Rational Self-Medication

Preliminary. Do not cite or distribute

Michael Darden and Nicholas W. Papageorge*

October, 2018

Abstract

Since Grossman (1972), economists have viewed health as a form of human capital that increases survival rates, raises productivity, and improves the quality of life. Rational individuals invest in their health until the long-term benefits of doing so cease to outweigh the upfront costs. In the absence of safe and effective medication, individuals in poor mental health may turn to substances with contemporaneous benefits but long-run costs (e.g., problem drinking). We develop a model which rationalizes this self-medication, and, in the context of depression, we study the dynamics of rational self-medication with respect to alcohol and tobacco consumption. Using forty years of longitudinal data from the Framingham Heart Study and the exogenous introduction of selective serotonin reuptake inhibitors (SSRIs), we demonstrate an economically meaningful reduction in alcohol for men when better medication (i.e., SSRIs) became available. Simulations of a dynamic model of alcohol, tobacco, and antidepressant medication behavior suggest clear a pattern of substitution away from alcohol, but this pattern is attenuated relative to the static, reduced-form result, mainly because of addiction. In our simulations, the reduction in alcohol consumption varies significantly by the severity of depression, with the most severely depressed individuals the least likely to substitute away from alcohol, which suggests that alcohol may be an important source of relief for the highly depressed. Our results suggest a role for rational self-medication in policy debates regarding "deaths of despair."

*Darden: Johns Hopkins University and NBER. Papageorge: Johns Hopkins University, IZA and NBER. We thank Tom Mroz, Matthew Harris, Melinda Pitts, and seminar participants at the Atlanta Federal Reserve Bank, the University of Tennessee, 2018 Workshop on the Economics of Risky Behavior, the 2018 American Society of Health Economists Conference, and the 2018 European Health Economics Association Conference

JEL Classification: I10; I12;

Keywords: Self-Medication; Depression; Mental Health; Dynamic System of Equations

1 Introduction

Since Grossman (1972), economists have viewed health as a form of human capital that increases survival rates, raises productivity, and improves the quality of life. Accordingly, behaviors that can improve health, such as exercise, healthy eating, abstaining from risky behavior, or medication usage, can be viewed as costly investments in human capital. Rational individuals invest in their health until the long-term benefits of doing so cease to outweigh the upfront costs. The basic model has been expanded upon to incorporate additional costs and benefits to capture the realities of many health-related decisions. Examples include side effects that discourage use of effective medications (Papageorge, 2016), addiction that encourages use of harmful substances (Darden, 2017), and technological substitution towards Statin pharmaceuticals (Kaestner *et al.*, 2014), among others.

This framework overlooks the idea that many individuals, lacking access to *good* medication, may take matters into their own hands, turning to substances that are potentially very harmful in the long-run (e.g., problem drinking or opioids) in an effort to seek relief from symptoms of illnesses, such as chronic pain or depression. Seen this way, many individuals who use harmful or addictive substances are rationally choosing to self-medicate, that is, to optimally make use of available technologies to alleviate symptoms, albeit at the cost of future poor health, possible addiction, and other negative consequences. Understanding how, and under what circumstances, people self-medicate is important because self-medication is socially costly, especially if it leads to addiction. However, treating use of dangerous substances as an error in judgment or an act of desperation - rather than as a rational attempt to mitigate health problems using prevailing technology - can lead to the wrong policy conclusions. Viewing problem drinking as an error suggests policies to curb drinking. Viewing it as rational self-medication would suggest that such policies could backfire if people substitute to potentially more harmful substances.[1] A better policy response would be to promote treatment innovations that obviate the need to self-medicate and thus induce rational actors to substitute towards less harmful substances.[2] For example, Powell *et al.* (2018) show that medical marijuana laws, and in particular the number of marijuana dispensaries, is associated with fewer opioid overdoses.

In this paper we examine self-medication in the form of problem drinking and smoking in the context of depression. Our work relates to the concept of self-medication, in which a depressed individual manages her depression outside of formal prescription medicine or therapy (Khantzian, 1985). We first develop a two-period model in which an agent makes investment decisions of alcohol, tobacco,

[1] For example, both Dinardo (2001) and Crost (2012) use minimum drinking age regulations to show clear substitution patterns between alcohol and marijuana. Furthermore, Bacolod *et al.* (2017) study minimum age drinking laws and show that the largest increase in drinking at age 21 (for those in the military) comes from the most depressed.

[2] For example, Anton *et al.* (2006) present results from the randomized-controlled trial COMBINE, the largest random intervention study of alcoholics, and show that the combination of medical management and Naltrexone significantly reduced the probability of relapse.

and antidepressant medications given her mental health state. Poor mental health generates symptoms which reduce utility in a consumption sense, and investments in mental health management have different contemporaneous (symptom relief) and intertemporal (mental health stock) effects. We show that self-medication can be rationalized purely by a convex symptom cost specification in the utility function. Furthermore, our model suggests that different bundles of investment may be appropriate depending on the stock of mental health.[3] Finally, our model emphasizes the importance of dynamics in the production of mental health. If antidepressant consumption causes individuals to substitute away from heavy alcohol and tobacco use, then a potential channel through which antidepressants improve both overall and mental health may lie in behavioral changes.

To investigate empirically, we turn to data from the Framingham Heart Study Offspring Cohort, which include longitudinal information on alcohol, tobacco, and antidepressant consumption, as well as measures of Center for Epidemiological Studies - Depression (CES-D) scores for roughly 5,000 individuals over a 40-year period. We exploit the arrival of a new technology, selective serotonin reuptake inhibitors (SSRIs), to test the rational self medication hypothesis. Conceptually, individuals who rationally self-medicate depression using a harmful substance may be more likely to substitute towards SSRIs, which are equally effective as previous generation antidepressants (i.e., Monoamine Oxidase Inhibitors (MAOI) and Tricyclic Antidepressants (TCA)) but which exhibit significantly less severe side-effects. Moreover, we expect there to be heterogeneity in substitution patterns depending on the stock of accumulated addictive capital.

The Framingham Heart Study follows individuals over nine waves from 1971 through 2011, straddling the Food and Drug Administration's approval of SSRIs in 1988. The average age at the initial wave of the Study is 36, ranging from 13 to 62. Over the panel, rates of antidepressant medication usage grow significantly while heavy alcohol consumption declines and tobacco use plummets. As a baseline specification, we estimate two-way individual and time fixed effects regressions of each behavior (light and heavy alcohol consumption and tobacco use) on a binary variable for contemporaneous antidepressant use.[4] Results suggest strong substitutability between alcohol and antidepressants: taking an antidepressant is associated with a statistically significant 5.5 (19.6%) percentage point increase in abstinence from alcohol, mainly driven by transitions from light to never drinking. However, this overall effect masks significant heterogeneity by gender. An antidepressant is associated with an 8.1 percentage point (40%) reduction in heavy drinking for men, yet women, who use antidepressants at significantly higher rates, drive the overall result of transitions from light to no drinking. We find no statistically significant effect of antidepressants on cigarette smoking for men or women. These results are robust

[3]Our model formalizes the argument that the type of substance abuse depends on the type and severity of mental health ailment(Khantzian, 1985)

[4]Typically, physicians discourage alcohol consumption with SSRIs, yet we observe both heavy alcohol and antidepressant use in our data frequently.

to a variety of specification tests, and trends in behaviors are statistically parallel for ever and never antidepressant users prior to SSRIs.

