How Does Medical Innovation Create Value?
Health, Human Capital and the Labor Market[*]

Nicholas W. Papageorge[†]

May 2, 2012

Abstract: This paper develops a dynamic framework to value medical innovation that takes explicit account of the link between health, human capital and the labor market. Using a characteristics approach, I model medical innovation as increased efficacy or reduced side effects. The framework is applied to HIV treatments, including a 1996 medical breakthrough (HAART) that transformed HIV infection from a virtual death sentence into a manageable, chronic condition. The main findings are (1) For an HIV+ individual, HAART introduction was worth between \$20,000 and \$340,000, with higher values corresponding to younger agents or agents with greater work experience. (2) A counterfactual version of HAART without side effects is worth up to an additional \$375,000. (3) When no treatment dominates others in terms of efficacy and side effects, agents optimally cycle among available options. In general, sicker agents choose effective treatments despite harsh side effects, but switch to less effective drugs with fewer side effects once their health improves. (4) Treatment cycling is partly determined by the labor market. If non-wage income declines, agents are compelled to work, increasing their frequency of switching to treatments with fewer side effects. More aggressive treatment cycling, in turn, diminishes average health and reduces survival probability.

1 Introduction

1.1 Overview

Biomedical research expenditures in the U.S. currently exceed $100 billion each year. Resulting medical innovation creates value through its impact on how patients live and how they work. Increased life-expectancy is an undeniable source of such value.[1] Less understood are the myriad mechanisms through which medical innovation improves the quality-of-life.[2] Quality-of-life includes how a patient feels, but it also reflects the impact of health on other facets of life, like consumption, income and employment.[3] Insofar as the labor market interacts with health and shapes medical decision-making, it helps to determine how medical innovation creates value.

This paper develops a general framework to measure the value of medical innovation, taking explicit account of the links between health, human capital and the labor market. A cornerstone of the approach is that treatments are measured along two dimensions: (a) effectiveness at improving underlying health, which governs both symptoms and life-expectancy and (b) propensity to cause side effects. Symptoms and side effects manifest as physical 'ailments' and these ailments affect patient utility.

Forward-looking consumers choose treatments fully anticipating the impact on their health. Effective treatments often entail harsh side effects and patient decisions reflect this trade-off. For example, a cancer patient might eschew a highly effective treatment (like chemotherapy) in favor of a less effective treatment with fewer side effects (like radiation therapy or even surveillance). This choice may prevent the immediate utility cost of side effects, but the patient risks recurrence later on, which could entail symptoms and a shorter lifespan.

The labor market plays a key role in this trade-off. Ailments induced by symptoms or side effects are potentially debilitating and painful unto themselves. They also influence the quality-of-life through their impact on labor choices and outcomes. In particular, poor health can lower productivity, depress income and exacerbate the utility cost of work. Forward-looking agents respond by shifting employment, fully aware that their decisions affect current consumption, the accumulation of human capital and future income streams.

---

[1] Murphy and Topel (2006) study the value of longevity gains attributed to biomedical research on heart disease. A more detailed discussion is found in the literature review (Section 1.2).

[2] An exception is the literature on quality-adjusted life-years (QALYs), which studies a broad array of interactions between health and quality-of-life. Relying on stated or elicited preferences, QALYs measure the value of a given health state, weighted by mortality and probabilities of reaching different health states in future periods. See, for example, Lipscomb et al. (2009)

[3] Grossman (1972) treats health like a form of capital stock in which agents invest. In his framework, good health is valuable since it increases agents' healthy time, which can be used to produce and consume market goods.

The dynamic trade-off between symptoms and side effects becomes a permanent fixture of agent decision-making in the case of chronic diseases or conditions that can be managed, but not cured (e.g. diabetes, multiple sclerosis, and HIV). Chronically ill, forward-looking agents develop complex employment and treatment plans to balance their long-term demand for consumption, physical comfort and, ultimately, longevity. I apply the framework developed in this paper to a chronic condition, studying treatment and employment decisions of HIV+ men. I use data from the Multi-Center AIDS Cohort Study, an ongoing investigation (beginning in 1984) of HIV infection in homosexual and bisexual men. Information is collected on objective hematological health measures like HIV status and CD4 count. Throughout this paper, underlying health is proxied by CD4 count, which measures immune system functionality. Data are also collected on individual reports of physical ailments like nausea and fever, medical treatment choices, employment decisions and labor market outcomes including income.

HIV and the AIDS epidemic provide a natural setting to study how medical innovation interacts with the labor market since identifying this link requires health status variation strong enough to affect employment. Untreated, HIV infection leads to immune system deterioration (known as AIDS) where fairly routine infections can lead to grave symptoms, illness and death. Absent treatment, a newly infected HIV+ subject lives an average of 11 years. Additionally, phases of the AIDS epidemic are distinguished by wide variability in treatments available to infected subjects. Key to identifying model parameters is that I observe treatment and employment choices both before and after a medical breakthrough known as HAART.[4] A treatment regime introduced in 1996, HAART effectively transformed HIV infection from a virtual death sentence into a chronic and manageable condition, though at the cost of harsh side effects.[5]

The econometric model is specified so that patients have preferences over physical ailments (manifestations of symptoms, side effects or both) and consumption.[6] They choose medical treatments to maximize utility as opposed to immune system health or survival probability. This distinction is crucial - if agents optimally avoid an effective treatment due to its harsh side effects, its value as a medical innovation is limited.[7] Moreover, each medical treatment

---

[4]Another study that exploits pre- and post-HAART observations to study how a medical breakthrough can affect outcomes is Duggan and Evans (2005).

[5]HAART stands for highly active anti-retroviral treatment. There is no vaccine or cure for HIV or AIDS, but HAART is the current standard treatment. In general, 1996 is marked as the year when two crucial clinical guidelines that comprise HAART came to be commonly acknowledged. First, protease inhibitors (made widely available towards the end of 1995) would be an effective HIV treatment. Second, several anti-retroviral drugs taken simultaneously could indefinitely delay the onset of AIDS.

[6]The structural framework also permits state-dependent utility, allowing health to shift both level utility and the marginal utility of consumption.

[7]This framework echoes earlier work on agent decision-making when faced with the risk of HIV infection. In particular, observed risky sexual decisions are not consistent with maximization of health or longevity, but

is modeled according to its propensity to produce health and physical ailments (through efficacy and side effects). This characteristics approach to estimating demand permits valuation of counterfactual treatment innovations, each defined as a bundle of characteristics. In particular, the framework allows a counterfactual innovation to be valued in light of its own effectiveness and side effects profile, the profiles of other existing or counterfactual treatments, agent preferences and the labor market environment.

The model is also designed to capture several interactions between health and employment. First, physical ailments can shift the utility cost of work. Second, health status is interacted with income, a link that likely works through health-induced productivity changes. Third, illness can induce costly gaps in employment. In particular, income is allowed to vary by experience and the model permits endogenous human capital accumulation, measured by observed periods of full-time work. Employment decisions therefore contain a forward-looking component: work experience can enhance human capital and future wages via on-the-job training. If ill agents do not work, they might arrest this process, which implies an additional avenue through which medical innovation can create value: by preventing health-induced employment gaps.

Reduced-form data analysis of treatment choices highlights an important pattern. Sicker agents tend to use more effective treatments like HAART and effective treatments are associated with better health. The dynamic choice model explains these patterns as part of a long-term optimal plan of treatment cycling. This behavior is driven by three principle factors: observed persistence in immune system health, a non-convexity in treatment choice due to discrete options and patient distaste for side effects. In general, cycling reflects how agents treat their underlying health as a form of capital stock. Agents in poor health choose effective treatments despite harsh side effects since they face a high marginal benefit of investing in their health 'stock'. Once their health improves, agents exploit persistence in health, switching to less effective drugs with fewer side effects. In essence, they reap the benefits of earlier health stock investments to enjoy periods free of costly physical ailments. In so doing, agents anticipate switching back to harsher treatments once their health deteriorates. I henceforth call this phenomenon Optimal Treatment Cycling or OTC. Cycling off of the most effective treatment available (sometimes known as taking a 'drug holiday') is often deemed to be suboptimal behavior that should be curbed. In the framework developed in this paper, such cycling behavior is part of a dynamically optimal treatment plan.[8]

A second and more puzzling pattern in the data involves the interaction between treatment choices and labor supply. Although full-time employment is associated with better health,

---

reflect trade-offs between current and future utility (Philipson and Posner, 1993).

[8]See, for example, Sabate (2003) on treatment non-compliance.

the use of HAART is negatively associated with employment. The dynamic choice model reveals that physical ailments are a crucial factor in determining employment decisions. For about half of agents (distinguished by their latent preference type) the utility cost of work is negatively affected by physical ailments arising from symptoms or side effects. When faced with a high probability of physical ailments, these agents reduce full-time employment. For example, a counterfactual policy experiment shows that if HAART had never been introduced, employment would have been about 30% lower after 1996. This finding reflects that, despite its harsh-side effects, HAART led to a net decrease in ailments through health improvements. The net reduction in ailments encouraged full-time employment, which increased average income and accelerated human capital accumulation.

Main results of this study begin with a valuation of HAART introduction. Modeling treatment and employment decisions for agents at different stages of their career and life-cycle reveals striking heterogeneity in the value of HAART introduction. For an HIV+ individual, HAART was worth between $25,000 and $350,000, with greater value associated with younger agents and those with high human capital.[9] Moreover, a decomposition exercise shows that up to 20% of the value of HAART is attributable to labor market factors. This portion is higher for younger agents since they face a long work life. It is also higher for agents with higher human capital since they earn more in each life-year gained.

The characteristics approach to estimating demand means that the model can be used to analyze counterfactual treatment innovations defined by their side effects and efficacy profiles. Three key results emerge. First, side effects reductions are potentially very valuable. Assuming HAART is the baseline technology, agents who are young or who have high human capital value the introduction of a version of HAART with no side effects at about $375,000. This result underscores the importance of including quality-of-life measures to fully understand the value of medical innovation. It also calls into question standard critiques of so-called "me-too" or "follow-on" drugs. Defined as treatments that are not clinically superior to existing treatments, "me-too" drugs are seen to diminish incentives for innovation without providing therapeutic benefit to consumers.[10] I find that a treatment that is therapeutically equivalent to HAART, but that entails fewer side effects, is valuable.

Second, young agents value a counterfactual, highly effective version of HAART at about $1,200,000. For forward-looking consumers, a portion of this value comes from the utility gained from foregoing that treatment, coupled with the possibility of choosing it when needed. In other words, the option value of a highly effective treatment consists in making milder treatments viable choices. The mechanism behind this phenomenon is optimal treatment

---

[9]Values throughout this paper are in 2003 dollars.

[10]For a critique of me-too drugs, see for example Angell (2000).

cycling: if agents anticipate a high probability of restored health by using effective treatment, cycling off effective treatment becomes part of a safe and viable dynamic plan. Third, the value of a potential medical innovation depends on existing treatments. For example, if a highly effective version of HAART already exists, the invention of a version of HAART with few side effects adds little value. The logic here is straightforward: if highly effective HAART is available, agents exploit its option value by cycling off treatment altogether when their health improves, thereby avoiding any and all side effects. In this context, a treatment with low side effects is fairly redundant.

Changes in the labor market environment influence patient outcomes through their impact on optimal treatment cycling. To highlight cyclical treatment choice behavior given varied labor market conditions, I simulate an environment in which agents face two counterfactual treatment technologies: one that has the same side effects as HAART, but is highly effective and another that has the same efficacy as HAART, but no side effects. Next, I simulate a decline in non-wage income, i.e., agents earn less if they choose to not work full-time. Since agents suffering HIV/AIDS are permitted to collect disability payments, this counterfactual scenario mimics an environment where public transfers are reduced and agents are more compelled to work. In response, agents in relatively good health increase their frequency of switching to treatments with fewer side effects. More aggressive treatment cycling does indeed lead to higher employment and income and accelerates the accumulation of human capital. However, it also entails worse health and decreased survival.

Finally, the cost of HIV treatment faced by consumers is much lower than the actual costs paid by both public and private insurances. A final policy experiment simulates agent choices and outcomes in an environment where treatment cost subsidies disappear. Unsurprisingly, agents engage in more aggressive treatment cycling, switching off treatment with increased frequency when in good health. More pronounced cycling does lower average health, though only slightly, and has little effect on survival probability. Perhaps more surprisingly, agents who dislike working with ailments are more likely to work since fewer periods on treatment entails a lower probability of suffering side effects. In general, both of the aforementioned policy experiments illustrate how chronically ill patients balance demand for consumption, physical comfort and longevity through their long-term treatment and employment plans. They also reflect the role of optimal treatment cycling in determining how agents invest in their health stock.

1.2 Literature Review

An extensive literature in economics asks if medical research investments are worthwhile, relying on two dominant approaches to measure the value of medical innovation. The first focuses on value accruing through increased life expectancy and finds large returns to biomedical research. In the case of heart disease treatment advances, Murphy and Topel (2006) show that the life-expectancy benefits far outweigh the research costs.[11] The second approach to valuing medical innovations goes beyond longevity to consider a much broader array of health-related outcomes, using stated or elicited preferences to compute the value of quality-adjusted life years or QALYs (Lipscomb et al., 2009). The framework I develop differs from both approaches in that it values both longevity and quality-of-life benefits of medical innovation, relying on revealed preferences (through observed choices) to estimate demand. This method avoids some of the problems associated with stated or elicited preferences. A key example is context dependence that can undermine generalizability. Moreover, estimating demand using observed choices permits a study of how treatment choices interact with other decisions, like employment. Designed this way, the estimated model can be used to perform counterfactual policy experiments that capture how changes to either the health or labor environment can affect both health and labor choices and outcomes.

Starting with Grossman (1972), the connection between health, human capital and labor has been examined in numerous studies. Recent research develops this line of inquiry, envisioning health as a form of human capital stock in terms of labor market outcomes (e.g.: Becker (2007), Heckman and Cunha (2007), Currie (2009) and Conti et al. (2010)). Currie and Madrian (1999) provide a survey of literature on the health-labor link, highlighting an empirical consensus that poor health has a negative impact on: productivity, earnings and labor supply. These findings are reflected in the context of HIV and HAART. Goldman and Bao (2004) consider joint employment and treatment decisions of HIV+ men after HAART is introduced, finding that HAART-use is associated with a higher likelihood of remaining employed.[12]

Health and labor market activity can be linked in two important ways. First, poor health can exacerbate the utility cost of work, thereby discouraging employment. This phenomenon is suggested by studies showing that medical technology improvements in the context of chronic illness can increase labor supply. Key examples are Kahn (1998) (who studies labor supply of diabetics facing insulin innovations) and Garthwaite (2010) (who considers nearly-elderly men's labor supply *vis a vis* Vioxx, a drug that reduces pain and inflammation).

---

[11] A collection of papers that take a similar approach is found in (Murphy and Topel, 2003).

[12] Further examples from other fields include Blalock et al. (2002), who studies whether HIV patients benefit psychologically from remaining employed.

Both studies are relevant to the present investigation since they highlight how labor supply is sensitive not only to underlying health *per se*, but also to how chronically ill or aging patients feel. A second link between health and labor is through wage, which can be affected by illness, since poor health potentially decreases productivity. For example, Levine et al. (1997) and Cawley (2004) study the negative impact on wages of smoking and obesity, respectively. In general, these links suggest specific mechanisms through which medical innovation creates value via its effect on labor market choices and outcomes.

Current employment decisions also contain a forward-looking component since work experience can enhance human capital and future wages via on-the-job training. Imai and Keane (2004) show that failure to account for returns to work accruing through human capital accumulation leads to erroneous estimation of the wage elasticity of labor supply. Endogenous human capital accumulation implies an important additional cost of health-induced employment gaps (Mincer and Polachek (1974) and Becker (1985)). Related to the present study, Eckstein and Wolpin (1989) model labor supply and allow for endogenous human capital accumulation in the context of females who face the decision of having children. Though pregnancy is certainly not a chronic illness, women contemplating child-bearing face a choice that is somewhat similar to chronically ill patients who choose effective medication with harsh side effects. Both decisions imply a future payoff (i.e., children or better health) at the cost of current-period consumption. However, it is crucial to capture that both decisions entail an additional cost in the form of employment gaps, which slow the process of human capital accumulation.

Symptoms or side effects can lower utility, altering patient treatment choices. Also important is how illness shifts the marginal utility of consumption, which could affect employment decisions. Failure to account for this shift could bias the estimated utility cost of working while ill.[13] Research on health-state dependent utility often exploits within-individual changes in utility proxies (such as measures of subjective well-being or of depression-like symptoms) at different health states (e.g.: Smith et al. (2005)). A potential drawback is that poor health might lower consumption (e.g.: due to reduced income or employment or increased medical expenditures) so that parameters measuring state-dependence could be confounded with differences in the marginal utility of consumption at different consumption levels (Smith, 1999)[14]. Finkelstein et al. (2008) counter these potential problems by focusing

---

[13]How consumption utility changes with health is *a priori* unclear. Agents who largely consume complements to health (e.g. outdoor activities or nightlife) would exhibit lower marginal utility of consumption as their health deteriorates; agents consuming mostly substitutes to good health (e.g. air-conditioning or taxi-cabs) would experience the opposite.

[14]Another approach uses survey data on the willingness-to-pay for lower risk of exposure to certain types of illness or injury at different wealth levels (e.g. Evans and Viscusi (1991), Viscusi and Evans (1990) and Sloan et al. (1998)).

on agents who are both retired and insured so that they would presumably avoid large consumption shifts induced by illness.[15] Since I account for employment decisions and resulting income shifts and include data on medical expenditures, I avoid bias by essentially controlling for the impact of health status on consumption levels.[16]

Previous work on how patients choose medication focuses largely on experimentation and learning, both of which are found to be important in explaining medical treatment decisions (Crawford and Shum (2005), Fernandez (2008), Chan and Hamilton (2006) and Chintagunta, Jiang and Jin (2009)). Common to this line of inquiry is that underlying health enters patients' utility function. The approach taken in the present investigation is different in that patients have preferences over consumption and physical comfort, both of which are affected by underlying health. This distinction is important. Patients might generally behave in accordance with a desire to improve their underlying health. However, counterfactual policy experiments aimed at capturing the health-labor link require a precise specification of patient preferences that captures the various reasons why health is valuable.[17]

The remainder of this paper is organized as follows: Section 2 describes the data; Section 3 presents the model; Section 4 explains the estimation procedure and discusses parameter estimates and model fit; Section 5 presents a valuation of HAART; Section 6 describes findings from counterfactual policy experiments and Section 7 concludes.

2   Data

2.1   The Multi-Center AIDS Cohort Study

I use the public data set from the Multi-Center AIDS Cohort Study (MACS), an ongoing study (beginning in 1984) of the natural and treated histories of HIV infection in homosexual and bisexual men conducted at four sites: Baltimore, Chicago, Pittsburgh and Los Angeles.[18]

---

[15]Their main finding is that poor health lowers the marginal utility of consumption by about 11%.

[16]Finkelstein et al. (2008) point out that a further potential drawback to this approach of inferring state dependence is that it is sensitive to assumptions on agents' bequest motives. The fact that subjects in my analysis are gay men suggests that they likely will not have children so that bequest motives may not play an important role in how consumption is allocated across health states.

[17]Chan and Hamilton (2006) take a step in this direction by modeling patient preferences over both side effects and underlying health. They show that patients - especially those in better health - are willing to forego treatment effectiveness for reductions in side effects. This finding is reflected in the results of the present study.

[18]Data in this manuscript were collected by the Multicenter AIDS Cohort Study (MACS) with centers (Principal Investigators) at The Johns Hopkins Bloomberg School of Public Health (Joseph B. Margolick, Lisa P. Jacobson), Howard Brown Health Center, Feinberg School of Medicine, Northwestern University, and Cook County Bureau of Health Services (John P. Phair, Steven M. Wolinsky), University of California, Los Angeles (Roger Detels), and University of Pittsburgh (Charles R. Rinaldo). The MACS is funded by the

At each biannual visit, data are collected on: treatment choices, employment decisions, labor market variables and health status, which includes CD4 count (a measure of immune system health) and subject reports of ailments. Ailments can be either symptoms of HIV or side effects of treatment. The data is a panel, which permits analysis of agent decisions both before and after HAART introduction and also allows the model to account for unobserved heterogeneity.

The full MACS data set contains information on 5,622 subjects at 41 possible visits for a total of 98,886 subject-visits. Since treatment information is somewhat sporadic before 1990, I focus on the years 1990 to 2003. In each period after 1990, agents report all medications they have used since their previous interview. As there are dozens of medications used to fight HIV infection, I follow medical literature on HIV treatment (see, for example, Detels et al. (2001)) in creating four broad and mutually exclusive treatment categories: no treatment, mono-therapy, combo-therapy and HAART.[19] To create a measure for experience, I use potential experience (current age minus 25) up until the start of the AIDS epidemic (corresponding to the baseline sample period) and thereafter construct employment histories using observed labor supply choices. Although more specific data is available for some sample periods, I model employment choices to be dichotomous: full time or not full time. Finally, in constructing the analysis sample, I drop all HIV− subjects or positive subjects while they are HIV−, observations lacking crucial health or employment information and observations where period-to-period transitions are unclear due to missed interviews. The resulting analysis sample is an unbalanced panel of 8,300 observations: 743 subjects over 26 visits.[20]

In subsequent analysis, HAART introduction is treated as a quasi-experiment.[21] Two

[19]An agent with the label "none" may take medications to fight opportunistic infections, such as pneumonia. Mono-therapy denotes a regimen consisting of a single nucleoside reverse transcriptase inhibitor (NRTI). Combo-therapy consists of several NRTIs. HAART has a more complex definition that includes several drug regimens, most of which include a protease inhibitor in combination with an NRTI or a non-nucleoside reverse transcriptase inhibitor (NNRTI).

