Medical Innovation and Health Disparities*

Barton H. Hamilton
hamiltonb@wustl.edu
Olin Business School, Washington University in St. Louis

Andrés Hincapié
andres.hincapie@unc.edu
Department of Economics, University of North Carolina at Chapel Hill

Emma C. Kalish
ekalish1@jhu.edu
Consumer Financial Protection Bureau

Nicholas W. Papageorge
papageorge@jhu.edu
Department of Economics, Johns Hopkins University and NBER

ABSTRACT: A long-standing question in health economics asks what drives heterogeneity in health investments across sociodemographic groups, which contribute to health disparities. We examine a novel source: variation by sociodemographic groups in the impact of medical treatment side effects. We estimate a lifecycle model of medication and labor supply decisions using data on men infected with human immunodeficiency virus (HIV) and with different levels of completed education. Agents in the model make dynamic medication and work choices as a function of expected income, health, and mortality. We also include utility costs of side effects, which can interact with the utility cost of work, and allow these parameters to vary by education. We use the estimated model to evaluate the disparate impacts of an effective HIV treatment innovation that had harsh side effects: HAART (which stands for highly active antiretroviral treatment). Measured in lifetime utility gains, HAART disproportionately benefited patients with more education in part due to differences in income and mortality, but also due to smaller impacts of side effects on labor supply. We also simulate the effects of a HAART treatment mandate, which mimics assignment to treatment in a clinical trial. The mandate improves health, which might be viewed as a success in a randomized trial. However, this masks the downsides of treatment, including lower labor supply due to side effects, especially for lower-education men. Broadly, viewing heterogeneity in health investments as solely a result of barriers to access is overly simplistic. Individually optimal investments in health may in part reflect the costs of managing side effects and work. If low health investments create negative externalities (e.g., due to increased use of publicly-funded healthcare), the benefits of policies that moderate resulting health-work tradeoffs could outweigh the costs.

KEYWORDS: Health Disparities, Health Behaviors, Dynamic Demand, Side Effects, Structural Models, HIV/AIDS.

JEL CLASSIFICATION: I12, I14, I20, J2, O31.

*First version: May 22, 2021. This version: May 14, 2026. We are grateful for helpful suggestions and comments from seminar participants in the Hopkins Economics Applied Micro Brownbag Seminar, the Annual Health Econometrics Workshop, the University of Copenhagen CEBI Workshop, and EAFIT–Colombia.

1 Introduction

Low investments in health contribute to a variety of poor health outcomes, such as high rates of infectious disease and chronic illness, both of which decrease mortality and reduce quality of life. Policy interest in health investments arises because poor health may lead to externalities by imposing costs on society, e.g., through public health expenditures, increases in insurance premia, or losses to productivity. Yet, policy options to address these externalities require a more complete understanding of key sources of variation in health investments. Indeed, a long-standing question in economics is why individuals invest differently in their health. Why do some people engage in risky health behaviors when others do not? Why do some people reject healthful eating, exercise, and preventative care when others embrace it? Why do some people adopt effective medical treatments when others avoid them?

Grossman (1972) provides a basic framework for analyzing health investments. The framework posits health as a form of human capital in which agents make costly investments in their health to increase their lifetime utility by improving the length and the quality of their lives. The model also allows health to improve productivity and earnings so that health investments generate labor market returns. According to the model, individuals invest in their health until the marginal benefits (including higher earnings) no longer exceed the marginal costs. The Grossman (1972) model has been used in numerous studies[1] and expanded upon to incorporate myriad factors relevant to many health contexts, e.g., addiction (Darden, 2017), side effects (Papageorge, 2016), mental health (Cronin et al., 2024), and relationships to other forms of human capital, such as schooling (Grossman, 2017), among many others.

However, the framework has trouble explaining a consistent pattern: low investments in health are often concentrated among low-income or otherwise under-resourced individuals, contributing to stubbornly persistent health disparities across sociodemographic groups. For example, people without a high school diploma are 6–7 times more likely to be in poor or fair health compared to those with a bachelor's degree or higher (Goldman and Smith, 2011). The Grossman (1972) model attempts to explain such patterns through the cost of health investments (e.g., the high price of new medications that low-income individuals cannot afford). While resource constraints help to explain some health disparities,[2] they tell only part of the story. The reason is that health disparities may persist even when health investments are low-cost or virtually free.[3] This suggests a need to further expand the Grossman (1972) model to better understand population variation in health investments and resulting health disparities.

[1] According to Google Scholar citations, as of January 1, 2025 Grossman (1972) has been cited over 11,000 times.
[2] See, e.g., Braveman (2006) for a review.
[3] See, e.g., Woolf and Braveman (2011) on non-financial barriers as sources of health disparities.

This study proposes a novel extension of the Grossman (1972) model to help explain differences across sociodemographic groups in health investments. The extension is designed to capture differences across sociodemographic groups in tradeoffs arising when physical ailments (symptoms of illness or side effects of medical treatment) potentially interfere with work. If treatments reduce symptoms, investments in health not only reduce mortality, but also improve how people feel, potentially encouraging labor supply. Alternatively, if treatments have side effects, a health-work tradeoff emerges because health investments can make people feel worse in the short run, potentially discouraging employment. Heterogeneity by sociodemographic groups in how individuals approach this type of health-work tradeoff can contribute to health disparities if, for example, side effects interfere with work relatively more for certain sociodemographic groups. Finally—and importantly—the model also captures more standard features of lifecycle health and work interactions that could help to explain differences in behavior, e.g., variation across demographic groups in labor income, the impact of medication on health, and the impact of health on mortality. Thus, the model allows us to weigh the importance of the side effects channel against other factors affecting work and medication choices.

We apply the model to examine differences in health investments across education groups in the context of HIV and a life-saving medical breakthrough, HAART, which entered the market in 1995.[4] We focus on medication usage among HIV-infected (henceforth *HIV-positive* or *HIV+*) men, comparing those with and without a college degree. This context provides a useful example to examine heterogeneity in health investments: When it entered the market, HAART was by far the most effective treatment—indeed, the only reliably life-saving treatment—for HIV. Yet its uptake, while high, was far from universal (as we show, it hovered around 70% during the period we study: 1995-2003). Moreover, non-uptake and non-use of HAART was more pronounced among individuals without a college degree compared to those with a college degree, consistent with variation in health investments often observed across sociodemographic groups.

Our analysis begins with descriptive evidence using rich data on the treatment choices, labor supply decisions, health outcomes, incomes, insurance status, and out-of-pocket medical expenditures of HIV+ men.[5] This analysis suggests that, in the context of HIV/AIDS and HAART in the U.S., monetary costs played a negligible role in explaining variation in treatment uptake, mostly because HIV medications are highly subsidized and out-of-pocket costs are low compared to income. Additionally, we find that individuals with higher education tend to work more and use

[4] HIV stands for human immunodeficiency virus, which is a virus that attacks the immune system leading to a condition known as acquired immunodeficiency syndrome (henceforth *AIDS*). HAART, which stands for "highly active antiretroviral treatment" was introduced in 1995. Further details on HIV/AIDS and HAART are in Section 2.

[5] Data for this study come from the Multi-Center AIDS Cohort Study, which has followed a sample of men who have sex with men starting in the 1980s, when the AIDS epidemic began in the U.S. We introduce the data set in Section 2.

more medication, suggesting that the net returns to both health investments and labor supply are relatively high. Yet, our preliminary analysis also shows that both the use of HAART and physical ailments are negatively associated with full-time work and that the negative interaction between HAART use and labor supply is stronger on for people with lower education.[6]

Armed with these empirical patterns, the centerpiece of our analysis is to construct and, using the same data set, estimate a structural lifecycle model that envisions agents making repeated medication and labor supply decisions to maximize their lifetime welfare. Broadly, the model can rationalize behavior that leads health to deteriorate since agents are not modeled as solely maximizing their survival or some other measure of health. Instead, the model permits agents to balance their incentive for better health with other factors, such as work and consumption, along with the disutility of ailments (e.g., side effects of medication), which in turn can affect labor supply. In this sense, we follow Hall and Jones (2007), who distinguishes between health and welfare. The model also allows utility parameters (along with state-to-state transition probabilities and outcome probabilities) to vary by education group. Doing so means health-work tradeoffs can differ by education group, giving rise to endogenous variation in health investments and, thus, health disparities. The estimated model reveals key differences between education groups. Individuals with lower education face higher mortality, a higher baseline likelihood of suffering physical ailments, and lower income. Parameters also imply that physical ailments are costly, leading to lower utility in general and a larger utility cost of work. While the disutility of physical ailments in general is similar across education groups, the negative utility of ailments interacted with work is larger for people with less education, which means they face a more drastic tradeoff between medication use and work.

Using the estimated model, we conduct three main counterfactual analyses. First, we measure the difference in lifetime utility in the first year after HAART entered the market compared to the year before HAART was invented. Health disparities existed during the early years of the AIDS epidemic, when treatments were largely ineffective. For example, in the early 1990s, the 6-month mortality rate for HIV+ college graduate men was almost half that for HIV+ men without a college degree. HAART drastically reduced mortality, improved health, and increased welfare. However, the side effects of HAART, including their interaction with the utility cost of work, meant that the benefits of the treatment disproportionately accrued to individuals with higher education. Converting value to 2000 dollars, a 30-year-old with a low CD4 count in the first post-HAART period would accept about $180,000 to return to a pre-HAART world. A similar individual with a college degree would accept about $315,000, a difference of $132,000. For a similar man with a high CD4 count, the corresponding values are $172,440 and $271,575, for a difference of about $99,000. In other words, when we account for a broad set of factors—including labor supply—that drive health

[6] We note that this final interaction is not always significant depending on specification and controls.

decisions, we find that the innovation increased inequality because it disproportionately benefited more educated patients.

Second, we decompose the value of HAART by education group to better understand how HAART generates value across sociodemographic groups. We begin with the total value of HAART and then replace parameters of low-education individuals with those of high-education individuals to understand to what degree different sets of parameters close the education gap in the value of HAART. Decompositions show that several differences by education group explain variation in their value of HAART, which we compute for people in dangerously poor versus relatively good health: income (40% and 44%), mortality (30% and 25%), the production of ailments (17% and 16%), and utility parameters capturing the interaction between work and side effects (13% and 14%).[7] The interpretation is that the low-education group values HAART less since they earn less in any life-year gained, are less likely to see reductions in mortality due to lower usage, are more likely to suffer ailments, and face a larger negative impact on the utility of work due to side effects. Indeed, parameters related to ailments account for roughly one-third of the differences in how individuals from different education groups value HAART. In contrast, differences in the evolution of health explain little of the difference in the value of HAART, as health parameters are similar across education groups. Similarly, and as expected given reported costs, differences in out-of-pocket treatment costs and insurance status explain none of the differences in the value of HAART between those with less and more education.

Third, we simulate the effects of a 6-month HAART treatment mandate. The mandate mimics what we might observe in a clinical trial that shows improvements to health among people assigned to treatment (and thus may be considered a success) but fails to consider tradeoffs associated with treatment and other economic outcomes. The simulation casts doubt on whether clinical trials offer the full picture of the impacts of a treatment. As expected, mandating an effective treatment improves health. Yet, due to side effects that manifest as physical ailments, it also reduces expected lifetime value. Lifetime utility declines more for those with less education (2.8% compared to 1.4% for those with a college degree), because they are less likely to use treatment before the mandate. This is reflected in larger decreases in labor supply (4.1% compared to 1.6%). This policy simulation underscores how individuals may optimally avoid treatments that generate health improvements if there is a high cost to usage. In the scenario we study, the costs of treatment are uneven, so mandating treatment is more costly for those with less education, exacerbating welfare inequality. Results from this counterfactual analysis highlight that health and welfare may not go

[7]Totals for the less healthy group do not sum to 100 because the low education groups benefit relatively more from utility parameters not related to health and work interactions, which means switching their parameters to those of the high education group reduces their value of HAART. Moreover, in principle, these numbers could be quite different depending on the order of the decomposition. In practice, we try all permutations and find small differences in the reported percentages. We elaborate on this point when presenting decomposition results in Section 4.

hand in hand, especially for low-education individuals.

This paper relates to a vast literature in public health and other fields that documents and examines the consequences of health disparities across socioeconomic or demographic groups (Adler and Rehkopf, 2008; Beer et al., 2011; Conti et al., 2010; Currie, 2009; Cutler et al., 2011; Goldman and Smith, 2011; Rubin et al., 2010; Williams and Jackson, 2005). In this literature, barriers to access are frequently identified as a source of health disparities.(Chang and Lauderdale, 2009; Lasser et al., 2006; Williams et al., 2010; Woolf et al., 2015). Another literature, however, discusses how health disparities may result from persistent differences in behavior across socioeconomic groups. Often, these differences are characterized as errors in judgment (e.g., impatience or present bias) and sometimes they are (implicitly) presented as reckless or careless choices (Adimora and Schoenbach, 2002; Robinson and Moodie-Mills, 2012). Our findings show that persistent behavioral differences across sociodemographic groups can lead to different health outcomes, but need not reflect biases or carelessness. Rather, they can reflect rational responses to prevailing circumstances, constraints, and market conditions.[8]

This paper complements the approach taken by the sociological literature on fundamental cause theory. This theory says that social factors such as socioeconomic status are "fundamental causes" of health that persist despite, or even because of, medical innovation (Link and Phelan, 1995; Phelan et al., 2010). Empirical studies based on this literature have found increased health disparities after innovations in cholesterol treatments (Chang and Lauderdale, 2009), respiratory treatments for infants (Frisbie et al., 2004), and HIV treatments (Rubin et al., 2010), among others (Glied and Lleras-Muney, 2008). While many of these papers point to access to medical care as a driver of this increased inequality, other factors, such as the difficulty of complying with a complicated regimen, are also considered (Chang and Lauderdale, 2009; Goldman and Lakdawalla, 2005). The structural model estimated in our paper allows us to quantify the impact of various factors, apart from the financial costs of medication, on treatment choices and welfare.

We also relate to literature examining ways to reduce health disparities. Often, a concern is that health disparities are not only unfortunate for the individual but can create negative externalities. Policies studied in this literature include lowering health care prices to expand access for low-income groups, paying people to take care of their health, providing information about the risks and benefits of specific health-related behaviors, and making health care more convenient (Avery et al., 2008; Gerber et al., 2005; Osborn et al., 2007; Sommers et al., 2012; Thornton et al., 2016; Wherry and Miller, 2016). While some policies have been effective, others have had mixed success or no impact at all (Gerber et al., 2005; Sommers et al., 2012; Thornton et al., 2016; Wherry and Miller, 2016). Mixed success likely reflects an incomplete understanding of the full set of factors

[8]This idea relates to Stafford and Wood (2017), a public health paper that discusses how health behaviors among homeless populations that can appear careless can often be explained by the need to prioritize food and shelter.

underlying differences in how people invest in their health, which thwarts efforts to effectively shift behavior and outcomes. We argue that taking account of a broader set of factors, such as relationships between health behaviors and labor market conditions and choices, could lead to more effective policy. To illustrate, behaviors that maximize health may also interfere with work, which means people may need to compromise their health in order to maintain their economic well-being (Cawley and Ruhm, 2012; Gilleskie, 1998).[9] Policies that help to break the link between medication side effects and difficulties working may be worthwhile social investments.

Our approach to studying health decisions has its origins in the view that health is a form of human capital in which individuals invest through their choices (Grossman, 1972; Becker, 2007). This framework is useful to understand health-related behaviors as it posits that individuals make health investments until the marginal costs of doing so exceed the benefits. Hence, risky health behaviors such as not using effective medications can be seen as *disinvestments* in health (Cawley and Ruhm, 2012). An implication is that rational individuals who face higher costs of health investments will exhibit worse health, and policies that lower these costs could reduce resulting health disparities by encouraging health investments (Gilleskie, 1998).

One way to operationalize the (Grossman, 1972) idea of treating health as human capital is to build a dynamic structural lifecycle model in which choices affect health (Arcidiacono et al., 2007; Chan et al., 2016; Chan and Hamilton, 2006; Crawford and Shum, 2005; Cronin, 2019; Cronin et al., 2024; Darden, 2017; Gilleskie, 1998; Papageorge, 2016). We follow this tradition. We are most closely related to papers applying insights from Grossman (1972) to health behaviors during the HIV epidemic (Chan et al., 2016; Hamilton et al., 2021; Lakdawalla et al., 2006; Papageorge, 2016). A goal of these papers is to investigate how medical innovation affects behavior and welfare. We use the same setting as in Papageorge (2016) and some of the modeling choices are also similar. However, our purpose is markedly different. While Papageorge (2016) aims to link non-uptake of effective medicine to labor supply and side effects, the goal of this paper is broader: to investigate distributional differences in the dynamic tradeoffs that characterize health decisions, along with the distributional consequences of medical innovation, which includes the idea that medical innovation can reinforce existing health disparities. A way to investigate these dynamics is by allowing the wedge between health- and welfare-maximizing choices to vary across sociodemographic groups.

The rest of the paper proceeds as follows. Section 2 describes the data and presents key empirical patterns. Section 3 presents the structural model and describes the estimation. Section 4 provides results on the value of innovation. Section 5 concludes.

[9] Our paper also relates to recent work (see, e.g., (Jones and Klenow, 2016)) in macroeconomics showing that cross-country differences in GNP may over- or understate differences in economic welfare once disparities in health and leisure are taken into account.

2 Data and Descriptive Patterns

We introduce the data set used in the project, discuss key variables, and provide summary statistics. Next, we conduct a preliminary data analysis and use the resulting relationships between variables to inform how we specify the structural model, introduced in Section 3.