Our fixed effects estimates establish an economically and statistically significant relationship between antidepressants and alcohol, but the fixed effects estimator ignores the inherent dynamics in health production. Alcohol may increase contemporaneous mental health but worsen future mental health. Similarly, a history of alcohol consumption may prevent substitution towards SSRIs in rational self-medicators. To address the dynamics of self-medication, we estimate a system of dynamic equations which approximate a more general structural model. Specifically, we estimate dynamic equations for alcohol, tobacco, and antidepressants jointly, along with depression, attrition, and mortality equations, to capture heterogeneity in uptake of antidepressants and to control for selective exit from the study. The empirical framework is similar to the dynamic seemingly unrelated regression (SUR) model in Darden *et al.* (2018), who use FHS data to study the effect of cigarette smoking on expected longevity. We allow for correlation in the permanent component of the error structure across equations to capture unobserved heterogeneity in the joint determination of these behaviors and outcomes (Heckman & Singer, 1984; Mroz, 1999).

Similar to the single-equation Arellano-Bond estimator (Arellano & Bond, 1991), identification of our dynamic system comes from time-varying exogenous characteristics. As noted above, our study utilizes the FDA's 1988 approval of SSRIs to identify substitution between risky behaviors and antidepressants. Relative to previous generation antidepressants, SSRIs represent a significant improvement in the side effects associated with antidepressants; furthermore, because of the lack of cardiovascular side effects, SSRIs made antidepressants newly available to certain subgroups (e.g., the elderly). Prescriptions for antidepressants increased dramatically following the FDA decision, and we argue that this technological innovation – SSRIs – caused a dramatic reduction in the full price of antidepressants.[5] Thus, identification comes from, in part, the exogenous improvement in antidepressants. Further exclusion restrictions – variables that shift behaviors but not outcomes – include the prices of alcohol and cigarettes during our time period. Finally, we model the initial conditions of alcohol and cigarettes, and we estimate these equations jointly with the per-period behaviors and outcomes.

We estimate our dynamic model with the same sample as in our fixed effects specification, and we simulate the model under two counterfactual scenarios. First, following the introduction of SSRI pharmaceuticals, we contrast a scenario in which we impose antidepressants on the entire sample to our baseline simulation. Heavy drinking declines by roughly four percentage points, but this decline is driven by men and by those simulated to *never* be depressed during our sample period. Indeed, we find no change in heavy alcohol consumption, in any period, for those simulated to be depressed. This result suggests that the marginal product of alcohol consumption may be higher for more severely depressed

[5]Horwitz (2010) document that the marketing of SSRIs as antidepressants, rather than anti-anxiety medication, was partly due to the poor public perception of anti-anxiety drugs.

individuals. Alternatively, antidepressants may lower the marginal utility of alcohol for both groups, but for severely depressed individuals with a large stock of addictive alcohol capital, the decrease in marginal utility was not sufficient to induce a significant change.

To investigate, our second simulation removes the dependence of past alcohol consumption in the contemporaneous alcohol demand equation. Overall, heavy alcohol consumption drops enormously and antidepressant usage (which is chosen endogenously) increases by 1.7% percentage points by the final exam. This substitution is driven primarily by women and those simulated to be depressed. We interpret these results to suggest that alcohol addiction may significantly hinder substitution away from alcohol. Overall, our results suggest that individuals who self-medicate, when facing a new and better medical technology, one that obviates the need to self medicate, substitute towards it. In other words, exploiting a large medical innovation, we find clear evidence of rational self-medication.

Our work relates to the medical literature on self-medication, which has largely studied cross-sectional correlations. For example, Bolton *et al.* (2009) outlines competing theories in which the direction of the causality between alcohol abuse and mental health problems is unclear. To our knowledge, we are the first paper to model the dynamics of investment into a mental health production function and to estimate the substitutability between inputs. Our paper also relates to a growing economic literature on the role of technological innovation on behavior (Kaestner *et al.*, 2014; Papageorge, 2016). Finally, our work has implications for how policy addresses antidepressant prescriptions and treatment of depression. Major Depressive Disorder (MDD) is the leading cause of disability worldwide. In the United States, MDD affects 8.1% of individuals over the age of 18, and there exists significant heterogeneity by socioeconomic status: while the rate of MDD for those below 100% of the Federal Poverty Line (FPL) was 15.8% between 2013 and 2016, the rate was only 3.5% for those at or above 40% of the FPL (Brody *et al.*, 2018). Not surprisingly, a large literature shows strong correlations between MDD and fundamental economic building blocks of productivity and health, including poor education outcomes, lack of marriage or early marriage, teen childbearing, martial functioning, employment status, cancer, cardiovascular disease, and diabetes.

More generally, providing evidence of rational self-medication has implications for understanding the dramatic increase in mortality rates of white non-Hispanic men since 1998, the so-called "deaths of despair," (Case & Deaton, 2015). The prevailing explanation for the increase in deaths due to drug overdoses, alcohol-related liver disease, and suicide, especially for those with a high school degree or less, is that birth-cohort specific cumulative exposure to poor labor market conditions affects health, marriage, and future labor market outcomes (Case & Deaton, 2017; Ruhm, 2018). Our theory is consistent with this explanation but goes further. In the context of depression, poor labor market conditions at labor market entry may induce alcohol consumption, and the cumulative exposure to alcohol (i.e., the addictive stock of capital) may leave individuals with low utility in levels but with a high marginal utility for

alcohol. Whereas a death of despair technically suggests a death due to a lack of hope, self-medication suggests the opposite: heavy alcohol use may represent an attempt to medicate. This suggests a new look at policy, especially with respect to how the judicial system treats addicts, but also with respect to innovation. If our theory better characterizes behavior, the appropriate policy response is to stop punishing people who use risky substances to self-medicate and instead work to develop treatments for chronic pain that are less addictive or more effective so that people can rationally substitute away from harmful self-medicating behavior.

This paper proceeds as follows. In Section 2, we discuss a simple, two-period theoretical model of rational self-medication. In Section 3, we provide background on depression and depression treatment, as well as the literature on self-medication. In Section 4, we present our main data, the Framingham Heart Study, and we document empirical evidence of a plausibly causal relationship between antidepressants and alcohol consumption. Section 5 presents our dynamic model, as well as parameter estimates, model fit, and simulation results. Section 6 discusses our results and concludes.

2 Theory

Agents solve a two-period problem, where periods are denoted t and $t+1$. Where possible, we drop time subscripts and denote $t+1$ variables with a "prime".

Agent i enters period t with state variable M_{it}, which is the stock of mental health. Lower values of M_{it} imply worse mental health. Agents choose whether or not to take an SSRI, denoted $D_{it} \in \{0, 1\}$ and how much alcohol to drink $A_{it} \in \mathbf{R}^+$. For ease of exposition, we assume that the agent i chooses non-zero alcohol consumption.

Agents have preferences over alcohol consumption A and SSRI consumption D, where the latter includes the price of SSRIs along with side effects, stigma and other non-pecuniary costs of SSRI use. They do not have preferences over mental health *per se*, but instead over symptoms of mental health S. Agents choose A and D to solve:

$$\max_{A,D} \Big(u(S, A, D) + \beta v(S') \Big) \tag{1}$$

where we assume that S and D enter negatively and A enters positively into both u and v. Period t is effectively a "terminal" period in which no decisions are made and $v(S')$ is thus a continuation payoff affected by period-t choices which thus provides dynamic incentives to improve mental health.

