs were missing and such information was discarded.

As presented in Table 16, there is an average loss over the follow-up period of 5.7% (median 6.8%) in package price, and an average annual loss of 1.1% resulting from a loss of 1.2% in prescription drugs and 0.02% in OTC drugs. Among products that have been in the market during the whole study period (not shown), the average total loss is 10% (median 12%) resulting from an average loss of 12% in prescription drugs and a gain of 0.6% in OTC drugs.

Table 16: Real Price Variation

	All		Prescription		OTC	
	Median	Average	Median	Average	Median	Average
Total during follow-up	-6.80%	-5.67%	-6.80%	-6.47%	-3.77%	-0.21%
Average per year in market	-1.52%	-1.06%	-1.53%	-1.21%	-1.01%	-0.02%

Note: OTC= Over the Counter Drugs

With respect to competitive pressure, analyzing at the three-digit level anatomic therapeutic category, there are, on average, 7.5 firms with around 2.6 products each. Although this suggests a significant level of competition, it should be noted that the average number of ICD per level 3 ATC class (ATC3)[32] is 6.4. Since initial price setting rules are based on identical or similar products defined according to (in order of importance) active principle, pharmaceutical form, strength and closest package, the "price competition" may be less severe than the number of firms/products per ATC3 would suggest.

Table 17: Competitive Indicators at the ATC3 Level

Year	Number of ATC3	Number of Firms per ATC3 (average)	Number of Products per ATC3 (average)	Number of Product per Firm per ATC3	Number of ICD by ATC3 (average)
1997	221	7.32	17	2.4	6.4
1998	219	7.37	18	2.4	6.4
1999	222	7.25	18	2.5	6.3
2000	221	7.28	18	2.5	6.3
2001	213	7.67	19	2.5	6.5
2002	213	7.76	22	2.8	6.4
2003	211	8.15	25	3.1	6.3

Note: OTC= Over the Counter Drugs; ICD= International Common Denominator; ATC=Anatomic Therapeutic Category

[32] Defined by grouping all drugs with the same therapeutic and pharmacological characteristics.

For the purpose of the analysis, we have excluded classes V (diagnostic agents) and T (various). These classes represented 0.3% of the sample and introduced a lot of noise with respect to relevant variables, such as units and dosage.

3.3.1. Product characteristics

Table 18 provides summary statistics of the observable product characteristics in the data. Hedonic regressions would ideally include the characteristics that are of intrinsic value to consumers and regulator, namely, therapeutic value, convenience (side-effects, intake frequency, brand associated quality). Moreover, relevant characteristics should include those that affect the physicians' decision to prescribe [71, 72], and the pharmacists' decision to have the product in stock (or to recommend the product in the case of OTC drugs). Since that kind of data is not available, we use proxies such as those summarized in Table 18. We fully realize that these characteristics may not be the determinant in the purchasing decision for several reasons. This is precisely the reason why it is important to account for unobservable characteristics in estimation.

Strength is defined as the quantity of active ingredient per IMS standard unit[33], which is a rough proxy to a dose. Since strength will depend on the active principle under consideration, the variable was normalized to the interval 0-1. This normalization is obtained, by subtracting the product strength from the highest strength available for that ICD, and diving this value by the difference between the highest and the lowest strength in that ICD. Assuming firms will place themselves at the middle (in the Hotelling sense) first, a value of 0.5 was assigned to those active principles with a single strength. Strength was then divided in three classes ($s < 0.33$; $0.33 <= s < 0.66$; $s > 0.66$). The majority of products (42%) are in the weak strength class, as shown in Table 18 .

Units were defined as the number of IMS standard units in the pack. This variable was also normalized to 0-1 in the same spirit. The option for a small pack is the most frequent (43%). In an unregulated market, package size should be inversely related to price per unit, if manufacturers pass on economies of scale to consumers. In the Portuguese pharmaceutical market this effect is mitigated by pack size regulation of drugs subject to reimbursement.

[33] The IMS standard unit is a proxy for the dose for each formulation, *e.g.* one tablet or capsule, 5ml. for liquids, etc.

Administration mode is an indicator of convenience and was therefore included among the relevant characteristics in the analysis[34].

Reimbursement rate classes are determined by the Government. Drugs can be reimbursed at 20%, 40%, 70% or 100%[35]. Generic products obtain an additional 10% reimbursement in the 40% and 70% classes. In the IMS databse we found no product with reimbursement level d (20%) and the majority of products are in the 70% or 40% classes, 36% and 33% respectively.

Prescription is a dummy variable for whether the product requires prescription. The variable was included to control for the fact physicians' preferences are also involved in the purchasing decision. Whether the product is a generic, or not, may affect perceived quality and is, thus, also included in the analysis. By the same token, because the reference price system involves products for which a generic reimbursed alternative is available in the market and to account for characteristics specific of products (and diseases) placed in therapeutic classes with a reference price, we have included a dummy variable for whether or not the product had a reference price in 2004.

Licensing year reflects age. While molecule age may be an inverse indicator of relative therapeutic value, assuming that recent molecules are on average more effective than older molecules, individual product's age will not be related to therapeutic value if new products are me-too versions of existing ones. Although, ideally both age indicators' (molecule and product) would be included in the analysis, because we do not have information on the first data each molecule was introduced, we focus on product age and include it as continuous variables in the analysis.

Dummy variables for ATC3 were also included in the analysis as a proxy for the disease, and thus of need, to consider the possibility that more serious diseases will be associated with a higher willingness to pay.

[34] In doing so, for dimensionally reasons we have grouped letters "e" into "d", "b" into "a", "f" into "g", "o" and "p" into "n", "j" and "k" into "I" and "q" into "v". Letters follow IMS Health definitions.

[35] Reimbursement rates have changed over the years. Currently, there are 4 levels: 95%, 69%, 37% and 15%. (Decree-Law n.º 48-A/2010, May 13th). The pharmacotherapeutic groups which integrate the different reimbursement levels have also been subject to changes over the years.

Table 18: Descriptive Statistics of Product Characteristics

Strength	**6,709**	**ATC1**		**6,880**	**Administration Mode**			**6,880**
Low	42%	a	Alimentary tract and metabolism	15%	Systemic			58%
Medium	29%	b	Blood and forming organs	2%		a	Oral solid	4%
High	29%	c	Cardiovascular system	19%		b	Oral solid retard	11%
IMS Standard Units	**6,880**	d	Dermatologicals	7%		d	Oral liquid	0%
Low	43%	g	Genito-urinary system and sex hormones	5%		e	Oral liquid retard	7%
Medium	26%	h	Systemic hormonal preparations	2%		f	Injectable	1%
High	31%	j	Anti-infectives systemic	13%		g	Injectable retard	2%
Reimbursement Class	**6,871**	l	Cytostatics	1%		h	Rectal, Excl. Hemorroidal	1%
0	20%	m	Muscular-skeletal system	9%		i	Nasal	
0.4	33%	n	Central nervous system	16%		j	Other	
0.5	2%	p	Parasitology	1%				
0.7	36%	r	Respiratory system	9%				
0.8	5%	s	Sensory organs	3%	Local			
1	5%					m	External topics	8%
Generic	**9%**					n	Ophtalmologics	3%
Prescription	**87%**					p	Otic	0%
Reference Price	**12%**					q	Nasal	1%
Licensing Year						r	Pulmonar administration	2%
<1970	5%					t	Vaginal	1%
1970-1979	5%					v	Others	1%
1980-1989	27%							
>=1990	63%							

Although fairly complete, the data set is far from perfect for the analysis at hand. 1997-2003 was a period of significant regulatory changes. New incentives were created for generic adoption, and the reference price system was introduced, just to name a few. The data is influenced by all of these factors, and isolating the effect of price regulation on product's survival becomes more difficult.

3.4. Consumer Preferences

In order to understand the determinants of product survival, it is necessary to account for patients/physicians preferences, so that robust quality indexes may be constructed. Such is the aim of the present section.

In 2005, Bajari and Benkard (BB) [53] proposed an approach to demand estimation, which in essence, consists of a two-step estimation. In the first step, unobservable characteristics are estimated non-parametrically, and in the second step, those estimates are included as an additional (now observable) characteristic. Gibbs sampling is, then, used to obtain structural demand parameters.

3.4.1. Estimation framework and results

Estimation is performed following the methodology described in detail in BB with exemplifying applications in Bajari and Khan [66, 67]. As such, we will briefly describe the framework referring the reader to the original papers for a complete discussion on the subject.

Let $j \in ב$ represent a product. Let the vector $\left(\mathbf{x}_j, \xi_j\right) = \left(x_{j1}, \ldots, x_{jK}, \xi_j\right) \in \mathbb{R}^{K+1}$ represent product attributes; $\mathbf{x}_j$ represents perfectly observable characteristics to both consumers and the econometrician; while ξ_j identifies the characteristic which is observable to consumers but not to the econometrician[36]. In our analysis there are $t = 1, \ldots 7$ markets. Let I_t be the set of all consumers in market t and $I = \cup_t I_t$ be the set of all consumers in all markets.

Assuming unit demand[37] for the product j, consumer I in market t solves the following maximization problem:

$$Max_{(j,c)} u_i(\boldsymbol{x}_j, \xi_j, c) \text{ subject to } p_{jt} + c \leq y_{it}$$

Where y_{it} is the individual i's after-taxes income at t, c is the composite good which price has been normalized to one and p_{jt} is the price of product j at time t.

The price function $\mathbf{p}_t\left(\mathbf{x}_j, \xi_j\right)$ is the equilibrium relationship between prices and characteristics in market t. It is a map from the set of product characteristics to prices that satisfy $p_{jt} = \mathbf{p}_t\left(\mathbf{x}_j, \xi_j\right), \forall_{j \in ב}$. This equilibrium price function depends on market primitives, thus, the t subscript. It does not inform about the price of a good that is not yet on the market, since its entry would affect the primitives and thus the whole price function.

In order to estimate the hedonic price function flexibly, we use a local linear model, whereby we assume that locally the hedonic price function $\mathbf{p}_t\left(\mathbf{x}_j, \xi_j\right)$ satisfies (ignoring the t subscripts):

$$p_j = \alpha_{o,j*} + \sum_k \alpha_{k,j*}\left(x_{j,k} - x_{j*,k}\right) + \xi_j$$

That is, we assume that in a neighborhood of $\left(x_{j*}, \xi_{j*}\right)$ the hedonic function is approximately linear. However, unlike a linear regression, where the relationship between the dependent and independent variables is globally linear, the relationship here is only locally linear. Thus, the

[36] The method allows for only a single-dimensional, vertically differentiated unobserved characteristic although the work has been extended in Benkard, C.L. and P. Bajari. J. Bus. and Econ. Statis., 2005. 23: p. 61-75.to account for a multidimensional vector of characteristics.

[37] See the discussion at the end of the present section.

coefficients have a subscript $\alpha_{\cdot,j*}$to emphasize that they will be specific to a particular bundle of characteristics $\left(x_{j*}, \xi_{j*}\right)$. This approach, followed by BB, is based on the techniques described by Fan and Gijbels [73].

For any $j^* \in \beth_t$, we use weighted least squares to estimate α_{j*} with weighting matrix:

$$W = diag\{K_h(\boldsymbol{x}_j - \boldsymbol{x}_{j*})\} = diag\left\{\frac{K\left(\frac{\boldsymbol{x}_j - \boldsymbol{x}_{j*}}{h}\right)}{h}\right\} = diag\left\{\frac{\prod_{k=1}^{K} N\left(\frac{\frac{x_{j,k} - x_{j*,k}}{h}}{\hat{\sigma}_k}\right)}{h}\right\}$$

, where $K(z) = \prod_{k=1}^{K} N\left(\frac{Z_k}{\hat{\sigma}_k}\right)$ and $K_h(z) = \frac{K\left(\frac{z}{h}\right)}{h}$. K is thus a product of standard Normal distributions evaluated, for each characteristic k, at $\frac{Z_k}{\hat{\sigma}_k}$ where $\hat{\sigma}_k$ is the sample standard deviation of characteristic z. h is the bandwidth chosen large enough to insure a smooth distribution function. Note that the kernel weights W are a function of the distance between product j^* and product j, thus assigning greater importance to observations near j.

Our estimates of the price function parameters allow us to recover an estimate of the unobservable product characteristic. This estimate is obtained as the residual of the local linear kernel regression model. Since the unobservable characteristic has no intrinsic units (and a separate regression is estimated for each product), normalization is required for identification. Following BB, we normalize such that the marginal distribution of $\xi \in U[0,1]$.[38]

In order to reduce the dimensionality, following BB we assume the price function to be additively separable in the ATC3 fixed effects[39] and proceed by first estimating these effects by ordinary least squares, then subtracting the ATC3 fixed effects from the price and applying the local linear kernel regressions described above to the price net of ATC3 effects. In a first step, we run a linear regression of price on strength, pack size, pharmaceutical form, administration mode, generic status, reimbursement rate, existence of reference price in class, licensing year and ATC3 fixed effects. ATC3 absorb a number of important attributes, such as severity of disease. We use dummies for categorical characteristics and the log of continuous

[38] In order to do this, simply associate the estimated omitted attribute with its percentile. I thank Patrick Bajari for this clarification.

[39] Such assumption allows to apply the nonparametric techniques to a 7-dimensional problem (instead of the approximately 300-dimensional one).

characteristics. Due to the domain of the variables strength and pack size variables, we do not transform these variables. Figure 11 provides the histogram of the estimated normalized unobservable characteristic.

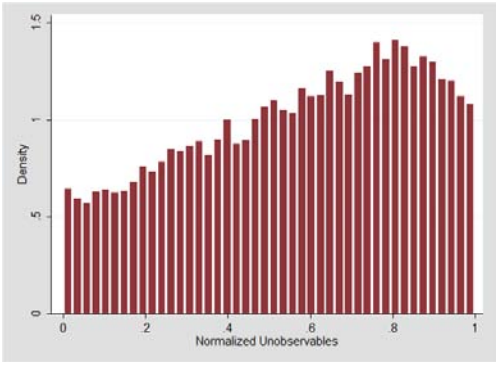


Figure 11: Histogram of normalized residuals from the local linear kernel regression

Once an estimate of the unobservable is obtained, it is included in the set of observable characteristics and preferences estimation, with only observable characteristics, follows. In what follows, ξ is thus be included in the x vector of product characteristics and we redefine $\boldsymbol{x}_j \equiv (x_{j1}, \dots, x_{jK}, \xi_j)$ and $\boldsymbol{\beta}_i = (\beta_{i,1}, \dots, \beta_{i,K}, \beta_{i,\xi})$.

As noted by BB, a parametric assumption on the utility function is required for identification. We assume a quasi-linear parametric function for utility, loglinear in continuous characteristics[40] and linear in discrete characteristics. The utility function is thus described as (omitting the t subscripts):

$$u_i(\boldsymbol{x}_j, y_i - p_j; \boldsymbol{\beta}_i) = log(\boldsymbol{x}_j^c)\boldsymbol{\beta}_i^c + \boldsymbol{x}_j^d \boldsymbol{\beta}_i^d + log(y_i - p_j)$$

, where $\boldsymbol{x}_j \equiv (\boldsymbol{x}_j^c, \boldsymbol{x}_j^d)$, $\boldsymbol{x}_j \in \mathbb{R}^{K+1}$, $\boldsymbol{\beta}_i \equiv (\boldsymbol{\beta}_i^c, \boldsymbol{\beta}_i^d)$, $\boldsymbol{\beta}_i \in \mathbb{R}^{K+1}$ and the coefficient on $y_i - p_j$ is normalized to one.

Since we do not have information on income distribution, we will assume the approach suggested by BLP, whereby a random draw is taken from the income distribution of the Portuguese population and we will assume the distribution of income to be independent of preferences. Given the link between health and income, as well as the agency problems in the market under consideration, the distribution of income for the entire population is not the distribution that should be taken into consideration. This issue is nonetheless mitigated by the

[40] The log specification allows product characteristics to have diminishing marginal utility.

fact that Portugal has National Health Insurance with null co-payment for most chronic diseases, where income effects are likely to be more significant.

The average value of after-tax equivalent income per deciles (Organization for Economic Co-operation and Development, OECD, modified scale) based on the National Health Survey 2005/2006 was used to obtain after-tax income distribution per equivalent adult[41]. In order to sample from that distribution in the simulation, we have adjusted a parametric distribution in RiskPalisade®, which is present in Figure 12.

If different segments of the population (retired individuals, for example) purchase more products than younger individuals, this National distribution of income should be adjusted to account for different weights of different income segments, with respect to pharmaceutical production consumption. Such information is not available, however, so the general distribution was assumed.

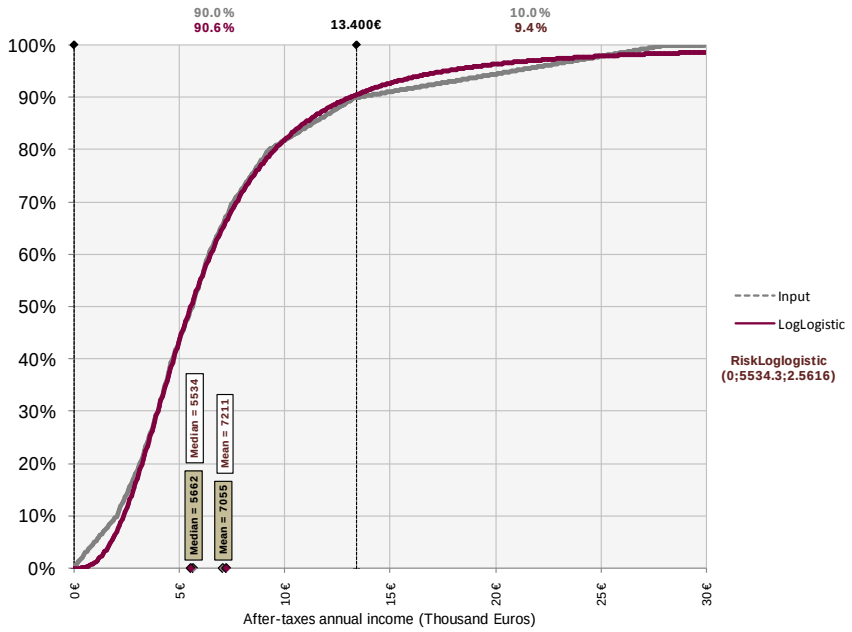


Figure 12: After-Taxes Income Distribution, per Equivalent Adult, in the General Population

The distribution of consumer preferences is processed by using the Gibbs sampling algorithm proposed by BB for the case where the commodity space is discrete. The consumer chooses from a finite set of $j = 1, \dots, J$ products. If the consumer i chooses product j it must be the case that:

[41] Data kindly provided by the National School of Public Health.

$$u(\boldsymbol{x}_j, y_i - p_j; \boldsymbol{\beta}_i) \geq u(\boldsymbol{x}_k, y_i - p_k; \boldsymbol{\beta}_i), \forall_{j \neq k}$$

The aim is therefore to estimate the set:

$$A_i = \{\boldsymbol{\beta}_i : u(\boldsymbol{x}_j, y_i - p_j; \boldsymbol{\beta}_i) \geq u(\boldsymbol{x}_k, y_i - p_k; \boldsymbol{\beta}_i), \forall_{j \neq k}\}$$

The likelihood function for this model is thus:

$$L(j|\boldsymbol{x}, y_i, \boldsymbol{\beta}_i) = \begin{cases} 1 \;\; if & u(\boldsymbol{x}_j, y_i - p_j; \boldsymbol{\beta}_i) \geq u(\boldsymbol{x}_k, y_i - p_k; \boldsymbol{\beta}_i), \forall_{k \neq j} \\ 0 \;\; if & otherwise \end{cases}$$

The Gibbs sampling algorithm, as any Bayesian analysis, requires an assumption on the prior distribution of the coefficient parameters. Following BB, we have assumed that the prior distribution for $\boldsymbol{\beta}_i$ has a uniform distribution defined by a set of conservative lower and upper bound for each taste coefficient. These limits were defined as 100 times the 95% confidence interval bounds of our initial guess of the corresponding willingness to pay of each variable (obtained by ordinary least squares). In that case, for each $l = 1, \ldots, K+1$, the conditional distribution $p(\beta_{i,l}|\boldsymbol{x}, p, C(i) = j, \boldsymbol{\beta}_{-l})$, will be $U[\beta_{i,l,min}, \beta_{i,l,max}]$, where

$$\beta_{i,l,min} = max\left\{min_{\beta_l|\beta_{-l}} B, max_{s.t.x_{l,j}>x_{l,k}}\left\{\frac{\sum_{m \neq l} \beta_{i,m}[log(x_{m,k}) - log(x_{m,j})] + log(y_i - p_j) - log(y_i - p_k)}{log(x_{l,j}) - log(x_{l,k})}\right\}\right\}$$

$$\beta_{i,l,max} = min\left\{max_{\beta_l|\beta_{-l}} B, min_{s.t.x_{l,j}<x_{l,k}}\left\{\frac{\sum_{m \neq l} \beta_{i,l}[log(x_{m,k}) - log(x_{m,j})] + log(y_i - p_j) - log(y_i - p_k)}{log(x_{l,j}) - log(x_{l,k})}\right\}\right\}$$

A estimate of the parameters' distribution may then be obtained with the following algorithm. Let $\boldsymbol{\beta}_i^{(0)} = \left(\beta_{i,1}^{(0)} \boldsymbol{\beta}_{i,-1}^{(0)}\right)$ be an arbitrary point of support, then,

1. Given $\boldsymbol{\beta}_i^{(0)}$ draw $\beta_{i,1}^{(1)}$ from the $U[\beta_{1,min}, \beta_{1,max}]$
2. Draw $\beta_{i,l}$ conditional on vector $\boldsymbol{\beta}_{i,-l}$ as in step 1 for $l = 2, \ldots, K$
3. Return to 1.

After discarding the first 50 draws, we simulated 1,000 taste coefficient draws per product for a total of 4,993 products in the market in 2003. Parameters were estimated for a single year due to the long running time of simulations[42]. We are thus not able to evaluate changes in

[42] Including the initial draws, this process took about one month running at the CEFAGE – Universidade de Évora server.

consumer preferences over time and thus assume those preferences to be constant over time. Moreover, as noted by BB, in the case of discrete product space, in general, it is not possible to obtain a point estimate for the βs but rather a bounded interval. As a point estimate, the midpoint of the interval was assumed.

Table 19: Summary statistics and OLS coefficients

	Average	St.Dev.	Min	Max	OLS Coef	P>|t|	[95% Conf.	Interval]
Strength	47.52	37.93	0	100	0.12	0.000	0.13	0.16
Package size	48.04	37.59	0	100	0.15	0.000	0.18	0.21
Age (ln)	1.83	1.14	0	3.989	-8.61	0.000	-8.94	-6.96
Reference Price	0.15	0.36	0	1	1.52	0.000	2.08	4.25
Prescription	0.89	0.31	0	1	10.74	0.000	11.33	15.52
Reimbursement rate	(0% is reference)							
40%	0.32	0.47	0	1	-3.68	0.018	-5.98	-0.57
70%	0.43	0.49			-10.71	0.000	-15.97	-9.73
100%	0.05	0.21	0	1	-4.74	0.301	-10.07	3.11
Generic	0.10	0.31	0	1	-14.73	0.000	-19.56	-15.07
Unobservable (ln)	0.56	0.28	2E-04	1	47.75	0.000	49.10	58.55
Constant					-10.93	0.000	-20.64	-16.25
					R2=0.3595; N=4,993			

Table 19 provides summary statistics and OLS coefficients for the variables included in the Gibbs sampling algorithm and Figure 13 graphs the (kernel smoothed) taste distributions for some characteristics, namely, strength (dosage), pack size (units), age and the unobservable characteristic converted to minimal willingness-to-pay. The estimated distribution of willingness-to-pay are centered around the mean value estimated by OLS. It should be noted that since demand for pharmaceuticals is influenced by such factors as insurance coverage, imperfect information, physician prescribing, switching costs, etc., these estimates are still implicit prices, but cannot be interpreted as marginal value to consumers.

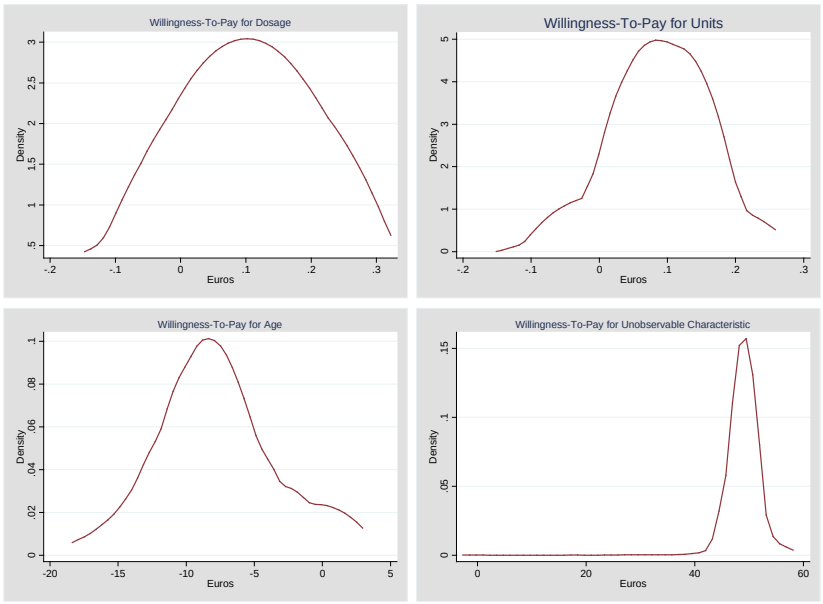


Figure 13: Willingness-to-pay for characteristics

The sequence of random draws may then be used to recover the distribution of tastes for the entire population of consumers since:

$$\lim_{S\to\infty} \frac{1}{S}\sum_{s=1}^{S} g\left(\boldsymbol{\beta}_i^{(s)}\right) = \int g\left(\boldsymbol{\beta}_i^{(s)}\right) Pr(\boldsymbol{\beta}_i | C(i), \boldsymbol{x}, p)$$

Given the aggregate nature of the data, the above expression has to be adjusted to account for the fact that observations represent more than one consumer. If N_j out of a population of N consumers chooses product j, then a weight of N_j/N is attached to each observation.

In performing such adjustment we made two simplifying assumptions. First, we assumed the market share of the outside option is zero. Demand estimation requires knowledge of market size, so that the market share of outside option (not buying) may also be estimated. The entire Portuguese population is, in principle, the market of interest, although ideally, and if available, market for each therapeutic class should be used. In the lack of such information and due to the annual structure of our data, we assume the market share of the outside option to be zero. This assumption is likely to be less relevant in prescription drugs where, for the most part, people see products essential and are thus less likely to forgo buying it if prescribed. The assumption is more relevant in the OTC market, where goods may be seen as less essential. Given the fact that we are interested in average value of preferences parameters and do not aim at welfare analysis, this unrealistic assumption becomes less of an issue.

Second, we assume that products are purchased individually and that no dynamic effects exist. Demand for differentiated products' models, in the BLP and BB literature assume unit demand or at least that the purchase of one product is independent of the purchase of another. In practice, this assumption is unlikely to hold for two reasons. First, on average, 2.12 drugs are obtained by prescription [74] and physicians' prescription habits would suggest that, for a given diagnosis, a physician will prescribe the same (approximate) bundle of products. Using data at the individual level, Dubé [75] has addressed this issue, suggesting how to account for multiple discreteness but such micro level data is not available to us. Secondly, patients learn from past experience and purchase of a product one month increases the changes of buying the same product next month. These learning effects are beyond the scope of this paper.

3.5. The Survival Model

In order to evaluate the determinants of product survival, namely price control measures, the preference estimates obtained in the previous section are used to construct indexes of quality, differentiation, and competition. We will first describe the indexes in detail and then present the results of the survival analysis using the Cox proportional hazard model.

3.5.1. Explanatory Variables

We begin with the quality index ψ, reflecting the value attached by consumers/physicians/pharmacists to the product of interest. Let $\overline{\boldsymbol{\beta}} = \left(\bar{\beta}_1, \dots, \bar{\beta}_{K+1}\right)$ be the weighted average of the willingness-to-pay coefficients estimated by the Gibbs sampling algorithm. Then, for each product j, in market t, the quality index is defined by:

$$Quality_{j,t} \equiv \psi_{j,t} = \sum_{k=1}^{K+1} \bar{\beta}_k x_{k,j}$$

Note that one of the characteristics in x is the estimated unobservable which may now be handled as any other (observable) characteristic.

We define market segments at the 3-digit ATC level. Consequently competition indicators, both intra and inter firms, are constructed at that level. Let $s_t = 1_t, \dots, S_t$ represent the set of market segments in year t and let $j^{s_t} = 1, \dots J^{s_t}$ represent the set of products in segment s at time t. Also, let $f^{s_t} = 1, \dots F^{s_t}$ represent the set of AIM owners in segment s at time t , $j^{f,s_t} =$

$1, \dots, J^{f,s_t}$ represent the set of products of firm f in segment s at time t and $j^{-f,s_t} = 1, \dots, J^{-f,s_t}$ represent the set of products not owned by firm f in segment s at time t.

We are interested in understanding whether price control measures have an impact on product survival. Ideally, we would like to compare the products' real and nominal price and understand if such difference (accumulated over the years since licensing) increases the probability of exit, when controlling for other factors such as quality, innovation and competition. In our sample, we have a significant portion of products that were already in the market on January 1st, 1997. Some products are licensed as far back as 1950. Since then, Portugal has had a revolution, the National Healthcare System was created, and rules have changed significantly. It is, therefore, not possible to estimate the real price of products dating back so far. As a proxy, we have used the average annual real price variation. That is, we estimated the (%) variation between the first nominal price observed in the sample and the last real price observed, and then divided that variation by the number of years the product was observed to be in the market during the study period.

Let π_t be the inflation rate in year t, $rp_{j,T}$ be the real price of product j in the last year (T) it was observed to be in the market ($T \in \{1, \dots, 7\}$, and $np_{j,1}$ be the first observed nominal price for product j. Then our proxy for the impact of price control measures is:

$$Price\ control_j = -\left(\frac{\frac{rp_{j,T} - np_{j,1}}{np_{j,1}}}{T}\right) * 100$$

, where $rp_{j,T} = np_{j,1} \prod_{t=2}^{T} \left(\frac{1+\left(\frac{np_{j,t} - np_{j,t-1}}{np_{j,t-1}}\right)}{1+\pi_{t-1}}\right)$

Product survival depends on the degree of competition and the ability of firms to respond to such competition by means of product differentiation and innovation. We, therefore, construct several indexes meant to account for such effects.

We use three measures of competition at the ATC3 level:

1. *Cannibalization*: the number of products owned the AIM owner J^{f,s_t}.
2. *Rivals' competition*: the number of products J^{-f,s_t} owned by other AIM owners
3. *Similar product competition*: the number of similar products (where similar is defined by same ICD, pharmaceutical form and strength)

To measure *rival quality differentiation*, following Stavins [50] and the literature thereafter, we construct a distance measure for each variant, with respect to rivals' variants in the segment, but excluding own firm's products in the corresponding segment.

$$Rival\ quality\ differentiation_{j^{f,s_t}} = \frac{\sqrt{\sum_{k^{-f,s_t}=1}^{K^{-f,s_t}} \left(\psi_{j^{f,s_t}} - \psi_{k^{-f,s_t}}\right)^2}}{J^{-f,s_t}}, \forall_{j \in} J^{f,s_t}$$

In our sample, because we have several segments with no competitors, after estimating the index for those who do have rival firms in the same segment, we normalize the variable to the 0-1 interval and assign a value of 1 to stand alone firms.

We construct an equivalent index among products of the same AIM owner in a given ATC3 in year t to evaluate within firm quality differentiation:

$$Within\ firm\ quality\ differentiation_{j^{f,s_t}} = \frac{\sqrt{\sum_{k^{f,s_t}=1}^{K^{f,s_t}} \left(\psi_{j^{f,s_t}} - \psi_{k^{f,s_t}}\right)^2}}{J^{f,s_t}}, \forall_{j \in} J^{f,s_t}$$

While a firm can only choose to differentiate a model when it is first launched, proliferation allows the firm to expand its coverage of the product spectrum. Requena-Silvena [52] suggests using the variance of product quality among products owned by the same firm, in segment s at time t as an indicator of proliferation. This index is thus at the firm level, per ATC3.

$$Proliferation_{f,s_t} = \frac{\sum_{k^{f,s_t}=1}^{K^{f,s_t}} \left(\psi_{k^{f,s_t}} - \bar{\psi}^{f,s_t}\right)^2}{J^{f,s_t} - 1}$$

,where $\bar{\psi}^{f,s_t}$ represents the average quality of products owned by firm f, in segment s at time t.

We characterize the firm's positioning strategy in the market with the following indexes:

a. Firm size: Number of product owned by the same AIM owner.
b. Firm atc3: Number of ATC3 groups in which the firm is present
c. Firm quality: Average quality of all products owned by the same AIM owner.

Lastly, we have included two additional variables: an indicator variable for whether prescription is required for product purchase and a binary variable for the 2001-2002 period. This last variable was included following Cardoso's finding that due to regulatory changes the hazard of exit was increased in that period [49].

3.5.2. Estimation framework

We are interested in understanding the determinants of product exit. Therefore parametric or semi-parametric analysis is in order. Because we do not aim at using the model to predict product survival but only to understand the impact of a set of covariates of interest on the hazard, the Proportional Hazard (PH) Cox proportional model [76] was considered.

Our covariates, described in the previous section, reflect market structure and product quality, which vary over time. These along with a proxy for the effect of price control and prescription requirements are our covariates of interest.

Since, in our data, approximately 56% of all products were already in the market before 1997, we have included these observations accounting for left-truncation[43]. Left-truncation occurs when the date the subject becomes at risk is known but the follow-up period begins at a delayed time and in this case, subjects with a short survival period are likely to be overlooked, causing overestimation of the survival rate [77-79]. Although we do not observe products over the complete lifespan, we do have information on when these products were licensed; so it is possible to correct the likelihood function to account for the fact that had the subjected failed, it would never have been observed.

The PH Cox model provides the required information without making any assumption on the hazard over time; what is assumed is that there exists a common baseline hazard function and that the hazard function for a product depends on the values of the covariates and the value of the baseline hazard. Given two products with particular values for the covariates, the ratio of the estimated hazards over time will be constant. This assumption can be tested using several methods as described in detail in Cleeves *et al.* [79] and if valid, the hazard rate may then be estimated, conditional on the obtained parameters. Moreover, random effects, left-truncation and time-varying covariates are easily included in the model with the available software packages, namely Stata11® which was used for estimation.

Three variants of the Cox proportional hazard model were considered and tested for the proportional hazard assumption. A model with ATC3 fixed effects, a model with ATCs random effects and a model stratified by ATC3 level. The stratified model differs from the fixed effects model in that the Cox model is estimated with an entirely separate hazard function for each

[43] Left-censoring means that it is only known that the subject failed at some point between the onset of risk and the beginning of the study (so the subject never enters), but exactly when, in that period, it occurred, cannot be ascertained. In their analysis, Requena-Silvente discard products that were already in the market before the study period, arguing that no method exists to account for left-censoring and time-varying covariates.

ATC3 while in the fixed effects version (that is, including binary variables for ATC3) a common baseline hazard is assumed and the hazard function for one ATC3 is a multiple of the hazard function for the other ATC3. A third approach is to treat ATC3 classes as a random effect. In the Cox proportional hazard model with shared frailty, the random effects enter multiplicatively on the hazard function [80]. The data is organized in $\mathrm{f} = 1, \ldots \mathrm{F}$ groups, with $\mathrm{j} = 1, .. \mathrm{J}^{\mathrm{f}}$ products. For the product j of firm f the hazard is:

$$h_{f,j} = h_0(t)\alpha_f exp(\boldsymbol{x}_{f,j}\boldsymbol{\beta})$$

, where α_f is the group random effect (usually named frailty in the context of survival analysis). Frailties are positive quantities assumed to have a Gamma distribution with mean one and variance θ. Large values of θ reflect heterogeneity among ATC3, and values of θ close to zero suggest little improvement from the traditional Cox model. ATC3 classes with $\theta > 1$ tend to have shorter survival times. To compare the Cox model and the Frailty model, we applied the likelihood ratio (LR) test.

Given the annual structure of our data and the fact that if time is continuous but one only observes it in grouped form, then the Discrete Time PH Model (complementary log-log link) is considered more appropriate [81, 82], and we also tested for this specification.

3.6. Results

Massive generic entry occurred at the end of the study period introducing significant noise unrelated to the main purpose of the analysis which is to understand whether price control measures and the associated real price decrease over the life cycle of the product have an impact on the probability of exiting. Consequently, in the analysis of the determinants of product withdrawal, generic products were excluded[44].

The summary statistics of the covariates described in 3.5.1 are presented in Table 20. The average loss, in real prices over the study period was 6.3%. After exclusion of generic products and 13 outliers, a total of 5,941 products are included in the analysis with a mean follow-up of 4.5 years, yielding a total of 26,918 observations.

[44] Moreover, while testing for the proportional hazard assumption, we found that this assumption would not hold if generic products were included. This non-proportionality was violated in covariates, such as the number of products owned by other firms in the segment or the number of products at the ATC3 level, thus reflecting the entrance of generic products at the end of the period due to a new regulatory environment. Although obviously of extreme interest to understand the impact of generic entrance in the pharmaceutical market, that effect is beyond the scope of this paper and introduces noise in the analysis of product replacement as a strategy to overcome price regulation. We could either have discarded the years of 2002-2003 or discarded generic products. We have chosen the latter, given the short time span and the percentage of generics in the sample.

Table 20: Descriptive statistics of covariates in the survival parametric analysis

		Mean	Std. Dev.	Min	Max
Price control	overall	1.224	3.408	-63.17	58.40
	between		3.408	-63.17	58.40
	within		0.000	1.22	1.22
Quality	overall	2.058	1.200	0.28	23.15
	between		1.248	0.35	20.21
	within		0.177	-9.55	9.22
Cannibalization	overall	4.585	3.909	1	23
	between		4.071	1	21
	within		1.017	-1.1	18.3
Rival's competition	overall	60.953	68.059	0.00	355.00
	between		76.914	0.00	354.00
	within		18.548	-30.71	175.45
Number of similar products	overall	3.783	4.579	1	37
	between		5.235	1	37
	within		1.153	-10.9	26.0
Rival quality differentiation	overall	0.028	0.118	0.00	1.00
	between		0.104	0.00	1.00
	within		0.045	-0.83	0.88
Within firm quality differentiation	overall	0.237	0.353	0	8.46
	between		0.340	0	6.43
	within		0.097	-3.5	2.27
Own firm proliferation	overall	0.381	2.886	0.00	89.50
	between		2.551	0.00	62.83
	within		1.121	-53.59	36.00
# of products	overall	51.758	41.027	1	197
	between		40.866	1	197
	within		6.415	28.8	101.6
# of ATC3	overall	17.349	11.859	1.00	52.00
	between		11.772	1.00	52.00
	within		1.128	12.21	24.21
Average quality	overall	2.058	0.475	0.59	6.18
	between		0.504	0.63	6.18
	within		0.127	0.65	3.76
Prescription	overall	0.857	0.350	0	1
	between		0.332	0	1
	within		0.000	0.86	0.86
N = 26,918; n = 5,941 ;T-bar = 4.5.					

Between 1997 and 2002, in the estimation sample, 1,334 market exits occurred. The Kaplan-Meier survival estimate is presented in Figure 14. The non-parametric analysis estimates a median survival age of 14 years ($95\%\ CI$: $[14; 15]$).

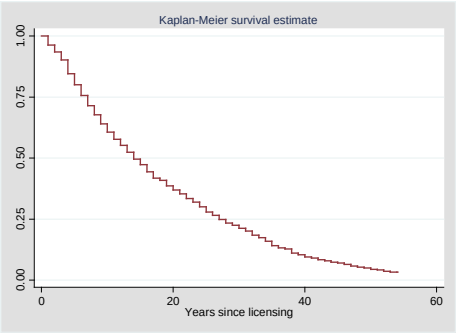


Figure 14: Kaplan-Meier Survival Estimates

The PH Cox model assume proportional hazards, that is, that the ratio of hazards is constant over time, so that assumption should be tested if such model is to be used. Following Cleves *et*

al. [79], we used the Grambsh and Therneau [83] test, based on Schoenfeld residuals, to evaluate the proportional hazard assumption for each variable and for the global model.

In the exploratory analysis, we found that a unique baseline hazard for all segments did not fit the data well and did not comply with the proportional hazard assumption. Once we allowed for different baseline hazards, by ATC3, proportional hazard assumption was satisfied and the analysis of scaled Schoenfeld residuals indicated a good fit. The graphical analysis of such residuals is presented in Appendix II. Table 21 presents the results of the PH test assumption for the PH Cox model stratified by ATC3.

Table 21: Test of Proportional Hazard Assumption

	rho	chi2	df	Prob>chi2
Price Control	-0.034	1.5	1	0.2203
Quality	-0.003	0.02	1	0.8991
Competion				
Cannibalization	-0.043	2.91	1	0.0882
Rival's competition	0.053	4.13	1	0.0421
Number of similar products	-0.008	0.1	1	0.7521
Rival quality differentiation	-0.040	2.79	1	0.0946
Within firm quality differentiation	-0.018	0.49	1	0.4829
Own firm proliferation	0.020	0.8	1	0.3722
Firm Characteristics				
Number of products	0.032	1.73	1	0.1885
Number of ATC3	-0.032	1.68	1	0.1953
Average quality	-0.033	1.75	1	0.1854
Prescription drug (Yes=1)	0.016	0.42	1	0.5186
Year 2001 or 2002	0.050	4.12	1	0.0423
Global Test		14.44	13	0.3439

We also tested the proportional hazard assumption using the graphical method of comparing the Kaplan-Meier estimate with the Cox estimate for each ATC3, and while a few misfits did exist (mainly in classes with few observations), the two models yields similar curves for the majority of ATC3. Figure 15 compares the Kaplan-Meier estimate with that obtained using the stratified Cox model for the first 2 ATC3 classes.

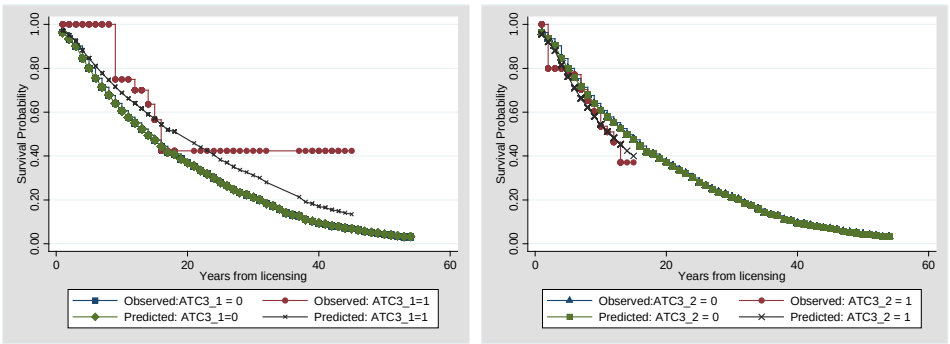


Figure 15: KM Versus Cox Null Model Stratified by ATC3

The Martingale residuals were used to graphically evaluate the functional form of each covariate and, when necessary, the variables were transformed. Fractional polynomial analysis was used to select the most adequate transformation. We use the base e logarithm for Cannibalization, Rival differentiation and Number of similar products and applied no transformation to the remaining variables. Appendix II provides the Martingale residuals analysis for the covariates.

Efficient score residuals were used to evaluate outliers and influential points. 13 outliers were discarded in the final version of the model. Appendix II provides the efficient score residuals graphs after excluding the identified outliers.

Lastly, as suggested in Cleves *et al.* [79] goodness of fit was evaluated using the cumulative hazard of the Cox-Snell residuals. As shown in Figure 16, the model has some degree of lack of fit since the cumulative hazard is slightly above the 45 degree line, but it does not diverge significantly in t.

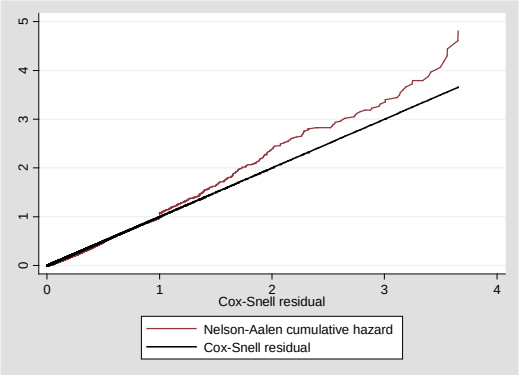


Figure 16: Testing for Goodness of Fit with Cumulative Hazard of Cox-Snell Residuals

While the θ in the frailty model was found to be statistically significant ($\theta = 0.3065$, likelihood-ratio test of $\theta = 0$: $\chi^2 = 160.38; p < 0.001$), neither the fixed effect nor the random effects models complied with the PH assumption ($p < 0.001\ in\ both\ models$) and were therefore discarded.

Given the tested requirement of stratification by ATC3 in the continuous time Cox model in order to comply with the proportional hazard assumption, the discrete time hazard mode
l [82] was also stratified by ATC3 . Results obtained were similar to those of the PH Cox model and are presented in Appendix II.

Table 22 provides the result of the Cox proportional hazard model stratified at the ATC3 level. Price control and quality are both significant at 1% significance level and have the expected sign. Indeed, a one percent increase in the average annual difference between the nominal and the real price, increases the hazard by 6.9%. This result suggests that indeed price control and real price decline over the lifespan of products creates an incentive for product withdrawal. Quality has a protective effect on the hazard; a unit increase in quality decreases the hazard by 50%.

Table 22: Stratified Cox Regression Results

Cox Regression Stratified by ATC3						
Breslow method for ties						
No. of subjects =	5,941					
No. of failures =	1,334			Number of obs =		26,918
Time at risk =	26,918			LR chi2(11) =		286.12
Log likelihood =	-2,968.83			Prob > chi2 =		0.0000
(Std. Err. adjusted for 5,941 clusters in id)						
_t	**Haz. Ratio**	**Std. Err.**	**z**	**P>z**	**[95% Conf.**	**Interval]**
Price Control	1.069	0.010	7.160	0.000	1.050	1.089
Quality	0.502	0.035	-9.880	0.000	0.437	0.575
Competion						
Cannibalization	1.011	0.013	0.880	0.380	0.987	1.036
Rival's competition	0.998	0.001	-1.890	0.059	0.995	1.000
Number of similar products	0.975	0.009	-2.800	0.005	0.957	0.992
Rival quality differentiation	0.359	0.458	-0.800	0.422	0.029	4.371
Within firm quality differentiation	0.403	0.088	-4.160	0.000	0.263	0.618
Own firm proliferation	1.053	0.016	3.510	0.000	1.023	1.084
Firm Characteristics						
Number of products	0.994	0.003	-2.200	0.028	0.988	0.999
Number of ATC3	1.005	0.010	0.560	0.574	0.987	1.025
Average quality	0.866	0.071	-1.750	0.080	0.736	1.018
Prescription drug (Yes=1)	1.259	0.167	1.740	0.083	0.971	1.632
Year 2001 or 2002 (Yes=1)	1.460	0.095	5.840	0.000	1.286	1.658

Notes: ATC3=Level 3 Anatomical Therapeutic Chemical Class.

With respect to competition measures, while the unadjusted for quality number of products owned by the firm in each segment (cannibalization) is not statistically significant, the within firm quality differentiation and own firm proliferation are both significant at 1% and have the expected signs. That is, quality differentiation among products of the same ATC3 owned by the firm has a protective effect while an increase in the firms' coverage of the quality spectrum in each ATC3 results in a higher hazard.

Unexpectedly, we find that the number of me-too products (same ICD, pharmaceutical form and strength) actually has a protective effect on the hazard rate. What is interesting is that the same result is obtained by Cardoso [49], who suggests that the protective effect may be

capturing the fact that markets with higher number of competing products are more profitable markets and, therefore, even with competition, products tend to last longer in those markets.

We also find that, although not significant at 5% but significant at 10%, the number of products owned by rival firms in the same ATC3 does not increase the probability of exiting. This variable was found not to be statistically significant in Cardoso's analysis.

At 10% significance level, prescription drugs have a hazard of exiting 1.26 times that of OTC drugs[45]. The estimated higher probability of survival of OTC drugs is also described in the analysis by Cardoso, who finds that OTC drugs have a hazard rate approximately 60% of the hazard of prescription drugs.

The number of products owned by firm f is a proxy for firm size and results suggest that bigger companies will have longer product life cycles. Qualitatively in line with the result obtained by Cardoso, the binary variable accounting for the 2001-2002 period of significant regulatory changes has the highest hazard ratio. The hazard rate is this period is 46% higher than in the remaining study period.

According to the Portuguese pharmaceutical price regulation, non-reimbursable OTC products' prices may be freely set and are, thus, not subject to annual price revision rules. We, therefore, analyze our results separately in the two market segments[46] and, indeed, find that, as expected, the price control index is not significant in the OTC market. Results suggest a crowding-out effect in the OTC market (captured by statistically significant hazard ratio of 2.8 in the rivals' quality differentiation) while such effect is not significant in the prescription market.

An increase in the number of products owned by the firm in the segment has a protective effect, suggesting possible economies of scope in the OTC market. Firm size is estimated to have some protective effect in the prescription market but not in the OTC market, possibly due scale economies in terms of marketing and market access procedures. As expected, the changes in the regulatory environment captured by the 2001-2002 indicator, are highly significant in the prescription market and not statistically significant in the OTC market suggesting that such regulatory changes has no impact on product survival in the OTC market.

[45] Equivalently, OTC drugs have 20% lower hazard of exiting than prescription drugs because prescription coefficient= Ln(1.26)=0.23; exp(-0.23)=0.79. Thus we estimate the hazard of OTC to be 80% of that of prescription drugs.

[46] Estimation for the OTC market was not stratified by ATC3 given the number of products in the OTC market (751), the number of ATC3 classes (219) and the fact that the model without stratification estimated on the OTC market binds with the proportional hazard assumption.

Table 23: Comparison of OTC Versus Prescription Drug Market

	Prescription Drugs*		OTC Drugs	
	Haz. Ratio	p-value	Haz. Ratio	p-value
Price Control	1.083	0.000	0.968	0.126
Quality	0.487	0.000	0.622	0.020
Competion				
Cannibalization	1.024	0.071	0.908	0.030
Rival's competition	0.998	0.051	1.005	0.065
Number of similar products	0.978	0.019	0.981	0.491
Rival quality differentiation	0.451	0.530	2.796	0.043
Within firm quality differentiation	0.476	0.002	0.541	0.550
Own firm proliferation	1.043	0.007	0.347	0.550
Firm Characteristics				
Number of products	0.992	0.016	1.004	0.582
Number of ATC3	1.010	0.379	0.964	0.130
Average quality	0.840	0.059	0.978	0.911
Year 2001 or 2002 (Yes=1)	1.504	0.000	1.077	0.608

Notes: ATC3=Level 3 Anatomical Therapeutic Chemical Class.; OTC=Over the Counter. *stratified by ATC3

3.7. Discussion

The aim of the present analysis was to understand whether price control measures have an impact on product life-cycle. While novel product introductions are expected to have a positive welfare impact due to therapeutic improvements, replacement of existing products carries costs to society (switching costs, marketing costs and licensing costs) with little or no benefit to consumers.

The reduced form analysis performed, suggests that, indeed, price control measures influence the decision of product withdrawal and, consequently, have an impact on turnover rates in the pharmaceutical market. According to INFARMED, in 2003, 108 ICDs were approved for reimbursement and, of these, only 15% were novel ICDs [84]. It could be argued that, even though the new products are not new ICDs, they still provide better quality from the consumers' point of view. We believe, therefore, it was important to account for product quality in estimation. Such quality, as perceived by consumers and physicians, will inevitably include characteristics that are not observed by the econometrists. If such unobservable characteristics are correlated with prices and are not accounted for, quality indexes will be biased. We have, thus, improved on the line of research followed by Requena-Silvente [52] by accounting for such unobservable characteristics when constructing quality and differentiation indexes. Moreover, we have included a proxy of price control rules among the determinants of product survival.

Given the fact that Cardoso [49] performs an analysis of the determinants of product exit in the Portuguese pharmaceutical market and that part of her analysis is performed in the 1996-2003 period which almost coincides with ours (1997-2003), it is of special interest to compare results obtained. In doing so, it should be noted that not only the focus of the analysis but also the methodology and the datasets used diverge in the two analysis. A major difference is that price and quantity information in which most of our explanatory variables were constructed were not available in their analysis. Even so, it is worth noting that for the most part comparable results obtained in the two analyses do not contradict each other.

We find that the number of similar products (independently of the owner) reduces the probability of exit. This result is in line with those obtained by Cardoso, namely that the number of competitors and own firm similar products decreases the probability of exit. Such finding could also be related to the reported anecdotal effect of generic products on price competition reported by Danzon and Chao [54] and consistent with the idea that "*in countries with strict regulation, generic competitors are predominantly either licensed co-marketers or ''new'' versions of old molecules that manufacturers introduce in order to obtain a price increase*". Although we do not incorporate products classified as generics in our analysis, the same argument may apply.

Cardoso also finds that the number of close substitutes (same therapeutic sub-class) owned by competitors is not a statistically significant determinant of exit. This result is in line with ours since we do not find a statistically significant impact of the number of competitors (close substitutes or identical products) on the hazard rate.

We do not find a statistically significant effect of cannibalization (close substitutes or identical products) on the exiting probability. This result may reflect the opposite effects captured in the Cardoso analysis, namely that identical products of the same firm have a hazard lower than one[47] while close substitutes of the same firm have a hazard higher that one. In accordance with our results, the authors find that products from larger firms (measured by the number of products) have a smaller hazard rate.

The present analysis has several limitations. In terms of data, we have a short span of data on an annual basis. Such structure limits the ability to evaluate product replacement, since gaps between exit and entry may exist due to licensing (and reimbursement) timings. Felix *et al.* [85]

[47] As noted by Cardoso (2010), usually, the range of own products at the same sub-market of chemical substitutes is not of self-competitors but different packages of exactly the same product. Therefore, it is expected that firms take advantage from scope economies of those portfolios.

report a median of 9.9 months for reimbursement requests processes in Portugal in 2007 after the licensing process has been completed, which will take months in itself. The period of analysis (1997-2003) is also not ideal for the analysis of the phenomenon under study, since several confounding factors occurred simultaneously[48], namely, new regulation on generic products which led to a substantial increase in the number of entries unrelated to the phenomenon at stake.