2.1 The MACS Data Set and Summary Statistics

Data for this paper come from the Multi-Center AIDS Cohort Study (MACS), an ongoing semi-annual study beginning in 1984 following a sample of men who have sex with men in four U.S. cities: Baltimore, Chicago, Pittsburgh, and Los Angeles.[10] The initial enrollment of the MACS included 4,954 men. Our analysis focuses on the roughly half of them who are HIV+ and uses data starting in 1991 through 2003. After removing observations with missing data, our analysis sample consists of 1,201 individuals comprising 11,290 observations across 13 years.[11] Our timing convention is as follows. State variables describe individuals as they enter period t from period $t-1$. Choices and outcomes (e.g., treatment options and side effects) affect people between periods t and $t+1$. State variables have the subscript $t-1$ while choice and outcome variables have the subscript t. We observe 432 deaths. In such cases, we observe state variables describing period $t-1$ to t , along with death, after which we do not observe choices and outcomes for the period t to $t+1$. This leaves us with 10,858 observations in which we observe state variables along

[10]Data in this manuscript were collected by the Multicenter AIDS Cohort Study (MACS). MACS (Principal Investigators): Johns Hopkins University Bloomberg School of Public Health (Joseph Margolick, Todd Brown), U01-AI35042; Northwestern University (Steven Wolinsky), U01-AI35039; University of California, Los Angeles (Roger Detels, Otoniel Martinez-Maza, Otto Yang), U01-AI35040; University of Pittsburgh (Charles Rinaldo, Lawrence Kingsley, Jeremy Martinson), U01-AI35041; the Center for Analysis and Management of MACS, Johns Hopkins University Bloomberg School of Public Health (Lisa Jacobson, Gypsyamber D'Souza), UM1-AI35043. The MACS is funded primarily by the National Institute of Allergy and Infectious Diseases (NIAID), with additional co-funding from the National Cancer Institute (NCI), the National Institute on Drug Abuse (NIDA), and the National Institute of Mental Health (NIMH). Targeted supplemental funding for specific projects was also provided by the National Heart, Lung, and Blood Institute (NHLBI), and the National Institute on Deafness and Communication Disorders (NIDCD). MACS data collection is also supported by UL1-TR001079 (JHU ICTR) from the National Center for Advancing Translational Sciences (NCATS) a component of the National Institutes of Health (NIH) and NIH Roadmap for Medical Research. The contents of this publication are solely the responsibility of the authors and do not represent the official views of the National Institutes of Health (NIH), Johns Hopkins ICTR, or NCATS. The MACS website is located at http://aidscohortstudy.org/.

[11]Specifically, we start with the full MACS sample of 139,288 observations for 7,175 individuals (including refresher samples). After restricting the sample to survey visits in the time period 1991–2003, we are left with 48,644 observations for 5,057 individuals. After removing HIV negative individuals and people with missing HIV status, we are left with 21,746 observations comprising 2,291 individuals. Removing individuals outside of the age range 30–65 leaves 20,937 observations from 2,185 people. We drop 1,098 observations from 390 non-white individuals from a refresher sample to the panel due to the sampling methodology used to select these individuals. Dropping observations with missing data leaves 11,352 observations for 1,203 individuals. Finally, we remove 62 observations with extreme outlier values for costs or who are recorded as using HAART before it was invented. This leaves 11,290 observations for 1,201 individuals.

with choice and outcome variables. Summary statistics for state variables thus use the sample of 11,290 observations, while choice and outcome variable statistics are computed for the 10,858 observations where we do not observe death prior at period t.

We now describe how we construct key variables used in our analysis. The MACS collects information on a rich set of health and labor market behaviors and outcomes. Moreover, the panel structure of the data set allows us to link behavior and outcomes during different phases of the AIDS epidemic, characterized by shifts in the characteristics of available treatments due to medical innovation. Beginning with treatment choices, MACS respondents report all medications (including HIV treatments) they have used in the six months leading up to each visit. While there are dozens of different HIV drugs, we follow Detels et al. (2001) and combine them into three categories: mono-therapy, combo-therapy, and HAART. Mono-therapy consists of a single medication, usually a nucleotide reverse transcriptase inhibitor (or NRTI, of which the drug AZT is an example). Combo-therapy is a set of medications, usually NRTIs. HAART is a particular set of medications that generally includes a mixture of NRTIs and a class of drugs, known as protease inhibitors (PIs). The key innovations that constituted HAART were the introduction of PIs and their combination with existing NRTIs. HAART was a major medical breakthrough. It was substantially more effective compared to earlier HIV treatments at improving health and lowering mortality. Since its emergence, HAART has been the standard treatment for people with HIV. In our analysis, each treatment category (along with the option of no treatment) is characterized by its likelihood to improve health, its propensity to lead to ailments (due to symptoms of illness or side effects of treatment), along with its average out-of-pocket costs, which respondents report.

Turning to key health measures, the MACS data contain an objective measure of immune system health known as the CD4 count, which is the number of white blood cells in a cubic meter of blood and is a common measure of immune system functionality. Typical CD4 counts range between 500 and 1,500 (with an average of 1,000) and HIV+ people are designated as having AIDS when counts fall below 250, which signals risk of an inability to recover from routine infections. Similar to Chan et al. (2016) and Papageorge (2016), health is defined as an indicator of AIDS-level CD4 count, above or below 250. Below this threshold of immune system health, mortality rates spike among individuals in our sample. We also examine patient reports of physical ailments. An individual is coded as suffering from physical ailments if he reports one of the following for at least three days since their previous semi-annual interview: fatigue, diarrhea, headache, fever, or drenching sweats.[12] As we explain below, when introducing the model, we will conceptualize ailments as arising from a combination of symptoms of illness and side effects of treatment.

Finally, the MACS data contain labor market information, including employment and income,

[12] In previous work, results are robust to allowing ailments to vary by type or frequency (Papageorge, 2016).

along with the highest completed level of education. These measures allow us to link labor market outcomes to treatment choices and health outcomes. We focus on binary employment decisions: individuals work full-time or not.[13] Work experience accumulated prior to the beginning of the survey equals the individual's potential experience: the number of years since graduation age given education status. After the beginning of the survey, work experience is obtained using the observed employment history. Income, which we convert to year 2000 dollars, is a categorical variable that grows in increments of \$10,000, with the highest value being "\$50,000 or more". Moreover, information on education means we can examine differences in the interaction between health and work across two socioeconomically distinct groups. To do so, in both our preliminary descriptive analysis and in our structural model, we distinguish between a higher-education group, consisting of respondents who obtained a four-year college degree or more, and a lower-education group, containing everyone else (which we refer to in tables and figures as *College*+ and <*College*, respectively).[14]

Summary statistics for our analysis sample are presented in Table 1 for seven different groups (labeled from (1)-(7) in the table). They are: (1) the full sample, (2) pre-HAART, (3) post-HAART, (4) <College, pre-HAART, (5) <College, post-HAART, (6) College+, pre-HAART, and (7) College+, post-HAART. The post-HAART era is set to start with survey visit 24, roughly the second semester of 1995.[15] In the full sample, 63% of the men have a college degree or more, and the average age is 44.[16] Individuals with less education are, on average, two years younger than those with a college degree. Due to the panel nature of the data, the population ages over time, with the average age increasing by 5 years between the pre- and post-HAART eras.

Table 1 also shows that treatment usage is 71% for the sample period, but is higher for people with more education and increases after HAART is introduced. For the less educated group, treatment usage increases from 54% to 75% and for the more highly educated group, it rises from 57%

[13]Individuals working part-time are classified as not working, as very few respondents work part time and their average incomes are comparable to those who do not work. Results do not change if we recode part-time workers as full-time workers or simply drop them from the analysis.

[14]The MACS data set contains several variables that in principle could provide additional supporting evidence (e.g., occupation, whether individuals had stopped medication due to side effects, or whether individuals had changed jobs due to HIV). However, data on occupations is coarse and measured only once. Data from the other questions are not answered enough to reliably draw conclusions (and do not get at the idea of stopping work due to side effects). In results available from the authors, we show that lower-education individuals are more likely to be in "manual" occupations, to have changed jobs due to HIV, and to have stopped medications for side effects. However, the magnitudes are small, and in general, the infrequency of answers makes it difficult to rely on these data either to generate moments to be matched or as strong supporting evidence.

[15]Results are robust to treating either survey visit 25 or 23 as the first post-HAART visit.

[16]College degree summary statistics are computed by individual and all individuals are observed at least once in the pre-HAART period, so 63% is also the proportion of pre-HAART college educated. Among individuals who survive to the post-HAART period, 68% have a college degree. If we compute college degree by observation, the proportion in the full sample and the pre- and post-HAART eras are 69%, 65% and 72%, respectively, reflecting selection into survival and generally more observations by education.

TABLE 1: Summary Statistics by Education and Era

	(1) Full Sample	(2) Pre-HAART	(3) Post-HAART	(4) <College pre-HT	(5) <College post-HT	(6) College+ pre-HT	(7) College+ post-HT
College +	0.63	0.63	0.68	0.00	0.00	1.00	1.00
Age	44	41	46	40	45	42	47
Treatment	0.71	0.56	0.82	0.54	0.75	0.57	0.85
Monotherapy	0.15	0.32	0.01	0.33	0.01	0.32	0.01
Combotherapy	0.20	0.24	0.17	0.21	0.16	0.26	0.18
HAART	0.36	.	0.64	.	0.58	.	0.66
Death	0.04	0.07	0.02	0.09	0.02	0.06	0.01
CD4 Count	452	396	499	377	485	406	504
High CD4 (≥ 250)	0.74	0.66	0.82	0.62	0.78	0.68	0.83
MOOP ($/half year)	245	185	293	159	210	198	326
Insurance	0.95	0.93	0.96	0.89	0.93	0.95	0.97
Ailments	0.40	0.41	0.39	0.44	0.42	0.39	0.39
Income ($/half year)	19938	19949	19929	15373	14780	22290	21902
Full Time Work	0.66	0.69	0.64	0.61	0.56	0.72	0.68
HAART if FT Work	.	.	0.61	.	0.53	.	0.64
HAART if not FT	.	.	0.69	.	0.64	.	0.71
Observations ($t-1$)	11290	5165	6125	1785	1708	3380	4417
Observations (t)	10858	4825	6033	1633	1672	3192	4361

Notes: The table provides descriptive statistics for our analytic sample of 1,201 individuals. There are 11,290 person-visit observations in total and 10,858 observations in which the individual does not die from one period to the next. The pre-HAART era contains observations from 1991 until mid-1995 (9 half-year periods), the post-HAART era contains observations from mid-1995 until mid-2003 (18 periods). Each entry represents the mean over person-visit observations for the given time period, except for education, which is measured once per person. Measures are proportions, except age, CD4 count, income, and medical expenditures (the latter two of which are in year 2000 dollars per half year). Given our timing conventions, state variables (college, age, death, and CD4 count) describe individuals until the end of period $t-1$. Summary statistics are thus computed for the larger sample that includes individuals who have died by period t. All other variables (including treatment choices, insurance, out-of-pocket payments, ailments, income, and work) describe behavior starting at period t and so summary statistics are computed for the 10,858 observations in which the individual is observed at $t-1$ and is alive in period t. Sample sizes for each set of observations are also found in the last two rows of the table.

to 85%. When available, average HAART use was 58% for the lower-education group and 66% for the higher education group (a difference of 14% or 8 percentage points). Uptake of HAART was accompanied by improvements to health and lower mortality, which we see for both education groups. However, HAART did not close health gaps between education groups. The probability of a CD4 count above 250 increases from 62% to 78% and from 68% to 83%, for the lower- and higher-education groups, respectively.

While HAART led to massive health improvements and its uptake was substantial, use of HAART was not universal, meaning that improved health, measured as CD4 count, was likewise not universal. We observe lower use (and worse health) among the less educated group. Out-of-pocket costs are a natural explanation for non-use of an effective, indeed, life-saving medical treatment. However, these costs are generally fairly low, in part because most respondents are insured. The post-HAART increases in out-of-pocket costs we observe in Table 1 reflect higher use of treatment. However, as we discuss below, the cost of HAART is not only low in general, but low compared to other treatments, and remains low among the uninsured, who can obtain medications directly from doctors or clinics. We thus consider a different factor, side effects, which manifest as physical ailments. The initial evidence that side effects might play a role is perhaps ironically the fact that, according to Table 1, reports of ailments remain fairly stable across time and within education groups in the pre- versus post-HAART eras. This stability reflects two countervailing dynamics. Higher medication usage leads to an increase in side effects (increasing reports of ailments) along with better health and fewer symptoms of illness (decreasing reports of ailments)—and these dynamics essentially cancel one another out.

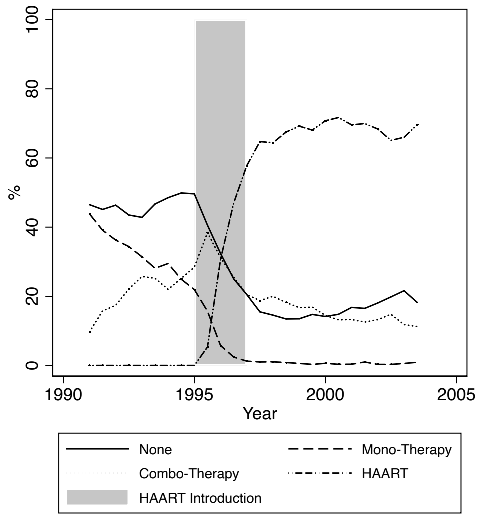
Ailments induced by side effects could discourage uptake of HAART. They can also affect behavior and welfare through labor supply, and we thus explore how HAART use relates to labor supply. In Table 1, we see that full-time work declined with the emergence of HAART, though the cohort we examine is aging towards retirement as time passes. We thus adjust for age in our analyses. Table 1 also shows a negative relationship between full time work and HAART usage in periods when HAART is available. These differences vary by education group. Among individuals with less education, 53% report HAART when working full time versus 64% when not working full time. Corresponding percentages for the higher education group are 64% and 71%. In other words, men with more education use HAART more, tend to work more, and also exhibit a relatively small HAART use-labor supply elasticity compared to individuals with less education (7 percentage points versus 11 percentage points). These statistics provide initial evidence of differences by group in the health-work tradeoff. In further analyses below, we explore these interactions to better assess how treatment use, ailments, and labor supply relate.

2.2 Preliminary Data Analysis

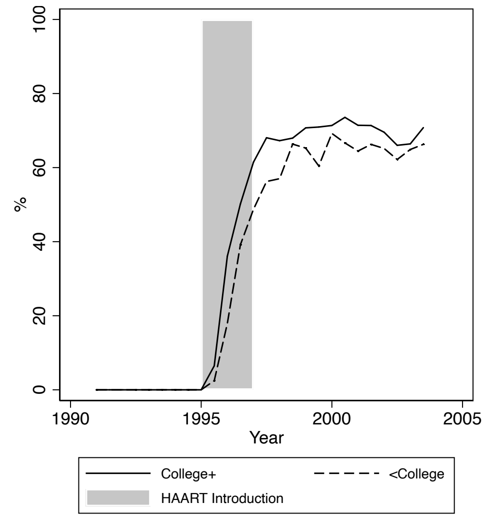
In what follows, we examine in greater depth key sets of variables related to treatment choices, health, and labor. We note that some tables discussed in this section are in Appendix A.

2.2.1 Treatment Choices

The introduction of HAART is associated with large changes in the treatment choices of HIV+ individuals. These shifts are illustrated in Figure 1a, which shows treatment choices over time. The



(A) Treatment



(B) HAART

FIGURE 1: Treatment and HAART Use Over Time. Panel (A) shows the probability of using each treatment in each period. Before HAART was introduced, many individuals did not use any treatment. HAART was adopted quickly after it was introduced, causing a decrease in the use of mono-therapy and in the share of individuals using no treatment. Panel (B) shows the probability of using HAART in each period by educational attainment. HAART was adopted more quickly by individuals with more education, and is used at higher rates by those individuals throughout the study period.

vertical bar (1995–1996) indicates HAART introduction.[17] Prior to the introduction of HAART, a large portion of the sample chose no treatment at all (46% of the lower-education and 43% of the higher-education group, using summary statistics from Table 1). Of patients who use a treatment in the pre-HAART era, most choose mono-therapy. Starting in the late 1980s, recommended treatment regimens tended to include multiple drugs, i.e., combo-therapy. There is little evidence (in this paper or elsewhere) that combo-therapy is more effective compared to mono-therapy, but it does induce more side effects. After HAART introduction, HIV+ men substitute away from other treatments and from the no-treatment option. Figure 1b also shows treatment use over time, but focuses on HAART use split by education group. The figure corroborates that those with less education use HAART less often compared to those with more education (compare the dashed line to the solid line) and suggests that a slower adoption of HAART among the less educated helps to explain lower overall use.

Table 2 provides further insight into patterns of use and non-use of HAART, both in general and across education groups. Higher-education individuals are more likely to ever be observed using

[17] Since interviews were staggered it is not possible to pinpoint at which exact survey visit individuals first had access to HAART.

treatment compared to lower-education people (88% versus 81%). They also spend more time on medication: higher-education individuals use medication in 72% of survey visits compared to 63% for lower-education people. Similarly, the higher-education group is more likely to use HAART at least once in the sample period and to spend more time over the sample period using it. The average survey visit at which those with a college degree start HAART is 27.25 (roughly the first semester of 1997). The corresponding number for those with less than a college degree is 28.07, approximately 5 months later. The difference in the probability of switching treatment in a given visit by education group is not statistically significant. However, those with a college degree are marginally more likely to have stopped HAART. Among those who ever stop HAART, 80% of those with less than a college degree and 73% of those with more restart (though these differences in means are not significant). Together, these patterns suggest that lower HAART usage among lower-education individuals arises primarily because they are less likely to start using HAART.

TABLE 2: Medication Spell Characteristics by Education for Analysis Sample

	<College	College +	p-value
Ever use treatment	0.81	0.88	0.002
Share of survey visits using treatment	0.63	0.72	0.000
Ever use HAART	0.67	0.81	0.000
Share of survey visits using HAART (when available)	0.48	0.58	0.001
First used HAART	28.1	27.3	0.025
Treatment transitions (per visit)	0.17	0.17	0.718
Ever stopped HAART	0.21	0.27	0.060
Ever started HAART after stopping	0.17	0.20	0.278
Ever started HAART after stopping (conditional on stopping)	0.80	0.73	0.363

Notes: This table considers individuals' treatment patterns over time (one observation per person for 1,201 individuals). Treatment in this context means using one of the three medications. "Ever use treatment" indicates whether the individual is observed using treatment during the sample period. The first visit where HAART was available is 24, so an individual first using HAART in visit 27 started approximately a year and a half after it was introduced. "Ever stopped HAART" indicates that the individual was observed using HAART and then later observed not using HAART.

Further evidence of medication usage patterns is found in Table A1 in Appendix A, which shows the transition matrix for treatment choices by education level and HAART era. In general, we see a high degree of persistence over time. For example, over 80% of individuals who are not using treatment will continue not to use treatment in the following period. Pre-HAART, 70% or more of those taking a given medication continue with the same medication in the following period, with a substantial minority switching between treatments. In the post-HAART era, however, many individuals switch from mono-therapy, combo-therapy, or no therapy at all to adopt HAART. Relatively few individuals go off treatment in any given period. In general, this transition matrix suggests it may be costly to switch treatments or to go on or off of treatment, which prompts the inclusion of switching costs in the model. Moreover, persistence suggests that while individuals switch treatments with some probability in each period (which leads to a substantial amount of ever switching over the sample period), we do not observe massive changes from one period to the next (the exception being the rapid uptake of HAART when it was introduced).