Mental health evolves according to the following production function

$$M_{t+1} = f_m(M_t, A_t, D_t) \tag{2}$$

where the argument M_t captures persistence in mental health stock, A_t captures how alcohol usage can have negative impacts on future mental health, perhaps through addiction stock increases and D_t

captures how SSRIs can improve long-run mental health. Period-t symptoms are a function of the same arguments so that:

$$S_t = f_s(M_t, A_t, D_t) \tag{3}$$

where symptoms are more likely to occur when M_t is lower. Alcohol can improve symptoms, which is the "self-medication" effect, and SSRIs can also improve symptoms.

To characterize self-medicating behavior, we use the model to make the following three points. One, we show conditions under which $D^\star = 1$. Two, we characterize optimal alcohol usage. Three, we discuss conditions under which lowering the costs associated with SSRI usage would lead to decreases in alcohol usage. The third point is consistent with a reduction in self-medication through alcohol when medication becomes the more attractive option.

To show optimal SSRI usage, assume optimal alcohol consumption A^* and A^{**}, when using SSRIs and not using SSRIs, respectively. Agents use SSRIs when the benefits of doing so exceed the costs:

$$\begin{aligned} u(S(D=1), A^*, D=1) &+ \beta v(S'(M'(D=1))) \geq \\ u(S(D=0), A^{**}, D=0) &+ \beta v(S'(M'(D=0))) \end{aligned} \tag{4}$$

To fix ideas, suppose we make the simplifying assumption on period-t utility that the costs of medication usage are additively separable from other utility components, e.g., $u(S, A, D) = \tilde{u}(S, A) - \phi(D)$ where $\phi(D=1) = \phi$ and $\phi(D=0) = 0$.[6] The agent uses SSRIs if and only iff

$$\begin{aligned} \tilde{u}(S(D=1), A^*) + \phi(D=1) &+ \beta v(S'(M'(D=1))) \geq \\ \tilde{u}(S(D=0), A^{**}) + \phi(D=0) &+ \beta v(S'(M'(D=0))) \iff \\ \tilde{u}(S(D=1), A^*) - \tilde{u}(S(D=0), A^{**}) &+ \beta[v(S'(M'(D=1))) - v(S'(M'(D=0)))] \geq \phi \end{aligned} \tag{5}$$

The last line implies that the benefits must outweigh the costs in order for SSRI usage to occur, where the benefits include current period utility of fewer symptoms along with discounted $t+1$ reductions in symptoms due to increased mental health stock. For a given level of SSRI effectiveness, SSRI usage increases if the flow utility costs decline, e.g., through side effects, stigma or price reductions. Moreover, as long as $\phi > 0$, SSRI usage only occurs if there are benefits in the form of improved symptoms, either currently or in the future.

Next, we characterize optimal alcohol consumption, in which the relevant first order condition is:

$$\frac{\delta u}{\delta S}\frac{\delta S}{\delta A} + \frac{\delta u}{\delta A} + \frac{\delta v}{\delta S'}\frac{\delta S'}{\delta M'}\frac{\delta M'}{\delta A} =!0 \tag{6}$$

[6] Additive separability implies that the marginal utility of alcohol is unaffected by SSRI usage. While this assumption is unrealistic, it simplifies the exposition for optimal SSRI usage, and it does not affect our comparative dynamics analysis presented below.

or

$$\frac{\delta u}{\delta A} + \frac{\delta u}{\delta S}\frac{\delta S}{\delta A} = -\beta \frac{\delta v}{\delta S'}\frac{\delta S'}{\delta M'}\frac{\delta M'}{\delta A} \tag{7}$$

The left hand side captures the marginal benefits of alcohol use, including both the enjoyment of alcohol along with reduction in symptoms from self-medicating. The right hand side captures marginal costs: higher A reduces M' and lower M' reduces continuation payoffs captured by v. Optimal alcohol usage occurs when the marginal benefit of an additional unit of A is equal to the marginal cost.

Finally, we use our simple model to derive conditions under which SSRI usage should lead to decreases in alcohol usage. It is convenient to define a function for the marginal utility of side effects for both periods as follows:

$$\frac{\delta v}{\delta S} = \frac{\delta u}{\delta S} \equiv \alpha(S) \tag{8}$$

For example, if $\alpha(S) = \alpha S$ and $\alpha > 0$, then utility is a concave function with increasingly negative marginal utility of S. Having done this, the first-order condition above can be rewritten as:

$$\frac{\delta u}{\delta A} = -\alpha(S)\left[\frac{\delta S}{\delta A} + \beta \frac{\delta S'}{\delta M'}\frac{\delta M'}{\delta A}\right] \tag{9}$$

If alcohol usage decreases with SSRIs, it must be the case SSRIs lead to the left hand side of the last equation falling or the right hand side rising. We do not allow the enjoyment of alcohol to be a function of symptoms, so the left hand side does not change. Thus, for SSRIs to lower alcohol usage, it must be the case that the right hand side rises or that -1 times the right hand side falls. In other words, we now examine when the following expression declines when symptoms decline:

$$\alpha(S)\left[\frac{\delta S}{\delta A} + \beta \frac{\delta S'}{\delta M'}\frac{\delta M'}{\delta A}\right] \tag{10}$$

There are four possibilities:

1. $\alpha(S)$ is lower when $D = 1$. Given that utility is a declining function of S, this suggests that costs of S rise with S. The implication is that medication leads to a decline in symptoms. This reduces the marginal cost of symptoms, which means that the marginal benefit of technology that reduces symptoms is lower.

2. A second possibility is that $frac\delta S\delta A$ is lower when $D = 1$. This could occur if alcohol is less productive at reducing symptoms at lower symptom levels.

3. The third possibility is that $\frac{\delta S'}{\delta M'}$ is smaller when $D = 1$. This means that improvements to mental health reduce symptoms more so when mental health is better.

4. Finally $\frac{\delta M'}{\delta A}$ is lower when $D = 1$ which suggests that alcohol reduces mental more so if mental

health is better.

3 Background

Depression is a chronic mental health condition that, while highly treatable, represents the leading cause of disability globally[7]. Depression produces symptoms that include feelings of sadness, pessimism, guilt, anxiety, and decreased energy, loss of interest in daily activities, and indecisiveness. Indeed, clinical diagnosis of major depressive disorder includes near daily symptoms plus some functional impairment with respect to family and peer relationships, school/work performance, and stress and anxiety level (O'Connor EA, 2009).[8] In the United States, in any given two-week period between 2013 and 2016, 8.1% of Americans suffered from depression, ranging from 5.5% for men to 10.4% for women. There exists a strong gradient between depression and income: 19.8% of women earning less that 100% of the Federal poverty line (FPL) exhibit depressive symptoms compared only 4.8% of women at or above 400% of the FPL (Brody *et al.*, 2018).

Unsurprisingly, depression is associated with a wide variety of mental and physical ailments, including sleep sleep problems, irritability, persistent physical pain, and suicidality (U.S. HHS, 2015). Beck *et al.* (2011) show that depression is associated with significantly lower fundamental economic building blocks such as workforce productivity, which they measure with the Work Productivity and Activity Impairment questionnaire, and Berndt *et al.* (1998) demonstrate that depressed workers have lower levels of perceived at-work productivity and performance. Furthermore, Kessler (2012) shows that MDD is associated with low educational attainment, teen pregnancy, martial disruption, unemployment, functional status, early mortality, and suicide.