Given the main purpose of the analysis, an additional important limitation is that our price indicator of real price decrease with product age is restricted to an average of (at the maximum) 7 years period. Administrative price decreases implemented by the government and significant fluctuations in inflation rates in the last decades provide a strong argument for the "recent average loss" not to be a perfect indicator, especially for products that have long been on the market.

In the pharmaceutical market, physicians act as patients' agents and, for the most part, select the products to be consumed by patients. Demand for pharmaceuticals is, thus, a mixture of physicians' and patients' preferences. While estimating demand parameters, it would be of interest to recover both sets of preferences, but taking into account the data available, only the mixed bundle was estimated. Our estimated preferences are thus capturing marketing efforts and other promotional activities performed by pharmaceutical companies to which physicians', as utility maximizing agents, may be sensitive. Since our aim is to investigate the determinants of product survival and given the fact that both patients' and physicians' preferences are relevant to the decision made by pharmaceutical companies to withdraw products from the market, the fact that we can only recover a mixed bundle of preferences does not *a priori* invalidate our results.

The fact that physicians' act as patient agents' does nonetheless aggravate the issue of unit demand, usually assumed in the demand for differentiated products literature, since the same bundles are more likely to be repeated. Furthermore, demand for pharmaceuticals are inherently dynamic, in the sense that a consumer who buys (or a physician who prescribes) product *j* today, is more likely to buy (or to prescribe) product *j* tomorrow. These issues were not accounted for in the present analysis.

[48] These are described and analyze in detail in Cardoso (2010) and include the introduction of a major public reimbursement for generics, relative to branded drugs in 2000; the introduction of the reference-price system of reimbursement for some products in December 2002 and the possibility of transforming branded drugs into generics in October 2003.

Such dynamic effects are also present on the supply side. Products' entry and exit are based on multiproduct firms, who optimize their portfolio over time. Only a dynamic structural model would allow for a full understanding of the impact of any given health policy. Such analysis is beyond the scope of the present paper. Such approach has been followed by Filson [86] who within a behavioral dynamic equilibrium model estimates the welfare impact of price controls in the pharmaceutical industry arguing that removing price controls increases firm value, R&D, the flow of new drugs, and consumer welfare globally, but reduces consumer welfare in the countries changing their policies.

The analysis here performed provides empirical evidence on the fact that, indeed, price control measures may lead to higher turnover rates. Further research is necessary on the welfare impact of such turnover, when switching costs, licensing costs and marketing costs are taken into consideration.

4. A Framework for Cost-Effectiveness Analysis of HIV Treatment in Portugal

A Framework for Cost-Effectiveness Analysis of HIV Treatment

In Portugal

Filipa Aragão*

ABSTRACT

We develop a microsimulation discrete events model (MSDEM) for cost-effectiveness analysis of Human Immunodeficiency Virus treatment, simulating individual paths from antiretroviral therapy (ART) initiation to death. Four driving forces determine the course of events: CD4+ cell count, viral load resistance and adherence.

A novel feature of the model with respect to the previous MSDEMs is that distributions of time to event depend on individuals' characteristics and past history. Time to event was modeled using parametric survival analysis. Events modeled include: viral suppression, regimen switch due virological failure, regimen switch due to other reasons, inpatient care episodes and AIDS events and death. Disease progression is structured according to therapy lines and the model is parameterized with cohort Portuguese observational data.

An application of the model is presented comparing the cost-effectiveness ART initiation with two nucleoside analogue reverse transcriptase inhibitors (NRTI) plus one non-nucleoside reverse transcriptase inhibitor(NNRTI) to two NRTI plus boosted protease inhibitor (PI/r) in HIV-1 infected individuals. We find 2NRTI+NNRTI to be a dominant strategy.

Results predicted by the model reproduce those of the data used for parameterization and are in line with those published in the literature.

JEL Classification: I18, I19, C65.

Keywords: Regulation, cost-effectiveness, microsimulation, discrete events, HIV, antiretroviral therapy.

*Support for this research has been provided by Fundação para a Ciência e Tecnologia, Programa POCTI - Formar e Qualificar - Medida 1.1., grant Praxis XXI/BD/19954/99

We are thankful to Julian Perelman, Miguel Gouveia Pinto, Pedro Pita Barros and Mónica Oliveira for helpful discussion. The opinions expressed are the authors alone.

Abbreviations

Abbreviation	Description
ACSS	Administração Central do Sistema de Saúde Healthcare System Central Administration
AE	Adverse Events
AIDS	Acquired Immune Deficiency Syndrome
ART	Antiretroviral Therapy
ARV	Antiretroviral
CAPS	Catálogo Público de Aprovisionamento da Saúde Catalogue of Health Procurement
CART	Combined Antiretroviral Therapy
CD4+	CD4+ T-Lymphocyte count per µL
CE	Cost-Effectiveness
CHC	Centro Hospitalar de Cascais Cascais Hospital Center
CI	Confidence Interval
CVEDT	Centro de Vigilância Epidemiológica das Doenças Transmissíveis Communicable Diseases and Epidemiological Surveillance Center
DCE	Discrete Choice Experiment
DES	Discrete-Event Simulation
DRG	Diagnosis-Related Group
EAS	Economic Evaluation Study
EQ-5D	EuroQol 5-D
HAART	Highly Active Antiretroviral Therapy
HCV	Hepatitis C Virus
HIV	Human Immunodeficiency Virus
HIV-1	Human Immunodeficiency Virus - Type 1
HIV-2	Human Immunodeficiency Virus - Type 2
HUI3	Health Utilities Index 3
ICD9	International Classification of Diseases, 9th Revision
ICD9-CM	International Classification of Diseases 9th Revision, Clinical Modification
ICER	Incremental Cost-Effectiveness Ratio
IDU	Injection Drug User
INE	Instituto Nacional de Estatística National Statistics Institute
INFARMED	Autoridade Nacional do Medicamento e Produtos de Saúde, I.P. National Authority of Medicines and Health Products, IP
IP/r	Boosted Protease Inhibitor
Log_{10}VL	Log_{10} HIV RNA copies per mL
LVHEM	Laboratório de Virologia do Hospital de Egas Moniz – Centro Hospitalar de Lisboa Ocidental, E.P.E. Egas Moniz Hospital Virology Laboratory - Western Lisbon Hospital Center
LY	Life Years
MOS-HIV	Medical Outcome Study-HIV Health Survey
MQoL-HIV	Multidimensional Quality of Life Questionnaire for Persons with HIV/AIDS
MRS	Marginal Rate of Substitution
N	Number of Observations/Individuals
NHS	National Healthcare Service
NNRTI	Non-Nucleoside Reverse Transcriptase Inhibitors
NRTI	Nucleoside Reverse Transcriptase Inhibitors
NS	Non-Suppressive
OI	Opportunistic Infections
P	Probability
PI	Protease Inhibitor
PLWHA	People Living With HIV/AIDS
QAdjLY	Quality Adjusted Life Years obtained using the QoLIndex
QALY	Quality Adjusted Life Years
QoL	Quality of Life
QoLIndex	Quality of Life based on DCE utility estimates
RNA	Ribonucleic Acid
SD	Standard Deviation
SE	Standard Error
SF-6D	Short-Form—6 Dimensions
SG	Standard Gamble
TTO	Time Trade-off
VF	Virological Failure
VL	Viral Load
WHO	World Health Organization
WHOQOL-HIV	World Health Organization Quality of Life Questionary for HIV Patients
WHOQOL-HIV-Bref	World Health Organization Quality of Life Questionary for HIV Patients, Brief Version

4.1. Introduction

According to the latest report by the Communicable Diseases and Epidemiological Surveillance Center (CVEDT) in December 31st, 2009, a total of 37,201 cases of infection have been notified in Portugal since the beginning of the epidemic in 1983 [87]. 24% of these have been notified to have died. Of the 28,388 notified cases alive, 29% have been diagnosed with AIDS, 11% have developed symptoms and the remaining 60% are in an asymptomatic stage of infection.

The number of infected individuals is unknown and likely to be significantly higher than the diagnosed and notified number of cases. According to the UNAIDS estimates [88], there were a total of 34,000 (20,000 to 63,000) infected individuals in Portugal by the end of 2007 with an adult prevalence of 0.5% (0.3% to 0.9%). Within the European region, as defined by the World Health Organization, Portugal had the second highest incidence of HIV infection (251.1 cases per 1,000,000) and the highest incidence of AIDS (79.5 cases per 1,000,000). Moreover, HIV/AIDS infection was the most prevalent cause of death among individuals aged 30 to 39.

In 2007 a National Program for the Control and Prevention of HIV/AIDS infection was defined by the Government [89]. This National plan aimed at (i) reducing the number of new infections by, at least, 25%; (ii) contribute, at an international level, to reduce transmission and improve care to those infected. Its implementation led to the establishment of a special financing regimen within NHS hospitals with regard to HIV infected individuals. Since 2007, hospitals admitted to the program receive an additional annual payment of 11,040€ per infected individual initiating antiretroviral therapy, conditional on satisfying the pre-defined inclusion and performance criteria.

According to the data recently published by the National Coordinator for HIV/AIDS infection, the number of patients on antiretroviral treatment (ART) increased 66% in the 2006-2009 period while the average annual cost per patient decreased by 7% [90]. In 2009, and according to the same source, there were a total of 22,409 HIV infected individuals on ART. Antiretroviral drug expenditure, in that year, was estimated as 193.23 million Euros. The financial burden of HIV infection is also reflected in the 11.4 million Euros spent on HIV related NHS hospitalizations[49], in the same year.

In 2009 pharmaceutical expenditure had reached 4,727.9 million Euros, with 24.2% referring to the NHS hospital market [91]. Between 2008 and 2009, pharmaceutical expenditure grew 7% in

[49] Data kindly provided by ACSS.

the NHS hospital market and antiretroviral drugs were a major driver in such increase [92], representing about 17% of the pharmaceutical expenditure in the hospital market. The country's excessive deficit in 2009 has lead to the proposal and implementation of a series of cost containment measures. Within that context, antiretroviral drugs' expenditure has been singled out as a target.

The present article aims at providing *a framework* for cost-effectiveness analysis of HIV treatment in Portugal. It does not aim at comparing A to B but instead it provides a model, general enough to allow, among others, for comparison of antiretroviral treatment options, of different cut-offs for treatment initiation, or different adherence promoting strategies.

While all efforts have been made to parameterize the model so as to reflect Portuguese clinical practice, given the unavailability of a national database complete enough for estimation of all parameters of interest, several databases were used. Some of the databases are local ones and, in that context, it may be argued that it does not reflect the national clinical practice or that the patient characteristics are not representative of the HIV infected population in Portugal. The databases used are nonetheless, the largest most complete electronic based sources currently existing in Portugal[50]. To our knowledge, this is the first attempt at creating a practical instrument that may be used to support a vast number of HIV treatment resource allocation decisions in the Portuguese context. Moreover, the instrument here developed may be re-parameterized with more representative samples, once these become available.

The present model should be regarded as *one*, among several, instruments to support resource allocation decisions. It is by no means a sufficient instrument, namely because it does not incorporate the externalities resulting from the infectious nature of the disease. Accounting for such externalities requires an epidemiological model, which although obviously of major relevance, was beyond the scope of this paper. Bearing in mind the importance of such externality, and the association between infectiousness and viral load, the model was designed to provide, as one of the outputs, the number of months without viral suppression. The goal, for future research, is to link such output to the epidemiological model.

The work here developed contributes to the existing literature in that the methodology applied, namely a microsimulation discrete events model has not, to our knowledge, been applied to cost-effectiveness of HIV treatment.

[50] In the scope of the HIV financing program, data are being collected nationally since 2007. This database is nonetheless unavailable even upon request for academic use. The EuroSIDA Portuguese cohort including 529 patients, 280 of which currently under follow-up. EuroSIDA.

The article is organized as follows. In Section 4.2, the Human Immunodeficiency Virus (HIV) infection and the impact of highly active antiretroviral therapy (HAART) are briefly described. This section draws the line in which the simulation model was built. In Section 4.3, a review of the existing literature in cost-effectiveness of HIV treatment is performed. In Section 4.4, the discrete events microsimulation model is described and Section 4.5 provides the econometric estimation details for the input parameters. In Section 4.6, model outputs are presented in the context of an application. In the last section, we compare model predictions to those published in the literature and discuss its limitations.

4.2. The Human Immunodeficiency Virus (HIV) Infection

4.2.1. Natural Evolution of HIV Infection

HIV's primary cellular targets are CD4+ T-Cells, the memory cells of the immune system. To reproduce itself, HIV converts its RNA genome to DNA, which is then imported into the host cell's nucleus and inserted into the host genome through the action of HIV integrase. Once integrated, HIV can remain dormant for the duration of these cells' lifetime. To actively produce the virus, certain cellular transcription factors need to be present, the most important of which is upregulated when T-cells become activated. This means that the cells most likely to be killed by HIV are those currently fighting infection [93].

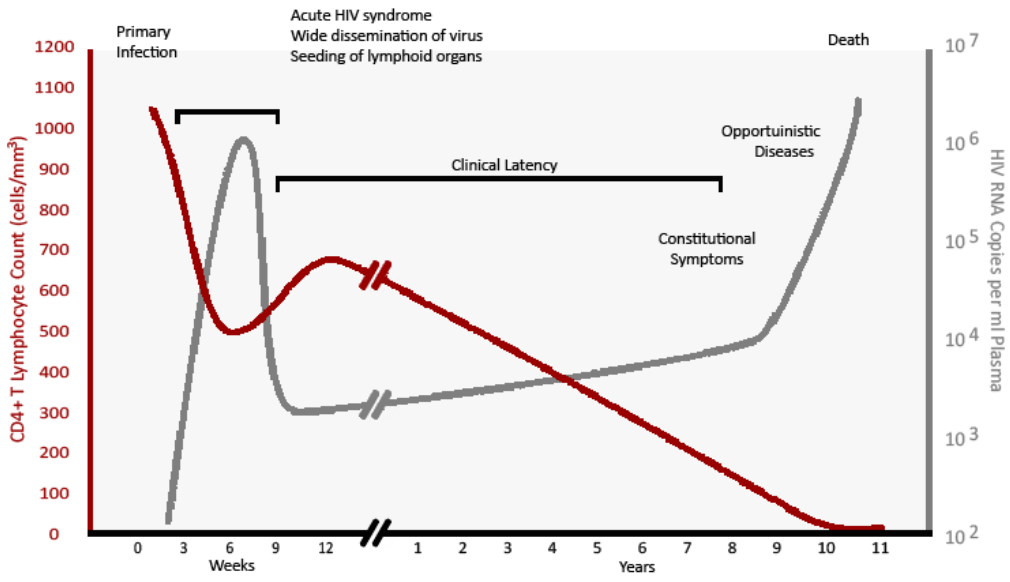


Figure 17: Natural Evolution of HIV Infection (Source: Pantaleo *et al.* 1993)

The natural progression of the HIV infection has been widely described in literature and may be classified in three main stages, as described in Figure 17 [94]:

1. Primary infection and acute HIV syndrome;
2. Asymptomatic period ("clinical latency") with progressive depletion of CD4+ T-cells;
3. Advanced immunodeficiency and AIDS.

In the absence of highly active antiretroviral therapy (HAART), progression from HIV infection to AIDS has been observed to occur at a median of nine to ten years and the median survival time after developing AIDS is only 9.2 months [95].

4.2.2. Evolution of HIV Infection with Antiretroviral Therapy

Since the introduction of HAART in 1996, many HIV-infected individuals have experienced remarkable improvements in their general health and quality of life, which has led to a large reduction in HIV-associated morbidity and mortality in the developed world [96-99]. The Antiretroviral Therapy Cohort Collaboration [100, 101] estimates that the life expectancy of an HIV infected individual at the age of 35 years old is 32 years from the time of infection, if ART is initiated at 350 CD4+ cells/mm^3.

Current standard of care treatment for antiretroviral treatment of HIV infection consists of HAART and so much so that the "Highly Active" (HA) part of the name has, for the most part, been dropped to simply Antiretroviral Treatment (ART). Recommended ART options are combinations (or "cocktails") consisting of at least three drugs belonging to at least two types, or "classes," of antiretroviral agents. Typically, these classes are two nucleoside analogue reverse transcriptase inhibitors (NARTIs or NRTIs), plus either a boosted protease inhibitor (PI/r) or a non-nucleoside reverse transcriptase inhibitor (NNRTI). New classes of drugs such as entry inhibitors and integrase inhibitors provide treatment options for patients who are infected with viruses already resistant to common therapies.

Despite its success in controlling HIV infection and reducing HIV-associated mortality, current drug regimens are unable to completely eradicate HIV infection. Many people on ART achieve viral suppression, HIV RNA drops below the limit of detection of standard clinical assays, for many years. Viral suppression minimizes the changes of developing resistant strains and slows down CD4+ cell count depletion and disease progression. Viral suppression, if achievable, is thus the aim of therapy.

However, upon withdrawal of ART or treatment failure, HIV RNA rebounds quickly with a subsequent decline in CD4+ T-Cells and possible clinical progression. In case of virological failure, the regimen should be changed, since the probability of developing resistant strains depends on the level of viral replication, on duration and on the regimen components [102].

When treatment failure occurs, a new ART regimen is selected based on: resistance assays, previous sequence (and duration) of ART medication, viral load, CD4+ cell count (and other clinical indicators); previous adverse events/intolerance; and personal characteristics, which influence adherence levels [102]. In the present model, each time a virological rebound occurs, the individual is said to start a new therapy line.

HIV infection may thus be described as a sequence of therapy lines, where each line starts with a detectable viral load (and in response to which a new regimen is selected) and ends with an ART regimen change as a result of a virological rebound. The first therapy line starts at the moment ART is introduced and ends when the regimen is switched in response to the first virological failure. The reasoning for subsequent therapy lines is analogous. The lag between virological rebound and therapy implies an inversion of the viral load trajectory before each line switch. When the lag between virological rebound and therapy change is long enough, a subsequent decrease in CD4+ may occur. Once viral load is undetectable, the CD4+ cell count starts recovering (more in earlier lines when the individual is fully susceptible to ART). This process is presented graphically in Figure 18.

Over time, as sequential virological rebounds occur and resistance develops, available medication will no longer be able to suppress viral load. At this stage, denominated non-suppressive therapy, incidence of opportunistic infections (OI) severely increases as CD4+ cell count decreases. Most patients at this stage, will die from OI or malignancies associated with the progressive failure of the immune system.

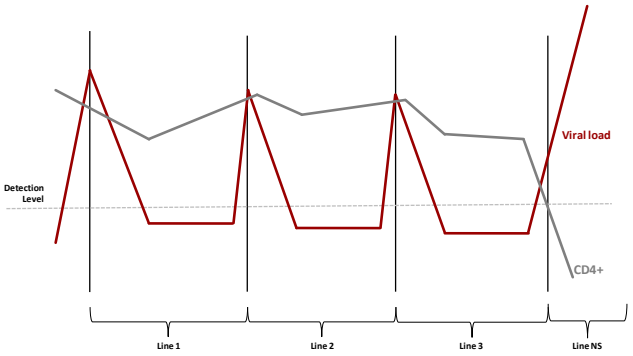


Figure 18: Therapy Line, CD4+ Cell Count and Viral Load over Time

4.3. Literature Review

HIV infection is a chronic disease where current treatment options not only have an impact on disease progression, but also have an impact on the future availability of treatment option. Moreover, individual characteristics and the complete history of past events are also important in determining present and future treatment strategies [102].

The economic evaluation literature regarding HIV prevention and treatment strategies is abundant, as recently shown in the systematic review performed by Hornberger *et al.* [103].

Markov models have been the standard framework for predicting long-term clinical and economic outcomes using the surrogate marker endpoints from clinical trials and cohort studies. Nevertheless, in the last decade, developments in computer capability have opened the gate to individual simulation models, which avoid some of the limitations of Markov models, namely (i) independence of transition probabilities from time spent in the current state, (ii) independence of transition probabilities from states visited before the current state and (iii) events timed to multiples of a fixed cycle time. An overview of the main available HIV related cost-effectiveness models may be found in Giguère [104].

Two papers have been published comparing Markov models to individual microsimulation models in cost-effectiveness analysis of HIV treatment. In 2007 Kuhne *et al.* [105] conclude that, for advanced treatment-experienced patients, the two types of models generated comparable estimates of life expectancy and costs. According to the authors, the cohort model minimizes computing time and simplifies the incorporation of probabilistic analysis of uncertainty, while microsimulation model allows a fuller set of model inputs, avoids over-simplifying assumptions and may have greater face validity among clinicians. In 2009, Simpson *et al.* [5], conclude that individual simulation models performed better than the Markov model in isolating long-term implications of small but important differences in crucial input data. The authors argue that, while more demanding in terms of data requirements, individual microsimulation models impose fewer restrictions, which may explain why these models consistently predict better survival.

To the best of our knowledge, four microsimulation models (along with plenty applications and derivations from those) have been developed, and are currently up to date, to perform cost-effectiveness analysis of HIV treatment.

In 2001, Freedberg *et al.* developed the Cost-Effectiveness of Preventing AIDS Complications (CEPAC) model [106], programmed in C[51]. In its original format, the model included three general categories of health and vital status (chronic illness, acute illness, and death), which were stratified according to important clinical characteristics (e.g., the CD4+ cell count, the HIV RNA level, and the history of clinical events). The characteristics of each patient (age, gender, CD4+ cell count, and HIV RNA level) are randomly drawn from distributions derived from clinical trials. Progression of HIV disease and risks of clinical events and death are derived from cohort data. HIV progression, effects of treatment, and use of resources are linked to both the CD4+ cell count and the HIV RNA level. Monte Carlo simulation is used to individually model the clinical courses of 1 million hypothetical patients, and outcomes are tracked on a monthly basis. The model tracks individual patient statistics regarding time spent in each health state, the sequence of clinical pathways, time on therapy, and length and quality of life; it then tallies the cohort's summary statistics on completion of all individual patient simulations. This model has been extensively applied by the authors and others [107-109].

In 2002, Ritcher *et al.* [110] developed the (now called) Antiretroviral Drug Valuation and Cost-Effectiveness (ADVANCE)[52] model, a microsimulation model, designed in Microsoft Excel®, to simulate the impact of differing drug sequencing strategies. The original version had 3-months cycles and three therapy lines after which the patient was assumed to follow the natural course of disease. In a recently published version/application (Kauf *et al.* in 2010 [111]), patients are eligible to receive up to five sequential antiretroviral regimens. The model is parameterized with a mix of cohort and clinical trial data. CD4 and viral load, both stratified, are the main model drivers. The model takes into account not only disease progression, but also short term adverse events of ART.

In 2010, Kuhne *et al.* [112] published the first full article application of the AntiRetroviral Analysis by Monte Carlo Individual Simulation (ARAMIS) model to the cost-effectiveness of maraviroc. This model is property of Pfizer and, unlike the CEPAC, no appendix is available with a detailed description of the model. From the information available, ARAMIS is structurally similar to the CEPAC model. ARAMIS is developed in TreeAge Pro 2007®, populated with data from clinical trials and validated against the CEPAC model in treatment-experienced patients.

In 2010, Johnston *et al.* [113] developed a new individual microsimulation model innovating in two ways. First, all baselines characteristics *and processes* were modeled using statistical

[51] The coding for the CEPAC model is not freely available.

[52] The name is attributed to the most recent version published by Kauf et al.

analyses of individual-level data for 1,895 individuals in British Columbia receiving ART treatment, that is, cohort observational data. Second, the microsimulation model describing the clinical and economic course of HIV is integrated with a previously built mathematical model describing transmission of HIV [114]. Resistance development is estimated using a survival analysis. Monthly cycles are assumed and, at each step, variables are updated. The software used is not described. The model was built to perform a comprehensive economic evaluation of a program to increase the uptake of HAART in British Columbia.

Our model differs from CEPAC in that (i) it is structured by therapy lines instead of general categories of health and vital status (ii) uses only observational data, (iii) uses survival analysis to estimate the parameters of interest and (iv) links progression of HIV disease, risks of clinical events, effects of treatment, and use of resources not only to CD4+ cell count, HIV RNA level and history of past events (as in theirs) but also to adherence and resistance. Our model shares an important feature with that of Ritcher *et al.*: the fact that it is structured by therapy line and not disease stage; it differs from ADVANCE in that it is based solely on cohort data and has adherence and resistance, along with CD4 and viral load as model drivers. Our model is much in line with that of Johnston *et al.* although we have not yet included the transmission part. At the time the present model was developed, we had no knowledge of the work being developed in British Columbia.

One main difference remains between our model and the five existing ones. The difference relies in cycle length. While applying an individual simulation approach, these models maintain a fixed cycle length structure. At each cycle, according to a random draw, individuals will either move to a new state or remain in the same. With such structure, random draws are necessary even for patients to simply stay in the same state. Moreover, only one event will occur in each cycle, and the shorter the cycle the more computationally demanding the model will become. In discrete-event simulation (DES), event occurrences determine time advances and outcomes are updated at the time of the event occurrence, not at the end of a predetermined time period. Moreover, DES is particularly useful when interaction between individuals is of relevance as is the case with infectious diseases[53] [115].

DES has been suggested and applied in a variety of settings [3, 116-118]. Of particular similarity to our model is that developed by Barton *et al. [3]* - The Birmingham Rheumatoid Arthritis Model (BRAM). The two models share important features, namely, they are both programmed

[53] This feature will be of benefit once the present model is linked to an epidemiological model.

in TreeAge Pro2009® and use the same method to determine timing of activities (in Section 4.5.3 the method is described in detail).

Our model differs from BRAM in that we model the parameters of the time to event distributions as a function of patient characteristics and history of past events while in the BRAM model, distribution parameters are assumed fixed. Moreover, it obviously relates to a different disease.

4.4. The Microsimulation Discrete Events Model

The Discrete Event Microsimulation Model was designed to fit the description of the disease presented in the Section 4.2.2. Figure 19 provides the model diagram; the main structure and assumptions are discussed next, and Section 4.5 provides details on model parameterization.

Following guidelines, we assume in the model that an individual switches line if viral suppression is not reached within 6 months of line initiation or if, after viral suppression has been achieved, virological failure is confirmed. Such assumption is imposed in the model when moving from first to second and from second to third line. In line three, where individuals agglomerate between the second failure and non-suppressive therapy, several virological failures are allowed; that is, regardless of the number of virological failures, the individual will remain in line 3 as long as his resistance level is not above 10 (resistance class 4).

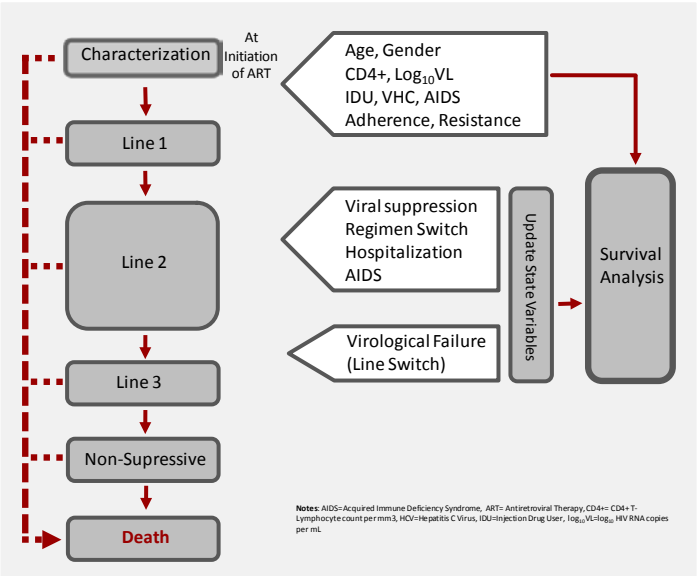


Figure 19: Discrete Events Microsimulation Model Designed in TreeAge ProSuite® 2009

When resistance level reaches a value above 10, the individual enters what is called non-suppressive therapy. In non-suppressive therapy, viral suppression is no longer possible with the available set of antiretroviral drugs. At this point, the treatment goal is to slow down the (inevitable) rate of depletion of CD4+ T-cell.

Individuals enter the model on the day they start antiretroviral therapy. At this moment, the individual is characterized in terms of the following variables: age, gender, mode of transmission, employment status, co-infection with hepatitis C virus (HCV), AIDS diagnosis, HIV resistance level, adherence level, CD4+ cell count, viral load and age of death due to non-HIV related causes. The individual specific values for each of these variables are obtained by sampling from the distributions observed in the data. Some of these variables are assumed constant over time, while others will change over time, conditioning and being conditioned by individual progression and sequence of events. Section 4.5.1, presents the distributions assumed for each of the above mentioned variables as well as the corresponding database source while Section 4.5.2 describes their evolution over time.

Once characteristics are assigned, “time to next occurrence” is sampled from each of the conditional distributions of the modeled events. The event with the shortest samples “time to next occurrence” is the event which will occur next [3]. Seven groups of events are modeled: (i) Viral suppression, (ii) Regimen switches without virological failure, (iii) Resistance development, (iv) Hospitalization, (v) AIDS-defining event, (vi) Line switch, and (vii) Death.

Section 4.5.3 describes each of the events and provides estimates for each distribution of time to event. These distributions depend on past and present individual characteristics, as well as on the therapy line numbers.

State variables, costs and benefits are updated at the occurrence of the event (Figure 20). The process is repeated until death becomes the next event to occur. Death determines the end of the individual simulation.

This microsimulation process is repeated for a large number of individuals (1 million) and the average values are used for cost-effectiveness results.

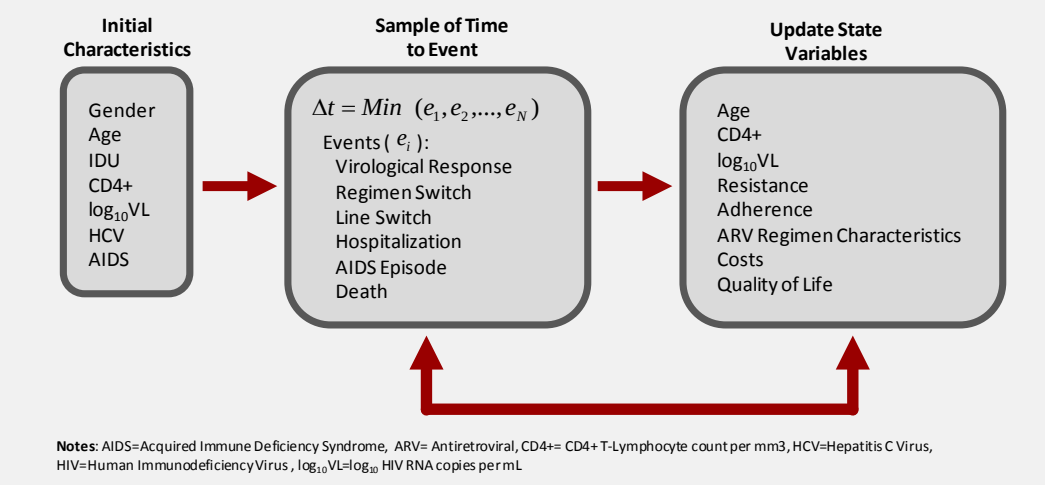


Figure 20: Discrete Event Microsimulation Process

4.5. Model inputs

Given the aim of the present model, all efforts were made to parameterize the model according to Portuguese data. When such parameterization was not possible due to lack of information, external (international) sources were used.

For the most part, the therapy model was parameterized according to five main databases:

1. The Cascais Hospital (*Centro Hospitalar de Cascais* - CHC) database, which provides information on 1,306 HIV-1 infected individuals on ART aged above 13 years-old, in an immunodeficiency treatment unit in the Lisbon area, between 2001 and 2008;
2. The Virological Laboratory of the Egas Moniz Hospital (*Laboratório de Virologia do Hospital Egas Moniz* - LVHEM) database, which provides information on 5,456 patients who have been tested for antiretroviral resistance, (almost) all over continental Portugal, between 2001 and 2008;
3. The Communicable Diseases and Epidemiological Surveillance Center (*Centro de Vigilância Epidemiólogica das Doenças Transmissíveis* - CVEDT) database, which is the national epidemiological database of HIV infection in Portugal and included 34,888 individuals in August 2009;
4. The National Institute of Statistics (*Instituto Nacional de Estatística* - INE) database, which provided information on HIV-related mortality and general population mortality, by age and gender, between 1988 and 2006.

5. The Central Administration of the Healthcare System (*Administração Central do Sistema de Saúde* – ACSS) database, which provides information on all HIV-related inpatient care episodes in 2009.

These databases are described, in detail, in Appendix III. Written consent for access to the databases was approved by the Ethics Committees and/or the entities responsible for the databases, as required.

Significant difference in treatment options exist for pediatric cases and adults/adolescents [119]. Moreover, in the CVEDT database, only 124 out of 26,066 cases alive in 2009 with known age, were children under 13 years-old. Consequently, individuals under the age of 13 were excluded from the analysis.

Due to differences in antiretroviral treatment options, HIV-2 individuals were also discarded from the analysis. Individuals infected with HIV-2 account for 3.4% of individuals registered in the CVEDT database and alive as of August 2009.

Regression models were estimated in Stata 11®. RiskPalisade5.5® was used to fit distribution when required. Selection of the distribution with the best fit was based on the Chi-Square ($\chi 2$) Goodness of Fit Test[54] [120]. Ungraph V5® software was used to recover values generating published graphs.

In this Section, model inputs are described in detail. We start with disease progression and treatment related parameters (distributions from which individual characteristics are sampled at the start of each microsimulation, driving forces of the model and time to event estimates) then move to quality of life and cost related parameters.

4.5.1. Characteristics at ART Initiation

The CVEDT database provides an overview of the (notified) cases in the whole country. It describes the situation at a given moment in time (dead or alive, HIV stage, age, etc.) and

[54] The test evaluates the null hypothesis that the data are governed by the assumed distribution against the alternative that the data are not drawn from the assumed distribution. Unlike the Kolmogorov-Smirnov test, the Anderson-Darling test or the root mean squared error, the Chi-Square goodness of fit test has can be used with sample input data, and any type of distribution function (discrete or continuous). While most critical values and P-values for the A-D and K-S fit statistics have been found by very detailed Monte-Carlo studies, not all distributions have been analyzed in enough detail and are therefore not reported by the software.

It should be noted that distribution parameters are defined differently in TreeAgeProSuite2009® and RiskPalisade5.5® (and for that matter Stata11®). Consequently, parameters obtained in RiskPalisade5.5® have been re-scaled to match the TreeAgeProSuite2009® definitions. The three software manuals provide the necessary information for such transformations.

although it does provide some information about the past, such as, the time of diagnosis it contains no information about ART therapy initiation. Consequently, its usefulness was limited in terms of parameterizing the model with respect to baseline characteristics at ART initiation.

Most of the variables (and distributions) described in this section were, therefore, taken from either the LVHEM database or the CHC database. The LVHEM database has the advantage of including individuals from (almost) all over the country but it has a selection bias, since this database tends to include the "difficult" cases – those individuals who needed a resistance assay. The database also includes resistance tests performed before ART initiation, although such practice is still not fully implemented. The CHC database, on the contrary, represents only a local sample of individuals but no a priori selection bias is expected.

Along with the common baseline characteristics assumed in microsimulation models, we have also included HCV co-infection, mode of transmission, employment status and AIDS diagnosis and regimen characteristics (number of protease inhibitors, daily frequency and number of pills per day).

Employment status was found to be a significant variable in determining quality of life (QoL) among HIV infected individuals (See Section 4.5.4.5.1) and was consequently included in the model in spite of no direct effect being modeled in terms of disease progression.

The disease stage at which individuals are diagnosed is a relevant policy variable, as patients diagnosed at a late stage have poorer prognosis [121-123] and there is evidence to suggest that efforts should be made towards earlier diagnosis in Portugal [124-126]. This variable was, therefore, included in the individual characteristics in the model.

As resistance develops and the infection advances, the required number of PIs for successful treatment increases. Such increase in the number of drugs results in a higher cost and possibly more toxic effects. This variable was, thus, included in the model to reflect modifications in treatment options.

The complexity of antiretroviral regimens means that adherence remains a problem [127]. Daily frequency [128, 129] and total number of pills per day [130, 131] have been shown to influence adherence level. These effects have also been evaluated in the Portuguese context by Reis *et al.* [132]. We have therefore, also included these variables in the model.

Table 24 provides the summary statistics of the distributions used in the simulation as well as the corresponding sources[55].

Table 24: Characteristics at ART initiation

Characteristics at ART initiation		Distribution assumed	Source
Age (median)	40.1	Table	Naive at CHC
Gender (female proportion)	28.4%	Bernoulli	CVEDT
Employment status	66.0%	Bernoulli	CHC and Reis et al. 2007**
HCV Co-infection	29.5%	Bernoulli	Naive at CHC
AIDS diagnosis	32.2%	Bernoulli	CVEDT/Naive at CHC
Model of transmission			
Injection Drug User	27.8%	Table	CVEDT, 2004-2008
Heterosexual	57.5%		
Homosexual	13.7%		
Other	1.0%		
CD4 cell count /μL (median)	239.5	Table	Naive at CHC
Log_{10} Viral Load	4.8	Table	Naive at CHC
Adherence (average)	86%	Table	Naive at CHC
Genotypic Sensitivity Score (inverted*)			
<1	91.2%	Table	Naive at LVHEM
1<=R<5	6.7%		
5<=R<10	1.6%		
R>10	0.5%		
Number of PIs in regimen (median)	1	Table	Naive at CHC
Daily frequency (median)	2	Table	Naive at CHC
Number of pills per day (median)	5	Table	Naive at CHC

Note: PI=Protease Inhibitor; *inverted to reflect resistance instead of sensitivity; **calculated

4.5.2. Evolving State Variables

In the model, most of these characteristics evolve over time and are updated upon event occurrence thereby influencing, along with the accumulated history of events, the distribution of time to next event occurrence. Employment status, HCV co-infection, mode of transmission and gender are assumed constant over time due to either being immutable or due to lack of information in the database.

[55] Table distributions are step-function distributions built based on the observed frequency of each value. To sample from a Table distribution, a uniform random number is generated and matched with the cumulative distribution to obtain the corresponding value of the variable of interest. When relevant, a discussion on the value assumed and a histogram from which the Table distribution was constructed is presented in Appendix IV.

.

4.5.2.1. *CD4+ Cell Count and Viral Load*

The average monthly variation of CD4+ in each therapy line was estimated by weighted non-linear least squares, based on the information provided by the CHC sample. Several functional forms were considered based on the following criteria: (i) possibility to explicitly condition variation on the initial CD4+ (at line start) and (ii) lowest Akaike's information criterion (AIC) [133]. The logarithmic function of the form $\mathrm{iCD4_n} + \mathrm{b_n} * \ln \mathrm{t}$, where $\mathrm{iCD4_n}$ is the CD4+ cell count at line *n* start and b_n is the estimated coefficient for line *n*, attained the lowest AIC in all lines and was therefore selected.

At start of a new therapy line, $\log_{10}$ of HIV RNA copies per mL ($\log_{10}$VL) is assumed to decrease at the average rate observed in the CHC sample. This average variation is conditional on line number, but not on the remaining state variables. The methodology used to estimate the average monthly variation of $\log_{10}$VL in each therapy line was identical to that used for CD4+. Table 25 provides the estimation results for both CD4+ and viral load[56].

Table 25: Monthly Variation of CD4+ and Viral Load, by Therapy Line

	b0	b1	SE	t	P>t	[95% CI]		LogLikelihood	AIC	N. Obs
Average CD4+ - Line 1	305.7	71.6	0.384	186.2	0.000	70.8	72.3	-12203.2	24408.4	2274
Average CD4+ - Line 2	390.6	56.2	0.176	319.5	0.000	55.9	56.6	-21835.0	43672.0	4464
Average CD4+ - Line 3	351.0	51.0	0.385	132.6	0.000	50.2	51.7	-14411.1	28824.2	2631
Average CD4+ - Line NS	516.7	-13.2	2.0	-6.7	0.000	-17.1	-9.3	-2616.1	5234.2	429
Average log10VL - Line 1	4.6	-0.4	0.014	-28.1	0.000	-0.4	-0.4	-1047.8	2097.6	787
Average log10VL - Line 2	4.1	-0.2	0.006	-35.0	0.000	-0.2	-0.2	-1389.0	2780.0	1532
Average log10VL - Line 3	3.8	-0.1	0.006	-11.4	0.000	-0.1	-0.1	-743.5	1489.1	1197
Average log10VL - Line NS	3.5	0.1	0.0	17.8	0.000	0.1	0.1	-203.8	409.5	511

Notes: CI=Confidence Interval, $\log_{10}$VL=$\log_{10}$ HIV RNA copies per mL, NS=Non-Suppressive, SE=Standard Error

Then, either the individual responds to the new regimen and reaches viral load suppression or it does not. If the individual does not respond, $\log_{10}$VL will continue the same average path until virological failure is declared due to lack of response. If suppression is achieved, it is assumed that the $\log_{10}$VL decreases instantaneously and remains constant (at the mean point between zero and the detection level) until a rebound occurs.

The (fitted) distributions of log10VL at first, second and third rebound did not differ significantly[57] in the CHC sample. Consequently, a common "log10VL rebound" distribution was assumed for all virological failures, regardless of the line in which it occurred. The fitted distribution is a Uniform distribution [1.70; 6.24] ($\chi^2 = 53.6, \mathrm{p} = 0.023$, Appendix IV, Figure 44). In non-suppressive therapy, most individuals will have high levels of resistance, this being

[56] In Section CHC of Appendix III, a graphically comparison of observed and fitted values is presented.

[57] The two-sample T test with unequal variance does not reject the null of equal means in either pair considered.

the cause for not achieving viral suppression. The (fitted) distribution of log10VL at "rebound"[58] in non-suppressive therapy, is a Normal distribution, with parameters μ=4.247 and σ=0.790 obtained from 3,757 observations in the LVHEM database ($\chi^2 = 183.7, \mathrm{p} < 0.001$, Appendix IV, Figure 45).

4.5.2.2. *Adherence Level*

Adherence is a broad concept which has recently been discussed and evaluated by Reis [132, 134] in the Portuguese context. Non compliance results in virological failure and resistance development over time [135-138]. It is, thus, a major factor in determining disease progression.

In the present analysis, adherence to pharmacy refill was used as a proxy for adherence to treatment[59]. Obviously, the usefulness of such proxy depends on the validity of the assumption that individuals do take all the given doses. This is a strong assumption, especially among problematic HIV infected individuals, such as drug addicts. Nonetheless, in the work by Reis [134] adherence to pharmacy refill along with general feeling of somatization, and the domain Environment of the World Health Organization Quality of Life Bref questionnaire (WHOQOL-Bref) were found to be the main determinants of adherence.

The most common described predictors of adherence in the literature are individual characteristics, disease and treatment characteristics and social factors. In the HIV context, adherence level is likely to depend on both individual characteristics (age, gender, race, life style, attitude towards infection, social integration, etc.) and on ART regimen characteristics (number of doses per day, number of pills, incidence of adverse effects, etc.).

Although the CHC database did contain more individual characteristics than those included in the model (age and gender), the number of missing values in those variables was significant and they were, therefore, not included. ART regimen characteristics included in the model were the number of pills per day and daily dosage frequency [129]. In addition, we included the regimen number to reflect the possible impact of regimen switches, and the therapy line to reflect disease progression. There is evidence in the literature that adherence tends to decrease over time among HIV infected individuals although a recent study by Cambiano *et al.* [139] suggests otherwise. We have, therefore, included time since ART initiation as an explanatory variable.

[58] Individuals in non-suppressive therapy do not have rebounds, in the sense that viral load does not jump from undetectable to detectable, but new resistances to the current regimen imply jumps in viral load and consequent regimen switches.

[59] A detailed explanation of the methodology used to estimate adherence using pharmacy refill data is provided in Appendix IV.

Given the fact that our adherence variable is a proportion, following Papke and Wooldridge [140, 141], we made use of a generalized linear model with a binomial family and a Logit link on relevant (and available) regressors. Results are presented in Table 26.

The average adherence level in the sample was 0.857 (95% CI: [0.851; 0.864]) and the predicted average is 0.86. In our analysis, we find that ART duration does not have a significant impact on the probability of adherence. We do find that disease progression (as measured by therapy line number) has a significant negative impact on adherence and the accumulated number of distinct regimens has the same effect. The total number of pills per day is not significant at 5% level, but the daily frequency of intake suggests better adherence to once-per-day regimens. Age at ART initiation has a significant and positive impact on adherence, possible reflecting the mode of transmission effect (mean age of diagnosis in the IDU mode is lower than that of the heterosexual class, and adherence is expected to be lower among IDUs). Female gender is also found to be associated with better adherence.

Table 26: Regression Model for the Probability of Adherence to ART

Generalized Linear Models

Optimization : ML		No. of obs=	3,205
Deviance =	690	Residual df=	3,192
Pearson =	734	Scale parameter=	1
Variance function: V(u) = u*[Binomial]		(1/df) Deviance=	0.21619
Link function : g(u) = ln(u/ [Logit]		(1/df) Pearson=	0.22984
Log pseudolikelihood =	-1,020	AIC=	0.645
		BIC=	-25,077

(Std.Err. adjusted for 1,302 clusters in id)

Probability of Adherence	Coef.	SE	z	P>z	[95% CI]		dP/dx	X
Number of Doses per Day (2 is reference)								
1	0.149	0.070	2.1	0.034	0.011	0.287	1.8%	0.249
3	-0.194	0.164	-1.2	0.235	-0.515	0.126	-2.5%	0.025
Number of Pills per Day	0.001	0.011	0.1	0.941	-0.021	0.023	0.0%	5.9
HCV	-0.081	0.069	-1.2	0.242	-0.216	0.055	-1.0%	0.484
Age at ART Initiation	0.011	0.004	3.2	0.002	0.004	0.018	0.1%	36.9
Gender (Female=1)	0.111	0.030	3.7	0.000	0.052	0.170	1.4%	2.7
Number of PIs	-0.102	0.037	-2.7	0.007	-0.175	-0.028	-1.2%	0.973
Regimen Number	-0.128	0.065	-2.0	0.051	-0.256	0.000	-1.6%	0.352
Line (1 is reference)								
Line 2	-0.137	0.063	-2.2	0.029	-0.261	-0.014	-1.7%	0.429
Line 3	-0.237	0.110	-2.2	0.031	-0.452	-0.022	-3.1%	0.129
ART Duration (years)	-0.001	0.016	-0.1	0.943	-0.032	0.030	0.0%	2.0
Monthly ART Cost	0.000	0.000	2.0	0.047	0.000	0.000	0.0%	824
Constant	1.1	0.193	5.8	0.000	0.743	1.5	Predicted mean=0.86	

Notes: ART=Antiretroviral Therapy, HCV=Hepatitis C-Virus, PI=Protease Inhibitors

In the simulation model, individual adherence level at ART initiation is calculated using the regression model, evaluated at the (randomly assigned) individual initial characteristics. This variable is then updated at each event occurrence.

4.5.2.3. *Resistance Level*

In order to parameterize the model regarding resistance over time, resistance level was defined as the sum of the inverted Genotypic Sensitivity Score (GSS) obtained from the REGA 8.0 algorithm. By inverted, we mean that while the GSS measures sensitivity to the drugs, we transform it to reflect resistance. The resistance score was grouped in four classes (R<1; 1<=R<5; 5<=R<10; 10<R<=25)[60]. For simplicity, it was assumed that individuals will sequentially follow the resistance class, that is, no jumps from non-adjacent classes were allowed. Time to resistance class change was estimated using survival analysis and it was included as one additional event in the model. At the time of event occurrence, the new resistance level assigned to each individual is sampled from a Uniform distribution within the class.

4.5.2.4. *AIDS Status*

An AIDS-defining event may occur at any point during HIV infection, and the individual is permanently classified as having AIDS from that point forward. Other episodes of AIDS-defining clinical conditions will occur over the life span of the individual, but the first of these sets the change from asymptomatic or symptomatic stages to the AIDS stage. In the model, the individual will remain in his initial AIDS status until an AIDS event occurs or the CD4+ cell count drops to less than 200 cells per mm^3.

4.5.2.5. *Regimen Characteristics*

The number of PIs reflects regimen characteristics over time. This variable, along with the frequency of daily dosing and the total number of pills per day are updated after each regimen switch or line switch. Upon event occurrence, a random draw from a Table distribution determines the new value for each variable in the new regimen. These Table distributions are conditional on line number.

4.5.3. Clinical Events

Model parameterization required estimating the relation between time to event and individual characteristics and accumulated history. Survival analysis was used to link the time to event to

[60] A full explanation is provided in Section LVHEM of Appendix III. Values used to create class limits are arbitrary although considered reasonable by Dr. Camacho who is in charge of the LVHEM database.

those characteristics, so that during the simulation process, the distribution from which time to event is sampled reflects characteristics and history of the simulated individual.

Seven major groups of events were considered: virological suppression, therapeutic modification without virological failure, line switch, resistance development, hospitalization, AIDS-defining events, and death.

Depending on data availability, explanatory variables considered in the survival analysis were:

1. Individual Characteristics: Age, Gender, Resistance Level, Adherence Level, IDU, HCV Co-Infection, Viral Load, CD4+ Cell Count;
2. Past History: Accumulated Number of Regimens, of Virological Failures, and AIDS status;
3. Regimen Characteristics: Number of PIs, Daily ART Cost and Year of First ART.

Several methods may be used to (randomly) determine timing of activities: (i) Determine event first, then time, (ii) Determine time first, then event or (iii) Sample times for each possible event and use the shortest [3]. In the latter case, which is the approach followed in the BRAM model and in the present analysis, a survival curve is required for each event, then a time is sampled for each event and the earliest time determines which event occurs next. Other events are taken as censored and may be discarded. This method has the advantage that the individual survival curves for the events can be calculated independently of each other. Nonetheless, when sampling "time to next occurrence", it is necessary to sample from conditional survival distributions, allowing for time already spent 'at risk' to be considered.

In the case of the Weibull distribution, such conditional survival distributions are easily handled due to the link between the Weibull and the Exponential distribution. A random variable X has a Weibull distribution with shape parameter p and scale parameter $\boldsymbol{\lambda}$ if $\left(\frac{\mathrm{x}}{\lambda}\right)^{\mathrm{p}}$ has an Exponential distribution with mean 1. The Exponential distribution has the property that the conditional distribution of time to event remains constant as long as no event has occurred. As such, to sample "time to next occurrence" (of a given event), let t be the time already spent at risk and μ be a value sampled from an Exponential distribution with mean 1. Then, time to next occurrence of a given event (v) is given by:

$$v = \lambda\left\{\left[\left(\frac{t}{\lambda}\right)^{p} + \mu\right]^{\frac{1}{p}}\right\} - t$$

In the present analysis, a Weibull distribution was used to parameterize survival curves for each event in the model. Obviously, the Weibull distribution with shape parameter p and scale parameter λ generalizes the Exponential distribution so that if $p = 1$ it equals the Exponential distribution, if $p < 1$, the risk decreases over time, and if $p > 1$, the risk increases over time.

Table 27 provides the estimated marginal effects of each covariate on median time to event occurrence as well as the estimated shape parameter, the source of information used, the median time to event predicted by the model (at the estimation sample mean values of the covariates) and the Kaplan Meier predicted median time to event. Full results of each regression are presented in Appendix V.

Table 27: Estimated Marginal Effects on Median Time to Event Occurrence and Adherence Level.

Event	Viral Suppression (Line 1&2)	Viral Suppression (Line 3)	Virological Failure (Line 1&2)	Virological Failure (Line 3)	Switch Without Virological Failure (All lines)	First Resistance	Resistance class switch	Hospitalization	AIDS	Death	Adherence Level
Predicted median* (months)	6		80	36.5	41	338	86	90		If AIDS: 116	Predicted Mean=0.86
KM estimated median (months)	7 (95% CI:[6;8])		72 (95% CI: [67;80])	31 (95% CI: [29;34])	35 (95% CI:[32;39])	Not Reached	87 (95% CI: [85;89])	72 (95% CI: [64;82])		If AIDS: 126	Sample Mean=0.86
Gender (Female=1)	-0.554		**7.653**	0.643	-2.702	8.59	-0.486	11.149		**1,523**	1.4%
Age	**-0.111**		0.559	**-0.080**	-0.079	**-5.18**	**-0.555**	**-2.502**		**-153**	0.1%
$Log_{10}VL$	**0.951**		**-19.563**	**-1.550**	**1.321**	**-84.45**	**-3.249**	**-18.216**			
CD4+	0.000		**0.026**		0.003			**0.146**			
HCV	**-1.680**		10.094		**-5.223**	-26.00		-32.008			-1.0%
Adherence	**-0.062**		0.184		**-0.940**	**2.91**		0.424			
Year 1st ART	0.022		2.448	**-5.188**	-1.189	-22.29	**0.020**	1.295			
Number of PIs	0.297		**-6.120**	-0.618	**-5.678**	-33.89	**6.850**	-3.453			-1.2%
Resistance Level	**-0.847**		-1.894		0.269			-7.927			
Regimen Number			**-8.963**	0.239	**1.071**	3.99	**-6.310**				-1.6%
Line (1 is reference)											
2	0.485		-10.658		-2.836			**-34.500**			-1.7%
3					-1.444			**-64.061**			-3.1%
Resistance Class (1 is reference)											
1<=R<5				**-6.225**							
5<=R<10				**-7.348**							
R>=10				**-7.021**							
Year of Diagnosis										**660**	
Transmission Group (Heterosexual is reference)											
IDU										**-3,692**	
Homosexual										**-900**	
Other										**-2,300**	
HIV Stage (Non-AIDS is reference)											
AIDS										**-18,143**	
p	1.030		**1.447**	**1.873**	**1.499**	**1.501**	**1.524**	**0.733**		**(stratified by AIDS status)**	
Number of Doses per Day (2 is reference)											
1											1.8%
3											-2.5%
Number of Pills per Day											0.0%
ART Duration (years)											-0.01%
Monthly ART Cost											0.00%
Source	CHC	Clinical Trails (POWER 1 &2)	CHC	LVHEM	CHC	CHC	LVHEM	CHC	May et al. (2004)	CVEDT	CHC

Note: *At estimation sample mean values. Coefficients significant at 5% are marked in bold and shaded. Coefficients significant at 10% are shaded but not bold. KM=Kaplan Meier, ART=Antiretroviral Therapy, $Log_{10}VL=Log_{10}$ HIV RNA copies per mL, PI=Protease Inhibitors, Resistance=Inverted Genotypic Sensitivity Score based on REGA 8.0 Algorith, AIDS= Acquired Immune Deficiency Syndrome, HIV=Human Immunodeficiency Virus, IDU=Injection Drug User, HVC=Hepatitis C-Virus

4.5.3.1. *Virological Suppression*

Time to virological suppression was only available in the CHC sample. That sample was used to estimate the time span from therapy line initiation to the first undetectable viral load registry.