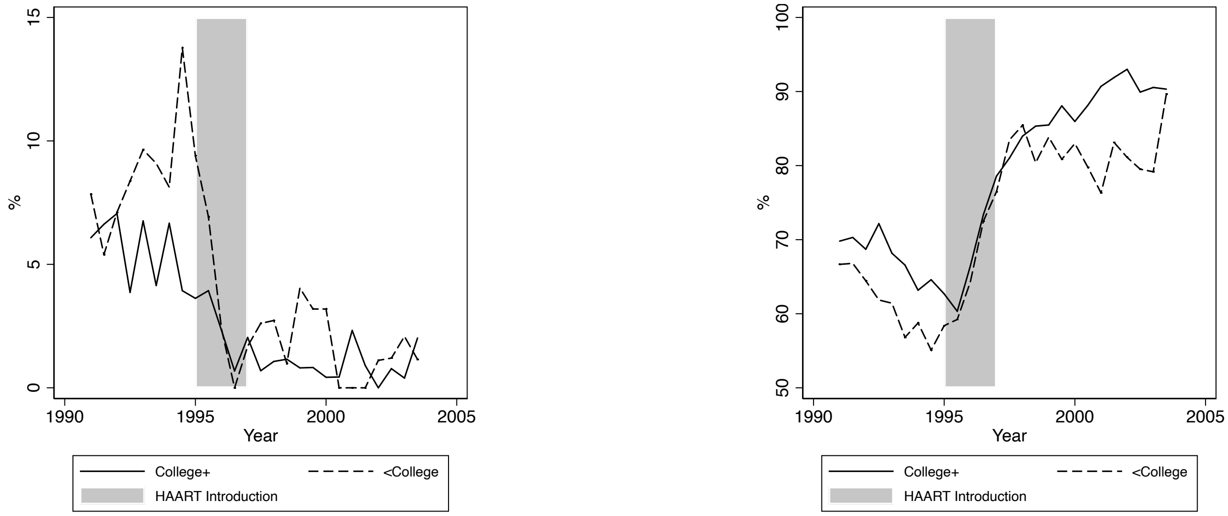
Finally, to investigate how different key variables relate to treatment choices, we estimate a multivariate logistic regression to explain the use of "any treatment." Here and throughout this section, we present descriptive regression tables that show estimates controlling for education and whether HAART was invented, and then separately by era and education group. In addition to the coefficients presented, these regressions include a second-order polynomial in age, along with an era-specific time trend. These models are similar to models used to recover parameters governing outcomes and state-to-state transition probabilities in the structural model.[18] Results from descriptive regressions are in Appendix A.[19] Estimates in Table A2 show that even after controlling for factors including health and employment, individuals with less education are less likely to use treatment. Overall, medication use is significantly lower before HAART is introduced. Unsurprisingly, individuals already in good health are less likely to use treatment across education categories and eras, though this effect is smaller after the introduction of HAART. In addition, we find that those who are working are also less likely to use treatment. Moreover, the negative relationship between work and treatment usage is larger in magnitude for those with less education, suggestive of a stronger tradeoff between health and work for this group.

2.2.2 Health and Mortality

We next investigate HAART, health, and mortality. The introduction of HAART drastically reduced mortality and improved health. Moreover, it did so for both education groups. Table 1 shows that the probability of dying in a given six-month period is higher for those with less education, regardless of whether HAART is available. However, HAART introduction coincided with drops in mortality from 9% to 2% for those with less than a college degree and from 6% to 1% for those with a college degree. Figure 2a shows how quickly mortality changed after HAART emerged. The figure depicts a precipitous drop in the probability of dying, which is evident for both education groups. It also shows that mortality is consistently higher for those with less education in both the pre- and post-HAART eras.

[18]The structural model is more parsimonious. In processes governing outcomes and transitions, we omit some interactions because they are insignificant but lead to small changes in results. Models containing all interactions do not appreciably or qualitatively change results. In some descriptive regressions presented in this section, coefficients on interactions we report are insignificant. However, we note that significance depends on control sets and how we stratify the sample. One reason is that some variables we add as controls capture mechanisms through which other variables relate to outcomes. For ease of exposition, when reporting results of descriptive regressions, rather than show multiple specifications with varying levels of significance, we alert readers when interactions are insignificant.

[19]Structural model parameter estimates are presented in Appendix B.



(A) Death

(B) High CD4

FIGURE 2: Death Probability and CD4 Level Over Time. Panel (A) shows the probability of death in each period by educational attainment. Individuals with less education are generally more likely to die, especially before the introduction of HAART. Panel (B) shows the probability of having above AIDS level CD4 (≥250) in each period by educational attainment. Individuals with less education are generally less likely to have high CD4 before and after the introduction of HAART.

Consistent with lower mortality rates, HAART introduction led to higher CD4 counts. From Table 1, overall, the average sample CD4 count is 452, well above 250, the threshold for transition from HIV to AIDS. For both education groups, CD4 counts are higher after HAART becomes available, increasing by 108 units for those with less education and 98 units for those with higher education. On average, those with less education are less healthy, but the differences across education groups are small compared to the differences across the pre- and post-HAART eras. These differences across time and groups are reflected in the aforementioned changes in the proportion of individuals above AIDS-level CD4 counts, which is higher after HAART was introduced and for those with a higher education.

To further understand the interactions between treatment and health, Table A3 presents coefficient estimates from logistic regressions for having a higher CD4 count in the current period when compared to the previous period. While Table 1 shows that less education is associated with worse health, those with less education are not significantly less likely to see health improvements. However, health improvements are much more likely after HAART was invented, and treatments are more effective at increasing health when HAART is accessible. In general, better health seems to predict health improvements. Somewhat surprisingly, however, the reverse is true after HAART,

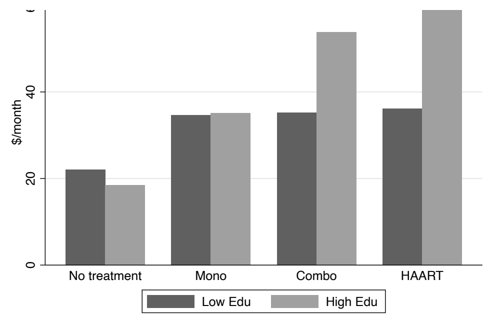


FIGURE 3: Out of Pocket Expenditure on Medication by Education and Treatment. This graph shows average expenditures per month on medications by treatment and education.

when a higher CD4 count predicts a lower likelihood of a one-period-ahead increase in CD4. This pattern reflects that CD4 is highly persistent and that HAART tends to function by maintaining higher CD4 counts or by increasing lower CD4 counts. As we explain when introducing the structural model, our approach to relating treatments to health will account for the complex relationship between treatment, CD4 count, and other outcomes.

2.2.3 Factors Predicting Non-Uptake of HAART: Prices and Ailments

A natural explanation for non-universal treatment usage is the financial cost of medications. However, out-of-pocket expenditures on treatments are relatively low, even for uninsured individuals (Gable et al., 1996). Table 1 shows that 95% of individuals are covered by health insurance in the sample, and even the group least likely to be insured, those with less education in the pre-HAART era, have an insurance coverage rate of 89%. While individuals in the sample do have some medical out-of-pocket expenditures, and these expenditures will be reflected in the model we estimate, Figure 3 shows that the level of expenditures is low compared to income. Individuals using treatments spend slightly more on medications, as do those with more education, which may reflect income-based subsidies (or the availability of medications using publicly financed programs (Bozzette et al., 1998). Across groups, average expenditures are less than $40 per month, suggesting that cost is not a major barrier to treatment use on average. Importantly, HAART is not meaningfully more costly to individuals compared to combo-therapy; indeed, for the lower-education group, out-of-pocket costs are similar for mono-therapy, combo-therapy, and HAART.

As out-of-pocket expenditures appear too low to explain non-uptake of HAART, we consider non-pecuniary costs: side effects, reported as physical ailments (fatigue, diarrhea, headache, fever, or drenching sweats). As we have mentioned, physical ailments can be symptoms of illness or side effects of medication. According to Table 1, 40% of the sample suffers from ailments at any given time. Those with less education are slightly more likely to suffer from ailments compared to

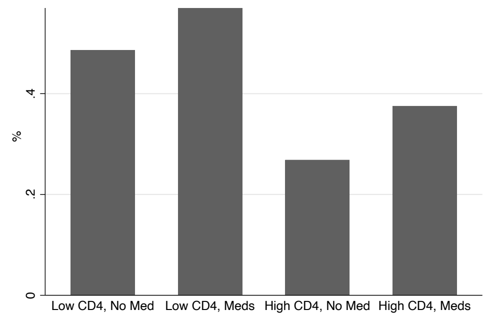


FIGURE 4: Ailment Probability by Health and Treatment. This graph shows the probability of experiencing ailments by CD4 count (above or below 250) and any treatment use.

those with a college degree (44% versus 39%, respectively). As discussed, the share of individuals suffering ailments does not differ significantly in the pre- and post-HAART eras, as decreases in symptoms were replaced by increases in side effects. Figure 4 shows the probability of experiencing ailments by treatment and health status. Consistent with the idea that ailments capture both symptoms of illness and side effects of medication, individuals using medication and with low CD4 counts are most likely to be experiencing ailments, i.e., more likely than both those with low CD4 and no treatment and those with high CD4 using treatment. The individuals least likely to experience ailments are those with high CD4 and no treatment use, who are suffering from neither medication-induced side effects nor symptoms related to AIDS.

To further understand what drives physical ailments, Table A4 in Appendix A presents coefficients from logistic regressions where the outcome variable is *not* suffering ailments. As with other regressions discussed in this section, we estimate the model for the full sample and then separately by education group and era. Across eras and education groups, treatment is associated with an increased risk of ailments. For all groups, a high CD4 count is negatively associated with ailments. These patterns provide support for the notion that ailments can be side effects of treatment, symptoms of illness, or both. There is no significant relationship between the HAART era and ailments, reflecting the countervailing dynamics discussed before (side effects replacing symptoms). Individuals with less education tend to report more ailments on average. Yet, there are differences in how treatment use and education interact to produce ailments. In the pre-HAART era, the relationship between treatment and ailments is stronger for the higher-education group, perhaps reflecting access to stronger medications or less skipping of medications. In the post-HAART era, however, the relationship between treatment and ailments is roughly the same across education groups, likely reflecting more standardized approaches to treatment. Put another way, treatment is associated with more ailments in the post-HAART versus the pre-HAART era, capturing that HAART was more effective but with stronger side effects compared to pre-HAART treatments. However, this pre-HAART versus post-HAART difference is larger for people with less education, suggesting they face larger downsides of HAART use compared to earlier treatments relative to individuals with more education.

2.2.4 Labor Market Decisions and Outcomes

Given that treatment decisions may affect or reflect employment decisions, we next consider what drives labor supply in the sample. On average, individuals work in 66% of the observed periods. Employment declines post-HAART for both groups as the cohort ages out of the labor market. In both the pre- and post-HAART eras, individuals with less education are approximately 11 percentage points less likely to be employed. A transition matrix (Table A5 in Appendix A) shows that unemployment is persistent. Among those with less education who work in a given period, 90.9% work in the following period; the corresponding number for higher-education individuals is 92.9%. Persistence in labor supply reflects persistence in health and treatment usage. While HAART introduction was quickly adopted and impacts on mortality were rapid, ailments remained similar as side effects replaced symptoms, and thus, we do not observe rapid or large movement in labor supply.

Regarding HAART use and labor supply, recall that Table 1 shows a negative relationship between full-time work and HAART use. Raw cross-tabulations show several points. Individuals with more education use HAART more and work more than people with less education. However, they also exhibit a smaller negative interaction between HAART and work. One interpretation is that individuals with more education face stronger incentives to work and invest in their health and, moreover, lower costs of working with side effects. For individuals with less education, there appears to be a stronger health-work tradeoff given the larger difference in HAART use by full-time work status. However, these cross-tabulations can reflect alternative factors, including different returns to work, health processes, and age patterns, among other factors. Therefore, we next explore the relationship between labor supply and medication use in a regression framework that allows us to control for such potential differences.

Table A6 in Appendix A presents coefficients from a logistic regression with employment as the outcome variable, controlling for a variety of factors. As before, we run the regression for the overall sample and separately by education groups and HAART era. Each regression includes health, ailments, treatment usage, the interaction between health and treatment usage, along with aforementioned controls: a second-order polynomial in age, along with an era-specific time trend. Individuals with less education are less likely to work and are less likely to work after HAART was introduced, even after controlling for age, i.e., beyond the impact of retirement. When we split the sample, we see that, across education groups and eras, a high CD4 count is associated with a higher probability of working. Ailments are associated with a lower probability of work, with a

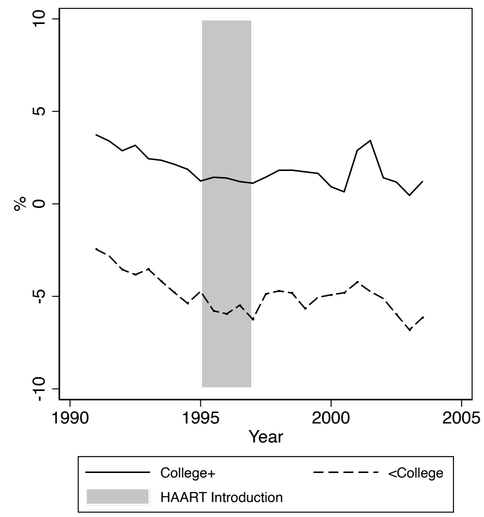


FIGURE 5: Residualized Income Over Time By Education. This graph shows average income over time by education, residualized on age and age squared.

stronger effect for individuals in the lower education group (though this interaction is insignificant depending on the specification and control sets). Before HAART was introduced, treatment use was associated with increased probability of employment, but this reverses in the post-HAART era. The negative relationship between HAART and work, even after controlling for health and ailments, is consistent with the idea that side effects manifesting as ailments are less costly when an individual is not working. This means that, when not working, individuals face stronger incentives to use HAART. The structural model introduced in the following section will formalize these incentives.[20] More generally, these patterns provide further preliminary evidence that working is especially difficult for those with less education who are experiencing ailments.

Evidence until now suggests that ailments from side effects can help to explain treatment and employment decisions. In particular, individuals may stop using effective treatment with side effects to work. We next ask whether income generated from work can help to explain these relationships. Individuals with less than a college degree earn, on average $15,373 and $14,780 in the pre- and post-HAART eras, respectively. Corresponding averages for the more highly educated group are $22,290 and $21,902.[21] Figure 5 shows income over time by education, residualized on age and age squared. According to the figure, incomes fall slightly for both groups post-HAART,

[20]After HAART is introduced, using medication is associated with a reduction in the probability of work. Treatment use interacted with ailments is associated with a reduction in employment and using treatment interacted with experiencing ailments is associated with a further reduction in the probability of work for those with less education. These latter relationships are not statistically significant when we split the sample as we have done here for ease of interpretation, but they are significant if we use the full sample with interactions.

[21]These numbers use averages of the upper and lower bounds of the reported income brackets, where we assume the upper bracket on those earning "more than $50,000" is $55,000. These amounts are then converted to year 2000 dollars.

likely because individuals age out of the workforce. However, the largest gap is across education categories and is persistent, which can help to explain different labor market behaviors, in particular, persistently lower labor supply for the lower-education group both before and after HAART was introduced. To further investigate income patterns, Table A7 in Appendix A presents linear regression results for income overall and by education group and era. In general, individuals earn more when they work and are healthy. Individuals with less education have lower incomes even after controlling for employment, experience, and health. When we split the sample to consider income by education group and era, the positive relationship between health and work is stronger in the post- versus pre-HAART era. This pattern is consistent with a negative relationship between treatment and work since sicker people are more likely to use treatment.

In general, the relationships presented here are difficult to analyze together if we rely solely on descriptive regressions absent a framework that posits how various pieces of suggestive evidence relate. The structural model described in the following section envisions labor supply and treatment decisions as being made jointly to maximize lifetime utility given their effects on health, mortality, income, side effects, and out-of-pocket payments. The model thus allows us to investigate these various health and work relationships in a unified framework.

3 Model

The empirical patterns in the previous section provide evidence of variation across education groups in health investments in the form of treatment use. HAART led to massive changes in medication usage and subsequent improvements to health and a precipitous drop in mortality. However, its uptake was not universal: it with lower among individuals with lower education, who adopted HAART later and generally exhibited lower rates of use over time. Further analysis of the data suggests that out-of-pocket monetary costs are too low compared to income to fully explain these patterns. We thus investigate a non-pecuniary cost: side effects manifesting as physical ailments. These ailments can lower utility directly. However, we also show that treatment use and ailments interact negatively with labor supply, with some evidence of stronger relationships for individuals with less education. These patterns are generally consistent with the idea that the labor market is a source of heterogeneity in health investments across demographic groups.

However, descriptive linear regressions make it difficult to understand the various dynamics at play and how they interact. Indeed, regressions showing a negative relationship between HAART use and work or between work and side effects suggest the presence of a novel mechanism that helps to explain health behavior differences and health disparities: disutility of working with side effects that may be stronger for low-education groups. However, this mechanism may be of negligible impact when compared to other factors that distinguish education groups, such as income or health differences, or how medication affects health. Empirical evidence of a potential interaction suggests a novel mechanism may be part of the story, but does not say how much.

Thus, in this section, we introduce a model designed to capture the various ways treatment choices interact with health, ailments, and labor supply. The model captures factors affecting choices to use medical treatments. While treatments can improve health and lengthen life, they cost money and may also have additional downsides, such as side effects, which can interfere with work. This leads to a tradeoff between health and work, which can vary across sociodemographic groups. We note that the model we develop in this section to capture this central tradeoff requires a set of assumptions, many based on available data, and thus has limitations that could be explored in future work. We mention some of these limitations when specifying the model and offer a more complete discussion in the Conclusion.