Antidepressants have existed since the initial Monoamine Oxidase Inhibitors (MOAI) developed in the in 1950s. MAOI antidepressants were effective at relieving symptoms of depression, but because these early drugs inhibit monoamine oxidase (i.e., the destruction of neurotransmitters) for several neurotransmitters, including dopamine, norepinephrine, serotonin, and tyramine, side effects of MAOIs included risk of stroke, cardiovascular ailments, sexual dysfunction, among others. Tricyclic antidepressants (TCA), which were developed in the 1960s, focused on the preventing the reuptake of just norepinephrine and serotonin. Despite the improvement in precision over MAOI antidepressants, side effects associated with TCA antidepressants can still be severe. Reflecting the side effects, as well as public stigma associated with antidepressants, only 2-3% of Americans used an anti-depressant through the middle 1980s.[9]

Selective Serotonin Reuptake Inhibitors (SSRIs) were approved by the Food and Drug Administration

[7] http://www.who.int/en/newsroom/fact-sheets/detail/depression

[8] In the middle 20th century, anxiety was the leading mental illness in the United States. **?** describes how, through a series of reclassification as well as the introduction of SSRIs, anxiety has given way to a focus and prevalence of depression.

[9] See Hillhouse & Porter (2015) for an excellent overview of the history on antidepressants.

in 1988, and, as the name suggests, effectively inhibit the reuptake of serotonin, making more serotonin available in the brain without affecting the levels of other neurotransmitters. SSRIs significantly altered the perception of antidepressants, reducing stigma, and expanding the set of individuals for whom an antidepressant is considered safe (i.e., the elderly). Rates of antidepressants have increased dramatically since 1988 – up to 12.7% of Americans were prescribed an antidepressant between 2011 and 2014, and of those taking an antidepressant, 25.3% have been taking an antidepressant for more than 10 years (Brody *et al.*, 2018). Researchers now use SSRI perscriptions to gauge rates of depression, mental health, and happiness. For example, Blanchflower & Oswald (2016) study the well-known u-shaped well-being curve with respect to age and show a similar pattern between antidepressants and age. Despite a significant literature that relates SSRIs to teen suicidality, Ludwig *et al.* (2009) shows that SSRIs actually reduce suicides across 25 countries after controlling for the intuitive selection of depressed individuals into antidepressant use.

Khantzian (1985) set forth the hypothesis that self-medication leads to addiction. Furthermore, the idea that the type of drug is not random but depends on the type of illness. Depressed individuals have a clear incentive to manage and maintain mental health, and these endogenous investments into the mental health production function may have important implications for a variety of outcomes, including labor market productivity and long-term health. For example, Figures 1a.-1d. present National Health and Nutrition Examination Survey (NHANES) data on the use of antidepressants and heavy alcohol for men and women by the tertile of the Patient Health Questionnaire (PHQ-9) depression score between 2007 and 2013. Not surprisingly, for both men and women, more severely depressed individuals are persistently and significantly more likely to engage in each behavior. The medical and public health literature document this cross-sectional correlation between each behavior and depression[10], but the dynamic interrelationship between these endogenous investments in mental health has not been studied.

The voluminous empirical literature on self-medication predominantly documents similar cross-sectional correlations to those in Figures 1a.-1d. For example, Harris & Edlund (2005) look at National Survey on Drug Use and Health and find that heavy alcohol use is associated with a lack of mental health services in the past year, but that illicit drugs (not marijuana) increased with unmet need for mental health care. Rather than rely on cross-sectional evidence to infer self-medication, Crum *et al.* (2013) directly asks survey participants if they self-medicate. Those authors show that mental health illness is a significant rationale for alcohol consumption, and that self-medication was associated with the development of alcohol use disorders. Finally, Deykin *et al.* (1987) were the first to demonstrate that major depressive disorder typically predates alcohol use disorders in adolescents, providing some evidence on the direction of causality.

[10]For example, see Bolton *et al.* (2009), who use nationally representative survey data form National Epidemiologic Survey on Alcohol and Related Conditions to document cross-sectional correlations between alcohol and drug use and a variety of mental helath conditions.

To summarize, major depressive disorder is the most common mood disorder in the United States, affecting over 16.2 million adults in 2016. SSRIs significantly expanded the choice set with respect to the management of depression, which is frequently medicated outside of the medical system with potentially harmful and addictive substances.

4 The Framingham Heart Study

To study self-medication empirically, we turn to the Offspring Cohort of the Framingham Heart Study (FHS). The Offspring Cohort data are ideal for our purposes as they include longitudinal information on alcohol, tobacco, antidepressant medication, and mental health over nine detailed health exams over 40 years. Begun in 1971, the Offspring Cohort includes roughly 5,000 offspring of the FHS Original Cohort, which began in 1948 in Framingham Massachusetts, and their spouses. Both cohorts of individuals have received detailed health examinations at 2-4 year intervals into the 21st century, and both cohorts have made significant contributions to the understanding of cardiovascular disease.[11]

Participants range from 13 to 62 years of age at the first exam, which reflects the wide age variation in the Original Cohort. The Original Cohort restricted its sampling to white residents of Framingham Massachusetts, and, while no restriction was placed on the ethnicity or residency of their offspring or spouses, data are not available on these characteristics. As FHS was not meant to be representative of any larger population, our final sample includes consistent information on 2,497 individuals.[12] To enter our sample, an individual must have completed at least the first three exams and must not have skipped exams in the subsequent periods. Following the third exam, individuals may leave the sample through either death or attrition.

Table 1 presents summary statistics of the Offspring Cohort at the initial exam (exam 1) by gender and by whether an individual is ever, over the nine waves, observed to be on any type of antidepressant. Of the 1,241 men in our sample, 12.17% are observed at some point to be taking antidepressants; that percentage for the 1,256 is considerably higher at 24.52%. At exams 3, 6, 7, and 9, we observe the Center for Epidemiological Services - Depression (CES-D) depression score, which aggregates 20 clinically verified depression questions (each on 0 to 3 Likert Scale) into a depression summary score (Radoff, 1977). The clinically verified threshold for depression is any score at or above 16. The first row of Table 1 reports the fraction of individuals ever observed, over waves in with CES-D is observed, with a CES-D score at or above 16. Not surprisingly, depression is significantly higher for both men and women for those ever observed to take an antidepressant, but we emphasize the sizable fraction of "never"

[11] See Mahmood *et al.* (2014) for a detailed history of the Study. See Darden *et al.* (2018) and Darden (2017) for economic studies of the Original and Offspring Cohorts, respectively.

[12] Kaestner *et al.* (2014) and Darden (2017) construct very similar samples from FHS Offspring Data.

depressed using antidepressants as foreshadowing of the heterogeneity results presented below.[13][14]

FHS asks respondents the number of 12oz beers, 5oz glasses of wine, and 1.5oz liquor drinks they typically consume per week. We aggregate these to a drinks per week measure, and we follow NIAAA guidelines for light and heavy alcohol consumption based on gender: light drinking is defined as up to seven drinks per day for women and 14 drinks per day for men; heavy drinking is any number above the gender-specific thresholds (National Institute on Alcohol Abuse and Alcoholism, n.d.). At the first exam, men drink more heavily than women (despite the higher threshold for heavy drinking), and rates of heavy drinking are higher for those ever-observed to take an antidepressant (although these differences are not statistically significant).