Time to viral suppression is expected to depend on disease stage and, namely, on the previous number of virological failures. Nonetheless, the log rank test for equality of survival functions (of the event "time to viral suppression") does not reject the null of equal survival functions among naive and non-naive individuals in the CHC sample ($p = 0.083$) and we have therefore used a common regression for lines 1 and 2. In the CHC sample, median time to viral suppression among naive individuals is 7 months and the median predicted by the model is 6. In this regard, it should be noted that while current guidelines indicate that virological failure should be declared if viral suppression is not reached within 6 months, 12 months was the threshold considered for the most part of our period of analysis. Given this fact, we have calibrated the model to allow for a 12 month period for viral suppression.

The shape parameter is not statistically significant different from one, indicating a constant rate over time. Viral Load, Adherence Level and Resistance Level are all statistically significant and have the expected signs. Co-infection with hepatitis C is found to decrease median time to viral suppression, which is unexpected especially because the prevalence of HCV co-infection, among HIV infected individuals, is highest among injection drug user. We do not find evidence that more recent therapies are more effective at reaching viral load suppression in a shorter time, which is likely due to period of analysis and the weight of no longer considered first choice drugs.

Line 3 in the model reflects patients at a symptomatic stage of infection; individuals who most likely have developed resistance. Therefore, the proportion of patients not reaching viral suppression is expected to be significantly higher relatively to previous lines. Given such characteristics of the patients, the LVHEM database should, in principle, be used as an estimate. Unfortunately, that database does not contain information on the timing for viral suppression (focusing instead on virological failure); consequently, occurrence of this event in the model was obtained from the international literature and is, thus, not conditioned to individual characteristics of the individual or past history.

Data from the POWER 1 and 2 clinical trials [142, 143], was used to parameterize the model. The POWER trails compare a recently introduced protease inhibitor (darunavir) to currently available protease inhibitors ("other PIs") in treatment of experienced patients. Patients included in the POWER trials, had at least two virological failures on regimens containing PIs and had characteristics similar to those modeled in line 3. The proportion of viral suppression in that arm was 45%, compared to only 10% in the "Other PIs" arm, so the darunavir arm was used to reflect the expected progress in ART therapy.

4.5.3.2. *Regimen Switch without Virological Failure*

While all line switches imply a regimen switch (due to virological failure), within the same therapy line, several regimen switches may occur. Treatment modification is here defined as changing at least one of the drugs in the regimen, adding a new drug or withdrawing a current one.

These switches are not associated with a virological rebound, and do not, therefore, originate a line change in the context of the model here developed. Because the consequences in terms of disease progression are different in nature from those switches generated by virological failure, that event was modeled separately.

Regimen switch without virological failure may occur for several reasons and, according to the literature, these reasons account for the majority of treatment modifications [144]. Information on regimen switches was available in both the CHC and the LVHEM samples, but the reason for switching was not clear. Treatment modification due to intolerance (serious adverse events causing regimen modification) is of special interest as it has become one of standard endpoints in clinical trials [145]. Given the impossibility to identify the cause of regimen switch due to reasons other than virological failure, it was assumed that the cause of regimen switch without virological failure was proportional to that observed in the Swiss Cohort [144][61]. It was, hence, assumed that 54% of those regimen switches occurring without virological failure were due to toxic effects, while the remaining 46% were due to other reasons. With this assumption, the model will reproduce the observed number of switches but it may not be correctly assigning costs and quality of life to those switches, since both are likely to depend on the cause of switch.

The log rank test for equality of survival functions (of the event "time to regimen switch without virological failure") does not reject the null of equal survival functions among naive and non-naive individuals in the CHC sample ($p = 0.826$). Moreover, switches without virological failure are likely to be unreported in the LVHEM database. Consequently, it was assumed that the distribution estimated from the CHC sample holds for all lines. Different values for the covariates reflecting characteristics in each line do, nevertheless, obviously generate different sampling distributions

The accumulated number of virological failure (i.e. the therapy line) has an impact on the probability of viral suppression and on the probability of virological failure, but it is not a priori

[61] A comparison of Elzi *et al.* results with those of the CHC sample is provided in the Section CHC of Appendix III .

clear how and if it has an impact on time to regimen switch due to reasons other than virological failure. Therapy line number was therefore included as a regressor to identify whether this distribution differs between therapy lines.

The Kaplan-Meier estimated median time to regimen switch without virological failure was 35 months ($95\%\ CI$: $[32;\ 39]$) and the predicted median is 6 months longer.

No evidence was found that time to regimen switch without virological failure varies with therapy line but the accumulated number of different regimens is statistically significant indicating that, patients who have switched a higher number of times will have a longer time to the next regimen switch. As expected, the higher the number of PIs and the adherence level, the shorter the time to regimen switch due to (among others) toxic effects.

4.5.3.3. *Line Switch - Virological Failure*

A line switch event is defined as a ART regimen modification as a result of virological failure. Virological failure may occur for several reasons; it may be due to non-adherence, to resistance development or to other reasons. Non-adherence and resistance are also linked, since, over time, one may cause the other. Reasons for virological failure were not available and are not directly modeled, instead a unique event "virological failure" is considered regardless of the cause.

It should be noted that the event modeled is regimen switch preceded by virological failure, that is, the time of event considered is the data of regimen switch and not the time of actual virological failure. While in principle, these two events should be close, unpublished work by Dr. Camacho suggests that a significant lag between the two exists.

Our results indicate no statistically significant difference in time to line switch among naïve and non-naïve individuals at 5% level although it is significant at 10%. To this regard, it should be noted that when considering the subsample of naïve individuals initiating therapy with the currently recommended regimen composition (2NRTI+NNRTI or 2NRTI+PI/r), time to line switch does differ among naïve and non-naïve individuals. The Kaplan-Meier estimated median time to regimen switch without virological failure was 72 months ($95\%\ CI$: $[67;\ 80]$) and the predicted median is 80 months.

In lines 1 and 2, gender has an impact on time to regimen switch due to virological failure, with longer time for virological failure among female HIV infected individuals. As expected, Log_{10}VL is found to increase the probability of the same event while adherence (significant at 10%)

decreases that probability. The number of PIs in the regimen has a negative marginal effect, which is in line with results published in the literature [146]. We find the same effect with regard to the accumulated number of regimens.

Resistance level is not statistically significant, which is an unexpected result since resistance is known to be a major determinant of virological failure. This result suggests that in the CHC sample, virological failure is mostly been driven by non-adherence or lack of viral response. One other explanation would be that those assumed not to have resistance (those without a resistance test), have indeed developed it but a resistance test is not available.

Indeed, in the estimation based on the LVEHM (used to model time to event in line 3), time to line switch is predicted to be shorter among individuals with high resistance levels. In that regression, significant coefficients have the expected sign but the difference between the Kaplan-Meier estimated median time to line switch and that predicted by the model (6 months higher), suggests a lack of fit. Median predicted time is less than half of that estimated for lines 1 and 2, reflecting cross-resistances and increased inability of currently available ART to contain viral replication as the panoply of susceptible drugs is reduced. The hazard rate is, as in earlier lines, increasing in time.

4.5.3.4. *Resistance Development*

Excluding those for whom information on ART therapy regimens is not available, the LVHEM provides a sample of individuals who have some positive level of resistance (and 76% are resistant to at least two drugs). This sample is, thus, a useful source of information in terms of characterizing disease progression amongst those who develop resistance. It cannot, on the other hand, be used to parameterize the model with respect to first resistance development[62], and that information was, therefore, obtained from the CHC sample.

Regarding first resistance development, adherence and viral load, which are the a priori main expected drivers of resistance development, the estimated coefficients are both significant and have the expected sign. An increase in adherence leads to longer time to resistance development while an increase in viral load has the opposite effect. Rigorously, no predetermined sign should be expected on the adherence marginal effect because the relationship between adherence and resistance is not a linear one and depends on the class of drug under analysis [135, 147-150]. Age is also found to be a determinant of both time to first

[62] The Kaplan-Meier estimates of time to first resistance development in each database is provided in the Section LVHEM of Appendix III

resistance development and to resistance class switch after the first resistance has occurred. The reasons for such are unclear.

Time to resistance class switch once the first resistance has developed was estimated on the LVHEM sample. In that sample we do not have information on adherence, consequently such relationship could not be estimated. Predicted median time to resistance is much longer for the first resistance than for when some level of resistance has already been developed (28 years versus 7 years) likely reflecting cross-resistance and cumulative exposure. An increase in viral load or in the accumulated number of regimens[63] reduces the median time to event. The year of ART initiation has a slight positive effect on decreasing the probability of resistance possibly capturing advances in ART. The number of PIs also has a protective effect, likely reflecting the documented more frequent emergence of drug resistance in the NNRTI class [151].

When the individual reaches non-suppressive therapy, he will be in the highest class of resistance (see Section 4.4). In order to account for the fact that resistance will eventually increase within class 4, we have used the model in Table 58 to sample time to resistance increase within line 4. Upon such event occurrence, the new resistance level is assumed to be the maximum between the current level of resistance and a random draw from a distribution fitted on resistance levels above 10 (among individuals in line 4 of the LVHEM sample). The best fit, when bounding the lower limit to 10, was a Triangular distribution between 10 and 22.5 and 20.75 as the most likely value ($\chi^2 = 12, \mathrm{p} = 0.213$, Appendix IV, Figure 46).

4.5.3.5. *Hospitalization*

The CHC database contains information on all inpatient care episodes and the corresponding Diagnoses Related Group in which the episode was classified. Duration of episodes is not known since the only information available is the discharge date. We use the discharge date as a proxy for the date of event occurrence. In defining the event, we have not considered those hospitalizations occurring right before ART initiation since these are related to the absence of ART (late diagnosis, late treatment initiation, etc).

The Kaplan-Meier estimated median time to hospitalization was 72 months ($95\%\ CI$: $[64; 82]$) and the predicted median is 90 months. As expected CD4+ cell count has a postponing effect on

[63] Which in that database is close to the accumulated number of line switches since switches due to reasons other than virological failure are likely to be sub-notified.

hospitalization events while viral load and age have the opposite effect. Median time to hospitalization is decreasing in therapy line reflecting disease progression.

4.5.3.6. *AIDS-defining Event*

As the CD4+ cell count decreases, the immune system's ability to react to infections deteriorates. Infections that, in a non-HIV infected individual, would be of no major consequence become serious events and for that reason they are denominated opportunistic infections (OI). Many OI, along with cancer related diseases are included in the list of AIDS-defining conditions established by the Center for Disease Control. In the symptomatic and AIDS stages of the disease, these conditions become more common, eventually culminating with the inability of the immune system to respond to such threats.

Bezerra and Roxo [152] report the number and type of OI in a sample of 185 HIV infected individuals on ART therapy, while Vieira *et al.* [153] report the number and type of AIDS-defining illnesses occurring in HIV infected individuals followed in a Portuguese hospital, between 1995 and 2003. Nonetheless, neither these references nor the CHC database provide enough information to allow for the estimation of time to AIDS.

Consequently, time to AIDS-defining conditions was parameterized with the prognostic model for survival of HIV positive patients treated with antiretroviral therapy, published by May *et al.* [154, 155] and later validated by the same authors on the CASCADE cohort [154]. May *et al.* use data from the ART Cohort Collaboration, to examine rates of progression to AIDS or death among HIV-1 positive patients[64]. The authors find that the best was a Weibull proportional hazards model stratified on CD4+ and transmission risk.

The authors considered the probability of progression to a combined endpoint of a new AIDS defining disease or death from any causes. There were 870 new AIDS events and 344 deaths. In our analysis, we would be interested in modeling the AIDS and death events separately, in order to be able to use Portuguese mortality data. Given this limitation, we parameterized the model under the assumption that all deaths in the May *et al.* model were preceded by an AIDS event. This fact is likely to overestimate the number of AIDS events, and the model should surely be improved on this aspect, once data becomes available.

[64] AIDS events were not considered as a separate endpoint because information on causes of death did not allow for the distinction between deaths due to AIDS-defining conditions and deaths from other causes.

4.5.3.7. *Immunological Failure*

Immunological failure is defined as a decrease in CD4+ cell count below the initial CD4+ at ART initiation. In non-suppressive therapy, when viral suppression is, by definition, not possible due to resistance to available drugs, maintenance of CD4+ cell count levels becomes the major goal of therapy. Occurrence of immunological failure is, thus, a relevant event in non-suppressive therapy. Because the evolution of CD4+ cell count over time is explicitly included in the model, this event was not considered per se, and it occurs as a result of the modeled path for CD4+ cell count.

4.5.3.8. *Mortality*

Mortality data was available from two distinct sources: the CVEDT dataset and HIV-related aggregate mortality data provided by the National Institute of Statistics (INE). While the first database accounts for all deaths among HIV-infected individuals (without specifying the cause of death), the second refers to deaths attributed to HIV. In spite of this fact, the number of deaths per year is significantly higher in the INE database. This difference may either be due to non-notification of death among those in the CVEDT database, or it may be that the vital status is correctly registered in the CVEDT sample, but not all HIV infected individuals are in the database[65].

Time to death was parameterized with a Weibull distribution stratified by disease stage, using the CVEDT database[66]. Although 3 disease stages (Asymptomatic, Symptomatic and AIDS) are currently the norm, the CVEDT database was constructed on the previous – AIDS/non-AIDS classification[67] and that was, therefore the strata used in the estimation. For the purpose of estimation we have dropped individuals less than 13 years old and those that were diagnosed after death (since those individuals did not initiate ART therapy).

The Kaplan-Meier estimate of median time to death in the AIDS group is estimated at 9.7 years ($95\%\ CI: [8.1; 12.1]$), and the model predicts a median time to death, among individuals diagnosed with AIDS, of 10.5. Median time to death in the non-AIDS group has not been reached and the predicted median is beyond that of the general population life expectancy. All

[65] A detailed comparison is provided in Section CVEDT of Appendix III.

[66] The INE database could not be used to estimate time to death, since it provides only aggregate information on the general population, deaths in the general population and deaths due to HIV. Information on the number of infected individuals, date of diagnosis or any other characteristics is unavailable.

[67] Only 10% of the observations are classified in the Symptomatic stage.

explanatory variables are highly significant and have the expected sign. IDU mode of transmission yields shorter time to death, as does age at diagnosis.

In order to account for the fact that mortality due to age may not yet be correctly reflected in the CVEDT database[68], the model assumes that the estimates provided by the CVEDT refer to HIV related deaths, and deaths due to other causes are included as a competing risk. It may be argued that death due to other causes is being double counted since CEVDT deaths already include deaths due to other causes. Since it is not possible to identify the cause of death, it is not possible to correct for double counting. Nonetheless, given the possible sub-notification of deaths in the CVEDT database suggested by the comparison with the INE data, we expect the double counting not to be of major relevance.

Parameterization of the model, regarding death due to other causes, was based on Portuguese general population mortality rates in the 2006-2008 period published by the INE [156]. The sampling process follows Barton *et al.* [3] and is described in the Section INE of Appendix III.

4.5.4. Quality of Life

Life expectancy of people with HIV is approaching that of people without the disease [100, 101, 157] suggesting that HIV infection is a approaching a chronic disease status with more impact on quality rather than quantity of life. In this context, quality of life parameters are of special relevance in economic evaluation studies of HIV treatment. As such, we start this Section with an overview of the methods available, move to a brief literature review of published HIV related quality of life estimates and, from there, we present the approach followed in the present analysis.

Due to the interference of "non-health" aspects, such as stigma and discrimination, on HIV infected individuals' quality of life, we use discrete choice experiment (DCE) to gain insight over the weight of such non-health related attributes. We do so by constructing and applying a questionnaire to a sample of 100 HIV infected individuals in Portugal. Multiplying the estimated utility of each attribute by the level of each attribute[69], we obtain a quality of life index which is used in the model to obtain quality of life adjusted years.

[68] A detailed discussion is provided in Section CVEDT of Appendix III.

[69] As explained in detail in Section 4.5.4.6., the level of each attribute was obtained from published references (using the World Health Organization Quality Of Live HIV instrument). Some attributes depend on CD4+ cell count while others, such as discrimination, as assumed constant over time.

Because, as discussed in detail below, DCE estimates do not generate QALYs, in the model we include not only the above mentioned index but also two distinct published utility estimates for HIV infected individuals based on the EuroQol quality of life instrument (EQ-5D) instrument (one based on a Portuguese sample and another based on an international sample).

4.5.4.1. *The QALY Framework*

The overall aim of economic evaluation is to aid decision makers to make efficient and equitable decisions, by comparing costs and benefits of health care interventions [158]. While it may seem obvious how to assign a cost to the direct resource consequences of a health intervention and even the less direct ones, the health consequences are less tangible and, hence, it is less obvious how to assess them [159]. Usually, treatment outcomes are scored according to the number of quality-adjusted life years (QALYs) that they yield.

In the QALY framework outcomes are measured in a common unit - a year in full health – encompassing both quality and quantity of life in a single score, thus allowing for comparisons among different populations, illnesses and interventions. Scores range from 0 (worst) to 1.0 (best) and are determined by having the respondent value health states against an external metric such as risk (usually of death) or time. These scores, thus, reflect society's, or the patients', willingness to exchange quality of life for survival. The QALYs obtained by each individual are calculated as a weighted average of years lived. The weights used represent the utility assigned to the health status, of each lived year. The total number of QALYs resulting from an intervention is, in turn, a sum of the QALYs obtained by each individual taking part in the intervention. Once estimated QALYs can be compared to costs, in the form of an incremental cost-effectiveness ratio (ICER), and comparisons across interventions can be made, thereby informing decisions as to whether an intervention can be considered value-for-money.

Some criticism has been presented with respect to the QALY [160]. One comment is that the model is based on expected utility theory, and there is evidence that expected utility theory is not valid as a descriptive theory of Decision making under risk [161, 162]. Another mentioned limitation is that QALY techniques focus solely on health outcomes [158, 163], and there is often a need to evaluate interventions that seek to improve an individual's quality of life beyond health. For example (complex) public health interventions seek to impact on broader aspects of quality of life, not just health, but also non-health outcomes such as discrimination, participation, amongst others, and QALYs are likely to underestimate the true benefits [164].

Sen's capability approach [165, 166] could provide a possible solution to the limitations discussed above, in that it expands the evaluation space to consider whether a program enhances an individual's capability [167]. In an attempt to operationalize the approach, Coast and colleagues are in the process of developing an index for capability using best-worst scaling within a discrete choice framework, specifically for use in the elderly [168, 169].

4.5.4.2. *Available Methodologies*

There are two major families of utility measures: direct and multi-attribute [158]. In the first, elicitation techniques are used to ask individuals to evaluate health states. Valuation techniques are related to measuring the patient's (or the community's) value for a specific component or attribute, either in absolute terms, or in relation to another attribute. The relative importance is identified by choices that inevitably require trading off one or more desirable outcomes (including price/co-payment, time, etc.) in a given area (or domain), in order to obtain a more desirable composite outcome. The most common elicitation techniques are visual analogue scale (VAS), standard gamble (SG) and time trade-off (TTO).

The indirect, or multi-attribute, approach is comprised of two elements: a questionnaire on self-assessed health status covering the different dimensions of health related quality of life, and a scoring algorithm reflecting general population values for the health states described by the instrument. To obtain the scoring, the health state values or utility weights for a sample of health states described by the instrument are typically derived from an adult general population sample using conventional valuation methods such as SG or TTO. Statistical modeling methods or multi attribute utility theory models [170, 171] are then employed to generate a scoring algorithm, which allows the estimation of utility weights for all possible health states defined by the instrument [159]. The most commonly used health preference measures, in this context, are the Health Utility Index, the Quality of Well-Being, and the EuroQol.

While in the direct method health states are valued by the individuals whose health is being assessed, in the second, health states are valued by a distinct sample of individuals. Nevertheless, in both cases stated preferences methods are employed to place a utility weight on each of the relevant health states being considered.

Stated preference techniques include the commonly used methods of preference weighting (e.g., rating scales, standard gamble, and time trade-off), contingent valuation and multi-

attribute valuation (conjoint analysis and choice modeling). Merino-Castelló [172] provides a detailed description of the main differences between contingent valuation, conjoint analysis and choice modeling[70]; Louviere *et al.* [173] discuss the difference between conjoint analysis and DCE[71] and a discussion of the methods of preference weighting, in general, may be found in Drummond *et al.* [158] or Dolan [160].

4.5.4.2.1. Indirect Methods

In terms of indirect methods, two factors ought to be addressed: instrument selection and source of preferences. The first issue relates to the utilization of a generic instrument (SF-6D, EQ-5D, HUI3, etc) versus a disease specific instrument (MOS-HIV, MQoL-HIV, WHOQOL-HIV, etc.). Generic measures are necessary to compare outcomes across different populations and interventions, particularly for economic evaluation. Disease-specific measures assess the special states and concerns of diagnostic groups and may thus be more sensitive for the detection and quantification of small changes that are important to clinicians or patients [174].

Regardless of the instrument used, valuation of health states requires the quantification of the relative weight of each dimension (and of each level on each dimension) in relation to some external metric. In general, these weights result from the application of TTO or SG on a random sample of the general population (possibly from another country), thus reflecting community preferences, instead of patient preferences.

The argument in favor of patient preferences is based on the notion that patients understand better the impact of the disease, and their valuation of health states includes the human being's ability to adapt. Empirical evidence suggests that patients' valuations are higher than those of the general population [175] and the HIV domain is no exception [176]. Schackman *et al.* [107] report that patient SG utilities (derived from a power transformation of rating scale) were higher than community SG utilities by 4% to 9%. Joyce *et al.* [177] report a weak correlation between average values obtained by undirected methods (based on community preferences) and those obtained by TTO e SG on a sample HIV infected individuals. In the

[70] In short, contingent valuation is a direct survey approach where respondents are asked to express their maximum willingness to pay for a hypothetical change in the level of provision of the good. Contingent valuation analyzes one attribute of the product at a time while multi-attribute valuation explores more than one attribute simultaneously. Within multi-attribute valuation, conjoint analysis requires the individual to rate or rank each alternative, while choice-based approaches ask the consumer to choose one of several alternatives. Conjoint data is typically analyzed using ordinary least squares (OLS), while choice-based approaches use the random utility function that represents the integrated behavioral theory of decision-making and choice behavior. The latter is composed of a deterministic component V_{ij} and of a stochastic one e_{ij}, and is usually estimated by maximum likelihood.

[71] According to Louviere et al. (2000), the theoretical framework of DCE is the main feature distinguishing DCEs from traditional conjoint analysis. "*Conjoint analysis is a theory about the behavior of sets of numbers in response to factorial manipulations of attributes which eventually may allow studying how holistic preferences for combinations of attribute levels are. DCEs are based on a sound, well-tested, integrated behavioral theory of decision-making and choice behavior: the random utility theory (RUT)".*

literature review performed by Tengs and Lin [176], the authors report that patients' preferences were used in 65% of the articles.

As discussed in Gosling [178] "*the encounter with the HIV/Aids virus may result in despair, hopelessness and ultimately death, or it may offer the opportunity of possible transformation, an opportunity to embrace a life lived productively and meaningfully with a chronic but manageable disease*". In fact, many patients with HIV/AIDS indicate that life is better post-diagnosis than it was pre-diagnosis [179-181]. There is evidence to support the idea that adaptation of patients with HIV/AIDS is manifested through patients' responses to both TTO and SG measures. Specifically, a substantial portion of patients seem to be unwilling to accept a (hypothetical) opportunity to live shorter-but-healthier lives or to take a risk of death in exchange for perfect health, a phenomenon that is referred to as ''*the will to live phenomenon*'' [179, 182]. Results reported by Kudel *et al.* [182] also indicate a greater disposition to gambling than to trading time in some classes of individuals.

4.5.4.2.2. Direct Methods

Although there is no uniformly accepted preferred method for eliciting health state values for the estimation of QALYs, the most commonly used methods are the choice-based valuation methods of SG and TTO [159]. The VAS method has been criticized for its direct and choice-less nature [183] as well as for the fact that data obtained through VAS may be subject to end point and context bias [184-186].

SG, which is founded on expected utility theory of decision making under uncertainty, has been argued to have the most rigorous theoretic foundation [160], while TTO is seen as a less complex alternative to SG because it overcomes the problems of explaining probabilities to respondent [187].

Some of the limitations of these two most commonly used methods include theoretical arguments, the significant cognitive burden and the difficulty to include "non-health" and "process" aspects.

First, with respect to theoretical arguments it has been argued that SG valuations may be influenced by factors other than a person's attitude towards the health state, including probability weighting and loss aversion [159, 188-190]. By the same token, there is evidence to suggest that duration effects and time preference effects can have an impact on the elicitation of TTO values [191-193]. Overall, none of these techniques simply reflect people's preference

for health states, instead including attitudes towards risk (in SG) or time preference (in TTO) [170].

Secondly, concerning the practical implementation of these methods, both methods place a considerable cognitive burden on respondents since respondents are asked to successively compare full health to a series of health states, until the point of indifference is found. Moreover, this cognitive burden usually implies the use of trained professional interviewers and may be difficult to apply to specific populations.

Thirdly, there is evidence to suggest that non-health related aspects of treatment receiving process are also important to individuals [194-196]. For the SG and TTO methods to represent utility fully, preferences over health and non-health attributes must be assumed independently [160] or else a significantly increase in the number of questions, (and in the complexity of each) would be required.

As an attempt to overcome such limitations, there has been an increasing interest in using ordinal approaches to estimate cardinal values for health states to calculate quality adjusted life years [159]. Such ordinal approaches include discrete choice experiments (DCE), which can be used to obtain data to estimate utilities of individual attribute levels (and their interactions, given larger designs), or the utility of a profile.

Several empirical studies have reported similar health state values using ordinal and cardinal approaches [197-199] while recent work by Ratcliffe *et al.* [200] found evidence that values estimated by DCE and ranking do differ from those obtained by TTO.

4.5.4.2.3. Discrete Choice Experiment (DCE)

DCE is an attribute-based survey method which has its theoretical basis in random utility theory [201, 202]. Based on Lancaster's economic theory of value, DCEs assume that individuals derive utility from the attributes of the commodity and that individuals' preferences are revealed through their choices [203]. In accordance with economic theory, it is assumed that the individual chooses the alternative with the highest level of utility. In reality, the individuals do not always make the optimal choice and/or their choice is based on factors unobservable by the researcher. Hence, the problem is inherently stochastic from the observer's view, which leads to the formulation of expressions for the probability of choice.

The random utility theory can be considered a more realistic representation of preferences; however, distributional assumptions about the random component are required in order to

make any predictions from the theory. It is usually operationalized in conjunction with a valuation function that relates the mean utility for a given health state with a set of explanatory variables [203, 204]. This operationalization of DCEs is usually conducted within a conditional (multinomial) logistic regression model and its generalizations [202].

Internal validity and convergent validity has been demonstrated with respect to standard gamble and willingness to pay [205]. The technique has been shown to be relatively insensitive to both the ordering and levels of attributes [206]. Theoretically, three key axioms underlying the technique— completeness, stability, and rationality—have been investigated with encouraging results [207-209].

DCEs have the advantage of allowing for analysis of preferences for complex multi-attribute goods and enable the quantification of the individual trade-offs between the different dimensions. DCEs are, thus, an appropriate framework for analyzing the impact of health and non-health attributes on health related quality of life.

Choice experiments not only give welfare consistent estimates [172], but they also provide information on the marginal rate of substitution between attributes, a key concept for economic analysis and efficient allocation of goods[72]. Moreover, DCEs may be implemented by self-administered questionnaires due to the low level of cognitive burden.
The main disadvantage of the DCE approach is that it estimates a latent health state utility value, but with arbitrary anchors whereas QALYs require health states to be valued on the full health (one) and being dead (zero) scale.

In recent years, different approaches to the issue of anchoring the utility scale obtained from DCE on a full health-dead scale have been published in the literature.

King *et al.* [210] include two survival attributes (duration and uncertainty) in the DCE. By using survival time as the matrix for measuring the welfare impacts of the change in health state, the authors argue that this approach can be used to determine the QALY weights. Nevertheless, because survival has a multiplicative relationship to QoL, its inclusion requires a far larger design and consequently longer questionnaires [211, 212].

[72] The efficient allocation in the Pareto sense for two private goods requires (1) to be on the frontier, (2) that MRS in consumption is the same for A and B (3) that the MRS equals the marginal rate of transformation. In the case of one public good and one private (1) is the same, (2) may differ as both consume equal amounts of the good (3) sum of MRS=MRT.

Ratcliffe *et al.* [200] re-scaled according to the highest and lowest predicted TTO values (corresponding to the best and worst SQOL health states, respectively) and found that this method yielded biased predictions relatively to TTO values.

Brazier and Cain [213] include a death state in the pairwise choice set and re-scale by dividing ß coefficients on each dimension level by the coefficient for being dead. This option may not be suitable for mild conditions, where the worse scenario would not be considered worse than death by any responder, as in Ryan *et al.* [214]. Moreover, as pointed out by Flynn *et al.* [212], the use of a statistical model such as conditional (multinomial) regression to anchor quality of life values from ordinal data to death, is inappropriate in the presence of respondents who do not conform to the assumptions of conventional random utility theory, namely those who think some states are worse than dead. For those individuals, their data tells us nothing about their strength of preference for QoL compared to quantity of life

While research continues on ways to obtain cardinal values from ordinal data, the adequacy of DCE methods with its current limitations will be context specific, depending on the aim of the research.

4.5.4.3. *Literature Review*

Utility estimates in the HIV research have aimed at a wide variety of themes. Reported utility estimates include the valuation of specific aspects, like the impact of HIV diagnosis [180], the impact of adverse events [215, 216], the impact of co-infection HIV/hepatitis [217], the impact of AIDS-defining events (ADEs) and non-AIDS serious adverse events (SAEs) [218] or the potential health gains from Kaposi sarcoma treatment [219]. Reported utility estimates also include valuation of health states, usually by means of several instruments simultaneously [220].

Within this last topic – valuation of health states – several papers have been published recently, suggesting that the mean valuation of health states (as measured by TTO and/or SG) significantly varies with the health state of the respondents [177, 221, 222], or even with a mixture of health state and attitudes of the respondents [179, 182]. This variability in health state valuations demonstrates that imposing the QALY model restrictions distorts valuations [161].

As reported in Tengs and Lin [176], there is a significant variation in published mean preference-based scores for HIV infected individuals. In the meta-analysis of HIV/AIDS utility estimates

performed by those authors to explain the considerable variation in the estimates reported, 22% of the utilities in the sample were elicited with the TTO method, 14% with the SG method, 22% with the rating scale method, 18% with the QWB Scale, and 27% using judgment. The authors find a positive significant impact of disease stage category (Symptomatic HIV) and a negative and statistically significant impact of Rating scale and respondent type (Nonpatients). Upper and lower bound scale label did not have a statistically significant impact on estimates reported. More recent articles report values ranging from 0.73 to 0.88, depending on the sample and instrument used [177, 220, 221]. Joyce *et al.* [177] compared scores obtained from direct and indirect methods in the same sample and found values ranging from 0.59 (HUI3) to 0.80 (TTO), although direct methods did not always yield the highest values. A strong correlation was found between indirect methods, but a weak correlation was found between scores obtained from direct and indirect methods.

The vast majority of the DCE applications in health economics have focused on the relative weight of different attributes of health care (including non-health and process attributes), rather than on valuing health *per se*. Some of these papers include price as an attribute in order to estimate the willingness to pay for each attribute. In addition, the technique has been applied to address a wide range of issues, including estimation of benefits within health technology assessments; analysis of patient/consumer and professional decision making; and developing prioritization frameworks (see [223-225] for a literature review of applications of DCEs in the field of health economics). In the area of HIV, Beusterien *et al.* [226] analyzed preferences with respect to HIV medication, Phillips *et al*. [227] estimated preferences for different attributes of HIV testing and Lee *et al.* [228] analyzed preferences with respect to hypothetical HIV vaccines.

A few studies have used DCE to estimate the values of different health states. Hakim and Pathak [229] compare rating scale and standard gamble with discrete choice modeling (including death in each set) for measuring EuroQol health state preferences, McKenzie *et al.* [230] estimated weights for asthma symptoms and Ryan *et al.* [214] estimate quality weights for social care of the elderly. In these studies, a value of zero is assigned to the worse case and a value of one to the health state with the highest level on all attributes, thus providing program-specific weights that do not allow for the comparison between different programs.

4.5.4.4. *Methodology selection*

In the absence of a gold standard method for valuation of health states, the selection of the method to use was based on the specific characteristics of the HIV infected population, on the resources available and on the aim of the present work.

Progresses in treatment of HIV, namely the introduction of ART, has not only reduced the gap between HIV infected individuals' life expectancy and that of the general population [101], but it has also prolonged the asymptomatic period of the disease and reduced the incidence of OI in the symptomatic period [231]. At present, one pill a day and a medical appointment with laboratory tests every 4-months may be all the health care that is required for an HIV infected individual in the asymptomatic period of the disease. At least during this period, which usually lasts for more than a decade [232], the infected individual is able to maintain the exact same life style he/she had prior to the infection and, in this sense, infection by HIV bears a lower disease burden than many other diseases. Nevertheless, even at the initial stage, the disease bears a significant social burden resulting from discrimination and stigma associated with HIV infection. HIV infection may, from the first moment, have an impact on family/love relationships, on access to credit and health insurance, on participation in the labor market, etc. and, thus, on the quality of life [233]. This is why, as described in the "National Program for Prevention and Control of the HIV/AIDS Infection" [89], health policies in HIV/AIDS require a coordination effort involving behavioral, social and health aspects.

In this context, it was considered of major relevance to include in the analysis of the quality of life of HIV infected individuals, not only health aspects but also "non-health" issues that affect and are affected by the utility assigned to health states. As discussed in Chapter 4 of Brazier *et al.* [159] "*Whether or not social activities are counted as health, health related quality of life or quality of life, the impact of social activities should form part of the description of the benefits of health care."*

This, in turn, led to the selection of DCE as the methodology to estimate utility weights for the population in analysis. DCEs are implemented with easily understood, self-administered questionnaires and by focusing on both health related and non-health related issues relevant to quality of life, the DCE estimated index can be used to compare across health and social interventions, which are all integrant parts of HIV/AIDS health policies.

Nevertheless, this possibility comes at a price. Since DCE utility weights are not anchored on the full health-death scale necessary for QALY calculations, the quality of life index constructed with

such weights may not be used to compare to programs in other areas of health care. While the utility weights here developed provide in-depth information about the relative weight assigned to each of the factors influencing quality of life among HIV infected individuals, it does not provide information on how these individuals would be willing to trade quantity for quality of life.

In order to allow for QALY calculations, we have also included in the model two alternative utility weights. The estimates used and the corresponding sources are detailed in Section 4.5.4.6.

4.5.4.5. *Application of DCE*

DCEs present respondents with samples of hypothetical choice sets drawn from all possible choice sets, according to statistical design principles. Each choice set comprises several alternatives, which vary in the levels of the attributes, and individuals are asked to choose the most (or least) preferred. The data generated is, then, modeled within a random utility maximization framework.

There are four sequential stages in conducting a DCE: (i) Identifying Attributes and Levels; (ii) Experimental Design to Determine Choices; (iii) Collecting Data and (iv) Model Estimation. The DCE methodology is explained in detail in Chapters 1 and 2 of Ryan *et al.* [234] and Coast and Horrocks [235] provide further guidance on attribute (and their levels) selection. The present analysis follows the guidelines therein. Consequently, this section will focus on explaining the application of the methodology to estimate the utility weights for quality of life of HIV infected individuals.

4.5.4.5.1. Identifying Attributes and Levels

Identification of the most relevant attributes was performed by gathering information from HIV infected individuals, HIV related associations and physicians who work with HIV infected individuals on a daily basis. The last section of the recently translated (to Portuguese) version of the WHOQoL-HIV-BREF questionnaire was used to inquire a (convenience) sample of 15 HIV infected individuals about the most relevant factors affecting their quality of life. This section includs 37 questions, where the respondents are asked to classify the factor in analysis, in terms of its degree of importance (on a 5-level scale). All factors assigned “highest importance”, by at least one of the individuals, were considered as potential attributes. Physicians and patients’ representative associations were asked, on a personal interview basis, about the attributes they

considered to be of relevance. The potential attributes identified were then grouped in order to attain a manageable number of attributes.

The number of attributes (and levels) included had an impact on the number of choice sets necessary to estimate the coefficients of interest and thus on the length of the questionnaire. This trade-off led to the selection of seven attributes: health and ability to perform daily activities, impact of treatment on life style, access to health care, discrimination, love/sexual relations, employment, and fear of premature death and/or suffering as a consequence of serious illnesses.

Employment was included as a proxy for social integration while the last attribute was included in an attempt to measure the value of being HIV infected *per se* even if the fact that a person is infected has absolutely no impact on quality of life at present (in any of the other dimensions), knowledge of HIV infection status bears with it a weight which the last attribute is meant to capture.

Each of these attributes has two levels. Since the number of health states increases exponentially with the number of levels, the two-level choice was, once again, based on the trade-off between length/complexity of the questionnaire and sample size. Although an initial version of the questionnaire with three levels for the first 5 attributes and two levels for the last two was tested on a small sample, respondents considered levels of attributes to be of minor relevance in the choice process. When faced with having to compare alternatives on the basis of 7 attributes, respondents confessed taking a two level approach in order to make the decision process easier. Consequently, a two level approach was selected even thought it implied the assumption of a linear marginal utility function, i.e. that the marginal utility of an attribute does not depend on the level that the attribute takes.

Table 28: Attributes and Levels Included in the Questionnaire

Attributes	Levels	
Health	Healthy, full of energy.	Sick, depressed and/or unable to self-care
Impact of Treatment on Lifestyle	Little impact and/or no severe adverse effects	Highly restricted activities and/or severe adverse effects
Access to Healthcare	Easy access to medication, routine blood tests and physician	Difficult access to medication, routine blood tests and physician
Discrimination	Never/rarely feels discriminated	Feels constantly discriminated
Impact of HIV Infection on Sexual and Love Relationships	Little or none	Unable to maintain a relationship
Employment Status	Employed	Unemployed
Fear About the Future	No worries about premature death and severe pain	Worries about premature death and severe pain
Coded Level	**1**	**0**

Notes: HIV=Human Immunodeficiency Virus

In an attempt to capture possible sources of variability within individuals which could lead to important valuation differences, as found in the empirical literature [177, 179, 182, 221, 222], the questionnaire included an initial characterization section where age, gender, transmission category and CD4+ cell count were asked to the respondent. Ideally, we would want to obtain DCE estimates by CD4+ stratum not only to allow for a comparison with estimates obtained with other methods but also for easier incorporation into the model.

Each choice set consisted of two alternatives with different values for all or some of the attributes. Because the choice between two hypothetical situations was found to be confusing to the respondents (who are used to questionnaires where they are asked to describe their own situation), the question was phrased by assigning a common Portuguese name to each alternative (in all choice sets) and asking whom they considered to be in a worse situation. Although it has been argued that using generic (unlabelled) experiments is preferred because it will focus the respondent on the attributes [173], the labels chosen were neutral enough to avoid attracting any attention from the respondent. As suggested by Miguel *et al.* [236], a clearly explained example was also included to help minimize choice inconsistencies.

Respondents were asked to choose who, in their opinion, was worse off. Asking the participants to choose the least preferred alternative, avoids having to choose "the best" of two undesirable options [214].

4.5.4.5.2. Experimental Design to Determine Choices

Seven varied attributes with two levels each, implied a total of 128 states. DCEs use only a subset of all possible combinations of attribute levels, called fractional factorial designs (FFDs). In the present context, a resolution 3 fractional factorial design (OMEP) obtained in Matlab R2009a® was used. The OMEP design ensures that estimates of all ($2 \times 7 = 14$) main effects are uncorrelated. Nevertheless, as noted in Ryan *et al.* [234], these designs have limitations, as they do not permit the estimation of interaction effects between any subsets of attributes, and main effects will be confounded with the interaction effects[73].

[73] A full factorial has the advantage of estimating both main effects of each attribute, as well as all the possible interactions between the attributes. But it is not tractable in practice, as it implies $(128 * 127)/2$ unique binary choice sets. A long version of the questionnaire, designed to be able to estimate interactions among attributes, was first considered. This version with 18 choice sets was, nonetheless, considered too long by physicians and patients. Although, using two versions of the questionnaire with 10 choice sets each could, in principle, be a solution, due to practical constraints it was not be possible to recruit enough respondents for two versions of the questionnaire and consequently, interactions could not be estimated.

The construction of experiment design, that is, the choice sets from the above orthogonal main-effects plan (OMEP), was performed considering the Huber and Zwerina criteria [234]: orthogonality, level balance, minimum overlap and utility balance, although these are neither necessary nor sufficient conditions for efficient designs [211]. The efficiency of a design is measured by the information it provides on the parameters, and this is measured by the Fisher information matrix of the model to be estimated. In order to produce a design with improved efficiency, it is desirable to maximize the information matrix. A design that achieves the optimum is called a D-optimal design. As discussed in *Ryan et al.* [234], in the context of the multinomial Logit model with two-level attributes and a choice set of size two, the D-optimal design is achieved by interchanging all 0s and 1s in the OMEP shown above. This design is 100% efficient.

The questionnaire thus consisted of 8 choice sets with two alternatives each. The size two of the choice was considered adequate, in that it facilitates the cognitive burden of the questionnaire. An additional dominant choice set was included for consistency tests. Appendix VI provides a blank copy of the questionnaire applied.

4.5.4.5.3. Collecting Data

A convenience sample was collected through self-administered questionnaires in two Immunodeficiency Treatment Units, while patients waited for a physician' appointment, and a non-governmental organization.

In the hospital setting patients were approached by a nurse or physician and asked whether they would like to participate. The Immunodeficiency Treatment Units are located in two hospitals in the Lisbon area (Hospital Santo António dos Capuchos, which is part of the Centro Hospitalar Lisboa Central, and Centro Hospitalar de Cascais).

Consent for application of the questionnaires was granted by the hospitals' Ethics Committees, as required and informed consent was signed by participants. The data from the two Immunodeficiency Treatment Units was collected between November 1st and January 31st, 2010, resulting in 59 respondents from Centro Hospitalar Lisboa Central, 33 from the Centro Hospitalar de Cascais. On randomly selected days during those three months and at the convenience of the nurses and doctors, patients waiting for an appointment were approached. No individual refused answering the questionnaire. Nevertheless, 10 respondents did not complete the questionnaire or failed the consistency question and were, therefore, discarded.

In an attempt to reach special subpopulations that may not be receiving appropriate care, such as illegal immigrants, individuals with drug addiction problems and/or homeless, data was also collected at SER+, a non-governmental organization located in Cascais. SER+ provides information, support and guidance in access to Social Security benefits to HIV infected individuals. At SER+, a technician approached potential respondents and explained the questionnaire to them. 23 individuals responded, 5 were discarded for not completing the questionnaire or failing the consistency question. The percentage, in each site, of complete questionnaires complying with the consistency question out of all questionnaires returned was thus higher in the hospital sample than in the SER+ sample (89% versus 78%).

Overall, 100 respondents were included in the estimation. Age ranged from 23 to 69 years old, with an average of 42.6 (standard deviation= 10.4) and 37% were female. With respect to CD4+ cell count, only 53% of respondents answered. Average CD4+ cell count was 571 per mm^3, (S.D. 349.5). The minimum observed value was 66 (maximum 1702) and 9.4% had CD4+ cell count per mm^3 below 200. With respect to mode of HIV transmission, 52% was sexual transmission (25% homosexual, 27% heterosexual), 31% was drug related. It is noteworthy that, in the sample, 13% of the individuals (62% of which are women) in the sample did not know (or became infected through other modes) and 4% did not answer.

Although the sample was never intended to be representative of the HIV population in Portugal, it is of interest to compare its composition to the characteristics of the declared HIV infected population alive in Portugal (as reported by the CVEDT database). In the sample, there is a sub-representation of individuals in the IDU class (31% versus 41%) and in the heterosexual class (27% versus 45%), and an over representation of individuals infected through homosexual contact (25% versus 12%). The most significant difference occurs, nonetheless, in the "Other/Unknown" mode, with 13% in the sample versus 2% in the National database. It is also noteworthy that, because the SER+ population is expected to have a higher prevalence of current IDUs and other HIV infected individuals living in less favorable socio-economic environments, the lower response rate observed in that center introduces a potential bias in results and suggests that alternative approaches (such as face-to-face questionnaires) may be required if a representative sample is to be obtained. Average age is one year lower in the sample and there is an over-representation of women (37% versus 28%).

4.5.4.5.4. Model Selection

The questionnaire consisted of eight choice sets with two alternatives each. Given this structure of the data, a binary choice model, accounting for panel structure of the data is appropriate. Most commonly, a random effects Probit/Logit model is used to model such type of data. Nevertheless, as previously discussed, there is empirical evidence in the HIV area to support the idea that health state valuations may depend on individual characteristics of individuals, such as health status and attitudes [177, 179, 182, 221, 222]. In such context, the coefficients associated with each attribute will vary from individual to individual, suggesting a random coefficients approach. Two models were, therefore, estimated: a fixed effects Logit model as a basis for comparison and a mixed Logit model[74].

The mixed Logit model overcomes the limitations associated with the Logit model, namely, it allows for random taste variation, unrestricted substitution patterns, and correlation in unobserved factors over time. Moreover, unlike Probit, it is not restricted to Normal distributions for all unobserved components of Utility, and is computationally simple, although the integral has to be (partially) solved by simulation because it does not have a closed-form. Train [237] provides an excellent guide for all steps required in such simulation and the analysis here performed follow those suggestions.

4.5.4.5.5. Results

Prior to model estimation, two transformations in the variables of the database were performed. First, the dependent variable was inverted (switching zero for one, and one for zero) so that the chosen alternative became the most preferred instead of the least preferred. Secondly, since the last attribute had been defined with a value of one for “worries about the future” and a value of zero for “no worries about the future”, the variable was multiplied by -1 to ensure a positive coefficient, in line with the remaining attributes. This is a standard procedure in discrete choice models, as described in Train [237].

Table 29 and Table 30 present the estimation results for the conditional (fixed effects) Logit Model and the Mixed Logit Model. In both specifications, the cluster-robust covariance estimator was used to allow for intra-correlation among observations of the same individual. All attributes are statistically significant at 5%. As expected (physical and mental) health is the attribute with the highest impact on utility, followed concerns about the future, by

[74] In Appendix VI the mixed Logit model is discussed in the context of the present application.

discrimination, employment and impact of treatment on lifestyle. These results hold for both specifications of the model.

Table 29: Conditional Logit Regression

Conditional (fixed-effects) logistic regression

		Number of obs =	1600
Log likelihood =	-476.0	Wald chi2(7) =	104.9
		Prob > chi2 =	0.000
(Std. Err. adjusted for 100 clusters in id)		McFadden R2	0.142
		% correct predictions =	68%

	Dep. Var: Best alternative	Coef.	Robust Std. Err.	z	P>z	[95% Conf.	Interval]
Mean	Health	0.70	0.10	6.97	0.000	0.50	0.89
	Lifestyle	0.23	0.07	3.23	0.001	0.09	0.36
	Healthcare	0.21	0.08	2.63	0.009	0.05	0.36
	Discrimination	0.38	0.07	5.08	0.000	0.23	0.52
	Love/sexual relationship	0.14	0.06	2.36	0.018	0.02	0.25
	Employment	0.30	0.09	3.44	0.001	0.13	0.48
	Future	0.38	0.10	3.83	0.000	0.19	0.58

Table 30: Mixed Logit Model

Mixed logit model

		Number of obs =	1600
Log likelihood =	-466.69	Wald chi2(7) =	98.9
		Prob > chi2 =	0.000
(Std. Err. adjusted for 100 clusters in id)		McFadden R2 =	0.148
		% correct predictions =	68%

	Dep. Var: **Best alternative**	Coef.	Robust Std. Err.	z	P>z	[95% Conf.	Interval]	P(x<0)
Mean	Health	0.82	0.12	7.05	0.000	0.59	1.05	
	Lifestyle	0.27	0.08	3.21	0.001	0.11	0.44	0%
	Healthcare	0.24	0.10	2.57	0.010	0.06	0.43	21%
	Discrimination	0.44	0.09	5.21	0.000	0.28	0.61	0%
	Love/sexual relationship	0.17	0.07	2.51	0.012	0.04	0.30	0%
	Employment	0.34	0.11	3.21	0.001	0.13	0.55	25%
	Future	0.46	0.13	3.66	0.000	0.21	0.71	28%
St. Dev.	Lifestyle	0.05	0.10	0.45	0.651	-0.15	0.25	
	Healthcare	0.30	0.25	1.24	0.215	-0.18	0.79	
	Discrimination	0.04	0.46	0.09	0.926	-0.86	0.95	
	Love/sexual relationship	0.04	0.04	0.89	0.373	-0.05	0.12	
	Employment	0.51	0.20	2.56	0.011	0.12	0.89	
	Future	0.77	0.22	3.5	0.000	0.34	1.20	

In the present analysis, convergence of the likelihood function was not achieved when a Lognormal distribution was assumed for the attributes expected to have the same sign for all individuals, namely, impact of treatment on life style, access to health care, discrimination, love/sexual relations; a Normal distribution was, therefore, assumed for all random coefficients. As shown in Table 30, even with the assumption of a Normal distribution for all random parameters, the coefficients have the expected sign. In fact, 0% of the individuals have a

negative sign on the attributes "impact of treatment on lifestyle", "impact of HIV infection on love/sexual relationships" and "discrimination". Moreover, the associated standard deviation coefficients are not statistically significant. 21% of the individuals are estimated to have a negative coefficient on the attribute "Access to healthcare" but the standard deviation coefficient is not statistically significant. Regarding the "Access to Health Care" attribute, it is worth mentioning that this attribute was found not to be statistically significant when the model was estimated on the sub-sample completing the questionnaires in the hospital context. Once individuals interviewed at SER+ were added the coefficient on this attribute becomes statistically significant likely reflecting the fact that people tend to value more what they do not have.

In line with prior beliefs, for some individuals, employment has a negative impact on utility (25%) and for 28% of the individuals results suggest that uncertainty about the future actually increases utility. This last finding is in line with "*the will to live*" phenomenon described in the literature [178]. Employment and concerns about the future where also the only two covariates with significant variability among respondents (as shown by the statistically significant standard deviations in Table 30) [75].

The overall significance of the models was evaluated by comparing the value of the pseudo-R2 or Likelihood Ratio Index of McFadden and the Wald statistics (shown in Table 29 and Table 30). The percentage of correctly predicted outcomes was also used as an additional measure of goodness of fit. The McFadden R^2 is slightly higher in the mixed model specification than in the conditional Logit, while in terms of the percentage of correct predictions the two models perform equally[76].

4.5.4.6. *Quality of Life Parameters for the CE of Therapy Model*

Results of the DCE analysis presented in the previous Section were used to construct a quality of life index (QoLIndex), which is used in the cost-effectiveness model to estimate the number

[75] While the estimated variances for all attributes, with the exception of employment and concerns about the future, are found not to be statistically significant, the model specification assuming fixed the coefficients associated with lifestyle, healthcare, discrimination and love/sexual relationships, yielded a lower McFadden R^2.

[76] The Mixed Logit Model was also estimated allowing for a correlation among coefficients and in that specification, few correlations were found to be statistically different from zero (Employment and relationships were correlated with discrimination and healthcare with lifestyle). This model specification yielded a slightly lower percentage of correct predictions.

of quality adjusted life years (QAdjLY) using a methodology analogous to that used to estimate QALYs (described in Section 4.5.4.1).

Nonetheless, as previously mentioned, since DCE utility weights are not anchored on the full health-death scale necessary for QALY calculations, the QoLIndex constructed with DCE utility estimates (and thus the resulting QAdjLY) may not be used to compare to programs in other areas of health care. In order to insure that the cost-effectiveness model produces an estimate of the number of QALYs for each alternative under consideration, we have also included two alternative utility weights.

Moreover, it is assumed in the model that inpatient care episodes, AIDS events and intolerance events are associated with a temporary reduction in quality of life. The remaining events (described in Section 4.5.3) are assumed not to have a specific disutility associated to them.

In this Section, we first present the estimates and corresponding sources of QALY weights used, then we describe the method used to obtain the QoLIndex based on the DCE results and at the end of the Section, the disutilities assumed for each modeled event and the corresponding source are presented.

4.5.4.6.1. QALY weights

In order to insure that the cost-effectiveness model produces an estimate of the number of QALYs for each alternative under consideration, we consider utility weights drawn from the literature. Both estimates were obtained using the EQ-5D with weights derived from the population in the United Kingdom. The first (Catarino 2010) was estimated on a sample of 152 HIV infected individual followed at Hospitais da Universidade de Coimbra (Portugal) [238] and the other (Simpson *et al.* 2004) was obtained from about 21,000 clinical trial patients [4]. Table 31 provides a summary of the results presented in these two studies[77].

Table 31: QALY weights obtained using EQ-5D

	Simpson et al. (2004)				Catarino (2010)
log_{10}HIV RNA->	**<2.6**	**2.6-4.3**	**4.3-5**	**>5**	**Any**
CD4<50	0.781				0.63
50<CD4<200	0.863	0.865	0.856	0.826	
201<CD4<350	0.929	0.931	0.933		0.72
351<CD4<500	0.934	0.931			0.83
CD4>500	0.954	0.938			0.85

[77] Catarino (2010) does provide estimates for suppressed and unsuppressed individuals. The author does not, however, provide those estimates desaggregated by CD4 cell count and we have consequently discarded the impact of viral load in favor of that of CD4.

4.5.4.6.2. The DCE-based Quality of life Index (QoLIndex)

For each individual, in each moment of time, the QoLIndex is calculated as a weighted average of the levels of each dimension. In our index, seven dimensions are considered: health and ability to perform daily activities, impact of treatment on life style, access to health care, discrimination, love/sexual relations, employment, and fear of premature death and/or suffering as a consequence of serious illnesses. The weights used are the (normalized) DCE utility estimates. The levels of each dimension are obtained from published results in the literature (Reis *et al.* [132, 134] and Canavarro *et al.* [239]).

The DCE model produces predicted valuations on an interval scale, such that meaningful comparisons of differences are possible, but the origins and units of the scale are defined arbitrarily [197, 240]. This means that the rank order of a set of health states will be the same under any positive affine transformation of the latent utilities, the utility from alternative *j* in choice situation *t* by person *n*, in its general form $U_{njt} = \alpha(\beta_n' x_{njt} + \varepsilon_{njt}) + \Phi$ and the predicted utility, conditional on the parameter values estimated by the model is $U_{njt} = \alpha(\beta_n' x_{njt}) + \Phi$. Consequently, in order to obtain the QoLIndex on a scale between zero and one, the estimated coefficients were normalized so that the state characterized by a low level on all domains yielded a utility of zero and the state characterized by a high level on all domains yielded a utility of one. This normalization is presented in Table 32.