3.1 Timing and Central Tradeoffs

At every period, forward-looking, HIV+ agents in our model learn their state variables (specified below) and then maximize their expected lifetime utility by choosing what antiretroviral treatment to consume and whether to participate in the labor market. When making choices, they take into account both upfront costs and long-run benefits. Their treatment choices affect their future health, which in turn affects their survival and future income. Although potentially beneficial for their future underlying health, treatment consumption can generate physical ailments (side effects) and monetary costs (medical expenditures). The latter are also affected by insurance coverage, which is determined through a stochastic process that depends on labor market participation and health. Regarding labor supply, agents in the model can work to earn income, which they can use for consumption, and also to gain work experience, but may also incur a utility cost of work. Ailments are allowed to interact with this utility cost. When making these choices, agents form expectations (e.g., the upfront cost of suffering ailments and the long-run impacts of treatment on health). We generally assume rational expectations, which means agents' expectations conform to estimated population processes discussed below. However, agents do not anticipate HAART nor time trends we estimate to capture aggregate shifts in state variables.[22] Finally, we allow for heterogeneity by education in utility parameters along with processes governing health, ailments, income, out-of-pocket costs, and income. As specified, the model is able to capture several channels that explain differences in health investments, including: existing differences in health, the impact of treatment on health, side effects, the disutility of side effects, the disutility of side effects while working, and differences in income lost from not working. A stylized description of the timing of the model is

[22] Hamilton et al. (2021) relax this assumption and show that even with some reasonable expectation of technological improvement that is based upon prior innovations, HAART was a surprisingly large improvement. In the Conclusion, we discuss how future work could address different assumptions about agents' information.

presented in Figure 6.

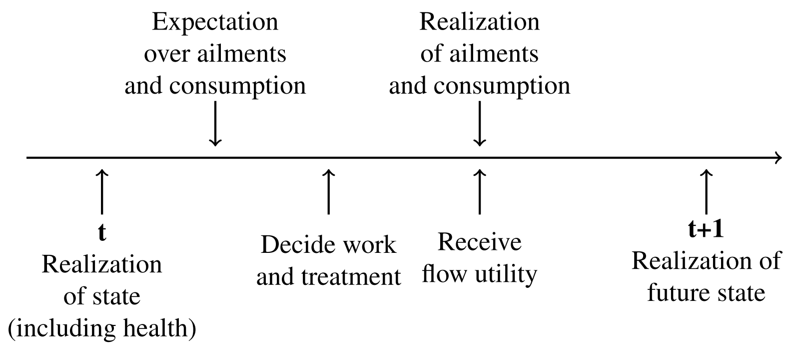


FIGURE 6: Timing of the Model. This figure shows the timing of the model, including state variables and decisions. Treatment and work choices are made simultaneously, not sequentially.

3.2 State Variables and Choices

Individuals are denoted by the subindex i. They enter period (semester) t with a vector z_{it} of state variables. z_{it} includes age a_{it-1}, which starts at $\underline{a} = 30$ and grows each period by 0.5 years until agents reach age $\bar{a} = 65$ after which agents receive a terminal value. z_{it} also includes completed education captured by the indicator $s_i \in \{0, 1\}$ that takes the value of 1 if they have college or more, labor market experience e_{it-1}, prior treatment decision d^m_{it-1}, and prior health status captured by the indicator $h_{it-1} \in \{0, 1\}$ that takes the value of 1 if their prior health was higher than AIDS level.[23] The health improvement variable Δh_{it} is an indicator constructed from the continuous measure of health (i.e., the CD4 count) and takes the value of one if $CD4_{it} \geq CD4_{it-1}$ and zero otherwise. In other words, the health improvement indicator captures health gains that can occur within a health state h_{it}. The period $\bar{t}$ at which HAART is introduced is the first semester of 1996. Aggregate components in the production of health and ailments, and in the processes for income and medical expenses are denoted ν_t.

At every period, individuals decide their labor market participation and their treatment. If they decide to work, the labor indicator d^l_{it} takes the value of 1. The introduction of HAART at $\bar{t}$ creates two treatment eras. In the first (pre-HAART) era ($t < \bar{t}$) there are three treatment alternatives: no treatment (d^m_{0it}), mono-therapy (d^m_{1it}) and combo-therapy (d^m_{2it}). In the second (post-HAART) era ($t \geq \bar{t}$) HAART becomes available as a fourth alternative (d^m_{3it}). Treatments

[23] While clearly restrictive, a binary variable for health captures how the effect of health on future survival, ailments, income and other outcomes is most starkly different between HIV+ individuals with AIDS and those without AIDS (Hamilton et al., 2021).

are mutually exclusive, and individuals must choose one. Hence, the collection of treatment-specific indicators $d^m_{rit} \in \{0,1\}$ satisfies $\sum_{r=0}^{2+\mathbf{1}\{t\geq\bar{t}\}} d^m_{rit} = 1$. The treatment decision vector $d^m_{it} \in \{0,1\}^{3+\mathbf{1}\{t\geq\bar{t}\}}$ collects all the treatment-specific indicators. Labor and treatment choices are made simultaneously. Therefore, there are six labor-treatment alternatives available at any $t < \bar{t}$ and eight at any $t \geq \bar{t}$. At every period, individuals observe their vector of alternative-specific preference shocks ε_{it} before making their choice. The preference shocks are distributed Type I Extreme Value (TIEV) and are independent and identically distributed across alternatives, individuals and over time.[24]

3.3 Processes Governing State Variable Transitions and Outcome Variables

3.3.1 State Variable Transitions

Health. Health h_{it} is determined in a two-step process. First, individuals draw a health improvement Δh_{it}, equal to one if CD4 increased or stayed the same, from a Bernoulli distribution with probability:

$$P[\Delta h_{it} = 1 | x^{\Delta h}_{it}] = \frac{exp(x^{\Delta h}_{it}\theta^{\Delta h})}{1 + exp(x^{\Delta h}_{it}\theta^{\Delta h})} \tag{1}$$

where $x^{\Delta h}_{it} \equiv [h_{it-1}, d^m_{it} \times h_{it-1}, a_{it-1}, a^2_{it-1}, s_i, \nu^{\Delta h}_t]$. The vector $x^{\Delta h}_{it}$ captures the efficacy of treatment alternatives d^m_{it} and its interaction with prior health and also permits variation by education and age. The scalar $\nu^{\Delta h}_t$ is an era-specific (pre-HAART and post-HAART) time trend that captures aggregate changes in health improvement at period t. Second, individuals transition into their next period health level according to:

$$P[h_{it} = 1 | x^h_{it}] = \frac{exp(x^h_{it}\theta^h)}{1 + exp(x^h_{it}\theta^h)} \tag{2}$$

where $x^h_{it} \equiv [h_{it-1} \times \Delta h_{it}]$. The vector x^h_{it} captures the effect of prior health and treatment (indirectly through the health improvement indicator) on the transition into future health. This two-step process parsimoniously captures the complex relationship between medication, CD4 count, and health outcomes (such as symptoms and mortality). Treatments function by increasing CD4 count from dangerous levels below 250, but also by maintaining CD4 counts that are in a safe range (on average between 400-500 for HIV+ individuals on effective treatment). Further, increases in CD4 count once an individual already has a non-AIDS-level CD4 count do not affect mortality or other

[24] We have experimented with different ways to allow for unobservable sources of dependence within individuals across time by allowing for multiple unobserved types in various processes (e.g., health and income) and find little impact on results and evidence that such types are not well identified. To capture dependence across time, we introduce "switching costs" in the utility function.

outcomes, i.e., the relationship between CD4 and observed health outcomes is highly non-linear. The first step in the process thus captures whether treatment leads to increases in CD4 count or whether it maintains it. The second step captures whether a non-AIDS-level CD4 count in the following period is due to a currently high CD4 count or a low CD4 count coupled with an increase due to treatment. While other specifications are possible, our parsimonious approach prevents us from misleadingly categorizing as ineffective treatments that function largely by keeping people healthy or by improving poor health, even though they may not necessarily improve health for HIV+ individuals who already have a high CD4 count.

Survival. Health has a direct effect on survival. At the beginning of every period, individuals face death with probability:

$$P[b_{it} = 1 | x_{it}^b] = \frac{exp(x_{it}^b \theta^b)}{1 + exp(x_{it}^b \theta^b)} \tag{3}$$

where $x_{it}^b \equiv [h_{it-1}, a_{it-1}, h_{it-1} \times a_{it-1}, s_i, v_t^b]$. Survival depends on the individual's health, age, and education level, as well as aggregate trends in survival (v_t^b). Current treatment consumption affects survival into next period indirectly through its effect on next period health h_{it}.

Additional state variables. Individuals enter the model with work experience calculated as potential experience given their age and education. In the model, agents endogenously accumulate work experience e_{it} in half-year increments as a function of their labor market participation:

$$e_{it} = e_{it-1} + 0.5 d_{it}^l \tag{4}$$

Age $a_{i,t-1}$ increases each period (semester) by one-half year. Hence, $a_{it} = a_{it-1} + 0.5$. The previous treatment choice, which is part of the state variable vector, is simply the realized previous choice. Education is pre-determined in the model and does not change over time. While the individuals are surprised by the aggregated components v_t, we model these components as following era-specific (pre- and post-HAART) linear time trends.

3.3.2 Outcomes

Depending on their labor and treatment choices, at every period individuals realize their ailments, income, insurance coverage, and medical expenditures. The vector y_{it} collects these outcomes.

Ailments. Individuals can suffer physical ailments for two main reasons: toxic treatments with strong side effects or symptoms caused by poor health. The production of ailments captures both these sources. Denote $y_{it}^{ailments}$ as the no-ailments indicator that takes the value of 1 if the individual does *not* suffer from ailments in period t. The probability of not suffering ailments is given by:

$$P[y_{it}^{ailments} = 1 | x_{it}^{ailments}] = \frac{exp(x_{it}^{ailments}\theta^{ailments})}{1 + exp(x_{it}^{ailments}\theta^{ailments})} \tag{5}$$

where $x_{it}^{ailments} \equiv [h_{it-1}, d_{it}^{m}, a_{it-1}, a_{it-1}^{2}, s_i, v_t^{ailments}]$. The vector $x_{it}^{ailments}$ captures the side effects of treatment alternatives and the symptoms generated by poor health. It also controls for education and aging. The scalar $v_t^{ailments}$ captures aggregate trends in side effects.

Income. Their labor market participation and experience determine their income according to:

$$y_{it}^{income} = x_{it}^{income}\theta^{income} + \varepsilon_{it}^{income} \tag{6}$$

where $x_{it}^{income} \equiv [d_{it}^{l}, h_{it-1}, a_{it-1}^{2}, a_{it-1}, e_{it-1}, e_{it-1}^{2}, s_i, s_i \times d_{it}^{l}, v_t^{income}]$ and $\varepsilon_{it}^{income}$ is an iid income shock with a conditional mean of zero. The biological processes of health evolution and aging thus affect individual productivity through a direct effect on income. These processes also affect income indirectly through the decision to participate in the labor market. The scalar v_t^{income} captures aggregate changes in income. By introducing labor market participation in our income process, equation (6) describes the individual income (labor and non-labor) for men who participate in the labor market and for those who do not.

Insurance coverage and medical expenditures. Health insurance in the model is a stochastic outcome that is realized after an individual makes his choice at t. An individual draws insurance coverage with probability:

$$P[y_{it}^{insurance} = 1 | x_{it}^{insurance}] = \frac{exp(x_{it}^{insurance}\theta^{insurance})}{1 + exp(x_{it}^{insurance}\theta^{insurance})} \tag{7}$$

where $x_{it}^{insurance} \equiv [h_{it-1}, a_{it-1}, a_{it-1}^{2}, e_{it-1}, e_{it-1}^{2}, d_{it}^{l}, s_i, s_i \times d_{it}^{l}, v_t^{insurance}]$. By controlling for labor market participation and experience, the vector $x_{it}^{insurance}$ captures the fact that health insurance is often employer-sponsored in the U.S. and that individuals with higher levels of experience may be more able to attain jobs that offer health insurance. Controlling for health helps to capture access through Medicaid, and controlling for age captures eligibility through Medicare.[25] The probability of insurance coverage also captures age, education, and health effects. The scalar

[25] In their Kaiser Family Foundation technical report, Dawson et al. (2023) show that "[m]ost Medicaid enrollees with HIV (...) qualify for coverage through a disability pathway". We are unable to control for geographic variation in public insurance provision as we do not have access to geographic data for individuals in our sample.

$v_t^{insurance}$ captures aggregate changes in insurance coverage. In turn, insurance coverage affects the amount of medical expenditures an individual pays out of pocket (MOOP) according to:

$$y_{it}^{expenses} = x_{it}^{expenses}\theta^{expenses} + \varepsilon_{it}^{expenses} \tag{8}$$

where $x_{it}^{expenses} \equiv [y_{it}^{income}, y_{it}^{insurance} \times d_{it}^{m}, h_{it}, y_{it}^{ailments}, h_{it} \times y_{it}^{ailments}, a_{it-1}, a_{it-1}^{2}, s_i, v_t^{expenses}]$ and $\varepsilon_{it}^{expenses}$ is an iid medical expenses shock with a conditional mean of zero. Importantly, treatment costs vary with insurance coverage; this is captured by the fact that the marginal effect of treatment consumption on out-of-pocket medical expenses depends on insurance status.

3.4 Utility and the Value Functions

Individuals use the entirety of their current income for consumption c_{it} and medical expenses. Hence, the individual's budget constraint for period t is given by:

$$c_{it} = y_{it}^{income} - y_{it}^{expenses} \tag{9}$$

There is no borrowing or saving in the model, which we discuss further in Section 5. We divide consumption, which is already in thousands of dollars, by 10 for ease of computation, i.e., in estimation, consumption is in tens of thousands of dollars.

Individuals with education level s draw utility from their consumption, their ailments, their treatment, and their employment choices according to:

$$\begin{aligned} u(y_{it}, d_{it}, d_{it-1}^{m}, s) = \prod_{f\in\{0,1\}} \Big[& c_{it} + \theta_{1sf}^{u} \cdot (1 - y_{it}^{ailments}) + \theta_{2sf}^{u} \cdot d_{it}^{l} \\ & + \theta_{3f}^{u} \cdot d_{0it-1}^{m} (1 - d_{0it}^{m}) \\ & + \theta_{4f}^{u} \cdot \left(1 - d_{0it-1}^{m}\right) \left(1 - (d_{it-1}^{m} \cdot d_{it}^{m})\right) \\ & + \theta_{5f}^{u} \cdot \left(1 - d_{0it-1}^{m}\right) d_{0it}^{m} + \varepsilon_{it}(d_{it}) \Big]^{\mathbf{1}[y_{it}^{ailments} = f]} \end{aligned} \tag{10}$$

As equation (10) suggests, we allow all of the utility parameters to vary by ailment status, and parameters θ_{1sf}^{u} and θ_{2sf}^{u} (which measure marginal utility of ailments by education) to vary by education level. Hence, people with different education levels in our model may experience different disutility from working while ill (or well). The flow utility captures direct utility from suffering ailments (θ_{1sf}^{u}), direct utility from work (θ_{2sf}^{u}) and switching costs for starting (θ_{3f}^{u}), changing (θ_{4f}^{u}) and stopping (θ_{5f}^{u}) treatment. The utility from not suffering ailments θ_{1s0}^{u} is normalized to zero for both education levels. The flow utility also contains the idiosyncratic, alternative-specific preference shock $\varepsilon_{it}(d_{it})$.

Value function. Recall that z_{it} denotes the observable part of the state vector. Upon reaching age $\bar{a}$, individuals no longer make choices and receive their terminal value equal to the monetary present value of perpetually receiving the flow utility obtained at $\bar{a}$; they compute this present value using their discount factor and their annual probability of survival, which remains fixed at its age $\bar{a}$ level. Let $z_{it}(a)$ be the state of an individual when he is of age a and let $\mathbb{K}_t$ denote the set of alternatives at period t. Hence, his conditional value function (net of the taste shock) from choosing alternative $k \in \mathbb{K}_t$ at age $\bar{a}$ is given by:

$$v_{kit}(z_{it}(\bar{a})) = \left(\frac{1}{1-\delta(z_{it}(\bar{a}))}\right) u_k(y_{it}, d_{it}, d^m_{it-1}, s) \tag{11}$$

where $\delta(z_{it}(\bar{a})) \equiv \beta(1 - P[b_{it} = 1 | x^b_{it}, age_{it} = \bar{a}])$ where β is the discount factor, and the term thus integrates the discounting that happens due to the annual possibility of death. For any age $a < \bar{a}$ his conditional value function is given recursively by:

$$v_{kit}(z_{it}(a)) = u_k(y_{it}, d_{it}, d^m_{it-1}, s) + \beta E_k[V_{it+1}(z_{it+1}(a+1)) | z_{it}(a)] \tag{12}$$

where E_k denotes the expectation of the state conditional on choosing alternative k. Given that the taste shocks are distributed TIEV, the ex-ante value function $V_{it}(z_{it}(a))$ is given by:

$$V_{it}(z_{it}(a)) = \gamma + \ln\left(\sum_{k' \in \mathbb{K}_t} \exp\{v_{k'it}(z_{it}(a))\}\right) \tag{13}$$

where γ is the Euler constant. If the agent dies, he receives no more utility, i.e., the value of death is zero. At any age $a \leq \bar{a}$, surviving individuals choose an alternative $k \in \mathbb{K}_t$ to solve the discrete maximization problem:

$$\max_{k \in \mathbb{K}_t} \{v_{kit}(z_{it}(a)) + \varepsilon_{kit}\} \tag{14}$$

3.5 Model Estimation and Identification

Let θ^{xy}, the vector that collects all parameters governing processes and transition probabilities, be $\theta^{xy} \equiv [\theta^{\Delta h}, \theta^h, \theta^b, \theta^{ailments}, \theta^{income}, \theta^{insurance}, \theta^{expenses}]$, and let θ^u be the vector of parameters of the flow utility function. We estimate the model parameters $\theta = [\theta^u, \theta^{xy}]$ following a nested procedure. In the inner step, given a set of proposed parameters, we use backwards induction to solve the dynamic programming problem for each set of observable state variables. This procedure generates choice probabilities that maximize utility given the parameters. The search algorithm in the outer step uses the probabilities generated by the inner step to search for the parameters that maximize the likelihood of the data. The likelihood contribution of each individual is:

$$L_i(\theta) = \prod_{t=1}^{T_i} P(d_{it}|X_{it};\theta) \times \prod_{t=1}^{T_i} f(X_{i,t+1}|X_{it}, d_{it}; \theta^{xy}) \tag{15}$$

where f denotes the density function derived from the processes. Because the log likelihood is additively separable, we estimate the processes separately from the utility parameters in a first step to reduce computational burden while retaining consistency. In the second step, we search for the utility parameters using the nested procedure described above.