Table 2 shows means and proportions of key variables over the nine exams. Our sample is defined such that an individual must participate in exams one through three, after which he or she can only leave the sample through death. As such, all trends presented in Table 2 are affected by both changes in behavior/outcomes and the changing composition of the sample, which we emphasize below in our dynamic system of equations model. Unfortunately, we do not observe antidepressant medication usage in exams one and two, however, the absence of these questions likely stems from the observed trends in their use: at exam three, only 1.2% of men and 1.9% of women used antidepressants. Antidepressant medication usage grows considerably within our sample over time for both men and women.

To identify the substitutability between antidepressants and risky behaviors, we begin by regressing alcohol and smoking behavior on a binary variable for any antidepressant usage at a given exam. Equation 11 presents our baseline empirical specification,

$$y_{it} = \alpha_i + x_{it}^{'}\beta + \delta d_{it} + \theta_t + \epsilon_{it}, \tag{11}$$

where y_{it} is risky behavior y for person i in year t, α_i represents an individual specific effect, x_{it} are time-varying individual characteristics characteristics, θ_t time controls, and ϵ_{it} and i.i.d. across individuals and time error component. Our variable of interest is d_{it}, which equals one if person i in exam t is taking an antidepressant. Table 3 presents results from Equation 11, in which we estimate linear probability models for each behavior listed on the top row, while interacting the antidepressant binary variable with a binary variable for female. Section 1 of Table 3 omits α_i, the individual-specific effect, exploiting both within and between variation in behaviors. Section 2 adds individual fixed effects (i.e., α_i). Results presented in sections 1 and 2 of Table 3 are conditional on age, education, and other health metrics, including blood pressure, obesity, cardiovascular disease, cancer, and exam fixed effects. Standard errors

[13] Wulsin *et al.* (2005) use FHS Offspring Cohort data to relate the exam three CES-D score to future health outcomes. They find that, relative the lowest tertile, CES-D score is statistically related to all-cause mortality but not coronary heart disease.

[14] Importantly, antidepressants are prescribed for a wide variety of conditions other than depression,including bipolar disorder, bulimia, fibromyalgia, insomnia, PTSD, and social anxiety disorder(for Medicare & Services, 2013).

are clustered at the individual level.

For men, Table 3 demonstrates initial evidence of substitutability between antidepressants and alcohol. Focusing on section 2, men are more likely to report no alcohol consumption when using an antidepressant. Alcohol cessation is driven largely from a 8.1 percentage point reduction in heavy drinking in the fixed effects specification. For each behavior, focusing on within individual variation only considerably attenuates the parameter estimates. For example, antidepressant medication is associated with a 8.0 percentage point *increase* in smoking in section 1 of the smoking results; however, this result is likely spurious as Table 1 demonstrates that, prior to antidepressant usage in exam 1, ever users were much significantly more likely to be smokers. Section 2 of the smoking results, which focus on within individual variation, shows no significant effect of antidepressants on smoking behavior for men. Table 3 also presents results on differential effects for women. Similarly, alcohol abstinence increases significantly with antidepressants, although somewhat less so for women. The negative and significant effect of antidepressants on heavy drinking is completely nullified for women, but women are less likely to drink lightly with an antidepressant (although this effect is only marginally statistically significant.)

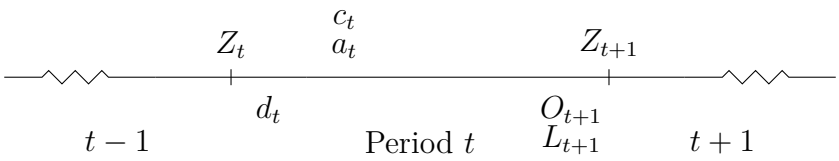
Results from Table 3 demonstrate some evidence of self-medication – during a period in which medication for depression became much better and more common, antidepressants were associated with declines in heavy drinking for men and increases in alcohol abstinence for both men and women. Our fixed effects results take a causal interpretation if there is no time-varying unobserved heterogeneity that affects both the decision to take antidepressants and behavior. While we cannot test this assumption, we interact our time fixed effects with a binary variable for ever being observed to take an antidepressant. Conditional on contemporaneous antidepressant usage, time-varying individual unobserved heterogeneity would likely generate different trends in behavior. The p-values of the F-tests that the interacted trend variables are all zero are presented at the bottom of section 3 of Table 3, along with point estimates and standard errors. For both men and women, none of the alcohol p-values suggest statistically significant different trends, and the point estimates, while slightly attenuated, are not significantly different from those in section 2.

Dynamic Empirical Model

Despite suggestive statistical evidence of self-medication, results from Table 3 are problematic in several important ways. First, a large and growing empirical literature recognizes the inherent dynamics in addictive goods (Arcidiacono *et al.*, 2007; Darden, 2017). Equation 11 is static in the sense that contemporaneous behavior is not allowed to depend on past behavior. Contemporaneous utility from alcohol and cigarettes is a function of past behavior, and the failure to model the dynamics of these behaviors will likely lead to an overestimate on the effect of antidepressants on behavior. Furthermore, while alcohol consumption, for example, may improve contemporaneous mental health, a large litera-

ture suggests that heavy alcohol consumption may harm future mental health. Second, the composition of our sample is changing over time through mortality and attrition. Especially because a.) the behaviors being modeled (drinking and smoking) may cause mortality or attrition and b.) significant antidepressant medication usage is not observed until the end of our sample period, selective mortality may significantly bias our results. Finally, estimation of each equation separately does not allow for correlation in unobserved heterogeneity across equations.

In the spirit of our two-period model presented above, and to address the limitations of our static empirical model, we estimate a dynamic system of equations of medication, alcohol, tobacco, depression, attrition and mortality. The model is an approximation of a more general structural model of behavior and outcomes, with the following timing:



Here, Z_t captures the period t state vector, which includes measures of addiction to cigarettes and alcohol, all of which have been updated following period $t-1$ behavior, as well as mental health. Period t begins with the decision to take antidepressants, d_t. Conditional on d_t, an individual chooses whether to smoke c_t and the intensity of alcohol consumption $a_t \in \{None, Light, Heavy\}$. Risky behaviors follow antidepressant medication to allow the marginal utility of alcohol and cigarettes to depend on antidepressant consumption. At the end of period t, a person may die, O_{t+1}, or attrit L_{t+1}, but conditional on remaining in the sample, the state variable Z updates.

While solution of such a model is beyond the scope of this paper, such a solution would generate demand equations for antidepressants, alcohol, and cigarettes, as well as outcome equations for mental health and mortality. Specifically, we estimate the following system:

$$p(d_{it} = d) = a(Z_{it}, X_{it}, \theta_i^d, \epsilon_{it}^d) \tag{12}$$

$$p(a_{it} = a) = a(Z_{it}, d_{it}, X_{it}, P_t, \theta_i^a, \epsilon_{it}^a) \tag{13}$$

$$p(c_{it} = c) = a(Z_{it}, d_{it}, X_{it}, P_t, \theta_i^c, \epsilon_{it}^c) \tag{14}$$

$$p(L_{it+1} = o) = a(Z_{it}, a_{it}, cit, d_{it}, M_{it}, X_{it}, \theta_i^o, \epsilon_{it}^o) \tag{15}$$

$$p(O_{it+1} = o) = a(Z_{it}, a_{it}, cit, d_{it}, M_{it}, X_{it}, \theta_i^o, \epsilon_{it}^o) \tag{16}$$

Under the assumption that each ϵ term takes an extreme value type 1 distribution, equations 12 through 16 become a system of dynamic logit equations. The equation for alcohol becomes a multinomial logit and the depression equation is a logit equation for the CES-D depression measure. X_{it} represents strictly

exogenous individual characteristics, which include both time varying and time invariant variables, and P_t presents a vector of prices of alcohol and cigarettes. The θ_i terms represent permanent unobserved heterogeneity, where the θs are correlated across equations, yielding the familiar seemingly unrelated regression framework.