For example, nowadays in Portugal, an HIV infected individual, asymptomatic, healthy, with no restrictions on lifestyle due to infection, employed and happily married, will have an estimated utility index between 0.63 [1-0.17-0.16] and 1 depending on the extent to which the individuals is affected by worries about the future and a feelings of discrimination.

Table 32: Quality Weights for Quality of Life Index (QoLIndex)

	Coefficient	Low level	High level	Quality weight	Low level	High level	MRS
Health	0.82	0	1	0.30	0	1	1.00
Lifestyle	0.27	0	1	0.10	0	1	0.33
Healthcare	0.24	0	1	0.09	0	1	0.30
Discrimination	0.44	0	1	0.16	0	1	0.54
Love/sexual relationship	0.17	0	1	0.06	0	1	0.20
Employment	0.34	0	1	0.12	0	1	0.42
Future	0.46	-1	0	0.17	0	1	0.56
Utility		-0.461	2.296	1.00	0	1	
Affine transformation of utility		0.000	2.757				

MRS=Marginal rate of substitution

The column labeled MRS, which is not affected by the affine transformation of the utility, reveals the rate at which an individual is willing to trade health for each of the remaining attributes, while maintaining the same level of utility. For example, say health of an HIV infected individual is measured on a 0-100 point scale and imagine an individual is on the top of the scale – that is, the individual is as healthy as he could be – and that he feels constantly discriminated for being HIV positive. This (average) individual would be willing to move down on the health scale from 100 to 46 (100-54) in exchange for not feeling discriminated.

At any point in time, the QoLIndex is calculated as a linear function $\mathrm{QoLIndex} = \boldsymbol{b}'\boldsymbol{X}$, where $\boldsymbol{b}$ is the vector of coefficients presented in the "Quality Weight" column of Table 32 and $\boldsymbol{X}$ is a vector containing level of each of the attributes (on a 0-1 scale) for the simulated individual, in that moment in time.

Quantification of the level of each attribute (x), on a 0-1 scale, should, ideally, be performed by an additional questionnaire to Portuguese HIV infected individuals, where each person would classify their current situation on each of the seven attributes included in the QoLIndex, in addition to providing information on their CD4+ cell count, viral load and socio-demographic characteristics. Although such information is not available, recent published literature on quality of life of HIV infected individuals in Portugal may be used as a proxy.

In 2008, Canavarro et al. published the results of the validation of the WHOQOL-HIV on a sample of 200 HIV infected individuals in Portugal [239]. In that that article, the authors provide information on the mean score (on a 0-100 scale) for each domain, by HIV stage (Asymptomatic, Symptomatic and AIDS). More recently, Reis et at. [132, 134] have collected data on a sample of 298 HIV infected individuals attending the Infectology Department of two Portuguese hospitals in 2009. The authors used the Portuguese versions of the WHOQOL-BREF [241], which also uses a 0-100 scale, plus the Instrumental-Expressive Social-Support Scale [242] to collect data on the socio-demographic profile of respondents. Reis et al. did not use the WHOQoL-HIV version of the questionnaire; consequently, questions regarding specific HIV related issues were not included[78]. A comparison of the results obtained in the two samples is provided in Appendix VI.

Both Canavarro *et al.* and Reis *et al.* present their results discriminated by disease stage (Asymptomatic, symptomatic and AIDS - CDC Classification System for HIV-Infected Adults and Adolescents [119]). When applying the results of the QoLIndex to estimate QAdjLY in the cost-

[78] Domain VI of the WHOQOL-HIV questionnaire includes questions 24 "Spirituality/Religion/ Personal Beliefs", 52 "Forgiveness and blame", 53 "Concerns about the future" and 54 "Death and dying"

effectiveness model, we have made the assumption that patients in lines 1 and 2 are asymptomatic, patients in line 3 are symptomatic and patients in non-suppressive therapy have a quality of life identical to that of patients diagnosed with AIDS. We have also assumed that, in any line, if a patient has been diagnosed with AIDS, his quality of life no longer depends on therapy line and it becomes until death equal to that of AIDS stage.

Table 33 provides a detailed explanation of how the QoLIndex was calculated. It specifies both the source and the corresponding facet or domain within each source used to proxy the level of each attribute included in our index. Although our QoLIndex (Table 33) is not directly comparable to the utility estimates obtained using EQ-5D (Table 31) due to the reasons previously discussed, results do suggest that non-health factors may indeed play a significant role in the quality of life of HIV infected individuals.

Table 33: Quality of Life Index (QoLIndex)

	Quality weight (DCE)	Line 1 and 2	Line 3	Line 4 or any line if AIDS	Source	Facet/Domain/Characteristic
Health	0.299	0.600	0.518	0.492	Canavarro *et al.* (2008)	Physical and Pscycological
Lifestyle	0.099	0.691	0.576	0.508	Canavarro *et al.* (2008)	Level of Independence
Healthcare	0.089	0.645	0.540	0.622	Reis *et al.* (2010)	Health and Social Care Access
Discrimination	0.161	0.630	0.630	0.630	Canavarro *et al.* (2008)	Social Inclusion
Love/Sexual Relationship	0.061	0.650	0.533	0.579	Canavarro *et al.* (2008)	Social Relationships
Employment*	0.124	0.660	0.660	0.660	Reis *et al.* (2010) & CHC	Percentage employed
Future	0.167	0.527	0.526	0.484	Canavarro *et al.* (2008)	Domain VI
Quality of Life Index (QoLIndex)		**0.616**	**0.564**	**0.552**		

Note: AIDS=Acquired immune deficiency syndrome. DCE=Discrete choice experiment. * In the cost-effectiveness model, employment level is a random draw from a Bernoulli distribution.

4.5.4.6.3. Decrement in utility due to clinical events

It is assumed in the model that hospitalizations, AIDS events and intolerance events are associated with a temporary reduction in quality of life. The remaining events (described in Section 4.5.3) are assumed not to have a specific disutility associated to them. The decrement in utility estimates where obtained from the Anis *et a*l. [218] who use data from OPTIMA (OPTions In Management with Antiretrovirals), a multinational, randomized, open, control, clinical management trial to estimate the impact of both AIDS-Defining Events and Non-AIDS Serious Adverse Events on quality of life using several instruments among which, the Health Utility Index (HUI3)[79]. Following their results, we assume a serious adverse event to have a variation in utility of -0.09 (95% C.I.[-0.14;-0.05]) for 4 weeks, followed by -0.10 (95% C.I.[-0.14;-

[79] The authors provide also estimates using the EQ-5D instrument. Nonetheless, as per discussion in the article, due to short recall period, EQ-5D estimates may have underestimated the impact of AIDS events. We have therefore selected the estimates obtained using the HUI3 instrument.

0.07]) for 8 weeks and a serious AIDS-defining event to cause a variation in utility of -0.09 (95% C.I.[-0.14;-0.04]) for 8 weeks followed by -0.7 (95% C.I.[-0.14;-0.01]) for 8 weeks. We assume an inpatient care event to generate a variation in utility similar to that of the last 8 weeks of the AIDS-defining event.

4.5.5. Costs

4.5.5.1. *Introduction*

While the relevant costs to be included will depend on the analysis at stake [158], the default model includes only direct costs associated with ART therapy, medical costs of follow-up and treatment of clinical events. No indirect costs were included in the model due to lack of available data.

Medication costs were valued according to the official table prices for Governmental health care products purchases (Catálogo de Aprovisionamento Público da Saúde) [243], when available. If more than one brand of a specific drug was available, the price used was the weighted average of available brands, with the weight set by the number of units bought by the Ministry of Health in 2008. If the drug was not available at that source, the price published by INFARMED [244] was used, and in that case, the unweighted average price was used. All other resources, (physician appointments, diagnosis tests and laboratory tests) were valued at the prices published in Portaria nº 132/2009 [245].

Resources were valued at 2009 prices. Given the fact that the database refers to the 2001-2008 period, these 2009 prices were not those in place when resource consumption occurred. Since relative prices may have changed, the prices at stake when resource consumption decisions were made should, ideally, be used to reflect the actual marginal rate of technical substitution. Because such price tables were not readily available for all resources, and in order to have a common criterion, 2009 prices were used for all resources. In the context of the Portuguese NHS, with some inelasticity of demand with respect to price (from the physicians' point of view), the error generated may not be too large. Moreover, given the fact that DRG tables do not yet reflect the actual production cost, although they have improved over time, 2009 tables may better reflect the true cost of the resources used.

4.5.5.2. *Antiretroviral Therapy Monthly Costs*

Antiretroviral drug prices used were those published in the Catalogue of Health Procurement (*Catálogo Público de Aprovisionamento da Saúde* [243]) in December 2009 and when necessary those published at INFARMED IP website. Prices were not available for some formulations (mostly, liquid presentation of antiretroviral drugs) and in that case, it was assumed the price of the solid equivalent formulation.

Antiretroviral drugs quantities were obtained directly from the CHC database. In line with the rational used to parameterize the model with respect to efficacy results, the CHC sample was used to estimate costs in first and second line, while the LVHEM sample was used for third line and non-suppressive therapy.

In the CHC sample, information was available on the exact package and formulation/dosage used, therefore the cost assigned to each regimen reflects that information. In the LVHEM sample, information was not available on the formulation used and in that case, it was assumed the price of the drug sold individually (that, is co-formulations were assumed not to exist). Moreover, the most frequent dosage for each drug was assumed.

Monthly regimen cost was assumed to vary according to individual characteristics. A Generalized linear model for panel data with a Gamma distribution with logarithm as the link function was used to estimate monthly ART cost per simulated individual. Specification tests were used to select the most appropriate link and family function according to the methods described in Glick [246]. Age and gender were not found to have an impact on monthly regimen cost. As expected, a higher CD4+ and a lower resistance score are associated with lower monthly regimen cost and; previous virological failures and year of first ART are associated with higher regimen costs.

Given the initial particular characteristics assigned to the simulated individual, a monthly regimen cost is consequently determined. That regimen cost will prevail, until the occurrence of a regimen switch event (either with or without virological failure). In the sample, the average monthly cost was 823€ ($95\%\ CI\ [814€; 832€]$) with a median of 804€

Table 34: Regression Model for Monthly Regimen Cost in Suppressive Therapy

GEE Population-Averaged model		Number of obs=	2684
Group variable:	id	Number of groups=	1076
Link:	log	Obs per group: min=	1
Family:	gamma	avg=	2.5
Correlation:	exchangeable	max=	11
Estimated within-id correlation:	0.286	Wald χ2(9)=	492.46
Scale parameter:	0.129	Prob> χ2=	0.000
Modified Park Test	0.728		
Pearson Test	0.020		
Pregibon Link test	0.173		

Monthly ART Cost	Coef.	Robust SE	z	P>z	[95% CI]		dy/dx
Gender (Female=1)	-0.029	0.018	-1.7	0.098	-0.064	0.005	-23.95 €
Age at 1st ART	0.000	0.001	0.470	0.637	-0.001	0.002	0.35 €
HCV	-0.015	0.017	-0.870	0.384	-0.048	0.019	-12.33 €
CD4+	-0.0001	0.000	-2.2	0.030	0.000	0.000	-0.05 €
Log_{10}VL	-0.030	0.006	-5.4	0.000	-0.041	-0.019	-24.76 €
Previous Virological Failures (0 is reference)							
VF=1	0.028	0.023	1.2	0.226	-0.017	0.074	23.27 €
VF>1	0.307	0.031	9.8	0.000	0.245	0.369	280.90 €
Resistance Level	0.039	0.007	5.5	0.000	0.025	0.052	31.86 €
Year 1st ART	0.035	0.003	13.4	0.000	0.030	0.040	28.81 €
Constant	-63.1	5.2	-12.1	0.000	-73.3	-52.8	Pred Mean= 826€

Notes: ART=Combined Antiretroviral Therapy, CD4+=CD4+ T-Lymphocyte count per µL, CI=Confidence Interval, HCV=Hepatitis C Virus, Log_{10} HIV RNA copies per mL

In non-suppressive therapy, the daily regimen cost assigned to each simulated individual is a random draw from the distribution of daily regimen costs fitted on the group of individuals in the LVHEM sample in class 4 of resistance. The fitted distribution was a Lognormal distribution with parameters80 $\mu = 6.807$ and $\hat{\sigma} = 0.276$ ($\chi^2 = 23.7, p = 0.0142$, Appendix IV, Figure 47) The observed average monthly cost is 939€ ($95\%\ CI$: [882€ ; 997€]) with a median of 904€, while the average of the fitted distribution is 937€. This value is likely to underestimate the current monthly cost in non-suppressive therapies given the availability of recent expensive drugs for end of line treatment.

4.5.5.3. *Outpatient Medical Monthly Costs*

According to the "Quality of follow-up Patterns" described in the Portuguese Recommendations for the HIV/AIDS infection [102], clinical follow-up should occur every three to four months. This follow-up is to include the physician appointment, complete blood tests, viral load and CD4+ cell count measurement. According to the prices published Portaria nº 132/2009 [245], such follow-up exam has a monthly cost of 53€.

[80] TreeAge® parameters for the Lognormal distribution are $\mu = \ln(median)$ and $\sigma = \sqrt{2\ln\left(\frac{\text{mean}}{\text{median}}\right)}$

In practice, costs may differ significantly from that value, since some individuals will not attend the pre-defined follow-up visits and others may need them on a more frequent basis. The CHC database contains information on the physician appointments, types of tests performed (blood samples, CD4+, viral load, etc); medication prescribed in conjunction with ART, and resources used on sporadic episodes (medication, other diagnosis exams and, in some cases, identification of the diagnosis) for the period of time that the patient attends the day hospital facility.

For modeling purposes, medication prescribed in conjunction with ART was used as a proxy for the cost of prophylaxis; physician appointments and fluids tests were assumed to be follow-up costs and resources associated with sporadic episodes were used as a proxy for the cost of light/moderate adverse events and non-severe OI.

These three sources of costs were added and the average monthly cost modeled as a function of individual characteristics. In the simulation, a random draw, conditional on individual characteristics at each event occurrence, provides the monthly outpatient follow-up cost.

Median monthly non-ART costs in the CHC was 108€ ($\bar{X} = 124€$, $95\%\ CI$: $[120€;\ 128€]$). The best fit to the generalized linear model for panel data has a Gamma distribution with logarithm as the link function[81]. Estimated coefficients are shown in Table 35. As expected, monthly non-ART costs are decreasing in CD4+ cell count and year of ART initiation and increasing in viral load, resistance level and disease progression. Interestingly, adherence has a positive marginal effect on cost, possibly reflecting the fact that patients who comply with medication will also attend physician consultations and perform fluid testing, as recommended by the physician.

[81] In estimation we have excluded 52 (1.6%) observations with monthly follow-up costs of more than 1,000€

Table 35: Regression Model for Outpatient Non-ART Cost

GEE Population-Averaged Model		Number of obs=	2472
Group variable:	id	Number of groups=	1149
Link:	log	Obs per group: min=	1
Family:	gamma	avg=	2.2
Correlation:	exchangeable	max=	8
Estimated within-id correlation:	0.105	Wald χ2(11)=	99.83
Scale parameter:	0.362	Prob> χ2=	0.0000
Modified Park Test	0.255		
Pearson Corr Test	0.599	Std. Err. adjusted for clustering on id	
Pregibon Link test	0.099		
Modified Hosmer Lemeshow Test	0.992		

Monthly Non-ART Cost	Coef.	Semi-Robust SE	z	P>z	[95% CI]		dy/dx	x
Gender (Female=1)	0.046	0.032	1.4	0.150	-0.017	0.110	6.06 €	35%
Age at 1st CART	0.000	0.002	-0.130	0.895	-0.003	0.003	-0.03 €	37.5
HCV	0.020	0.031	0.630	0.530	-0.042	0.081	2.56 €	52%
CD4+	-0.0001	0.000	-2.3	0.023	0.000	0.000	-0.02 €	435.3
Log_{10}VL	0.028	0.011	2.5	0.013	0.006	0.050	3.59 €	3.0
Previous Virological Failures (0 is reference)								
VF=1	0.000	0.033	-0.020	0.988	-0.065	0.064	-0.06 €	54%
VF>1	0.166	0.046	3.6	0.000	0.076	0.256	22.64 €	20%
Resistance Level	0.021	0.009	2.4	0.015	0.004	0.038	2.74 €	0.5
Year 1st CART	0.042	0.006	6.6	0.000	0.030	0.055	5.45 €	2002
Adherence	0.010	0.001	10.4	0.000	0.008	0.011	1.23 €	85.6
Constant	-80.2	12.8	-6.3	0.000	-105.3	-55.2	Predicted mean=130€	

Notes: ART=Antiretroviral Therapy, CART=Combined Antiretroviral Therapy, CD4+=CD4+ T-Lymphocyte count per µL, HCV=Hepatitis C Virus, Log_{10} HIV RNA copies per mL, VF=Virological Failure

4.5.5.4. *Cost of events*

Antiretroviral therapy costs and outpatient medical costs are assumed to be incurred continuously. These costs accumulate over time, even when no event occurs. Due to the discrete events nature of the model, values are updated at event occurrence but the update is performed considering the time lag since the last event occurrence (that is, since the last update) and the increment is the accumulated value. The cost of events is incurred upon event occurrence. Both types of costs are, nonetheless, dependent on individual characteristics and past history. In this section we provide our estimates for the cost of events.

4.5.5.4.1. Cost of Regimen Switch with and without Virological Failure

Portuguese Guidelines for HIV/AIDS infection [102] recommend a complete patient evaluation at ART initiation. The recommended resources, as well as their respective costs [245], are presented in Table 36.

Changing regimens, whether due to intolerance, virological failure or other, requires, in general, utilization of resources, such as physician appointments and laboratory tests. Given the fact that regimen switch frequency may be an important determinant cost-effectiveness analysis of

different treatment options, costs associated with regimen switch was modeled separately. In order to avoid double counting, physician appointment costs and laboratory tests' costs performed at the time of regimen switch were not included in estimation of outpatient medical monthly costs and instead accounted as switching costs.

Table 36: Cost of Antiretroviral Therapy Initiation

Cost of ART Initiation	DRG	Cost
Physician Appointment		31.0 €
Laboratory Tests		
Confirmation of HIV Status	26029	99.5 €
2x Viral Load Test	26322/26328	174.2 €
2x CD4+ Cell Count Test	25719	83.8 €
Complete Blood Test	See note 1	40.7 €
Resistance Assay	26340	263.1 €
Other Recommended Tests	See note 2	79.2 €
Chest X-Ray	10405	10.4 €
Electrocardiogram	40301	7.5 €
Gynecological Exam and Colpocytology (Women only)	48190	36.6 €
Anal Cytology Test for Screening of Pre-Cancerous Lesions	51280	15.7 €
Tuberculosis Skin Test	81365	7.1 €
Dentist Appointment		31.0 €
Total		879.8 €

Source: Portuguese Guidelines for HIV/AIDS infection (2007) and Ministério da Saúde, Portaria n.º 132/2009 de 30 de Janeiro

Notes: ART=Antiretroviral Therapy, CD4+=CD4+ T-Lymphocyte count per µL, DRG=Diagnosis-Related Group , HIV=Human Immunodeficiency Virus

Note 1: DRGs:4209+24380+22076+22949+21620+21623+22271+21217+21220+21665+22035+21935+21196+21559+21539+21545+22920+22954+22679+22076.

Note 2: RPR/VDRL, serology test for Taxoplasma Gondii, CMV and Hepatites A, B, C.

According to the Portuguese Guidelines for HIV/AIDS infection, it is assumed in the model that a regimen switch requires utilization of the following resources: two physician appointment, a viral load test, a CD4+ cell count test and a complete blood test. In case of virological failure, confirmation of such failure and a resistance assay should also be performed. As such, the cost of regimen switch with and without virological failure was calculated to be 582€ and 495€, respectively.

4.5.5.4.2. Cost of Adverse Events

Light adverse events costs are accounted for in "Outpatient Medical Monthly Costs". The cost of intolerance (adverse events causing a regimen switch) was obtained in the literature. Llibre-Codina *et al.* [247] estimate the annual cost of toxicity associated with NRTIs of 1,268 HIV-1 infected patients in several centers in Spain (RECOVER study), in the year 2005.

In their cost estimation Llibre-Codina *et al.* include both direct and indirect cost of intolerance, reporting that direct costs account for 45% of the total cost and events considered are classified in three categories: light, moderate and severe. The average (direct) cost of an intolerance episode of moderate or severe level, reported by Llibre-Codina *et al.*, is 1,126€

(95% CI: [1,124€; 1,129€]) at 2005 prices. 2009 prices were obtained using the inflation rates reported by the UNECE Statistical Division Database [70].

4.5.5.4.3. Cost of Hospitalization and AIDS-Defining Events

The CHC database contains information on all inpatient care episodes and the corresponding Diagnoses Related Group in which the episode was classified. Duration of episodes is nonetheless not known. Due to lack of information on lengths of stay, it was not possible to use generalized linear models for panel data link event costs to patient characteristics and past history. We neither obtained a reasonable fit nor were model variables found to be statistically significant. This link between individual characteristics and costs was the main advantage of using the CHC database. Once the option becomes the average cost, the ACSS database becomes a better source of information since it includes duration, DGR code and corresponding ICD-9 main diagnosis.

According to the ACSS database, the average cost of an inpatient care episode classified with the 042 ICD9-CM[82] code in 2009, was 4,765€ (95% CI: [4,597€ ; 4,933€]) and the average cost of the remaining episodes was 4,742€ (95% CI: [4,480€ ; 5,003€]). As shown in Table 37, the corresponding average duration was 19.3 (95% CI: [18.2 ; 20.5]) and 14.9 (95% CI: [13.5; 16.2]), respectively[83]. Although no difference was found in the average cost of the episode, the average duration was lower in the case of non-AIDS defining events.
In the model we have used assumed the cost of an AIDS event to be identical to that of inpatient care episodes classified with ICD-CM code 042 and that the average cost of hospitalization to be the average cost of the remaining episodes.

Because AIDS-defining episodes may imply inpatient care, our model will be double-counting the cost (and utility decrement) of AIDS-defining episodes in the cases where these events result in hospitalization. Once data becomes available, these assumptions will need improvement. Meanwhile, the overall impact is likely to be minimized by the fact that the CHC database contains mainly patients in early treatment lines (this being the reason for using the LVHEM database in line 3 and 4). The hospitalization rate observed in the CHC sample is, therefore, unlikely to reflect the higher hospitalization frequency observed in patients at an advanced stage of the disease.

[82] The 042 "Human immunodeficiency virus [HIV] disease" includes AIDS and symptomatic infection.

[83] As discussed in Section ACSS of Appendix III, Pina (2007) performed a similar analysis with a larger sample, over a longer period, broken down by hospital, which allowed for a parameterization according to the CHC cost pattern. Nevertheless, that analysis was based on previous DRG tables, which are no longer valid. Consequently, the 2009 sample was used.

Table 37: Cost of Hospitalization and AIDS Event

	AIDS Episode*		Hospitalization	
	Cost	Duration (days)	Cost	Duration (days)
Average	4,765 €	19.3	4,742 €	14.9
Standard Deviation	3,474 €	24.9	3,689 €	18.9
Median	3,353 €	12.0	3,353 €	8
N	1641	1641	763	757

Source: ACSS database 2009.*identified by the ICD9-CM code. **Notes**: ACSS= Administração Central do Sistema de Saúde, AIDS=Acquired Immune Deficiency Syndrome, N=Number of Observations/Individuals

Since hospitalization costs include the cost of ART, to avoid double counting in the model, this cost is subtracted from the cost of the event, assuming the average duration for the event as reported in the DRG table. This ART subtracted cost is the one incurred by the simulated individual at the time of event occurrence.

4.5.6. Summary of sources for model input parameters

The present framework for cost-effectiveness analysis of HIV treatment would, ideally, be based on a single complete national database. The availability of such database would allow for nationally representative parameter estimates and also for adequate modeling for the linkage between the different clinical parameters.

While the National Coordination for HIV/AIDS infection has already taken the first step towards such database – through the SI. VIDA software ("*Sistema Informático do VIH/sida*") – such project is yet at a pilot stage and facing several operational difficulties, mainly due to due to lack of resources.

In the absence of such database, the present model, as discussed in detail in Section 4.5, makes use of several databases and published international literature. Table 38 summarizes the sources of information used to estimate each required parameter.

Table 38: Summary of Sources for Model Input Parameters

Parameters for:	Source
Characteristics at ART Initiation	Detailed in Table 24 CHC, LVHEM, CVEDT, Reis *et al.* (2007)
Evolving State Variables	
CD4+ Cell Count and Viral Load	CHC, LVHEM
Adherence Level	CHC
Resistance	Endogenous from resistance event
AIDS Status	Endogenous from AIDS event and CD4+
Regimen Characteristics	CHC
Clinical Events	Detailed in Table 27
Virological Suppression	CHC, Clotet *et al.* (2007) (POWER Clinical trails)
Regimen Switch without Virological Failure	CHC, Elzi *et al.* (2010) (Swiss cohort)
Line Switch - Virological Failure	CHC, LVHEM
Resistance Development	CHC, LVHEM
Hospitalization	CHC
AIDS Event	May *et al.* (2004)
Immunological Failure	Endogenous
Mortality	CVEDT, INE
Quality of Life	
DCE Index	
Quality Weights	DCE questionnaire (CHC, CHLC and SER+)
Dimension scores	Reis *et al.* (2010), Canavarro *et al.* (2008), CHC
QALY weights	Simpson *et al.* (2004) , Catarino (2010)
Costs	
Antiretroviral Therapy	CHC, LVHEM, CPAS, INFARMED
Outpatient non-ART costs	CHC, Portaria nº 132/2009
Regimen Switch	Portuguese Guidelines HIV/AIDS infection, Portaria nº 132/2009
Intolerance	Llibre-Codina *et al.* (2007) (RECOVER Study)
Hospitalization	ACSS inpatient data 2009
AIDS-Defining Events	ACSS inpatient data 2009

Notes: CHC=Centro Hospitalar de Cascais, LVHEM=Egas Moniz Hospital Virology Laboratory - Western Lisbon Hospital Center, CVEDT=Communicable Diseases and Epidemiological Surveillance Center, INE=National Statistics Institute, CAPS=Catalogue of Health Procurement , ACSS=Healthcare System Central Administration, DCE=Discrete choice experiment

4.6. Model outputs in the context of an application

The aim of the present section is to introduce the reader to the outputs produced by the model. Such outputs will be presented in the context of an example: the cost-effectiveness analysis of two NRTIs plus one NNRTI versus two NRTIs plus one PI/r as a first line antiretroviral therapy regimen. It should be noted that this analysis has been previously presented at the HIV10 Glasgow conference, 2010 [248][84].

Portuguese Guidelines for treatment of HIV/AIDS infection [102] recommend initiating treatment with a regimen composed of 2NRTI+NNRTI or 2NNRTI+PI/r as do most of the Guidelines (Table 39). However, these recommendations do not take into consideration cost-

[84] The differences in results refer to the fact that in the previous analysis, we did not have inpatient data in the CHC database and have, therefore, used values reported in the literature. Moreover, once inpatient data become available, slight corrections were performed to the database to correct for inconsistencies.

effectiveness analysis. The objective of this study was to analyze the cost-effectiveness of the two strategies in a cohort of HIV-1 infected naïve patients.

While NNRTI and PI/r are considered clinically equivalent, in the sense that elements of both classes are considered as first choices in clinical recommendations, the average daily cost of PI/r is significantly higher than that of NNRTIs. Consequently, in a search for a more efficient use of resources, especially given the increasing cost containment pressure in Portugal, it is relevant to estimate the long term impact of each treatment option and compare the alternatives within a cost-effectiveness framework.

In the present analysis, we consider four therapy lines, a lifetime horizon and a 5% annual discount rate. Given the unavailability of reliable data to estimate indirect costs, we have assumed the National Health Service perspective considering only direct medical costs.

Table 39: Third Agent Recommendations for Initial ART treatment

	DHHS[a] 2011	IAS-USA[b] July 2010	BHIVA[c] June 2008	EACS[d] 2009	Rec. PT[e] Jan11	GESIDA[f] January 2011
Preferred /Recommended	EFV	EFV	EFV	EFV	EFV	EFV
				NVP#	NVP#	NVP* #
	ATV/r	ATV/r		ATV/r	ATV/r	ATV/r
	DRV/r	DRV/r		DRV/r	DRV/r	DRV/r
	RAL	RAL		LPV/r		LPV/r*
				SQV/r bid		RAL*
Alternative	NVP		NVP			
					FPV/r	FPV/r
	FPV/r	LPV/r	PI/r	SQV/r id	LPV/r	SQV/r
	LPV/r	FPV/r		FPV/r	SQV/r	MVC
		MRV		RAL	RAL	

Notes: *Not all panel experts agreed with this level of recommendation. # CD4 Restrictions.
a Panel on Antiretroviral Guidelines for Adults and Adolescents. Guidelines for the use of antiretroviral agents in HIV-1-infected adults and adolescents. Department of Health and Human Services. January 10, 2011; 1–166. Available at http://www.aidsinfo.nih.gov/ContentFiles/AdultandAdolescentGL.pdf
b Thompson M, et al. JAMA. 2010;304(3): 321-333
c. British HIV Association guidelines for the treatment of HIV-1-infected adults with antiretroviral therapy 2008. HIV Med. 2008; 9:563–608.
d. European AIDS Clinical Society (EACS) Guidelines for the Clinical Management and Treatment of HIV Infected Adults in Europe. November 2009. Available online at http://www.europeanaidsclinicalsociety.org/guidelinespdf/EACS-EuroGuidelines2009FullVersion.pdf
e. Recomendações Portuguesas para oTratamento da Infecção VIH/SIDA. Available online at http://www.sida.pt
f. Documento de consenso de Gesida y PNS sobre el tratamiento antirretroviral del adulto (enero 2011)

4.6.1. Input data

In order to perform the analysis, we considered the sub-sample of naïve individuals in the CHC database who initiated ART with either 2NRTI+NNRTI or 2NRTI+PI/r. 317 out of the 366 naïve individuals who initiated therapy during the period of analysis, met the criterion and were included.

Table 40 provides the summary characteristics of the two groups at ART initiation. The only statistically significant difference relates to year of ART initiation. It should nonetheless be noted that patients were *not* randomly assigned to each group.

In the simulation, initial individual values for each of the characteristic are randomly drawn from Table distributions generated from the sample in analysis. Obviously, the number of PIs was set to two in the 2NRTI+PI/r and to zero in the 2NRTI+NNRTI. The Table distribution of the number of PIs in line 2 was redefined to become conditional on the number of PIs in the first regimen.

Table 40: Initial Characteristics of the 2NRTI+NNRTI and 2NRTI+PI/r Groups

	2NRTI+NNRTI	2NRTI+PI/r	p-value*
N *(not randomized)*	158	159	
Year of ART initiation	2003	2005	<0.001
Female	32.3%	34.0%	0.751
HCV	29.8%	29.6%	0.971
Resistance* >1	2	1	0.79
Median [IQR]			p-value*
Age	39 [33 ; 46]	39 [33 ; 50]	0.649
CD4+ cell count/µL	234 [128 ; 349]	219 [108 ; 350]	0.500
Log10 HIV RNA/mL	4.9 [4.3 ; 5.4]	5.1 [4.3 ; 5.5]	0.056
Adherence (%)	89 [71 ; 98]	88 [73 ; 96]	0.327
NRTI Pair			
AZT+3TC	55%	52%	
TDF+FTC	19%	23%	
TDF+3TC	9%	9%	
ABC+3TC	3%	11%	
Others	15%	5%	
3rd Agent			
EFV	64%		
NVP	36%		
LPVr		73%	
Other PI/r		27%	

Note: Inverted Genotype Sensitivity Score based on REGA 8 Algorithm; HCV= hepatitis C virus; Adherence measured by pharmacy refills in first regimen. *P-value of the test for difference in means/medians. NRTI=Nucleoside Reverse Transcriptase Inhibitor; NNRTI=Non-nucleoside Reverse Transcriptase Inhibitor; PI/r=Boosted Protease Inhibitor.

The two groups were tested for differences in time to event using the log-rank test for equality of the survival functions (Figure 21) and CD4+ variation and viral load variation. Figure 22 shows the variation in CD4 cell count over time in the two groups. The only statistically significant difference was found in time to regimen switch without virological failure.

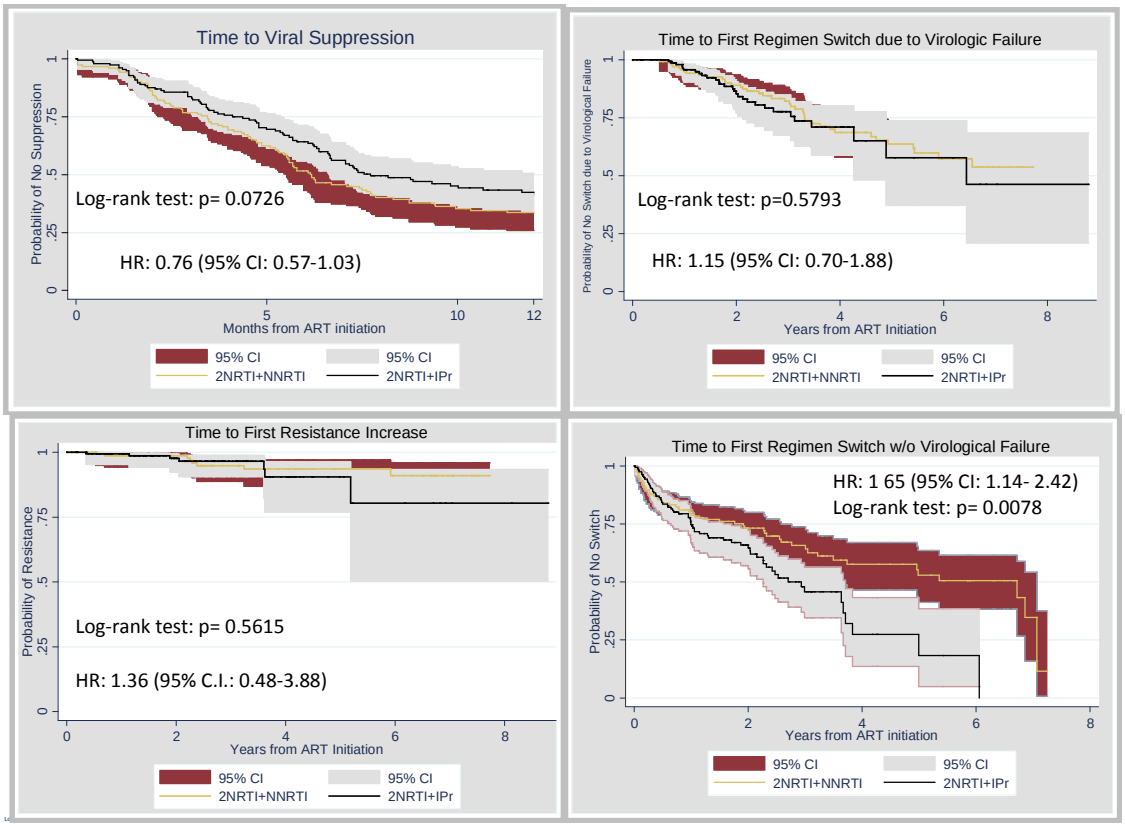


Figure 21: Kaplan Meier Time Estimates in the NNRTI and PI/r groups, (a) Time to Viral Suppression, (b) Time to Regimen Switch due to Virological, (c) Time to First Increase in Resistance and (d) Time to Regimen Switch without Virological Failure

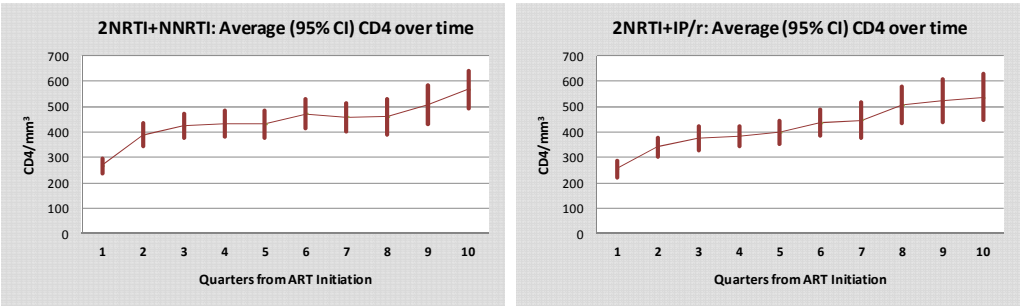


Figure 22: Average CD4+ over time in the 2NRTI+NNRTI and 2NRTI+PI/r Groups

First line time to event regression models[85] for time to viral suppression, line switch and regimen switch were re-estimated on the sample in analysis. The variable number of PIs identifies the group in those regressions. The estimated parameters replaced those used in the

[85] And consequently, second line corresponding regressions.

default version of the model. Due to reduced number of events observed, time to first resistance parameters were maintained at the default values.

First line non-ART costs in each group were assumed not to depend on individual characteristics. Instead the sample values presented in Table 41 were used. The ART cost of the first regimen, for each group, was that presented in the same Table.

Table 41: Observed Monthly Costs in the 2NRTI+NNRTI and 2NRTI+PI/r Groups

	Average [95% C.I.]	p-value	Median [IQR]	p-value
First Regimen Monthly Cost				
2INRT+INNRT	**705€** [658€ ; 752€]	p<0.001	**613€** [525€ ; 797€]	p<0.001
2INRTI+IP/r	**994€** [966€ ; 1022€]		**1057€** [853€ ; 1057€]	
First Line Outpatient Monthly Costs				
2INRT+INNRT	**82€** [63€ ; 101€]	p=0.05	**63€** [24€ ; 106€]	p=0.008
2INRT+IP/r	**109€** [89€ ; 129€]		**98€** [41€ ; 137€]	
First Line Inpatient Monthly Costs				
2INRT+INNRT	**92€** [51€ ; 134€]	p=0.62	**0€** [0€ ; 0€]	
2INRT+IP/r	**109€** [57€ ; 161€]		**0€** [0€ ; 46€]	
First Line Inpatient Monthly Rate				
2INRT+INNRT	**0.0162** [0.0097 ; 0.0227]	p=0.9492	**0** [0 ; 0]	
2INRT+IP/r	**0.0159** [0.0094 ; 0.0224]		**0** [0 ; 0.0133]	

NRTI=Nucleoside Reverse Transcriptase Inhibitor; NNRTI=Non-nucleoside Reverse Transcriptase Inhibitor; PIr=Boosted Protease Inhibitor.

With respect to mortality, we estimated time to death among the subsample of individuals diagnosed in the HAART era (1996 onwards) and considered it as the relevant data for individuals in suppressive therapy. In non-suppressive therapy we considered the full sample to capture the decreased benefits of ART at this stage. The remaining parameters of the model were kept at the respective default value.

4.6.2. Microsimulation Results

In the context of individual microsimulation models, it is important to distinguish between variability associated with individual characteristics and second-order variability. The first source of variability would exist even if we knew with absolute certainty all parameter values, simply because individuals are different, while the second occurs because we are uncertain about the exact value of the parameters in the population. Even if individuals were all identical, probabilistic analysis accounting for parameter uncertainty would still be pertinent.

Microsimulation outputs presented in this section were obtained by simulating 1 million individuals for each comparator, using the average values of all parameters. Such results are equivalent to those obtained in the deterministic analysis of a Markov model. Parameter

uncertainty is evaluated in Section 4.6.3 by simulating 50,000 individuals for each set of 133 different (randomly sampled) combinations of parameter values, for each comparator.

4.6.2.1. *Clinical outcomes*

Relevant clinical outcomes produced by the model include the number of months before viral suppression is achieved as well as the percentage of individuals achieving it. This information is relevant not only because of the risk of disease progression and resistance development but also because infectiousness is highly correlated with viral load [249].

CD4+ cell increase is also a common indicator reflecting the recovery of the immune system as a result of ARV therapy. In addition to these two indicators, the average number of each type of clinical event modeled is also reported.

The number of regimen switches (for reasons other than virological failure) reflects regimens' tolerability and safety.

Time spent in each therapy line provides information on the average time to virological failure. Virological failure may occur because viral suppression was not reached or because viral load has becomes detectable (in two consecutive tests) after a period of viral suppression. The model reports the average number of each event per simulated individual.

As expected, with 1 million simulations, simulated initial characteristics average values are close to those observed in the distributions from which those characteristics were sampled (Table 24 and Table 40).

With respect to clinical outcomes, little difference is predicted by the model in the two groups. These results are expected given the pre-analysis on the Kaplan-Meier estimates of time to event. The statistically significant difference in time to regimen switch without virological failure is estimated to reduce by 17% the total number of switches over the individuals' lifetime therapy.

Table 42: Clinical outcomes

Clinical outcomes	2NRTI+ NNRTI	2NRTI+ PIr	Absolute Difference	Percent Difference
Months w/o viral suppression (among those who reach it)				
Line 1	7.55	7.55	0.00	0.0%
Line 2	7.17	7.25	-0.08	-1.2%
Line 3	17.95	17.66	0.29	1.7%
% attaining viral suppression				
Line 1	68%	68%	0.00	-0.1%
Line 2	55%	55%	0.00	-0.4%
Line 3	46%	46%	0.00	0.0%
Variation in CD4 cell count				
Line 1	265	264	0.28	0.1%
Line 2	180	181	-0.66	-0.4%
Line 3	101	102	-0.19	-0.2%
Line 4	-49	-50	0.76	-1.5%
% Reaching each line				
Line 2	0.76	0.75	0.01	1.8%
Line 3	0.49	0.48	0.00	0.6%
Line 4	0.33	0.32	0.01	4.0%
Life time events				
Regimen switch	4.16	5.02	-0.86	-17.2%
Hospitalization	4.55	4.70	-0.15	-3.1%
Virological failure	3.53	3.50	0.03	0.7%
Failure after supression	1.53	1.52	0.01	0.4%
Supression not achieved	2.00	1.98	0.02	1.0%

Note: 1 million individuals simulated. Values undiscounted. CD4+=CD4+ T-Lymphocyte count per µl. OI=Opportunistic Infections

The remaining differences reflect the impact of the number of PIs and of the accumulated number of regimens on the distribution of time to event. For example, the slightly lower number of hospitalizations events in the NNRTI group reflects the negative marginal effect (Table 27) of the number of PIs on the median time to such event occurrence. It should be noted that, although 1 million individuals simulated should be enough to estimate the average value with precision, small differences may remain reflecting nothing but first-order variability.

The estimated average time in first line (Table 43) is around 9 years and a 5 months difference, with longer time for the NNRTI group, is predicted in terms of the average time spent in first line. This reflects the negative impact of both the number of PIs and the number of accumulated regimens on the median time to virological failure in lines 1 and 2. Overall effectiveness differences are small with an estimated overall difference of 2.5 months in undiscounted life expectancy.

With respect to the time of death, it is estimated that 24% of the individuals will die during first line, 28% will die during second line and 16% after two virological failures and before reaching

non-suppressive therapy. The remaining 33% will die in non-suppressive therapy. The percentage of individuals reaching non-suppressive therapy is likely to be high when compared to that expected from current available treatments[86] but it reflects the fact that the model was parameterized with 2000-2008 data.

4.6.2.2. *Costs, life years and ICERs*

As explained in Section 4.5.4.6, effectiveness is evaluated using three different indicators: Life years, life years adjusted for quality according to the quality of life index obtained by DCE and QALYs (with two different estimates on utility weights).

Cost, reported in Euros (€), are subdivided in ART costs and other costs (clinical events and follow-up) and further discriminated by therapy line. Incremental cost-effectiveness ratios (ICER) are also reported. Following the economic evaluation literature, when dominance occurs, ICERs are replaced by information on whether the strategy under consideration is dominant or dominated.

The main difference between the two groups relies on the total life time cost (Table 43). The model predicts that initiating ART with 2NRTI+NN results in savings of 28,178€ over the lifetime of the individual, that is, around 1,200€ per year. This annual difference is 11% of the value paid to the hospital per patient in the HIV financing program. These savings occur not only in terms of ART costs but also in terms of non-ART costs. The most significant cost differences are, obviously, in first line since treatment in subsequent lines is only slightly conditional on first line ART in each group (through the number of PIs and the number of accumulated different regimens). ART costs account for 79% of all costs reflecting the fact that progress in ART has improved overall health, decreasing the frequency and severity of OI/ AIDS-defining episodes but such progress is being paid for in ART costs.

Annual non-ART costs are increasing in therapy line indicating the need for more intense utilization of resources as the disease progresses. ART-costs are also increasing in therapy line reflecting the need for more complex and expensive drugs as resistance develops. Non-suppressive therapy is the exception. Annual ART costs in non-suppressive therapy are estimated to be lower than those for third line which is an unrealistic result but is a direct consequence of the data available. Indeed, the observed average monthly cost of 939€ among individuals with high resistance levels (in the LVHEM database) is unlikely to reflect current

[86] This observation is based on clinical expert opinion, not on published literature given its unavailability.

monthly costs among these individuals, given the new more expensive drugs that have meanwhile become available.

Table 43: Costs and Life Years Microsimulation Results

	Discounted at 5% per year				Undiscounted			
	2NRTI+NNRTI	2NRTI+PIr	Absolute Difference	Percent Difference	2NRTI+NNRTI	2NRTI+PIr	Absolute Difference	Percent Difference
Costs								
ART costs	135,406 €	151,158 €	-15,753 €	-10.4%	273,294 €	295,529 €	-22,235 €	-7.5%
Line 1	54,229 €	70,450 €	-16,221 €	-23.0%	84,775 €	107,044 €	-22,269 €	-20.8%
Line 2	41,752 €	41,880 €	-128 €	-0.3%	86,810 €	88,450 €	-1,640 €	-1.9%
Line 3	31,812 €	31,070 €	742 €	2.4%	70,431 €	68,287 €	2,144 €	3.1%
Line 4	7,613 €	7,758 €	-145 €	-1.9%	31,277 €	31,747 €	-470 €	-1.5%
Monitoring and	37,336 €	41,157 €	-3,821 €	-9.3%	73,998 €	79,941 €	-5,943 €	-7.4%
Line 1	14,166 €	17,991 €	-3,825 €	-21.3%	20,199 €	25,759 €	-5,560 €	-21.6%
Line 2	11,422 €	11,515 €	-92 €	-0.8%	22,567 €	23,191 €	-625 €	-2.7%
Line 3	8,704 €	8,534 €	170 €	2.0%	18,629 €	18,123 €	506 €	2.8%
Line 4	3,044 €	3,118 €	-74 €	-2.4%	12,604 €	12,867 €	-264 €	-2.0%
Total cost	172,742 €	192,315 €	-19,573 €	-10.2%	347,292 €	375,470 €	-28,178 €	-7.5%
Effectiveness								
DCEAdjLY	7.82	7.76	0.06	0.8%	15.26	15.03	0.23	1.5%
QALYs[##]	10.21	10.09	0.12	1.2%	20.87	20.50	0.37	1.8%
QALYs*	11.84	11.70	0.14	1.2%	22.84	22.43	0.41	1.8%
Life Years	15.69	15.57	0.12	0.8%	24.35	24.14	0.21	0.9%
Line 1	8.13	8.06	0.07	0.9%	9.41	9.15	0.26	2.8%
Line 2	4.41	4.41	0.00	0.0%	7.62	7.59	0.03	0.4%
Line 3	2.58	2.52	0.06	2.5%	4.55	4.59	-0.04	-0.8%
Line 4	0.57	0.59	-0.01	-2.0%	2.77	2.81	-0.04	-1.4%
Incremental Cost effectiveness Ratio								
ICER (€/QALY or DCE AdjLY)	2NRTI+NNRTI is Dominant				2NRTI+NNRTI is Dominant			
ICER (€/LY)	2NRTI+NNRTI is Dominant				2NRTI+NNRTI is Dominant			

Note: 1 million individuals simulated. ART=Antiretroviral therapy, CD4+=CD4+ T-Lymphocyte count per µl, LY=Life Years; DCE Adj = quality of life adjusted by the DCE estimates, ## based on Catarino (2011), *based on Simpson et al. (2004).

The model predicts a life expectancy of around 24 years, which means that the average age of death is 65. When accounting for quality of life estimates differ significantly based on the instrument (and sample) used to estimate utility weights. Recalling that both Simpson *et al.* [4] and Catarino [238] use the same instrument (EQ-5D) to estimate quality of life among HIV-infected individuals, it is worth noting that the estimates obtained on a Portuguese sample yield a 8% reduction in the total number of QALYs. This result suggests that HIV infected individuals have a lower quality of life in Portugal when compared to international standards. As expected, the total quality adjusted life years estimated by the model using DCE estimates, is significantly

lower that the total number of QALY. Accounting for non-health aspects of HIV infection that have an impact on quality of life, such as discrimination, reduces the estimate utility associated with a given health status.

With a lower total cost and a slight increase in life expectancy, ART initiation with 2NRTI+NNRTI is, in the present analysis, a dominant strategy when compared to ART initiation with 2NRTI+PI/r. Given the small differences estimated regarding the total number of QALYs in the two strategies, it is of special interest to evaluate the impact of the assumed parameter values on the results. Such probabilistic analysis is discussed in the next section.

4.6.3. Probabilistic Analysis

Although uncertainty resulting from individual variability is included in the microsimulation outputs presented in the previous section, the uncertainty (from the statistician point of view) about the parameters' values assumed in the model is not accounted for. The microsimulation outputs in the previous section were obtained with all parameters at their mean values.

In the present section, we discuss the model outputs designed to evaluate the robustness of the results with respect to parameter uncertainty. Such uncertainty, denominated second-order uncertainty, is incorporated in the probabilistic analysis. Probabilistic analysis is performed by sampling a value for each parameter from its distribution and then running the microsimulation for that set of parameter values.

The model provides two instruments for probabilistic analysis: the cost-effectiveness plane (Figure 23) and the acceptability curve (Figure 24). In the cost-effectiveness analysis of 2NRTI+NNRTI versus 2NRT+PI/r, 50,000 individuals were simulated for each of the 133 possible sets of parameters. Each point represented in the cost-effectiveness plan represents the pair (difference in average cost, difference in average effectiveness) for a given set of randomly sampled parameter values.

Results presented use the QALY values based on Simpson *et. al.* [4]. 57% of the simulations yielded ART initiation with 2NRTI+NNRTI as dominant strategy while 9% resulted in such strategy being dominated.

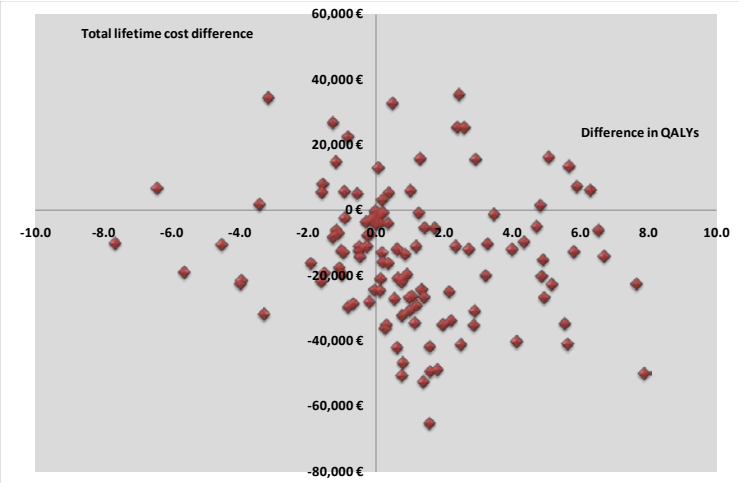


Figure 23: CE Plane of 2NRTI+NNRTI versus 2NRTI+PI/r (Discounted)

The probability of the treatment being cost-effective, obviously, depends on the willingness to pay for each unit gained. The acceptability curve summarizes how that probability varies in response to changes in willingness to pay. Such curve is represented in Figure 24, where a unit of gain equals a QALY. The probability of being cost-effective, when quality of life is taken into consideration, is about 80% at a threshold of 20,000€. It should be noted, that neither indirect costs nor new infections averted due to treatment, are included in the analysis.

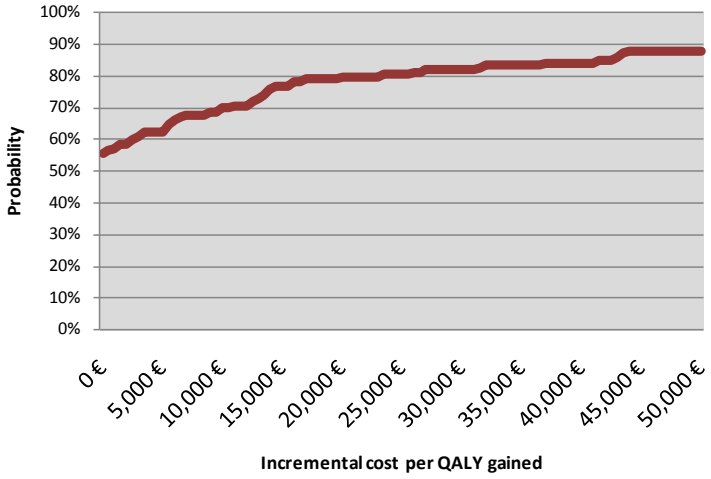


Figure 24: Probabilistic sensitivity analysis: cost-effectiveness acceptability curve for 2NRTI+NNRTI compared with 2NRTI+IPr as initial antiretroviral therapy

4.7. Discussion

Although by definition a simplification of reality, to be of use, a model should be able to reproduce the main aspects under consideration. Its validity may then be tested by applying the model to different samples and verifying whether it is capable of reproducing those results as well.

While the second step has not yet been performed, in order to evaluate model predictions, it is of interest to compare them, when possible, to those observed in the CHC sample and those published in the literature. We consider first clinical outcomes then the ICER related main variables. Model limitations are discussed next and we conclude with a final overall appreciation.

Sixty eight percent of the individuals are predicted to reach viral suppression in the first line and those who do, will attain it, on average, 7.6 months after ART initiation. This percentage (68%) is lower than clinical trial results of ARV drugs used by the CHC sample (for example, in Study 934 [250] 70% reach HIV RNA <50 copies/µL at 48 weeks in the zidovudine arm, while 80% do so in the tenofovir arm), but this is likely to reflect the difference between efficacy and effectiveness.