[20]Employment histories are constructed using all available data, including observations when positive agents were observed negative, if applicable, and observations with up to two missed subsequent visits, in which case I assume that agents engage in the same employment status as in the last observed period. For analysis, however, model specification requires that for each observation the subject be observed in the following period. Dropping observations that do not fit these criteria leads to an unbalanced panel.

[21]This assumption implies the need for caution in applying the framework developed in this paper to cases where medical innovation is anticipated. A possible generalization of the current model would permit agents to expect an improvement in medical technology with a small probability in each period. Such an extension would be most salient for chronic conditions where agents are expected to remain alive for many periods so that they would reasonably expect to be alive to enjoy an innovation in the distant future and might alter their current behavior accordingly. For HIV/AIDS, pre-HAART life expectancy was so low that this was

observations justify this approach. First, HAART was not a single medication developed and improved over time such that subjects might update their beliefs and anticipate higher future efficacy. Rather, HAART introduction was abrupt and many components of HAART already existed prior to 1996. The key insight involved the union of several existing technologies, none of which was particularly effective on its own. Second, subject reports from survey questions asking about their hopefulness about the future are not consistent with anticipation of HAART. Specifically, one in a battery of questions meant to assess depression asks subjects how often in the week preceding their interview they felt hopeful about the future.[22] Figure 1 plots the probability that subjects answer, "all or most of the time" over time. Note the pre-HAART flat (or even downward) trend followed by a break and reversal coinciding with HAART introduction. Importantly, if the effectiveness of HAART had been anticipated, this upward shift in hopefulness should have occurred before HAART introduction.[23]

To illustrate the variation used to identify structural parameters, Figures 2-4 depict the impact of HAART on a variety of measures-of-interest. Figure 2 plots death probability over time among HIV+ subjects in the analysis sample and, for comparison, HIV− subjects in the MACS data set. HAART introduction coincided with a drastic decrease in deaths so that the death rate for HIV+ subjects plunged to levels not far above those of HIV− subjects. Figure 3 plots treatment choices over time. It appears that agents substitute HAART for other treatments and that those who refrained from using earlier, less effective treatments chose HAART after 1996. This figure also hints at a high degree of persistence in medication usage. A transition matrix (Table 1) verifies this persistence, though does indicate some cycling behavior both among treatments and on and off treatments. Figures 4 depicts labor supply decisions over time. HAART coincides with a break in the decreasing trend of full-time employment in the aging sample. To underline the significance of this break, I extrapolate the pre-HAART full-time employment trend until 2001.[24] This exercise suggests that a counterfactual world without HAART may have witnessed lower employment among HIV+ men. Finally, Table 2 shows transitions in and out of the labor force. Transitions highlight a higher degree of persistence in employment decisions, suggesting possible heterogeneity in

likely not the case.

[22]Questions are from a depression screening tests known as the Center for Epidemiological Studies Depression (or CES-D) scale. See, for example, Ostrow et al. (1989), for an example of CES-D scale use with the MACS data set.

[23]One concern is that hopefulness is highly correlated with health so that the trend reversal simply reflects HAART-induced health improvements. To account for this, I control for a polynomial in CD4 count and age in a regression where the regressand is a dichotomous variable for being hopeful about the future 'most or all of the time'. Plotted residuals show a similar trend reversal at, but not before, HAART introduction.

[24]In particular, I regress pre-HAART employment decisions on age, age-squared and a linear time trend and then use these parameters to predict employment decisions in the post-HAART era, taking post-HAART age profiles as given.

the utility cost of work.

Summary statistics for the full sample and then separately for the periods before and after HAART and by work status are presented in Table 3. In the full sample, subjects are about 32-years-old at the start of the AIDS epidemic (1984). The main health measure is CD4 count, a measure of immune system functionality. Healthy HIV− individuals generally present a CD4 count of between 500 and 1500 units per mm^3 of blood. The sample average is slightly lower: about 450. For subsequent analysis, I construct a binary measure of health that indicates whether patient CD4 count is low enough to indicate AIDS (< 250). According to this measure, about one quarter of observations exhibit AIDS-level CD4. Subjects also report a number of ailments, which may indicate symptoms of AIDS, side effects of medications (or both). I construct an indicator variable for symptoms and side effects, which takes the value 0 if agents report persistently experiencing one of the following ailments: fatigue, diarrhea, headaches, fever or drenching sweats. About 60% of subjects report that they are free of such ailments (i.e. $F = 1$). Finally, death probability is about 4% over the entire sample period.

Considering the pre-HAART and post-HAART eras separately, there are important differences. Foremost are improvements in health measures (CD4 count) although the same proportion of agents suffer ailments. This is likely due to the increase in subjects using harsh treatment so that side effects replace symptoms. The increase in treatment costs supports this. There is some post-HAART movement towards public insurance (in favor or private or no insurance) which could reflect changes to laws governing disability or an aging population. In either case, this suggests that changes in insurance provision are an important factor to account for in subsequent analysis. Finally, death probability plunges after HAART introduction. Next, comparing agents by their employment status, there is a clear indication that good health (both CD4 count and lack of ailments) predict employment. Curiously, the likelihood of HAART use is higher among non-workers than workers, which could reflect a number of factors, such as an aging cohort. It may also reflect that effective treatments entail side effects that discourage employment even though good health ultimately encourages full-time work.

Annual individual gross income is presented in Table 3 (as reported and in 2003 dollars).[25] Average reported income is about $38,000. Income reported by non-workers (about $26,500), though lower than that reported by full-time workers (about $44,000), may seem high, but reflects that HIV/AIDS is considered a disability, which opens up the possibility of social security disability payments and private pensions. In support of this possibility, HIV+ agents' non-wage income is positively correlated to their wage income in periods be-

---

[25]Income is a categorical variable. To convert per-period income into dollars, I take the midpoint of each category and then divide it by 2.

fore they were infected with HIV.[26] In the model, allowing non-working agents to collect high income essentially assumes that public and private disability payments can be collected with no transaction costs.[27]

2.2   Labor Supply and Treatment Choices

Reduced form analysis of labor supply, pre-HAART treatment and post-HAART treatment choices are found in Tables 4 - 6. Results suggest that agents are highly persistent in their labor supply choices. Experience and good health, including both CD4 count and absence of ailments, predict employment. Post-HAART estimates relating labor and treatment choices (Column (5)) suggest a negative relationship between treatments and labor supply. These estimates suggest a puzzling dynamic. First, working agents who go onto HIV treatment are less likely to work in the following period. Second, treatments improve health. Third, healthy agents are more likely to work.

Results from (multinomial logit) regressions of treatment choices for before and after HAART introduction are found in Tables 5 and 6, respectively. Column (1) of Table 5 regresses treatment choice on CD4 count and previous-period choices For both the pre- and post-HAART eras, estimates suggest that treatment choice exhibits high persistence and that agents with low CD4 count are more likely to go onto treatment. The same holds for subjects with public insurance and higher human capital. In general, these results point to a cyclical component of treatment choice: agents who are less healthy go onto treatment; treatments improve health; and agents who are healthier are less likely to use treatment.

Consider Column (4) of Table 6, where employment is included as a regressor. Again, there is a negative relationship between employment and HIV treatment and this relationship is significant in the case of HAART. The result is puzzling since although health and employment are positively correlated, effective treatment and employment are negatively correlated. A possible explanation is that although improved health encourages employment by reducing symptoms, side effects have the opposite effect. Also, these results suggests that the cyclical component of treatment-use may have an impact on employment decisions. The structural model is designed to shed light on these patterns, explaining how agents formulate dynamic

---

[26]Note that part-time work is counted as "not-full-time", but excluding individuals reporting part-time work does not appreciably affect these averages.

[27]How public disability payments relate to labor supply has been studied in detail in a number of contexts. Singleton (2011) finds fewer post-HAART reports of disability among HIV+ agents. He also finds persistence among agents who were receiving disability payments prior to HAART introduction and suggests that this reflects difficulties in applying for such payments. French and Song (2010) finds that agents receiving disability payments subsequently supply less labor than similar agents who are not accepted into the program. In the present investigation, transaction costs associated with disability payments are captured by preference parameters for labor supply.

treatment and employment plans given the trade-off between symptoms and side effects.

3   Model

In each period, agents enjoy flow utility, a function of current choices and state variables. Before retirement at age 65, agents choose treatments and employment at each period. Agents are forward-looking, so their choices maximize the present discounted value of future utility. Agents retire at age 65 and cease making employment decisions, but continue to make treatment decisions. Period t state variables are a function of previous-period states and choices. I solve the dynamic programming problem of agents using backward induction.

3.1   Choices and Flow Utility

At each period t and until retirement agents choose a pair $d_{it} \equiv (d_{it}^L, d_{it}^M)$, where d_{it}^L represents the employment choice and d_{it}^M the treatment choice. In particular, the possible choices on each dimension are:

$$d_{it}^L = \begin{cases} 0 & \text{Not full time work} \\ 1 & \text{Full time work} \end{cases} \quad \text{and} \quad d_{it}^M = \begin{cases} 0 & \text{No Treatment} \\ 1 & \text{Mono-therapy} \\ 2 & \text{Combo-therapy} \\ 3 & \text{HAART (only after 1996)} \end{cases} \tag{1}$$

Note that the set of choice pairs, which I denote D_t, is time-dependent since HAART is available only after 1996. Specifically, denoting as D_t^L and D_t^M the set of labor and treatment options available at period t, respectively, $D_t \equiv D_t^L \times D_t^M$. Ailment status is given by $F_{it} \in \{0, 1\}$, where 1 signifies being free of ailments and 0 signifies suffering ailments. Flow

utility is given by:

$$
\begin{aligned}
U(C_{it}, F_{it}, d_{it}) =\ & u\left(C_{it}, F_{it}, \gamma(F_{it})\right) \\
+\ & \theta_1 \times \mathbf{1}\{F_{it} = 0\} \\
+\ & \theta_2 \times \mathbf{1}\{d_{it}^L = 1\} \\
+\ & \theta_3 \times \mathbf{1}\{d_{it}^L = 1\} \times \mathbf{1}\{F_{it} = 0\} \\
+\ & \theta_4 \times \mathbf{1}\{d_{i,t-1}^M = 0\} \times \mathbf{1}\{d_{it}^M \neq 0\} \times \mathbf{1}\{F_{it} = 0\} \\
+\ & \theta_5 \times \mathbf{1}\{d_{i,t-1}^M \neq d_{it}^M\} \times \mathbf{1}\{F_{it} = 0\} \\
+\ & \theta_6 \times \mathbf{1}\{d_{i,t-1}^M \neq 0\} \times \mathbf{1}\{d_{it}^M = 0\} \times \mathbf{1}\{F_{it} = 0\} \\
+\ & \theta_7 \times \mathbf{1}\{d_{i,t-1}^M = 0\} \times \mathbf{1}\{d_{it}^M \neq 0\} \times \mathbf{1}\{F_{it} = 1\} \\
+\ & \theta_8 \times \mathbf{1}\{d_{i,t-1}^M \neq d_{it}^M\} \times \mathbf{1}\{F_{it} = 1\} \\
+\ & \theta_9 \times \mathbf{1}\{d_{i,t-1}^M \neq 0\} \times \mathbf{1}\{d_{it}^M = 0\} \times \mathbf{1}\{F_{it} = 1\} \\
+\ & \epsilon(d_{it})
\end{aligned}
\tag{2}
$$

This first term on the right-hand-side represents individual utility over consumption (C_{it}). In particular, I assume that the marginal utility of consumption varies by ailment status and that $u(\cdot)$ is a CRRA utility function with parameter $\gamma(F_{it})$ so that

$$
u\left(C_{it}, F_{it}, \gamma(F_{it})\right) = \frac{1}{1 - \gamma(F_{it})} C_{it}^{1 - \gamma(F_{it})}.
\tag{3}
$$

The second term on the right-hand-side of Equation 2 represents the utility cost of suffering ailments. The third term captures the utility cost of full-time work. The fourth stands for the additional cost of working while suffering ailments.

Period t treatment choices affect long-term value (through their effect on health as measured by CD4 count) and current period flow utility (through ailments induced by side effects). Both of these processes will be explained in the following section. Treatment choices also enter flow utility directly via switching costs, captured by the fifth through tenth terms of the right-hand-side of Equation 2. Finally, $\epsilon(d_{it})$ is a choice-specific utility-shifter, which captures factors that affect agent choices, but that are not observable to the econometrician. In particular, $\epsilon_{it} : D_t \to \mathbb{R}$ and I use $\epsilon(d_{it})$ to denote the utility shifter associated with choice d_{it}. Moreover, $\epsilon_{it}(d_{it})$ are independent across time, agents and choices and Extreme Value Type I distributed.[28]

Switching costs capture factors - beyond preferences over ailments and long-term health - that affect agent treatment decisions, including doctors' orders, treatment protocols and

---

[28]This assumption (along with assumptions on conditional independence of states and outcomes, which will be formally stated later) follows assumptions maintained in Rust (1987).

14

the social benefits of antiretrovirals. In particular, effective HIV treatments lower viral loads (the amount of virus in a patient's blood), which renders patients less infectious to HIV− sex partners. Note that switching costs are generic, i.e., not specific to any particular treatment. Instead, agents experience a cost of starting, switching or ending treatment. Moreover, switching costs vary by ailment status.

Note that this specification of preferences amounts to a characteristics approach to modeling treatment. In other words, patients do not have preferences over a specific treatment like HAART. Pioneered by: Stigler (1945), Lancaster (1966), Rosen (1974) and Gorman (1980), this approach permits the valuation of counterfactual treatment innovations, each defined by the probability distribution it implies over CD4 count and ailments.[29,30] The processes according to which choices and states generate ailments and consumption are described in the following section.

3.2   States and Transitions

Upon entering period t, the agent learns his vector of period-t state variables (denoted S_{it}), but he still faces uncertainty about ailments (F_{it}) and consumption (C_{it}), both of which are realized only after he makes his labor supply and treatment decision. Therefore, the agent evaluates expected flow utility conditional on his current choice d_{it} and his vector of period-t state variables, formally:

$$\mathrm{E}[U(C_{it}, F_{it}, d_{it})|S_{it}]. \tag{4}$$

The agent's treatment and labor supply decision has a direct impact on the stochastic generation process of F_{it} and C_{it}. Finally, choices and current states jointly determine period-$(t+1)$ state variables.

State variables (S_{it}) include a vector of observables, denoted $\mathcal{X}_{it}$, and a vector of unobservable utility shifters (ϵ_{it}). Specifically, $\mathcal{X}_{it} \equiv [H_{i,t-1}, A_{i,t-1}, E_{i,t-1}, v_{t-1}]$, where:

$$
\begin{aligned}
H_{i,t-1} &\in \{0,1\} &:& \quad \text{High (non-AIDS) CD4 count at } t \\
A_{i,t-1} &\in \{25, 25.5, 26, \ldots, 65\} &:& \quad \text{Age at } t \\
E_{i,t-1} &\in \{10, 20, \ldots, 50\} &:& \quad \text{Semesters of full-time experience at } t \\
v_{t-1} &\in \{1, \ldots, 15\} &:& \quad \text{Period-}t \text{ dummy}
\end{aligned}
$$

---

[29]Moreover, note that similarly to the approach in Lancaster (1966) ailments can be interpreted as outputs of the activity 'treatment consumption'.

[30]Further contributions using this approach include: Trajtenberg (1989) Bresnahan and Stern (1997) Petrin (2004) and Chan (2006).

Recall from Section 2 that HIV infection leads to a low CD4 count, which means that the patient's immune system is compromised.[31] S_{it} also includes the unobserved, choice-specific utility shifters ($\epsilon(d_{it})$'s) defined in the previous section.

Next, the agent forms expectations on F_{it} and C_{it}, which are collected into a vector denoted $\mathcal{Y}_{it}$ so that: $\mathcal{Y}_{it} = [F_{it}, C_{it}]$.[32] I assume conditional independence of $\mathcal{Y}_{it}$, i.e., outcomes are independent of realizations of unobservable flow utility shifters. Formally:

$$\mathrm{E}[\mathcal{Y}_{it}|\mathcal{X}_{it}, d_{it}, \epsilon_{it}] = \mathrm{E}[\mathcal{Y}_{it}|\mathcal{X}_{it}, d_{it}]. \tag{5}$$

Ailments F_{it} evolve according to:

$$\mathrm{P}\left[F_{it} = 1|X_{it}^F; \theta^F\right] = \frac{\exp(X_{it}^F \theta^F)}{1 + \exp(\cdot)} \tag{6}$$

where $X_{it}^F \equiv \left[H_{i,t-1}, v_{i,t-1}, H_{i,t-1} \times d_{it}^M\right]$ and θ^F is a vector of parameters governing the process generating ailments.

Consumption is equal to income (I_{it}) minus out-of-pocket treatment costs (p_{it}).[33] Formally,

$$C_{it} = I_{it} - p_{it} . \tag{7}$$

Evaluating expected consumption requires several steps since the agent faces uncertainty on both income and treatment costs. Agent income uncertainty reflects unanticipated shocks. For example, an agent may fall ill at some point before the end of period t and incur an income loss for missing work days. Agents also form expectations on out-of-pocket treatment costs (p_{it}. These are a function of underlying health at the end of period t (H_{it}) and period-t insurance provision (N_{it}), both of which are unknown at t. This setup reflects that, after choosing a treatment category at t, out-of-pocket treatment costs will depend on their (as yet unrealized) health state throughout the period. In summary, to derive expected consumption given period-t choices and states, the agent must form expectations on income (I_{it}), insurance (N_{it}), CD4 count (H_{it}) and out-of-pocket treatment costs (p_{it}). Each of these stochastic

---

[31]To reduce the size of the state space, H_{it} is binary, but captures the most salient feature of CD4 count, namely, whether it is low enough to suggest AIDS (i.e. <250). Transitions between binary health states are fairly low, reflecting persistence in continuous CD4 count. However, dichotomous health does not reflect that agents with CD4 counts near 250 face a higher probability of switching health states. An extension of the current model would capture this difference by permitting H_{it} to take on more values.

[32]Note that F_{it} and C_{it} are not state variables so do not belong to S_{it}, but do affect utility. Such variables are often deemed 'payoffs' or 'outcomes'.

[33]Agents in the model cannot save. The potential impact of this assumption is discussed as results are presented.

processes is explained in turn. Income is modeled as

$$I_{it} = X_{it}^I \theta^I + \epsilon_{it}^I \tag{8}$$

where $X_{it}^I \equiv \left[(E_{i,t-1}, E_{i,t-1}^2, A_{i,t-1}, H_{i,t-1}, v_{i,t-1}) \times d_{it}^L \right]$, $\epsilon_{it}^I \sim N\left(0, \sigma_I^2\right)$ and θ^I denotes a vector of parameters governing the income process.[34],[35] Note that state variables affecting the income process are interacted with period-t employment decisions. This reflects that an agent's current state can affect wage and non-wage income in different ways.

Insurance status (N_{it}) affects treatment costs and is also modeled as a process determined by state variables and labor supply decisions.[36] Formally,

$$P\left[N_{it}|X_{it}^N; \theta^N\right] = \frac{\exp(X_{it}^N \theta^N)}{1 + \exp(\cdot)}, \tag{9}$$

where $X_{it}^N = \left[H_{i,t-1}, E_{i,t-1}, E_{i,t-1}^2, A_{i,t-1}, A_{i,t-1}^2, v_{i,t-1}, d_{it}^L\right]$ and θ^N is a vector of parameters governing the insurance process.

Underlying health, as measured by CD4 count, is affected by treatments. The salient features to be captured are (a) whether treatment (or lack thereof) moves CD4 above or below AIDS levels and (b) possible persistence in CD4 count. First, ΔH_{it} indicates whether an agent's CD4 increased (versus either decreased or remained unchanged) between periods t and $t+1$. ΔH_{it} evolves according to

$$P\left[\Delta H_{it} = 1|X_{it}^{\Delta H}, d_{it}^M; \theta^{\Delta H}\right] = \frac{\exp(X_{it}^{\Delta H} \theta^{\Delta H})}{1 + \exp(\cdot)}, \tag{10}$$

where $X_{it}^{\Delta H} \equiv \left[H_{i,t-1}, v_{i,t-1}, d_{it}^M \times H_{i,t-1}\right]$. In other words, both treatments and period-t CD4 count determine if CD4 count increases or not. Then, period-t CD4 count and the direction of change ΔH_{it} determine whether CD4 is above or below AIDS levels in $t+1$. In particular, for parameters θ^H, the CD4 count process is modeled as:

$$P\left[H_{it} = 1|X_{it}^H, d_{it}^M; \theta^H\right] = \frac{\exp(X_{it}^H \theta^H)}{1 + \exp(\cdot)} \tag{11}$$

---

[34]Income is a function of health at the beginning of the period $H_{i,t-1}$. This modeling choices reflects the timing of income offers and employment decisions. After learning his health status, the agent faces income offers for full-time employment. Employers know agent productivity, which is a function of health and human capital. The employer does not, however, know which medications will be chosen, so the income offer is not a function of expected ailment status.