Our first stage processes are identified by their data counterparts, e.g., we regress income, expenditures, and treatments onto the variables discussed previously. Flow utility parameters are identified from observed choice probabilities conditional on different sets of state variables. Additionally, HAART created quasi-experimental variation as it emerged in the middle of our sample period and led to variation in individuals' choice sets, which helps to identify preference parameters. In particular, HAART had different characteristics compared to existing medications, and the likelihood that individuals switched to it (or chose not to) helps to identify preferences governing the tradeoff between ailments and medication effectiveness. We assume that agent expectations about outcomes and transitions align with the population averages implied by the processes governing state variable transitions and outcome variables. Essentially, we are assuming rational expectations with the caveat that agents are surprised by the emergence of the treatment breakthrough (HAART) and also by the time trends in the state and outcome variables. Further, we assume the discount factor $\beta = \sqrt{0.95}$ (consistent with the semestral frequency of the data) and we normalize the parameter $\theta^u_{1,f=1}$ to zero.[26] Given data on transitions, outcomes, and choices, the choice of β, assumptions on expectations, the distributional assumption of taste shocks, and the flow utility normalization, Magnac and Thesmar (2002) show there is a unique parameter vector that maximizes the likelihood function.

To provide further intuition on identification of utility parameters, we consider the parameter governing the difference by education group in the utility of working while suffering ailments, which is a key preference that underlies some of our subsequent results on health differences by education group. Differences in medication usage have implications for ailments, health, mortality, and treatment costs. We account for these through our processes governing state variables and

[26]Given our setup, this means that the utility of not suffering ailments, not working, no consumption, and no switching costs is zero. As the value of death is also assumed to be zero (agents do not obtain more utility once they die), there is a "double normalization," which is not an attractive assumption. However, experimentation with alternative models in which we allow agents to obtain a non-zero (and potentially negative) utility the period prior to dying is not identified. If we assume the same value that is estimated in Chan et al. (2016) from a sample from the same data set (identified off risky sex decisions that potentially lead to HIV-infection), we find only negligible impacts on our results.

transitions. Additional differences load onto the utility cost of ailments. Lower use of medication when an agent is working (or, alternatively, lower labor supply if an agent is taking medication) identifies a higher utility cost of working while suffering ailments. Differences by education group in the aforementioned differences in treatment and labor supply decisions identify variation in the utility cost of work while suffering ailments across education groups.

3.6 Structural Parameter Estimates

This section presents estimates of preference parameters and parameters governing outcomes and transitions. Table 3 presents estimates of the utility parameters θ^u along with standard errors calculated using the delta method. $\theta^u_{1,f=0}$ represents the disutility of ailments for those with less than a college degree (-2.13), while the sum of $\theta^u_{1,f=0}$ and $\theta^u_{1,f=0,s=1}$ (-2.13-0.22=-2.35) represents the disutility of ailments for those with a college degree. The utility cost of ailments alone is not significantly different across education groups, though parameter estimates suggest a higher disutility of ailments for those with more education. Regardless of educational attainment, working without ailments enters positively in the utility function, which captures that agents (once we account for earnings) would prefer working to not working and may, on average, enjoy their job (rows 4 and 5).[27]

An important difference—and one that helps to drive subsequent results—is between the disutility of work while suffering from ailments between education groups (rows 6 and 7). For the college educated, working with ailments is far worse than working without ailments. However, the relationship is even more negative for individuals without a college degree. Given that the latter group earns less, this utility cost may not be overcome by earnings and could thus further incentivize lower-education individuals to work less or to take less medication. For those with less than a college degree, the disutility of working with ailments is -2.73 ($\theta^u_{2,f=0}$) (more than doubling the cost of ailments), while for those with more education, the cost is less, -1.97 ($\theta^u_{2,f=0}+\theta^u_{2,f=0,s=1}$), which adds to the disutility of ailments, but does not double it. This is in line with our descriptive evidence showing larger impacts of ailments on work for those with less education.[28]

In addition to flow utility parameters, we estimate parameters for the health, survival, and additional processes included in y_{it}. The estimated parameters for these processes are presented in Appendix B. Table B1 shows the estimates for the two-step health process. In the first step of the

[27] It is also possible that top-coding of income data means that a substantial portion of individuals are earning more than what the model imposes. If so, to rationalize work, the utility parameter for work may be biased upward. However, experimentation with higher income for individuals reporting the highest bracket do not affect main results on the tradeoff between health and work by education group.

[28] Remaining parameters capture, in general, stickiness in changing medication, which may be due to doctors' orders. In general, if any individual is not suffering ailments, the cost of switching or stopping treatments is very high, capturing how individuals are unlikely to change their medication if it happens to be working well for them.

TABLE 3: Estimated Structural Utility Parameters

Definition	Parameter	Estimate	Standard Error
No ailments	$\theta^u_{1,f=1}$	0	–
Ailments	$\theta^u_{1,f=0}$	-2.13	0.13
Ailments, college +	$\theta^u_{1,f=0,s=1}$	-0.22	0.14
Work, no ailments	$\theta^u_{2,f=1}$	1.03	0.06
Work, no ailments, college +	$\theta^u_{2,f=1,s=1}$	0.11	0.07
Work, ailments	$\theta^u_{2,f=0}$	-2.73	0.09
Work, ailments, college+	$\theta^u_{2,f=0,s=1}$	0.76	0.10
Start treatment, no ailments	$\theta^u_{3,f=1}$	0.25	0.49
Start treatment, ailments	$\theta^u_{3,f=0}$	-1.64	0.38
Change treatment, no ailments	$\theta^u_{4,f=1}$	-4.98	0.12
Change treatment, ailments	$\theta^u_{4,f=0}$	1.49	0.14
Stop treatment, no ailments	$\theta^u_{5,f=1}$	-9.39	0.47
Stop treatment, ailments	$\theta^u_{5,f=0}$	3.18	0.45

Notes: Parameters are for equation (10). $\theta_{1,f=1}$ is set to 0. f=1 means no ailments. Standard errors calculated using the delta method.

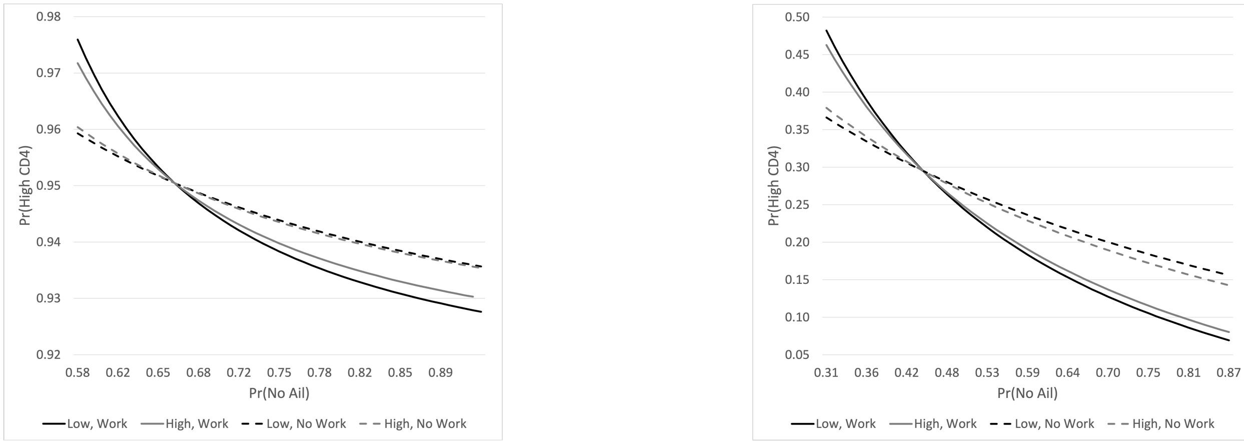
process, the estimated parameters suggest that all medications increase the probability of a health improvement, with HAART being the most effective. The difference between HAART and other medications is especially pronounced for individuals in poor health. Conditional on the health improvement, in the second stage of the process, both previous-period health and health improvements increase the likelihood of high current health. There is no statistically significant relationship between health and education, though the parameter estimate suggests a positive relationship between education and health improvements.

Table B2 shows parameter estimates for the death process. Individuals with a high CD4 count are much less likely to die, though that relationship weakens as individuals age. Before HAART, the aggregate trend in the probability of dying had a positive slope, but post-HAART, the slope is negative, capturing aggregate improvements in survival-inducing care for HIV+ individuals. The parameter estimates for the no ailments process are shown in Table B3. Individuals in good health are less likely to suffer ailments. All three treatment options reduce the probability of no ailments, but HAART produces more ailments than mono-therapy or combo-therapy. Furthermore, individuals with less education are more likely to suffer ailments.

Parameter estimates from the income process show that employment and education have strong positive relationships with income (Table B4). In addition, health generates higher income, and the aggregate trend in income is negative. Age and experience are also associated with higher income, though the returns are decreasing. Table B5 shows parameter estimates for the insurance process. Employment is associated with a significant increase in the probability of insurance coverage, as is higher education. A high CD4 count is associated with a decrease in the probability of insurance, perhaps capturing slight adverse selection. In all, the vast majority of those in the sample are covered by insurance. Medical expenditures increase with income and decrease when insured (Table B6). Unsurprisingly, both medication usage and ailments generate higher medical costs. In line with the results shown in Table 1, individuals with less education have lower medical expenditures, but the difference is not large. Out-of-pocket costs for HAART are lower for people without insurance, reflecting that medication can be obtained from publicly financed programs, e.g., clinics (Bozzette et al., 1998).

To understand differences in preferences by education group, work status, and CD4 counts, we plot indifference curves in Figures 7a (for men with high CD4 counts) and 7b (for men with low CD4 counts). In both figures, we consider the value of a 30-year-old HIV+ individual with 10 years of work experience at the first post-HAART visit and examine four different groups: high and low education workers and non-workers. For each group, when calculating value, we impose sample average probabilities that HAART induces a one-period-ahead high CD4 count and ailments. We then normalize this value across the four groups in each figure to be the same at those probabilities (where curves thus cross). Finally, using model parameters specific to each group, we increase the probability of no ailments and assess what change to the probability of a one-period-ahead high CD4 count would keep each kind of agent indifferent. The resulting curves show changes to the probability of no ailments (x-axis), accompanied by changes to the probability of a high CD4 count (y-axis) that leave agents indifferent.

In both figures and for all employment and education groups, the downward slope reflects that an improvement along one dimension of treatment quality leaves agents indifferent when accompanied by worse quality on the other dimension. The most noticeable pattern is the difference between the solid lines (for workers) versus the dashed lines (non-workers). The solid lines are steeper, which means that for a reduction in the probability of no ailments, agents need a larger improvement in the probability of a higher CD4 count to remain indifferent. The steepest line, for working agents with a low education, indicates less substitutability, which is in line with utility parameter estimates: in terms of lifetime utility, low-education workers must be compensated with relatively large improvements to treatment effectiveness to remain indifferent when facing a rise in the likelihood of ailments.



(A) High CD4 **(B)** Low CD4

FIGURE 7: Simulated Indifference Curves for Medications. This figure shows simulated indifference curves for medications that generate combinations of health and ailment probabilities. Panel (A) shows curves for agents with high CD4, Panel (B) shows curves for agents with low CD4. The simulated agent is 30 years old, has 10 years of work experience, and is on visit 26 (1996). We then consider four sub-groups of agents, distinguished by education and labor supply. Finally, we equalize value at average quality of HAART (where the curves thus cross). To create curves, we then plot which increases in the probability of a high one-period-ahead CD4 count accompanied by a higher likelihood of ailments keep an agent indifferent. In general, the steepest curve is for working, low-education individuals, who would need to be compensated relatively more in terms of drug effectiveness to accept an increase in the likelihood of side effects.

3.7 Model Fit

Overall, the model predictions match the data (Table B7) in Appendix B. Both the data and the model predict that agents will be employed 66% of the time. The results are similar for HAART use, with the model predicting a HAART rate of 33% and the data showing a HAART use rate of 36% (using the full sample, including when HAART was not available). The model continues to match the data reasonably well if we condition on health and education, especially the employment shares, although it over-predicts HAART use for the unhealthy and under-predicts use for the healthy.[29] The model does not capture how HAART use increases in the data for individuals in better versus worse health, and in fact shows the opposite pattern. One reason is that fit considers all periods, including the period corresponding to the introduction of HAART. In the data, HAART use increased rapidly, but the model setup means uptake could be immediate, which could lead the model to over-predict HAART for the low-CD4 count group. Switching costs and time trends

[29] In results available by request, we show that this pattern is robust to various specification changes such as the inclusion of additional or different switching costs, changes to the process inputs and alternate specifications of the utility from consumption. We also show that specifications including unobserved heterogeneity in the utility function produce very small probabilities of a second type (below 5%), evidence against the inclusion of unobserved heterogeneity in utility.

mitigate over-prediction by capturing persistence in choices and later improvements to HAART that do not emerge immediately when HAART is introduced. However, these same parameters that mitigate over-prediction of HAART for low-CD4 count individuals then lead us to under-predict HAART usage, especially in the first periods after its emergence, for the high CD4 count group.

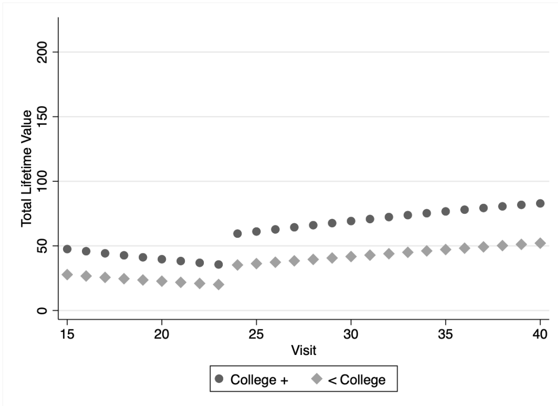
4 Results

This section uses the estimated model to generate a series of results, beginning with a quantification of the value of HAART along with a decomposition of its value for both education groups. Moreover, we consider a counterfactual HAART treatment mandate to illustrate differences across sociodemographic groups in the welfare consequences of being assigned an effective treatment with side effects that interfere with work.

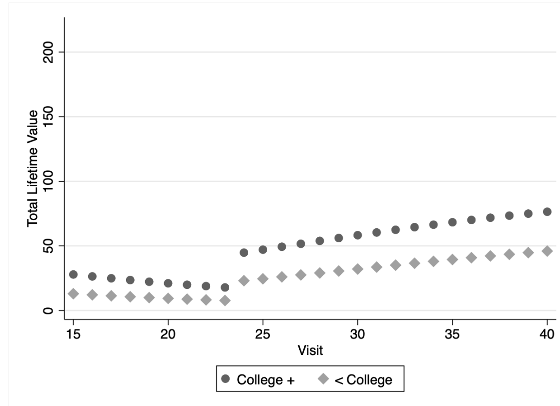
4.1 The Value of HAART

To illustrate the impact of HAART on welfare, Figures 8a and 8b show the total expected lifetime value for a 30-year-old individual on the best available treatment in each period by education status and for those with high and low CD4 counts, respectively. These figures show declining welfare prior to HAART as patients grew older and sicker, followed by a jump when HAART emerged. We observe this increase for both education and health groups. However, we also see larger increases for people with more education. As we will argue below, this is in part because they are more likely to use (and therefore benefit from) HAART. Post-HAART increases in welfare are also apparent and reflect time trends in the model that capture small but steady improvements to HAART as patients and doctors learned more about it, and that are not otherwise captured in the model.

In our model, HAART is an unanticipated innovation, so by comparing expected lifetime value just before and just after it was introduced, we can calculate the impact it has on 30-year-old agents with 10 years of work experience. In particular, Table 4 shows expected lifetime value in 1995 and 1996 by health and education status on the most effective available treatment at the time. For ease of comparison, we normalize to 100 the welfare of high-educated and high-CD4 count men. Prior to HAART, notice that having less than a college degree or having a low CD4 count each cut welfare in half, the former due to lower lifetime income and the latter due to lower life expectancy given largely ineffective medication. Consistent with these differences, individuals with a low CD4 count and relatively low education have roughly one-fifth the lifetime utility compared to those with better health and higher education, suggesting a complementarity between health and education. Put another way, low health and low education combine to lower lifetime utility more than each separately.



(A) High CD4



(B) Low CD4

FIGURE 8: Expected Total Lifetime Value on the Best Available Treatment. This figure shows the expected total lifetime value by visit and education for an agent with a high CD4 count (Panel A) or low CD4 count (Panel B) on the best available treatment. Before HAART was introduced in visit 24, combo-therapy was the best available treatment. The simulated agent is 30 years old and has 10 years of work experience. Total value increases for both education levels when HAART is introduced, but the value change is larger for those with more education.

The introduction of HAART had large and positive impacts on lifetime utility for all groups, which we see comparing the second to the first row of Table 4. We report the absolute gain in lifetime utility in the third row, along with the percent gain in the fourth row. For agents with a low CD4 count, HAART's introduction is associated with a 176.6–236.1% increase in expected lifetime value, compared to an increase of 76.3–85.7% for those with a high CD4 count. In absolute terms, the gains were larger for those with more education because they are more likely to take advantage of the innovation. However, in percentage terms, the innovation increased expected lifetime value somewhat more for those with less education because their expected lifetime value was lower to begin with.

To contextualize these gains, we also provide the dollar equivalent, similar to calculating a compensating variation. We accomplish this by comparing lifetime utility using the most effective available treatment in 1995 (combo-therapy) versus 1996 (HAART) and then calculating how much money paid to the first (1995) group would make them indifferent to the second (1996) group in terms of lifetime utility. The amount is how much an agent would need to be compensated to remain in the 1995 state versus transitioning to the 1996 state. The monetary gains reflected here range from $172,440 to $315,098. The lowest value is for individuals with a high CD4 count and low education. These individuals value the option of HAART, but may not choose to use it until their health deteriorates, so this value is discounted according to how long they expect not to use HAART. The value is also relatively low since low education predicts relatively low use (and thus few) benefits of HAART. Individuals without a college degree but with low CD4 counts see a larger value ($182,693) since their potential use of HAART is much sooner. Individuals with high CD4 counts and high education see a value of $271,575, reflecting a higher expectation to use HAART and a high value of each life-year gained due to higher expected earnings. The highest value of HAART ($315,098) is for highly educated people with low CD4 counts, individuals who have an immediate need for HAART, expect thus to use it immediately, do not expect it will interfere with their work as much as it may for lower-educated groups, and see greater utility gains for each life-year gained as income enters positively into utility.