An average of 7.6 months to viral suppression matches the input data. Greenbaum *et al.* [251] estimates a median time to first undetectable HIV-1 RNA level of 5.4 months, but average and median are not directly comparable. Moreover, the event modeled is “HIV RNA<50 copies/mL test result” which depends not only on ART efficacy but also on testing frequency. The model predicts time to viral suppression to be identical in both groups, reflecting the fact that the difference was not found to be statistically significant (Figure 22). Results in the literature are contradictory with respect to such comparison [251, 252].

The percentage of individuals who respond to therapy and reach viral suppression decreases with the number of previous therapy failures, as expected, and is in line with clinical trial results of naïve and experience patients. After the first virological failure 55% will respond and, after two virological failures, 46% will reach viral suppression before switching to non-suppressive therapy. These values are similar to those observed in clinical trials [142] (from which third line data was drawn).

With active ARV medication blocking the virus from reproduction, CD4+ cell count tends to increase. Such increase is predicted to be 265, 180 and 129 in first, second and third line,

respectively[87]. This implies an annual increase (were the rate to be constant) of 28, 24 and 21 per year respectively. This values are similar to those estimated by Phillips *et al.* [253] on the EuroSIDA cohort. In non-suppressive therapy, CD4+ cell count is predicted to decrease at a rate of 17 cells per mm^3 per year, which is slightly lower that the rate estimated by Plato collaboration [254]. At virological failure the viral load increases above the detection value, reflecting the failure of ART. According to the model, at this point the viral load increases to about 83% of the level at ART initiation. This value is close to the estimates reported in the literature by Sanders *et al.,* who report that rebound viral load levels are less than the original levels by about 10% [255-257].

The total number of regimen switches without virological failure is predicted to be 4.2 in the NNRTI group and PI/r group. This is the variable where the highest percent difference between groups occurs. Results are in alignment with input parameters (namely, the statistically significant difference found in the survival curves of such event) and this difference has also been found in clinical trials results [145, 258].

The ART Cohort Collaboration [100, 101] estimates that life expectancy of HIV infected individuals at the age of 20 is about two-thirds of that in the general population in each country and that overall, at the age of 35, life expectancy is 31.7. A recent analysis on mortality among all HIV-infected persons receiving care in Denmark [157] indicates a higher life expectancy but as noted by the authors, it is likely to be country specific. In Portugal, according to INE [156] the general population life expectancy[88] at the age of 39 is 39.5, thus two-thirds would be 26.6, which is close to the 24.3 predicted by the model. The difference found is reasonable given the higher proportion of both IDUs and AIDS diagnosis (two factors associated with higher mortality rates) in our sample compared to the sample used in the ART Cohort Collaboration study.

It should, nevertheless, be noted that mortality was estimated with data from both the CVEDT database, a national database of all registered HIV infections in Portugal, and the INE HIV mortality records. Discrepancies among sources indicate that the national database is either incomplete or not fully updated. Moreover, neither of those sources provides a link to clinical variables. Changes in the epidemiology of the disease (namely age of infection and transmission

[87] If the individual is monitored on a regular basis, routine checkups will detect an increase in viral load above the expected level, before it has an impact on CD4+ cell count and, in that case, CD4+ cell count is not expected to decrease at virological failure. Such was assumed in the model, although, in practice, because patient do not always comply with the recommended follow-up scheme, CD4+ cell count will, in fact, decrease before virological failure is detected. Since such variation in CD4+ cell count was not modeled, it is estimated that CD4+ cell count will, on average, increase in all lines except in non-suppressive therapy.

[88] Weighted average according to the observed gender distribution of HIV infected individuals.

mode) are also likely to have a significant impact on mortality predictions. Life expectancy predicted by the model should, thus, be considered with caution.

Of these 24 years, the microsimulation model predicts that an average of 9 years will be spent in first line. This value is identical to the predicted value obtained using a Markov model parameterized with the same CHC database [248], but it should be noted that while matching the data, the average time in first line with *currently* recommended regimens is perceived in general, by physicians, to be longer than 9 years. The estimated time in first line is significantly lower than that implied from the data published by Beck *et al.* [8], but their definition of end of line differs from ours[89] so the results are not comparable.

We predict a 5 months difference in the average time spent in first line in the two groups, with favoring the 2NRTI+NNRTI option. Once again our results contrast with those of Beck *et al.* but coincide with those of Geretti *et al.* [146] who uses a definition of failure identical to ours. The predicted average number of years spent in each subsequent line is decreasing in line number. This result is in accordance with the available literature [259, 260].

The total number of QALYs predicted by the model using estimates on a Portuguese sample is lower than the total number of QALYs predicted using Simpson *et. al.* [4] estimates. A number of reasons could explain this result. In Simpson *et al.* estimates are presented for pairs of CD4 and viral load, while only CD4 was considered in Catarino's [238] estimates. Sample sizes are also significantly different, and one study uses clinical trial data while the other uses observational data. It may also be that HIV infected individuals in Portugal have a lower quality of life, for the same health status, than in other countries. While further research is need, namely in terms of obtaining results on a more representative and larger sample, the work developed by Catarino raises an important issue, of extreme relevant to cost-effectiveness analysis of HIV strategies.

The total number of quality adjusted life years obtained using DCE estimates is significantly lower than QALY estimates obtained by standard methods. Non-health related factors affecting quality of life are likely to be the main source of the difference found. In the context of a disease, which is as much a "medical" condition as a "social" condition, restricting the economic evaluation to health related quality of life may lead to biased conclusions.

[89] In Beck et al. failure was defined as any change made to the ART containing regimen, which included intensification of regimen by adding any anti-retroviral drug to the regimen or swapping the NNRTI or a PI to another anti-retroviral drug class. changing. Our definition relates to virological failure.

Visible (identifiable by others) adverse events of ART, such as lipodystrophy, have an impact that goes beyond its clinical consequences, as it may, for example, decrease the probability of finding a job. Issues such as regimen simplification take a dimension beyond "convenience" if they reduce the probability of discrimination by facilitating adherence to ART without the embarrassing question of "why are you always taking pills?".

Certainly when considering HIV infection but also in other diseases, there is an urgent need for more research on how to account for non-health aspects in a framework that allows for comparison of resource allocation in different diseases. Meanwhile, quality adjusted life years based on DCE-based utilities may complement the information provided by QALYs, when comparing alternative uses of resources devoted to HIV infection.

The overall cost of treating an HIV infected individual from time of ART initiation to death is estimated to be 375,470€ in the 2NRTI+PI/r group and 7.5% less in the 2NRT+NNRTI group. 79% of the costs are ART drug costs. This 79% weight seems reasonable given the increasing weight of ART cost in the overall treatment cost of HIV infected individuals reported by Krentz *et al.* [261]. According to Krentz *et al.,* this weight rose from 50% in 1996/97 (at the time HAART was introduced) to 70% in 2000/2001 (at the time PIs became available). The sources of direct cost of HIV care do, nonetheless, differ significantly by country (and CD4 cell count) as documented in the literature review by Levy *et al.* [262].

The predicted average annual cost of ART in first line is 9,010€ in the 2NRTI+NNRTI groups and 23% higher in the 2NRTI+PI/r suggesting that the 11,040€ established in the HIV financing program for ART naïve individuals covers first line ART costs, although not first line total costs. It should also be noted that the analysis includes (less expensive) regimens no longer recommend as first choice drugs.

Annual ART treatment cost is, as expected, increasing in therapy line, reflecting resistance development and the need for more expensive ART options[90]. Monitoring and treatment costs also increase in the therapy line, reflecting a need for more health care as the disease progresses.

In the present analysis, ART initiation with 2NRTI+NNRTI, when clinical viable, is a dominant or at least, equally effective cost-saving strategy compared to ART initiation with 2NRTI+PI/r. Two cost-effectiveness analyses have been performed comparing NNRT to PI/r as a third agent in the

[90] As previously discussed, an exception occurs in non-suppressive therapy reflecting the likely underestimation of the current cost.

initial treatment of HIV-1 infected individuals[91]. Both studies, although performed in very distinct settings (Beck *et al.* in the United Kingdom [8] and Walensky *et al.* [7] in Africa) find that initiation with 2NRTI+NNRTI, when clinically viable, is a dominant[92] option when compared to 2NRTI+PI/r and are thus in accordance with our results.

It should nonetheless be noted that the analysis, as those of Walensky *et al.* and Beck *et al.,* was performed using data from a cohort of individuals whose regimens include antiretroviral drugs no longer recommended as a first choice (although, still significant in clinical practice, according to IMS Health data 2011). This is certainly a limitation of our analysis.

Overall, the model predicts values close to those observed in the CHC sample from which the model was (for the most part) parameterized. Individual parameter values predicted by the model (such as, CD4+ cell count growth, % reaching viral suppression, first line duration, etc.) are also close to those published in the literature, when such comparison is possible to be performed.

Developments in antiretroviral therapy have allowed for an increase in life expectancy among HIV infected individuals, as well a reduction of the morbidity associated with the disease. Such improvements are happening on a daily basis and it is important to have an instrument that will easily incorporate such progress, quantifying the costs and benefits associated to it. The goal of the present analysis was to create an instrument able to provide relevant information for an efficient use of resources devoted to HIV infection.

The HIV therapy model is flexible enough to allow for analysis of small details, such as the impact of the daily regimen frequency, as well as more comprehensive approaches. In fact, the model has already been used or is currently being updated and improved to perform several distinct economic evaluation analyses, namely, cost-effectiveness of the initial NRTI pair, cost-effectiveness of one adherence improving strategy and cost-effectiveness of early treatment.

As with any model, the present model has limitations resulting from both the available data and assumptions made. The model may be improved on both dimensions and, in fact, as discussed

[91] Ritcher et al. original model back in 2002, actually perform a similar analysis although the authors compare regimen sequences, rather that first line regimens. Regimens included in that analysis are obsolete and it assumes a 5 year time horizon.

[92] In the work by Beck *et al.* 2NRTI+NRTI is a dominant strategy in patients with CD4+ > 200 cell/µL. In patients with CD4+<200 cell/µL 2NRTI+PI/r yields higher effectiveness but with an incremental cost-effectiveness ratio above that considered acceptable by the National Institute for Health and Clinical Excellence (NICE). Walensky *et al.* find 2NRTI+NNRTI to be a dominant strategy even in areas of high prevalence of NNRTI resistance (up to 76%).

below, some of the limitations are already being address in a new version of the model under development.

With respect to the data available, while all efforts have been made to parameterize the model with Portuguese data, it remains true that the sample used is not a nationally representative sample. Estimated results are based on two main sources of information: the CHC database and the LVHEM database. As previously discussed, the CHC database represents *one* Portuguese hospital, while the LVHEM represents an approximate national sample of HIV individuals: those who have been tested for resistance.

Moreover, due to sample size restrictions, the complete observation period was utilized for estimation. This means that estimates obtained reflect clinical practice for the 2000-2008 period. Due to rapid progress in HIV therapy, it is likely that a closer reflection of reality would be obtained with a larger sample in a more recent period. The underestimation of ART costs in non-suppressive patients is a relevant example; another is the proportion of zidovudine containing regimens in the CHC sample.

When a national database with epidemiological, socio-demographic, therapeutic and clinically relevant information becomes available, the model should be re-parameterized. In the meanwhile, an approximation was obtained using the available databases and we are currently working on updating and improving the existing ones.

With respect to the assumptions made and methodological approaches followed, there is certainly room for improvement.

We have used a very simple approach to the estimation of CD4+ and viral load variation and we have not accounted for the correlation between the two in estimation. Following the literature, Johnston *et al.* [113] applied generalized additive models; an approach that is likely to provide a better fit to the data and avoid simplifying assumptions made in our model, such as the assumption that CD4+ remains constant while the individual is not suppressed, increasing only when viral suppression has been achieved.

Moreover, the main drivers of the model (CD4+ cell count, viral load, resistance, adherence and socio-demographic characteristics) should be jointly modeled. In 2005, Brathwaite *et al.* [263] developed an individual simulation model incorporating adherence and the development of

antiretroviral resistance (from a biological approach[93]). The model was designed to predict the duration of effectiveness of treatment, and distinguish deaths attributable to HIV from deaths that are not attributable to HIV[94]. While not a cost-effectiveness model, this model is unique in that the adherence-resistance relationship estimated by the model is an emergent property of the underlying biological principles instantiated in the model, rather than resulting from particular *a priori* beliefs about the form of this relationship, or from a statistically modeled relationship between adherence and resistance. Viral load and CD4 paths, conditional on adherence and resistance level, were then estimated and validated on large clinical cohorts. Acknowledging the value of this more structural approach, we are incorporating it in the new version of our framework for cost-effectiveness analysis of HIV treatment in Portugal.

One additional limitation regarding resistance is that we have considered a single resistance score and, consequently, a single resistance event. There is evidence in the literature suggesting that the likelihood of resistance development (and more recently, the level of adherence required to avoid resistance development) is drug specific, or at least, drug class specific [148]. Separately modeling resistance to each class has the important advantage of allowing for the selection of the future ART regimens to be based on resistance history. Both Braithwaite *et al.* [263] and Johnston *et al.* [113] (although using different approaches) follow that path and we have also moved in that direction in the new version.

We have assumed that a virological failure is *immediately* followed by a regimen switch. Although in theory such should be the case, in practice a significant lag is often observed in clinical practice[95]. If the regimen is switched immediately after virological failure, less resistance will development and the impact of a detectable viral load on CD4+ is less likely to occur. In our model we have, therefore, assumed CD4 does not decrease upon virological failure, but in clinical practice, this assumption may be unrealistic. If so, costs will be underestimated and

[93] *"At the start of the simulation, patients have "wild-type" HIV virus. With each passing day, combination therapies may give rise to HIV mutations by means of selection pressures on viral replication. Each HIV mutation may or may not result in resistance to one or more antiretroviral drugs. As resistance accrues, the viral replication rate increases, which in turn increases the probability that subsequent mutations will develop. Adherence, viral resistance, and other patient characteristics together determine the level of effectiveness of combination therapies, as manifested by changes in CD4 count and viral load. Adherence not only affects the extent to which combination therapies reduce viral replication, but it also directly affects the selection pressures for new mutations."*

[94] *The model, programmed in Decision Maker for Windows (Version Beta 0.99.11.12a, New Brunswick, NJ), was calibrated and validated using data from large clinical cohorts. A cohort of 10,000 antiretroviral-naive patients was simulated in the original article. Initial patients' characteristics were identical but each patient proceeded through the model with a distinct clinical trajectory (determined by individual time-varying characteristics), reflecting the heterogeneity in actual patients.*

[95] Work in progress by Dr. Ricardo Camacho

QALYs overestimated. In the new version of our framework, this issue has been corrected by including an event which is "physician appointment" and imposing that a regimen switch may only occur upon such event.

Models such as CEPAC [109] focus on OI and complications, thereby extensively discriminating between different OIs and modeling, in detail, the relationship between clinical parameters and the probability of each OI. Our model is more in the spirit of that of Ritcher *et al.* [110], focusing on potential impact of different durations of HIV treatment effectiveness and we, therefore, simplify in terms of OIs. Due to lack of data on OI and AIDS events in the CHC and LVHEM databases, we model inpatient care episodes (that is the consequence rather than the cause) and use data from May *et al.* [154, 155] to model an generic AIDS event. An alternative approach to modeling AIDS events is to use the EuroSIDA risk score which has the advantage of depending on ART duration, socio-demographic characteristics and clinical parameters.

Progresses in treatment of HIV, namely the introduction of ART, has prolonged the asymptomatic period of the disease and reduced the incidence of OI in the symptomatic period [231]. In this context, adverse events of ART become at least as relevant as OIs. Drug toxicity is, currently, the most common reason for treatment discontinuation [144], and substitution of individual ART components to address adverse events is believed not to be associated with an increased risk of virologic failure. Accounting for the cause of regimen switch when performing cost-effectiveness analysis of HIV treatment has thus become an essential feature, but it is still not considered in several cost-effectiveness analysis recently published [264, 265]. In our model, as well as that of Kauf *et al.* [111][96], the cause of regimen switch is accounted for.

Nonetheless, two limitations regarding adverse event modeling remain in our model. First, we did not have information on adverse events of ART. Consequently, it was assumed that, with a given probability (taken from the Swiss cohort), a regimen switch without virological failure was due to toxicity. We, thus, aggregate adverse events in a single event and link its frequency to regimen class through the frequency of regimen switch. When comparing specific drugs, this approach will not correctly fully capture differences in drugs' toxicities. Secondly, our model (as well as all others in the literature) does not explicitly model long-term safety. As infected individuals are living longer, the long-term impact of ART therapy is becoming more evident. Issues such as renal disease, cardiovascular risk and premature aging have an impact on both costs and QALYs, and will certainly require cost-effectiveness models to adapt to these new

[96] Kauf et al. ADVANCE model is, in essence, Ritcher *et al*. model.

challenges [266]. While CD4 and viral load continue to be the main markers for disease progression, other clinical parameters will need to be modeled (lipid profile, creatinine, bone density, etc) and its consequences quantified.

Our model considers four therapy lines. While it may be argued that after two virological failures individuals have similar characteristics with respect to disease progression, allowing for more therapy lines (as in Kauf *et al.* [111]) would permit a clearer evaluation of actions affecting the subset of HIV infected individuals. The consequences of our simplification are thus more pronounced when evaluating treatment strategies for experienced patients.

Individual characteristics considered in our model are in line with those of other currently available models. Nonetheless, the application of the model to the cost-effectiveness of 2NRTI+NNRTI versus 2NRTI+PI/r has raised the question on whether we are accounting for all relevant characteristics. In the sample of naïve to ART individuals used for that analysis, essentially no difference was found in baseline characteristics of each group. Nonetheless, this individuals were not randomized, which suggests that relevant characteristics (behind the assignment to one or the other group) are not being considered.

Our model was developed in TreeAge®. This software is inexpensive and may be considered user-friendly. Nonetheless, experience has shown us that for extensive model utilization, it may not be the most adequate software. We have faced significant problems with both the random number generator (for example, when randomly sampling we often obtain 10 identical values – up to Microsof Excel® precision) and unwarned /unexplained crashes (which force program shut-down and imply a return to a previous version of the model). The new version of the framework for cost-effectiveness analysis of HIV is being developed in AnyLogic®, which seems to avoid those problems but is far from being user-friendly, since it requires extensive Java coding.

We finalize this discussion on the limitations of the model by looking *at the big picture* in terms of future research, and in the context of providing an instrument to support Health Policy decisions regarding HIV infection. In doing so, we point out to three essential, but for the most part, unexplored topics, which are not covered in our model.

The first concerns the infectious nature of the disease. There is an increasing awareness of HIV infection as a public health matter fomented, namely, by the acknowledged preventive role of ART among infected individuals [249] and, most recently, when used as pre-exposure prophylaxis (PrEP) prevention [267]. Infectious disease models are commonly used in

epidemiology to model the epidemiological impact of interventions. Since these interventions consume scarce healthcare resources, decision makers are increasingly interested in combining such models with pharmacoeconomic techniques in order to investigate whether or not interventions provide good value for the resources invested [268]. Johnston *et al.* [113] provides an excellent example of the need to account for avoided transmission due to ART when determining the benefits of access to ART in British Columbia and Aragão *et al.* [269] provide evidence that early ART initiation may be cost saving when the transmission is reduced due to earlier suppression of HIV infected individuals.

The second concern the perspective adopted in cost-effectiveness analysis. While data on indirect costs of HIV infection are not readily available and especially difficult to obtain given the social implications of the disease, a most recent (May 2011) analysis performed on behalf of the National Coordination for HIV/AIDS infection [270] provides evidence of the non-negligible impact of HIV infection on labor conditions.

The third concerns the need for structural models in HIV. In a recent review of mathematical models of HIV/AIDS interventions and their implications for policy, Johnson *et al.* [271] list numerous areas where such models may be of use to policy makers. Since structural models, where behavioral changes associated with each intervention is explicitly considered, are required to account for the full impact of health policy interventions, this will certainly be an area of prominent research.

Although the number of new diagnosis reported to the Communicable Diseases and Epidemiological Surveillance Center is decreasing[97], the trend on new infections is unknown. Overall, a significant and increasing amount of resources is used to prevent, treat and improve life expectancy and quality of life of HIV infected individuals in Portugal. In such context, it is important to have an instrument able to quantify the tradeoff between costs and benefits associated with resource allocation decisions.

The goal of the present analysis was to create an instrument able to provide relevant information for an efficient use of resources devoted to HIV infection in Portugal. While more

[97] Both in terms of the number of notifications each year and in terms of new diagnoses notified in the year of diagnose.

tasteful ingredients will result in a tastier meal, the recipe has value on its own. The instrument developed is able to reproduce the patterns observed in the data available, innovating not only in terms of the methodology used but also in terms of being an in depth readily available instrument to support resource allocations within the context of the National Program for Prevention and Control of the HIV/AIDS Infection.

5. Overall conclusion

In this dissertation we reinforce the need for context specific pre-analysis of the impact of health care regulation. Through a series of essays applied to distinct areas of health economics, we aim at analyzing issues of relevance in the national context. While research has value of its own no matter the country at focus, we believe it to be important to contribute to the (relatively) less abundant research in, or relevant to, the Portuguese context.

The first chapter focuses on health insurance in the United States, where private health insurance plays a major role financing healthcare, learning from their experience will allow us to be better prepared for the already growing weight of private health insurance in Portugal. It is estimated that 2 million people had a private health insurance health plan in Portugal in 2009 [33] and access to private health care services has tripled in the last 30 years[98]. Our analysis suggests that there are potential efficiency gains to be obtained by adequate regulation of this growing subsector. Namely, we find that adequate pricing and premium payment system regulation are likely to draw sounder bases in which the sector may grow.

In Chapter 2 we transpose the ideas developed in the first chapter to a distinct sector which we analyze with national data. We believe that this approach is valuable; the ability to apply existing concepts to different settings and contexts extents the utility and benefit from the already established knowledge. By building on the product survival literature, we provided evidence of price regulation as a determinant to product withdrawal thus suggesting that a more in-depth analysis should be performed in order to estimate the welfare impact of such regulation and possible alternatives to the current price regulation in the pharmaceutical market in Portugal.

Chapter 3 draws attention to the role of economic evaluation as a support instrument for regulation in the health sector. By expressing results in a "common unit" (the incremental cost per QALY or per life year) economic evaluation is a useful instrument for efficient resource allocation. While a theoretical, long lasting and valuable discussion goes on as to whether the incremental cost per QALY is the adequate measure, truth is decisions are being made based on such an indicator. In such context, cost-effectiveness framework of HIV treatment may not be the ultimate instrument but it is certainly a necessary and useful one. Chapter 3 provides a

[98] Diário de Notícias, June 13th, 2009. http://dn.sapo.pt/inicio/portugal/interior.aspx?content_id=1306442

model, general enough to allow, among others, for comparison of antiretroviral treatment options, of different cut-offs for treatment initiation or of different adherence promoting strategies. Parameterized with observational data, this framework provides a global picture for decision makers, highlighting reality beyond clinical trials. Cohort data is crucial to correct factors that are overestimated in clinical trial settings, such as adherence.

As a whole, we believe that the work developed is able to emphasize relevance of economic analysis as an adequate and supportive instrument for health care regulation. Moreover, it underlines the broad spectrum of such analysis. We built our evidence on three approaches: a theoretical model, a reduced-form econometric analysis and an economic evaluation framework. These distinct methodologies, regardless of respective advantages and limitations, provide relevant and valuable information for decision makers.

It contributes to the existing literature in two main aspects. First, in each chapter, we have attempted to move one *small* step forward in terms of methodology either by developing a new model, by gathering statistical methods from different research areas within health economics or improving on the existing modeling approaches. Secondly, we have focused on relevant issues in the Portuguese context with an effort to provide information on country specific data. Lack of national data is a long lasting and recognized problem in Portugal which limits the country's ability to understand its reality, act upon evidence and define preventive strategies.

Nonetheless our efforts, it was not our intention to create an all-embracing and global solution to any of the problems analyzed. Health policy decision requires even more in-depth analysis, namely, dynamic equilibrium models. This occurs because, as noted by Lucas back in 1976 [272], regression coefficients are unlikely to remain constant when the policy regime changes substantially because agent behavior changes in response. Thus, for policy analysis, it is necessary to use structural models, on both the demand and supply side, with "deep" parameters and behavioral rules (such as value maximization) that are unlikely to change when the policy regime changes.

That was not our goal either. The purpose of this research work was to highlight the relevance of health economics as a support instrument for the regulator role most would agree the Government should have in the Health sector. As in any other aspect of life, the ability to anticipate the impact of current and future actions helps us make the correct choices or at least, helps us make an informed guesses about what the right choices are, given the information available.

Although imbedded in a common goal, this dissertation is composed of three essays. It is therefore of relevance to discuss the main findings, the relation to the existing literature, the limitations and pointers to future work within the context of each essay. We dedicate the remaining of this section to such chapter specific discussion.

Death Spirals, Switching Costs, and Health Premium Payment Systems

The existing insurance market literature has generally modeled equilibrium patterns of adverse selection without trying to model the dynamics of health plan entry and exit while in the non-health literature, a variety of models examine dynamic pricing strategies with switching costs. We develop a model of health plan competition and pricing in order to understand the dynamics of health plan entry and exit, in the presence of switching costs and alternative health premium payment systems.

Several empirical published papers support the idea of death spiral, biased selection and switching costs in the health insurance market [23-29, 31]. Pauly *et al.* [30], on the other hand, argue that the phenomenon interpreted as death spiral may instead be an adjustment towards more preferred products which would have occurred even in the absence of adverse selection. Our model provides a rational for death spirals even in the presence of identical plans.

We use the model to simulate how premium cost sharing schemes, risk-adjustment policies and the imposition of price floors, affects entry and exit, as well as health plan pricing and enrollee choice patterns. These simulations allow for several insights. First, plan deaths impose much higher switching costs than those generated by voluntary switches, since higher prices induce the lowest switching cost enrollees to switch plans, while plan deaths force those with the highest switching costs to involuntarily change plans. Second, in the presence of switching costs, cost sharing is potentially harmful because it increases plans' market power. Third, by contradicting the forces driving prices up (i.e., the increase in costs due a deterioration of the mix of enrollees), risk adjustment creates the right incentives for plans not to follow a death spiral pattern. This result is in line with the recently published analysis on the determinants of plan switching by Cutler *et al.* [25]. By slowing down death spirals, risk adjustment also decreases switching costs, thus increasing consumers' welfare with no impact on profits.

Among our most interesting findings is the possibility that, a first period price floor strategy, can potentially be welfare improving. In our simulations, price floors were found to be extremely effective at reducing switching costs.

We are perfectly aware of the fact that our simple model may not capture many relevant aspects and certainly the model we built is limited in a number of ways that could be relaxed and studied in future research. The first, is that it assumes the same (uniform) distribution of switching costs for healthy and sick enrollees. Second, we focus on the simple case in which there are only two types of consumers, healthy and sick, whereas in the real world there are a continuum of types, with different demand responsiveness and treatment costs. Third, we focus on the case in which plans are all identical, and the only switching cost is a fixed cost for changing plans. An alternative framework, that might be interesting to contemplate, is one that is closer to a matching model, in which consumers derive utility each period that depends on the quality of their match with a health plan. Finally, there is no moral hazard problem in our paper, and plans do not explicitly adopt strategies to affect health plan choices other than through pricing. If service distortion and/or explicit dumping are permitted, it would clearly change the optimal pricing, entry, and exit decisions.

Price Regulation and Product Survival in the Portuguese Pharmaceutical Market

The aim of the analysis in Chapter 2 was to understand whether price control measures have an impact on product life-cycle. While novel product introductions are expected to have a positive welfare impact due to therapeutic improvements, replacement of existing products carries costs to society (switching costs, marketing costs and licensing costs) with little or no benefit to consumers.

In 2000, Danzon and Chao [54] find evidence of lack of competition among manufactures of the same compound found in more regulated markets. In interpreting the results, the authors suggest that price regulation, by driving down real prices over time and consequently creating an incentive for manufacturers to introduce me-too products, may be an explanation for observed lack of competition. Our analysis supports that interpretation by finding evidence that indeed price control measures influence the decision of product withdrawal and consequently may have an impact on turnover rates in the pharmaceutical market. Moreover, we do not find evidence of such effect in the non-prescription drug market where price control rules are not

present. This result may partially explain the increased hazard rate of prescription drugs versus over-the-counter drugs estimated by Cardoso [49].

In our analysis of the determinants of product withdrawal, we account for unobservable characteristics which influence perceived quality. By including preferences for observable and unobservable characteristics when constructing indexes of quality, regulation and competition we not only improve on the existing product survival literature but also provide further evidence of the impact of price regulation on product withdrawal.

While only a dynamic structural model of entry would allow for a full understanding of the impact of price regulation measure or any given health policy, the analysis here performed provides empirical evidence on the fact that indeed price control measures have an impact on product withdrawal and may thus lead to higher turnover rates. We also find quality to be a major determinant of higher survival rates, reinforcing the value of controlling for (observed and unobserved) quality when analyzing determinants of survival.

Given the fact that Cardoso [49] performs an analysis of the determinants of product exit in the Portuguese pharmaceutical market and that part of their analysis is performed in the 1996-2003 period which almost coincides with ours (1997-2003), it is of special interest to compare results obtained. In doing so, it should be noted that not only the focus of the analysis but also the methodology and the datasets used diverge. A major difference is that price and quantity information in which most of our explanatory variables were constructed were not available in their analysis. Even so, it is worth noting that comparable results obtained in the two analyses do not contradict each other.

We find that the number of similar[99] products (owned by the firm or competitors) reduces the probability of exit while the number of competitors (similar or close substitutes) in the same market segment (ATC3) does not have a statistically significant impact on the hazard rate. These results are in line with those obtained by Cardoso [49].

We do not find a statistically significant effect of the number of own firm products in the market segment (ATC3) on the exiting probability but we do find that quality differentiation among those products has, as expected, a protective effect on the probability of withdrawal while a broadening of quality spectrum covered by the firm increases the hazard rate. If similar products were synonymous to identical quality, these results would be in contrast with those of

[99] Defined as having the same ICD, dosage and pharmaceutical form. These are named “identical” in Cardoso analysis.

Cardoso, who estimates that (own firm) similar products have a hazard ratio lower than one while (own firm) close substitutes have a hazard ratio higher than one.

Our analysis and that of Cardoso estimate a protective effect of firm size on product exit and an increased hazard in the 2001-2002 period of major regulatory changes.

The analysis in Chapter 2 has several limitations. First, in terms of data, we have a short span with data on an annual basis. Such structure limits the ability to evaluate product replacement, since gaps between exit and entry may exist due to licensing (and reimbursement) timings. Félix *et al.* [85] report a median of 10 months for reimbursement requests processes in Portugal in 2007 after the licensing process has been completed, which itself will take months. The period of analysis (1997-2003) is also not ideal for the analysis of the phenomenon under study since several confounding factors occurred simultaneously, namely, new regulation on generic products which led to a substantial increase in the number of entries unrelated to the phenomenon at stake [36, 37, 49].

Second, in the pharmaceutical market, physicians act as patients' agents and, for the most part, select the products to be consumed by patients. Demand for pharmaceutical is thus a mixture of physicians' and patients' preferences. While estimating demand parameters it would be of interest to recover both sets of preferences but with the data available only the mixed bundle was estimated. Moreover, the fact that physicians' act as patient agents aggravates the issue of unit demand usually assumed in the demand for differentiated products literature, since the same bundles are more likely to be repeated [75]. Furthermore, demand for pharmaceuticals in inherently dynamic in the sense that a consumer who buys (a physician who prescribes) product j today is more likely to buy (to prescribe) product j tomorrow.

Third, dynamic effects are also present on the supply side and products' entry and exit decisions are based on multiproduct firms, who optimize their portfolio over time. None of these issues were accounted for in the present analysis. Only a dynamic structural model would allow for a full understanding of the impact of any given health policy and that is beyond the scope of the present paper. Further research is thus necessary not only to estimate the welfare impact of such turnover (when switching costs, licensing costs and marketing costs are taken into consideration) but also to develop a model where alternative price regulation policies may be compared. While not explicitly accounting for switching costs, within a behavioral dynamic equilibrium model, Filson [86] estimates the welfare impact of price controls in the pharmaceutical industry arguing that removing price controls increases firm value, research and

development, the flow of new drugs, and consumer welfare globally, but reduces consumer welfare in the countries changing their policies.

A Framework for Cost-Effectiveness Analysis of HIV Treatment in Portugal

Although the number of new diagnosis reported to the Communicable Diseases and Epidemiological Surveillance Center is on a decreasing trend, the trend on new infections is unknown. Overall, a significant and increasing amount of resources is used to prevent, treat and improve life expectancy and quality of life of HIV infected individuals in Portugal. In such context, it is important to have an instrument able to quantify the tradeoff between costs and benefits associated with resource allocation decisions. The goal of the analysis in Chapter 3 was thus to create an instrument able to provide relevant information for an efficient use of resources devoted to HIV infection in Portugal.

We are confident that microsimulation model is the most adequate approach for cost-effectiveness analysis of HIV treatment and likely the one to prevail in the future research in HIV infection. Three most recent articles published on broad cost-effectiveness issues relating to HIV such as routine screening [273], expanding access to HAART [113] or HIV testing referral strategies [274] make use of microsimulation models. Moreover, Lamontagne *et al.* [275] point out the availability of microsimulation models as a promising path towards the current evaluation of the macroeconomic impact of HIV.

The instrument developed is able to reproduce the patterns observed in the data available, innovating not only in terms of the methodology used but also in terms of being an in depth readily available instrument to support resource allocations within the context of the National Program for Prevention and Control of the HIV/AIDS Infection.

Moreover, the results obtained in the model application example – the cost-effectiveness analysis of 2NRTI+NNRTI versus 2NRTI+PI/r as a first line antiretroviral therapy regimen are in line with those obtained by Walensky *et al.* [7] in Africa and Beck *et al.* [8] in the United Kingdom. All three analysis reinforce the British HIV Association 2008 [9] recommendations, NNRTIs are explicitly recommended as third agent of choice for treatment naïve patients unless clinical factors (such as resistance) prevent such option.

In developing the framework for cost-effectiveness analysis of HIV treatment, it was our goal that the model could effectively be used to support regulatory decisions, namely complementing the clinical evaluation of the options available provided by guidelines. As such, we hope the framework will, in the future, be available to both the medical community and the National Coordination for HIV/AIDS Infection for any analysis of national relevance.

The optimal timing for therapy initiation, in terms of CD4 cell count, is an issue under debate [276-278] and given the limitation of the Markov model approach followed in our previous work [248], we are currently working on applying the microsimulation discrete events model here developed to a cost-effectiveness analysis of early treatment initiation.

We are also using it to estimate the cost-effectiveness of adherence improvement. The four papers published on the cost-effectiveness of adherence or adherence interventions [279-282] indicate that strategies to improve adherence in developed countries are cost-effective (although not cost saving). Nonetheless, results depend highly on the effectiveness of the interventions effectiveness and associated costs, the epidemiology parameters and surely, on country specific clinical practice and costs. Moreover, there is an ongoing debate on the required level of adherence to obtain maximal benefits from currently available drugs which has not been included in previous studies. The debate is based on the fact that clinical trials' levels of adherence are not observed in real life and yet viral suppression is maintained thereby suggesting that the assumed 95% threshold may no longer be required given the developments in antiretroviral therapy [147].

As with any model, the framework developed in Chapter 3 has several limitations. The first limitation refers to the data available, namely, the need for make use of several unlinked databases, the non-representativeness of the CHC dataset and an analysis period from 2000 to 2008 which, due to rapid progress in HIV therapy, may no longer reflect current practice. The second limitation concerns simplifying assumptions made, such as a single (non-class specific) resistance event and the abstraction from long-term safety issues. The third limitation relates to its unaccountability for the infectious nature of the disease, which may only be fully captured by an epidemiological module. The last, but certainly not least limitation, of the model concerns the fact the present model is, by no means, a sufficient instrument for health policy analysis,

because evaluating the impact of most health policy interventions requires a structural model where behavioral changes associated with each intervention are accounted for[100].

Given these limitations, pointers to further research include (i) obtaining nationally representative cohort data for model validation or, at the least, updating the existing databases, (ii) explicitly modeling the selection of the components of next regimen based on the reason for switching and the development of resistance based on viral replication and adherence and (iii) accounting for antiretroviral therapy long-term safety issues, (iv) developing an epidemiological model to be merged to the present framework, and (v) reprogramming it in a more adequate software. These small steps forward are already under way. The ultimate goal is to move from a regression-based approach to a structural one where the welfare impact of HIV related health policies may indeed be evaluated.

The research work here developed aimed at highlighting the relevance of health economics as a support instrument for Governmental regulation in the Health sector. It drew attention to potential and future problems that may be averted through proper regulation; it identified potential problems with current regulatory options, and generated a usable instrument to support management decisions within one disease area. While certainly a small step in a long path, this dissertation will have served its purpose if it inspires others to perform, produce or make use of more and better economic analysis aiming towards more efficient health regulation.

[100] For example, it may happen that less discrimination changes the mixing pattern, in terms of sexual activity, between infected and non-infected individuals, which in turn will have an effect on the number of new infections and thus on costs and effectiveness of a given “anti-discrimination policy”.

References

1. Mark, A., J.D. Hall, et al. *U.S. Senate Committee on Finance, Hearing on "47 Million and Counting: Why the Health Care Marketplace is Broken"*. 2008 [cited 2010 May 1st]; Available from: http://finance.senate.gov/imo/media/doc/061008MHTest.pdf.
2. Lake Research Partners. *Reducing Enrollee Churning in Medicaid, Child Health Plus, and Family Health Plus*. 2009 [cited 2010 May 1st]; Available from: http://www.nyshealthfoundation.org/userfiles/file/LakeResearch_2_2009.pdf.
3. Barton, P., P. Jobanputra, et al., *The use of modelling to evaluate new drugs for patients with a chronic condition: the case of antibodies against tumour necrosis factor in rheumatoid arthritis.* Health Technology Assessment 2004. **8**(11): p. 1-91.
4. Simpson, K.N., M.P. Luo, et al., *Cost-effectiveness of lopinavir/ritonavir versus nelfinavir as the first-line highly active antiretroviral therapy regimen for HIV infection.* HIV Clin Trials, 2004. **5**(5): p. 294-304.
5. Simpson, K.N., A. Strassburger, et al., *Comparison of Markov model and discrete-event simulation techniques for HIV.* Pharmacoeconomics, 2009. **27**(2): p. 159-65.
6. Aragão, F., I. Vaz Pinto and J. Vera, *Cost-effectiveness of 2NRTI+NNRTI versus 2NRTI+PIr as the initial cART regimen.* Journal of the International AIDS Society 2010. **13**(Suppl 4): p. P237.
7. Walensky, R.P., M.C. Weinstein, et al., *HIV drug resistance surveillance for prioritizing treatment in resource-limited settings.* AIDS, 2007. **21**(8): p. 973-82.
8. Beck, E.J., S. Mandalia, S. Lo Gary, et al., *Cost-Effectiveness of Early Treatment with First-Line NNRTI-Based HAART Regimens in the UK, 1996-2006.* PLoS ONE, 2010. **6**(5): p. e20200. doi:10.1371/journal.pone.0020200.
9. Gazzard, B.G., *British HIV Association guidelines for the treatment of HIV-1-infected adults with antiretroviral therapy 2008.* HIV Medicine, 2008. **9**: p. 563–608.
10. Cutler, D.M. and S.J. Reber, *Paying for Health Insurance: the Tradeoff Between Competition and Adverse Selection.* Quarterly Journal of Economics, 1998. **113**(2): p. 443-466.
11. Rothschild, M. and J.Stiglitz, *Equilibrium in Competitive Markets: an essay on the Economics of Imperfect Information.* Quarterly Journal of Economics, 1976. **90**: p. 629-649.
12. Jenson, G., R. Feldman and B. Dowd, *Corporate benefit policies and health insurance costs.* J Health Econ, 1984. **3**(3): p. 275-296.
13. Dowd, B.E. and R. Feldman, *Premium Elasticities of Health Plan Choice.* Inquiry, 1997. **31**(4): p. 438-44.
14. Bradford, W.D., *Efficiency in Employment-based health insurance: The potential for supra-marginal cost pricing.* Economic Inquiry, 1996. **34**(2): p. 341-56.
15. Dowd, B., R. Feldman and J. Klein *What do employers really want in a health plan?* Bus Health, 1987. **4**(3): p. 44-48.
16. Cutler, D.M. and R.J. Zeckhauser, *The anatomy of Health Insurance*, in *Handbook of Health Economics*, A.J. Culyer and J.P. Newhouse, Editors. 2000, Elsevier: ,Amsterdam.
17. Neipp, J. and R. Zeckhauser, *Persistence in the Choice of Health Plans*, in *Advances in Health Economics and Health Services Research: Biased Selection in Markets for Health Care*, R. Scheffler and L. Rossiter, Editors. 1985. p. 47-72.
18. Keeler, E.B., G.M. Carter and J.P. Newhouse, *A model of the impact of reimbursement schemes on health plan choice.* J Health Econ, 1998. **17**: p. 297-320.
19. Vistnes, J.P., P.F. Cooper and G. Vistnes, *Employer Contribution Methods and Health Insurance Premiums: Does Managed Competition Work?* International Journal Health Care Finance Economics, 2001. **1**(2): p. 159-87.
20. Ellis, R.P. and C.-t.A. Ma, *Health Insurance, Cost Expectations, and Adverse Job Turnover.* Health Econ, 2011. **20**(1): p. 27-44.

21. Beggs, A. and P. Klemperer, *Multi-Period Competition with Switching Costs.* Econometrica, 1992. **60**: p. 651–666.
22. Gale, D. and R. Rosenthal, *Price and Quality Cycles for Experience Goods.* RAND Journal of Economics, 1994. **25**(4, Winter): p. 590-607.
23. Strombom, B.A., T.C. Buchmueller and P.J. Feldstein, *Switching Costs, Price Sensitivity and Health Plan Choice.* J Health Econ, 2002. **21**(1): p. 89-116.
24. Grönqvist, E., *Does Adverse Selection Matter? - Evidence from a Natural Experiment*, in *SSE/EFI Working Paper Series in Economics and Finance No. 575*. 2004, Stockholm School of Economics.
25. Cutler, D.M., B.Lincoln and R. Zeckhauser, *Selection stories: Undestanding movements across health plans.* J Health Econ, 2010. **29**: p. 821-838.
26. Thomasson, M.A., *Early evidence of an adverse selection death spiral? The case of Blue Cross and Blue Shield* Explorations in Economic History, 2004. **41**(4): p. 313-328.
27. Castano, R. and A. Zambrano, *Biased selection within the social health insurance market in Colombia.* Health Policy, 2006. **79**(2): p. 313-324.
28. Cebul, R., J.B. Rebitzer, et al., *Unhealthy Insurance Markets: Search Frictions and the Cost and Quality of Health Insurance* NBER Working Paper Series, 2008. **w14455**.
29. Frank, R.G. and K. Lamiraudb, *Choice, price competition and complexity in markets for health insurance.* Journal of Economic Behavior & Organization, 2009. **71**(2): p. 550-562.
30. Pauly, M., O. Mitchell, and P. Zeng, *Death Spiral or Euthanasia? The Demise of Generous Group Health Insurance Coverage.* Inquiry, 2007. **44**(4): p. 412-427.
31. Handel, B., *Adverse Selection and Switching Costs in Health Insurance Markets: When Nudging Hurts*, in *Department of Economics*. 2009, Northwestern University.
32. Glazer, J. and T.G. McGuire, *Optimal Risk Adjustment in Markets with Adverse Selection: An Application to Managed Care.* American Economic Review, 2000. **90**(4(Sep., 2000)): p. 1055-1071
33. Ramos, A. and C. Silva. *Evolução do seguro de saúde em Portugal [Analítico de Periódico]*. 2009 [cited 2011 August 1st]; Available from: http://www.isp.pt/winlib/cgi/winlibimg.exe?key=&doc=18235&img=2837.
34. Yu, W., R.P. Ellis and A.S. Ash, *Risk selection in the Massachusetts state employee health insurance program.* Health Care Management Science, 2001. **4**(4): p. 281-7.
35. Organisation for Economic Co-operation and Development. *OECD Health data 2011*. 2011 [cited 2011 August 1st.]; Available from: http://www.oecd.org/document/16/0,3343,en_2649_34631_2085200_1_1_1_1,00.html.
36. Oliveira, M.D. and C.G. Pinto, *Health care reform in Portugal: an evaluation of the NHS experience.* Health Econ, 2005. **14**(Suppl 1): p. S203-20.
37. Barros, P.P. and L.C. Nunes, *The impact of pharmaceutical policy measures: a structuralbreak approach*, in *FE: EG - Working Papers*. 2010, Universidade Nova de Lisboa: Lisbon.
38. Danzon, P.M. and A.J. Epstein, *Effects of Regulation on Drug Launch and Pricing in Interdependent Markets.* NBER Working Paper Series, 2008(w14041).
39. Portaria n.º 300-A/2007 de 19 de Março de 2007, in *DR I Série*, n.º 55, de 19 de Março de 2007 - Ministério da Economia e da Inovação e da Saúde.
40. Decreto-Lei n.º 134/2005 de 16 de Agosto de 2005, in *DR I Série A*, nº 156, de 16 de Agosto de 2005 - Ministério da Saúde.
41. Portaria n.º 29/1990 de 13 de Janeiro de 1990, in *DR I Série B*, n.º , de 13 de Janeiro de 1990 - Ministérios da Economia e da Saúde.
42. Decreto-Lei n.º 65/2007 de 14 de Março de 2007, in *DR I Série* n.º 52, de 14 de Março de 2007 - Ministério da Saúde: .

43. Decreto-Lei n.º 184/2008 de 5 de Setembro de 2008, in *DR I Série*, n.º 172, de 5 de Setembro de 2008 - Ministério da Saúde.
44. Portaria n.º 57/94 de 24 de Janeiro de 1994, in *DR I Série B*, n.º 19 de 24 de Janeiro de 1994, 348 - Ministros das Finanças, da Saúde e do Comércio e Turismo.
45. Portaria n.º 73/99 de 29 de Janeiro de 1999, in *DR I Série B*, n.º 24, de 29 de Janeiro de 1999 - Ministérios da Economia e da Saúde.
46. Portaria n.º 1279/2001 de 14 de Novembro de 2001, in *DR I Série B*, n.º 264, de 14 de Novembro de 2001, 7261-7262 - Ministérios da Economia e da Saúde.
47. Danzon, P.M. and L.W. Chao, *Cross-national price differences for pharmaceuticals: how large, and why?* J Health Econ, 2000. **19**(2): p. 159-95.
48. Danzon, P.M., Y. R. Wang and L. Wang *The impact of price regulation on the launch delay of new drugs - evidence from twenty-five major markets in the 1990s.* Health Econ, 2004. **14**(3): p. 269 - 292.
49. Cardoso, C., *Essays on Product Portfolios in Pharmaceutical Markets*, in *Faculdade de Economia.* 2010, Universidade do Porto: Porto.
50. Stavins, J., *Model Entry and Exit in a Differentiated-product Industry: the Personal Computer Market.* Review of Economics and Statistics, 1995. **77**(4): p. 571–584.
51. Figueiredo, J.M. and M.K. Kyle, *Surviving the Gales of Creative Destruction: The Patterns of Innovative Activity in the Desktop Laser Printer Industry.* Strategic Management Journal 2005. **27**(3): p. 241 - 264.
52. Requena-Silvente, F. and J. Walker, *Competition and survival in differentiated product markets: the case of the UK automobile market (1971-1998).* The Manchester School, 2009. **77**(3): p. 288 - 316.
53. Bajari, P. and L. Benkard, *Demand Estimation with Heterogeneous Consumers and Unobserved Product Characteristics:A Hedonic Approach.* Journal of Political Economy, 2005. **113**(6): p. 1239-1276.
54. Danzon, P.M. and L.-W. Chao, *Does regulation drive out competition in pharmaceutical markets?* Journal of Law & Economics, 2000. **43**(2): p. 311-21.
55. Lanjouw, J.O., *Patents, Price Controls and Access to New Drugs: How Policy Affects Global Market Entry.* NBER Working Paper Series, 2005(w11321).
56. Kyle, M.K., *The Role of Firm Characteristics in Pharmaceutical Product Launches.* RAND Journal of Economics, 2006. **37**(3): p. 602-618.
57. Greenstein, S.M. and J.B. Wade, *The Product Life Cycle in the Commerical Mainframe Computer Market, 1968-1982.* RAND Journal of Economics, 1998. **29**(4): p. 772-789.
58. Asplund, M. and R.D. Sandin, *The Survival of New Products.* Review of Industrial Organization, 1999. **15**(3): p. 229–237.
59. Ruebeck, C.S., *Model exit in a vertically differentiated market: interfirm competition versus intrafirm cannibalization in the computer hard disk drive industry.* Review of Industrial Organization, 2005. **26**: p. 27-59.
60. Berry, S., J. Levinsohn, and A. Pakes, *Automobile Prices in Market Equilibrium.* Econometrica, 1995. **63**(4): p. 841-89.
61. Ackerberg, D. and M. Rysman, *Unobserved Product Differentiation in Discrete Choice Models: Estimating Price Elasticities and Welfare Effects.* RAND Journal of Economics, 2005. **36**(4): p. 771-788.
62. Berry, S., *Estimating Discrete-Choice Models of Product Differentiation.* RAND Journal of Economics, 1994. **25**(2): p. 242-262.
63. Nevo, A., *Measuring Market Power in the Ready-To-Eat Cereal Industry.* Econometrica, 2001. **69**(2): p. 307-342.
64. Petrin, A., *Quantifying The Benefits of New Products: The Case of the Mini-van.* Journal of Political Economy, 2002. **110**(4): p. 705-729.

65. Rosen, S., *Hedonic prioces and implicit markets: Product differentiation in pure competition.* Journal of Political Economy, 1974. **82**: p. 34-55.
66. Bajari, P. and M.E. Kahn, *Estimating Housing Demand with an Application to Explaining Racial Segregation in Cities.* Journal of Business and Economic Statistics, 2005. **23**(1): p. 20-33.
67. Bajari, P. and M.E. Kahn, *Estimating Hedonic Models of Consumer Demand with an Application to Urban Spraw.* Hedonic Methods in Housing Markets, 2008. **II**: p. 129-155.
68. Portugal. Instituto Nacional da Farmácia e do Medicamento. *Prontuário Terapêutico Online*. 2004 [cited 2004 January 4th]; Available from: http://www.infarmed.pt/infomed/inicio.php.
69. Portaria n.º 914/2003 de 1 de Setembro de 2003, in *DR I Série B*, n.º 201, de 1 de Setembro de 2008 - Ministério da Saúde.
70. United Nations Economic Commission for Europe. *UNECE STATISTICAL DATABASE, Economic Statistics, Price Index*. Consumer Price Index 2010 [cited May 18th, 2010]; Available from: http://w3.unece.org/pxweb/.
71. Iizuka, T., *Experts' Agency Problems: Evidence from the Prescription Drug Market in Japan* RAND Journal of Economics, 2007. **38**(3): p. 844-862.
72. Ridley, D.B., *Payments, Promotion, and the Purple Pill*. 2005, Duke University Working Paper.
73. Fan, J. and I. Gijbels, *Local Polynomial Modelling and Its Applications*, in *Monographs on Statistics and Applied Probability*. 1996, Chapman & Hall: London.
74. Portugal. Instituto Nacional da Farmácia e do Medicamento. *Monitorização da Prescrição por DCI e Implementação da Receita Médica*. 2005 [cited 2009 January 5th]; Available from: http://www.infarmed.pt/portal/page/portal/INFARMED/MONITORIZACAO_DO_MERCADO/OBSERVATORIO/INTRODUCAO_DE_FICHEIROS/receita_medica.pdf.
75. Dubé, J.-P., *Multiple discreteness and product differentiation: Demand for carbonated soft drinks.* Marketing Science, 2004. **23**(1): p. 66-81.
76. Cox, D.R., *Regression models and life-tables (with discussion).* Journal of the Royal Statistics Society, Series B, 1972. **30**: p. 248-275.
77. Hidemichi, Y., H. Nobuyuki and M. Keitaro *Censoring and Truncation in Survival Analysis.* Biotherapy (Tokyo), 1999. **13**(7): p. 811-816.
78. Carvalho, M., V. Andreozzi, et al., *Análise de Sobrevida: teoria e aplicações em Saúde*, ed. F.O. Cruz. 2005: Editora FioCruz.
79. Cleves, M., W. Gould, et al., *An introduction to survival analysis using Stata*. Second edition. 2008: Stata Press.
80. Dias, S.S., V. Andreozzi, et al., *Predictors of mortality in HIV-associated hospitalizations in Portugal: a hierarchical survival model.* BMC Health Serv Res, 2009. **9**: p. 125.
81. Jenkins, S.P., *Easy estimation methods for discrete-time duration models.* Oxford Bulletin of Economics and Statistics, 1995. **57**(1): p. 129-138.
82. Jenkins, S.P., *sbe17: Estimation Methods for Discrete-Time (Grouped Duration Data) Proportional Hazard Models: PGMHAZ.* Stata Technical Bulletin, 1997. **39**: p. 1-12.
83. Grambsch, P.M. and T.M. Therneau, *Proportional hazards tests and diagnostics based on weighted residuals* Biometrika, 1994. **81**: p. 515-526.
84. Alves, A. and C. Furtado. *Mercado de Medicamentos em Portugal: Análise da Evolução da Utilização e da Despesa entre 2003 e 2006*. 2007 [cited 2008 December 15th]; Available from: http://www.infarmed.pt/portal/page/portal/INFARMED/MONITORIZACAO_DO_MERCADO/OBSERVATORIO/INTRODUCAO_DE_FICHEIROS/APES-MT-2003-2006.pdf.