[35]Main results of the paper are robust to different methods to capture progressive taxation.

[36]Insurance could also be modeled as a choice. However, MACS includes no data on insurance options. Also, insurance provision is highly persistent in the data and largely dependent on employment, so I model insurance provision as a process that agents indirectly control through their labor supply decisions.

where $X_{it}^H \equiv [\Delta H_{it} \times H_{i,t-1}]$.

Out-of-pocket treatment costs are modeled as

$$p_{it} = X_{it}^P \theta^P + \epsilon_{it}^P \tag{12}$$

where $X_{it}^P \equiv \left[H_{it} \times F_{it}, I_{it}, N_{it} \times d_{it}^M, v_{it} \right]$, $\epsilon_{it}^P \sim N\left(0, \sigma_P^2\right)$ and θ^P is a vector of parameters.[37] Given the processes specified above, expected consumption is formally defined as:

$$\mathrm{E}\left[C_{it} | \mathcal{X}_{it}, d_{it}\right] = \mathrm{E}\left[I_{it} | I_{it} \geq 0, \mathcal{X}_{it}, d_{it}\right] - \mathrm{E}[p_{it} | p_{it} \geq 0, \mathcal{X}_{it}, d_{it}]. \tag{13}$$

Note that both income and treatment costs are assumed to be non-negative.

Until now, I have described the stochastic processes governing each component of flow utility. The model is dynamic in the sense that, in making his current decision, the agent must also evaluate how his choices and current state affect the distribution over future states. Formally, define the state-to-state distribution function for current (observable) state $\mathcal{X}_{it}$, current choice d_{it} and period-$t+1$ (observable) state $\mathcal{X}_{i,t+1}$ as

$$G_X(\mathcal{X}_{i,t+1} | \mathcal{X}_{it}, d_{it}). \tag{14}$$

I further assume that the distribution over future states is independent of current unobservable state variables $\epsilon(d_{it})$ conditional on current observable state variables and choices. Formally,

$$\mathrm{E}[\mathcal{X}_{i,t+1} | \mathcal{X}_{it}, d_{it}, \epsilon_{it}] = \mathrm{E}[\mathcal{X}_{i,t+1} | \mathcal{X}_{it}, d_{it}]. \tag{15}$$

Furthermore, note that $H_{i,t-1} \in \mathcal{X}_{it}$ evolves according to Equation 11.

Full-time work experience at t, $E_{i,t-1}$ increases by 0.5 for each period of full-time employment. Formally, $E_{it} = E_{i,t-1} + 0.5 \times \mathbf{1}\left[d_{it}^L = 1\right]$. Next, age at t $A_{i,t-1}$ and the time dummy $v_{i,t-1}$ evolve deterministically. Specifically, $A_{it} = A_{i,t-1} + 0.5$ and $v_{it} = v_{i,t-1} + 1$.

Finally, the probability of dying between periods t and $t+1$ is denoted

$$\mathrm{P}\left[B_{it} = 1 | X_{it}^B; \theta^B\right] = \frac{\exp(X_{it}^B \theta^B)}{1 + \exp(\cdot)} \tag{16}$$

where $X_{it}^B = [H_{i,t-1}, A_{i,t-1}, H_{i,t-1} \times A_{i,t-1}]$, B_{it} is an indicator function for death and θ^B is a vector of parameters that govern death probability. Note that current period decisions do not affect the probability of dying. In other words, upon entering the period and learning his

---

[37]Note that the costs process includes I_{it} to account for the possibility that treatments are subsidized according to income.

state variable realizations, the agent either continues on to enjoy period-t flow utility or dies, in which case he receives flow utility 0 forever.

3.3 Parameters and Unobserved Heterogeneity

Flow utility parameters from Equation 2 are collected into a vector denoted θ^U. Parameters governing processes and transition probabilities are denoted θ^{XY} so that

$$\theta^{XY} \equiv \left[\theta^F, \theta^I, \theta^N, \theta^H, \theta^P, \theta^B \right]. \tag{17}$$

Collect these parameters into a vector θ so that $\theta \equiv \left[\theta^U, \theta^{XY} \right]$.

Unobserved heterogeneity is introduced in preference parameters via latent types, of which there is a finite number K, and for the remainder of the paper, I set $K = 2$. I allow the following preference parameters to vary by type: the utility cost of work, the cost of ailments, the interaction between the two and the marginal utility of consumption for each ailment status. This modeling decision arises from high observed persistence in labor supply choices within individuals over time. This is consistent with heterogeneity in distaste for work.[38] Let θ^k denote type-k parameters, where $k \in \{1, \ldots, K\}$. Denote agent i's parameters as θ_i. Type probabilities are given by

$$\pi_k \equiv P[\theta_i = \theta^k]. \tag{18}$$

The subject knows his type k, but the econometrician does not, which means that the distribution over types must be integrated out and the π_k's jointly estimated. Finally, collect all parameters to be estimated into a vector Θ, where

$$\psi = \left[\theta^1, \ldots, \theta^K, \pi_1, \ldots, \pi_K \right]. \tag{19}$$

This concludes the specification of the theoretical model. The following section describes how ψ is estimated.

---

[38]In principle, all parameters could vary by latent type. A variety of specifications have permitted unobserved heterogeneity in parameters governing both health and labor market processes. However, I cannot reject that these parameters are the same for both types.

4   Estimation

The vector of parameters ψ is estimated using a nested procedure.[39] At the "inner" step and given a proposed a set of parameters (denoted $\psi^{(g)}$), the dynamic programming problem is solved via backward induction for each set of observed state variables $\mathcal{X}_{it}$. This yields a set of transitions and choice probabilities, which maximize utility. At the "outer" step, the algorithm searches for parameters that maximize a likelihood function computed from the data

4.1   Value Functions

The structure of the value functions for retired and non-retired agents differs. First, I introduce retirement value functions and then I describe value functions for non-retired agents. The value of retirement is an infinite stream of average expected flow utility. Agents no longer work, but do continue to optimally choose treatment. For retired agents and each treatment choice d_{it}^M, the probability of choosing that treatment is denoted $P^R(d_{it}^M)$. Exploiting distributional assumptions on utility shifters, this probability is given by:

$$P^R\left(d_{it}^M|\mathcal{X}_{it}\right) \equiv \frac{\exp(\tilde{U}(C_{it}, F_{it}, d_{it}^M, d_{it}^L = 0|S_{it}))}{\sum_{d^M \in D_t^M} \exp(\tilde{U}\left(C_{it}, F_{it}, d^M, d_{it}^L = 0|\mathcal{X}_{it}\right))}, \tag{20}$$

where $\tilde{U}(\cdot)$ is flow utility as defined in Equation 2 with the utility-shifter netted out. Using these treatment probabilities, in each post-retirement period, agents receive:

$$\sum_{d_{it}^M \in D_t^M} U\left(C_{it}, F_{it}, d_{it}^M, d_{it}^L = 0|S_{it}\right) \times P^R\left(d_{it}^M|\mathcal{X}_{it}\right). \tag{21}$$

Agents receive this flow utility at all post-retirement ages, though in each period weighted by the discount factor β and the probability of dying conditional on state variables at retirement $P[B_{it} = 1|\cdot]$.[40] Therefore, total retirement value is equal to an infinite sum, given by

$$\begin{aligned} V^R(A_{i,t-1} = 65, S_{it}) &= \\ \frac{P[B_{it}|\cdot]}{1 - \beta\, P[B_{it}|\cdot]} &\times \sum_{d_{it}^M \in D_t} U\left(C_{it}, F_{it}, d_{it}^M, d_{it}^L = 0|S_{it}\right) \times P\left(d_{it}^M|\mathcal{X}_{it}\right). \end{aligned}$$

---

[39]I employ estimation methods developed by Rust (1987) and Hotz and Miller (1993) and surveyed in Aguirregabiria and Mira (2010).

[40]The discount factor β is set to 0.975 per semester.

Note that I assume that agents remain in the same health state at which they retire. This is a reduced-form way to capture that good health is valuable at retirement.[41]

Let us now turn attention to non-retired agents. In every period t, they choose $d_{it} \in D_t$ to maximize

$$\mathrm{E}\left[\sum_{j=0}^{T_i-1} \beta^j U(C_{i,t+j}, F_{i,t+j}, d_{i,t+j}|\mathcal{X}_{it}) + \beta^{T_i} V^R(A_{i,t-1} = 65, S_{it})\right] \qquad (22)$$

where $T_i \equiv (65 - A_{it}) \times 2$ represents the number of periods until retirement. Using the Bellman principle, we can define the value function for periods before retirement as follows:

$$V(S_{it}) = \max_{d_{it} \in D_t}\left\{\mathrm{E}[U(C_{it}, F_{it}, d_{it})] + \beta \int V(S_{i,t+1})\mathrm{d}G_X(\mathcal{X}_{i,t+1}|\mathcal{X}_{it}, d_{it})\right\} \qquad (23)$$

where $G_X(\mathcal{X}_{i,t+1}|\mathcal{X}_{it}, d)$ is defined in Equation 14. Choice-specific value functions can be written as:

$$v(S_{it}, d_{it}) \equiv \mathrm{E}[U(C_{it}, F_{it}, d_{it})] + \beta \sum_{\mathcal{X}_{i,t+1}} \bar{V}(\mathcal{X}_{i,t+1})g_X(\mathcal{X}_{i,t+1}|d, \mathcal{X}_{it}), \qquad (24)$$

where $\bar{V}(\cdot)$ is the expectation of the value function taken over the distribution of $\epsilon(d_{it})$ and $g_X(\cdot)$ is the transition density of $\mathcal{X}_{it}$ corresponding to transition distribution function $G_X(\cdot)$. Notice that $\bar{V}(\cdot)$ takes the form of an expected maximization since the agent does not know future realizations of ϵ_{it}.

Given this setup, I obtain choice probabilities for each set of observable variables via backward induction.[42] For example, suppose that agent i enters period t at age 64.5, so that $A_{i,t-1} = 64.5$. Then, each choice will imply a probability distribution over $\mathcal{X}_{i,t+1}$, from which I compute expected retirement value. Given state-specific retirement value, I compute choice-specific value functions for each state at age 64.5. Once I have obtained choice and state-specific value functions for age 64.5, I can compute choice and state specific value functions for age 64 using Equation 24 and so on until age 30. Note that I do not observe

---

[41]Allowing V^R to be a function of both S_{it} and A_{it} is a slight abuse of notation since A_{it} is an element in the vector dS_{it}. Strictly speaking, S_{it} in this case refers to the vector of observable state variables without A_{it}.

[42]Experience (E_{it}) is measured at five grid points, but estimation requires evaluating value functions between these grid points. For example, if an agent with 10 periods of experience decides to work in period t, his period-$t+1$ experience will be 11. I use linear-spline interpolation (see Judd (1998)) to compute value functions for state variable values that lie between grid points.

$\epsilon_{it}(d)$, but know that its distribution implies the following choice probabilities:

$$P(d_{it}|\mathcal{X}_{it}) = \frac{\exp\{\tilde{V}(\mathcal{X}_{it}, d_{it})\}}{\sum_{d'_{it} \in D_t} \exp\{\tilde{V}(\mathcal{X}_{it}, d'_{it})\}} \tag{25}$$

where $\tilde{V}(\cdot)$ is the net-of-error choice specific value function (i.e., Equation 24 minus $\epsilon_{it}(d_{it})$):

$$\tilde{V}(\mathcal{X}_{it}, d_{it}) = \mathrm{E}[U(C_{it}, F_{it}, d_{it})] + \beta \sum_{\mathcal{X}_{i,t+1}} \bar{V}(\mathcal{X}_{i,t+1}) g_X(\mathcal{X}_{i,t+1}|d, \mathcal{X}_{it}). \tag{26}$$

Finally, note that in the preceding derivations, I have omitted notation identifying type-specific parameters. For each set of suggested parameters $\psi^{(g)}$, the estimation routine includes solving the dynamic programming problem to obtain choice probabilities for each set of type-specific parameters θ^k.

4.2 The Likelihood Function

The likelihood contribution of individual i is:

$$\begin{aligned} L_i(\theta) &= \sum_{k=1}^{K} \pi_k [\Pi_{t=1}^{T_i} P(d_{it}|\mathcal{X}_{it}; \theta^k) \times \Pi_{t=1}^{T_i} g_Y(\mathcal{Y}_{it}|\mathcal{X}_{it}, d_{it}; \theta^{XY}) \\ &\times \Pi_{t=1}^{T_i-1} g_X(\mathcal{X}_{i,t+1}|\mathcal{X}_{it}, d_{it}; \theta^{XY})], \end{aligned} \tag{27}$$

where g_Y denotes the density function derived from processes governing F_{it} and C_{it} and θ^{XYk} denotes type-specific θ^{XY}. Recall, however, that only utility parameters are allowed to vary by type. This means that the second two portions of Equation 27 can be pulled out of the summation over k and the likelihood function can be rewritten as:

$$\begin{aligned} L_i(\theta) &= \sum_{k=1}^{K} \pi_k [\Pi_{t=1}^{T_i} P(d_{it}|\mathcal{X}_{it}; \theta^k)] \\ &\times \Pi_{t=1}^{T_i} g_Y(\mathcal{Y}_{it}|\mathcal{X}_{it}, d_{it}; \theta^{XYk}) \\ &\times \Pi_{t=1}^{T_i-1} g_X(\mathcal{X}_{i,t+1}|\mathcal{X}_{it}, d_{it}; \theta^{XYk}). \end{aligned} \tag{28}$$

Since the final two portions of the likelihood function are outside of the summation over k, the log likelihood function consists of three additively-separable components that can be separately maximized. As such, parameters governing transitions probabilities and the realizations of F_{it} and C_{it} can be estimated in a separate first step, which does not involve solving the dynamic programming problem. In fact, given modeling assumptions, this step requires estimation of a set of Tobit, logistic and multinomial logistic regressions, all of which can be accomplished with standard statistical software. In the second step, I only need to search for preference parameters θ^{Uk}, which denotes type-specific θ^U, along with probability

masses π_k. This decreases the number of iterations, which greatly reduces computation burden since each iteration includes solving the dynamic programming problem.

4.3   Identification

This section discusses how moments in the data identify estimated model parameters. In the data, each period-t choice and state combination implies a probability distribution over period-$t+1$ states and these moments identify parameters governing state-to-state transitions and outcomes. Parameters in the flow utility function govern choice probabilities and are identified through agents' observed state-dependent choices probabilities. The CRRA coefficient γ, which measures the curvature in the utility function, is identified by differences in how agents choose both treatments and employment at different consumption levels. Employment decisions imply large changes in consumption and treatment choices, inducing variability in medical expenditures, imply small changes in consumption. Insofar as choice probabilities for given state variables change at different rates for different consumption levels, these choices trace out the marginal utility of consumption. State-dependent utility parameters ($\gamma(F_{it})$ are identified through differences in γ across health-status. Finally, parameters describing the distribution of latent types and type-specific parameters are identified through repeated observed choices of the same subject over time and given different values of state variables.

4.4   Parameter Estimates

This section presents parameter estimates from first-step regressions of state-to-state transitions and outcomes (θ^X and θ^{XY}). Next, I present preference parameter estimates (θ^U) and go on to compute a posterior type probability for each individual in the dataset, which permits investigation of how latent types relate to variables not included in the model. Finally, I provide results from a model fit exercise.

According to a regression of the probability of a CD4-count increase (versus a decrease) on treatment choices (see Table 7), HAART is the most effective treatment, especially for agents with low CD4.[43] Mono-therapy and combo-therapy, though less effective than HAART, are more effective than no treatment. Next, predicted values from this regression are used to explain period-$t + 1$ CD4 count. Specifically, Table 8 contains results from a logistic regression of dichotomous CD4 (high verus low). Results indicate that high CD4 at t and a

---

[43]Tables contain results from multiple specifications. Parameter estimates from the right-most (final) specification in each table correspond to the parameter values used in solving the dynamic programming problem.

CD4 increase independently predict high period-$t+1$ CD4 count.[44] The former result reflects persistence in underlying immune system health. Next, an absence of ailments is associated with higher CD4, which reflects that agents in better health are less likely to suffer ailments, i.e., symptoms (see Table 9 and recall that $F_{it} = 1$ indicates the absence of ailments during period t). Treatments also cause ailments via side effects and it generally holds that more effective treatments like HAART imply the harshest side effects. High CD4 count drastically reduces death probability. Since higher age signals good health among HIV+ subjects, I interact age with high CD4. The positive coefficient indicates that HIV+ subjects with high CD4 counts face higher death probability as they age (see Table 10).

Income is modeled to be a function of: high CD4, experience, experience-squared, age, a time-trend and current employment fully interacted with these variables. Income increases with human capital (as measured by experience), but at a decreasing rate. Age independently predicts a lower wage and good health is associated with higher income. The positive relationship between experience and income for non-full-time workers is consistent with increased non-wage income (e.g: disability payments) given a longer work history. For full-time workers, and with the exception of health, these effects are more pronounced. The effect of health on income is weaker for employed workers, which likely reflects that AIDS counts as a disability so that unemployment benefits are high for sick agents (see Table 11).

Health insurance (public, private or no insurance, the latter being the base category) is modeled as a function of CD4 count, age, labor supply and experience. Low CD4 and higher age predict higher probability of public insurance, which may again reflect that AIDS is considered a disability and that medicare eligibility is age-dependent. Also, full-time employment predicts private insurance provision, but predicts a lower probability of public insurance (see Table 12). Finally, treatment costs are a function of treatment choice, health status and insurance. Estimates indicate that HAART is more expensive than other treatments; costs increase over time; healthier subjects spend less on their medications and higher income and private insurance are associated with higher treatment costs (see Table 13).

Preference parameter estimates (found in Table 14) reveal that both Type I agents (constituting 52% of the population) and Type II agents (48% of the population) experience a utility cost of ailments, but an insignificant utility cost of working while not suffering ailments.[45] The key difference between the two types lies in the utility cost of working with

---

[44]This two-step method of obtaining transition probabilities among health states is meant to capture two salient features of treatment technology without increasing the size of the state space, namely, treatment effectiveness at increasing CD4 count and whether this increase brings agents to non-AIDS immune system health.

[45]Standard errors are computed using numerical derivatives to form the outer product measure of the covariance matrix. See Wooldridge (2002) pp. 395-396.

ailments. For Type I agents, this cost is near zero; for Type II agents, this cost is large. This difference has far-reaching consequences for agent behavior in counterfactual treatment and labor market environments: given their preferences, Type II agents can essentially attenuate the utility cost of ailments by choosing not to work.[46] The marginal utility of consumption increases with ailments, meaning that these agents consume goods that they value more when they are sicker. Type II agents experience a slight decrease in the marginal utility of consumption when they experience ailments, which means that they tend to consume goods that are complements to good health.[47] Finally, switching costs vary by ailment status: it is costly to begin treatment and to switch treatment. For agents with ailments, beginning treatment is costly, but ending treatment carries a benefit after controlling for the impact of this choice on other components of agents' utility, including health.

To gain further insight into the labor market heterogeneity captured by modeling two latent types, I compute 'posterior' type probabilities for each individual in the sample.[48] This permits an analysis of the correlation between unobserved latent type and observable labor market factors not included in the state space. Results are presented in Table 17. The general message is that latent Type II corresponds to lower education attainment and to occupations that are less flexible and more strenuous, which may explain why they exhibit a disutility of working with ailments. For example, individuals with a college degree are Type I with probability 0.56 and Type II with probability 0.47; for high school graduates corresponding probabilities are 0.41 and 0.54. Regarding broad occupation categories: individuals with professional specialties (e.g.: lawyers, doctors or teachers) are marginally more likely to be Type I than those in other occupations (54% versus 51%).[49] Those with a professional craft are also like to be Type I than those who are not (74% versus 52%). Waiters and individuals working in extractive industries (e.g.: mining) or transport are less likely to be Type I.

Results from a model fit exercise are found in Tables 15 and 16 for labor supply and treatment choices, respectively. Taking current states as given, agent choices are simulated

---

and then compared with state-dependent choices found in the data. Notice in Table 15 that the model successfully matches dynamics found in the data. According to Table 16, the model seems to confuse mono-therapy with combo-therapy. This is not unexpected since their profiles of characteristics are similar.

5   The Value of HAART

This section converts the value of HAART into a measure of willingness-to-accept-payment in dollars (henceforth: WTAP) and goes on to discuss what portion of this value can be attributed to labor market factors. Finally, I examine how a dynamically optimal treatment policy exhibits cycles. Specifically, sicker agents choose effective treatments despite harsh side effects and switch to less effective drugs with fewer side effects once their health improves. I refer to this phenomenon as Optimal Treatment Cycling or OTC.[50]

5.1   Willingness-To-Accept-Payment for HAART

To obtain WTAP for HAART, I compare computed value function values in the first post-HAART period to analogous values under the counterfactual scenario in which HAART is never introduced. Rather, a counterfactual treatment technology is introduced with the same effectiveness and side effects profile as combo-therapy. Next, I compute what per-period payment (similar to an annuity) under the counterfactual scenario is required to make agents indifferent to HAART introduction. Finally, I compute the present discounted value of this annuity using expected years of life, which is also simulated with estimated model parameters. Results are presented in Figure 5, which graphs the present value of future per-period payments for Type I agents of different ages and levels of human capital.[51] Two key findings emerge. First, HAART is very valuable, worth over $340,000 for a 38-year-old with 15 years of experience. Second, there is striking heterogeneity in the value of HAART. Older agents value HAART less since their life horizon is shorter, implying fewer years during which they benefit from HAART. This effect is compounded for younger agents since health gains made earlier in life persist over time. Finally, agents with higher human capital value HAART more since each life-year gained entails higher consumption. For example, a 45-year-old with

---

[50]Valuations of both factual and counterfactual treatment introductions capture present discounted value of utility in all future periods, i.e., gains in years-of-life weighted by consumption over time. Transfers between periods should not drastically change total value. Therefore, modeling savings behavior is not expected to appreciably affect results.