As noted above, we do not observe the CES-D measure of depression in every wave, and, we only observe it in two consecutive waves (6 and 7). Thus, rather than model the CES-D measure as a dynamic process, we estimate an binary logit model for whether an individual is ever observed to be depressed over the sample period (using our 4 waves of depression data) jointly with Equations 12 through 16. Specifically, we estimate:

$$p(e_i = e) = e(X_i, \theta_i^e, \epsilon_i^e) \tag{17}$$

Importantly, by estimating Equation 17 jointly with the other equations, we are able to connect the distribution of permanent unobserved heterogeneity (i.e., θ) to an observable measure of depression (i.e., e). The depression equation is only a function of observables X_i at the beginning of the sample, and we estimate initial conditions (at exam 1) equations for drinking and smoking as a function the same variables.

Identification of the system comes from four sources. First, prices of cigarettes and alcohol, which we interact with age to generate cross-sectional variation, appear only in the demand equations for cigarettes and alcohol.[15] The assumption is that any effect of prices on our depression and mortality outcomes works through alcohol and cigarettes. Second, as discussed above, the FDA's approval of SSRIs dramatically lessened the side-effects of taking an antidepressant and opened antidepressants to new demographic markets (e.g., the elderly). We argue that the full price of antidepressants shifted exogenously between exams 3 and 4 as a result of this innovation. Third, following Arellano & Bond (1991), time-varying exogenous variables serve as implicit instruments for behavior. Finally, functional form assumptions (i.e., logit) help to identify the system (as is common in the structural econometric literature).

Following Heckman & Singer (1984) and Mroz (1999), rather than make a parametric distributional assumption for the joint distribution of the θ terms, we assume estimate a step-function for an assumed number of points of support for each term. Joint with each point of support, we estimate the probability of each type. For example, Table 5 presents the estimated points of support for each distribution in each equation, subject to the normalization that the first point of support is zero in all equations. While we cannot interpret exactly what each "type" represents, type 4 individuals of the permanent support, for example, are less likely to use exhibit any of the behaviors or outcomes relative to type 1. Type 4 individuals constitute 34.8% of our estimated sample. We estimate the system via maximum likelihood, integrating over the estimated support of the type distribution for θ. Table 4 presents estimates of the

[15] Nearly all of our sample remain in Massachusetts, so the only variation in average prices is temporal.

multinomial logit equation for both light and heavy drinking alcohol. For each behavior, we present results with and with the unobserved heterogeneity correlation. Results for light and heavy drinking are relative to not drinking. The individual parameters are difficult to interpret given the interactions and dynamics, but for both light and heavy drinking, results point to a reduction in drinking when taking an antidepressant.

Table 5 presents the estimated points of support of the distribution of unobserved heterogeneity, as well as the probability of each associated type.

To evaluate our model, we simulate both the extent to which our model can recover the time path of each behavior/outcome and the extent to which it can capture transitions between behaviors. To proceed, we replicate the baseline sample, complete with their baseline characteristics, 50 times. For men, this implies a simulated sample of 50*1,497=74,850 simulated observations. Using the estimated distribution of θ_i, we endow each simulated individual with a complete set of draws of the error structure (including ϵ). We simulate behavior and outcomes forward from the baseline period, taking care to update the state vector with endogenous terms. For example, when an individual is simulated to smoke, his or her next period "years of smoking" is updated accordingly, regardless of if the person actually smoked. Figure 2 presents the time path of each behavior/outcome for men and women. In all cases, our model produces the observed patterns quite well.

Table 6 presents simulated transitions for each behavior along with the analogous transition proportion in the data for both men and women. For example, conditional on drinking heavily in period $t-1$, 61.8% of men are simulated to be drinking heavily in period t. In the data, that percentage is 59.4%. Capturing transitions is more difficult than capturing averages, yet our model does a good job of recovering the transitions in the data.

Simulation

We simulate our estimated dynamic model under various counterfactual scenarios to shed light on the dynamics behavior and outcomes. As a natural first step, evaluate a counterfactual in which all sample participants take an antidepressant as soon as SSRIs become available and there onward (i.e., exam 4 through 9). Figure 3a presents results for the entire sample. Heavy drinking declines by approximately four percentage points and remains lower through the rest of the simulated sample. Figures 3b and Figures 3c break the results from Figure 3a by gender, which demonstrates that men are primarily changing heavy drinking. Figures 3d and Figure 3e break the results from Figure 3a by whether the person is simulated to ever experience depression. Surprisingly, the reduction in heavy drinking associated with antidepressants is entirely driven by those simulated to never be depressed.

As addiction may be important in substitution away from alcohol, Figure 4 presents results in which we simulate our model assuming that the parameters on all terms reflecting past alcohol consumption

are set to zero. Not surprisingly, in panels b., heavy alcohol consumption plummets. However, the drop in heavy alcohol consumption is accompanied by only a modest 1.7 percentage point increase in antidepressant usage by the end of the simulated panel. Interestingly, the increase in antidepressant usage in a simulation without alcohol addiction is driven by women and those who are simulated to ever be depressed. This finding suggests that addiction is preventing substitution away from alcohol.

Discussion

Conclusion

References

Anton, Raymond, O'Malley, Stephanie, Ciraulo, Domenic, & et al. 2006. Combined Pharmacotherapies and Behavioral Interventions for Alcohol Dependence. The COMBINE Study: A Randomized Controlled Trial. *JAMA*, **295**(17), 2003–2017.

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

Arellano, Manuel, & Bond, Stephen. 1991. Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations. *The review of economic studies*, **58**(2), 277–297.

Bacolod, Marigee, jesse Cunha, & Shen, Yu-Chu. 2017. The Impact of Alcohol on Mental Health, Physical Fitness, and Job Performance.

Beck, Arne, Crain, A. Lauren, Solberg, Leif, & et al. 2011. Severity of Depression and Magnitude of Productivity Loss. *Annals of Family Medicine*, **9**(4), 305–312.

Berndt, Ernst R., Finkelstein, Stan N., Greenberg, Paul E., & et al. 1998. Workplace performance effects from chronic depression and its treatment. *Journal of Health Economics*, **17**, 511–535.

Blanchflower, David G., & Oswald, Andrew J. 2016. Antidepressants and age: A new form of evidence forU-shaped well-being through life. *Journal of Economic Behavior and Organization*, **127**, 46–58.

Bolton, James M., Robinson, Jennifer, & Sareen, Jitender. 2009. Self-medication of mood disorders with alcohol and drugs in the National Epidemiologic Survey on Alcohol and Related Conditions. *Journal of Affective Disorders*, **115**(3), 367 – 375.

Brody, Debra J., Pratt, Laura, A., & Hughes, Jeffery P. 2018. *Prevalence of Depression Among Adults Aged 20 and Over: United States, 2013–2016*. Data Brief 303. National Center for Health Statistics.