TABLE 4: Expected Total Lifetime Value

	<College		College+	
	Low CD4	High CD4	Low CD4	High CD4
Combo-therapy (1995)	21.7	56.5	50.1	100.0
HAART (1996)	73.0	104.9	138.6	176.3
Absolute gain	51.3	48.4	88.5	76.3
Percent gain	236.1	85.7	176.6	76.3
Dollar gain	$182,693	$172,440	$315,098	$271,575
Δ by College	.	.	$132,405	$99,135

Notes: Estimates from model simulations. Total expected lifetime value for a 30-year-old male on best available treatment, 1995 (visit 23) versus 1996 (visit 26). High education, high CD4, combo-therapy normalized to 100. Dollar gain represents the dollars of consumption needed to reach the same value increase as the value increase from *Combo-therapy (1995)* to *HAART (1996)* in the table, or equivalently, to compensate an agent for staying at the 1995 state rather than transitioning to the 1996 state.

4.2 Decomposing Differences in the Value of HAART

According to Table 4, the difference by education group in the value of HAART for low-CD4 count men is $132,405 and for high-CD4 count men is $99,135. In showing large differences in the value of HAART due to health and education, the prior section enumerated several reasons, including links to the labor market. To more precisely show the various channels through which HAART had different impacts on lifetime utility, we perform a decomposition. We begin with the results in Table 4 and then progressively and cumulatively replace model parameters capturing lower-education agents with those of higher-education agents. Results are presented in Table 5, where panel A replicates results from Table 4, i.e., the different impact of HAART across the two health (high and low CD4) and education (college or more versus less than college) groups. The first row of each panel shows the 1995 combo-therapy value, normalized again at 100 for high education individuals with a high CD4 count. The second row of each panel shows the value of HAART under different parameters, i.e., as we assign high education parameters to the low education group. The third row shows the absolute gain in lifetime utility. The fourth row shows the total gain for each group in units of $10,000. The fifth row shows the percent of the total difference in gains across education groups that is accounted for by increasingly equalizing parameters. We note that the decompositions here could lead to different results depending on the order that we change parameters. In our case, different orderings have small effects on the reported magnitudes.[30]

Panel B shows the lifetime value of HAART assuming the lower-education group has the same income process as the higher-education group, but is otherwise unchanged. This closes 40% of the gap for low CD4 individuals and 44% for high CD4 individuals. The reason is that higher income means higher lifetime value for each year gained. It also means that individuals face stronger incentives to use HAART to be healthier and thus more productive, further increasing the value. In contrast, in panel C, when we assign parameters governing expenditures and insurance, there is little change to the lifetime utility gains of the low-education group; in fact, the gaps widen a bit because lower-education individuals in the context we study tend to pay less for medication. In general, this result underscores how medication costs are simply not an important factor driving differences in health disparities and the value of HAART in the context we study.

Changing the health production function parameters (panel D) also leads to little difference in the value of HAART across education groups. The health processes are quite similar across education groups, i.e., differences in outcomes do not appear to be driven by natural health processes, but instead by health behaviors. However, if we assign to lower-education individuals the same ailments process parameters as high-education individuals (panel E), we see somewhat larger relative gains in the welfare impact of HAART for the lower-education group. This reflects how higher-education individuals tend to report fewer ailments due to HAART, perhaps due to better dosing or guidance. The account of the gap in gains between education groups rises to 57% (or an additional 17% of the gap) for low CD4 count men and 60% (or an additional 16%) for high CD4-count men. In panel F, we change the mortality process parameters, which leads to larger changes (rising to explain 86% and 85% for low- and high-CD4, respectively, i.e., another 30% of the gap for low-CD4 count men and 25% for high-CD4 count men). In other words, if lower-education individuals had the same mortality process as higher-education individuals, they would value HAART substantially more. Higher education individuals expect to live longer and thus expect more years during which to value HAART.

Finally, panel G shows the effects of changing the structural utility parameters, capturing the

[30] We experimented with different possible orderings and chose the ordering generating the results here that delivers neither the largest or smallest explanatory power for each set of parameters. In Appendix C, we show that the value accruing to low-education agents if we change their utility parameter governing working with ailments to match that of high-education agents is remarkably stable. We show this by assessing the value of changing it after changing each other set of parameters. The dollar value remains remarkably stable at about $19,000 for low-CD4 count men and $15,000 for high-CD4 count men.

TABLE 5: Value Decomposition

	<College		College+	
	Low *CD*4	*High* *CD*4	*Low* *CD*4	*High* *CD*4
Panel A: All Education Differences				
Combo-therapy (1995)	21.7	56.5	50.1	100.0
HAART (1996)	73.0	104.9	138.6	176.3
Absolute Gain	51.3	48.4	88.5	76.3
Absolute Gain (in $10k)	18.3	17.2	31.5	27.2
Percent of Gap	0%	0%	.	.
Panel B: Same Income Process Only				
Combo-therapy (1995)	31.2	76.2	50.1	100.0
HAART (1996)	97.2	136.8	138.6	176.3
Absolute Gain	66.0	60.6	88.5	76.3
Absolute Gain (in $10k)	23.5	21.6	31.5	27.2
Portion of Gap	39.4%	43.7%	.	.
Panel C: Add Insurance & MOOP				
Combo-therapy (1995)	31.1	76.0	50.1	100.0
HAART (1996)	97.0	136.5	138.6	176.3
Absolute Gain	65.9	60.5	88.5	76.3
Absolute Gain (in $10k)	23.4	21.5	31.5	27.2
Portion of Gap	39.1%	43.3%	.	.
Panel D: Add Health				
Combo-therapy (1995)	31.4	76.7	50.1	100.0
HAART (1996)	97.9	137.4	138.6	176.3
Absolute Gain	66.4	60.8	88.5	76.3
Absolute Gain (in $10k)	23.7	21.6	31.5	27.2
Portion of Gap	40.7%	44.3%	.	.
Panel E: Add Ailments				
Combo-therapy (1995)	34.3	83.7	50.1	100.0
HAART (1996)	106.7	148.8	138.6	176.3
Absolute Gain	72.4	65.1	88.5	76.3
Absolute Gain (in $10k)	25.8	23.2	31.5	27.2
Portion of Gap	56.7%	59.9%	.	.
Panel F: Add Survival				
Combo-therapy (1995)	46.5	93.5	50.1	100.0
HAART (1996)	129.9	165.7	138.6	176.3
Absolute Gain	83.4	72.2	88.5	76.3
Absolute Gain (in $10k)	29.7	25.7	31.5	27.2
Portion of Gap	86.4%	85.5%	.	.
Panel G: Add College×Ailments×Work Utility Parameter				
Combotherapy (1995)	51.9	102.1	50.1	100.0
HAART (1996)	140.9	178.4	138.6	176.3
Absolute Gain	88.9	76.3	88.5	76.3
Absolute Gain (in $10k)	31.7	27.2	31.5	27.2
Portion of Gap	101.2%	100.1	.	.
Panel H: Add All Utility Parameters				
Combo-therapy (1995)	50.1	100.0	50.1	100.0
HAART (1996)	138.6	176.3	138.6	176.3
Absolute Gain	88.5	76.3	88.5	76.3
Absolute Gain (in $10k)	31.5	27.2	31.5	27.2
Portion of Gap	100%	100%	.	.

Notes: Estimates from model simulations. This table decomposes the value of HAART by education group by gradually changing processes and parameters to give the agents with lower education the processes and parameters of those with higher education. Panel A shows results with all differences intact, as in Table 4. Total expected lifetime value for a 30-year-old male on best available treatment, 1995 (visit 23) vs 1996 (visit 26). High education, high CD4, combo-therapy normalized to 100.

disutility of work due to ailments. Adding these more than closes the gap, accounting for 101% and 100% of the total valuation gap for the low and high education groups, respectively. By construction, the total amount of the gap declines to 100% for both groups (see Panel H) when we change the remaining parameters because other utility parameters imply larger benefits for the low education group. Returning to Panel G, differences in the disutility of working while suffering ailments account for about 15% of the total difference in value of HAART for both education groups. Recalling total value differences, this means that the side effects-work disutility parameter accounts for about $19,000 of the total HAART value difference by education group of $132,405 for low-CD4 count men. It accounts for about $15,000 of the total HAART value difference of $99,135 for high-CD4 count men.

Since preferences are often deep individual characteristics, unlike with the other processes, it is less clear whether policy can close this part of the gap. In our case, however, many preference parameters are quite similar across education groups, but there is a difference in the utility of working with ailments. This may reflect differences in work conditions, which are modifiable by policy. Doing so would help to close the gap in the value of innovation across demographic groups.

In summary, this exercise shows that the large gap in the value of an effective medical innovation with side effects is driven mostly by differences across groups in: income, mortality, ailments, and preferences over ailments and work. It is driven by income since each life-year gained yields more consumption utility; by ailments, which lead to lower utility; by mortality, as the innovation leads to more life years in which better health can be enjoyed; and by preferences as individuals with a lower utility of cost of work with ailments are more likely to use and thus to benefit from the innovation. Taken together, parameters associated with ailments account for about one-third of the total value difference. Notably, differences are not driven by the direct impact of ailments on utility. Similarly, gaps in the lifetime benefit by education group are not driven by differences in how health evolves (with or without medication) or by price, both of which are often discussed as key sources of health inequality.

4.3 Treatment Effects on Health versus Welfare

Our last exercise uses the estimated model to illustrate how improved health through assignment to treatment, akin to what might be shown in a clinical trial, does not necessarily translate to welfare improvements. The idea behind this policy experiment is not that it corrects a market failure or externality. Rather, we aim to highlight why individuals optimally choose to forgo effective medication in light of their labor market concerns. Higher treatment uptake may be a social policy goal for a host of reasons. Yet, assignment to effective treatment, while it improves health, can also lead to uncomfortable side effects that interfere with work. As we show, a treatment mandate not only decreases individual welfare, but also increases welfare inequality as side effects are relatively more costly for individuals with less education. These results help to explain why we do not observe universal uptake of HAART in the data.

We construct expected lifetime value under the simulation by assigning agents to use HAART, but allowing them to choose whether or not to work full time. It is important to mention how we assign HAART has implications for welfare calculations. We thus provide results on several options and in each case we find that the mandate improves health and reduces welfare for some individuals, with declines concentrated among those with lower education. One possibility simply reduces the set of options from eight choices (four treatment options times two work options) to two choices (one treatment option times two work options), which results in a mechanical decrease in value because the expected value of the utility shocks falls with the number of draws from the distribution. This represents the loss of value associated with the loss of options. Even for agents where HAART appears to be an obvious choice, a small share of individuals will forgo it due to idiosyncratic variation in preferences. By mandating HAART, we eliminate this possibility, which reduces expected value. On the other extreme, we can allow four independent draws of preference shocks while maintaining the HAART mandate, which increases welfare on average because individuals face the same number of utility shocks as in the baseline, but also face the deterministic value of HAART, which is higher than other options and adds to the expected value of the maximum of all four utility shocks.

We opt for the first method because removing the choice-specific draws from unavailable choices more closely represents the scenario faced by individuals mandated to use treatments in settings like clinical trials. Table 6 shows expected lifetime value, the probability of a high CD4 count, and the probability of employment across several 6-month mandate scenarios.[31] Panel A shows results from the model with all choices intact, so agents can choose among three treatment options or choose not to use treatment. Panel B shows the effects of a HAART mandate where all agents are forced to use HAART for 6 months. Panel C shows a treatment mandate where agents can choose among the three treatments but cannot forgo treatment. Panel D shows a no-treatment mandate where agents are not able to use any treatments. We integrate over the distribution of states observed in the data in visit 30 (1998) to reflect the composition of the sample population.

Mandating HAART (panel B) reduces expected lifetime value. Value declines less for those with a low CD4 count (0.3–0.9%) compared to those with a high CD4 count (1.6–3.1%) because they are likely to use HAART even without the mandate. Similarly, value declines more for those with less education (2.8% compared to 1.4%) because they are less likely to be using treatment be-

[31] Appendix Table C2 shows the results of this simulation compared with our preferred methodology. Under this simulation, lifetime value increases for all groups on average, but a substantial share of individuals, especially those with high CD4 and those with less education, are made worse off.

Table 6: Treatment Mandate

	All	<College			College +		
		All	Low CD4	High CD4	All	Low CD4	High CD4
Panel A: Lifetime value, no mandate	100.0	65.6	52.8	68.6	113.6	88.1	117.6
Pr(High $CD4_{t+1}$), no mandate	0.84	0.82	0.27	0.94	0.86	0.28	0.95
Pr(Treatment), no mandate	0.87	0.83	0.94	0.80	0.89	0.97	0.88
Pr(HAART), no mandate	0.65	0.61	0.80	0.57	0.66	0.87	0.63
Pr(Work), no mandate	0.66	0.58	0.40	0.62	0.69	0.54	0.71
Panel B: Lifetime value, HAART mandate	98.3	63.8	52.3	66.4	111.9	87.8	115.8
Percent change	-1.7	-2.8	-0.9	-3.1	-1.4	-0.3	-1.6
Pr(High $CD4_{t+1}$), HAART mandate	0.85	0.83	0.31	0.95	0.86	0.30	0.95
Percent change	1.2	1.7	14.4	0.9	1.0	8.7	0.6
Pr(Work), HAART mandate	0.64	0.56	0.39	0.59	0.68	0.54	0.70
Percent change	-2.2	-4.1	-2.7	-4.3	-1.6	-1.0	-1.7
Panel C: Lifetime value, treatment mandate	99.5	64.9	52.6	67.7	113.1	88.0	117.1
Percent change	-0.5	-1.1	-0.3	-1.2	-0.4	-0.1	-0.4
Pr(High $CD4_{t+1}$), treatment mandate	0.85	0.82	0.28	0.95	0.86	0.29	0.95
Percent change	0.4	0.6	3.8	0.4	0.3	2.3	0.2
Pr(Work), treatment mandate	0.65	0.56	0.39	0.60	0.68	0.54	0.70
Percent change	-1.5	-2.8	-1.8	-3.0	-1.0	-0.7	-1.0
Panel D: Lifetime value, no treatment	93.2	59.7	44.9	63.2	106.4	78.5	110.9
Percent change	-6.8	-9.0	-15.0	-7.9	-6.3	-10.9	-5.7
Pr(High $CD4_{t+1}$), no treatment	0.80	0.77	0.09	0.92	0.81	0.09	0.92
Percent change	-5.5	-6.1	-66.7	-2.1	-5.3	-68.5	-2.3
Pr(Work), no treatment	0.73	0.67	0.52	0.70	0.76	0.64	0.77
Percent change	11.1	15.4	29.3	13.4	9.6	18.4	8.6

Notes: Represents model predictions matched to data on observable state variables for visit 30. Value for "all, no mandate" normalized to 100.

fore the mandate. While the mandate reduces welfare, it increases health because people are forced to use health-improving treatments. Health improves more for those with a low CD4 count (8.7–14.4%) because those with a high CD4 count were likely to remain healthy regardless. Health increases more for those with less education (1.7% compared to 1.0%), reflecting the larger increase in HAART use, which comes with significant ailments that inhibit work. Employment declines when HAART is mandated, especially for those with less education (4.1% compared to 1.6%). This is because those with less education find it more difficult to work while experiencing side effects of treatment. Turning to panel C, allowing agents to choose between treatments reduces the expected value loss from the mandate, but also reduces the health improvements because some individuals choose sub-standard treatments. Panel D shows that preventing people from accessing treatment reduces both value and health, but increases employment through the reduction in ailments that comes from forgoing treatment.

Appendix Tables C3 and C4 show results removing the effect of medications on the ailment process. This allows us to show that the negative effects of the mandate are largely driven by the ailments (side effects) caused by the treatment. Table C3 shows results from a counterfactual scenario where none of the available treatments increase the probability of ailments, while Table C4 removes the ailment effects of HAART only. Removing the ailments of medication reduces the loss in value associated with a medication mandate. It also reduces the health improvements by increasing the baseline probability of medication use pre-mandate. Under this scenario, the medication mandate has almost no effect on labor supply because the agents do not face additional work-inhibiting ailments. In fact, Table C4 shows that in a scenario where HAART does not cause ailments but other treatments do, a HAART mandate actually slightly increases employment as agents switch from ailment-producing treatments to HAART.

The bottom line from this exercise is that assigning people to an effective treatment with harsh side effects can reduce welfare, especially for disadvantaged groups, given the interaction between labor supply and side effects. Results from clinical trials that show health improvements among people who are randomly assigned to receive treatment would be unlikely to provide information on these types of indirect costs that can affect patients' welfare and ultimately their decisions to take the medication outside of a clinical trial.

5 Conclusion and Limitations

5.1 Discussion

This paper develops a model for assessing variation in the value of medical innovation across sociodemographic groups. The model incorporates several reasons that can explain variation in health and labor decisions, including differences in income, the impact of medicine on health, and mortality. The model incorporates novel features that can further explain differences in how people make medical decisions, including side effects of medication that can interfere with work differently across sociodemographic groups.

We use the model to study an HIV treatment innovation, HAART, which was introduced in the mid-1990s. While HAART was far more effective than earlier treatments, it had harsh side effects that interfered with employment, especially so for patients with less than a college degree compared to those with a college degree or more. In general, we find that HAART provided less value (measured as gains in lifetime utility) for people with less education, thus exacerbating existing inequality in terms of lifetime welfare. From a decomposition exercise, we find individuals with less education value HAART relatively less since they earn less in any life-year gained, are less likely to see reductions in mortality due to lower usage of HAART, are more likely to suffer ailments, and face a larger negative impact on the utility of work due to side effects. Parameters related to ailments explain about one-third of why individuals with less education value HAART differently from those with more education. Variation in the utility cost of working while suffering side effects explain 13%-14% of the education differences in the value of HAART.

More broadly, the model we develop captures how health-maximizing and welfare-maximizing behaviors can be at odds and, moreover, that health-welfare tradeoffs can be particularly salient for disadvantaged individuals because various features of their lives—such as their work conditions—make health investments particularly costly. This approach to modeling patient behavior can further our understanding of how the contexts, circumstances, and tradeoffs agents face can lead them to optimally choose behaviors that exacerbate health disparities. Related, our approach illustrates that it is perhaps too simplistic to explain heterogeneity in health investments (often leading to health disparities) as arising from either barriers to access or from differences in preferences. This view creates a false dichotomy that masks a somewhat more complex story about how people, given their constraints, can optimally choose behavior that generates health disparities and, potentially, costs to society. Policy analysis and further research could tackle the question of whether relaxing certain constraints (e.g., changing work conditions to moderate the health-work tradeoff) is worthwhile.

We view the AIDS epidemic as a useful historical analogy that provides lessons for other health contexts. A recent example is COVID-19. Like HIV, COVID-19 has unequal consequences in part because protective health actions imply different costs for different groups. In the case of HIV, effective medication has side effects that make work difficult, especially for less educated people. In the case of COVID-19, staying at home has been more difficult for people in cramped housing or who could not telework. In both contexts, a useful starting point—one that can contribute to risk mitigation—is to understand the health-welfare tradeoffs that people in different circumstances face and to design policy accordingly.