[51]Type II agents value HAART slightly less since they can essentially attenuate the utility costs of ailments by not working.

high human capital values HAART at over \$300,000, whereas a lower human capital agent values it at about \$50,000.

5.2 Decomposing the Value of HAART

This valuation exercise points to another question about the value of medical innovation, namely, "How much of the value of medical innovation can be attributed to labor market factors?". One can address this question on both an individual and a societal level. The latter perspective will be addressed in Section 7. The focus of this section is to decompose the value created by HAART for an HIV+ individual. The full impact of HAART includes its effect on: longevity, utility and labor market factors, the latter component working through the utility cost of work and productivity-induced wage changes.

I compare the HAART-induced value increases under two scenarios: where health can affect both longevity and utility via level effects and the marginal utility of consumption and where health can also affect labor disutility and productivity. I then compare value gains under both scenarios, the latter indicating how much additional value is created via the labor market. The principle finding is that up to 20% of of the value of HAART (again depending upon age and human capital) can be attributed to labor market factors. Figures 6 and 7 illustrate this point by graphing value decompositions for Type I agents of different age and experience profiles. I normalize the value of HAART to 1 for a 30-year-old with 5 years of experience. For young agents and agents with higher levels of human capital, up to about 22% of the value of HAART works through the labor market. For a 30-year-old with low human capital, the impact of HAART working purely through labor market factors is worth 20% of \$250,000, or about \$50,000. For older agents or agents with less work experience, this proportion is much smaller. This result suggests that the key labor market mechanism through which medical innovation creates value is increased productivity and higher wage. This benefits younger agents (who face a longer work history) and more experienced, higher-income agents who are more likely to work.

For Type II agents (not shown), the decomposition exercise shows that only a small portion of the value of HAART is attributable to the labor market. This result might seem surprising since Type II agents exhibit a utility of cost of work that is sensitive to their ailment status. However, recall that HAART side effects are harsh. Therefore, when Type II agents use HAART, they face ailments due to side effects rather than symptoms. Another key point emerges from this decomposition, namely, that the value of medical innovation attributable to labor market factors is inherently self-limiting given estimated model parameters. A treatment cannot reduce the utility cost of labor without improving a representative

patient's estimated utility parameters. This is most clear for Type II agents since an ailment reduction makes them better off regardless of their employment decisions, but increases utility only if an agent decides to work.

5.3   Optimal Treatment Cycling

When no available treatment dominates along both dimensions of quality (efficacy and side effects), agents optimally choose to cycle among available treatments. In other words, an optimal treatment path is a non-stationary closed loop driven by three factors: (1) persistence in health; (2) a non-convexity in discrete treatment choices; and (3) the disutility of ailments induced by side effects. Details of cycling behavior indicate that agents with AIDS-level CD4 count are highly likely to switch to the more effective treatment (in this case HAART), akin to a phase of investment in health 'stock'. While on HAART, agents face a higher probability of health improvements. Once their health improves, agents switch back to less effective treatment with fewer side effects (in this case, no treatment). During this phase, agents essentially exploit previous investments in their health stock, trading a higher probability of diminished future health for several periods with fewer side effects.

As will become evident in the following section, optimal treatment cycling is the key mechanism through which counterfactual environments affect agent choices and outcomes. Consider figure 8, which shows the anatomy of optimal treatment cycling for simulated Type I agents in an environment where available treatments correspond to options in the factual post-HAART world: no treatment, mono-treatment, combo-treatment or HAART. Agents respond to counterfactual environments primarily through shifts in the frequency of the first step of the cycle, i.e., how often they switch to less effective treatments when in relatively good health (11% in this case).[52] Generally, once agents are off the most effective treatment, most remain there until their health diminishes, at which point most resume effective treatment.

Treatment cycling is often considered to be a form of suboptimal non-compliance that should be curbed (Sabate, 2003). Switching off of treatment is sometimes referred to as a 'drug holiday' and some medical literature points to the dangers (both social and private) of engaging in such behavior (Meredith, 1996). In contrast, I find that a cyclical treatment pattern can be the result of optimal forward-looking behavior and refer to the phenomenon as Optimal Treatment Cycling (OTC). Recent medical research on long-term, chronic illness suggests adapting current treatments to patient responses to previous treatment (Murphy, 2005).[53] OTC is similar in that current decisions reflect previous treatment outcomes, though

---

[52]For Type II agents, this figure is largely similar, though they switch off treatment with lower frequency (7-8%) since they can attenuate utility loss from ailments by not working.

[53]Specifically, this line of research suggests designing medical trials involving multiple randomizations to

it is driven by patient decision-making. OTC is also consistent with findings that medical doctors, despite advocating highly effective treatments for their patients, often opt for less effective drugs with fewer side effects when faced with similar medical conditions (Ubel et al., 2011).

6 Policy Experiments

The framework developed in this paper - in particular, the characteristics approach to modeling treatments - is well-suited to evaluate health and labor outcomes in counterfactual treatment and labor market environments. A key component of the framework is that treatment innovations are modeled along two-dimensions: efficacy and side effects. Policy simulations examine how these dimensions affect agent decision-making, health and labor market characteristics. I present results on four exercises. First, I compute value (WTAP) for a variety of counterfactual treatment innovations. Then, I study agent decisions and outcomes under counterfactual: treatment technology, labor market and treatment cost environments. I focus on Type II agents: since they exhibit a large utility loss from working with ailments, their health and labor decisions are more tightly linked.

6.1 The Value of Treatment Innovation

Suppose that after HAART is introduced, patients are faced with an improvement on HAART along one or both dimensions of drug-quality. One possibility is a version of HAART without side effects. Computing WTAP as in Section 3.2, I compute agent valuations of such an innovation, presenting results for low-human capital patients of different ages in Figure 9.[54] Note that black bars in the figure depict the value of HAART introduction. The white bars show the value of HAART plus the value of an innovation supposing HAART already exists. In other words, given HAART, the value of HAART without side effects is the difference between the black and white bars. A version of HAART without side effects creates enormous value: about $375,000 for a 30-year-old. As in previous results, older agents value the innovation less since they have fewer periods to enjoy it.[55] This result is especially striking since the innovation entails no improvement on underlying health or longevity. In this sense, a version of HAART without side effects could be seen as a "me-too" treatment since, by definition, it is therapeutically equivalent to an existing treatment. Contrary to arguments that such

---

better formulate decision rules for adaptive treatments.

[54]I compute WTAP for Type I agents. Type II agents value treatments similarly.

[55]High-experience agents value the innovation higher than low-experience agents according to results that are not shown.

innovations offer little benefit to consumers, I find that a treatment that is therapeutically equivalent to HAART, but entails fewer side-effects, is valuable.[56]

Suppose that instead of a reduction in side effects, HAART is improved along the efficacy dimension. In particular, low-CD4 agents who use HAART face a 32% probability of non-AIDS CD4 in the following period. Under the counterfactual improvement, this probability is doubled to 67%. Figure 10 depicts how agents value this innovation. As in the previous figure, the black bars show the value of HAART. For comparison, the grey bars show the value of the side effects innovation, given that HAART already exists. The white bars show the value of the efficacy innovation (again using HAART as the baseline technology) and this value is large: upwards of $1,000,000 for a 30-year-old. Agents would be expected to place high value on a life-improving and life-saving technology. What is more surprising is that optimal treatment cycling underlies some portion of this value. In general, switching onto milder treatments is risky since the full treatment cycle includes periods where CD4 count is low and death probability is high. If a highly effective version of HAART exists, however, agents anticipate fewer such periods. They respond by cycling more aggressively, i.e., by more frequently switching to milder treatments once their CD4 count is high. In other words, the value of an effective treatment includes the implied option value of optimally cycling off of it in periods of relatively good health.

Another key finding is that the value of counterfactual treatments depend on existing treatments. Suppose that the two aforementioned innovations (side effects and efficacy) occur simultaneously in separate treatments, so that two new drugs are introduced. Figure 11 depicts how agents value these innovations. Notice that, as compared to the efficacy innovation, the two innovations create little additional value. This finding is striking: a side effects innovation is valuable absent an efficacy innovation, but creates little value compared with a regime where HAART is highly effective. Again, the underlying mechanism is optimal treatment cycling: if a highly effective version of HAART exists, agents can simply cycle off of treatment altogether (avoiding all side effects), retaining the option value of resuming treatment once their health deteriorates. A drug without side effects adds little additional value in such a scenario.[57] Nonetheless, combining an efficacy and side effects innovation into a single drug does imply additional value since it permits agents to live without side effects, but to avoid risks associated with cycling. Such an innovation starts to approximate a cure and its value reflects this: about $1,400,000 for a 30-year-old (see Figure 12).

---

[56]See, for example, Angell (2000) on why "me-too" drug development should be curtailed.

[57]This does not necessarily imply that a private pharmaceutical firm would not profit from investing in marginal improvements on either dimension of drug quality since a high proportion of patients would presumably switch to the improved treatment despite the small implied value increase.

6.2 Treatment Innovation, Choices and Outcomes

To gain insight into how counterfactual medical innovations create value, the next simulation traces agent decisions, longevity, health and labor market characteristics from the time of HAART introduction until the end of the sample period under regimes distinguished by available treatment technologies.[58] I compare three of the treatment scenarios outlined above. The first is the baseline factual regime where HAART is introduced in 1996. In the second, a treatment identical to combo-therapy is introduced at the time of HAART introduction. This scenario mimics a continuation of the pre-HAART world in the sense that a new treatment becomes available, but does not improve upon existing technology. In the third scenario, two counterfactual treatments are simultaneously introduced: HAART with no side effects along with a highly effective version of HAART with HAART-level side effects. This final scenario illustrates behavior when innovations separately occur along two dimensions of treatment quality. Note that agents are modeled to make optimal choices given each counterfactual environment.[59]

Simulated death rates are found in Figure 13. A world without technology improvements sees only a slight decline in death rates. Death rates under the observed introduction of HAART are lower and death rates were HAART to be improved hover around 1%, i.e., survival is essentially that of HIV− subjects.[60] The probability of suffering ailments (Figure 14) exhibits a more complex pattern since ailments reflect both symptoms of disease and side effects of treatments. For several periods into the simulation, agents facing counterfactual treatment technology improvements are free of ailments *less* often than their counterparts facing existing HAART technology. This occurs since agents are more likely to choose a more effective version of HAART and therefore suffer side effects with higher probability. However, as time goes on, this dynamic reverses: ever more agents exhibit higher CD4 and so move towards less effective drugs with fewer side effects, in effect living off of their health stock.

This pattern of physical ailments is reflected in employment decisions over time, which are

---

[58]Policy experiments take the distribution of state variables at the time of HAART introduction as given, with the exception that all agents are modeled to have chosen "no treatment" in the previous period. This facilitates comparisons across counterfactual treatment policies.

[59]In what follows, results are presented for Type II agents whose utility cost of work is exacerbated by ailments. Results for Type I agents (available upon request) are generally similar in terms of treatment choice and health measures. The key difference is that Type I agents are much less likely to change their employment decisions in response to medical innovation since their utility cost of work remains essentially zero in spite of ailments.

[60]The fact that death rates decline in a world where HAART is not introduced might seem surprising given that the cohort is aging. This dynamic is found in the data. In particular, transition probabilities indicate that older agents with low CD4 count have a slightly higher survival probability, a finding that might reflect selection by age. In other words, low-CD4 agents who survive until they are older might be less susceptible to AIDS-related illness than those who die when they are younger.

depicted in Figure 15.[61] In general, Type II agents exhibit a fairly low probability of working full-time (between 25% and 35% versus 80% or more for Type I agents). This low probability arises, in part, due to the timing of labor supply decisions: agents choose whether or not to work for a full period before they know their ailment status. A high enough probability of suffering ailments coupled with a high disutility of labor while suffering ailments, means that Type II agents will often avoid employment. Further, Figure 15 illustrates that HAART brought increases in labor supply. Keeping in mind that working agents pay taxes, this result suggests that medical innovation creates value not just for the individual, but for society in general. In particular, in 2002 HAART implies an increase in full-time employment from 22% to 32%.

Next, it is unclear whether improvements in treatment technology compared to HAART will increase labor supply. Given the third scenario where two treatment improvements are introduced, a higher proportion of Type II agents use the version of HAART with greater effectiveness, which induces heavier side effects, thereby reducing their employment. However, as their health improves and these agents face a lower probability of symptoms-induced ailments, they increase their employment. Finally, this pattern of treatment and employment decisions is reflected in expected income (Figure 16). Although average income increases as treatments improve, human capital accumulation decreases. The latter results from composition effects: agents who do not work due to symptoms are more likely to die than agents who do not work due to side effects. Therefore, effective treatment with heavy side effects implies a greater number of living, but non-working agents.

6.3 A Decline in Non-Wage Income

The introduction of HAART occurred under very specific circumstances since HAART treats a condition that is legally considered a disability, giving patients access to disability payments. Therefore, income remains fairly high for agents who choose not to work.[62] The goal of the following experiment is to ascertain agent choices and outcomes in a counterfactual environment where non-wage income is lower. To highlight the role of cycling, available treatments are: no treatment, HAART and the two counterfactual improvements upon HAART

---

[61]Would modeling savings behavior change the result that Type II agents tend to work more under a medical technological improvement? Suppose the negative ailment-employment interaction partly captured a tendency to save while working. Then, even a slight improvement in ailment status should imply even larger gains to employment, in effect strengthening this result.

[62]Under the Americans with Disabilities Act, people living with HIV/AIDS qualify for social security disability payments. These payments cover both symptoms of AIDS and side effects of treatment. Moreover, limited benefits can continue even if agents return to work, reflecting the cyclical nature of chronic disease. For more information, see http://ssa.gov/pubs/10019.html. For payment calculations, also see: http://www.ssa.gov/policy/docs/statcomps/supplement/2011/.

discussed in the third scenario of the previous section. In the simulated environment, agents face reductions in non-wage income, operationalized via decreased parameters of the income process for agents not choosing full-time employment. In effect, non-wage income is reduced 20%, 40% and 60%.[63]

Facing lower non-wage income, Type II agents engage in more pronounced optimal treatment cycling. In particular, the probability that high-CD4-count agents switch off of highly effective treatment rises from 8% to 15%, 30% and 42% as non-wage income declines.[64] The mechanism is as follows: Type II agents, facing lower unemployment income, are compelled to work. However, to avoid the high utility cost of working with ailments, they quickly cycle onto treatments with fewer side effects when their CD4 count is high. More pronounced OTC means they are more likely to be free of ailments (see Figure 17). Labor market outcomes also improve. In particular, agents are much more likely to work: a 60% reduction in non-wage income increases employment from about 30% to about 83% (see Figure 18). Average income also increases (Figure 19). However, more pronounced optimal treatment cycling carries some risk. Figure 20 compares death rates under full non-wage income versus a 60% reduction and shows that this reduction ultimately increases death probability. This policy experiment highlights the long-term trade-offs between comfort, consumption and longevity. In particular, agents are willing to trade reduced longevity for increased consumption when the labor market environment compels them to.

6.4 Unsubsidized Treatment Costs

Recall that HIV+ agents pay on average about $500 per year for treatment. However, the actual costs, paid by insurance (both public and private) is much more. A year of HAART therapy costs about $12,000. Combo-therapy costs $8,000 and mono-therapy $6000. What would happen to agent choices and outcomes if they were compelled to pay these unsubsidized costs? The following scenario answers this question, simulating environments where agents would pay 33%, 66% or the full cost of treatment. Again, more pronounced optimal treatment cycling is the key mechanism through which changes in the environment affect patient choices. Specifically, facing high costs, agents are more likely to switch off of HAART once their health improves. Recall that this occurs with a frequency of about 7% in the factual environment, but increases to 32%, 58%, and 74% under the three counterfactual treatment costs. Agents

---

[63]As noted in Section 3, observed non-wage income is too high to solely reflect public disability payments. Therefore, the experiment does not simulate an environment where non-wage income is driven to zero. The simulated reductions are chosen to mimic an environment with reduced public payments.

[64]Type I agents' response to low non-wage income is to slightly increase already high levels of employment. They do not, however, appreciably shift their treatment cycling behavior since they do not experience a utility cost of working with ailments.

face lower average health, as seen in Figure 21, though survival probability remains largely unchanged. Agents do exhibit an improvement in their side effects status, shown in Figure 22, which encourages an increase in employment (Figure 23), resulting in higher income (Figure 24).[65]

These findings allude to an important aspect of medical decision-making, namely, the role of estimated switching costs. A simulation where switching costs are set to zero results in more pronounced treatment cycling, in the sense that patients are more likely to switch off harsh treatment when their health improves. Suppose we accept that, beyond mere persistence or transaction costs faced by agents switching treatments, switching costs also capture medical protocols or doctors' orders. Switching costs may also reflect the societal benefit of remaining on effective treatment. For example, there is some evidence that treatment cycling lowers drug effectiveness. Also, recent literature points out that consistent use of antiretrovirals lowers the infectiousness of HIV+ individuals.[66] Then, doctors function as imperfect agents, essentially compelling agents to reduce their frequency of switching to less effective, lower side effects treatment when in better health. Therefore, counterfactual unsubsidized costs reflect an environment where the cost of disobeying doctors' orders simply outweighs the cost of following them. Moreover, more aggressive cycling does not lead to lower survival probability, suggesting that disobeying doctors' orders may not harm a patient in the long run, even though it will lead to better labor market outcomes.

7   Conclusion

This project develops a general framework to value medical innovation, taking explicit account of the link between health, human capital and the labor market. The framework is applied to a chronic condition: HIV. In particular, I study the treatment choices and employment decisions of HIV+ men. I find that a medical innovation known as HAART was worth between $25,000 and $350,000, depending on age and human capital. I also find that agents, when faced with treatments differing in their efficacy and side effects profiles, optimally choose to cycle among available options. Exploiting the characteristics approach to modeling treatments, I also evaluate counterfactual treatment innovations and find that agents place high value on a version of HAART without side effects.

I also investigate how health and labor interact and show that OTC underlies dynamic

---

[65]How would modeling savings change this result? A potential pattern would be that agents save while working (and not purchasing treatment) and then use savings to consume and purchase harsh treatments while not working. Under this scenario, the link between not-working and harsh treatments would likely be stronger. However, we would still expect less treatment use, resulting in higher employment.

[66]For an economic study on this phenomenon from the perspective of HIV− subjects, see Chan et al. (2011).

decision-making. Specifically, if non-wage income is so low that agents are compelled to work, they switch to less effective drugs with higher probability when in relatively good health. Although this behavior improves their labor market outcomes, it reduces both their underlying average health and survival probability. In another simulation, I show that agents facing unsubsidized drug costs quickly cycle off treatment once their health improves. This behavior, though it reduces health measures, does not appreciably decrease survival probability. However, it does reduce ailments induced by side effects. As a result, agents are more likely to work, earning higher income and accelerating the accumulation of human capital.

A decomposition exercise shows that, from the perspective of the individual, up to 20% of the value of HAART is created through labor market factors like increased employment and higher income. However, policy experiments show significant behavior changes under counterfactual environments and some of these changes would entail important social benefits. For example, the HAART-induced 30% increase in employment among HIV+ Type II agents presumably entailed an increase in income tax revenues. In other words, from a societal perspective, the labor market portion of the value of HAART may have been much higher than from the individual perspective. On the other hand, although OTC is optimal from the individual perspective, there are important externalities that arise from implied lower average CD4 counts, including higher infectiousness (due to higher HIV viral load). An extension of the current model could address this possibility by including uninfected agents' decision-making when faced with the risk of contracting HIV.

Given data on employment and treatment decisions, applying the framework developed in this paper to other illnesses is straightforward, though some caveats apply. For example, a long-term illness such as diabetes might prove problematic unless agents are allowed to form and update expectations of an innovation. Further research could also extend the characteristics approach to other dimensions of treatment quality. For example, in the context of HIV, convenience is of ever greater importance. Early versions of HAART required that multiple medications be taken in specific sequences. More recently, HAART often amounts to taking a single tablet each day. Convenience, perhaps less important in the face of life-threatening illness, becomes more salient once treatments are effective, side effects are manageable and patients demand further improvements to their quality-of-life.

References

AGUIRREGABIRIA, V. AND P. MIRA, "Dynamic Discrete Choice Structural Models: A Survey," *Journal of Econometrics* 156 (2010), 38–67.

ANGELL, M., "The Pharmaceutical Industry To Whom Is It Accountable?," *New England Journal of Medicine* 342 (2000), 1902–1904.

BECKER, G., "Human Capital, Effort, and the Sexual Division of Labor," *Journal of Labor Economics* 3 (1985), 33–58.

———, "Health as Human Capital: Synthesis and Extensions," *Oxford Economic Papers* 59 (2007), 379–410.

BLALOCK, A., J. MCDANIEL AND E. FARBER, "Effect of Employment on Quality of Life and Psychological Functioning in Patients with HIV/AIDS," *Psychosomatics* 43 (2002), 400.