Case, Anne, & Deaton, Angus. 2015. Rising morbidity and mortality in midlife among white non-Hispanic Americans in the 21st century. *Proceedings of the National Academy of Sciences*, **112**(49), 15078—15083.

Case, Anne, & Deaton, Angus. 2017. Mortality and Morbidity in the 21st Century. *Brookings Papers on Economic Activity*, **Spring**, 397–476.

Crost, Benjamin. 2012. The Effect of Alcohol Availability on Marijuana Use: Evidence from the Minimum Legal Drinking Age. *Journal of Health Economics*, **31**(1), 112–121.

Crum, Rose M., Mojtabai, Ramin, Lazareck, Lamuel, & et al. 2013. A Prospective Assessment of Reports of Drinking to Self-medicate Mood Symptoms With the Incidence and Persistence of Alcohol Dependence. *JAMA Psychiatry*, **70**(7), 718–726.

Darden, Michael. 2017. A Dynamic Stochastic Model of Lifetime Smoking Behavior. **125**(4), 1465–1522.

Darden, Michael, Gilleskie, Donna, & Strumpf, Koleman. 2018. Smoking and Mortality: New Evidence from a Long Panel. **59**(3), 1571–1619.

Deykin, Eva, Levy, Janice, & Wells, Victoria. 1987. Adolescent Depression, Alcohol and Drug Abuse. *American Journal of Public Health*, **77**(2), 178–182.

Dinardo, John. 2001. Alcohol, Marijuana, and American Youth: The Unintended Consequences of Government Regulation. *Journal of Health Economics*, **20**(6), 991–1010.

Fluharty, Meg, Taylor, Amy E., meryem Grabski, & Munafo, Marcus. 2017. The Association of Cigarette Smoking With Depression and Anxiety: A Systematic Review. *Nicotine and Tobacco Research*, **19**(1).

for Medicare, Center, & Services, Medicaid. 2013. *Antidepressant Medications: Use in Adults*. Report.

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

Harris, Katherine M., & Edlund, Mark J. 2005. Self-Medication of Mental Health Problems: New Evidence from a National Survey. *Health Services Research*, **40**(1), 117–134.

Heckman, James J., & Singer, Burton. 1984. A Method for Minimizing the Impact of Distributional Assumptions in Econometric Models for Duration Data. *Econometrica*, **52**(2), 271–320.

Hillhouse, Todd, & Porter, Joseph. 2015. A brief history of the development of antidepressant drugs: From monoamines to glutamate. *Experimental Clinical Psychopharmacology*, **23**(1), 1–21.

Horwitz, Allan. 2010. How an Age of Anxiety Became an Age of Depression. *The Millbank Quarterly*, **88**(1), 112–138.

Hughes, John R., Stead, Lindsay F, & Lancaster, Time. 2011. Antidepressants for smoking cessation. *Cochrane Database of Systematic Reviews*, **8**.

Kaestner, Robert, Darden, Michael, & Lakdawalla, Darius. 2014. Are Investments in Disease Prevention Complements? The Case of Statins and Health Behaviors. *Journal of Health Economics*, **36**, 151–163.

Kessler, Ronald C. 2012. The Costs of Depression. *Psychiatric Clinics of North America*, **35**(1), 1–14.

Khantzian, Edward. 1985. The Self-Medication Hypothesis of Addictive Disorders: Focus on Heroin and Cocaine Dependence. *The American Journal of Psychiatry*, **142**(11), 1259–1264.

Lembke, Anna. 2012. Time to Abandon the Self-Medication Hypothesis in Patients with Psychiatric Disorders. *The American Journal of Drug and Alcohol Abuse*, **38**(6), 524–529.

Ludwig, Jens, Marcotte, Dave E., & Norberg, Karen. 2009. Anti-depressants and suicide. *Journal of Health Economics*, **28**, 659–676.

Mahmood, SS., Levy, D., Vasan, RS., & TJ., Wang. 2014. The Framingham Heart Study and the epidemiology of cardiovascular disease: a historical perspective. *Lancet*, **383**(9921).

Mroz, Thomas. 1999. Discrete Factor Approximations in Simultaneous Equation Models: Estimating the Impact of a Dummy Endogenous Variable on a Continuous Outcome. *Journal of Econometrics*, **92**(2), 233–274.

National Institute on Alcohol Abuse and Alcoholism. *Drinking Levels Defined*. Report.

O'Connor EA, Whitlock EP, Gaynes B et al. 2009. Screening for Depression in Adults and Older Adults in Primary Care: An Updated Systematic Review. **115**(3), 367 – 375.

Papageorge, Nicholas W. 2016. Why medical innovation is valuable: Health, human capital, and the labor market. *Quantitative Economics*, **7**(3), 671–725.

Powell, David, Pacula, Rosalie Liccardo, & Jacobson, Mireille. 2018. Do medical marijuana laws reduce addictions and deaths related topain killers? *Journal of Health Economics*, **58**, 29–42.

Radoff, Lenore Sawyer. 1977. The CES-D Scale: A Self-Report Depression Scale for Research in the General Population. *Applied Psychological Measurement*, **1**(3), 385–401.

Ruhm, Christopher J. 2018. DEATHS OF DESPAIR OR DRUG PROBLEMS? *NBER Working Paper*.

Wulsin, Lawson R., Evans, Jane C., & et al. 2005. Depressive Symptoms, Coronary Heart Disease, and Overall Mortality in the Framingham Heart Study. *Psychosomatic Medicine*, **67**(1), 697–702.

Wuslin, Lawson R., Evans, Jane C., Vasan, Ramachandran S., & et al. 2005. Depressive Symptoms, Coronary Heart Disease, and Overall Mortality in the Framingham Heart Study. *Psychosomatic Medicine*, **67**, 697–702.

Main Tables

Table 1: Baseline Characteristics by Gender and Ever Anti-Depressant Usage

	Men = 1,241			Women = 1,256		
	Never (87.83%)	Ever (12.17%)	p-value	Never (75.48%)	Ever (24.52%)	p-value
Ever Depressed	0.166	0.391	0.000	0.243	0.523	0.000
Drinking Never	0.098	0.099	0.964	0.155	0.169	0.566
Light	0.657	0.596	0.142	0.662	0.643	0.529
Heavy	0.245	0.305	0.114	0.182	0.188	0.819
Smokes	0.408	0.530	0.005	0.368	0.484	0.000
Ever Has Cance	0.336	0.285	0.212	0.214	0.091	0.000
Ever Has CVD	0.414	0.411	0.941	0.343	0.276	0.030
Dies Before Exam 9	0.372	0.397	0.540	0.203	0.234	0.243
Age	37.152	36.199	0.286	36.993	33.396	0.000
Education Less than HS	0.017	0.026	0.440	0.006	0.003	0.528
HS Grad.	0.304	0.272	0.419	0.379	0.390	0.732
Some College	0.423	0.404	0.659	0.461	0.481	0.551
College or More	0.185	0.252	0.053	0.098	0.078	0.290
BMI	26.799	27.170	0.230	24.391	24.509	0.702
Obese	0.162	0.199	0.263	0.114	0.120	0.767

Notes: $n = 2,497$. Statistics are calculated from Exam 1, which took place between 1971 and 1975. The sample is constructed such that an individual must be present for exams 1 through 3. Rows for never and ever antidepressant usage reflect whether the person was ever observed to take an antidepressant. Ever depressed is a binary variable for whether an individual is ever observed with a CES-D score over 15 over the nine waves. Light drinking is defined as seven or fewer drinks per week for women and 14 or fewer drinks per week for men. Heavy drinking is defined as more than seven drinks per week for women and more than 14 drinks per week for men.