5.2 Limitations and Future Work

While the model used in our analysis captures the role health plays in an individual's life and its interaction with sociodemographic factors through a rich set of mechanisms, it abstracts from several considerations. We close our discussion with suggestions for future research, which could use the basic framework we propose to better understand heterogeneity in health investments and the role of policy when resulting health disparities arise.

Some omissions from the model suggest obvious room for improvement and are due to lack of data. We omit savings and investments, as there are no data that would allow us to identify such behaviors. Second, we assume linear consumption utility. Conceptually, both omissions suggest we overestimate utility differences from consumption levels for workers versus non-workers. To the degree this is more pronounced for higher-education (and thus higher-income) groups compared to lower-education groups, incorporating savings and risk aversion would lead to larger differences in structural parameters across education groups, suggesting that the estimates we provide are a lower bound on true differences across groups. Nevertheless, we do not think true differences are substantially larger than the estimates we produce since we have experimented with several alternative modeling assumptions and find little change to our results.

Further—and we believe especially fruitful—research would enrich the model in two key directions. The first relates to information and expectations. It is possible that different beliefs or information about medications and their impact could explain variation in health behaviors, including differences across socioeconomic groups, such as those we find here. In the case of HIV, the impacts of medications were well-publicized and thus likely to have been understood. Moreover, while belief differences might explain the lack of medication use in general, it is not obvious why they would explain interactions between medication use and work, which is a focus of this paper. Yet, we live in a world in which people do not trust institutions, misinformation abounds, and individuals may not fully understand the implications of using (or failing to use) different kinds of medications. A decision not to use medication may be due to actual side effects or instead perceptions about side effects held by specific subsets of the population (consider, e.g., vaccine hesitancy due to perceived links to adverse health consequences). In general, hesitation to use medication by certain groups could be due to an over-expectation of side effects or an under-estimation of medication effectiveness, both of which can be of grave concern if they nudge people away from medications from which they would benefit. An enriched model could use data on expectations to capture alternative information structures that depart from rational expectations.

Second, to further explore the health-work link, future research could include a richer model of occupational characteristics and work conditions that are policy-modifiable. This type of extension would allow the researcher to examine the impact of counterfactual policies that reduce the costs of working while also engaging in health-maximizing behaviors. As we have shown in our paper, health and work can be at odds in some contexts (and for some groups), and useful policy might help to reduce health disparities (and their associated social costs) by minimizing this type of conflict.

References

Adimora, Adaora and Victor Schoenbach, "Contextual Factors and the Black-White Disparity in Heterosexual HIV Transmission," *Epidemiology*, 2002, *13* (6), 707–712.

Adler, Nancy and David Rehkopf, "U.S. Disparities in Health: Descriptions, Causes, and Mechanisms," *Annual Review of Public Health*, 2008, *29*, 235–252.

Arcidiacono, Peter, Holger Sieg, and Frank Sloan, "Living Rationally Under the Volcano? An Empirical Analysis of Heavy Drinking and Smoking," *International Economic Review*, 2007, *48* (1), 37–65.

Avery, Rosemary, Donald Kenkel, Dean Lillard, Alan Mathios, and Hua Wang, "Health Disparities and Direct-to-Consumer Advertising of Pharmaceutical Products," *Advances in Health Economics and Health Services Research*, 2008, *19*, 71–94.

Becker, Gary, "Health as Human Capital: Synthesis and Extensions," *Oxford Economic Papers*, 2007, *59* (3), 379–410.

Beer, Linda, Christine Mattson, Heather Bradley, and Jacek Skarbinski, "Understanding Cross-Sectional Racial, Ethnic, and Gender Disparities in Antiretroviral Use and Viral Suppression Among HIV Patients in the United States," *Medicine*, 2011, *95* (13), 1–9.

Bozzette, Samuel A, Sandra H Berry, Naihua Duan, Martin R Frankel, Arleen A Leibowitz, Doris Lefkowitz, Carol-Ann Emmons, J Walton Senterfitt, Marc L Berk, Sally C Morton et al., "The Care of HIV-Infected Adults in the United States," *New England Journal of Medicine*, 1998, *339* (26), 1897–1904.

Braveman, Paula, "Health Disparities and Health Equity: Concepts and Measurement," *Annual Review of Public Health*, 2006, *27* (1), 167–194.

Cawley, John and Christopher Ruhm, "The Economics of Risky Health Behaviors," *Handbook of Health Economics*, 2012, *2*, 95–199.

Chan, Tat Y. and Barton H. Hamilton, "Learning, Private Information, and the Economic Evaluation of Randomized Experiments," *Journal of Political Economy*, 2006, *114* (6), 997–1040.

_ , _ , and Nicholas W. Papageorge, "Health, Risky Behaviour and the Value of Medical Innovation for Infectious Disease," *Review of Economic Studies*, 2016, *83* (4), 1465–1510.

Chang, Virginia and Diane Lauderdale, "Fundamental Cause Theory, Technological Innovation, and Health Disparities: The Case of Cholesterol in the Era of Statins," *Journal of Health and Social Behavior*, 2009, *50*, 245–260.

Conti, Gabriella, James Heckman, and Sergio Urzua, "The Education-Health Gradient," *American Economic Review: Papers and Proceedings*, 2010, *100*, 234–238.

Crawford, Gregory S. and Matthew Shum, "Uncertainty and Learning in Pharmaceutical Demand," *Econometrica*, 2005, *73* (1), 1137–1173.

Cronin, Christopher, "Insurance-Induced Moral Hazard: A Dynamic Model of Within-Year Medical Care Decision Making Under Uncertainty," *International Economic Review*, 2019, *60* (1), 187–218.

_ , Matthew Forrstrom, and Nicholas Papageorge, "What Good Are Treatment Effects Without Treatment? Mental Health and the Reluctance to Use Talk Therapy," Forthcoming at, Review of Economic Studies 2024.

Currie, Janet, "Healthy, Wealthy, and Wise: Socioeconomic Status, Poor Health in Childhood, and Human Capital Development," *Journal of Economic Literature*, 2009, *47* (1), 87–122.

Cutler, David, Adriana Lleras-Muney, and Tom Vogl, "Socioeconomic Status and Health: Dimensions and Mechanisms," *The Oxford Handbook of Health Economics*, 2011.

Darden, Michael, "Smoking, Expectations, and Health: A Dynamic Stochastic Model of Lifetime Smoking Behavior," *Journal of Political Economy*, 2017, *125* (5), 1465–1522.

Dawson, Lindsey, Jennifer Kates, Tatyana Roberts, and Priya Chidambaram, "Medicaid and People with HIV," Technical Report, KFF 2023.

Detels, Roger, Patrick Tarwater, John Phair, Joseph Margolick, Sharon Riddler, and Alvaro Muñoz, "Effectiveness of Potent Antiretroviral Therapies on the Incidence of Opportunistic Infections Before and after AIDS Diagnosis," *AIDS*, 2001, *15* (3), 347.

Frisbie, Parker, SeungEun Song, Daniel Powers, and Julie Street, "The Increasing Racial Disparity in Infant Mortality: Respiratory Distress Syndrome and Other Causes," *Demography*, 2004, *41* (4), 773–800.

Gable, Carol, Jonathan Tierce, Diane Simison, Douglas Ward, and Katy Motte, "Costs of HIV+/AIDS at CD4+ Counts Disease Stages Based on Treatment Protocols," *Journal of Acquired Immune Deficiency Syndromes and Human Retrovirology*, 1996, *12* (4), 413–420.

Gerber, Ben, Irwin Brodsky, Kimberly Lawless, Louanne Smolin, Ahsan Arozullah, Everett Smith, Michael Berbaum, Paul Heckerling, and Arnold Eiser, "Implementation and Evaluation of a Low-Literacy Diabetes Education Computer Multimedia Application," *Diabetes Care*, 2005, *28* (7), 374–378.

Gilleskie, Donna, "A Dynamic Stochastic Model of Medical Care Use and Work Absence," *Econometrica*, 1998, *66* (1), 1–45.

Glied, Sherry and Adriana Lleras-Muney, "Technological Innovation and Inequality in Health," *Demography*, 2008, *45* (3), 741–761.

Goldman, Dana and Darius Lakdawalla, "A Theory of Health Disparities and Medical Technology," *Contributions to Economic Analysis and Policy*, 2005, *4* (1), 1–32.

_ **and James P. Smith**, "The Increasing Value of Education to Health," *Social Science and Medicine*, 2011, *72*, 1728–1737.

Grossman, Michael, "On the Concept of Health Capital and the Demand for Health," *The Journal of Political Economy*, 1972, *80* (2), 223–255.

_ , "The Correlation between Health and Schooling," in "Determinants of Health: An Economic Perspective," Columbia University Press, 2017, pp. 128–189.

Hall, Robert E and Charles I Jones, "The Value of Life and the Rise in Health Spending," *The Quarterly Journal of Economics*, 2007, *122* (1), 39–72.

Hamilton, Barton, Andrés Hincapié, Robert Miller, and Nicholas Papageorge, "Innovation and Diffusion of Medical Treatment," *International Economic Review*, 2021, *62* (3), 953–1009.

Jones, Charles I. and Peter J. Klenow, "Beyond GDP? Welfare across Countries and Time," *American Economic Review*, September 2016, *106* (9), 2426–57.

Lakdawalla, Darius, Neeraj Sood, and Dana Goldman, "HIV Breakthroughs and Risky Sexual Behavior," *The Quarterly Journal of Economics*, 2006, *121* (3), 1063–1102.

Lasser, Karen, David Himmelstein, and Steffie Woolhandler, "Access to Care, Health Status, and Health Disparities in the United States and Canada: Results of a Cross-National Population-Based Survey," *American Journal of Public Health*, 2006, *96* (7), 1300–1307.

Link, Bruce and Jo Phelan, "Social Conditions as Fundamental Causes of Disease," *Journal of Health and Social Behavior*, 1995, pp. 80–94.

Magnac, Thierry and David Thesmar, "Identifying Dynamic Discrete Decision Processes," *Econometrica*, 2002, *70* (2), 801–816.

Osborn, Chandra, Michael Paasche-Orlow, Terry Davis, and Michael Wolf, "Health Literacy: An Overlooked Factor in Understanding HIV Health Disparities," *American Journal of Preventive Medicine*, 2007, *33* (5), 374–378.

Papageorge, Nicholas W., "Why Medical Innovation is Valuable: Health, Human Capital, and the Labor Market," *Quantitative Economics*, 2016, *7* (3), 671–725.

Phelan, Jo, Bruce Link, and Parisa Tehranifar, "Social Conditions as Fundamental Causes of Health Inequalities: Theory, Evidence, and Policy Implications," *Journal of Health and Social Behavior*, 2010, *51*, 28–40.

Robinson, Russel and Aisha Moodie-Mills, "HIV/AIDS Inequality: Structural Barriers to Prevention, Treatment, and Care in Communities of Color," Technical Report, Center for American Progress and Berkeley Law 2012.

Rubin, Marcie, Cynthia Colen, and Bruce Link, "Examination of Inequalities in HIV/AIDS Mortality in the United States From a Fundamental Cause Perspective," *American Journal of Public Health*, 2010, *100* (6), 1053–1059.

Sommers, Benjamin, Katherine Baicker, and Arnold Epstein, "Mortality and Access to Care among Adults after State Medicaid Expansions," *The New England Journal of Medicine*, 2012, *367* (11), 1025–1034.

Stafford, Amanda and Lisa Wood, "Tackling Health Disparities for People Who Are Homeless? Start with Social Determinants," *International Journal of Environmental Research and Public Health*, 2017, *14.*

Thornton, Rachel, Crystal Glover, Crystal Cené, Deborah Glik, Jeffrey Henderson, and David Williams, "Evaluating Strategies For Reducing Health Disparities By Addressing The Social Determinants Of Health," *Health Affairs*, 2016, *35* (8), 1416–1423.

Wherry, Laura and Sarah Miller, "Early Coverage, Access, Utilization, and Health Effects of the Affordable Care Act Medicaid Expansions: A Quasi-Experimental Study," *Annals of Internal Medicine*, 2016, *164* (12), 795–803.

Williams, David and Pamela Jackson, "Social Sources of Racial Disparities in Health," *Health Affairs*, 2005, *24* (2), 325–334.

_ , Selina Mohammed, Jacinta Leavell, and Chiquita Collins, "Race, Socioeconomic Status and Health: Complexities, Ongoing Challenges and Research Opportunities," *Annals of the New York Academy of Sciences*, 2010, *1186*, 69–101.

Woolf, Steven H and Paula Braveman, "Where Health Disparities Begin: The Role of Social and Economic Determinants—and Why Current Policies May Make Matters Worse," *Health Affairs*, 2011, *30* (10), 1852–1859.

Woolf, Steven, Laudan Aron, Lisa Dubay, Sarah Simon, Emily Zimmerman, and Kim Luk, "How Are Income and Wealth Linked to Health and Longevity?," Technical Report, The Urban Institute and the Center on Society and Health 2015.

For Online Publication

Appendix to: Medical Innovation and Health Disparities

By: Barton H. Hamilton, Andrés Hincapié, Emma Kalish and Nicholas W. Papageorge

A Additional Descriptive Evidence

This appendix provides additional descriptive evidence on choices and outcomes.

A.1 Choices and Outcomes

TABLE A1: HIV Treatment Choice Transitions

		Time $t+1$				
		No treatment	Mono-therapy	Combo-therapy	HAART	Observations
			Pre-HAART, <College			
Time t	No treatment	85.9	9.9	4.2	–	838
	Mono-therapy	8.3	73.2	18.6	–	628
	Combo-therapy	7.5	19.5	73.1	–	319
			Pre-HAART, College+			
Time t	No treatment	88.1	8.6	3.3	–	1,477
	Mono-therapy	6.5	69.6	23.9	–	1,141
	Combo-therapy	3.4	21.4	75.2	–	762
			Post-HAART, <College			
Time t	No treatment	82.8	0.9	5.3	11.0	463
	Mono-therapy	9.4	28.1	40.6	21.9	35
	Combo-therapy	2.7	0.7	67.6	29.1	308
	HAART	2.6	0.1	4.1	93.2	902
			Post-HAART, College+			
Time t	No treatment	80.6	0.5	6.4	12.4	758
	Mono-therapy	1.0	30.0	33.0	36.0	103
	Combo-therapy	2.1	1.3	67.5	29.0	858
	HAART	1.5	0.2	4.3	94.1	2,698

Notes: This table presents a transition matrix for medication choices by HAART era and education level. Before HAART, medication use/non-use was highly persistent. After HAART, individuals using other medications often switch to HAART.

TABLE A2: Treatment Use by Education

	Overall	Pre-HAART Low Edu	Pre-HAART High Edu	Post-HAART Low Edu	Post-HAART High Edu
Health	-1.228***	-1.373***	-1.516***	-0.677***	-0.695***
	(0.064)	(0.123)	(0.097)	(0.172)	(0.147)
Insurance	0.950***	0.737***	0.814***	1.194***	1.126***
	(0.095)	(0.178)	(0.172)	(0.210)	(0.202)
Work	-0.271***	-0.082	-0.022	-0.638***	-0.530***
	(0.052)	(0.115)	(0.090)	(0.127)	(0.103)
<College	-0.357***				
	(0.050)				
HAART Era	0.798***				
	(0.093)				
Constant	-0.581	2.784*	-0.997	-4.823**	-0.420
	(0.792)	(1.673)	(1.376)	(2.306)	(1.776)
Observations	10,858	1,633	3,192	1,672	4,361

Notes: Parameters are from logistic regressions for 1,210 individuals in the sample. Standard errors in parentheses. Treatment is an indicator for whether the individual is taking any medication. Health indicates CD4≥250. Regressions include era-specific time trends (*Pre- and Post-HAART*), age, and age^2.

TABLE A3: Next Period Higher CD4 by Education

	Overall	Pre-HAART Low Edu	Pre-HAART High Edu	Post-HAART Low Edu	Post-HAART High Edu
Treatment	0.373***	0.056	0.238***	0.800***	0.640***
	(0.047)	(0.110)	(0.079)	(0.118)	(0.087)
Health	-0.127***	0.185	0.155*	-0.306**	-0.632***
	(0.049)	(0.117)	(0.086)	(0.128)	(0.089)
<College	-0.040				
	(0.043)				
HAART Era	0.815***				
	(0.082)				
Constant	-0.893	-0.563	-1.183	0.533	-0.894
	(0.678)	(1.587)	(1.314)	(2.000)	(1.278)
Observations	10,858	1,633	3,192	1,672	4,361

Notes: Parameters are from logistic regressions for 1,210 individuals in the sample. Standard errors in parentheses. Increase in health is defined as CD4$\geq CD4_{t-1}$. Regressions include era-specific time trends (*Pre- and Post-HAART*), age, and age^2.

TABLE A4: No Ailments by Education

	Overall	Pre-HAART Low Edu	Pre-HAART High Edu	Post-HAART Low Edu	Post-HAART High Edu
Health	0.909***	1.335***	0.961***	0.749***	0.698***
	(0.049)	(0.117)	(0.085)	(0.127)	(0.086)
Treatment	-0.478***	-0.205*	-0.552***	-0.464***	-0.525***
	(0.049)	(0.111)	(0.081)	(0.122)	(0.094)
<College	-0.208***				
	(0.044)				
HAART Era	-0.020				
	(0.084)				
Constant	2.962***	1.330	2.604*	3.341	4.986***
	(0.702)	(1.592)	(1.356)	(2.062)	(1.364)
Observations	10,858	1,633	3,192	1,672	4,361

Notes: Parameters are from logistic regressions for 1,210 individuals in the sample. Standard errors in parentheses. No ailments is an indicator for if the individual suffered no ailments in the period. Health indicates CD4≥250. Regressions include era-specific time trends (*Pre- and Post-HAART*), age, and age^2.

TABLE A5: Employment Choice Transitions

		Time $t+1$		
			<College	
		Not Working	Working	Observations
Time t	Not working	90.8	9.3	1,377
	Working	9.1	90.9	2,002
			College+	
		Not Working	Working	Observations
Time t	Not working	87.9	12.1	2,239
	Working	7.1	92.9	5,369

Notes: This table presents a transition matrix for employment choices by education level. Working is defined as full time work.