BRESNAHAN, M., T. TRAJTENBERG AND S. STERN, "Market Segmentation and the Sources of Rents from Innovation: Personal Computers in the Late 1980s," *RAND Journal of Economics* 28 (1997), 17–44.

CAWLEY, J., "The Impact of Obesity on Wages," *Journal of Human Resources* 39 (2004), 451.

CHAN, T., "Estimating a Continuous Hedonic-Choice Model with an Application to Demand for Soft Drinks," *RAND Journal of Economics* 37 (2006), 466–482.

CHAN, T. AND B. HAMILTON, "Learning, Private Information, and the Economic Evaluation of Randomized Experiments," *Journal of Political Economy* 114 (2006), 997–1040.

CHAN, T., B. HAMILTON AND N. PAPAGEORGE, "AIDS, Antivirals and Risky Behavior," mimeo (October 2011).

CHINTAGUNTA, P., R. JIANG AND G. JIN, "Information, Learning, and Drug Diffusion: The Case of Cox-2 Inhibitors," *Quantitative Marketing and Economics* 7 (2009), 399–443.

CONTI, G., J. HECKMAN AND S. URZUA, "The Education-Health Gradient," *American Economic Review* 100 (2010), 234–38.

CRAWFORD, G. AND M. SHUM, "Uncertainty and Learning in Pharmaceutical Demand," *Econometrica* (2005), 1137–1173.

CURRIE, J., "Healthy, Wealthy, and Wise: Socioeconomic Status, Poor Health in Childhood, and Human Capital Development," *Journal of Economic Literature* 47 (2009), 87–122.

CURRIE, J. AND B. MADRIAN, *Health, Health Insurance and the Labor Market*, volume 3 of *Handbook of Labour Economics*, chapter 50 (Elsevier, 1999), 3309–3416.

DETELS, R., P. TARWATER, J. PHAIR, J. MARGOLICK, S. RIDDLER AND A. MUÑOZ, "Effectiveness of Potent Antiretroviral Therapies on the Incidence of Opportunistic Infections Before and after AIDS Diagnosis," *AIDS* 15 (2001), 347.

DUGGAN, M. AND W. EVANS, "Estimating the Impact of Medical Innovation: A Case Study of HIV Antiretroviral Treatments," Technical Report, National Bureau of Economic Research, 2005.

ECKSTEIN, Z. AND K. WOLPIN, "Dynamic Labour Force Participation of Married Women and Endogenous Work Experience," *The Review of Economic Studies* 56 (1989), 375–390.

EVANS, W. AND W. VISCUSI, "Estimation of State-Dependent Utility Functions Using Survey Data," *The Review of Economics and Statistics* 73 (1991), 94–104.

FERNANDEZ, J., "An Empirical Model of Learning under Ambiguity: The Case of Clinical Trials," *mimeo* (2008).

FINKELSTEIN, A., E. LUTTMER AND M. NOTOWIDIGDO, "What Good is Wealth without Health? The Effect of Health on the Marginal Utility of Consumption," mimeo (2008).

FRENCH, E. AND J. SONG, "The Effect of Disability Insurance Receipt on Labor Supply," *Federal Reserve Bank of Chicago Working Paper* (2010).

GARTHWAITE, C., "Economic Benefits of Medical Innovations," mimeo (2010).

GOLDMAN, D. AND Y. BAO, "Effective HIV Treatment and the Employment of HIV+ Adults," *Health Services Research* 39 (2004), 1691–1712.

GORMAN, W., "A Possible Procedure for Analysing Quality Differentials in the Egg Market," *The Review of Economic Studies* 47 (1980), 843–856.

GROSSMAN, M., "On the Concept of Health Capital and the Demand for Health," *The Journal of Political Economy* 80 (1972), 223–255.

HECKMAN, J. AND F. CUNHA, "The Technology of Skill Formation," *American Economic Review* 97 (2007), 31–47.

HOTZ, V. AND R. MILLER, "Conditional Choice Probabilities and the Estimation of Dynamic Models," *The Review of Economic Studies* 60 (1993), 497–529.

IMAI, S. AND M. KEANE, "Intertemporal Labor Supply and Human Capital Accumulation," *International Economic Review* 45 (2004), 601–641.

JUDD, K., *Numerical Methods in Economics* (The MIT press, 1998).

KAHN, M., "Health and Labor Market Performance: The Case of Diabetes," *Journal of Labor Economics* 16 (1998), 878–899.

LANCASTER, K., "A New Approach to Consumer Theory," *The Journal of Political Economy* 74 (1966), 132–157.

LEVINE, P., T. GUSTAFSON AND A. VELENCHIK, "More Bad News for Smokers-The Effects of Cigarette Smoking on Wages," *Industrial and Labor Relations Review* 50 (April 1997), 493–509.

LIPSCOMB, J., M. DRUMMOND, D. FRYBACK, M. GOLD AND D. REVICKI, "Retaining, and Enhancing, the QALY," *Value in Health* 12 (2009), S18–S26.

MEREDITH, P., "Therapeutic Implications of Drug Holidays," *European Heart Journal* 17 (1996), 21.

MINCER, J. AND S. POLACHEK, "Family Investments in Human Capital: Earnings of Women," *The Journal of Political Economy* 82 (1974), 76–108.

MURPHY, K. AND R. TOPEL, eds., *Measuring the Gains from Medical Research: An Economic Approach* (University of Chicago Press, 2003).

MURPHY, K. AND R. TOPEL, "The Value of Health and Longevity," *The Journal of Political Economy* 114 (2006), 871–904.

MURPHY, S., "An Experimental Design for the Development of Adaptive Treatment Strategies," *Statistics in Medicine* 24 (2005), 1455—1481.

OSTROW, D., A. MONJAN, J. JOSEPH, M. VANRADEN, R. FOX, L. KINGSLEY, J. DUDLEY AND J. PHAIR, "HIV-Related Symptoms and Psychological Functioning in a Cohort of Homosexual Men," *American Journal of Psychiatry* 146 (1989), 737.

PETRIN, A., "Quantifying the Benefits of New Products: The Case of the Minivan," *Journal of Political Economy* 110 (2004), 705–729.

PHILIPSON, T. AND R. POSNER, *Private Choices and Public Health: The AIDS Epidemic in an Economic Perspective* (Cambridge: Harvard University Press, 1993).

ROSEN, S., "Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition," *The Journal of Political Economy* 82 (1974), 34–55.

RUST, J., "Optimal Replacement of GMC Bus Engines: An Empirical Model of Harold Zurcher," *Econometrica* 55 (1987), 999–1033.

SABATE, E., "Adherence to Long-Term Therapies: Evidence for Action," Technical Report, World Health Organization, Available: www.who.int/chp/knowledge/publications/adherence_introduction.pdf. Accessed 4 Oct 2011., 2003.

SINGLETON, P., "The Role of Health in the Demand for Disability Benefits: Evidence from HIV Antiretroviral Treatments," mimeo (June 2011).

SLOAN, F., W. KIP VISCUSI, H. CHESSON, C. CONOVER AND K. WHETTEN-GOLDSTEIN, "Alternative Approaches to Valuing Intangible Health Losses: The Evidence for Multiple Sclerosis1," *Journal of Health Economics* 17 (1998), 475–497.

SMITH, D., K. LANGA, M. KABETO AND P. UBEL, "Health, Wealth, and Happiness: Financial Resources Buffer Subjective Well-Being after the Onset of a Disability," *Psychological Science* 16 (2005), 663–666.

SMITH, J., "Healthy Bodies and Thick Wallets: The Dual Relation between Health and Economic Status," *The Journal of Economic Perspectives* 13 (1999), 145–166.

STIGLER, G., "The Cost of Subsistence," *Journal of Farm Economics* 27 (1945), 303–314.

TRAJTENBERG, M., "The Welfare Analysis of Product Innovations with an Application to CT Scanners," *Journal of Political Economy* 97 (1989), 444–479.

UBEL, P., A. ANGOTT AND B. ZIKMUND-FISHER, "Physicians Recommend Different Treatments for Patients Than They Would Choose for Themselves," *Archives of Internal Medicine* 171 (2011), 630–634.

VISCUSI, W. AND W. EVANS, "Utility Functions that Depend on Health Status: Estimates and Economic Implications," *The American Economic Review* 80 (1990), 353–374.

WOOLDRIDGE, J., *Econometric Analysis of Cross Section and Panel Data* (MIT Press, 2002).

8 Tables and Figures

		TIME $(t+1)$			
		None	Mono	Combo	HAART
		PRE-HAART			
	None	0.87	0.10	0.04	.
TIME (t)	Mono	0.09	0.68	0.23	.
	Combo	0.06	0.24	0.70	.
		POST-HAART			
	None	0.82	0.02	0.05	0.12
TIME (t)	Mono	0.03	0.63	0.09	0.25
	Combo	0.03	0.03	0.66	0.29
	HAART	0.02	0.04	0.01	0.93

Table 1: TRANSITION MATRIX for HIV treatment choices.

		TIME $(t+1)$	
		Not Full Time	Full Time
		PRE-HAART	
TIME (t)	Not full time	0.86	0.14
	Full Time	0.10	0.90
		POST-HAART	
TIME (t)	Not full time	0.91	0.09
	Full Time	0.06	0.94

Table 2: TRANSITION MATRIX for employment decisions.

	FULL SAMPLE	PRE-HAART	POST-HAART	FULL-TIME	NOT FULL-TIME
Age at Baseline (Years)	32.71	33.16	32.30	32.40	33.32
CD4 count	448.76	392.54	500.37	490.57	366.12
High CD4	0.73	0.65	0.81	0.81	0.59
No Ailments	0.59	0.59	0.59	0.68	0.41
Death Probability	0.04	0.07	0.02	0.02	0.09
Treatment: None (t+1)	0.30	0.45	0.18	0.33	0.25
Treatment: Mono (t+1)	0.20	0.32	0.09	0.20	0.19
Treatment: Combo (t+1)	0.17	0.23	0.12	0.17	0.17
Treatment: HAART (t+1)	0.33	.	0.62	0.30	0.39
Income ($(2003)/year)	32,348.08	28,040.50	36,301.32	37,168.07	22,818.58
Out-of-Pocket Cost (2003$/year)	455.02	309	580.68	430.82	506.66
Experience (Years)	12.56	11.78	13.28	12.89	11.91
No Insurance	0.05	0.07	0.03	0.04	0.08
Private Insurance	0.81	0.84	0.79	0.92	0.59
Public Insurance	0.14	0.09	0.18	0.04	0.33
Not full time (t+1)	0.34	0.32	0.36	0.08	0.89
Full time (t+1)	0.66	0.68	0.64	0.92	0.11
N [time - t variables]	8,300	3,972	4,328	5,512	2,788
N [time - $(t+1)$ variable]	7,954	3,694	4,260	5,335	2,500

Table 3: SUMMARY STATISTICS for the full analysis sample, for the pre- and post-HAART eras and by current employment status at time - t. The number of observations (N) for time - $(t+1)$ variables is smaller than the number of time - t observations reflecting agents who die between t and $(t+1)$. Summary statistics are for 6-month periods (corresponding to the frequency of MACS interviews) and units are (%/100) unless otherwise indicated.

	EMPLOYMENT CHOICE ($t+1$)				
	(1)	(2)	(3)	(4)	(5)
				Pre-HAART	Post-HAART
Full time (t)	4.26***	3.97***	3.97***	3.41***	4.47***
Experience	0.11***	0.17***	0.17***	0.2***	0.14***
Exper. Squared	-0.001***	-0.0005	-0.0005	-0.0005	-0.0002
Age	-0.11	-0.15	-0.15	-0.29	0.13
Age^2	0.0000414	-0.001	-0.001	-0.0007	-0.004*
High CD4	.	0.98***	0.95***	0.94***	0.83***
No Symptoms	.	0.56***	0.55***	0.61***	0.47***
Haart Available	.	1.02***	1.11***	.	.
Treatment: Mono (t+1)	.	.	-0.008	0.26	-0.24
Treatment: Combo (t+1)	.	.	-0.27**	0.2	-0.74**
Treatment: HAART (t+1)	.	.	-0.18	.	-0.64**
Treatment: Mono (t)	.	.	.	-0.23	-0.06
Treatment: Combo (t)	.	.	.	-0.41*	0.22
Treatment: HAART (t)	.	.	.	.	0.36
Observations	7954	7954	7954	3694	4260

Table 4: REDUCED-FORM LOGISTIC REGRESSION of dichotomous employment choices.

	TREATMENT CHOICE ($t+1$)			
	(1)	(2)	(3)	(4)
Mono:Treatment: Mono (t)	4.12***	4.12***	4.11***	4.08***
Mono:Treatment: Combo (t)	3.47***	3.47***	3.45***	3.49***
Mono:High CD4	-0.52***	-0.51***	-0.53***	-0.59***
Mono:No Symptoms	.	-0.04	-0.05	-0.1
Mono:Age	.	-0.004	-0.006	-0.19***
Mono:Private Insurance	.	.	0.5**	0.42*
Mono:Public Insurance	.	.	0.51*	0.68**
Mono:Income/10000	.	.	0.03	-0.02
Mono:Full time (t)	.	.	.	0.19
Mono:Experience	.	.	.	0.1***
Combo:Treatment: Mono (t)	4.03***	4.05***	4.06***	4.04***
Combo:Treatment: Combo (t)	5.52***	5.56***	5.54***	5.58***
Combo:High CD4	-0.66***	-0.71***	-0.76***	-0.82***
Combo:No Symptoms	.	0.2	0.16	0.12
Combo:Age	.	-0.005	-0.007	-0.07
Combo:Private Insurance	.	.	0.08	0.02
Combo:Public Insurance	.	.	-0.02	0.07
Combo:Income/10000	.	.	0.09**	0.05
Combo:Full time (t)	.	.	.	0.28
Combo:Experience	.	.	.	0.04
Observations	3694	3694	3694	3694

Table 5: REDUCED-FORM LOGISTIC REGRESSION of pre-HAART treatment choices.

	Treatment Choice $(t+1)$			
	(1)	(2)	(3)	(4)
Mono:Treatment: Mono (t)	4.94***	4.93***	4.92***	4.94***
Mono:Treatment: Combo (t)	1.83***	1.88***	1.91***	1.98***
Mono:High CD4	-0.36*	-0.3	-0.21	-0.22
Mono:No Symptoms	.	-0.56***	-0.49***	-0.44***
Mono:Age	.	0.04***	0.04***	0.11***
Mono:Private Insurance	.	.	1.04*	1.13**
Mono:Public Insurance	.	.	2.17***	2.00***
Mono:Income/10000	.	.	0.05	0.11**
Mono:Full time (t)	.	.	.	-0.31
Mono:Experience	.	.	.	-0.05**
Combo:Treatment: Mono (t)	3.37***	3.37***	3.38***	3.39***
Combo:Treatment: Combo (t)	5.43***	5.44***	5.45***	5.45***
Combo:High CD4	0.07	0.07	0.24	0.27
Combo:No Symptoms	.	-0.16	-0.07	-0.06
Combo:Age	.	0.004	0.001	0.004
Combo:Private Insurance	.	.	0.62	0.65
Combo:Public Insurance	.	.	1.54***	1.47***
Combo:Income/10000	.	.	-0.09*	-0.06
Combo:Full time (t)	.	.	.	-0.24
Combo:Experience	.	.	.	-0.002
HAART:Treatment: Mono (t)	0.86***	0.85***	0.85***	0.86***
HAART:Treatment: Combo (t)	1.09***	1.11***	1.14***	1.17***
HAART:High CD4	-0.5***	-0.43***	-0.32**	-0.3**
HAART:No Symptoms	.	-0.49***	-0.4***	-0.38***
HAART:Age	.	0.01**	0.01	0.04**
HAART:Private Insurance	.	.	0.95***	1.01***
HAART:Public Insurance	.	.	2.09***	1.99***
HAART:Income/10000	.	.	0.02	0.07**
HAART:Full time (t)	.	.	.	-0.29**
HAART:Experience	.	.	.	-0.02*
Observations	4260	4260	4260	4260

Table 6: Reduced-form logistic regression of post-HAART treatment choices.

	CD4 Count Increase		
	(1)	(2)	(3)
Treatment: Mono (t+1)	0.24***	0.22***	.
Treatment: Combo (t+1)	0.38***	0.36***	.
Treatment: HAART (t+1)	0.91***	0.9***	.
High CD4	.	-0.1*	0.07
Low CD4 (t) × mono	.	.	0.33**
Low CD4 (t) × combo	.	.	0.25*
Low CD4 (t) × HAART	.	.	1.41***
High CD4 (t) × mono	.	.	0.24***
High CD4 (t) × combo	.	.	0.46***
High CD4 (t) × HAART	.	.	0.59***
Time trend	.	.	0.02***
Observations	7954	7954	7954

Table 7: TRANSITIONS: CD4 increase versus decrease

	High CD4 Count		
	(1)	(2)	(3)
High CD4	4.33***	4.59***	5.51***
CD4 Increase (Predicted)	.	4.10***	.
CD4 Increase × Low CD4	.	.	4.69***
CD4 Increase × High CD4	.	.	2.70***
Observations	7954	7954	7954

Table 8: TRANSITIONS: High CD4 at $(t+1)$

	No Ailments			
	(1)	(2)	(3)	(4)
High CD4	0.99***	0.94***	1.41***	1.41***
Treatment: Mono (t+1)	.	-0.44***	.	.
Treatment: Combo (t+1)	.	-0.48***	.	.
Treatment: HAART (t+1)	.	-0.46***	.	.
Low CD4 (t) × mono	.	.	0.06	0.07
Low CD4 (t) × combo	.	.	-0.11	-0.1
Low CD4 (t) × HAART	.	.	0.12	0.07
High CD4 (t) × mono	.	.	-0.58***	-0.58***
High CD4 (t) × combo	.	.	-0.56***	-0.57***
High CD4 (t) × HAART	.	.	-0.59***	-0.66***
Time trend	.	.	.	0.006
Observations	7954	7954	7954	7954

Table 9: TRANSITIONS: Absence of ailments (symptoms or side effects).

	Death Probability		
	(1)	(2)	(3)
High CD4	.	-3.17***	-6.16***
Age	0.009	-0.001	-0.01
Age $\times$ High CD4	.	.	0.07***
CD4 count	-0.01***	.	.
CD4-squared	3.06e-06***	.	.
Observations	8300	8300	8300

Table 10: TRANSITIONS: Death probability.

	Income		
	(1)	(2)	(3)
High CD4	2606.55***	1237.73***	1794.46***
Experience	1476.84***	884.60***	838.61***
Exper. Squared	-6.28***	-4.65***	-5.99***
Age	-2050.10***	-1066.34***	-802.67***
Time since combo	705.46***	443.96***	389.73***
Full time (t+1)	.	6861.13***	30923.38***
Full time (t+1) $\times$ High CD4	.	.	-1128.23**
Experience	.	.	205.08**
Exper. Squared	.	.	2.52**
Full time (t+1) $\times$ Age	.	.	-718.44***
Full time (t+1) $\times$ Time trend	.	.	156.45***
Observations	7954	7954	7954

Table 11: OUTCOMES: Semester income.

	Insurance Choice		
	(1)	(2)	(3)
Private:Full time (t+1)	1.22***	1.27***	1.16***
Private:High CD4	.	-0.28**	-0.5***
Private:Experience	.	.	0.16***
Private:Exper. Squared	.	.	-0.002***
Private:Age	.	.	-0.3**
Private:Age^2	.	.	0.002*
Private:Time trend	.	.	0.07***
Private:Time since HAART	.	.	0.02
Public:Full time (t+1)	-1.30***	-1.17***	-0.88***
Public:High CD4	.	-0.69***	-1.05***
Public:Experience	.	.	-0.03
Public:Exper. Squared	.	.	-0.0000149
Public:Age	.	.	0.17
Public:Age^2	.	.	-0.0008
Public:Time trend	.	.	0.1***
Public:Time since HAART	.	.	-0.05
Observations	7945	7945	7945

Table 12: Outcomes: Insurance type (public, private or no insurance).

	Treatment Costs			
	(1)	(2)	(3)	(4)
Treatment: Mono (t)	224.33***	201.32***	204.50***	438.73***
Treatment: Combo (t)	255.40***	223.65***	225.94***	86.38
Treatment: HAART (t)	405.09***	380.45***	390.35***	-62.94
High CD4	.	-62.93***	-66.59***	-108.28***
No Symptoms	.	-137.27***	-142.73***	.
Side-Effects (no AIDS)	.	.	.	151.79***
Symptoms (AIDS)	.	.	.	123.70***
Income/10000	.	49.23***	44.69***	38.81***
Private Insurance	.	.	-11.86	-62.23
Public Insurance	.	.	-80.62**	73.04
Private × Mono	.	.	.	-206.55**
Public × Mono	.	.	.	-441.04***
Private × Combo	.	.	.	159.28
Public × Combo	.	.	.	-136.07
Private × HAART	.	.	.	350.54***
Public × HAART	.	.	.	52.93
Time trend	.	.	.	15.62***
Observations	8300	8300	8300	8300

Table 13: Outcomes: HIV treatment costs.

| | $F_{it} = 1$ | | $F_{it} = 0$ | |
| | (No Ailments) | | (Ailments) | |
	Coefficients	Error	Coefficients	Error
TYPE I				
Constant	0	.	-17.28	0.33
CRRA	0.90	0.54	0.84	0.02
Labor Disutility	-0.13	0.16	0.84	0.25
Begin Treatment	-1.85	0.42	-26.10	0.34
Change Treatment	-5.24	0.15	3.50	0.17
End Treatment	-0.69	0.25	16.65	0.25
TYPE II				
Prob[Type II]	0.48	0.02		
Constant	0	.	-11.03	0.17
CRRA	0.89	0.17	0.90	0.50
Labor Disutility	0.10	0.02	-6.04	0.19
LL	-12498.9			

Table 14: STRUCTURAL ESTIMATES of preference parameters.

| | LABOR SUPPLY (%) | |
	Data	Model
Full Sample	0.66	0.65
Pre-HAART	0.68	0.64
Post-HAART	0.64	0.66
Good health	0.72	0.71
Poor health	0.46	0.48
High Experience	0.70	0.68
Low Experience	0.63	0.63
Older than 50	0.59	0.63
50 or younger	0.68	0.66

Table 15: MODEL-FIT EXERCISE: Employment decisions.