85. Félix, J., S. Rabiais and J. Almeida, *Financiamento Público dos Medicamentos em Portugal: Acessibilidade aos medicamentos com novas moléculas ou novas indicações terapêuticas.*, APIFARMA, Editor. 2010.
86. Filson, D. *A Markov-Perfect Equilibrium Model of the Impacts of Price Controls on the Performance of the Pharmaceutical Industry*. 2007 [cited 2008 July 30th]; Available from: http://www.claremontmckenna.edu/econ/dfilson/Impacts%20of%20Price%20Controls.pdf.
87. Portugal. Centro de Vigilância Epidemiológica de Doenças Transmissíveis. *INFECÇÃO VIH/SIDA, A SITUAÇÃO EM PORTUGAL 31 de Dezembro de 2009*. 2010 [cited 2010 June 6th]; Available from: http://www.sida.pt/aaaDefault.aspx?f=1&back=1&codigono=56265720611461986200AAAA.
88. UNAIDS. *AIDS epidemic update*. 2009 [cited 2010 February 2nd]; Available from: http://data.unaids.org/pub/Report/2009/JC1700_Epi_Update_2009_en.pdf.
89. Portugal. Coordenação Nacional para a Infecção VIH/sida. *Programa Nacional de prevenção e controlo da infecção VIH/Sida 2007-2010*. 2007 [cited 2009 January 7th]; Available from: http://www.portaldasaude.pt/NR/rdonlyres/FE59B2DA-A6B4-4EB1-B524-9B0C658C2493/0/planoprevencaosida.pdf.
90. Portugal. Coordenação Nacional para a Infecção VIH/Sida. *A infecção VIH em Portugal, 2009. Dados hospitais 2009.* 2011 [cited 2011 August 1st]; Available from: http://www.sida.pt/.
91. Associação Portuguesa da Industria Farmacêutica. *A Industria Farmacêutica em Números*. 2010 [cited 2011 January 8th]; Available from: http://www.apifarma.pt/BackofficeApifarma/uploads/IFNumeros2010.pdf.
92. Portugal. Instituto Nacional da Farmácia e do Medicamento. *Análise do Mercado de Medicamentos no âmbito do Serviço Nacional de Saúde*. 2010 [cited 2010 May 6th]; Available from: http://www.infarmed.pt/portal/page/portal/INFARMED/MONITORIZACAO_DO_MERCADO/OBSERVATORIO/ANALISE_MENSAL_MERCADO/ANALISE_MERCADO_MEDICAMENTOS_CHNM/2009/Rel_CHNM_200912.pdf.
93. The Body - The Complete HIV/AIDS Resource. *The HIV Life Cycle*. 2009 [cited 2009 December, 15th]; Available from: http://www.thebody.com/content/whatis/art40989.html.
94. Pantaleo, G., C. Graziosi and A.S. Fauci, *New concepts in the immunopathogenesis of human immunodeficiency virus infection.* N Engl J Med, 1993. **328**(5): p. 327-35.
95. Morgan, D., C. Mahe, et al., *HIV-1 infection in rural Africa: is there a difference in median time to AIDS and survival compared with that in industrialized countries?* AIDS, 2002. **16**(4): p. 597-603.
96. Palella, F.J., K.M. Delaney, et al., *Declining morbidity and mortality among patients with advanced human immunodeficiency virus infection. HIV Outpatient Study Investigators.* N Engl J Med, 1998. **338**(13): p. 853-60.
97. Chene, G., J.A. Sterne, et al., *Prognostic importance of initial response in HIV-1 infected patients starting potent antiretroviral therapy: analysis of prospective studies.* Lancet, 2003. **362**(9385): p. 679-86.
98. Wood, E., R.S. Hogg, et al., *Is there a baseline CD4 cell count that precludes a survival response to modern antiretroviral therapy?* AIDS, 2003. **17**(5): p. 711-20.
99. Schackman, B.R., K.A. Gebo, et al., *The lifetime cost of current human immunodeficiency virus care in the United States.* Med Care, 2006. **44**(11): p. 990-7.

100. The Antiretroviral Therapy Cohort Collaboration, *Life expectancy of individuals on combination antiretroviral therapy in high-income countries: a collaborative analysis of 14 cohort studies.* Lancet, 2008. **372**(9635): p. 293-9.
101. The Antiretroviral Therapy Cohort Collaboration, *Mortality of HIV-infected patients starting potent antiretroviral therapy: comparison with the general population in nine industrialized countries.* International Journal of Epidemiology, 2009. **38**(6): p. 1624-1633.
102. Portugal. Coordenação Nacional para a Infecção VIH/sida. *Recomendações Portuguesas para o Tratamento da Infecção VIH/sida (em discussão)*. 2011 [cited 2011 February 12th]; Available from: http://www.portaldasaude.pt/NR/rdonlyres/030A05A2-12E7-4710-832D-BCE84E17C93B/0/i006008.pdf.
103. Hornberger, J., M. Holodniy, et al., *A systematic review of cost-utility analyses in HIV/AIDS: implications for public policy.* Med Decis Making, 2007. **27**(6): p. 789-821.
104. Giguère, P., *Cost Effectiveness Analysis of Abacavir/Lamivudine Versus Tenofovir/Emtricitabine Combination Therapy as Part of Highly Active Antiretroviral Therapy in Treatment Naive HIV-infected Patients*, in *Department of Epidemiology and Community Medicine*. 2010, University of Ottawa: Ottawa.
105. Kuhne, F.C., J. Chancellor, et al., *Microsimulation or Cohort Modelling? A Comparative Case Study in HIV Infection. .* iHEA 2007 6th World Congress: Explorations in Health Economics Paper. Working Paper Series. Available at SSRN: http://ssrn.com/abstract=994349, 2007.
106. Freedberg, K.A., E. Losina, et al. *The cost effectiveness of combination antiretroviral therapy for HIV disease*. N Engl J Med 2001 Mar 15 [cited 344 11]; 2001/03/15:[824-31]. Available from: http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=11248160.
107. Schackman, B.R., S.J. Goldie, et al., *Cost-effectiveness of earlier initiation of antiretroviral therapy for uninsured HIV-infected adults.* Am J Public Health, 2001. **91**(9): p. 1456-63.
108. Weinstein, M.C., S.J. Goldie, et al., *Use of genotypic resistance testing to guide hiv therapy: clinical impact and cost-effectiveness.* Ann Intern Med, 2001. **134**(6): p. 440-50.
109. Goldie, S.J., Y. Yazdanpanah, et al., *Cost-effectiveness of HIV treatment in resource-poor settings--the case of Cote d'Ivoire.* N Engl J Med, 2006. **355**(11): p. 1141-53.
110. Richter, A., B. Hauber, et al., *A Monte Carlo simulation for modelling outcomes of AIDS treatment regimens.* Pharmacoeconomics, 2002. **20**(4): p. 215-224.
111. Kauf, T., R.A. Farkouh, et al., *Economic Efficiency of Genetic Screening to Inform the Use of Abacavir Sulfate in the Treatment of HIV.* Pharmacoeconomics, 2010. **28**(11): p. 1025-1039.
112. Kuhne, F.C., J. Chancellor, et al., *A microsimulation of the cost-effectiveness of maraviroc for antiretroviral treatment-experienced HIV-infected individuals.* HIV Clin Trials, 2010. **11**(2): p. 80-99.
113. Johnston, K.M., A.R. Levy, et al., *Expanding access to HAART: a cost-effective approach for treating and preventing HIV.* AIDS, 2010. **24**(12): p. 1929-35.
114. Lima, V.D., K. Johnston, et al., *Expanded access to highly active antiretroviral therapy: a potentially powerful strategy to curb the growth of the HIV epidemic.* J Infect Dis, 2008. **198**(1): p. 59-67.
115. Barton, P., S. Bryan and S. Robinson, *Modelling in the economic evaluation of health care: selecting the appropriate approach.* Journal of Health Services Research & Policy, 2004. **8**(2): p. 110-118.

116. Barth-Jones, D.C., A.L. Adams and J.S. Koopman. *Monte Carlo simulation experiments for analysis of HIV vaccine effects and vaccine trial design,*. in *2000 Winter Simulation Conference (WSC'00)*. 2000.
117. Vieira, I.T., P.R. Harper, et al., *Mother-to-Child Transmission of HIV: A Simulation-Based Approach for the Evaluation of Intervention Strategies* The Journal of the Operational Research Society, 2003. **54**(7): p. 713-722.
118. Ying, H., F. Lin, et al., *A self-learning fuzzy discrete event system for HIV/AIDS treatment regimen selection.* IEEE Trans Syst Man Cybern B Cybern, 2007. **37**(4): p. 966-79.
119. Centers for Disease Control and Prevention, *1993 revised classification system for HIV infection and expanded surveillance case definition for AIDS among adolescents and adults.* MMWR Recommendation Report 1993. **18**(41(RR-17)): p. 1-19. .
120. Palisade Corporation. *Guide to Using @Risk. Risk Analysis and Simulation Add-In for Microsoft® Excel*. Version 5.5 2009 [cited 2010 December 3rd]; Available from: http://www.palisade.com.
121. Chadborn, T.R., K. Baster, et al., *No time to wait: how many HIV-infected homosexual men are diagnosed late and consequently die? (England and Wales, 1993-2002).* AIDS, 2005. **19**(5): p. 513-20.
122. May, M., J.A. Sterne, et al., *Prognosis of HIV-1-infected patients up to 5 years after initiation of HAART: collaborative analysis of prospective studies.* AIDS, 2007. **21**(9): p. 1185-97.
123. The Antiretroviral Therapy Cohort Collaboration. *Risk calculator for HIV positive patients starting antiretroviral therapy*. 2009 [cited 2009 February 7th]; Available from: http://www.epi.bris.ac.uk/art-cohort/hiv_form.html.
124. Carvalho, C., *Acesso, Práticas e Barreiras ao Teste VIH.* Arquivos de Medicina, 2009. **23**(2): p. 68-69.
125. Vaz Pinto, I., *Late Diagnosis of HIV Infection in patients followed at the HIV Unit in Centro Hospitalar de Cascais: a retrospective analysis from 2001-2007*. 2009, University of Barcelona: Barcelona.
126. Alfaiate, D., A.C. Miranda, et al., *The impact of late presentation: analysis of a cohort of 313 Portuguese patients.* Journal of the International AIDS Society 2010. **13**(Suppl 4): p. P174.
127. Barnett, P.G., J.L. Sorensen, et al., *Effect of incentives for medication adherence on health care use and costs in methadone patients with HIV.* Drug Alcohol Depend, 2009. **100**(1-2): p. 115-21.
128. Carey, D., *Once-daily therapies.* J HIV Ther, 2003. **8**(1): p. 7-11.
129. Parienti, J.J., D.R. Bangsberg, et al., *Better adherence with once-daily antiretroviral regimens: a meta-analysis.* Clin Infect Dis, 2009. **48**(4): p. 484-8.
130. Bangalore, S., G. Kamalakkannan, et al., *Fixed-dose combinations improve medication compliance: a meta-analysis.* Am J Med, 2007. **120**(8): p. 713-9.
131. Airoldi, M., M. Zaccarelli, et al., *One-pill once-a-day HAART: a simplification strategy that improves adherence and quality of life of HIV-infected subjects.* Patient Prefer Adherence, 2010. **4**: p. 115-25.
132. Reis, A., M. Guerra, and L. Lencastre, *Treatment adherence, quality of life and clinical variables in HIV/AIDS Infection.* Journal of the International AIDS Society, 2010. **13**(suppl 4): p. 74.
133. Akaike, H., *A new look at the statistical model identification.* IEEE Transactions on Automatic Control, 1974. **19**(6): p. 716-723.
134. Reis, A., *Adesão à terapêutica na Infecção pelo Vírus da Imunodeficiência Humana*, in *Faculdade de Psicologia e de Ciências*. 2007, Universidade do Porto: Porto. p. 159.
135. Bangsberg, D.R., A.R. Moss and S.G. Deeks, *Paradoxes of adherence and drug resistance to HIV antiretroviral therapy*. J Antimicrob Chemother, 2004. **53**(5): p. 696-9.

136. Maggiolo, F., D. Ripamonti, et al., *The effect of HIV-1 resistance mutations after first-line virological failure on the possibility to sequence antiretroviral drugs in second-line regimens.* Antivir Ther, 2006. **11**(7): p. 923-9.
137. González, M., J. Cuevas, et al., *A Prospective Study of Adherence and Virologic Failure in HIV-Infected Patients: Role of a Single Determination of Plasma Levels of Antiretroviral Medications.* J Int Assoc Physicians AIDS Care (Chic Ill), 2007. **6**: p. 245.
138. Nachega, J.B., M. Hislop, et al., *Adherence to nonnucleoside reverse transcriptase inhibitor-based HIV therapy and virologic outcomes.* Ann Intern Med, 2007. **146**(8): p. 564-73.
139. Cambiano, V., F.C. Lampe, et al., *Long-term trends in adherence to antiretroviral therapy from start of HAART. .* AIDS, 2010. **15**(24(8)): p. 1153-62.
140. Papke, L.E. and J. Wooldridge., *Econometric methods for fractional response variables with an application to 401(k) plan participation rates.* Journal of Applied Econometrics, 1996. **11**: p. 619–632.
141. McDowell, A. and N.J. Cox. *How do you fit a model when the dependent variable is a proportion?* 2001 August 2004 [cited 2008 June 20th]; Available from: http://www.stata.com/support/faqs/stat/logit.html.
142. EMEA - European Medicines Agency, A. *Report for Prezista - Scientific Discussion.* 2009 [cited 2008 May 25th]; Available from: http://www.emea.europa.eu/docs/en_GB/document_library/EPAR_-_Scientific_Discussion/human/000707/WC500041754.pdf.
143. Clotet, B., N. Bellos, et al., *Efficacy and safety of darunavir-ritonavir at week 48 in treatment-experienced patients with HIV-1 infection in POWER 1 and 2: a pooled subgroup analysis of data from two randomised trials.* Lancet, 2007. **369**(9568): p. 1169-78.
144. Elzi, L., C. Marzolini, et al., *Treatment modification in human immunodeficiency virus-infected individuals starting combination antiretroviral therapy between 2005 and 2008.* Arch Intern Med, 2010. **170**(1): p. 57-65.
145. Sax, P.E., C. Tierney, et al., *Abacavir-lamivudine versus tenofovir-emtricitabine for initial HIV-1 therapy.* N Engl J Med, 2009. **361**(23): p. 2230-40.
146. Geretti, A.M., C. Smith, et al., *Determinants of virological failure after successful viral load suppression in first-line highly active antiretroviral therapy.* Antivir Ther, 2008. **13**(7): p. 927-36.
147. Bangsberg, D.R., D.L. Kroetz and S.G. Deeks, *Adherence-resistance relationships to combination HIV antiretroviral therapy.* Curr HIV/AIDS Rep, 2007. **4**(2): p. 65-72.
148. Bangsberg, D.R., E.P. Acosta, R. Gupta, et al., *Adherence-resistance relationships for protease and non-nucleoside reverse transcriptase inhibitors explained by virological fitness.* AIDS, 2006. **20**(2): p. 223-31.
149. Bangsberg, D.R., T.C. Porco, et al., *Modeling the HIV protease inhibitor adherence-resistance curve by use of empirically derived estimates.* J Infect Dis, 2004. **190**(1): p. 162-5.
150. Braithwaite, R.S., M.J. Kozal, et al., *Adherence, virological and immunological outcomes for HIV-infected veterans starting combination antiretroviral therapies.* AIDS, 2007. **21**(12): p. 1579-89.
151. Gupta, R., A. Hill, et al., *Emergence of drug resistance in HIV type 1-infected patients after receipt of first-line highly active antiretroviral therapy: a systematic review of clinical trials.* Clin Infect Dis, 2008. **47**(5): p. 712-22.
152. Bezerra, V. and F. Roxo. *Esquemas HAART numa consulta de imunodeficiência - estudo multifactorial retrospectivo.* 2º Congresso Virtual VIH/SIDA 2001 [cited 2008 June 1st]; Available from:

http://www.aidscongress.net/Modules/WebC_AidsCongress/CommunicationHTML.aspx?Mid=31&CommID=84.

153. Vieira, F.R., M. Mota, et al. *A mulher e a infecção VIH no Centro Hospitalar de Gaia.* . 4º Congresso Virtual VIH/SIDA 2003 [cited 2009 June 3rd]; Available from: http://www.aidscongress.net/Modules/WebC_AidsCongress/CommunicationHTML.aspx?Mid=33&CommID=237.
154. May, M., K. Porter, et al., *Prognostic model for HIV-1 disease progression in patients starting antiretroviral therapy was validated using independent data*, in *J Clin Epidemiol*. 2005. p. 1033-41.
155. Egger, M., M. May, et al., *Prognosis of HIV-1-infected patients starting highly active antiretroviral therapy: a collaborative analysis of prospective studies.* Lancet, 2002. **360**(9327): p. 119-29.
156. Portugal. Instituto Nacional de Estatística. *Tábua Completa de Mortalidade para Portugal 2006-2008*. 2009 [cited 2010 March 12th]; Available from: http://ue2007.ine.pt/xportal/xmain?xpid=INE&xpgid=ine_destaques&DESTAQUESdest_boui=56512592&DESTAQUESmodo=2.
157. Lohse, N., A.B. Hansen, et al., *Survival of persons with and without HIV infection in Denmark, 1995-2005.* Ann Intern Med, 2007. **146**(2): p. 87-95.
158. Drummond, M., M. Sculpher, et al., *Methods for the Economic Evaluation of Health Care Programmes*. 3rd ed. 2005: Oxford University Press.
159. Brazier, J., J. Ratcliffe, et al., *Measuring and valuing health benefits for economic evaluation.* 2007: Oxford University Press.
160. Dolan, P., *The measurement of Health Related Quality of Life*, in *Handbook of Health Economics*, A.J. Culyer and J.P. Newhouse, Editors. 2001, Elsevier: Amesterdam.
161. Viney, R. and E. Savage, *Health care policy evaluation: empirical analysis of the restrictions implied by Quality Adjusted Life Years.* CHERE Working Paper 2006(2006/10).
162. Bleichrodt, H. and J. Quiggin, *Life-cycle preferences over consumption and health: when is cost-effectiveness analysis equivalent to cost-benefit analysis?* J Health Econ, 1999. **18**(6): p. 681-708.
163. Gold, M., J. Siegel,, et al., *Cost-Effectiveness in Health and Medicine*. 1996: New York: Oxford University Press.
164. Kelly, M., D. McDaid, et al. *Economic appraisal of public health inteventions*. 2005 [cited 2009 March 30th]; Available from: http://www.nice.org.uk/nicemedia/documents/Economic_appraisal_of_public_health_interventions.pdf.
165. Sen, A., *Commodities and Capabilities*. 1985, New York: Elsevier.
166. Sen, A., *Capability and well-being*, in *The Quality of Life*, S.A. Nussbaum M, editors, Editor. 1993, Claredon Press: Oxford. p. 30-54.
167. Lorgelly, P.K., K. Lorimer, et al., *Outcome Measurement in Economic Evaluations of Public Health Interventions: A Role for the Capability Approach?* International Journal of Environmental Research and Public Health 2010(Special Issue "Health Economics").
168. Coast, J., T. Flynn, et al. *An Index of capability for Older People in the UK: An Empirical Investigation*. 2006 [cited 2008 May 9th]; Available from: http://www.capabilityapproach.com/pubs/2_5_Coast.pdf.
169. Flynn, T.N., J. Louviere, et al., *Best--worst scaling: What it can do for health care research and how to do it.* J Health Econ, 2007. **26**(1): p. 171-89.
170. Brazier, J. and J. Roberts, *Methods for developing preference-based measures of health*, in *The Elgar Companion to Health Economics*, Andrew Jones ed., Editor. 2006, Edward Elgar Publishing Limited: Cheltenham.

171. Feeny, D., *The multi-attribute utility approach to assessing health-related quality of life*, in *The Elgar Companion to Health Economics* A. Jones, Editor. 2006, Edward Elgar Publishing Limited.
172. Merino-Castelló, A., *Eliciting Consumers Preferences Using Stated Preference Discrete Choice Models: Contingent Ranking versus Choice Experiment*, in *Department of Economics and Business*. 2003, Universitat Pompeu Fabra: Barcelona.
173. Louviere, J.J., T.N. Flynn and R.T. Carson, *Discrete Choice Experiments Are Not Conjoint Analysis.* Journal of Choice Modelling, 2000. **3**(3): p. 57-72.
174. Patrick, D.L. and R.A. Deyo, *Generic and disease-specific measures in assessing health status and quality of life.* Med Care, 1989. **27**(3 Suppl): p. S217-32.
175. Boyd, N.F., H.J. Sutherland, et al., *Whose utilities for decision analysis?* Med Decis Making, 1990. **10**(1): p. 58-67.
176. Tengs, T.O. and T.H. Lin, *A meta-analysis of utility estimates for HIV/AIDS.* Med Decis Making, 2002. **22**(6): p. 475-81.
177. Joyce, V.R., P.G. Barnett, et al., *Health-related quality of life in a randomized trial of antiretroviral therapy for advanced HIV disease.* J Acquir Immune Defic Syndr, 2009. **50**(1): p. 27-36.
178. Gosling, J., *HIV/Aids: an encounter with death or a journey into life?* J Anal Psychol, 2008. **53**(2): p. 261-9.
179. Tsevat, J., S.N. Sherman, et al., *The will to live among HIV-infected patients.* Ann Intern Med, 1999. **131**(3): p. 194-8.
180. Honiden, S., V. Sundaram, et al., *The effect of diagnosis with HIV infection on health-related quality of Life.* Qual Life Res, 2006. **15**(1): p. 69-82.
181. Szaflarski, M., P.N. Ritchey, et al., *Modeling the effects of spirituality/religion on patients' perceptions of living with HIV/AIDS.* J Gen Intern Med, 2006. **21 Suppl 5**: p. S28-38.
182. Kudel, I., S.L. Farber, et al., *Patterns of responses on health-related quality of life questionnaires among patients with HIV/AIDS.* J Gen Intern Med, 2006. **21 Suppl 5**: p. S48-55.
183. Bleichrodt, H. and M. Johannesson, *Standard gamble, time trade-off and rating scale: experimental results on the ranking properties of QALYs.* J Health Econ, 1997. **16**(2): p. 155-75.
184. Torrance, G.W., D. Feeny and W. Furlong, *Visual analog scales: do they have a role in the measurement of preferences for health states?* Med Decis Making, 2001. **21**(4): p. 329-34.
185. McCabe, C. and K. Stevens, *Visual Analogue Scales: do they have a role in the measurement of preferences for Health States?", .* The University of Sheffield Discussion Paper Series, 2004(04/6).
186. Doctor, J.N., H. Bleichrodt and H.J. Lin, *Health utility bias: a systematic review and meta-analytic evaluation.* Med Decis Making, 2010. **30**(1): p. 58-67.
187. Torrance, G.W., *Measurement of health state utilities for economic appraisal.* J Health Econ, 1986. **5**(1): p. 1-30.
188. Richardson, J., *Cost utility analysis: what should be measured?* Soc Sci Med, 1994. **39**(1): p. 7-21.
189. Bleichrodt, H., *A new explanation for the difference between time trade-off utilities and standard gamble utilities.* Health Econ, 2002. **11**(5): p. 447-56.
190. Bleichrodt, H., C. Herrero and J.L. Pinto, *A proposal to solve the comparability problem in cost-utility analysis.* J Health Econ, 2002. **21**(3): p. 397-403.
191. Bleichrodt, H., J.L. Pinto and J.M. Abellan-Perpinan, *A consistency test of the time trade-off.* J Health Econ, 2003. **22**(6): p. 1037-52.

192. Dolan, P. and P. Stalmeier, *The validity of time trade-off values in calculating QALYs: constant proportional time trade-off versus the proportional heuristic.* J Health Econ, 2003. **22**(3): p. 445-58.
193. Craig, B.M. and J.J. Busschbach, *The episodic random utility model unifies time trade-off and discrete choice approaches in health state valuation.* Popul Health Metr, 2009. **7**: p. 3.
194. Ryan, M. and J. Hughes, *Using conjoint analysis to assess women's preferences for miscarriage management.* Health Econ, 1997. **6**(3): p. 261-73.
195. Ratcliffe, J. and M. Buxton, *Patients' preferences regarding the process and outcomes of life-saving technology. An application of conjoint analysis to liver transplantation.* Int J Technol Assess Health Care, 1999. **15**(2): p. 340-51.
196. Johansson, G., B. Stallberg, et al., *Asthma treatment preference study: a conjoint analysis of preferred drug treatments.* Chest, 2004. **125**(3): p. 916-23.
197. Salomon, J.A., *Reconsidering the use of rankings in the valuation of health states: a model for estimating cardinal values from ordinal data.* Popul Health Metr, 2003. **1**(1): p. 12.
198. McCabe, C., J. Brazier, et al., *Using rank data to estimate health state utility models.* J Health Econ, 2006. **25**(3): p. 418-31.
199. Craig, B.M., J.J. Busschbach and J.A. Salomon, *Modeling ranking, time trade-off, and visual analog scale values for EQ-5D health states: a review and comparison of methods.* Med Care, 2009. **47**(6): p. 634-41.
200. Ratcliffe, J., J. Brazier, et al., *Using DCE and ranking data to estimate cardinal values for health states for deriving a preference-based single index from the sexual quality of life questionnaire.* Health Econ, 2009. **18**(11): p. 1261-76.
201. Thurstone, L.L., *A law of comparative judgement.* Psychological Review, 1927. **34**: p. 278–286.
202. McFadden, D., *Conditional Logit Analysis of Qualitative Choice Behavior*. Frontiers in Econometrics, ed. P. Zarembka. 1973: Academic Press.
203. Lancaster, K.J., *A New Approach to Consumer Theory.* Journal of Political Economy, 1966. **74**(1): p. 132.
204. McFadden, D. and K. Train, *Mixed MNL models of discrete response.* Journal of Applied Econometrics, 2000. **15**: p. 447–470.
205. Ryan, M., D. Scott, et al., *Eliciting public preferences for health care: a systematic review of techniques.* Health Technology Assessment, 2001. **5**(5): p. 1-194.
206. Ryan, M. and S. Wordsworth, *Sensitivity of willingness to pay estimates to the level of attributes in discrete choice experiments.* Scottish Journal Political Economy, 2000b. **47**: p. 504-24.
207. Johnson, F.R., K.E. Mathews and M.F. Bingham, *Evaluating welfare: theoretic consistency in multiple-response, stated-preference surveys.* Triangle Economic Research. , 2000(Working paper T-0003).
208. Miguel, F.S., *Testing the assumptions of completeness, stability and rationality of preferences using discrete choice experiments*. 2000, University of Aberdeen: Aberdeen.
209. Ryan, M. and F.S. Miguel, *Testing for consistency in willingness to pay experiments.* J Econ Psychol, 2000. **21**: p. 305-17.
210. King, M., R. Viney, et al., *Men's preferences for treatment of early stage prostate cancer: Results from a discrete choice experiment.* CHERE Working Paper 2006(2006/14).
211. Street, D. and L. Burgess, *The Construction of Optimal Stated Choice Experiments: Theory and Methods*. 2007, New York: Wiley.
212. Flynn, T.N., J.J. Louviere, et al., *Rescaling quality of life values from discrete choice experiments for use as QALYs: a cautionary tale.* Popul Health Metr, 2008. **6**: p. 6.

213. Brazier, J. and T. Cain. *Utilising rank and DCE data to value health status on the 'QALY' scale using conventional and Bayesian methods*. 2008 [cited 2008 April 9th]; Available from: www.shef.ac.uk/content/1/c6/07/15/12/John%20Brazier.ppt.
214. Ryan, M., A. Netten, et al., *Using discrete choice experiments to estimate a preference-based measure of outcome--an application to social care for older people.* J Health Econ, 2006. **25**(5): p. 927-44.
215. Kauf, T.L., N. Roskell, et al., *A predictive model of health state utilities for HIV patients in the modern era of highly active antiretroviral therapy.* Value Health, 2008. **11**(7): p. 1144-53.
216. Braithwaite, R.S., J. Goulet, et al., *Quantifying the decrement in utility from perceived side effects of combination antiretroviral therapies in patients with HIV.* Value Health, 2008. **11**(5): p. 975-9.
217. Tillmann, H.L., T. Kaiser, et al., *Differential influence of different hepatitis viruses on quality of life in HIV positive patients.* Eur J Med Res, 2006. **11**(9): p. 381-5.
218. Anis, A.H., B. Nosyk, et al., *Quality of life of patients with advanced HIV/AIDS: measuring the impact of both AIDS-defining events and non-AIDS serious adverse events.* J Acquir Immune Defic Syndr, 2009. **51**(5): p. 631-9.
219. Harris, A.H., R.H. Osborne, et al., *Quality of life and Kaposi sarcoma: using preference techniques to value the health gains from treatment.* Support Care Cancer, 2002. **10**(6): p. 486-93.
220. Stavem, K., S.S. Froland and K.B. Hellum, *Comparison of preference-based utilities of the 15D, EQ-5D and SF-6D in patients with HIV/AIDS.* Qual Life Res, 2005. **14**(4): p. 971-80.
221. Mrus, J.M., K.E. Sherman, et al., *Health values of patients coinfected with HIV/hepatitis C: are two viruses worse than one?* Med Care, 2006b. **44**(2): p. 158-66.
222. Lara, A.M., B.N. Wakholi, et al., *Utility assessment of HIV/AIDS-related health states in HIV-infected Ugandans.* AIDS, 2008. **22 Suppl 1**: p. S123-30.
223. Ryan, M. and K. Gerard, *Using discrete choice experiments to value health care programmes: current practice and future research reflections.* Appl Health Econ Health Policy, 2003. **2**(1): p. 55-64.
224. Belkar, R. and D.G. Fiebig, *Use of discrete choice experiments in health economics: An update of the literature.* University of New Working Papers Series, 2004.
225. Guttmann, R., R. Castle and D. G. Fiebig, *Use of Discrete Choice Experiments in health economics: An update of the literature.* CHERE Working Paper, 2009(2009/2).
226. Beusterien, K.M., K. Dziekan, et al., *Understanding patient preferences for HIV medications using adaptive conjoint analysis: feasibility assessment.* Value in Health, 2005. **8**(4): p. 453-61.
227. Phillips, K.A., T. Maddala and F.R. Johnson, *Measuring preferences for health care interventions using conjoint analysis: an application to HIV testing.* Health Serv Res, 2002. **37**(6): p. 1681-705.
228. Lee, S.J., R.A. Brooks, et al., *HIV vaccine acceptability among immigrant Thai residents in Los Angeles: a mixed-method approach.* AIDS Care, 2008. **20**(10): p. 1161-8.
229. Hakim, Z. and D.S. Pathak, *Modelling the EuroQol data: a comparison of discrete choice conjoint and conditional preference modelling.* Health Econ, 1999. **8**(2): p. 103-16.
230. McKenzie, L., J. Cairns and L. Osman, *Symptom-based outcome measures for asthma: the use of discrete choice methods to assess patient preferences.* Health Policy, 2001. **57**(3): p. 193-204.
231. Olsen, C.H., J. Gatell, et al., *Risk of AIDS and death at given HIV-RNA and CD4 cell count, in relation to specific antiretroviral drugs in the regimen.* AIDS, 2005. **19**(3): p. 319-30.
232. Sowadsky, R. *What Is the Approximate Timeline of Testing and Symptoms for HIV/AIDS?* The Complete HIV/AIDS Resource 2010 [cited 2010 March 9th]; Available from: http://www.thebody.com/content/art2309.html#timeline.

233. Yadav, S., *Perceived social support, hope, and quality of life of persons living with HIV/AIDS: a case study from Nepal.* Qual Life Res, 2010. **19**(2): p. 157-66.
234. Ryan, M., K. Gerald and M. Amaya-Amaya, *Using Discrete Choice Experiments to Value Health and Health Care*. The Economics of Non-Market Goods and Resources. Vol. 11. 2008: Mabel.
235. Coast, J. and S. Horrocks, *Developing attributes and levels for discrete choice experiments using qualitative methods.* J Health Serv Res Policy, 2007. **12**(1): p. 25-30.
236. Miguel, F.S., M. Ryan and M. Amaya-Amaya, *'Irrational' stated preferences: a quantitative and qualitative investigation.* Health Econ, 2005. **14**(3): p. 307-22.
237. Train, K.E., *Discrete Choice Methods with Simulation*. 2nd ed. 2009, Cambridge: Cambridge University Press.
238. Catarino, J.M., *Qualidade de Vida Relacionada com VIH*, in *Faculdade de Economia*. 2010, Universidade de Coimbra: Coimbra. p. 120.
239. Canavarro, M.C., M.S. Pereira, et al., *Estudos Psicométricos da Versão Portuguesa (Portugal) do Instrumento de Avaliação de Qualidade de Vida na Infecção VIH da Organização Mundial de Saúde (WHOQOL-HIV).* Psicologia, Saúde & Doenças, 2008. **9**(1): p. 15-28.
240. Stevens, S., *On the theory of scales of measurement.* Science, 1946. **103**: p. 677-680.
241. Vaz Serra, A., M.C. Canavarro, et al., *Estudos psicométricos do instrumento de avaliação da qualidade de vida da Organização Mundial de Saúde (WHOQOL-100) para Português de Portugal.* Psiquiatria Clínica, 2006. **27**(1): p. 31-40.
242. Guerra, M., *Uma Escala de Avaliação do Suporte Social: Sua Aplicação numa população seropositiva ao vírus HIV.* Avaliação Psicológica: Formas e Contextos, 1995. **3**: p. 25-34.
243. Portugal. Administração Central do Sistema de Saúde IP. *Catálogo de Aprovisionamento Público da Saúde*. 2009 [cited 2009 January 11th]; Available from: http://www.catalogo.min-saude.pt/CAPS/publico/default.asp.
244. Portugal. Instituto Nacional da Farmácia e do Medicamento. *Prontuário Terapêutico Online*. 2009 [cited 2009 February 8th]; Available from: http://www.infarmed.pt/infomed/inicio.php.
245. Portaria n.º 132/2009 de 30 de Janeiro de 2009, in *DR I Série B*, n.º 21, de 30 de Janeiro de 2009, 660-758 - Ministério da Saúde.
246. Glick, H. *Methods for Cost Estimation in CEA: the GLM Approach - practical advice for ways implement GLM models*. 2008 [cited 2009 March 30th]; Available from: http://www.uphs.upenn.edu/dgimhsr/documents/acadhlth.glick.061008.pdf.
247. Llibre-Codina, J., M.Á. Casado-Gómez, et al., *Cost of nucleoside analogue reverse transcriptase inhibitorrelated toxicity in HIV-1-infected patients.* Enfermedades Infecciosas Y Microbiologia Clinica, 2007. **25**(2): p. 98-107.
248. Aragão, F., I. Vaz Pinto and J. Vera, *Avaliação Económica do Início Precoce de Tratamento Antiretrovírico em Doentes Infectados pelo VIH-1 em Portugal*, in *Arquivos de Medicina*. 2010. p. 211.
249. Wilson, D.P., M.G. Law, et al., *Relation between HIV viral load and infectiousness: a model-based analysis.* Lancet, 2008. **372**(9635): p. 314-20.
250. Gallant, J.E., E. DeJesus, et al., *Tenofovir DF, emtricitabine, and efavirenz vs. zidovudine, lamivudine, and efavirenz for HIV.* New England Journal of Medicine, 2006. **354**(3): p. 251-60.
251. Greenbaum, A.H., L.E. Wilson, et al., *Effect of age and HAART regimen on clinical response in an urban cohort of HIV-infected individuals.* AIDS, 2008. **22**(17): p. 2331-9.
252. Liddy, J., *Time to Viral Suppression and Subsequent Failure in ART-Naive Subjects Receiving HAART: Results from the HIV-NAT 006 Cohort*, in *College of public Health Sciences*. 2007, Chulalongkorn University. p. 107.

253. Phillips, A.N., B. Ledergerber, et al., *Rate of change in CD4 counts in patients with stable HIV viremia. .* Journal of the International AIDS Society, 2008. **11**(Suppl 1): p. O17.
254. Ledergerber, B., J.D. Lundgren, et al., *Predictors of trend in CD4-positive T-cell count and mortality among HIV-1-infected individuals with virological failure to all three antiretroviral-drug classes.* Lancet, 2004. **364**(9428): p. 51-62.
255. Deeks, S.G., F.M. Hecht, et al., *HIV RNA and CD4 cell count response to protease inhibitor therapy in an urban AIDS clinic: response to both initial and salvage therapy.* AIDS, 1999. **13**(6): p. F35-43.
256. Le Moing, V., G. Chene, et al., *Predictors of virological rebound in HIV-1-infected patients initiating a protease inhibitor-containing regimen.* AIDS, 2002. **16**(1): p. 21-9.
257. Sanders, G.D., A.M. Bayoumi, et al., *Cost-effectiveness of screening for HIV in the era of highly active antiretroviral therapy.* N Engl J Med, 2005. **352**(6): p. 570-85.
258. Daar, E.S., *ACTG 5202: ATV/r or EFV in Combination with ABC/3TC or TDF/FTC in Antiretroviral Naïve Patients*, in *17th Conference on Retroviruses and Opportunistic Infections February 16-19. .* 2010, Oral Presentation 59LB: Moscone Center West San Francisco, CA.
259. Robbins, G.K., B. Daniels, et al., *Predictors of antiretroviral treatment failure in an urban HIV clinic.* J Acquir Immune Defic Syndr, 2007. **44**(1): p. 30-7.
260. Beck, E.J., S. Mandalia, et al., *Treatment outcome and cost-effectiveness of different highly active antiretroviral therapy regimens in the UK (1996-2002).* Int J STD AIDS, 2008. **19**(5): p. 297-304.
261. Krentz, H.B., M.C. Auld and M.J. Gill, *The changing direct costs of medical care for patients with HIV/AIDS, 1995-2001.* Canadian Medical Association Journal, 2003. **169**(2): p. 106-10.
262. Levy, A.R., D. James, et al., *The direct costs of HIV/AIDS care.* Lancet, 2006. **6**(3): p. 171-7.
263. Braithwaite, R.S., A.C. Justice, et al., *Estimating the proportion of patients infected with HIV who will die of comorbid diseases.* Am J Med, 2005. **118**(8): p. 890-8.
264. Brogan, A., et al., *US cost effectiveness of darunavir/ritonavir 600/100 mg bid in treatment-experienced, HIV-infected adults with evidence of protease inhibitor resistance included in the TITAN Trial.* Pharmacoeconomics, 2010. **28**(Suppl 1): p. 129-46.
265. Brogan, A.J., S.E. Talbird and C.Cohen, *Cost-Effectiveness of Nucleoside Reverse Transcriptase Inhibitor Pairs in Efavirenz-Based Regimens for Treatment-Naïve Adults with HIV Infection in the United States.* Value in Health, 2011(ahead of print).
266. Cambiano, V. and A.N. Phillips, *Modelling the impact of treatment with individual antiretrovirals.* Curr Opin HIV AIDS, 2011. **6**(2): p. 124-30.
267. Cohen, M.S., et al., *Prevention of HIV-1 Infection with Early Antiretroviral Therapy.* N Engl J Med.
268. Jit, M. and M. Brisson, *Modelling the epidemiology of infectious diseases for decision analysis: a primer.* Pharmacoeconomics, 2011. **29**(5): p. 371-86.
269. Aragão, F., J. Vera, et al., *Economic evaluation of early antiretroviral treatment for HIV-1 infected individuals in Portugal*, in *International AIDS Society Conference. July 17-20th. Rome*. 2011. p. TUPE416.
270. Portugal. Coordenação Nacional para a Infecção VIH/Sida. *Diagnósticos da Infecção VIH/SIDA: Representações e efeitos na condição laboral*. 2011 [cited 2011 July 30th]; Available from: http://www.sida.pt/.
271. Johnson, L.F. and P.J. White, *A review of mathematical models of HIV/AIDS interventions and their implications for policy.* Sex Transm Infect, 2011.
272. Lucas, R.E., *Econometric policy evaluation: A critique* Carnegie-Rochester Conference Series on Public Policy, 1976. **1**: p. 19-46

273. Yazdanpanah, Y., C.E. Sloan, et al., *Routine HIV screening in France: clinical impact and cost-effectiveness.* PLoS One, 2010. **5**(10): p. e13132.
274. Uhler, L.M., N. Kumarasamy, et al., *Cost-effectiveness of HIV testing referral strategies among tuberculosis patients in India.* PLoS One, 2010. **5**(9).
275. Lamontagne, E., M. Haacker, et al., *Macroeconomic impact of HIV.* Curr Opin HIV AIDS, 2010. **5**(3).
276. Kitahata, M.M., S.J. Gange, et al., *Effect of early versus deferred antiretroviral therapy for HIV on survival.* N Engl J Med, 2009. **360**(18): p. 1815-26.
277. Sax, P.E. and L.R. Baden, *When to start antiretroviral therapy--ready when you are?* N Engl J Med, 2009. **360**(18): p. 1897-9.
278. Sterne, J.A., M. May, et al., *Timing of initiation of antiretroviral therapy in AIDS-free HIV-1-infected patients: a collaborative analysis of 18 HIV cohort studies.* Lancet, 2009. **373**(9672): p. 1352-63.
279. Goldie, S.J., A.D. Paltiel, et al., *Projecting the cost-effectiveness of adherence interventions in persons with human immunodeficiency virus infection.* Am J Med, 2003. **115**(8): p. 632-41.
280. Freedberg, K.A., L.R. Hirschhorn, et al., *Cost-effectiveness of an intervention to improve adherence to antiretroviral therapy in HIV-infected patients.* J Acquir Immune Defic Syndr, 2006. **43 Suppl 1**: p. S113-8.
281. Munakata, J., J.S. Benner, et al. *Clinical and economic outcomes of nonadherence to highly active antiretroviral therapy in patients with human immunodeficiency virus*. Med Care 2006 Oct [cited 44 10]; 2006/09/27:[893-9]. Available from: http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=17001259.
282. Zaric, G.S., A.M. Bayoumi, et al., *The cost-effectiveness of counseling strategies to improve adherence to highly active antiretroviral therapy among men who have sex with men.* Med Decis Making, 2008. **28**(3): p. 359-76.
283. The EuroGuidelines Group for HIV resistance, *Clinical and laboratory guidelines for the use of HIV-1 drug resistance testing as part of treatment management: recommendations for the European setting. .* AIDS, 2001. **15**(3): p. 309-20.
284. Vercauteren, J., K. Deforche, et al., *The incidence of multidrug and full class resistance in HIV-1 infected patients is decreasing over time (2001-2006) in Portugal.* Retrovirology, 2008. **5**: p. 12.
285. Van Laethem, K., A. De Luca, et al., *A genotypic drug resistance interpretation algorithm that significantly predicts therapy response in HIV-1-infected patients.* Antivir Ther, 2002. **7**(2): p. 123-9.
286. Pina, M.F., *Análise Descritiva e Geográfica dos Internamentos Hospitalares por Infecções pelo Vírus da Imunodeficiência Humana (GCD 25) ocorridos no período de 2000 a 2006*. 2007, Faculdade de Medicina da Universidade do Porto.

Appendix I

Proof on the Expressions for Healthy and Sick

Claim: $X_t \equiv \begin{bmatrix} N_t \\ M_t \end{bmatrix} = \begin{bmatrix} F_t l^{t-1} k^{t-1} N_1 \\ F_t l^{t-1}\{M_1 + [1-k^{t-1}]N_1\} \end{bmatrix}$

Remark: $f_1 \equiv F_1 \equiv 1; f_{T+1} \equiv F_{T+1} \equiv 0$

Proof:

$$\begin{bmatrix} N_t \\ M_t \end{bmatrix} = \begin{bmatrix} f_t lkN_{t-1} \\ f_t l(M_{t-1} + sN_{t-1}) \end{bmatrix} <=>$$

$$\begin{bmatrix} N_t \\ M_t \end{bmatrix} = \begin{bmatrix} \dfrac{W - P_t - \widehat{P_{Min}}}{W - P_{t-1} + \widehat{P_{Min}}} lk \left(\dfrac{W - P_{t-1} + \widehat{P_{Min}}}{W - P_{t-2} + \widehat{P_{Min}}} lkN_{t-2} \right) \\ \dfrac{W - P_t - \widehat{P_{Min}}}{W - P_{t-1} + \widehat{P_{Min}}} l \left\{ \dfrac{W - P_{t-1} + \widehat{P_{Min}}}{W - P_{t-2} + \widehat{P_{Min}}} l(M_{t-2} + sN_{t-2}) + s \left(\dfrac{W - P_{t-1} + \widehat{P_{Min}}}{W - P_{t-2} + \widehat{P_{Min}}} lkN_{t-2} \right) \right\} \end{bmatrix} <=>$$

$$\begin{bmatrix} N_t \\ M_t \end{bmatrix} = \begin{bmatrix} \dfrac{W - P_t - \widehat{P_{Min}}}{W - P_{t-2} + \widehat{P_{Min}}} l^2 k^2 N_{t-2} \\ \dfrac{W - P_t - \widehat{P_{Min}}}{W - P_{t-2} + \widehat{P_{Min}}} l^2 (M_{t-2} + sN_{t-2} + skN_{t-2}) \end{bmatrix} <=>$$

$$\begin{bmatrix} N_t \\ M_t \end{bmatrix} = \begin{bmatrix} \dfrac{W - P_t - \widehat{P_{Min}}}{W - P_{t-3} + \widehat{P_{Min}}} l^3 k^3 N_{t-3} \\ \dfrac{W - P_t - \widehat{P_{Min}}}{W - P_{t-3} + \widehat{P_{Min}}} l^3 \{M_{t-3} + s(N_{t-3} + kN_{t-3} + k^2 N_{t-3})\} \end{bmatrix} <=>$$

Recursively substituting up to period 1 results in:

$$\begin{bmatrix} N_t \\ M_t \end{bmatrix} = \begin{bmatrix} \dfrac{W - P_t - \widehat{P_{Min}}}{W} l^{t-1} k^{t-1} N_1 \\ \dfrac{W - P_t - \widehat{P_{Min}}}{W} l^{t-1} \{M_1 + s(1 + k + k^2 + \cdots + k^{t-2}) N_1\} \end{bmatrix} <=>$$

$$\begin{bmatrix} N_t \\ M_t \end{bmatrix} = \begin{bmatrix} F_t l^{t-1} k^{t-1} N_1 \\ F_t l^{t-1}\{M_1 + [1-k^{t-1}]N_1\} \end{bmatrix}$$

Proof on the Cumulative Probability of Not Switching

Claim: $F_t \equiv \prod_{i=1}^{t} f_i = \frac{W - P_t + \widehat{P_{Min}}}{W}$

Switching costs $w \sim U[0,1]$ and consumers will switch if $P_t > \widehat{P_{Min}} + w$

Thus, the probability of not switching in period t, considering that they did not switch in the previous period is:

$$f_t = P(\overline{Sw_i}|\overline{Sw_{i-1}}) = P\left(w > P_t - \widehat{P_{Min}}|w > P_{t-1} - \widehat{P_{Min}}\right) = \frac{W - \left(P_t - \widehat{P_{Min}}\right)}{W - \left(P_{t-1} - \widehat{P_{Min}}\right)}$$

So,

$$F_t = \prod_{i=2}^{t} P(\overline{Sw_i}|\overline{Sw_{i-1}}) = \frac{W - P_t + \widehat{P_{Min}}}{W - P_{t-1} + \widehat{P_{Min}}} \frac{W - P_{t-1} + \widehat{P_{Min}}}{W - P_{t-2} + \widehat{P_{Min}}} \dots \frac{W - P_2 + \widehat{P_{Min}}}{W}$$

$$= \frac{W - P_t + \widehat{P_{Min}}}{W}$$

Proof of the Average Cost Expression

Claim: $AVC_t = \frac{C_N k^{t-1} N_1 + C_M \left[M_1 + \left(1 - k^{t-1}\right) N_1\right]}{N_1 + M_1}, \forall_{t=2,\dots,T}$

Proof:

$$AVC_t = \frac{C_N N_t + C_M M_t}{N_t + M_t} = \frac{C_N (F_t l^{t-1} k^{t-1} N_1) + C_M (F_t l^{t-1} \{M_1 + [1 - k^{t-1}] N_1\})}{N_t + M_t}$$

$$<=> AVC_t = \frac{F_t l^{t-1} [C_N k^{t-1} N_1 + C_M (M_1 + [1 - k^{t-1}] N_1)]}{F_t l^{t-1} [k^{t-1} N_1 + (M_1 + [1 - k^{t-1}] N_1)]}$$

$$<=> AVC_t = \frac{[C_N k^{t-1} N_1 + C_M (M_1 + [1 - k^{t-1}] N_1)]}{[k^{t-1} N_1 + (M_1 + [1 - k^{t-1}] N_1)]}$$

$$<=> AVC_t = \frac{[C_N k^{t-1} N_1 + C_M (M_1 + [1 - k^{t-1}] N_1)]}{N_1 + M_1}$$

Proof of N_1^*

Claim: $N_1^* = \frac{1}{(1-kl)\sum_{t=1}^{T} k^{t-1} l^{t-1} F_t^*}$

Proof:

$$N_1 = NE + IHS + VHS$$

$$<=> \mathrm{N_1} = 1 + \mathrm{klN_t} + \sum_{t=1}^{T-1} k^t l^t \mathrm{F_t}(1 - \mathrm{f_{t+1}})\mathrm{N_1} <=> 1 + \mathrm{k^T l^T F_T N_1} + \sum_{t=1}^{T-1} k^t l^t \mathrm{F_t}(1 - \mathrm{f_{t+1}})\mathrm{N_1} <=>$$

$$<=> \mathrm{N_1} = 1 + \sum_{t=1}^{T} k^t l^t \mathrm{F_t}(1 - \mathrm{f_{t+1}})\mathrm{N_1} <=> 1 + \mathrm{k^T l^T F_T N_1} + \sum_{t=1}^{T-1} k^t l^t \mathrm{F_t}(1 - \mathrm{f_{t+1}})\mathrm{N_1} <=>$$

$$<=> \mathrm{N_1^*} = \frac{1}{1 - \sum_{t=1}^{T} k^t l^t (\mathrm{F_t^* - F_{t+1}^*})} <=> \mathrm{N_1^*} = \frac{1}{1 - \{\mathrm{k^0 l^0 F_1^*} + \sum_{t=1}^{T} (k^t l^t - k^{t-1} l^{t-1})\mathrm{F_t^*}\}}$$

$$<=> \mathrm{N_1^*} = \frac{1}{\sum_{t=1}^{T} (k^t l^t - k^{t-1} l^{t-1})\mathrm{F_t^*}}$$

Note that: $N_1^* = \frac{1}{(1-kl)\sum_{t=1}^{T} k^{t-1} l^{t-1} F_t^*} \Rightarrow N^* = \frac{1}{(1-kl)}$

Proof of M_1^*

Claim: $M_1^* = \frac{N_1^* \sum_{t=1}^{T} [l^t(1-k^t) - l^{t-1}(1-k^{t-1})] F_t^*}{(1-l)\sum_{t=1}^{T} l^{t-1} F_t^*}$

Proof:

$$M_1 = M_T l + slN_T + \sum_{i=1}^{T-1} (M_i l + slN_i)(1 - f_{i+1})$$

$$<=> M_1 = F_T l^T \{M_1 + [1 - k^{T-1}]N_1\} + sF_T l^T k^{T-1} N_1 + \sum_{t=1}^{T-1} (F_i l^t \{M_1 + [1 - k^{T-1}]N_1\} + sF_t l^t k^{t-1} N_1)(1 - f_{t+1})$$

$$<=> M_1 = F_T l^T\{M_1 + [1 - k^{T-1} + sk^{T-1}]N_1\} + \sum_{t=1}^{T-1} l^t(M_1 + [1 - k^{t-1}]N_1 + sF_t l^t k^{t-1} N_1)(F_t - F_{t+1})$$

$$<=> M_1 = F_T l^T\{M_1 + [1 - k^T]N_1\} + \sum_{i=1}^{T-1} l^t(M_1 + [1 - k^t]N_1)(F_t - F_{t+1})$$

$$<=> M_1 = N_1\left\{F_T l^T(1 - k^{T-1}) + \sum_{i=1}^{T-1} l^t(1 - k^t)(F_t - F_{t+1})\right\} + M_1\left\{F_T l^T + \sum_{i=1}^{T-1} l^t(F_t - F_{t+1})\right\}$$

$$M_1^* = \frac{N_1^*\{F_T^* l^T(1 - k^{T-1}) + \sum_{i=1}^{T-1} l^t(1 - k^t)(F_t^* - F_{t+1}^*)\}}{\{1 - F_T^* l^T - \sum_{i=1}^{T-1} l^t(F_t^* - F_{t+1}^*)\}}$$

$$<=> M_1^* = \frac{N_1^*\{F_T^* l^T(1 - k^T) + [-l^{T-1}(1 - k^{T-1})F_T^* + \sum_{i=1}^{T-1}[l^t(1 - k^t) - l^{t-1}(1 - k^{t-1})]F_t^*]\}}{\{1 - F_T^* l^T - [1 - l^{T-1}F_T^* + \sum_{i=1}^{T-1}(l^t - l^{t-1})F_t^*]\}}$$

$$<=> M_1^* = \frac{N_1^*\{F_T^* l^T(1 - k^T) + [-l^{T-1}(1 - k^{T-1})F_T^* + \sum_{i=1}^{T-1}[l^t(1 - k^t) - l^{t-1}(1 - k^{t-1})]F_t^*]\}}{\{1 - F_T^* l^T - [1 - l^{T-1}F_T^* + \sum_{i=1}^{T-1}(l^t - l^{t-1})F_t^*]\}}$$

$$<=> M_1^* = \frac{N_1^*\{F_T^*[l^T(1 - k^T) - l^{T-1}(1 - k^{T-1})] + \sum_{i=1}^{T-1}[l^t(1 - k^t) - l^{t-1}(1 - k^{t-1})]F_t^*\}}{\{-F_T^* l^T + l^{T-1}F_T^* + \sum_{i=1}^{T-1}(l^{t-1} - l^t)F_t^*\}}$$

$$<=> M_1^* = \frac{N_1^*\{F_T^*[l^T(1 - k^T) - l^{T-1}(1 - k^{T-1})] + \sum_{i=1}^{T-1}[l^t(1 - k^t) - l^{t-1}(1 - k^{t-1})]F_t^*\}}{\{F_T^*(l^{T-1} - l^T) + \sum_{i=1}^{T-1}(l^{t-1} - l^t)F_t^*\}}$$

$$<=> M_1^* = \frac{N_1^* \sum_{i=1}^{T}[l^t(1 - k^t) - l^{t-1}(1 - k^{t-1})]F_t^*}{\sum_{i=1}^{T}(l^{t-1} - l^t)F_t^*}$$

Note that: $M_1^* = \frac{N_1^* \sum_{t=1}^{T}[l^t(1-k^t) - l^{t-1}(1-k^{t-1})]F_t^*}{(1-l)\sum_{t=1}^{T} l^{t-1}F_t^*} \Rightarrow M^* = \frac{l}{1-l} - \frac{kl}{(1-kl)}$

Proof of Optimal P_t

Claim: $P_t = \frac{(W+\widehat{P_{Min}})}{2} + \frac{AVC_t}{2}$

Proof: Because of the simplification made to the model, profit in any period $t = 2, \dots, T$ does not depend on any past period prices. Therefore, maximizing Π_t yields the same result as maximizing $\sum_{i=t}^{T} \Pi_i$, so the problem resumes to static maximization. For $t = 2, \dots, T$

$$\Pi_t = (P_t - C_N)N_t + (P_t - C_M)M_t - FC$$

$$<=> \Pi_t = (P_t - C_N)\frac{W - P_t + \widehat{P_{Min}}}{W} l^{t-1}k^{t-1}N_1 + (P_t - C_M)\frac{W - P_t + \widehat{P_{Min}}}{W} l^{t-1}(M_1 + (1 - k^{t-1})N_1) - FC$$

$$<=> \Pi_t = \left[-P_t^2 + P_t\left(W - \widehat{P_{Min}} + C_N\right) - C_N\left(W + \widehat{P_{Min}}\right)\right]\frac{l^{t-1}k^{t-1}N_1}{W} + \left[-P_t^2 + P_t\left(W - \widehat{P_{Min}} + C_M\right) - C_M\left(W + \widehat{P_{Min}}\right)\right]\frac{l^{t-1}}{W}(M_1 + (1 - k^{t-1})N_1) - FC$$

$$<=> \Pi_t = \frac{l^{t-1}}{W}\left\{-P_t^2[N_1 + M_1] + P_t\left[\left(W + \widehat{P_{Min}}\right)(N_1 + M_1) + C_N k^{t-1}N_1 + C_M(M_1 + (1 - k^{t-1})N_1)\right] - \left(W + \widehat{P_{Min}}\right)[C_N k^{t-1}N_1 + C_M(M_1 + (1 - k^{t-1})N_1)]\right\} - FC$$

$$<=> \Pi_t = \frac{l^{t-1}}{W}\left\{-P_t^2[N_1 + M_1] + P_t\left[\left(W + \widehat{P_{Min}}\right)(N_1 + M_1) + \frac{\widehat{TC_t}}{l^{t-1}}\right] - \left(W + \widehat{P_{Min}}\right)\frac{\widehat{TC_t}}{l^{t-1}}\right\} - FC$$

$$<=> \Pi_t = \frac{l^{t-1}}{W}(N_1 + M_1)\left\{-P_t^2 + P_t\left(W + \widehat{P_{Min}}\right) + AC_t - \left(W + \widehat{P_{Min}}\right)AC_t\right\} - FC$$

$$\frac{\vartheta \Pi_t}{\vartheta P_t} = 0 <=> \frac{l^{t-1}}{W}(N_1 + M_1)\left\{-2P_t^2 + \left(W + \widehat{P_{Min}}\right) + AC_t\right\} = 0$$

$$<=> P_t = \frac{\left(W + \widehat{P_{Min}}\right)}{2} + \frac{AVC_t}{2}$$

Proof of Optimal Profit Expression

Claim: $\prod_t = \frac{l^{t-1}}{4W}(N_1 + M_1)\left[\left(W + \widehat{P_1}\right) - AVC_t\right]^2 - FC$

Proof:

$$\Pi_t = \frac{l^{t-1}(M_1+N_1)}{W}\left[P_t\left(\left\{\widehat{P_1}+W\right\}+AC_t\right)-\left\{\widehat{P_1}+W\right\}AC_t - P_t^2\right]-FC$$

$$\Leftrightarrow \Pi_t = \frac{l^{t-1}(M_1+N_1)}{W}\left[\left(\frac{W+\widehat{P_1}}{2}+\frac{AC_t}{2}\right)\left(\left\{\widehat{P_1}+W\right\}+AC_t\right)-\left\{\widehat{P_1}+W\right\}AC_t-\left(\frac{\left\{\widehat{P_1}+W\right\}}{2}+\frac{AC_t}{2}\right)^2\right]-FC$$

$$\Leftrightarrow \Pi_t = \frac{l^{t-1}(M_1+N_1)}{W}\left[\frac{1}{2}\left(\left\{\widehat{P_1}+W\right\}+AC_t\right)^2-\left\{\widehat{P_1}+W\right\}AC_t-\frac{1}{4}\left(\left\{\widehat{P_1}+W\right\}+AC_t\right)^2\right]-FC$$

$$\Leftrightarrow \Pi_t = \frac{l^{t-1}(M_1+N_1)}{W}\left[\frac{1}{4\lambda}\left(\left\{\widehat{P_1}+W\right\}+AC_t\right)^2-\left\{\widehat{P_1}+W\right\}AC_t\right]-FC$$

$$\Leftrightarrow \Pi_t = \frac{l^{t-1}(M_1+N_1)}{4W}\left(\left\{\widehat{P_1}+W\right\}-AC_t\right)^2-FC$$

Proof of the Optimal Entry Price

Claim: $P_1 = AVC_1 + \frac{\sum_{t=1}^{T}\rho^{t-1}FC}{N_1+M_1} - \frac{1}{4w}\sum_{t=2}^{T}\rho^{t-1}l^{t-1}\left[\left(W+\widehat{P_1}\right)-AVC_t\right]^2$

Proof: Bertrand competition prior to entry forces $\sum_{t=1}^{T}\rho^{t-1}\prod_t = 0$ so:

$$\Pi_1 + \sum_{t=2}^{T}\rho^{t-1}\Pi_t = 0$$

$$\Leftrightarrow \left[(P_1-C_N)N_1+(P_1-C_M)M_1-FC+\sum_{t=2}^{T}\rho^{t-1}\left(\frac{N_1+M_1}{4W}l^{t-1}\left[\left(W+\widehat{P_1}\right)-AC_t\right]^2-FC\right)=0\right]=0$$

$$\Leftrightarrow \left[P_1N_1-C_NN_1+M_1P_1-M_1C_M-FC'+\frac{N_1+M_1}{4W}\sum_{t=2}^{T}\rho^{t-1}l^{t-1}\left[\left(W+\widehat{P_1}\right)-AC_t\right]^2-\sum_{t=2}^{T}\rho^{t-1}FC\right]=0$$

$$\Leftrightarrow \left[P_1(N_1+M_1)-(C_NN_1+M_1C_M)+\frac{N_1+M_1}{4W}\sum_{t=2}^{T}\rho^{t-1}l^{t-1}\left[\left(W+\widehat{P_1}\right)-AC_t\right]^2-\sum_{t=1}^{T}\rho^{t-1}FC\right]=0$$

$$\Leftrightarrow P_1 = AC_1 - \frac{1}{4W}\sum_{t=2}^{T}\rho^{t-1}l^{t-1}\left[\left(W+\widehat{P_1}\right)-AC_t\right]^2 + \frac{\sum_{t=1}^{T}\rho^{t-1}FC}{N_1+M_1}$$

Determination of the Optimal Steady-State Live Span

Our model is a free terminal time discrete optimal control problem[101]. Unlike its continuous counterpart, there is no terminal condition uniquely defining T, the optimal time period [102].