TABLE A6: Full Time Work by Education

	Overall	Pre-HAART Low Edu	Pre-HAART High Edu	Post-HAART Low Edu	Post-HAART High Edu
Ailments	-0.851***	-0.825***	-0.792***	-1.005***	-0.896***
	(0.086)	(0.164)	(0.136)	(0.229)	(0.201)
Health	0.834***	0.664***	0.777***	1.217***	0.935***
	(0.051)	(0.119)	(0.091)	(0.141)	(0.091)
Treatment	-0.093	0.304*	0.249**	-0.458***	-0.511***
	(0.066)	(0.160)	(0.120)	(0.166)	(0.132)
Ail X Treat	-0.144	-0.577***	-0.254	-0.264	0.083
	(0.100)	(0.222)	(0.175)	(0.260)	(0.214)
<College	-0.496***				
	(0.047)				
HAART Era	-0.492***				
	(0.091)				
Constant	-3.477***	-2.861*	-1.461	-6.998***	-5.936***
	(0.737)	(1.699)	(1.456)	(2.121)	(1.406)
Observations	10,858	1,633	3,192	1,672	4,361

Notes: Parameters are from logistic regressions for 1,210 individuals in the sample. Standard errors in parentheses. Health indicates CD4≥250. Regressions include era-specific time trends (*Pre- and Post-HAART*), age, and age^2.

TABLE A7: Income by Education

	Overall	Pre-HAART Low Edu	Pre-HAART High Edu	Post-HAART Low Edu	Post-HAART High Edu
Work	9.461***	8.823***	9.387***	10.492***	9.304***
	(0.165)	(0.409)	(0.331)	(0.380)	(0.259)
Health	1.034***	1.004**	0.320	2.086***	1.553***
	(0.179)	(0.394)	(0.309)	(0.460)	(0.327)
<College	-5.184***				
	(0.181)				
HAART Era	-2.029***				
	(0.310)				
Constant	-9.767***	-10.018	-21.364***	-16.811**	-7.946
	(2.793)	(7.784)	(6.412)	(7.681)	(5.197)
Observations	10,858	1,633	3,192	1,672	4,361
R-squared	0.356	0.270	0.243	0.360	0.272

Notes: Parameters are from linear regressions for 1,210 individuals in the sample. Standard errors in parentheses. Income is in thousands of year 2000 dollars per half year. Health indicates CD4≥250. Regressions include era-specific time trends (*Pre- and Post-HAART*), age, age^2, experience, and experience^2.

B Structural Parameter Estimates and Model Fit

TABLE B1: Health Process

	Δh_{it}
Health (h_{it-1})	0.071
	(0.111)
Health X Monotherapy $\left(h_{it-1} \cdot d^m_{1it}\right)$	0.108
	(0.077)
Health X Combotherapy $\left(h_{it-1} \cdot d^m_{2it}\right)$	0.443
	(0.067)
Health X HAART $\left(h_{it-1} \cdot d^m_{1it}\right)$	0.632
	(0.063)
Low Health X Monotherapy $\left((1-h_{it-1}) \cdot d^m_{1it}\right)$	0.016
	(0.136)
Low Health X Combotherapy $\left((1-h_{it-1}) \cdot d^m_{2it}\right)$	0.414
	(0.127)
Low Health X HAART $\left((1-h_{it-1}) \cdot d^m_{3it}\right)$	1.356
	(0.136)
Age (a_{it-1})	0.043
	(0.030)
Age^2 (a^2_{it-1})	-0.001
	(0.000)
Pre-HAART ν_t	-0.043
	(0.009)
Post-HAART ν_t	-0.001
	(0.005)
<College	-0.019
	(0.044)
Constant	-1.375
	(0.684)
	Health h_{it}
Health (h_{it-1})	5.332
	(0.313)
Health Booster X Low Health $\left(\hat{\Delta} h_{it} \cdot (1-h_{it-1})\right)$	4.649
	(0.362)
Health Booster X Health $\left(\hat{\Delta} h_{it-1} \cdot h_{it}\right)$	3.164
	(0.532)
Constant	-4.065
	(0.207)
Observations	10,858

Notes: Parameter estimates are from logistic regressions for the two-step health process in the structural model. Health is defined as CD4$\geq$250.

TABLE B2: Death Process

	Death b_{it+1}
Health (h_{it})	-6.881
	(1.034)
Age (a_{it})	0.020
	(0.008)
Age X Health ($a_{it} \cdot h_{it}$)	0.083
	(0.022)
<College	0.417
	(0.107)
Pre-HAART v_t	0.028
	(0.019)
Post-HAART v_t	-0.111
	(0.019)
Constant	-2.799
	(0.368)
Observations	11,290

Notes: Parameter estimates are from a logistic regressions for survival process in structural model. Estimates are based on a larger sample as because observations for individuals who die between t & t+1 are not included in the estimation of other processes.

Table B3: No Ailment Process

	No Ailments $\left(1-y_{it}^{ailments}\right)$
Health (h_{it-1})	0.908
	(0.049)
Monotherapy	-0.355
	(0.066)
Combotherapy	-0.508
	(0.060)
HAART	-0.563
	(0.063)
Age (a_{it-1})	-0.109
	(0.031)
Age^2 (a_{it-1}^2)	0.001
	(0.000)
Pre-HAART ν_t	-0.022
	(0.009)
Post-HAART ν_t	0.010
	(0.005)
<College	-0.213
	(0.044)
Constant	2.823
	(0.701)
Observations	10,858

Notes: Parameter estimates are from a logistic regression for the ailment process in structural model. The outcome is equal to one if the individual does *not* experience ailments.

TABLE B4: Income Process

	Income y_{it}^{income}
Experience (e_{it-1})	0.144
	(0.039)
Experience^2 (e_{it-1}^2)	-0.003
	(0.001)
Age (a_{it-1})	0.743
	(0.144)
Age^2 (a_{it-1}^2)	-0.005
	(0.002)
Health (h_{it-1})	1.211
	(0.199)
Work (d_{it}^l)	9.934
	(0.217)
<College	-6.665
	(0.306)
<College X Work	1.540
	(0.365)
Pre-HAART ν_t	-0.027
	(0.035)
Post-HAART ν_t	-0.112
	(0.019)
Constant	-8.634
	(3.022)
Variance of y_{it}^{income}	67.322
	(0.984)
Observations	10,858

Notes: Parameter estimates are from a linear regression for the income process in structural model. Income is in thousands of year 2000 dollars per half year.

TABLE B5: Insurance Process

	Insurance $y_{it}^{insurance}$
Health (h_{it-1})	-0.596
	(0.112)
Age (a_{it-1})	-0.236
	(0.083)
Age^2 (a_{it-1}^2)	0.003
	(0.001)
Experience (e_{it-1})	0.061
	(0.020)
Experience^2 (e_{it-1}^2)	-0.001
	(0.000)
Work (d_{it}^l)	0.787
	(0.126)
<College	-0.868
	(0.136)
<College X Work	-0.025
	(0.175)
Pre-HAART ν_t	-0.033
	(0.018)
Post-HAART ν_t	0.057
	(0.012)
Constant	7.257
	(1.729)
Observations	10,858

Notes: Parameter estimates are from a logistic regression for the insurance process in structural model. Insurance is a binary indicator for health insurance coverage.

TABLE B6: Medical Out-of-Pocket Process

	Medical OOP Expenses $y_{it}^{expenses}$
Income $\left(y_{it}^{income}\right)$	0.008
	(0.001)
Insurance $\left(y_{it}^{insurance}\right)$	-0.081
	(0.035)
Insurance X Monotherapy $\left(y_{it}^{insurance} \cdot d_{1it}^{m}\right)$	0.173
	(0.018)
Insurance X Combotherapy $\left(y_{it}^{insurance} \cdot d_{2it}^{m}\right)$	0.249
	(0.016)
Insurance X HAART $\left(y_{it}^{insurance} \cdot d_{3it}^{m}\right)$	0.261
	(0.016)
No insurance X Monotherapy $\left((1-y_{it}^{insurance}) \cdot d_{1it}^{m}\right)$	0.274
	(0.074)
No insurance X Combotherapy $\left((1-y_{it}^{insurance}) \cdot d_{3it}^{m}\right)$	-0.187
	(0.077)
No insurance X HAART $\left((1-y_{it}^{insurance}) \cdot d_{3it}^{m}\right)$	-0.164
	(0.067)
Health (h_{it})	-0.037
	(0.018)
No Ailments $(1-y_{it}^{ailments})$	-0.096
	(0.022)
Health X Ailments $\left(h_{it} \cdot y_{it}^{ailments}\right)$	-0.039
	(0.025)
Age (a_{it})	0.005
	(0.008)
Age^2 (a_{it}^2)	-0.000
	(0.000)
<College	-0.051
	(0.012)
Pre-HAART ν_t	0.002
	(0.002)
Post-HAART ν_t	0.009
	(0.001)
Constant	-0.197
	(0.187)
Variance of $y_{it}^{expenses}$	0.278
	(0.004)
Observations	10,858

Notes: Parameter estimates are from a linear regression for the medical out-of-pocket expenditure process in the structural model. Medical out-of-pocket expenses (MOOP) are in thousands of year 2000 dollars per half year.

TABLE B7: Model Fit

		Pr(Work)		Pr(HAART)	
Healthy	College+	Data	Model	Data	Model
0	0	0.39	0.42	0.25	0.29
1	0	0.66	0.64	0.31	0.26
0	1	0.52	0.56	0.31	0.38
1	1	0.75	0.73	0.40	0.35
All		0.66	0.66	0.36	0.33

Notes: This table compares moments simulated from the estimated model with data from the analysis sample.

C Additional HAART Mandate Simulations

This appendix provides results from decompositions of differences in the value of HAART using alternative orders of parameter changes. We also provide additional HAART mandate simulations performed under different assumptions compared to the simulation we present as part of our main results.

TABLE C1: Value Decompositions—Different Ordering

		Low CD4	High CD4
Income	Combotherapy (1995)	31.2	76.2
	HAART (1996)	97.2	136.8
+ Parameter	Combotherapy (1995)	35.3	84.1
	HAART (1996)	106.6	148.9
Dollar value of parameter change ($10k)	Combotherapy (1995)	1.46	2.82
	HAART (1996)	3.38	4.31
Income, Insurance, MOOP	Combotherapy (1995)	31.1	76.0
	HAART (1996)	97.0	136.5
+ Parameter	Combotherapy (1995)	35.2	83.9
	HAART (1996)	106.5	148.6
Dollar value of parameter change ($10k)	Combotherapy (1995)	1.46	2.82
	HAART (1996)	3.38	4.31
Income, Insurance, MOOP, Health	Combotherapy (1995)	31.4	76.7
	HAART (1996)	97.9	137.4
+ Parameter	Combotherapy (1995)	35.6	84.6
	HAART (1996)	107.4	149.6
Dollar value of parameter change ($10k)	Combotherapy (1995)	1.47	2.84
	HAART (1996)	3.40	4.33
Income, Insurance, MOOP, Health, Ailments	Combotherapy (1995)	34.3	83.7
	HAART (1996)	106.7	148.8
+ Parameter	Combotherapy (1995)	38.4	91.3
	HAART (1996)	115.8	160.3
Dollar value of parameter change ($10k)	Combotherapy (1995)	1.46	2.73
	HAART (1996)	3.25	4.09
Income, Insurance, MOOP, Health, Ailments, Survival	Combotherapy (1995)	46.5	93.5
	HAART (1996)	129.9	165.7
+ Parameter	Combotherapy (1995)	51.9	102.1
	HAART (1996)	140.9	178.4
Dollar value of parameter change ($10k)	Combotherapy (1995)	1.93	3.09
	HAART (1996)	3.89	4.54
All Same, original value for high education	Combotherapy (1995)	50.1	100.0
	HAART (1996)	138.6	176.3

*Notes:*Estimates from model simulations. This table complements Table 5 in the main text. As in that table, we gradually change processes and parameters to give the agents with lower education the processes and parameters of those with higher education. However, in each case we then add the parameter governing the negative interaction between work and ailments. We show that the amount of value this change accounts for is about the same regardless of the order in which it is changed. For example, if we change the income process of low-education men to be the same as that of high-education men, the dollar value of changing the work-ailments utility parameter is about $33,800 or $43,100 for low and high CD4 count agents, respectively. If instead we change income, insurance, out-of-pocket costs, the health process and survival parameters (again so those of low-education agents equal those of high-education men), the dollar value of then changing the work-ailments parameters is $38,900 for low CD4 count agents and $45,400 for high CD4 count agents.

TABLE C2: Treatment Mandate Simulation—Comparing Shocks

	All	<College			College +		
		All	Low CD4	High CD4	All	Low CD4	High CD4
Lifetime value, no mandate	100.0	65.6	52.8	68.6	113.6	88.1	117.6
Lifetime value, HAART mandate (2 draws)	98.3	63.8	52.3	66.4	111.9	87.8	115.8
Percent change	-1.7	-2.8	-0.9	-3.1	-1.4	-0.3	-1.6
Unit change	-1.7	-1.8	-0.5	-2.2	-1.6	-0.3	-1.8
Lifetime value, HAART mandate (8 draws)	101.0	66.5	55.1	69.2	114.7	90.5	118.5
Percent change	1.0	1.4	4.3	0.8	1.0	2.8	0.8
Unit change	1.0	0.9	2.3	0.6	1.1	2.4	0.9
Difference, mandate with 2 and 8 shocks	2.7	2.7	2.7	2.7	2.7	2.7	2.7

Notes: Model predictions averaged over observable state variables for visit 30. The value for "all, no mandate" normalized to 100.

TABLE C3: Treatment Mandate Simulation—Treatments Do Not Ever Cause Ailments

	All	<College			College +		
		All	Low CD4	High CD4	All	Low CD4	High CD4
Panel A: Lifetime value, no mandate	118.6	81.7	67.2	85.0	133.2	104.6	137.7
Pr(High $CD4_{t+1}$), no mandate	0.85	0.82	0.29	0.95	0.86	0.29	0.95
Pr(Treatment), no mandate	0.92	0.89	0.97	0.87	0.93	0.98	0.93
Pr(HAART), no mandate	0.74	0.73	0.92	0.69	0.75	0.93	0.72
Pr(Work), no mandate	0.73	0.67	0.51	0.70	0.75	0.64	0.77
Panel B: Lifetime value, HAART mandate	117.5	80.5	67.0	83.7	132.0	104.4	136.4
Percent change	-1.0	-1.4	-0.3	-1.6	-0.9	-0.1	-0.9
Pr(High $CD4_{t+1}$), HAART mandate	0.85	0.83	0.31	0.95	0.86	0.30	0.95
Percent change	0.7	0.9	5.6	0.6	0.6	4.4	0.4
Pr(Work), HAART mandate	0.73	0.67	0.51	0.70	0.75	0.64	0.77
Percent change	-0.02	-0.03	-0.02	-0.03	-0.01	-0.01	-0.01
Panel C: Lifetime value, treatment mandate	118.4	81.4	67.1	84.6	133.0	104.5	137.5
Percent change	-0.2	-0.4	-0.1	-0.4	-0.1	0.0	-0.2
Pr(High $CD4_{t+1}$), treatment mandate	0.85	0.83	0.30	0.95	0.86	0.29	0.95
Percent change	0.3	0.4	1.8	0.3	0.2	1.1	0.2
Pr(Work), treatment mandate	0.73	0.67	0.51	0.70	0.75	0.64	0.77
Percent change	-0.01	-0.01	-0.01	-0.01	0.00	0.00	0.00
Panel D: Lifetime value, No treatment	109.3	73.4	57.7	77.0	123.4	93.5	128.2
Percent change	-7.9	-10.1	-14.1	-9.4	-7.3	-10.6	-6.9
Pr(High $CD4_{t+1}$), No treatment	0.80	0.77	0.09	0.92	0.81	0.09	0.92
Percent change	-6.0	-6.9	-69.3	-2.4	-5.6	-69.8	-2.5
Pr(Work), No treatment	0.73	0.67	0.52	0.70	0.75	0.64	0.77
Percent change	0.1	0.1	0.3	0.1	0.1	0.2	0.1

Notes: Model predictions averaged over observable state variables for visit 30. The value for "all, no mandate" in Table 6 normalized to 100.

TABLE C4: Treatment Mandate Simulation—HAART Does Not Ever Cause Ailments

	All	<College			College +		
		All	Low CD4	High CD4	All	Low CD4	High CD4
Panel A: Lifetime value, no mandate	117.6	80.6	66.6	83.8	132.3	104.2	136.7
Pr(High $CD4_{t+1}$), no mandate	0.85	0.82	0.29	0.95	0.86	0.29	0.95
Pr(Treatment), no mandate	0.92	0.89	0.97	0.86	0.93	0.98	0.92
Pr(HAART), no mandate	0.80	0.79	0.93	0.75	0.81	0.95	0.78
Pr(Work), no mandate	0.72	0.66	0.51	0.69	0.75	0.64	0.76
Panel B: Lifetime value, HAART mandate	117.0	79.9	66.5	83.0	131.6	104.1	136.0
Percent change	-0.5	-0.8	-0.2	-1.0	-0.5	-0.1	-0.5
Pr(High $CD4_{t+1}$), HAART mandate	0.85	0.83	0.31	0.95	0.86	0.30	0.95
Percent change	0.6	0.8	4.5	0.5	0.5	3.1	0.4
Pr(Work), HAART mandate	0.73	0.67	0.51	0.70	0.75	0.64	0.77
Percent change	1.1	1.4	0.8	1.5	1.0	0.4	1.1
Panel C: Lifetime value, treatment mandate	117.4	80.2	66.6	83.4	132.1	104.2	136.5
Percent change	-0.2	-0.4	-0.1	-0.5	-0.2	0.0	-0.2
Pr(High $CD4_{t+1}$), treatment mandate	0.85	0.83	0.30	0.95	0.86	0.30	0.95
Percent change	0.3	0.5	1.8	0.4	0.3	1.1	0.2
Pr(Work), treatment mandate	0.72	0.66	0.51	0.69	0.75	0.64	0.76
Percent change	-0.1	-0.2	0.0	-0.3	-0.1	0.0	-0.1
Panel D: Lifetime value, no treatment	108.4	72.3	57.1	75.8	122.6	93.1	127.3
Percent change	-7.9	-10.3	-14.3	-9.6	-7.3	-10.7	-6.9
Pr(High $CD4_{t+1}$), no treatment	0.80	0.77	0.09	0.92	0.81	0.09	0.92
Percent change	-6.1	-7.0	-69.6	-2.5	-5.7	-70.1	-2.6
Pr(Work), no treatment	0.73	0.67	0.52	0.70	0.75	0.64	0.77
Percent change	1.2	1.5	1.1	1.6	1.1	0.6	1.2

Notes: Model predictions averaged over observable state variables for visit 30. The value for "all, no mandate" in Table 6 normalized to 100.