	TREATMENT (%)			
	NONE	MONO	COMBO	HAART
Full Sample				
Data:	0.30	0.20	0.17	0.33
Model:	0.31	0.11	0.21	0.37
Pre-HAART				
Data:	0.45	0.32	0.23	.
Model:	0.44	0.20	0.36	.
Post-HAART				
Data:	0.18	0.09	0.12	0.62
Model:	0.19	0.03	0.09	0.69

Table 16: MODEL-FIT EXERCISE: Treatment choices.

	P[TYPE I]	
	Yes	No
EDUCATION		
$\geq$ College degree	0.56	0.47
$\leq$ High school diploma	0.41	0.54
OCCUPATION		
Prof. specialty	0.54	0.50
Administrative	0.52	0.53
Waiter	0.46	0.53
Craft / Construction	0.74	0.52
Extractive	0.41	0.53
Transport	0.45	0.53

Table 17: POSTERIOR LATENT TYPE PROBABILITIES. For clarification on reading this table, consider the first category (college). Among individuals with a college degree (or higher) the posterior probability of latent Type I is 0.56. Among non-college graduates, i.e., agents with less than a college degree, the posterior probability of latent Type I is 0.47.
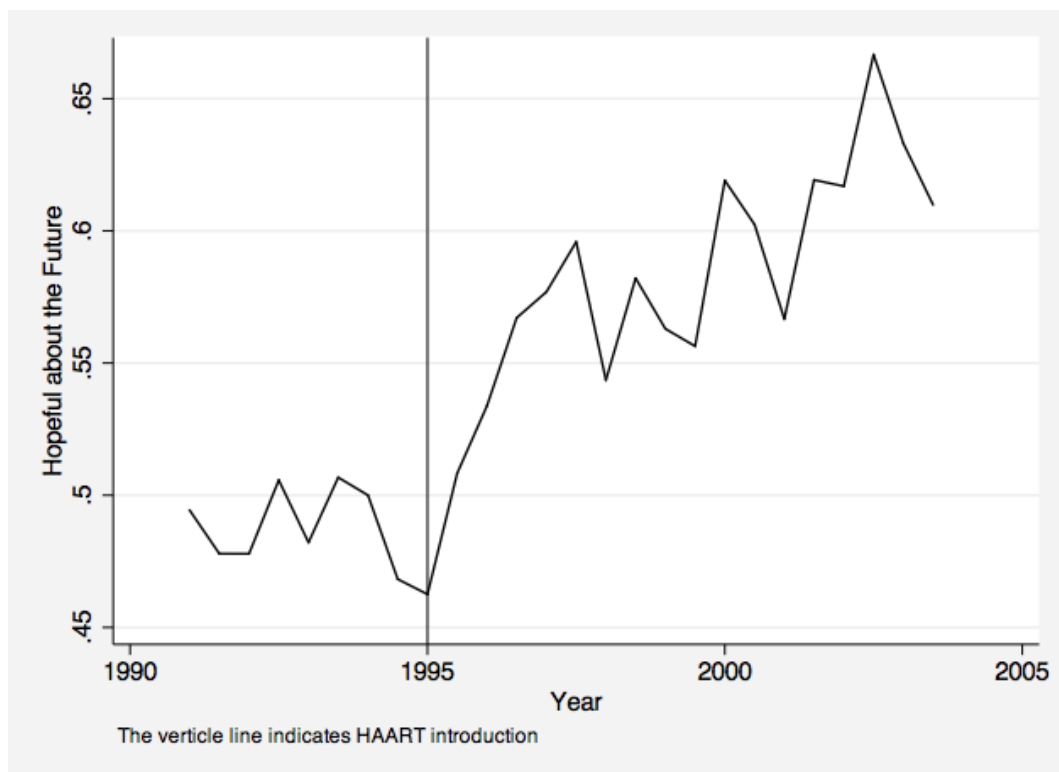
Figure 1: TRENDS: Probability of reporting being hopeful about the future 'most or all of the time' during the week prior to MACS interview.
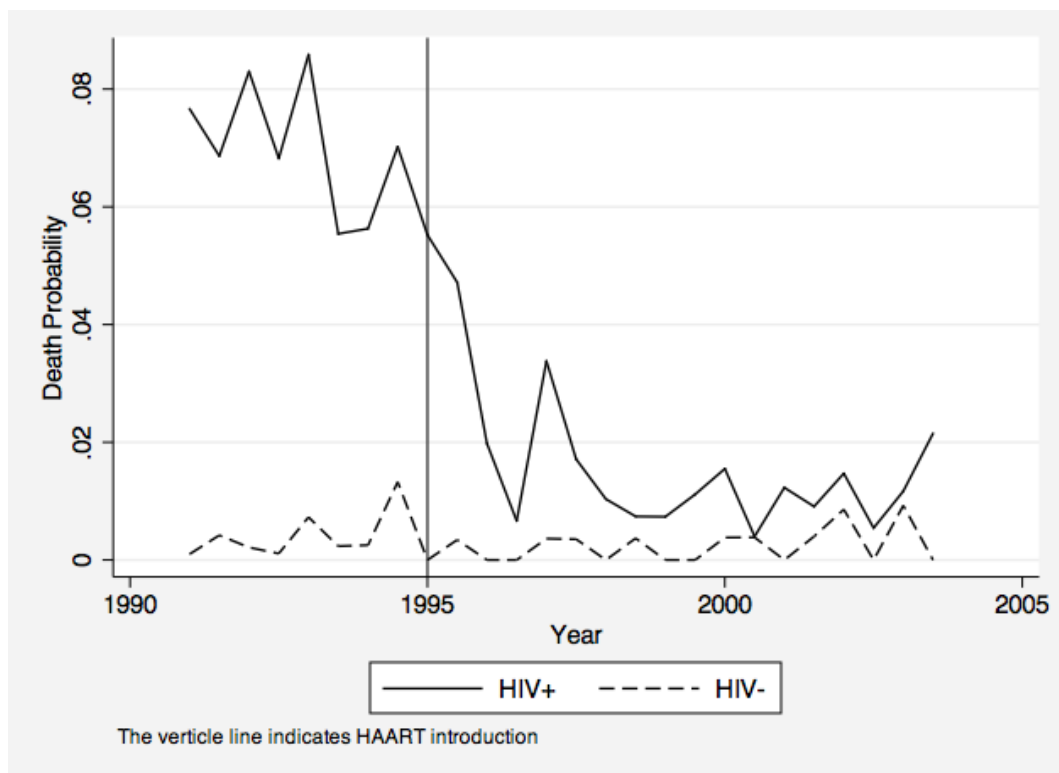
Figure 2: TRENDS: Death Probability over Time

50
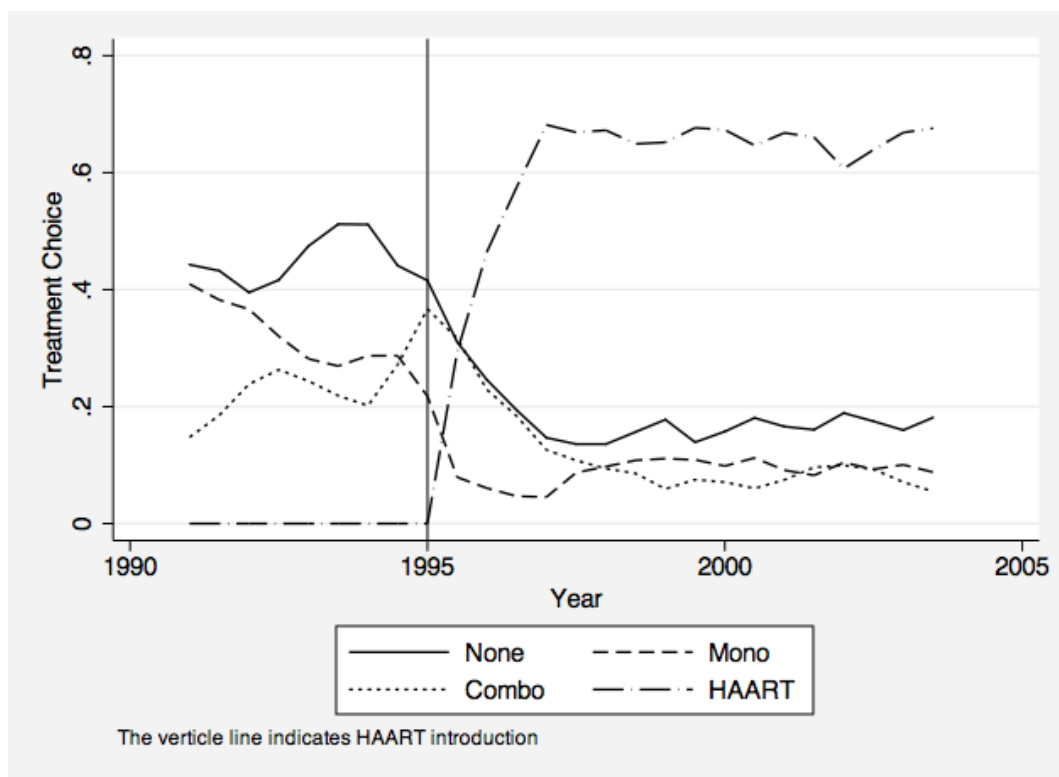
Figure 3: TRENDS: Treatment Choices over Time
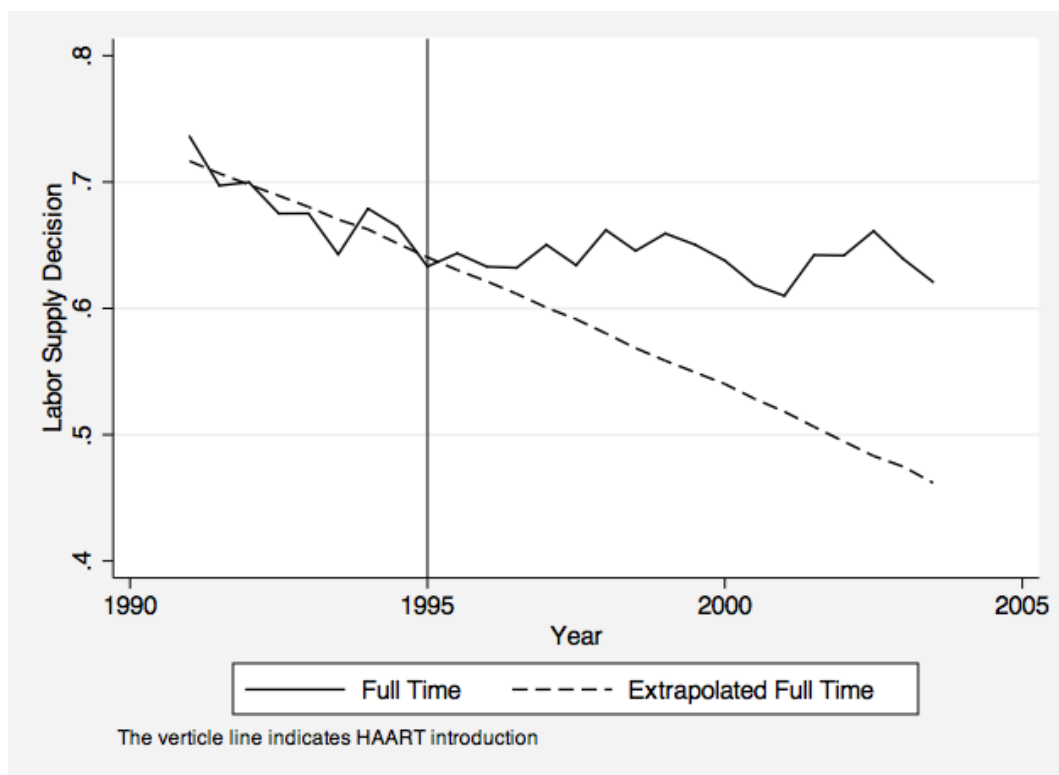
Figure 4: TRENDS: Labor Supply over Time
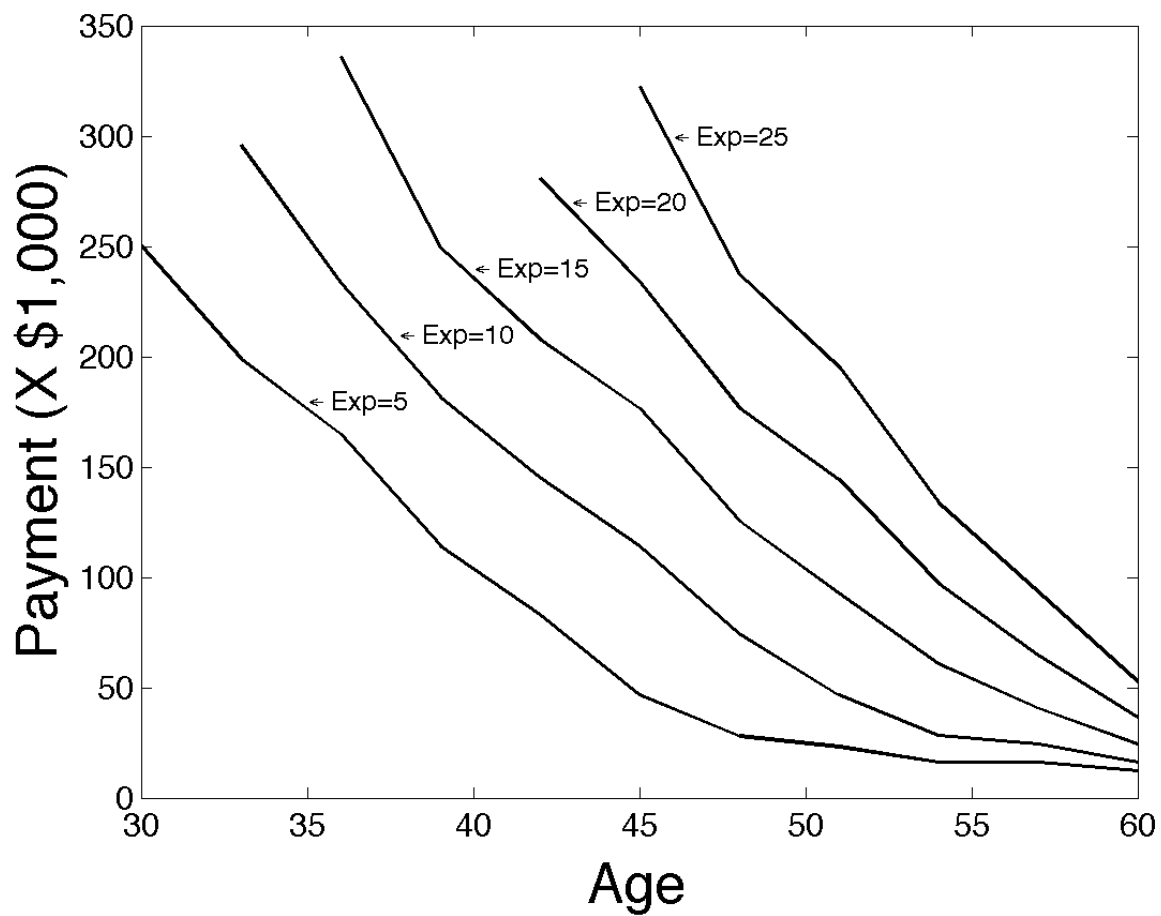
Figure 5: Willingness-to-Accept Payment: HAART - by Age and Human Capital (× $1,000)
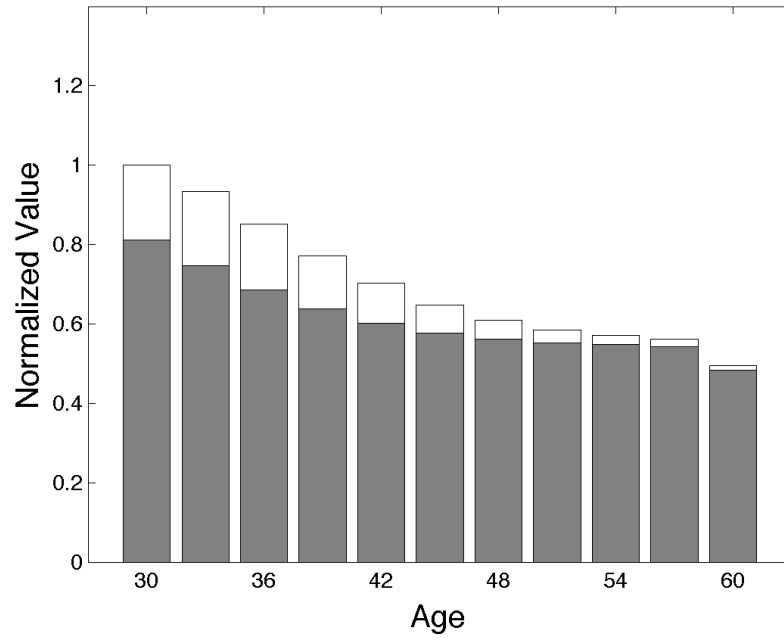
Figure 6: DECOMPOSITION: The Value of HAART: Low Experience
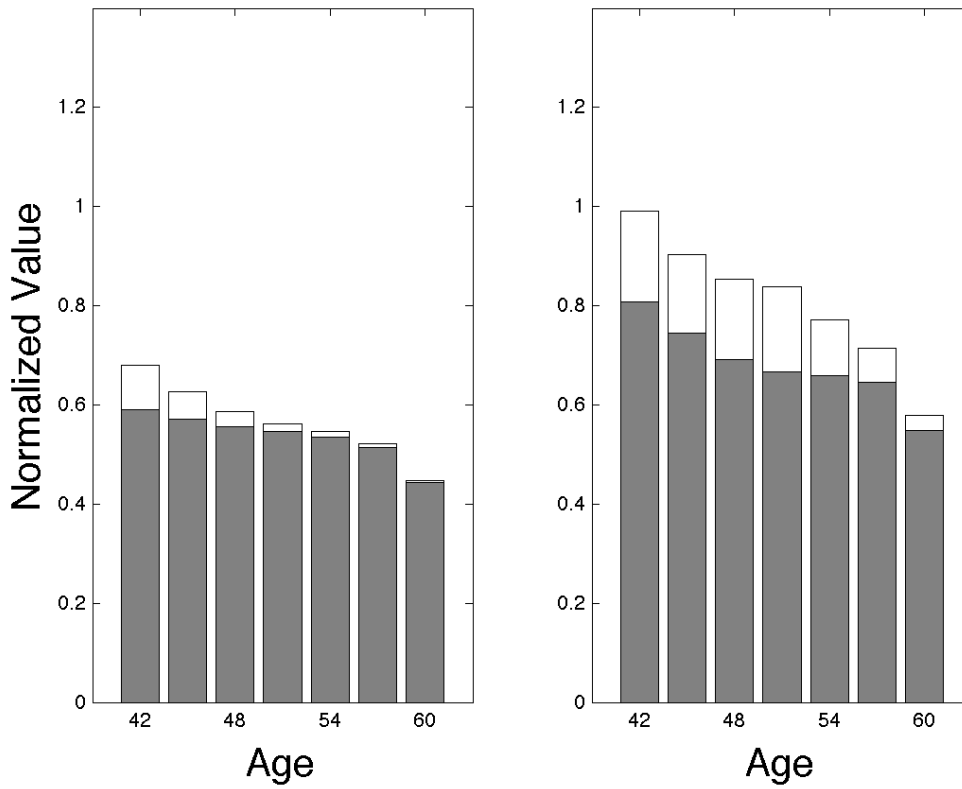
Figure 7: DECOMPOSITION: The Value of HAART: Low versus High Experience