In the presence of switching costs, plan owners know they will have market power over consumers once these are "captured". This implies that plans are willing to incur losses to attract potential consumers because they will be able to recover those losses in future years. If no restriction is imposed on first period profits (as it occurs in our price floor scenario), it is fairly intuitive to expect a pattern of profits beginning with negative profits in the first year, in order to establish market share, and positive profits in all subsequent years (recall that, by assumption, any plan unable to reach expected positive profits in the forthcoming year, exits at the begging of that year). Moreover, the longer plans live, the more profitable their market share will become. Thus, we expect P_1^* to be decreasing in T. Section 2.6, confirms this result.

Claim: T is optimal life span in steady-state equilibrium if the $HPPC$ is not satisfied at $T+1$, given the optimal price sequence for a life span of T and if the optimal sequence of profits for any life span $\tilde{T} > T$ yields a negative profit for some $t > 1$.

First note that, if given the optimal price sequence for a T life span, $\exists_{P_{T+1}}: \prod_{T+1} > 0$, there would be an incentive to offer the plan one more year and T would not be optimal to begin with. This is, thus, a necessary, although no sufficient, condition for equilibrium.

Now, suppose T was the optimal life span in the steady-state but the optimal sequence of prices and profits for plans living $\tilde{T}$ yielded positive profits in all periods beyond the first. Each year, the employer selects one plan from among the potentially many entrants. This "*pre-entrance*" competition implies that, if a plan announces a first period premium yielding zero intertemporal profits for a life span of T, some other plan could announce a (lower) first period price, equally generating zero intertemporal profits and satisfying the $HPPC$ but over a longer life span. And this last plan would be selected. Thus T could not be the optimal life span.

[101] We chose a discrete time model to capture the fact that health plan pricing decisions and commercial plan enrollment decisions are each made only once a year.

[102] For a discussion on terminal conditions in discrete optimal problems, we refer the reader to Ried, W. Health Economics, 1996. 5: p. 447-468, Grossman, M., The Human Capital Model, in Handbook of Health Economics, J.A. Culyer and J.P. Newhouse, Editors. 2000, Elsevier: Amsterdam and Eisenring, C., Health Economics, 2000. 9(8): p. 669-680.
.

Appendix II

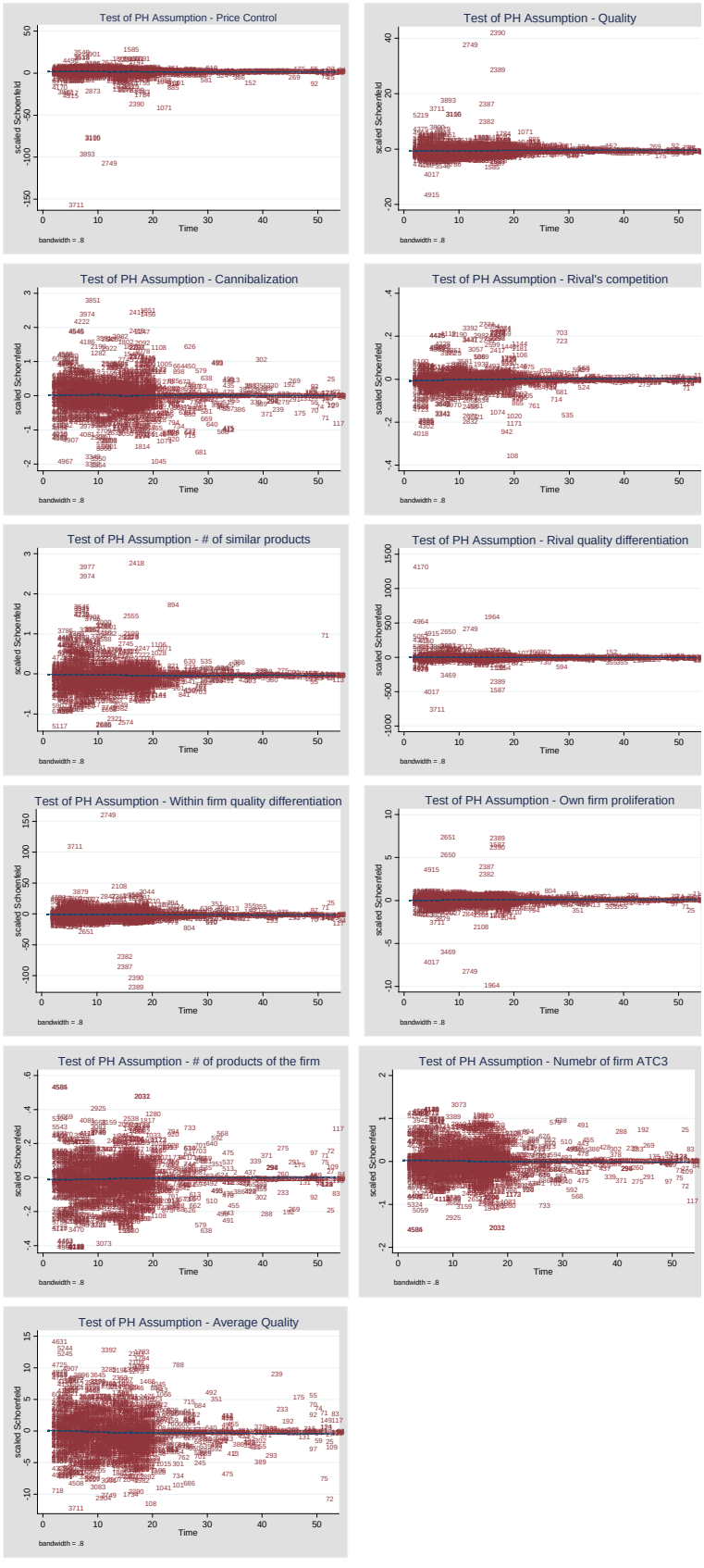


Figure 25: Graphic Analysis of Proportional Hazard Assumption

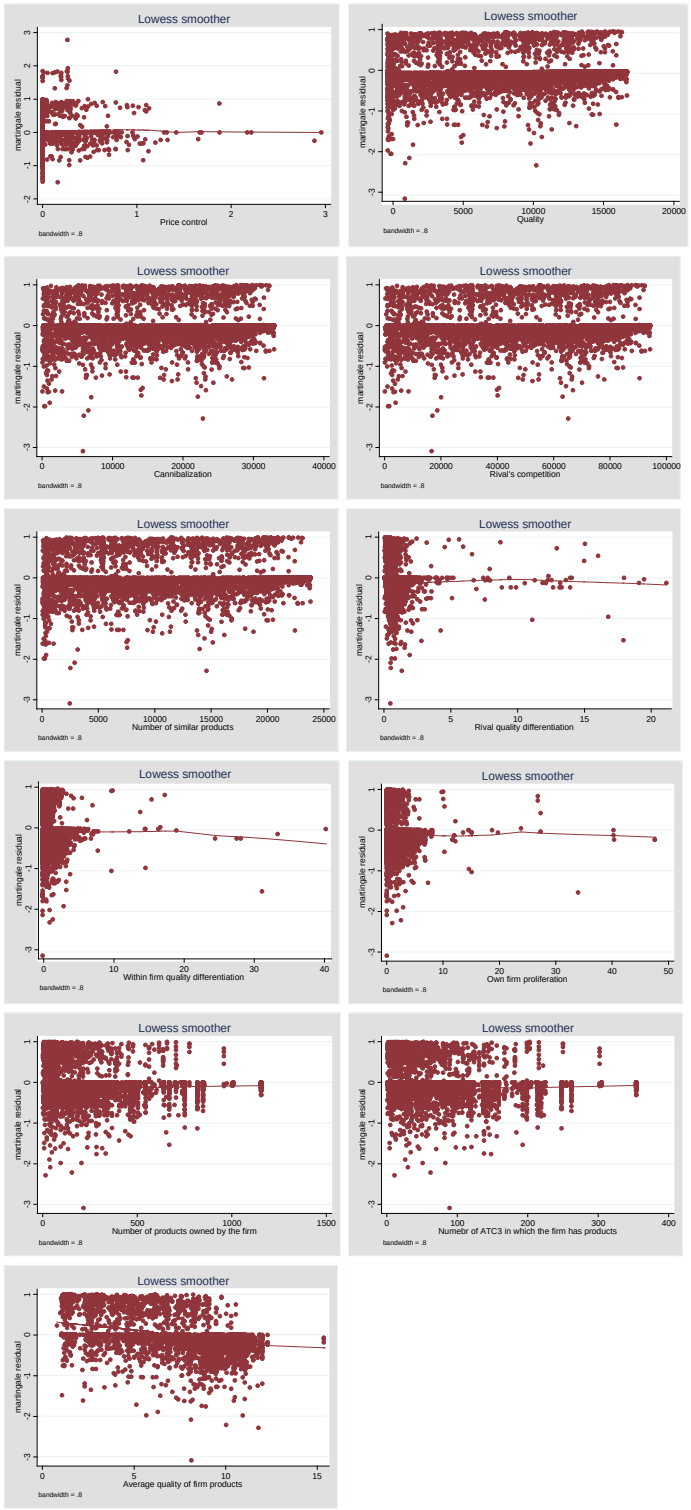


Figure 26: Functional Form Graphic Evaluation using Martingale Residuals

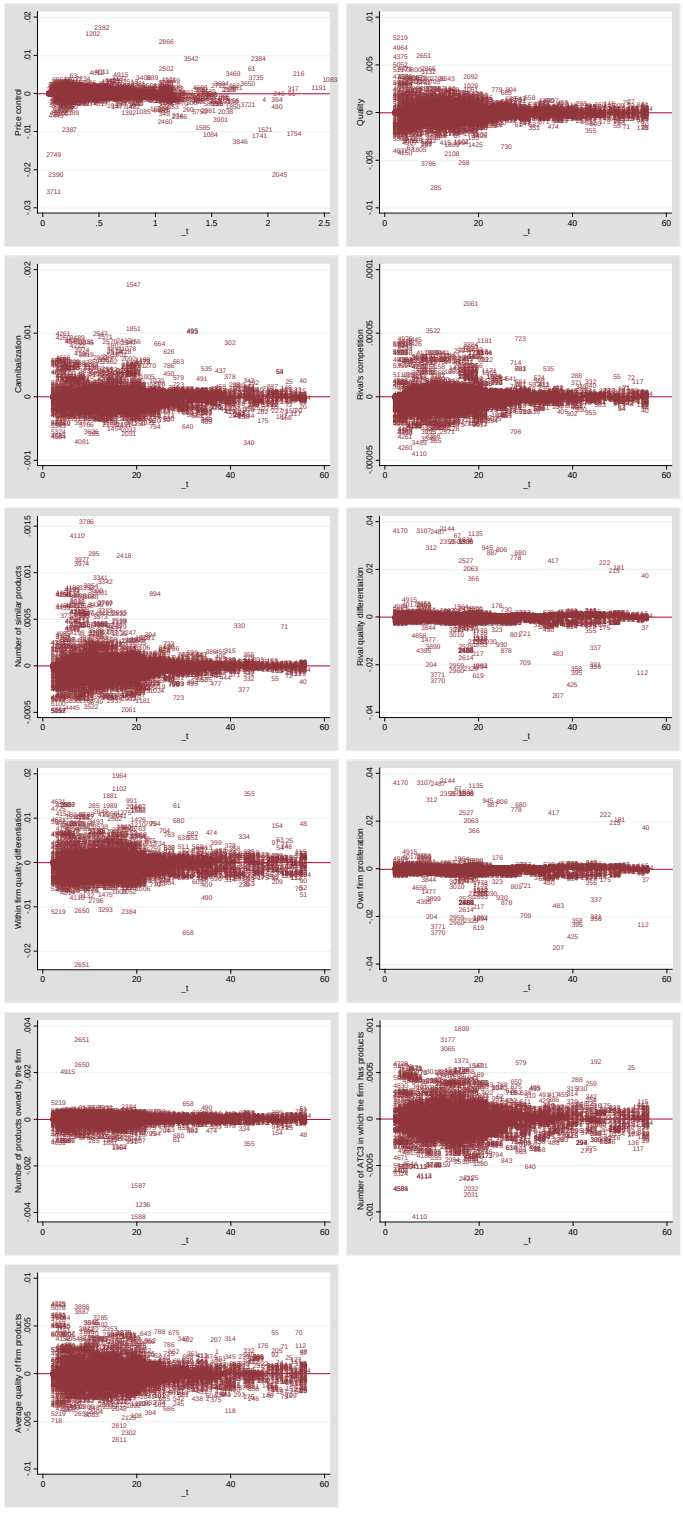


Figure 27: Graphical Analysis of Outliers Using Score Residuals

Table 44: Discrete time Proportional Hazard Model Estimation Results

Discrete Time Proportional Hazard Regression						
Residual df=	26,434			No. of obs=		26,973
Pearson X2=	24,342			Deviance=		8,648.93
Dispersion=	0.92			Dispersion=		0.327189
	Haz. Ratio	**Std. Err.**	**z**	**P>z**	**[95% Conf. Interval]**	
Price Control	1.059	0.012	5.02	0.000	1.036	1.083
Quality	0.801	0.033	-5.450	0.000	0.739	0.867
Competion						
Cannibalization	0.974	0.058	-0.450	0.652	0.867	1.093
Rival's competition	0.992	0.058	-0.140	0.890	0.885	1.112
# of similar products	1.002	0.000	5.030	0.000	1.001	1.003
Rival quality differentiation	0.782	0.114	-1.68	0.092	0.588	1.041
Within firm quality differentiation	0.591	0.096	-3.230	0.001	0.429	0.813
Own firm proliferation	1.002	0.001	3.610	0.000	1.001	1.004
Firm Characteristics						
# of products	0.913	0.007	-11.95	0.000	0.899	0.926
# of ATC3	0.986	0.009	-1.510	0.131	0.968	1.004
Average quality	0.908	0.068	-1.280	0.201	0.784	1.052
Prescription drug (Yes=1)	1.391	0.073	6.280	0.000	1.255	1.542

Log likelihood (-0.5*Deviance) = -4324.1311. Cf. log likelihood for intercept-only model (Model 0) = -5460.6541
Chi-squared statistic for Model (1) vs. Model (0) = 2273.0459. Prob. > chi2(538) = 2.91e-211 Note: Interaction terms omitted

Appendix III

The databases

CHC - Centro Hospitalar de Cascais

The CHC database is used by physicians at the CHC to assist medical care during physician appointments. It is, therefore, extremely complete. It provides information on socio-demographic characteristics, clinical relevant variables over time, medication prescribed and delivered (both ART and non-ART), other diagnosis, and clinical events.

Inpatient information is not included in the database. It was obtained by requesting the hospital Board for all inpatient episodes of the individuals included in the database. The information provided by the Board includes the attributed DRG code and date of discharge, but information on admission date is missing so episode duration could not be calculated.

This database contains information on 3,085 individuals followed at the CHC immunodeficiency treatment unit. The vast majority of the individuals began the follow-up between 2001-2008, but 804 started before 2001. The average follow-up time of the total sample is 3.4 years. Of the 3,085 individuals, 1,744 were not HIV infected, had not started antiretroviral therapy, or such information was missing. For those whose ART information was available, 7 were under 13 years old and 28 were HIV-2. The remaining 1,306 individuals were included in the analysis, from the first prescription date to the censoring date (December 31st, 2008) or death. Table 45 presents the main descriptive statistics on the most relevant variables while Table 46 provided the un-weighted proportions of each antiretroviral drug prescribed.

Table 45: Descriptive Statistics of CHC Sample

Descriptive Statistics of CHC Sample	N	Average/Freq.	Stand.Dev./ %	Median	Minimum	Maximum
HIV-1 infected individuals on ART	1,306					
Year of entry in cohort	1,306	2002	2.7	2002	1994	2008
Year First (in cohort) ART	1,306	2003	2.9	2003	1994	2008
Follow-up Period (Years)	1301*	4.1	3.0	3.3	0	13.5
Annual Mortality Rates	96					
2002	5	1%				
2003	7	1%				
2004	15	2%				
2005	21	3%				
2006	21	2%				
2007	13	1%				
2008	14	1%				
Socio-Demographic Characteristics						
Age at cohort entry	1,303	36.3	10.3	35	14	75
Age at ART initiation	1,303	37.8	10.0	36	14	76
Gender (F)	1,306	35%				
Race	1,167					
Caucasian		79%				
Black		19%				
Clinical Information						
% of HCV Co-Infection	1,306	43%				
% of HBV Co-Infection	1,306	5%				
Initial Distribution	1,053	355	295	286	2	1983
Initial Log_{10} Viral Load	961	4.0	1.5	4.4	1.7	6.9
Resistance Score	124	3.8	4.2	0.8	0.8	18.0

Notes: Only HIV-1 is included; ART=Antiretroviral Therapy, CD4+=CD4+ T-Lymphocyte count per µL, HBV=Hepatitis B Virus, HCV=Hepatitis C Virus, log_{10} Viral Load=log_{10} HIV RNA copies per mL, N=Number of Observations/Individuals; Resistance score= inverted Genotypic Sensitivity Score, where 1 out of 1 means full resistance for the drug and is based on the REGA algorithm.
*5 are single observations

Table 46: Antiretroviral Drugs Used in the CHC Sample.

Nucleoside Reverse Transcriptase Inhibitors (NRTIs)	
Lamivudina	61.32%
Zidovudina	38.52%
Tenofovir	37.62%
Emtricitabina	22.39%
Abacavir	15.66%
Estavudina	14.73%
Didanosina	13.39%
Zalcitabina	0.53%
Non-Nucleoside Reverse Transcriptase Inhibitors (NNRTIs)	
Efavirenze	23.70%
Nevirapina	17.84%
Protease Inhibitors	
Ritonavir	43.44%
Lopinavir	26.41%
Indinavir	9.59%
Atazanavir	6.63%
Nelfinavir	4.36%
Saquinavir	3.15%
Fosamprenavir	2.80%
Darunavir	0.65%
Tipranavir	0.22%
Amprenavir	0.06%
Integrase Inhibitors	
Raltegravir	0.06%
Entry Inhibitors	
Aplaviroc	0.12%
Enfuvirtida	0.62%

Since individuals may have been transferred from other hospitals where they had already started ART, they were not necessarily naive at the first prescription date in the CHC database. Dr. Inês Vaz Pinto identified naïve, on a case by case basis. 366 naive individuals were identified as having initiated therapy at the CHC between 2001 and 2008 and this sub-sample was the

main source of information used to parameterize the model with respect to initial characteristics. Table 47 shows the main descriptive statistics on the most relevant variables.

Table 47: Descriptive Statistics of Naive Sample in the CHC Database

	N	Average/Freq.	Stand.Dev./ %	Median	Minimum	Maximum
Socio-Demographic Characteristics						
Initial Age	366	40.2	11.2	38.1	19	75
Gender (F=1)	366	36%				
Race	359					
Caucasian		71%				
Black		28%				
Other/unknown		1%				
Transmission	365					
Homosexual		8%				
Heterosexual		65%				
IDU		24%				
Other/unknown		3%				
Late Diagnosis**	365	40%				
Clinical Information						
Year of First ART	366	2004	2.09	2005	2000	2008
% of HCV Co-Infection	366	30%				
Initial CD4+ Distribution	366	269	198	239.5	2	1334
Classe CD4+ Inicial						
0– 50	366	38	10%			
51– 100		33	9%			
101– 200		72	20%			
201– 350		128	35%			
351– 500		53	14%			
>500		42	11%			
Log_{10} Viral Load	366	4.7	0.895	4.8	1.7	6.9

Notes: ART=Antiretroviral Therapy, CD4+=CD4+ T-Lymphocyte count per µL, HCV=Hepatitis-C Virus, Log_{10} Viral Load=Log_{10} HIV RNA copies per mL, N=Number of Observations/Individuals, IDU=Injection Drug User.**defined as either having a CD4+ cell count at the time of diagnosis of less than 200 cells/mm^3 or having AIDS diagnosed concomitantly with HIV infection or in the 12 months thereafter.

CD4 and Viral load Over Time – graphical analysis.

Weighted least squares regression was used to estimate CD4+ and log_{10}VL trajectories, conditional on therapy line. Figure 28 and Figure 29 provide a graphical comparison of the observed and estimated average values over time.

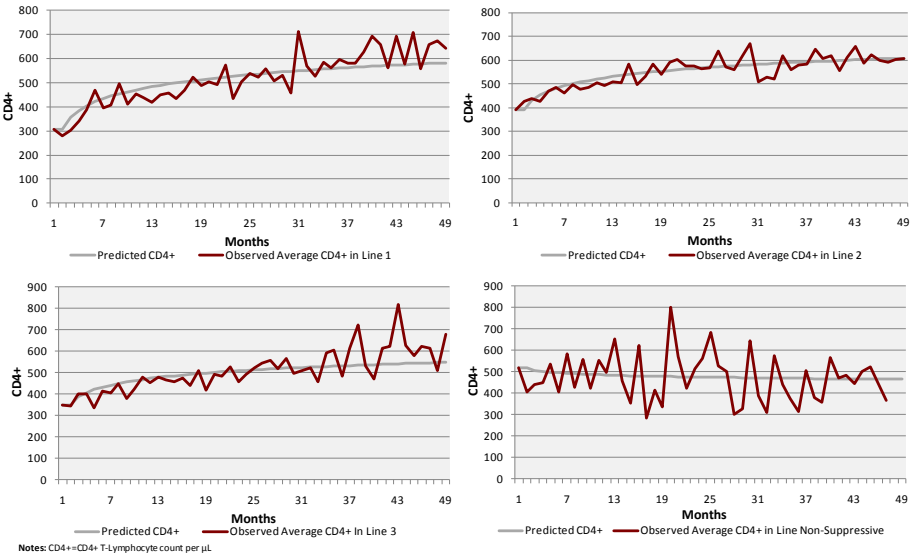


Figure 28: Variation of CD4+ over Time, per Line.

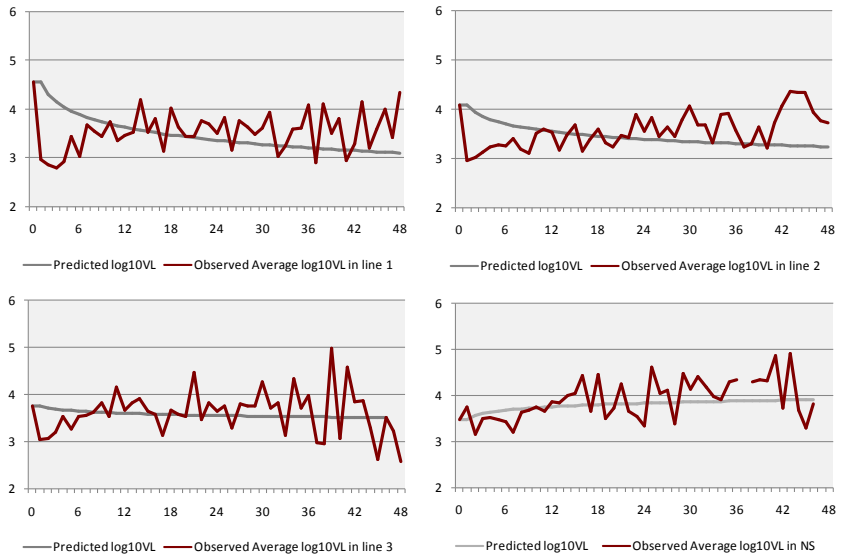


Figure 29: Average Monthly Variation in Log_{10}VL, by Line.

Regimen switches in the CHC sample versus the Swiss cohort.

A recent analysis of "time to the first treatment modification" in the first year of ART, as the primary end point, was performed on 1,318 naive individuals from the Swiss HIV Cohort Study, in the period 2005-2008 [144]. The included regimens were, for the most part, those recommended by Portuguese current guidelines. In that analysis individuals starting ART, 391

(29.7%) modified their treatment[103] during the first year, half of them within the first 3 months. The most frequent reasons for treatment modification were toxic effects (46.6%), followed by a physician's decision (22.8%), the patient's choice (16.9%), and treatment failure (5.9%). During the first year of ART, 208 individuals (15.8%) modified their treatment because of drug intolerance and/or drug toxicity, mostly (64.2%) during the first 3 months.

Comparing the results in Elzi *et al.* with those of the CHC sample, out of the 1,306 (not necessarily naive individual in the database), 339 (26%) switched regimen during the first year of follow-up and 117 switched in the first three months. This switching rate is, thus, close to the one found in the Swiss HIV Cohort. During the first year, virological failure accounted for 20% (68 out of 339) regimen switches, which is significantly higher than the 6% found in the Swiss cohort, but it should be noted that the CHC sample does not contain only naive patients.

If we consider only those 366 naive individuals, 98 (26%) switched regimen in the first year and of these 18 switched due to virological failure. Virological failure thus accounted for 18% of regimen switches in the first year which is a percentage significantly above that reported by Elzi *et al.* The reason for this fact may be related to the fact that, as noted by Elzi *et al*., combined zidovudine and lamivudine was associated with higher rates of treatment modification[104] and in the Swiss cohort 23.4% were on such NRTI pair, while the corresponding percentage is 54.1 in the CHC sample.

LVHEM - Laboratório de Virologia do Hospital Egas Moniz

Since January 2001, European guidelines recommend resistance testing in case of treatment failure [283] and such recommendation has been routinely implemented in Portugal. The vast majority of samples were tested at Hospital Egas Moniz, which is the laboratory of reference [284].

The LVHEM database provides information on 5,456 individuals older than 12 years old, from 22 hospitals located in the Portuguese mainland and the Madeira Archipelago, who were tested

[103] Treatment modification was defined as a switch or discontinuation of ART within the first year. A switch to another regimen was defined as changing 1 or more drugs within 4 weeks after stopping ART. Discontinuation was defined as stopping any antiretroviral drug for at least 4 weeks.

[104] Results are not provided for modification due to virological failure.

for resistance from March 2001 to March 2009[105]. This database includes around 80% of all resistance tests performed by HIV infected individuals in continental Portugal.

Each time a resistance test is performed, the physician in charge of the database – Dr. Ricardo Camacho – requests information on the complete ART. Such information includes drugs used, starting date, ending date and reason for termination. This data is, thus, retrospectively collected. Complete information regarding ART history is available for 2,136 individuals, incomplete information is available for 1,216 individuals, and no information of ART regimens is available for 2,104 individuals.

Since 2001 it is recommended to perform resistance tests at ART initiation. In the LVHEM sample, of the 2,104 individuals without any information on ART regimens, 1,960 performed a resistance test. These are most likely naive patients tested prior to ART initiation and such was assumed in the analysis.

At the time of the test, which for most cases coincides with therapy failure, information on viral load is registered. In some cases, CD4+ cell count is also registered. Table 48 provides a summary of the main characteristics of the LVHEM sample.

Resistance levels recorded in the database use the new REGA 8.0 algorithm [285] and the following drugs are tested: enfuvirtide, raltegravir, atazanavir, nelfinavir, etravirine efavirenz, nevirapine, tenofovir DF, emtricitabine, abacavir, stavudine, lamivudine, didanosine, zidovudine, darunavir/r, tipranavir/r, atazanavir/r, lopinavir/r, fosamprenavir/r, indinavir/r, saquinavir/r. Results of the REGA 8.0 algorithm should be interpreted as: 0= high resistance level, 0.5/0.75= some resistance level (0.75 is for booster PIs) and 1/1.5= susceptible to the drug (1.5 is for boosted PIs). This classification was inversed to generate an indicator that increases with resistance. The current resistance level, as calculated from REGA 8.0, was subtracted from the maximum level attainable by that drug. After such transformation 0= susceptible and 1/1.5=high resistance. This yields a resistance score between 0 and 24.5.

[105] Resistance level information is not available for 201 individuals, therefore, other criteria may apply for inclusion in the database.

Table 48: Descriptive Statistics of the LVHEM Sample

	N	Average/Freq.	Stand.Dev./ %	Median	Minimum	Maximum
Complete ARV Information	2,136					
First Resistance Level	2,057	5.2	4.3	4.8	0.750	21.5
Incomplete ARV Information	1,216					
First Resistance Level	1,156	6.3	5.5	4.8	0.750	22.8
No ARV Information	2,104					
First Resistance Level	1,960	1.1	1.7	0.8	0.750	22.8
Whole LVHEM Sample						
Socio-Demographic Characteristics						
Gender (F)	5,103	0.343				
Geographic Origin	2,729					
Africa		540				
Asia		15				
Estern Europe		2				
Europe		2,119				
South America		53				
Mode of Transmission	721					
IDU		298				
Heterossexual		260				
Homossexual/Bissexual		92				
Other		71				
Clinical Information						
CD4+	3,277	310.4	250.4	265.0	1.0	2,953.0
% with CD4+<200	3,277	0.382				
Log_{10} Viral Load	5,374	4.3	0.818	4.2	1.3	7.0
Number of Regimens	3,352	2.2	1.3	2.0	1.0	13.0

Notes: ART=Antiretroviral Therapy, ARV=Antiretroviral, CD4+=CD4+ T-Lymphocyte count per µL, IDU=Injection Drug User, Log_{10} Viral Load=log_{10} HIV RNA copies per mL, LVHEM=Laboratório de Virologia do Hospital de Egas Moniz, N= Number of Observations/Individuals

In order to estimate time to event models (conditional on therapy line and regimen number, among others) it is necessary to have complete ART history; consequently the sub-sample of 2,123[106] individuals was considered in the parameterization of the model with respect to time to event. No information is provided on the alive/dead status, so all individuals are assumed censored at the date of the last record.

Figure 30 provides the Kaplan-Meier estimate of time to development of first resistance[107] to at least one drug, and the difference between the LVHEM database and the CHC one is clear.

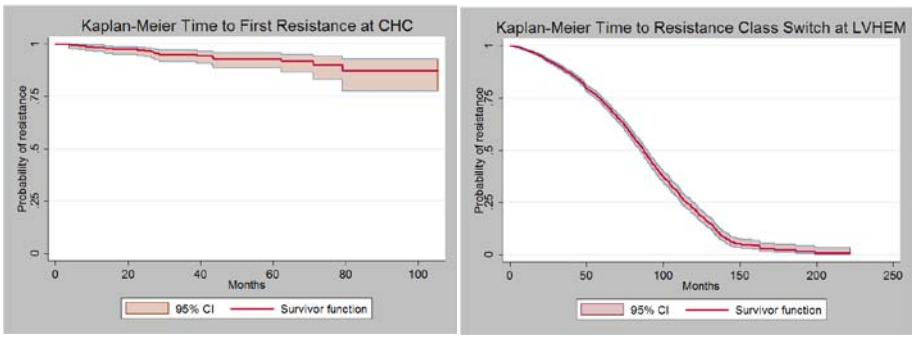


Figure 30: K-M Estimate of Time to First Resistance Development in the CHC and LVHEM Samples

[106] In the 13 cases where the first and last observations coincide, the individuals were discarded from the analysis.

[107] It should be noted that information is not available for those individuals who have not been tested. Given that resistance development reflects itself though detectable levels of viral load without response to treatment, it was assumed that those who were not tested had not yet developed resistance.

CVEDT - Centro de Vigilância Epidemiológica de Doenças Transmissíveis

The CVEDT database is the official national list of HIV infected individuals. Detailed official statistics referring to that database are published by the Communicable Diseases and Epidemiological Surveillance Center on a regular basis, consequently, only a short summary is described in Table 49.

Information obtained from this database consists of a "picture" taken in August, 28th, 2009. It characterizes infection at that moment in time and, thus, unlike the above mentioned databases, it is not in a panel form. Some past information is included, such as probable year of infection, the date of diagnosis, date of notification, date of death (if it has occurred), and date and type of first symptoms. It provides information on the current stage of infection (asymptomatic, symptomatic or AIDS), but not on the dates of transition between stages. Socio-demographic variables provided include: place of birth, nationality, age and gender. HIV related variables provided include: type of infection (HIV-1 or HIV-2) and mode of transmission.

Given that, in the CVEDT database, only 0.8% of the notified HIV cases in Portugal are due to mother/child transmission and that only 124 out of 26,066 cases alive in 2009 with known age, were children under 13 years-old, the present model focuses on the adolescent/adult population.

Table 49: Descriptive Statistics of the CVEDT Database

	Total		Alive in August 2009	
	N	Average/Freq.	N	Average/Freq.
Number of Cases	34718*		26,200	
Age at Diagnosis	34,431	35.4	25,933	34.99
Final Age*	34,431	42.3	25,933	43.57
Gender (F=1)	34,703	25.4%	26,186	28.4%
HIV-1	34,431	96.4%	25,965	96.6%
Mode of Transmission	34,070		25,829	
Heterosexual		41%		45%
Homosexual/Bisexual		13%		12%
IDU**		44%		41%
Other		2%		2%
Year of Infection (most likely)	4,502		4,212	
1968-1989		3%		3%
1990-1999		34%		32%
2000-2008		63%		65%
Year of Diagnosis	34,718		26,200	
1983-1989		3%		2%
1990-1999		46%		40%
2000-2008		52%		58%
Year of Death	8,513			
1983-1989		3%		
1990-1999		52%		
2000-2008		45%		
Number of Years Between Diagnosis and Symptoms (Year of Symptoms-Year of Diagnosis)	9,383	-0.781	5,521	-0.725
HIV Stage***	34,718		26,200	
Assymptomatic		47%		60%
Symptomatic		10%		11%
AIDS		43%		29%

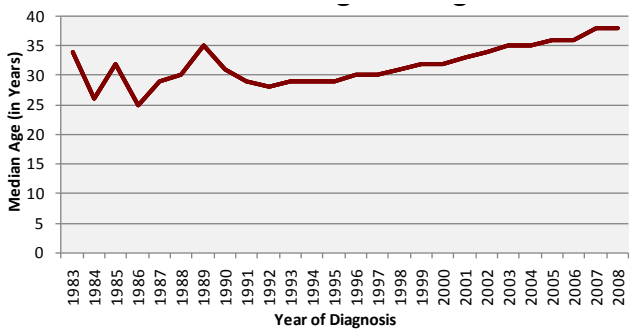
Notes: AIDS=Acquired Immune Deficiency Syndrome, HIV=Human Immunodeficiency Virus, HIV-1=Human Immunodeficiency Virus - Type 1, IDU=Injection Drug User, N=Number of Observations/Individuals

*Only 13 years old or older at death or in 2009 if alive are included;

** Includes the "Homosexual/IDU" Cases;

*** At Death or in 2009 if Alive

The median age at diagnosis in this sample is 33 years-old (average 35.4, 95% CI [35.3; 35.5]. This median is, nonetheless, increasing over time, as shown in Figure 31.



Notes: CVEDT=Centro de Vigilância Epidemiológica das Doenças Transmissíveis

Figure 31: Median Age at Diagnosis in the CVEDT Sample

This pattern is related to the change in the distribution of transmission mode, presented in Figure 32.

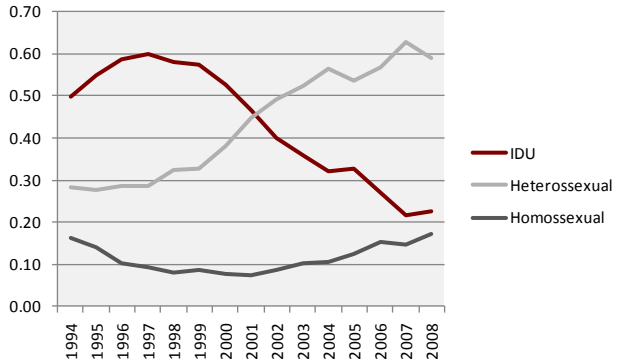


Figure 32: Proportion of Diagnosis by Transmission Group

This change in the in the epidemiology of the disease have an impact on mortality since a statistically significant difference is found in the survival curves by transmission group as shown in Figure 33

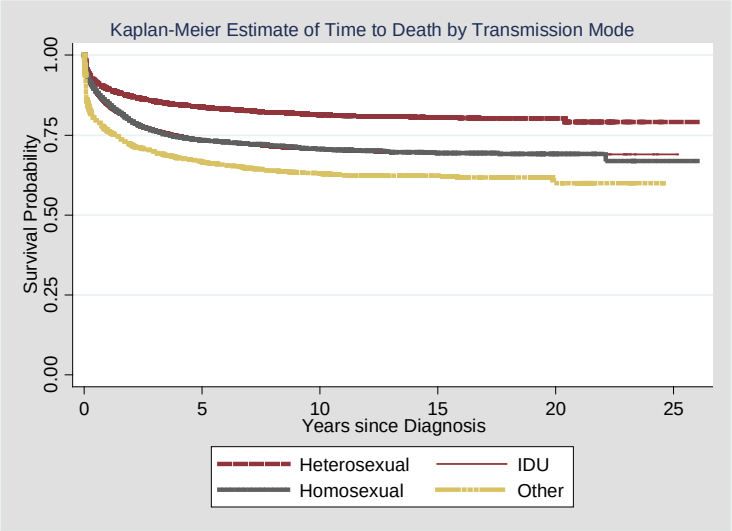


Figure 33: Kaplan Meier Estimates of Time to Death by Mode of Transmission

Progresses in ART, efforts to take infected individuals into therapy and changes in the distribution of individuals by transmission group are among the reasons for the reduction the mortality observed in Figure 34.

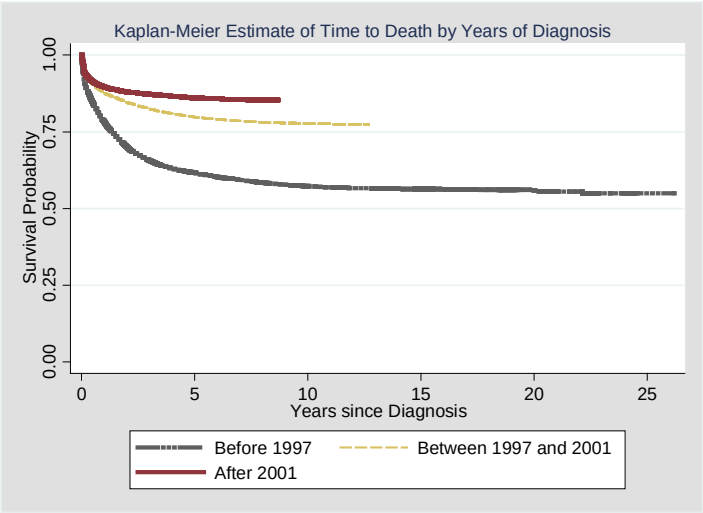
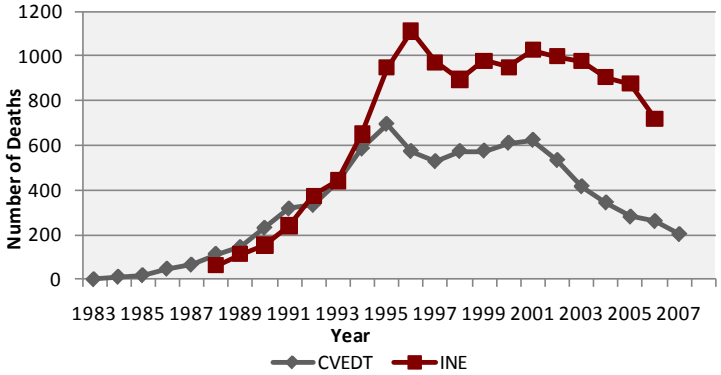


Figure 34: Kaplan Meier Estimate of Time to Death By Year of Diagnosis

Mortality data was available from two distinct sources: the CVEDT dataset and HIV-related aggregate mortality data provided by the National Institute of Statistics (INE). While the first database accounts for all deaths among HIV-infected individuals (without specifying the cause of death), the second refers to deaths attributed to HIV. In spite of this fact, as shown in Figure 35, the number of deaths per year is significantly higher in the INE database. This difference may either be due to non-notification of death among those in the CVEDT database, or it may be that the vital status is correctly registered in the CVEDT sample, but not all HIV infected individuals are in the database. In fact, in the CVEDT database, 59% of those who died were diagnosed no sooner than one year before death, which suggests that there may be a severe problem of late diagnosis of infections. Moreover, the median time between diagnosis and notification (inclusion in the database) is 2.7 months but 21% of notifications occurred more than one year after diagnosis.



Notes: CVEDT=Centro de Vigilância Epidemiológica das Doenças Transmissíveis, INE=Instituto Nacional de Estatística

Figure 35: Deaths over Time, by Source of Information

The distribution of the age of death among HIV infected individuals has itself evolved significantly overtime, as shown by the solid lines in Figure 36. Even if we ignore the 2007 curve, which is clearly out of pattern and most likely reflects an incomplete update, we observe that the most frequent age of death interval has changed from 35-39 in 2002 to 40-44 in 2006, reflecting a change in the epidemiology of the disease and the progress in ART therapy. Nevertheless, comparing the distribution by age of death among HIV infected individuals and that of the general population (the dotted line), the two remain significantly different.

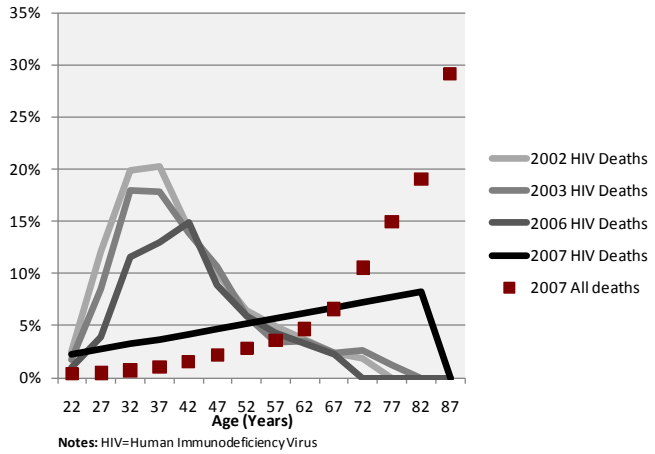


Figure 36: Distribution of Age of Death due to HIV and all Causes. Source: INE Data

In spite of this fact, the Kaplan-Meier estimate on the CVEDT sample (Figure 37), indicates a survival probability of 75% 25 years after diagnosis (that is, roughly at a median age of 58 since the median age at diagnosis is 33 years). In the general population, the proportion of survivors at the age of 58 is 91% (Figure 38). There is, thus, an increased mortality risk due to infection, but possibly not as high as it would be expected from distributions in Figure 36.

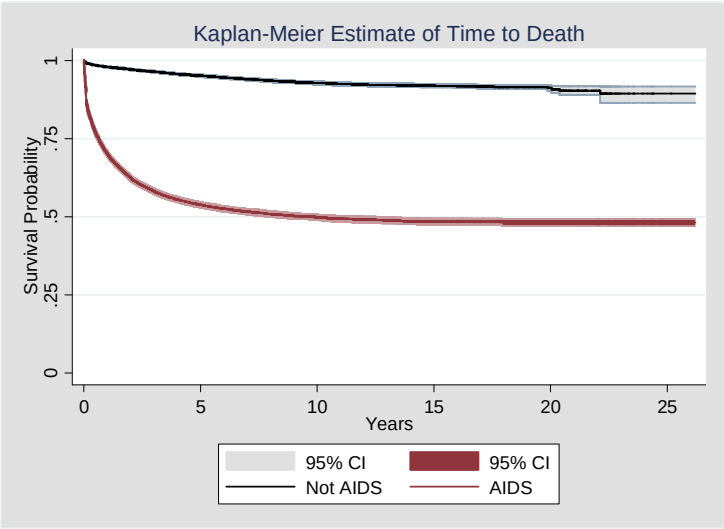


Figure 37: Kaplan-Meier Survival Curve among HIV Infected Individuals

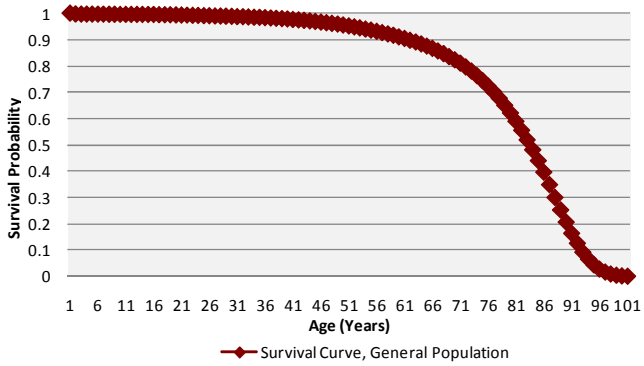


Figure 38: Survival in the General Population. Source: INE 2009.

This increased risk of death due to age is not, at its full length, yet captured in the CVEDT sample. In fact, about 90% of the diagnosis in the CVEDT sample occurred after 1992 and 51% occurred after 1999, when the median age of diagnosis was 32. The median age of HIV infected individuals alive in 2009 is 42 years old. Consequently, even if the alive/dead variable in the CVEDT database is correctly updated, the individuals in the database have not reached an age where the risk of death is significantly increased due to age.

ACSS - Administração Central do Sistema de Saúde

The ACSS database is a national database, at the patient level, of inpatient care related to HIV-infection in all hospitals within the National Healthcare Service in 2009. All records with HIV related codes, either as the primary diagnosis or subsequent ones, are included.

This database contains information on date of birth, data of admission, date of discharge, destination at discharge, DRG code, all diagnosis codes (ICD9-CM classification) and other causes for inpatient care.

In 2009, the average age for HIV-related inpatient care was 43 years-old (median 41, range 14-87). The majority of cases were associated with AIDS diagnosis (68.9%) and 14% resulted in the patients' death (Table 50).

Table 50: Descriptive Statistics of the ACSS Database

		N	Average (SD)/Freq.	SD/ %	Median	Minimum	Maximum
Age		2,404	43.2 (11.6)	11.6	41	14	87
Duration (Days)		2,404	17.9 (23.3)	23.3	11	0	339
Main Diagnosis		2,381					
ICD9	**Description**						
42	AIDS		68.9%				
486	Pneumonia, Organism Unspecified		1.8%				
4829	Bacterial Pneumonia Unspecified		1.2%				
1363	*Pneumocystis carinii*		1.0%				
1300	Meningoencephalitis due to Toxoplasmosis		0.8%				
1175	Cryptococcosis		0.7%				
4660	Acute Bronchitis		0.6%				
64761	Other Viral Diseases		0.6%				
481	Pneumococcal Pneumonia		0.5%				
5712	Alcoholic Cirrhosis of Liver		0.5%				
Discharge Destination		2,404					
	Home		76%				
	Other Healthcare Facility with Inpatient Care		5%				
	Home Based Healthcare		1%				
	Against Physicians Recommendation		5%				
	Death		14%				

Notes: AIDS=Acquired Immune Deficiency Syndrome, ICD9=International Classification of Diseases-9th Revision, SD=Standard Deviation

As shown in Table 51, no statistically significant difference was found in the cost per episode between AIDS and non-AIDS diagnosis ($p = 0.442$), but a statistically significant difference was found in the average duration ($p < 0.001$).

Table 51: Cost of Inpatient Care in the ACSS Database

Main Diagnosis	**ICD9=042 (AIDS)**		**ICD9 *not* 042**		
Most Frequent DRG					
	DRG Codes	Frequency (>5%)	DRG Codes	Frequency (>5%)	
	714	30.1%	714	17.0%	
	710	18.8%	710	6.9%	
	716	8.8%	712	5.8%	
	709	7.7%	557	5.1%	
	715	6.0%			
	712	6.0%			
Cost of Episode					***p-value***
Average	4,764.8 €		4,741.6 €		0.442
Standard Deviation	3,474.5 €		3,688.7 €		
Duration					
Average	19.3		14.9		<0.001
Standard Deviation	24.9		18.9		

Notes: AIDS=Acquired Immune Deficiency Syndrome, DRG=Diagnosis-Related Group, ICD9=International Classification of Diseases-9th Revision

It is of interest to compare the results obtained from the 2009 ACSS database with those reported in the literature, using the same (although more complete) database in different years. Dias *et al.* [80] analyze the predictors of mortality in HIV-associated hospitalizations in Portugal, on a sample of 12,078 adult discharges, from patients with HIV infection diagnosis assisted at Portuguese hospitals from 2005–2007 and registered on the diagnosis-related groups' database. The authors find gender, age, urgent/programmed inpatient care, surgical/medical diagnosis and co-morbidities to be associated with hospital mortality of HIV-infected individuals. The most common HIV-related complications were: Pneumonia (27%), Tuberculosis (25%), PCP (10%), Hepatitis B (5%) and Hepatitis C (32%)[108]. The authors report a median length of hospital stay of 19.2 days, while the median duration observed in the 2009 sample is 11 days.