53

CD4: High
Stay off
HAART

97%

7%

CD4: High
Switch off
HAART

CD4: High
to Low if
Off Haart

11%

CD4: High
On HAART

97%

CD4: Low
Switch onto
HAART

CD4: Low
to High if
On HAART

33%

99%

CD4: Low
Stay on
HAART
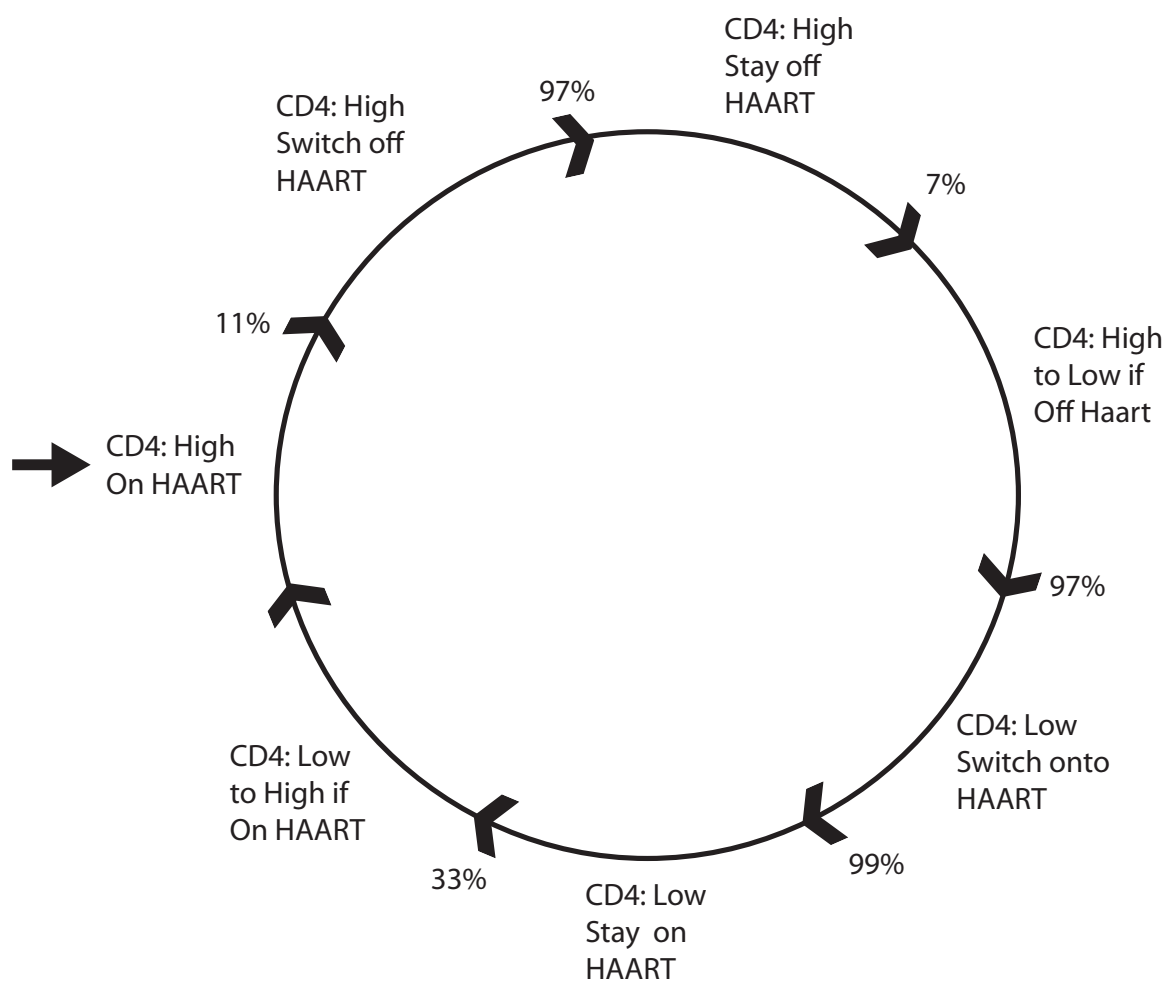
Figure 8: ANATOMY OF OPTIMAL TREATMENT CYCLING
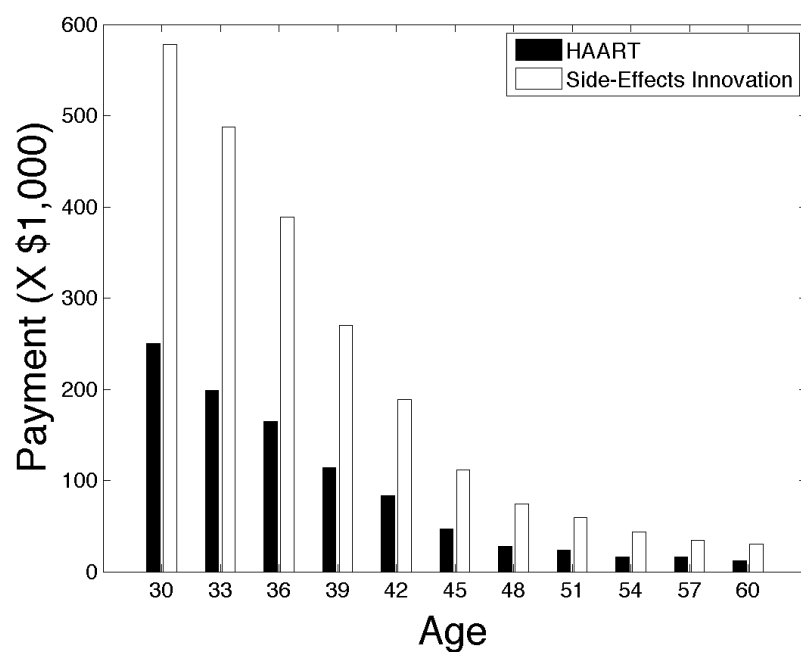
Figure 9: WILLINGNESS-TO-ACCEPT PAYMENT for counterfactual treatment innovations and HAART at different ages holding human capital constant (experience=5 years)
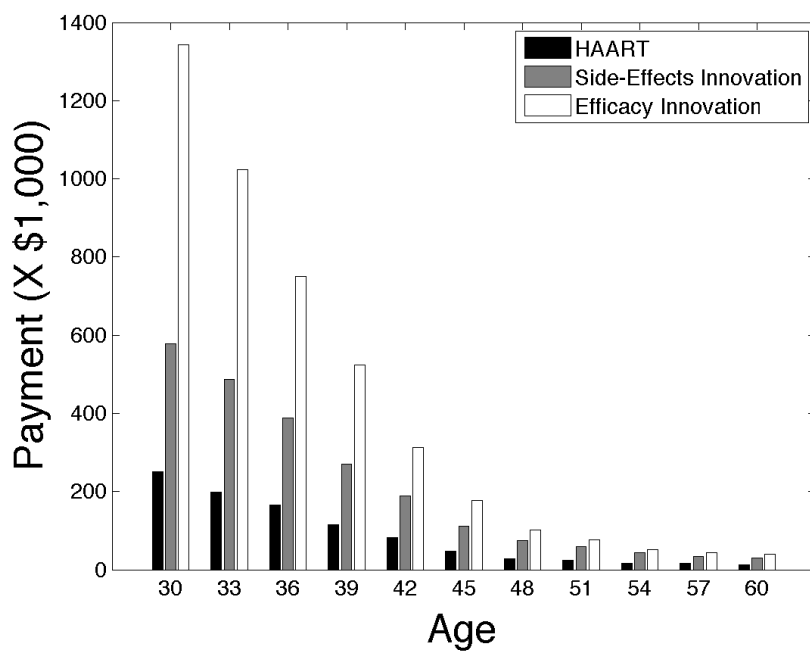
Figure 10: WILLINGNESS-TO-ACCEPT PAYMENT for counterfactual treatment innovations and HAART at different ages holding human capital constant (experience=5 years)
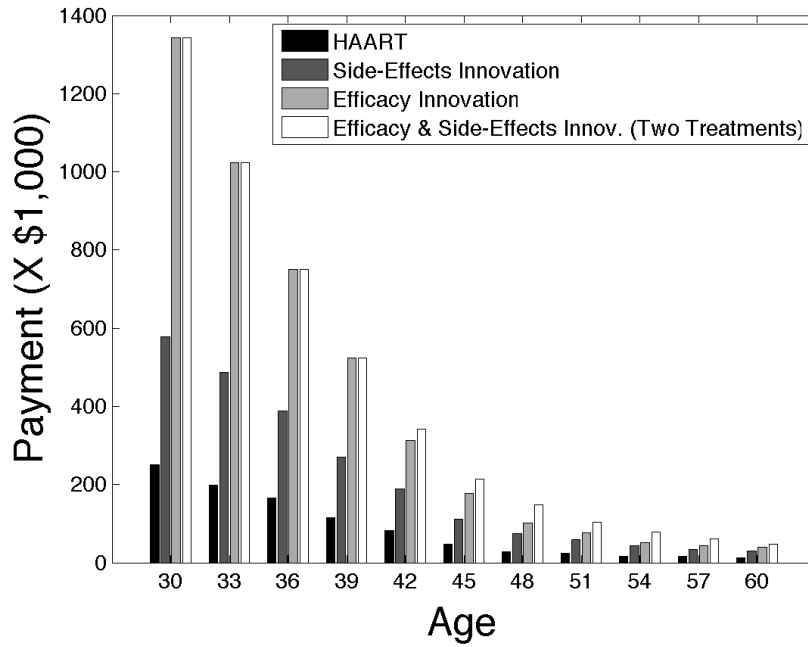
Figure 11: WILLINGNESS-TO-ACCEPT PAYMENT for counterfactual treatment innovations and HAART at different ages holding human capital constant (experience=5 years)
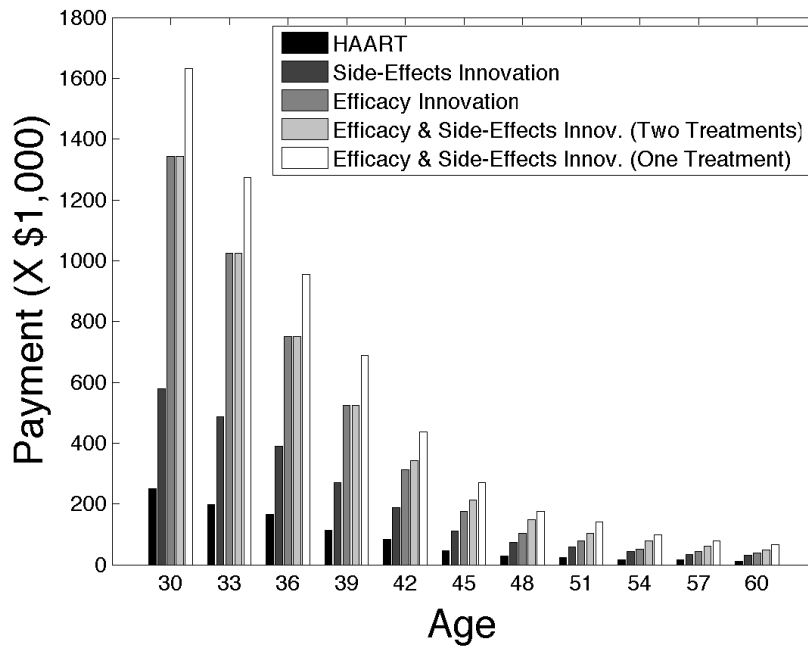
Figure 12: WILLINGNESS-TO-ACCEPT PAYMENT for counterfactual treatment innovations and HAART at different ages holding human capital constant (experience=5 years)
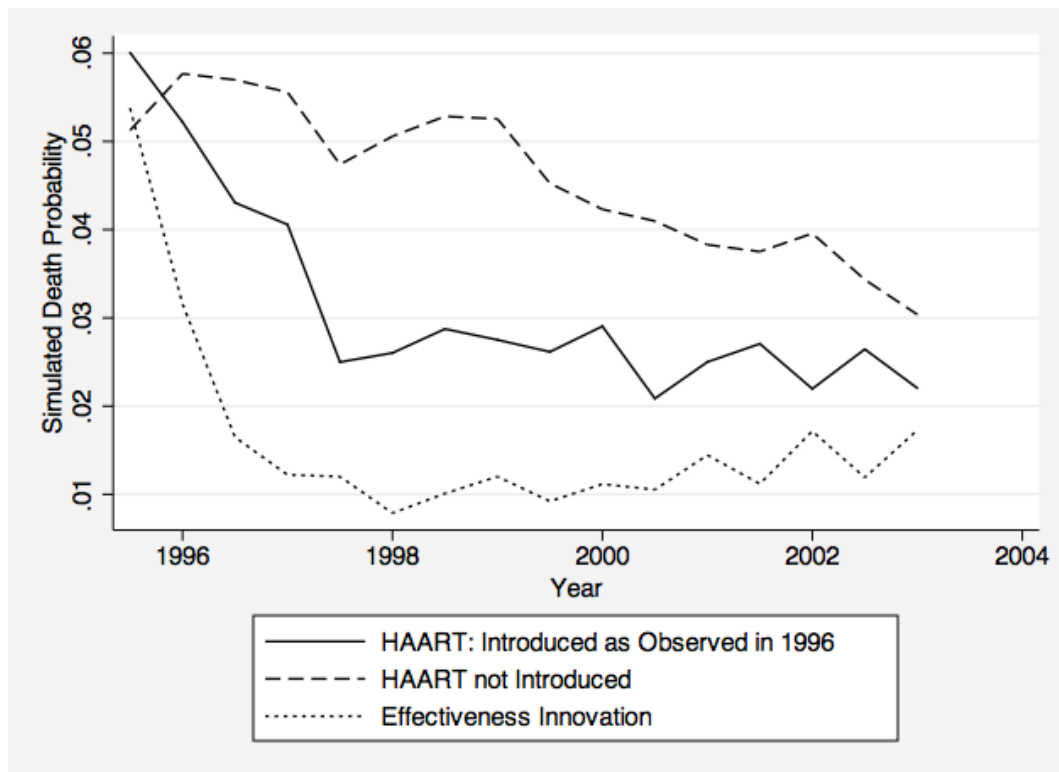
Figure 13: SIMULATION - COUNTERFACTUAL TREATMENTS: Death Probability, Type II
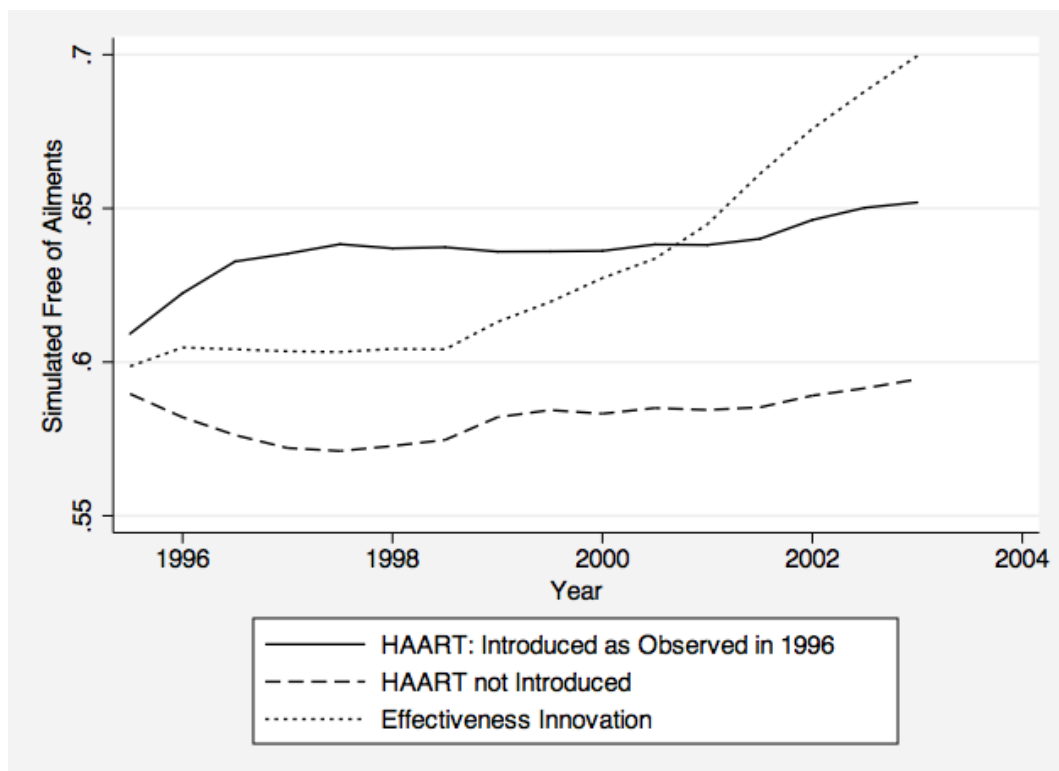
Figure 14: SIMULATION - COUNTERFACTUAL TREATMENTS: No Ailments Probability, Type II
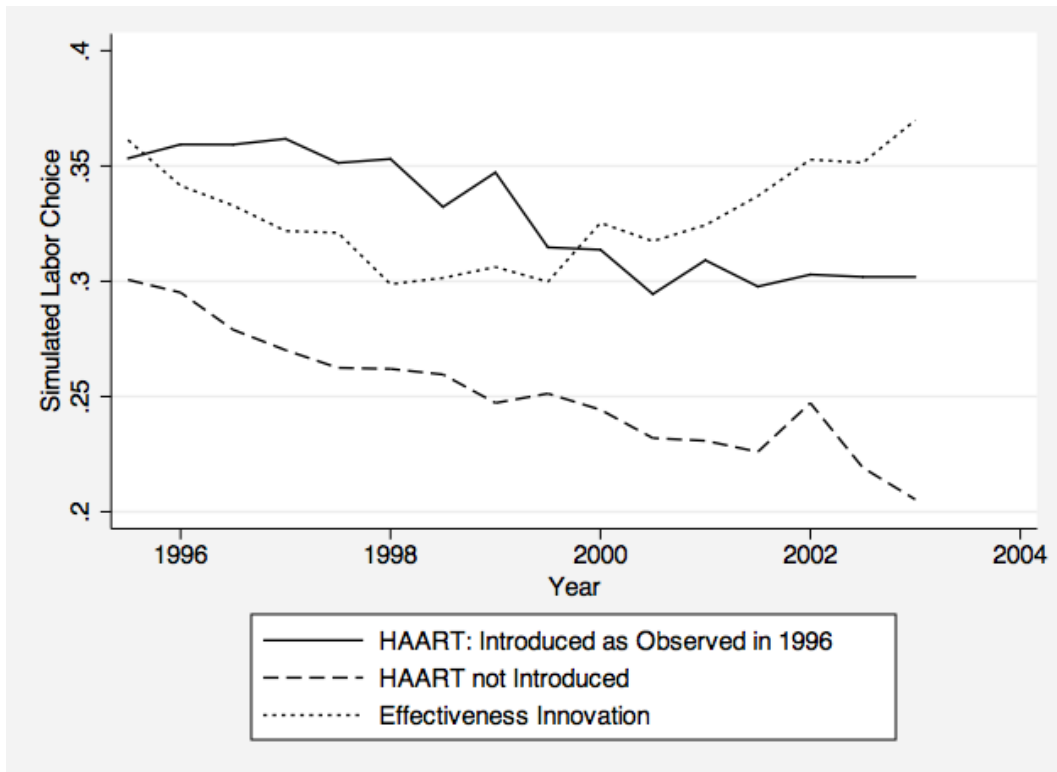
Figure 15: SIMULATION - COUNTERFACTUAL TREATMENTS: Full Time Employment, Type II
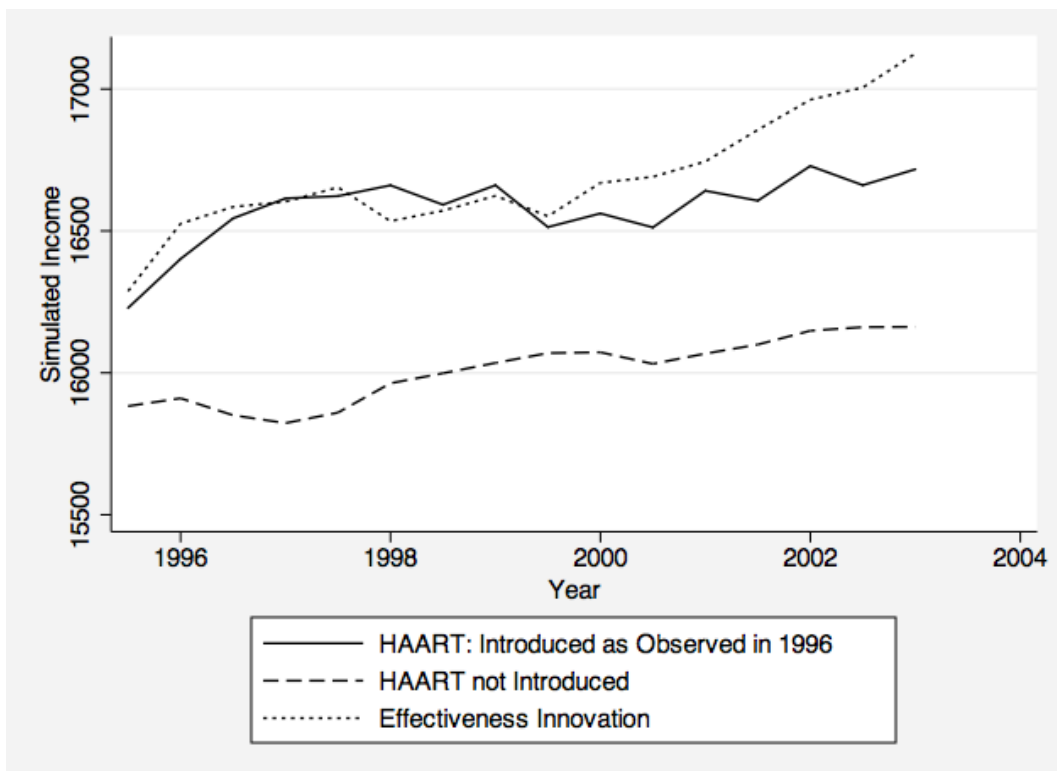
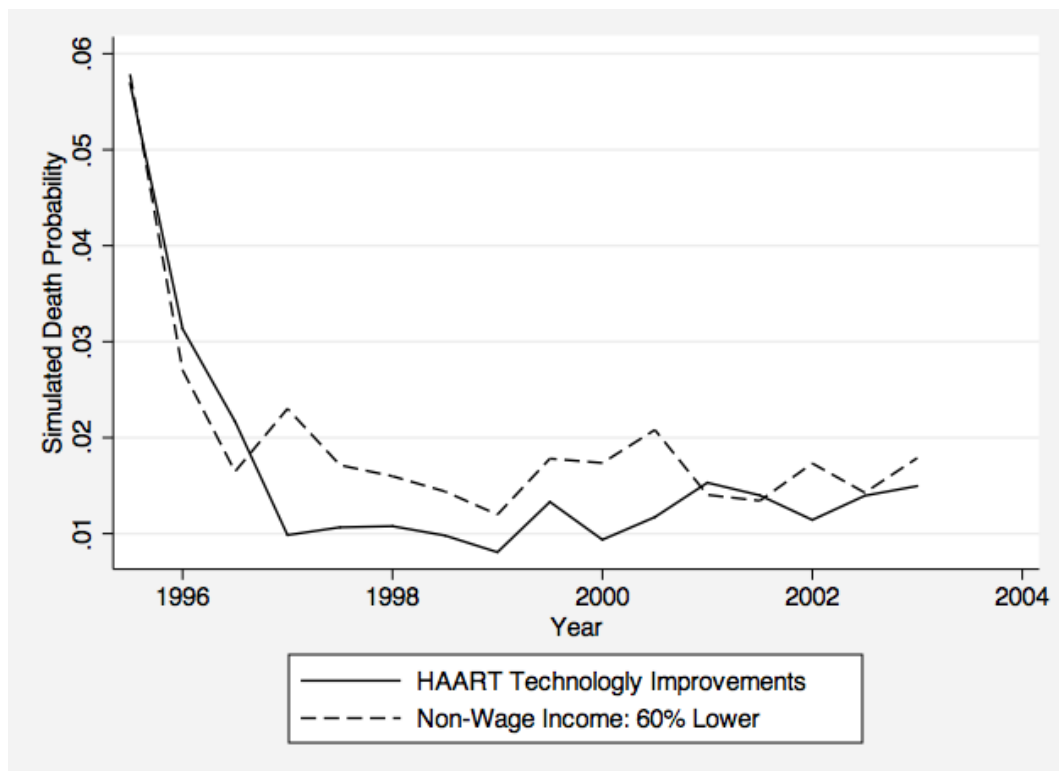**Figure 16:** SIMULATION - COUNTERFACTUAL TREATMENTS: Income, Type II
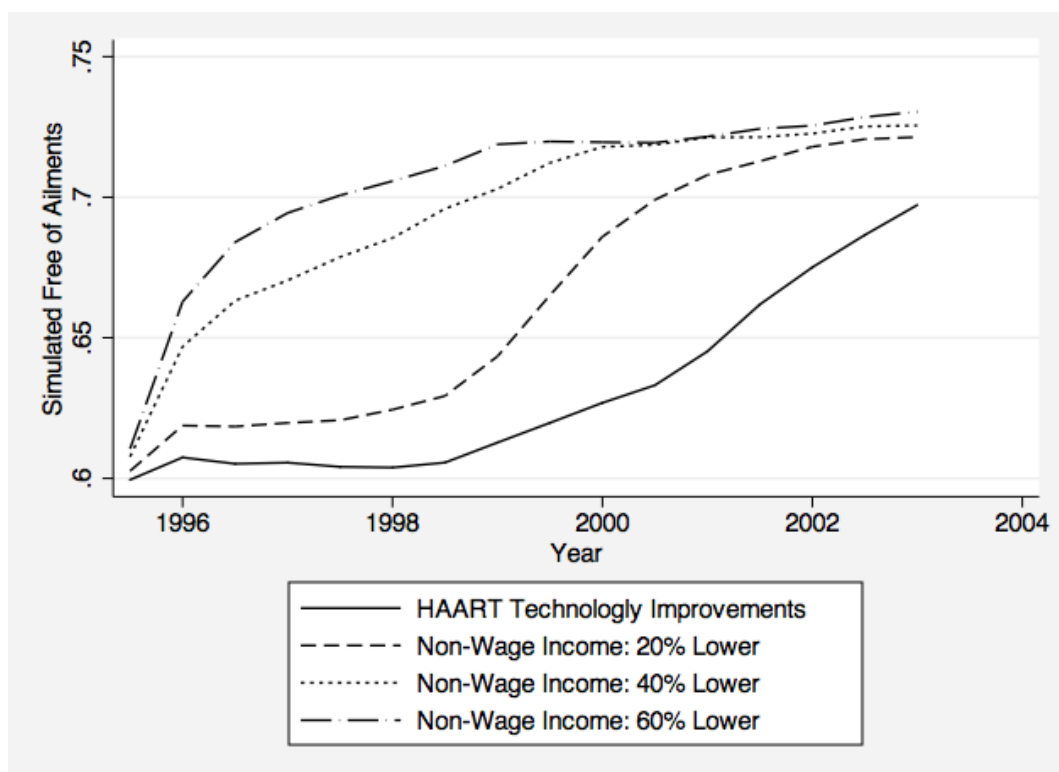
Figure 17: SIMULATION - COUNTERFACTUAL TREATMENTS: No Ailments Probability, Type II
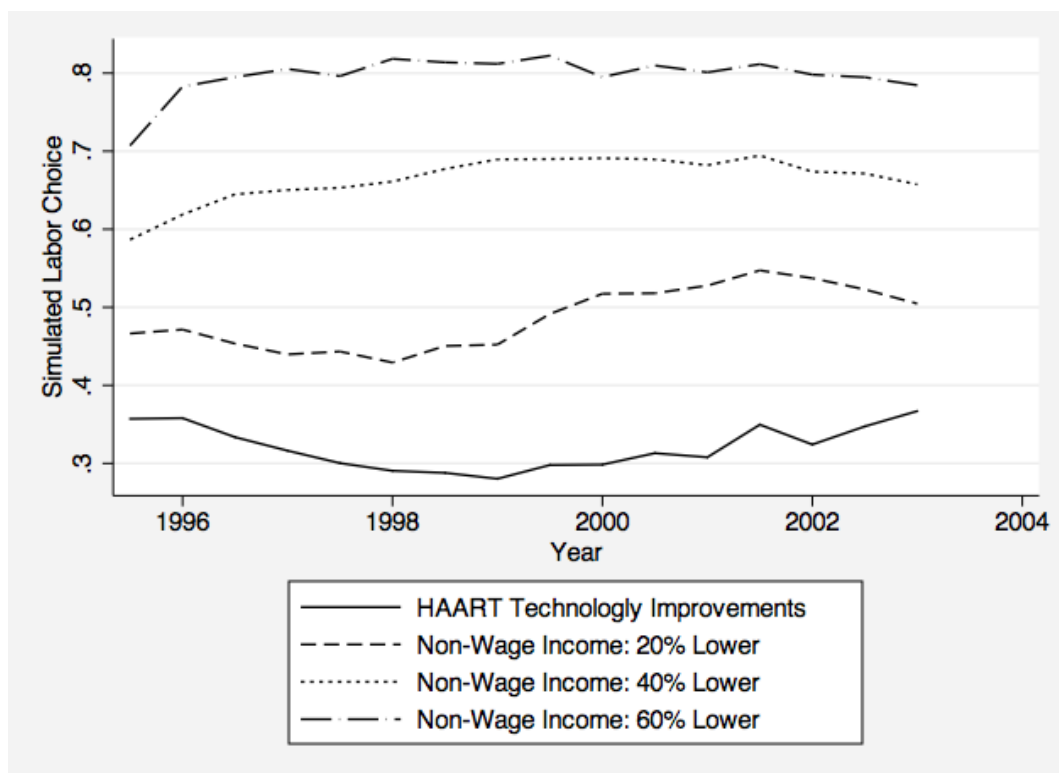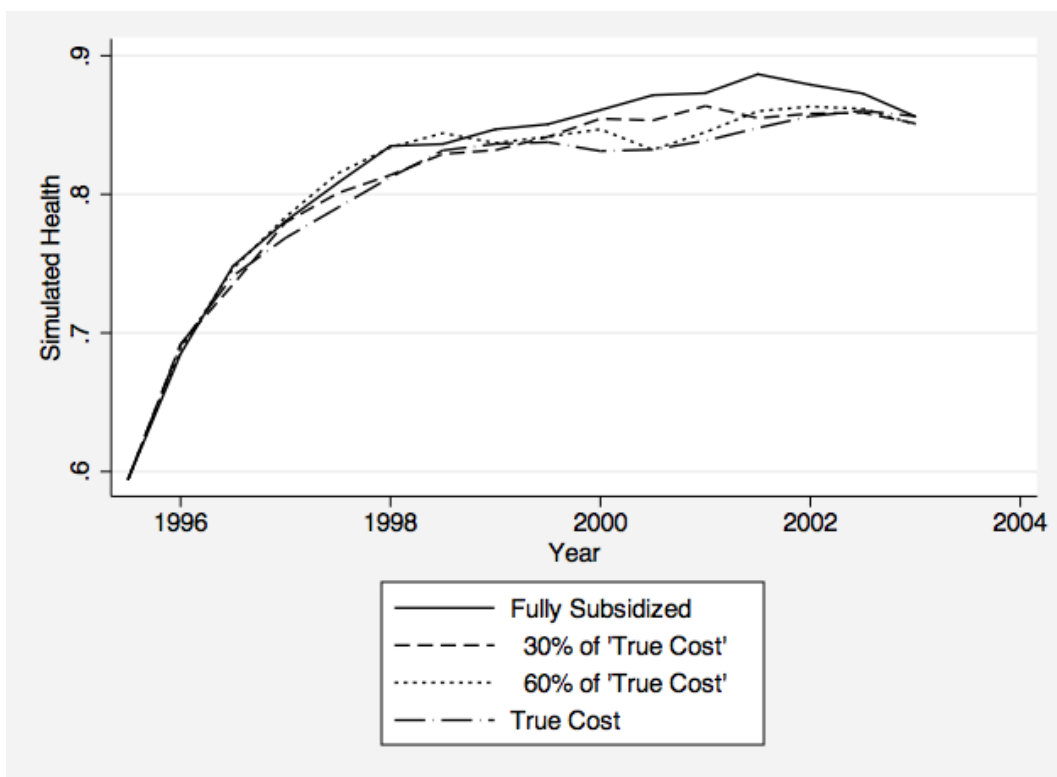

Figure 18: SIMULATION - COUNTERFACTUAL TREATMENTS: Full Time Employment, Type II
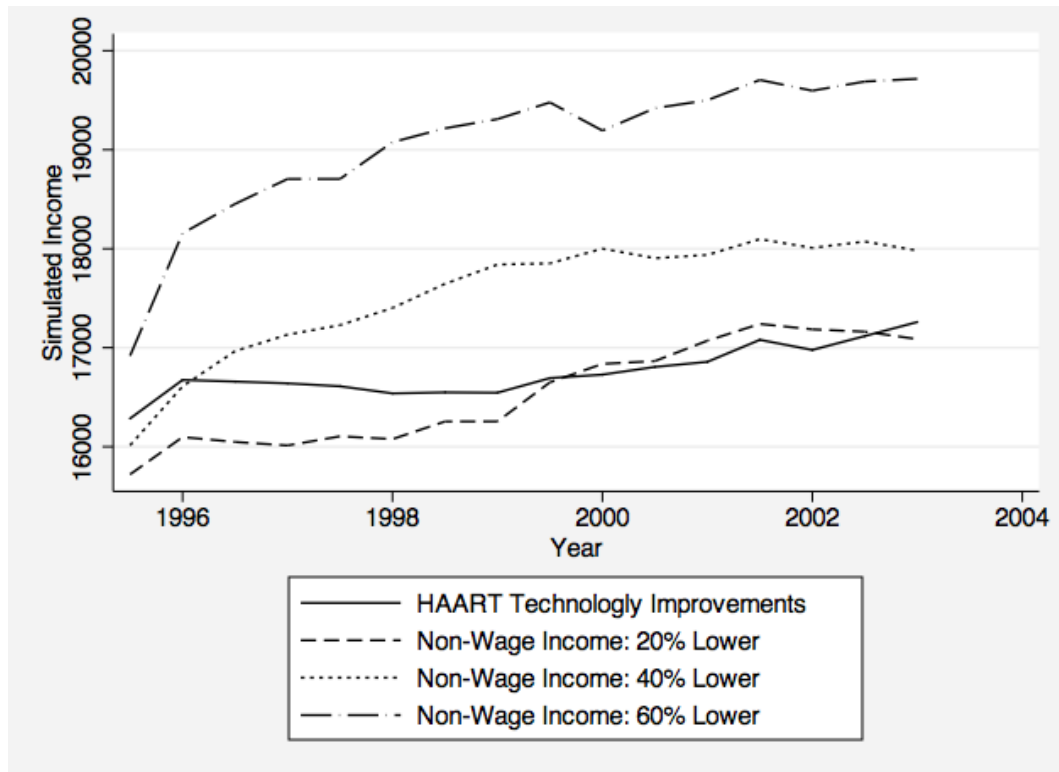
Figure 19: SIMULATION - COUNTERFACTUAL TREATMENTS: Semester Income, Type II

Figure 20: SIMULATION - COUNTERFACTUAL TREATMENTS: Death Probability, Type II

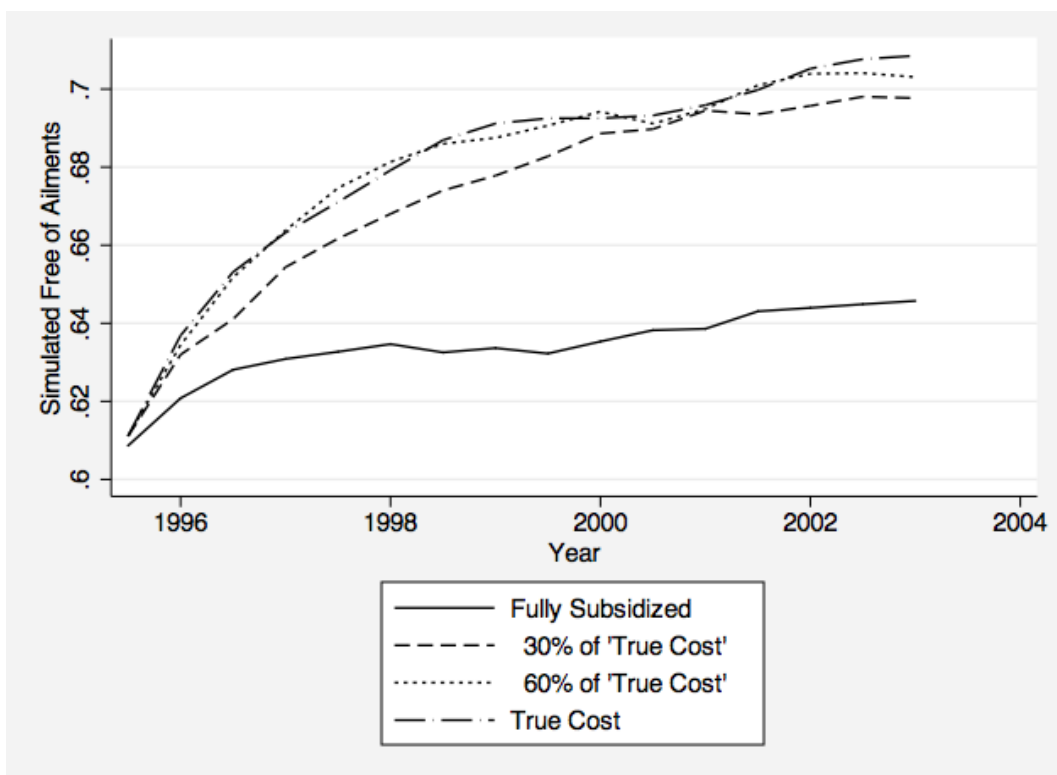**Figure 21:** SIMULATION - NON-SUBSIDIZED DRUG COSTS: High CD4 Probability, Type II

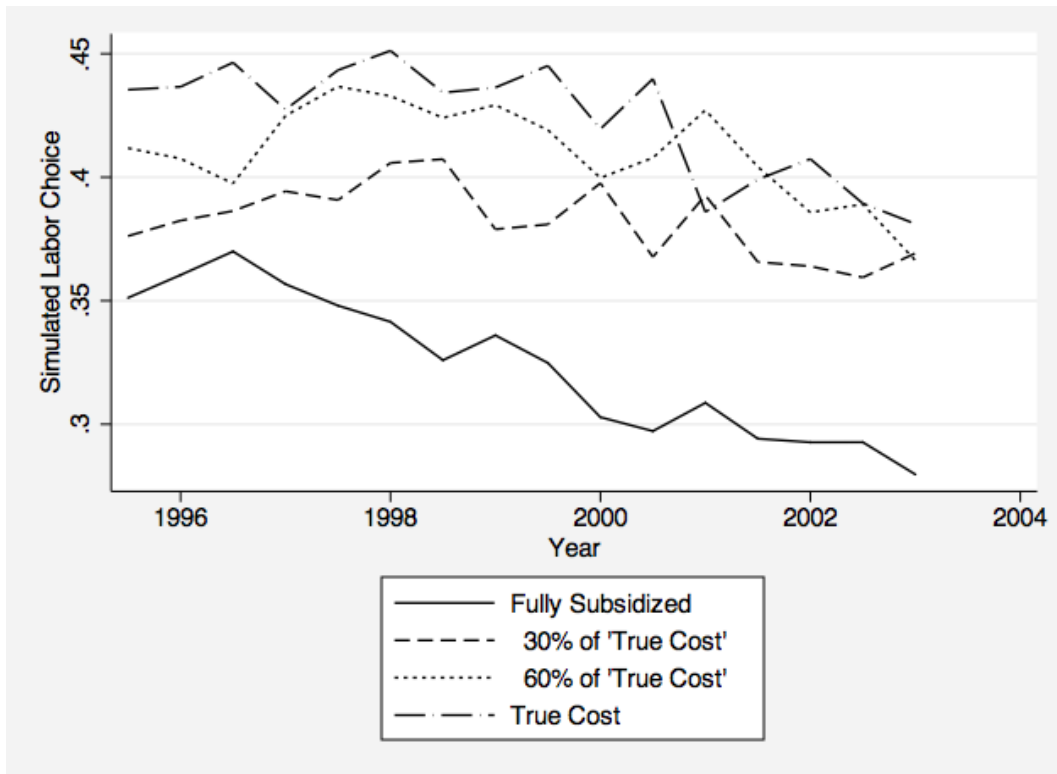**Figure 22:** SIMULATION - NON-SUBSIDIZED DRUG COSTS: No Ailments Probability, Type II

Figure 23: Simulation - Non-Subsidized Drug Costs: Full Time Employment, Type II
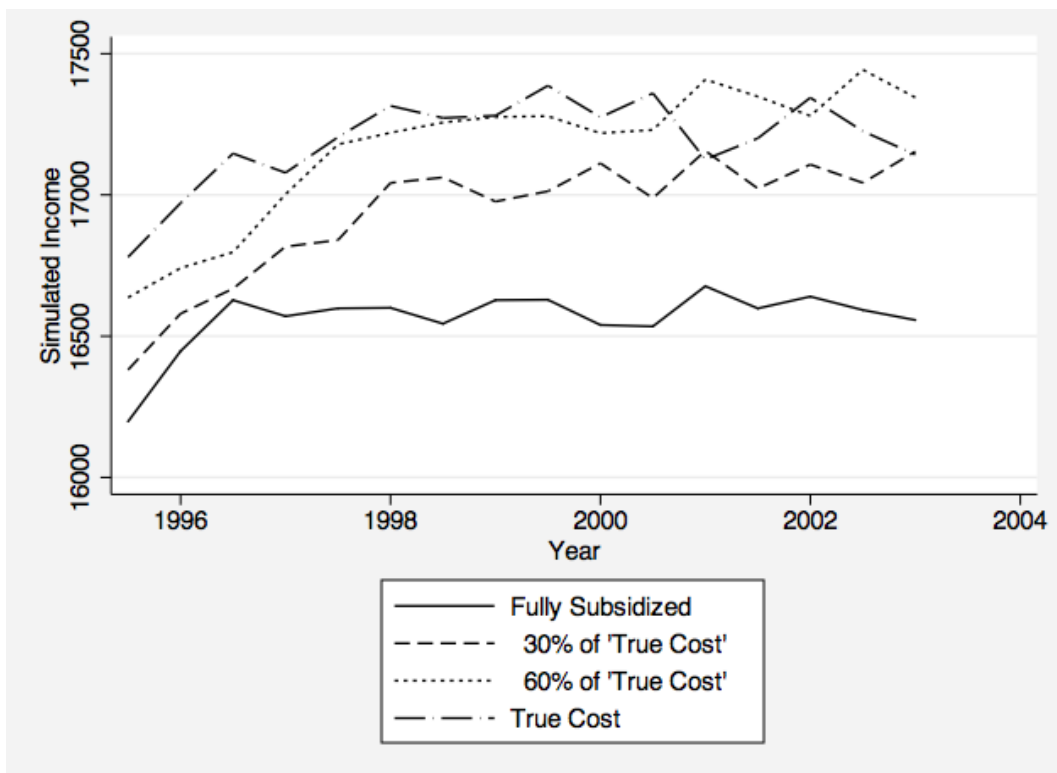

Figure 24: Simulation - Non-Subsidized Drug Costs: Semester Income, Type II

62