Pina [286] performs a statistical analysis of HIV related hospitalizations in the period 2000-2006 by year and hospital. The average length of stay reported by Pina for the CHC hospital is 18 days (the national average is not directly provided) which is close to the average in the 2009 sample. A total of 794 hospital admissions were registered. At the time the analysis was performed, only three HIV-related DRG codes existed. Consequently, results on the average cost are not directly comparable. Moreover, all data is disaggregated by hospital or region. Nevertheless, the average cost reported by Pina for the CHC hospital was 5,326€

[108] PCP is an AIDS-defining illness, while pneumonia and tuberculosis may or may not be an AIDS-defining event, depending on the severity (recurrent pneumonia and disseminated tuberculosis).

INE - Instituto Nacional de Estatística

The National Institute of Statistics (INE) kindly provided data, on the number of deaths due to HIV and in the general population, over the period 1988 – 2006, and by gender and age group.

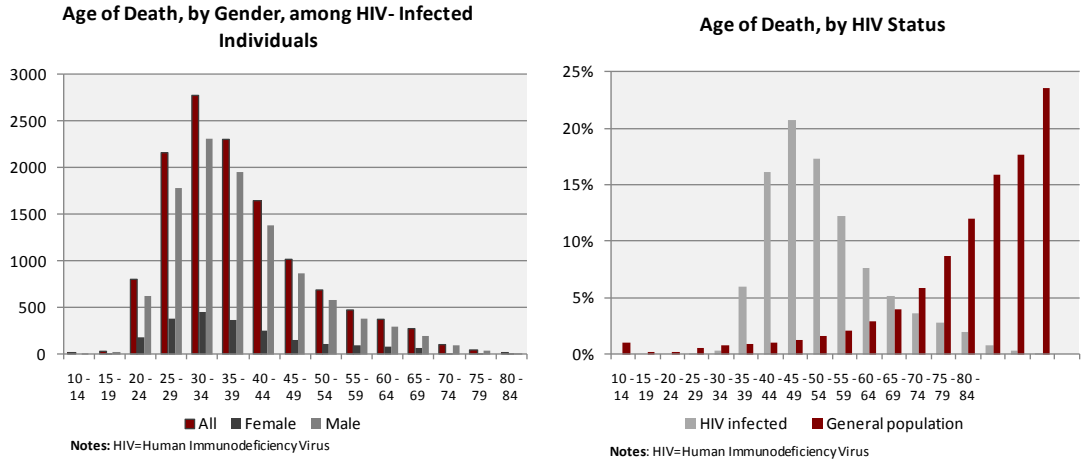


Figure 39: Age of Death According to HIV Status and Gender

Figure 39 summarize the information provided by INE. Although international studies indicate that the gap between life expectancy among HIV infected and the general population is narrowing down [101], the average age of death among HIV individuals is still far from the average observed in the general Portuguese population, as shown in Figure 40.

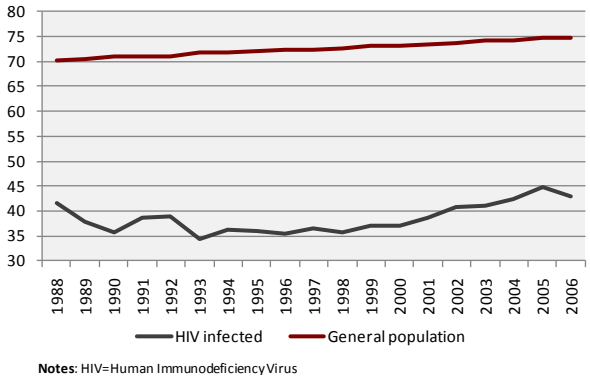


Figure 40: Average Age of Death among HIV Infected Individuals versus General Population

The Sampling Process for Age of Death

At model entrance, individual characteristics are sampled from the distributions described in Section 4.5.1. Following that process, the age of death, conditional on having lived up to the age of ART initiation is sampled from the gender specific general population mortality tables provided by INE [156]. The sampling process, conditional on age, is as follows: assume that the

general population probability, for a female, of surviving to at least the age of 33 is 0.98961 and to the age of 58 is 0.58778. Then, the conditional probability that a woman drawn from the general population who has reached the age of 33 will survive for at least another 25 years is $\frac{0.58778}{0.98961} = 0.59395$. If that woman is HIV-infected and the increased risk of death due to the condition is 1.5, then from standard risk analysis [3] the conditional probability is taken to the power of the hazard ratio yielding $\left(\frac{0.58778}{0.98961}\right)^{1.5} = 0.4577$. As such, if α is the probability of death at current age, the age of death may be obtained by first calculating $\beta = \alpha\, p^{\frac{1}{HR}}$, where p is a random draw from a Uniform distribution and HR is the hazard ratio, and then using the mortality tables to convert the probability β, back to age of death.

A hazard ratio of 1 was assumed since the random draw is meant to reflect death due to age. Although set at 1 in the base case, this parameter is useful for model calibration.

Appendix IV

Baseline Individual Characteristics Distributions

The histogram of **age at ART initiation** is shown in Figure 41.

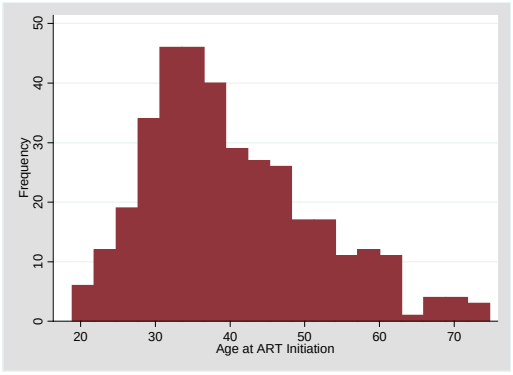


Figure 41: Age at ART Initiation

The initial distribution of CD4+ cell count is taken from the CHC database, as it is the only one which included this particular data. For each individual, the last observed CD4+ value recorded prior to the date of the first ART prescription is assumed as the initial CD4+ count. A Table distribution was fit to the observed frequencies (Figure 42). In the CHC naive sample, 74% are below 350, 15% are between 351 and 500 and 11% are above 500.

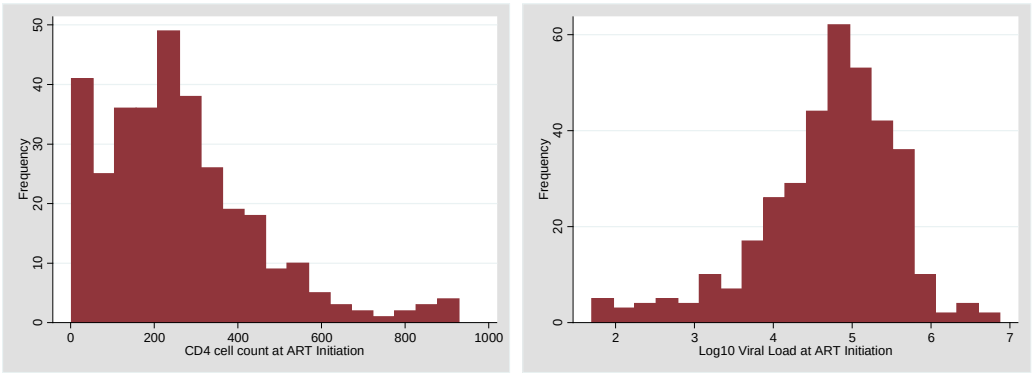


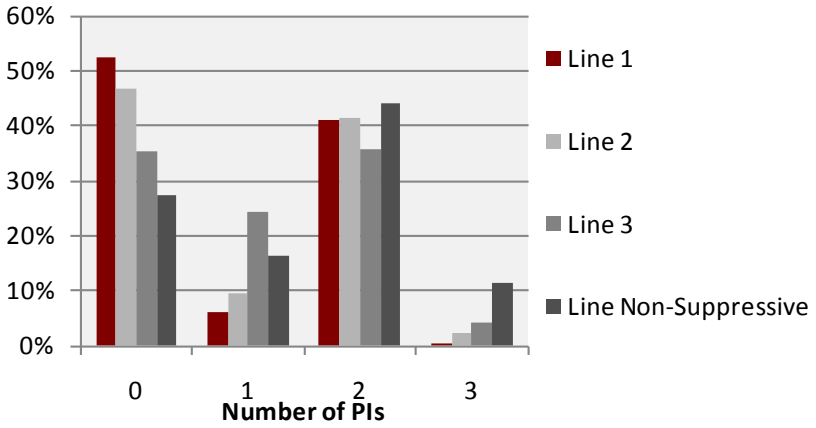
Figure 42: Distribution of CD4+ and Log10VL at ART Initiation

Data on adherence was available in the CHC sample for the period 2002-2008. The database provides information on the number of days medication delivered lasts, and on the day the

individual refilled the prescription. Considering that the individual may get a refill while still having some left, the number of days that the individual may not have taken medication (unless it was obtained from somewhere else) is added over the course of the regimen. The proportion of non-missing days out of the total number of days between the first and last refill the adherence indicator used in the model. We have ignored the last refill since we do not have information on last month's adherence. Two adherence indicators were constructed for each individual: an overall adherence level and a regimen specific adherence level.

Employment status was found to be a significant variable in determining quality of life (QoL) among HIV infected individuals (See Section 4.5.4.5.1) and was consequently included in the model in spite of no direct effect being expected on disease progression. The amount of information available on this regard is scarce, as it is only available for 42% of the CHC database and most likely not updated. In the CHC sample 31% of the individuals are inactive (not employed or retired) while Reis *et al.* [132, 134] report 48% of inactive individuals in their sample. The employment status is sampled from a Bernoulli distribution with $\mathrm{p} = \frac{703}{1065}$ (66%), reflecting the joint results of CHC and Reis *et al.*

In the model, the number of PIs in each regimen was sampled from the Table distribution of the corresponding therapy line. Table distributions for lines 1 and 2 were obtained from the observed frequencies in the CHC sample, while those of lines 3 and non-suppressive came from the LVHEM data. Figure 43 describes the observed frequencies of the number of PIs.



Notes: PI=Protease Inhibitor

Figure 43: Number of PIs, per Therapy Line

159 out of 1,306 individuals in the CHC sample where tested for resistance and 64 of these tests occurred in naive patients. In the LVHEM sample, 1,960 naive[109] individuals were tested for

[109] Assuming naive patients are those for whom no information on ART regimens is available (i.e. tested before ART initiation)

resistance. Table 52 shows the frequency in each resistance level among naive patients in both cases. Given the relative sample sizes, the LVHEM information was used to parameterize the model with respect to initial resistance level. A Table distribution was assumed using the values summarized in the last column of Table 52.

Table 52: First Line Resistance Level in the CHC and LVHEM Samples

Class of Resistance Level	CHC Naive	LVHEM Naive
<1	81.3%	91.2%
1<=R<5	10.9%	6.7%
5<=R<10	7.8%	1.6%
R>10	0.0%	0.5%
N	64	1,960

Notes: CHC=Centro Hospitalar de Cascais, LVHEM=Laboratório de Virologia do Hospital de Egas Moniz – Centro Hospitalar de Lisboa Ocidental, E.P.E., N=Number of Observations/Individuals

The CVEDT database does not provide information on the date of ART initiation, so it is not possible to know how many were classified as having AIDS at that time. The database does not provide information on AIDS status at diagnosis, it only informs on such status as of August 2008. However, considering the sub-sample of individuals diagnosed in 2008 and assuming the AIDS/not-AIDS status did not change until August 2009, we obtain a proxy of the proportion of individuals who are diagnosed with AIDS at ART initiation. This value is obviously a proxy, since it is not possible to know the percentage of diagnosed cases that started ART on the year of diagnosis. This value will be close to 100 in the AIDS and Clinically Symptomatic cases, but will be lower for asymptomatic cases. Moreover, there is a significant lag in notification, so these values are likely to change in the near future. Considering all limitations, it was assumed a Bernoulli distribution with $p = \frac{387}{1{,}201}$ (32%). This value is in accordance with the 31.6% of late presenters (that is, individuals who reach medical care at AIDS stage) in the CHC sample between 2001 and 2007, reported by Vaz Pinto [125] although significantly lower than the 42% reported by Alfaiate *et al.* [126].

Fitted distributions

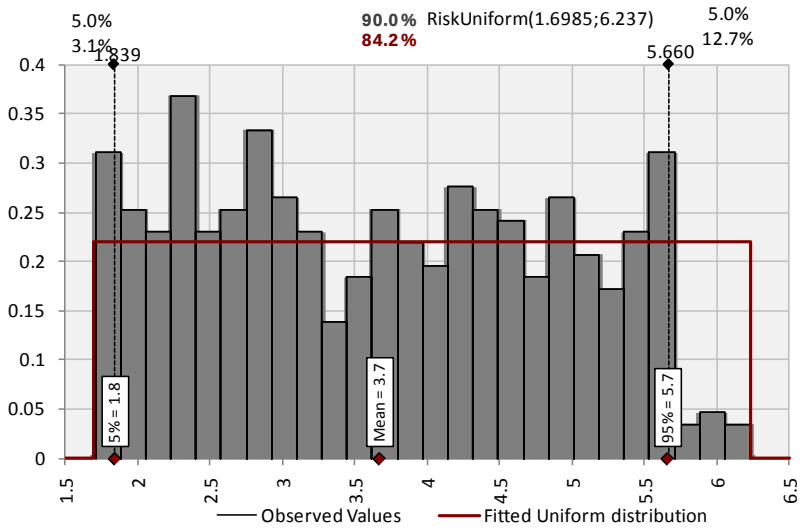


Figure 44: Fit Comparison for Log_{10}HIV RNA at rebound

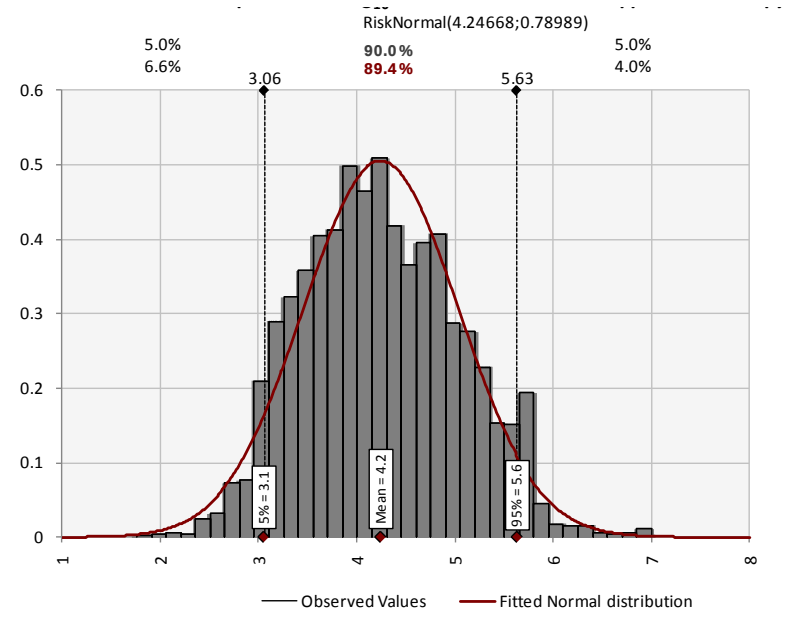


Figure 45: Fit Comparison for log_{10}HIV RNA at rebound in non-suppressive therapy

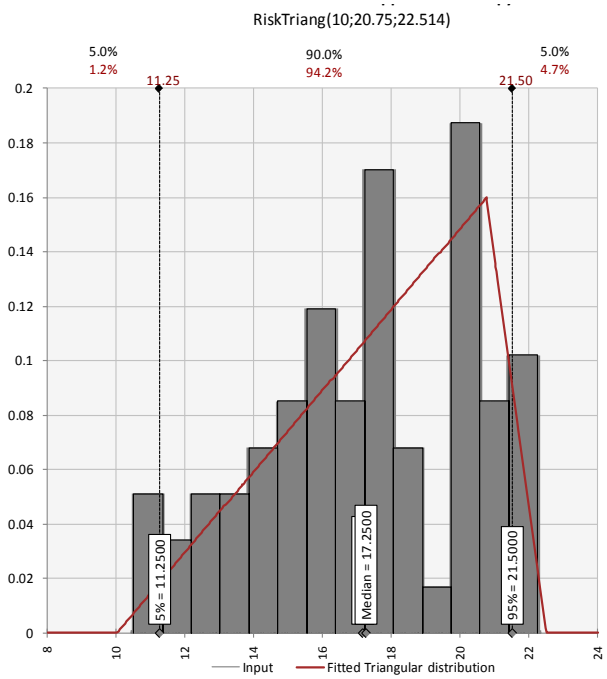


Figure 46: Fit Comparison of Resistance level in Non-suppressive Therapy

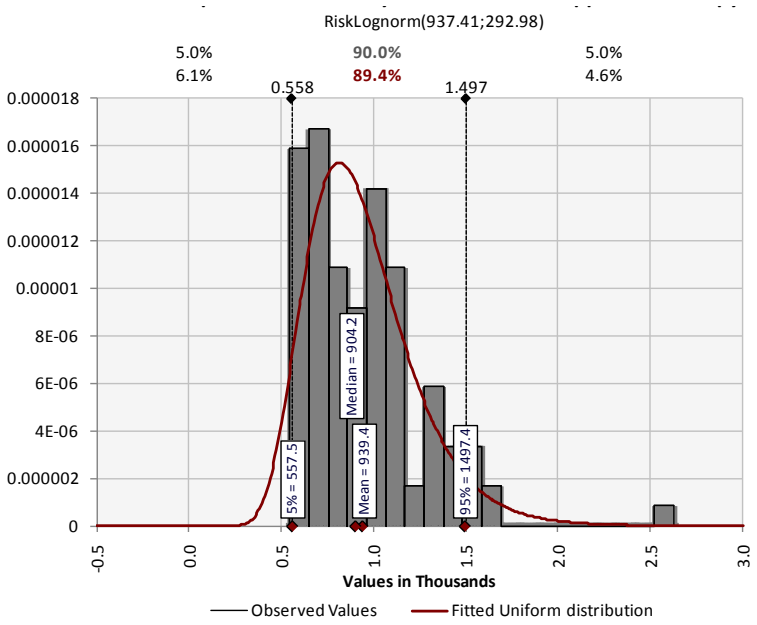


Figure 47: Fit comparison for monthly antiretroviral costs in non-suppressive therapy

Appendix V

Regression Models for Time to Event

Table 53: Regression Model for Time to Viral Suppression

Weibull Regression -- Log Relative-Hazard Form

No. of subjects =	1,276	Number of obs=	1,562
No. of failures =	831	Wald χ2(9)=	75.1
Time at risk =	19,730	Prob > χ2=	0.000
Log pseudolikelihood =	-1,965		
AIC=	3,957	(Std. Err. adjusted for 1,276 clusters in id)	
BIC=	4,026		

Time to Viral Suppression Line 1 & Line 2	Coef.	Robust SE	z	P>z	[95% CI]		dy/dx
Gender (Female=1)	0.090	0.086	1.1	0.294	-0.078	0.257	-0.554
Age	0.018	0.004	4.2	0.000	0.010	0.026	-0.111
$Log_{10}VL$	-0.152	0.032	-4.7	0.000	-0.215	-0.089	0.951
CD4+	0.000	0.000	0.5	0.655	0.000	0.000	0.000
HCV	0.273	0.089	3.1	0.002	0.098	0.447	-1.680
Adherence	0.010	0.003	3.7	0.000	0.005	0.015	-0.062
Year 1st ART	-0.004	0.017	-0.2	0.836	-0.037	0.030	0.022
Number of PIs	-0.048	0.047	-1.0	0.313	-0.140	0.045	0.297
Resistance Level	0.136	0.059	2.3	0.021	0.020	0.251	-0.847
Line (1 is reference)							
2	-0.077	0.093	-0.8	0.407	-0.260	0.105	0.485
Constant	3.8	34.4	0.1	0.912	-63.6	71.2	
/ln_p	0.030	0.026	1.1	0.255	-0.021	0.080	7 [6;8]
p	1.030	0.027			0.979	1.084	Pred. median=6

Notes: ART=Antiretroviral Therapy,HVC=Hepatitis C-Virus, $Log_{10}VL=log_{10}$ HIV RNA copies per mL, PI=Protease Inhibitor

Table 54: Regression Model for Time to Virological Failure in Line 1 and 2

Weibull Regression -- Log Relative-Hazard Form

No. of subjects =	1,302	Number of obs=	3,205
No. of failures =	575	Wald χ2(11)=	129.95
Time at risk =	63,356	Prob > χ2=	0
Log likelihood=	-1,170	(Std. Err. adjusted for 1,302 clusters in id)	
AIC=	2,366		
BIC=	2,445		

Time to Virological Failure	Coef.	Robust SE	z	P>z	[95% CI]		dy/dx
Gender (Female=1)	-0.137	0.056	-2.5	0.014	-0.246	-0.028	7.65
Age	-0.010	0.006	-1.6	0.121	-0.023	0.003	0.56
$Log_{10}VL$	0.362	0.100	3.6	0.000	0.166	0.559	-19.56
CD4+	0.000	0.000	-2.6	0.010	-0.001	0.000	0.03
HCV	-0.182	0.111	-1.6	0.102	-0.400	0.036	10.09
Adherence	-0.003	0.002	-1.7	0.093	-0.007	0.001	0.18
Year 1st ART	-0.044	0.024	-1.9	0.063	-0.090	0.002	2.45
Number of PIs	0.110	0.049	2.2	0.027	0.013	0.206	-6.12
Resistance Level	0.034	0.028	1.2	0.228	-0.021	0.089	-1.89
Regimen Number	0.160	0.036	4.4	0.000	0.089	0.232	-8.96
Line (1 is reference)							
Line 2	0.191	0.102	1.9	0.062	-0.010	0.392	-10.7
Constant	81.5	47.1	1.7	0.084	-10.9	173.8	
/ln_p	0.370	0.032	11.7	0.000	0.308	0.432	72 [67;80]
p	1.447	0.046	1.4	1.540			Pred. median=80

Notes: ART=Antiretroviral Therapy, HVC=Hepatitis C Virus, $Log_{10}VL=log_{10}$ HIV RNA copies per mL, PI=Protease Inhibitor

Table 55: Regression Model for Time to Virological Failure in Line 3

Weibull Regression -- Log Relative-Hazard Form

No. of subjects =	1,864	Number of obs=	2,184
No. of failures =	1,932	Wald χ2(12)=	1,017
Time at risk =	62,290	Prob > χ2=	0.000
Log likelihood=	-961	(Std. Err. adjusted for 1,864 clusters in id)	
AIC=	1,944		

Time to Virological Failure	Coef.	Robust SE	z	P>z	[95% CI]		dy/dx
Gender (Female=1)	-0.033	0.045	-0.720	0.470	-0.122	0.056	0.643
Age	0.004	0.002	2.1	0.039	0.000	0.008	-0.080
$Log_{10}VL$	0.079	0.027	2.9	0.004	0.026	0.133	-1.550
Year 1st ART	0.266	0.009	29.7	0.000	0.249	0.284	-5.188
Regimen Number	-0.012	0.030	-0.410	0.680	-0.070	0.046	0.239
Number of PIs	0.032	0.026	1.220	0.221	-0.019	0.083	-0.618
Resistance Class (1 is reference)							
1<=R<5	0.330	0.064	5.2	0.000	0.205	0.455	-6.225
5<=R<10	0.393	0.065	6.0	0.000	0.266	0.521	-7.348
R>=10	0.385	0.065	5.9	0.000	0.257	0.513	-7.021
Constant	-540.2	18.0	-30.0	0.000	-575.6	-504.9	
/ln_p	0.627	0.023	27.2	0.000	0.582	0.673	31 [29;34]
p	1.873	0.043			1.8	2.0	Pred. median= 36.5

Notes: ART= Antiretroviral Therapy, $Log_{10}VL=log_{10}$ HIV RNA copies per mL, PI=Protease Inhibitor, R=Genotypic Sensitivity Score (inverted).

Table 56: Regression Model for Time to Regimen Switch without Virological Failure

Weibull Regression -- Log Relative-Hazard Form

No. of subjects =	1,302	Number of obs=	3,205
No. of failures =	1,328	Wald χ2(12)=	208
Time at risk =	63,356	Prob > χ2=	0.000
Log likelihood=	-2,259	(Std. Err. adjusted for 1,302 clusters in id)	
AIC=	4,547		
BIC=	4,632		

Time to Switch Without VF	Coef.	Robust SE	z	P>z	[95% CI]		dy/dx
Gender (Female=1)	0.073	0.083	0.880	0.379	-0.090	0.236	-2.7
Age	0.006	0.004	1.5	0.148	-0.002	0.014	-0.079
$Log_{10}VL$	-0.100	0.034	-2.9	0.003	-0.167	-0.033	1.3
CD4+	0.000	0.000	0.230	0.815	0.000	0.000	0.003
HCV	0.246	0.083	3.0	0.003	0.083	0.408	-5.2
Adherence	0.030	0.003	8.9	0.000	0.023	0.037	-0.940
Year 1st ART	0.034	0.019	1.8	0.072	-0.003	0.070	-1.2
Number of PIs	0.199	0.041	4.9	0.000	0.119	0.279	-5.7
Resistance Level	-0.055	0.035	-1.6	0.119	-0.124	0.014	0.269
Regimen Number	-0.097	0.034	-2.9	0.004	-0.164	-0.030	1.1
Line (1 is reference)							
Line 2	0.096	0.092	1.0	0.299	-0.085	0.276	-2.8
Line 3	-0.072	0.124	-0.580	0.561	-0.314	0.170	-1.4
Constant	-76.2	37.4	-2.0	0.041	-149.5	-3.0	
/ln_p	0.405	0.040	10.2	0.000	0.327	0.483	35 [32;39]
p	1.5	0.060			1.4	1.6	Pred. median=41

Note: ART=Antiretroviral Therapy, CD4+=CD4+ T-Lymphocyte count per µL, HVC=Hepatitis C-Virus, $Log_{10}VL=log_{10}$ HIV RNA copies per mL, PI=Protease Inhibitor, VF=Virological Failure

Table 57: Regression Model for Time to First Resistance

Weibull Regression -- Log Relative-Hazard Form

No. of subjects =	1,298	Number of obs=	3,094
No. of failures =	60	Wald $\chi2(8)$=	40.7
Time at risk =	61,754	Prob > $\chi2$=	0.000
Log Pseudolikelihood=	-248.3		
AIC=	98.0	(Std. Err. adjusted for 1,298	
BIC=	144.3	clusters in id)	

Time to First Resistance	Coef.	Robust	z	P>z	[95% CI]		dy/dx
Gender (Female=1)	-0.04	0.290	-0.13	0.896	-0.606	0.530	8.59
Age	0.02	0.012	1.98	0.048	0.000	0.046	-5.18
$Log_{10}VL$	0.37	0.093	4.05	0.000	0.193	0.556	-84.45
HCV	0.12	0.289	0.40	0.690	-0.451	0.682	-26.00
Adherence	-0.01	0.006	-2.21	0.027	-0.024	-0.001	2.91
Year 1st ART	0.10	0.056	1.76	0.078	-0.011	0.209	-22.29
Number of PIs	0.15	0.108	1.39	0.165	-0.062	0.362	-33.89
Regimen Number	-0.02	0.154	-0.12	0.908	-0.319	0.284	3.99
Constant	-208.48	112.970	-1.85	0.065	-429.891	12.941	
/ln_p	0.41	0.186	2.18	0.029	0.041	0.772	Pred.
p	1.50	0.280	1.04	2.164			Median=338

Notes: ART=Antiretroviral Therapy, HVC=Hepatitis C Virus, $Log_{10}VL=log_{10}$ HIV RNA copies per mL, PI=Protease Inhibitor

Table 58: Regression Model for Time to Resistance Class Switch

Weibull Regression -- Log Relative-Hazard Form

No. of subjects =	2,323	Number of obs=	2,713
No. of failures =	1,236	Wald $\chi2(6)$=	186
Time at risk =	74,712	Prob > $\chi2$=	0.000
Log likelihood=	-1,238	(Std. Err. adjusted for	
AIC=	1,145	2,323 clusters in id)	

Time to Resistance Class Switch	Coef.	Robust SE	z	P>z	[95% CI]		dy/dx
Gender (Female=1)	0.013	0.065	0.2	0.838	-0.114	0.141	-0.486
Age	0.015	0.003	5.3	0.000	0.010	0.021	-0.555
$Log_{10}VL$	0.089	0.036	2.4	0.015	0.018	0.160	-3.249
Year 1st ART	-0.001	0.000	-7.0	0.000	-0.001	0.000	0.020
Regimen Number	0.172	0.037	4.7	0.000	0.100	0.245	-6.310
Number of PIs	-0.187	0.034	-5.5	0.000	-0.254	-0.120	6.850
Constant	-6.7	0.289	-23.1	0.000	-7.3	-6.1	
/ln_p	0.421	0.029	14.4	0.000	0.364	0.479	87 [85;89]
p	1.524	0.045			1.4	1.6	Pred. median=86

Note: ART=Antiretroviral Therapy, CI=Confidence Interval, $Log_{10}VL=Log_{10}$ HIV RNA copies per mL, PI=Protease

Table 59: Regression Model for Time to Hospitalization

Weibull Regression -- Log Relative-Hazard Form

No. of subjects =	1,301	Number of obs	41,721
No. of failures =	600	Wald χ2(10)=	98.4
Time at risk =	63,282	Prob > χ2=	0.000
Log pseudolikelihood=	-1,621		
AIC=	3,267	(Std. Err. adjusted for	
BIC=	3,379	1,301 clusters in id)	

Time to hospitalization Event	Coef.	Robust SE	z	P>z	[95% CI]		dy/dx
Gender (Female=1)	-0.088	0.162	-0.5	0.587	-0.405	0.229	11.149
Age	0.020	0.008	2.5	0.011	0.005	0.036	-2.502
$Log_{10}VL$	0.146	0.053	2.8	0.006	0.042	0.251	-18.216
CD4+	-0.001	0.000	-3.6	0.000	-0.002	-0.001	0.146
HCV	0.255	0.136	1.9	0.061	-0.012	0.523	-32.008
Adherence	-0.003	0.003	-1.1	0.278	-0.010	0.003	0.424
Year of 1st ART	-0.010	0.021	-0.5	0.617	-0.051	0.030	1.295
Number of PIs	0.028	0.074	0.4	0.709	-0.118	0.173	-3.453
Resistance Level	0.064	0.035	1.8	0.070	-0.005	0.133	-7.927
Line (1 is reference)							
2	0.289	0.132	2.2	0.029	0.030	0.548	-34.500
3	0.711	0.242	2.9	0.003	0.236	1.186	-64.061
Constant	16.3	41.5	0.4	0.694	-65.1	97.7	
/ln_p	-0.311	0.057	-5.4	0.000	-0.423	-0.199	72 [64;82]
p	0.733	0.042			0.655	0.820	Pred. median=90

Note: ART=Antiretroviral Therapy,HVC=Hepatitis C-Virus, $Log_{10}VL=log_{10}$ HIV RNA copies per mL, PI=Protease Inhibitor

Table 60: Time to Death among HIV Infected Individuals

Weibull Regression -- Log Relative-Hazard Form

No. of subjects =	34,185	Number of obs=	34,185
No. of failures =	8,278	Wald χ2(7)=	3,807
Time at risk =	2,875,986	Prob > χ2=	0.000
Log likelihood=	-30,183		
AIC=	60,387		
BIC=	60,471		

Time Death	Coef.	Robust SE	z	P>z	[95% CI]		dy/dx
Time							
Gender (Female=1)	-0.189	0.033	-5.8	0.000	-0.253	-0.125	1,523
Age at Diagnosis	0.021	0.001	17.3	0.000	0.019	0.024	-153
Transmission Group							
Heterosexual (reference)							
IDU	0.523	0.031	16.8	0.000	0.462	0.584	-3,692
Homosexual	0.140	0.039	3.6	0.000	0.063	0.217	-900
Other	0.516	0.059	8.7	0.000	0.399	0.632	-2,300
Year of Diagnosis	-0.092	0.003	-34.1	0.000	-0.097	-0.086	660
HIV Stage							
Non-AIDS (reference)							
AIDS	2.7	0.066	41.1	0.000	2.597	2.9	-18,143
Constant	176.9	5.4	33.0	0.000	166.4	187.5	
ln_p							
HIV Stage							
Non-AIDS (reference)							
AIDS	-0.205	0.026	-7.9	0.000	-0.255	-0.154	
Constant	-0.726	0.025	-29.5	0.000	-0.774	-0.678	

Notes: AIDS=Acquired Immune Deficiency Syndrome, HIV=Human Immunodeficiency Virus, IDU=Injection Drug User

Appendix VI

The Mixed Logit Model

A Mixed Logit Model is any model whose choice probabilities can be expressed in the form

$$P_{ni} = \int L_{ni}(\boldsymbol{\beta}) f(\boldsymbol{\beta}) d\beta$$

where,

$$L_{ni}(\boldsymbol{\beta}) = \frac{e^{V_{ni}(\boldsymbol{\beta})}}{\sum_{j=1}^{J} e^{V_{nj}(\boldsymbol{\beta})}}$$

There are, hence, two sets of parameters to be estimated in a mixed Logit model: the parameters **β** and the parameters that describe the density of β, $f(\boldsymbol{\beta})$.

In our analysis, the coefficient associated with physical and mental health was assumed fixed as the common main attribute associated with quality of life among HIV infected individuals, while the remaining factors were assumed random, since their weight in quality of life may vary significantly among individuals. This allows us to quantify the dispersion in the relative weight of the remaining factors with respect to health.

A Normal distribution was assumed for the distribution of the random parameters. Employment was assumed to follow a Normal distribution, to allow for the possibility that employment may be regarded as undesirable by some individuals. Fear of future consequences of the disease and/or premature death was also assumed to follow a Normal distribution to allow for the possibility of the "*will to live*" phenomenon described in the literature [178]. The coefficients associated with the remaining attributes are expected to be positive for every individual and to ensure estimation results which are in line with that assumption, a Lognormal distribution is often used.

In the present analysis, utility was assumed linear in β, consequently:

$$P_{ni} = \int \left(\frac{e^{\boldsymbol{\beta}' \boldsymbol{x}_{ni}}}{\sum_j e^{\boldsymbol{\beta}' \boldsymbol{x}_{nj}}} \right) f(\boldsymbol{\beta}) d\boldsymbol{\beta}$$

In this case, the mixed Logit Probability is a weighted average of the Logit formula evaluated at different values of **β**, with the weights given by the density f (**β**)[110].

In the questionnaire, respondents were asked to make repeated choices. Each individual was asked to choose the worse alternative in each of the 9 choice sets presented. Consequently, the data obtained has a panel structure. Given that all choices were performed in a single moment in time, it was assumed that the β parameters, although varying among people, were constant over choices of the same individual[111]. In this context, utility from alternative j in choice situation t by person n is $U_{njt} = \beta_n x_{njt} + \varepsilon_{njt}$, with ε_{njt} being independent and identically distributed (*i.i.d.)* extreme value over time, people, and alternatives.

Conditional on **β**, the probability that the person makes this sequence of choices $i = (i_1, \ldots, i_T)$ is:

$$L_{ni}(\boldsymbol{\beta}) = \prod_{t=1}^{T} \left[\frac{e^{\boldsymbol{\beta}_n{}' \boldsymbol{x}_{nit}}}{\sum_j e^{\boldsymbol{\beta}_n{}' \boldsymbol{x}_{njt}}} \right]$$

since the $\varepsilon_{nj\,t}$'s are independent over time. The uncondicional probability is, thus, the integral of this product over all possible values of β:

$$P_{ni} = \int L_{ni}(\boldsymbol{\beta}) f(\boldsymbol{\beta}) d\boldsymbol{\beta}$$

It should be noted that the *i.i.d*. assumption on ε_{njt} does not imply the independence from irrelevant alternatives (IIA) assumption, since the ratio of mixed Logit probabilities, $\frac{P_{ni}}{P_{nj}}$, depends on all the data, including attributes of alternatives other than i or j (since the denominators of the Logit formula are inside the integrals, they do not cancel one another)

In fact, in the random coefficient specification, utility is always correlated among alternatives. To observe this fact, the coefficients $\boldsymbol{\beta}_n$ can be decomposed into their mean **α** and deviations μ_n, so that $U_{nj} = \boldsymbol{\alpha} \mathbf{x}_{nj} + \boldsymbol{\mu}_n{}' \mathbf{x}_{nj} + \varepsilon_{nj}$, where $\boldsymbol{\mu}_n{}' \mathbf{x}_{nj} + \varepsilon_{nj}$, represents the unobservable part of utility. The covariance between alternative i and alternative j is, therefore: $E(\boldsymbol{\mu}_n \mathbf{x}_{ni} + \varepsilon_{ni})(\boldsymbol{\mu}_n \mathbf{x}_{nj} + \varepsilon_{nj}) = \mathbf{x}_{ni} W \mathbf{x}_{nj}$, where W is the covariance of μ_n. Utility is correlated over alternatives even when, as in most specifications, the error components are independent, such as when W is diagonal.

[110] In the statistics literature, the weighted average of several functions is called a mixed function, and the density that provides the weights is called the mixing distribution.

[111] Identification issues arise with regard to socio-demographic variables, as these do not vary among alternatives and in choice models only differences in utility matter.

Along with the percentage of correct predictions, the overall significance of the model is evaluated by comparing the value of the log likelihood function with the base model, where the base case coefficients were all set to zero [234]. Two test statistics were used, both based on comparison of the likelihood function:

The pseudo-R2 or Likelihood Ratio Index of McFadden:

$$McFaddenR^2 = 1 - \frac{LL_{estimated}}{LL_{base}}$$

And the Wald statistics

$$Wald\ \chi^2 = -2(LL_{base} - LL_{estimated}) \sim \chi^2_{estimated\ -base}$$

The WHOQoLHIV and WHOQoL-BREF – a comparison

Table 61 provides a summary of the average scores on each domain of the questionnaires used by Canavarro *et al.* [239] and Reis *et al.* [132, 134]. Average scores are not directly comparable, since the instrument used is not the same. In Canavarro *et al.* the WHOQoL-HIV is used while in Reis *et al.* the WHOQoL-BREF was selected. Table 62 provides a comparison between the facets included in each.

Table 61: Average WHOQOL-HIV and WHOQOL-BREF Scores in Portugal

Source	Canavarro *et al.* (2008)							Reis *et al.* (2007)							
Instrument	WHOQOL-HIV							WHOQOL-BREF							
HIV Stage	Asymptomatic		Symptomatic		AIDS			Asymptomatic		Symptomatic		AIDS		All	
N	87		37		39			168		7		15		190	
	Average	SD	Average	SD	Average	SD	p-value	Average	SD	Average	SD	Average	SD	Average	SD
Physical	0.557	0.193	0.453	0.155	0.428	0.180	0.000	0.587	0.027	0.443	0.003	0.488	0.016	0.574	0.026
Pscycological	0.643	0.195	0.583	0.151	0.555	0.176	0.020	0.588	0.026	0.411	0.015	0.449	0.044	0.570	0.029
Level of Independence	0.691	0.190	0.576	0.155	0.508	0.150	0.000								
Social Relationships	0.650	0.177	0.533	0.151	0.579	0.178	0.002	0.540	0.043	0.393	0.015	0.405	0.030	0.524	0.043
Environment	0.598	0.150	0.543	0.111	0.526	0.135	0.010	0.543	0.025	0.393	0.020	0.445	0.054	0.530	0.028
Spirituality	0.527	0.192	0.526	0.152	0.484	0.188	0.440								
Overall	0.597	0.211	0.492	0.171	0.474	0.174	0.001	0.541	0.201	0.306	0.189	0.394	0.258	0.521	0.213

Notes: In Canavarro *et al.* (2008) of the 200 individuals, 36 did not known their HIV-stage. HIV=Human Immunodeficiency Virus, N=Number of Observations/Individuals, SD=Standard Deviation, WHOQOL-HIV=World Health Organization Quality of Life Questionary for HIV Patients, WHOQOL-BREF=World Health Organization Quality of Life Questionary for HIV

With respect to sample variability, the Canavarro et al. sample has a higher proportion of women (40% versus 31%) and more educated individuals (18% with college degree versus 8%). Age, which is an important factor in HIV infection characterization, namely due to its link to the

mode of transmission, is not reported in the Canavarro et al. sample and is 40 years-old in the Reis et al. sample. The Canavarro sample does not provide information on where the questionnaires took place.

Table 62: Correspondence between the WHOQOL-HIV / WHOQOL-BREF Facets and the DCE Attributes

Model Domains	WHOQOL-BREF FACET	WHOQOL-HIV FACET	Description
Health	3	1	Pain and Disconfort
	10	2	Energy and Fatigue
	16	3	Sleep and Rest
		50	Symptoms of PLWHA
	5	4	Positive Feelings / Enjoy Life
	7	5	Thinking, Learning, Memory and Concentration
	19	6	Self-Esteem
	11	7	Bodily Image and Appearance
	26	8	Negative Feelings
Lifestyle	15	9	Mobility
	17	10	Activities of Daily Living
	4	11	Dependence on Medication or Treatments
	18	12	Work Capacity
Love / Sexual Relationship	20	13	Personal Relationships
	22	14	Social Support
	21	15	Sexual Activity
Discrimination		51	Social Inclusion
	22	16	Physical Safety and Security
	21	17	Home Environment
	19	18	Financial Resources
Healthcare	24	19	Health and Social Care: Accessibility and Quality
	13	20	Access to Relevant Information
	14	21	Participation in Recreation / Leisure Activities
	23	22	Physical Environment
	25	23	Transport
Future	6	24	Spirituality / Religion / Personal Beliefs
		52	Forgiveness and Blame*
		53	Concerns about the Future*
		54	Death and Dying*

Notes: * HIV Specific Questions

HIV=Human Immunodeficiency Virus, WHOQOL-BREF: World Health Organization Quality of Life Questionary for HIV Patients, Brief Version, WHOQOL-HIV: World Health Organization Quality of Life Questionary for HIV Patients, PLWHA=People Living With HIV/AIDS

The Discrete Choice Questionnaire

VIH+
Que qualidade de vida?

A melhoria da qualidade de vida das pessoas infectadas pelo VIH implica conhecer os factores que determinam essa mesma qualidade de vida. Preencher este questionário é um passo nesse sentido.

Este questionário é anónimo. Muito obrigada pela sua colaboração.

Esta participação será voluntária, pelo que poderá interrompê-la a qualquer momento.

Filipa Aragão, Escola Nacional de Saúde Pública.

Os dados recolhidos serão utilizados exclusivamente no âmbito da tese de Doutoramento.
Os resultados agregados da análise serão apresentados publicamente.

Consentimento informado

Gostaríamos de saber se aceita colaborar neste estudo respondendo a algumas questões.
As suas informações são estritamente confidenciais pois os resultados serão codificados e utilizadas apenas neste estudo. Esta participação será voluntária , pelo que poderá interrompê-la a qualquer momento.

Data ...../...../.....
Assinatura do Investigador
_____

Data ...../...../.....
Assinatura do participante
_____

Caracterização

Sexo: Masculino ☐ Feminino ☐

Idade: ☐

Categoria de transmissão:

- Heterossexual ☐
- Homossexual ☐
- Toxicodependência ☐
- Outra/desconhecida ☐

CD4+: ☐ (Aproximadamente)

Instruções

As respostas a este questionário consistem em **escolher**, para cada par de situações hipotéticas, **quem, na sua opinião, está em pior situação**. Por exemplo:

				Na sua opinião, quem está em *pior situação?*
Imagine que a Maria:	Está bem de saúde	Tem uma excelente relação amorosa/ sexual	Tem um bom ordenado	
Imagine que a Isabel:	Está doente	Tem uma péssima relação amorosa/sexual	Está desempregada	X

Como a Isabel está pior que a Maria, deverá ser colocada uma cruz no quadrado em frente à situação da Isabel.

1.

Na sua opinião, quem está em *pior situação?*

Imagine que o Zé:	Sente-se doente, deprimido ou incapaz de cuidar de si próprio	Está impedido de fazer o que gosta por causa do tratamento e/ou a medicação provoca-lhe efeitos adversos graves	Tem dificuldade em levantar a medicação, em fazer análises, em falar com seu o médico assistente quando necessita.	Nunca se sente discriminado, a infecção por VIH não tem qualquer estigma social	A infecção pelo VIH pouco ou nada afecta a sua relação amorosa/sexual	Tem emprego	Não se preocupa com o facto de poder morrer precocemente ou de sofrer com doenças graves
Imagine que o Manuel:	Sentre-se saudável, com energia.	Tem o tratamento perfeitamente integrado no seu estilo de vida e/ou a medicação não lhe provoca efeitos adversos graves	Tem acesso fácil à medicação, às análises de rotina e ao médico assistente sempre que necessita	No dia-a-dia sente-se constantemente discriminado	Não consegue ter ou manter uma relação amorosa/sexual pelo facto de ser VIH+.	Está desempregado	Vive preocupado com com o facto de poder morrer precocemente ou sofrer com doenças graves

2.

Na sua opinião, quem está em *pior situação?*

Imagine que o Zé:	Sente-se doente, deprimido ou incapaz de cuidar de si próprio	Tem o tratamento perfeitamente integrado no seu estilo de vida e/ou a medicação não lhe provoca efeitos adversos graves	Tem acesso fácil à medicação, às análises de rotina e ao médico assistente sempre que necessita.	No dia-a-dia sente-se constantemente discriminado	Não consegue ter ou manter uma relação amorosa/sexual pelo facto de ser VIH+.	Tem emprego	Não se preocupa com o facto de poder morrer precocemente ou de sofrer com doenças graves
Imagine que o Manuel:	Sentre-se saudável, com energia.	Está impedido de fazer o que gosta por causa do tratamento e/ou a medicação provoca-lhe efeitos adversos graves	Tem dificuldade em levantar a medicação, em fazer análises, em falar com seu o médico assistente quando necessita.	Nunca se sente discriminado, a infecção por VIH não tem qualquer estigma social	A infecção pelo VIH pouco ou nada afecta a sua relação amorosa/sexual	Está desempregado	Vive preocupado com com o facto de poder morrer precocemente ou sofrer com doenças graves

3.

Na sua opinião, quem está em *pior situação?*

Imagine que o Zé:	Sentre-se saudável, com energia.	Está impedido de fazer o que gosta por causa do tratamento e/ou a medicação provoca-lhe efeitos adversos graves	Tem acesso fácil à medicação, às análises de rotina e ao médico assistente sempre que necessita.	No dia-a-dia sente-se constantemente discriminado	A infecção pelo VIH pouco ou nada afecta a sua relação amorosa/sexual	Está desempregado	Não se preocupa com o facto de poder morrer precocemente ou de sofrer com doenças graves
Imagine que o Manuel:	Sente-se doente, deprimido ou incapaz de cuidar de si próprio	Tem o tratamento perfeitamente integrado no seu estilo de vida e/ou a medicação não lhe provoca efeitos adversos graves	Tem dificuldade em levantar a medicação, em fazer análises, em falar com seu o médico assistente quando necessita.	Nunca se sente discriminado, a infecção por VIH não tem qualquer estigma social	Não consegue ter ou manter uma relação amorosa/sexual pelo facto de ser VIH+.	Tem emprego	Vive preocupado com com o facto de poder morrer precocemente ou sofrer com doenças graves

4.

Na sua opinião, quem está em *pior situação?*

Imagine que o Zé:	Sentre-se saudável, com energia.	Tem o tratamento perfeitamente integrado no seu estilo de vida e/ou a medicação não lhe provoca efeitos adversos graves	Tem dificuldade em levantar a medicação, em fazer análises, em falar com seu o médico assistente quando necessita.	Nunca se sente discriminado, a infecção por VIH não tem qualquer estigma social	Não consegue ter ou manter uma relação amorosa/sexual pelo facto de ser VIH+.	Está desempregado	Não se preocupa com o facto de poder morrer precocemente ou de sofrer com doenças graves
Imagine que o Manuel:	Sente-se doente, deprimido ou incapaz de cuidar de si próprio	Está impedido de fazer o que gosta por causa do tratamento e/ou a medicação provoca-lhe efeitos adversos graves	Tem acesso fácil à medicação, às análises de rotina e ao médico assistente sempre que necessita.	No dia-a-dia sente-se constantemente discriminado	A infecção pelo VIH pouco ou nada afecta a sua relação amorosa/sexual	Tem emprego	Vive preocupado com com o facto de poder morrer precocemente ou sofrer com doenças graves

5.

Na sua opinião, quem está em *pior situação?*

Imagine que o Zé:	Sentre-se saudável, com energia.	Tem o tratamento perfeitamente integrado no seu estilo de vida e/ou a medicação não lhe provoca efeitos adversos graves	Tem acesso fácil à medicação, às análises de rotina e ao médico assistente sempre que necessita.	Nunca se sente discriminado, a infecção por VIH não tem qualquer estigma social	A infecção pelo VIH pouco ou nada afecta a sua relação amorosa/sexual	Tem emprego	Não se preocupa com o facto de poder morrer precocemente ou de sofrer com doenças graves
Imagine que o Manuel:	Sente-se doente, deprimido ou incapaz de cuidar de si próprio	Está impedido de fazer o que gosta por causa do tratamento e/ou a medicação provoca-lhe efeitos adversos graves	Tem dificuldade em levantar a medicação, em fazer análises, em falar com seu o médico assistente quando necessita.	No dia-a-dia sente-se constantemente discriminado	Não consegue ter ou manter uma relação amorosa/sexual pelo facto de ser VIH+.	Está desempregado	Vive preocupado com com o facto de poder morrer precocemente ou sofrer com doenças graves

6.

Na sua opinião, quem está em *pior situação?*

Imagine que o Zé:	Sente-se doente, deprimido ou incapaz de cuidar de si próprio	Está impedido de fazer o que gosta por causa do tratamento e/ou a medicação provoca-lhe efeitos adversos graves	Tem acesso fácil à medicação, às análises de rotina e ao médico assistente sempre que necessita.	Nunca se sente discriminado, a infecção por VIH não tem qualquer estigma social	Não consegue ter ou manter uma relação amorosa/sexual pelo facto de ser VIH+.	Está desempregado	Vive preocupado com com o facto de poder morrer precocemente ou sofrer com doenças graves
Imagine que o Manuel:	Sentre-se saudável, com energia.	Tem o tratamento perfeitamente integrado no seu estilo de vida e/ou a medicação não lhe provoca efeitos adversos graves	Tem dificuldade em levantar a medicação, em fazer análises, em falar com seu o médico assistente quando necessita.	No dia-a-dia sente-se constantemente discriminado	A infecção pelo VIH pouco ou nada afecta a sua relação amorosa/sexual	Tem emprego	Não se preocupa com o facto de poder morrer precocemente ou de sofrer com doenças graves

7.

Na sua opinião, quem está em *pior situação?*

Imagine que o Zé:	Sente-se doente, deprimido ou incapaz de cuidar de si próprio	Tem o tratamento perfeitamente integrado no seu estilo de vida e/ou a medicação não lhe provoca efeitos adversos graves	Tem dificuldade em levantar a medicação, em fazer análises, em falar com seu o médico assistente quando necessita.	No dia-a-dia sente-se constantemente discriminado	A infecção pelo VIH pouco ou nada afecta a sua relação amorosa/sexual	Está desempregado	Vive preocupado com com o facto de poder morrer precocemente ou sofrer com doenças graves
Imagine que o Manuel:	Sentre-se saudável, com energia.	Está impedido de fazer o que gosta por causa do tratamento e/ou a medicação provoca-lhe efeitos adversos graves	Tem acesso fácil à medicação, às análises de rotina e ao médico assistente sempre que necessita.	Nunca se sente discriminado, a infecção por VIH não tem qualquer estigma social	Não consegue ter ou manter uma relação amorosa/sexual pelo facto de ser VIH+.	Tem emprego	Não se preocupa com o facto de poder morrer precocemente ou de sofrer com doenças graves

8.

Na sua opinião, quem está em *pior situação?*

Imagine que o Zé:	Sentre-se saudável, com energia.	Está impedido de fazer o que gosta por causa do tratamento e/ou a medicação provoca-lhe efeitos adversos graves	Tem dificuldade em levantar a medicação, em fazer análises, em falar com seu o médico assistente quando necessita.	No dia-a-dia sente-se constantemente discriminado	Não consegue ter ou manter uma relação amorosa/sexual pelo facto de ser VIH+.	Tem emprego	Vive preocupado com com o facto de poder morrer precocemente ou sofrer com doenças graves
Imagine que o Manuel:	Sente-se doente, deprimido ou incapaz de cuidar de si próprio	Tem o tratamento perfeitamente integrado no seu estilo de vida e/ou a medicação não lhe provoca efeitos adversos graves	Tem acesso fácil à medicação, às análises de rotina e ao médico assistente sempre que necessita.	Nunca se sente discriminado, a infecção por VIH não tem qualquer estigma social	A infecção pelo VIH pouco ou nada afecta a sua relação amorosa/sexual	Está desempregado	Não se preocupa com o facto de poder morrer precocemente ou de sofrer com doenças graves

9.

Na sua opinião, quem está em *pior situação?*

Imagine que o Zé:	Sentre-se saudável, com energia.	Tem o tratamento perfeitamente integrado no seu estilo de vida e/ou a medicação não lhe provoca efeitos adversos graves	Tem acesso fácil à medicação, às análises de rotina e ao médico assistente sempre que necessita.	Nunca se sente discriminado, a infecção por VIH não tem qualquer estigma social	A infecção pelo VIH pouco ou nada afecta a sua relação amorosa/sexual	Tem emprego	Vive preocupado com com o facto de poder morrer precocemente ou sofrer com doenças graves
Imagine que o Manuel:	Sente-se doente, deprimido ou incapaz de cuidar de si próprio	Está impedido de fazer o que gosta por causa do tratamento e/ou a medicação provoca-lhe efeitos adversos graves	Tem dificuldade em levantar a medicação, em fazer análises, em falar com seu o médico assistente quando necessita.	No dia-a-dia sente-se constantemente discriminado	Não consegue ter ou manter uma relação amorosa/sexual pelo facto de ser VIH+.	Está desempregado	Não se preocupa com o facto de poder morrer precocemente ou de sofrer com doenças graves

Muito obrigada pela sua colaboração!!