Table 2: Sample Behaviors over Time by Gender.

Men, $n = 9,586$								
Exam	Count	Year	Age	Antidepressant	Never	Light	Heavy	Smoke
1	1241	1973	37.036	0.000	0.098	0.649	0.252	0.423
2	1241	1981	44.911	0.000	0.180	0.571	0.248	0.418
3	1241	1985	49.267	0.010	0.212	0.555	0.233	0.269
4	1198	1989	52.422	0.013	0.264	0.539	0.197	0.234
5	1122	1993	55.603	0.020	0.266	0.546	0.188	0.178
6	1043	1997	59.301	0.036	0.291	0.548	0.161	0.129
7	1005	2000	61.867	0.056	0.276	0.554	0.170	0.116
8	845	2007	67.424	0.088	0.249	0.591	0.161	0.090
9	650	2011	71.462	0.105	0.269	0.554	0.177	0.055
Women, $n = 10,169$								
Exam	Count	Year	Age	Antidepressant	Never	Light	Heavy	Smoke
1	1256	1973	36.111	0.000	0.158	0.658	0.184	0.396
2	1256	1981	43.994	0.000	0.284	0.512	0.204	0.314
3	1256	1985	48.362	0.021	0.350	0.473	0.177	0.278
4	1225	1989	51.740	0.036	0.343	0.507	0.150	0.219
5	1183	1993	55.173	0.049	0.332	0.525	0.143	0.174
6	1131	1997	59.034	0.084	0.450	0.417	0.133	0.141
7	1107	2000	61.822	0.112	0.388	0.451	0.162	0.114
8	972	2007	67.418	0.186	0.321	0.515	0.164	0.099
9	783	2011	71.775	0.217	0.354	0.469	0.178	0.056

Notes: $n = 19,755$. Statistics are calculated from nine exams, which took place between 1971 and 2011. The sample is constructed such that an individual must be present for exams 1 through 3, after which some individuals are lost to death. Light drinking is defined as seven or fewer drinks per week for women and 14 or fewer drinks per week for men. Heavy drinking is defined as more than seven drinks per week for women and more than 14 drinks per week for men.

Table 3: Reduced-Form Estimates of Antidepressants on Behavior

	Never Drinking	Light Drinking	Heavy Drinking	Smoking
Section 1: Time Fixed Effects Only				
Antidepressant	0.132***	-0.062	-0.070**	0.080**
	(0.042)	(0.040)	(0.027)	(0.038)
*Female	-0.055	-0.014	0.069**	-0.054
	(0.050)	(0.046)	(0.033)	(0.043)
Section 1: Individual and Time Fixed Effects				
Antidepressant	0.076**	0.005	-0.081***	0.000
	(0.032)	(0.032)	(0.025)	(0.022)
*Female	-0.031	-0.056	0.086***	-0.030
	(0.037)	(0.037)	(0.029)	(0.027)
Section 1: Individual and Time Fixed Effects, Separate Trends				
Antidepressant	0.056*	0.024	-0.080***	0.018
	(0.033)	(0.033)	(0.026)	(0.023)
*Female	-0.031	-0.055	0.086***	-0.029
	(0.037)	(0.037)	(0.029)	(0.027)
p-value	0.146	0.124	0.455	0.010

Notes: n =19,755 person/year observations in all regressions. All regressions include controls for age binary variables, education, cardiovascular disease indicators, cancer indicators, and body mass index. * $p < 0.1$, ** $p < 0.05$ *** $p < 0.01$.

Table 4: Selected Parameter Estimates

	Light Drinking				Heavy Drinking			
	Beta	S.E.	Beta	S.E.	Beta	S.E.	Beta	S.E.
Medication	-0.404	0.171	-0.425	0.238	-1.083	0.278	-1.211	0.369
Medication*	0.200	0.186	0.210	0.254	0.721	0.296	0.726	0.391
Female								
Medication*	-0.119	0.174	-0.106	0.230	0.366	0.272	0.519	0.352
Depression								
L. Heavy Drinking	2.381	0.044	1.196	0.063	3.544	0.137	2.307	0.161
L. Light Drinking	2.794	0.084	1.679	0.118	6.464	0.150	4.089	0.183
L. Smoking	-0.111	0.071	0.154	0.099	0.084	0.097	0.326	0.138
Depression	-0.179	0.049	-0.022	0.107	-0.182	0.071	0.078	0.144
Years Smoking	0.002	0.002	-0.007	0.003	0.013	0.003	-0.001	0.004
Years Smoking Cessation	0.007	0.002	0.010	0.003	0.013	0.003	0.022	0.004
Female	-0.278	0.045	-0.704	0.076	-0.248	0.063	-0.824	0.108
Age								
(35, 40]	-0.326	0.128	-0.206	0.150	-0.337	0.176	-0.248	0.204
(40, 45]	-0.466	0.123	-0.411	0.146	-0.424	0.171	-0.316	0.202
(45, 50]	-0.345	0.128	-0.280	0.154	-0.382	0.179	-0.242	0.214
(50, 55]	-0.500	0.136	-0.433	0.166	-0.579	0.193	-0.434	0.237
(55, 60]	-0.471	0.152	-0.472	0.186	-0.663	0.219	-0.645	0.272
(60, 65]	-0.475	0.176	-0.548	0.218	-0.657	0.257	-0.739	0.322
(65, 70]	-0.480	0.206	-0.633	0.255	-0.757	0.305	-0.949	0.384
(70, 75]	-0.664	0.241	-0.970	0.299	-1.078	0.360	-1.549	0.455
>75	-0.764	0.302	-1.339	0.373	-1.482	0.454	-2.369	0.572
Education								
High School	0.203	0.089	0.372	0.149	0.123	0.125	0.292	0.229
Some College	0.422	0.089	0.868	0.149	0.339	0.124	0.738	0.222
College or More	0.578	0.102	1.005	0.176	0.514	0.141	0.803	0.259
MA Alcohol Tax * Age	-0.007	0.004	0.005	0.004	-0.004	0.005	0.017	0.007
Constant	-1.083	0.154	1.623	0.291	-4.227	0.249	-1.522	0.429
θ_1			0.000	0.000			0.000	0.000
θ_2			-4.605	0.234			-3.256	0.365
θ_3			-1.114	0.309			0.906	0.372
θ_4			-2.054	0.168			-4.460	0.385

Notes: $n = 19,755$. Selected parameter estimates are from models estimated on data in exams 2-9.

Table 5: Unobserved Heterogeneity Distribution

	Probability	Ever Depressed	Anti-depressants	Light Drinking	Heavy Drinking	Smoking	Attrition	Death
θ_1	0.275	0.000	0.000	0.000	0.000	0.000	0.000	0.000
θ_2	0.153	0.589***	0.377*	-4.605***	-3.256***	0.371**	-0.077	0.063
θ_3	0.224	0.115	0.087	-1.114***	0.906***	0.456***	0.318	0.535**
θ_4	0.348	0.333*	-0.019	-2.054***	-4.460***	0.303**	-0.228	0.143

Notes: $n = 19,755$. Selected parameter estimates are from models estimated on data in exams 2-9. * $p < 0.1$, ** $p < 0.05$ *** $p < 0.01$.

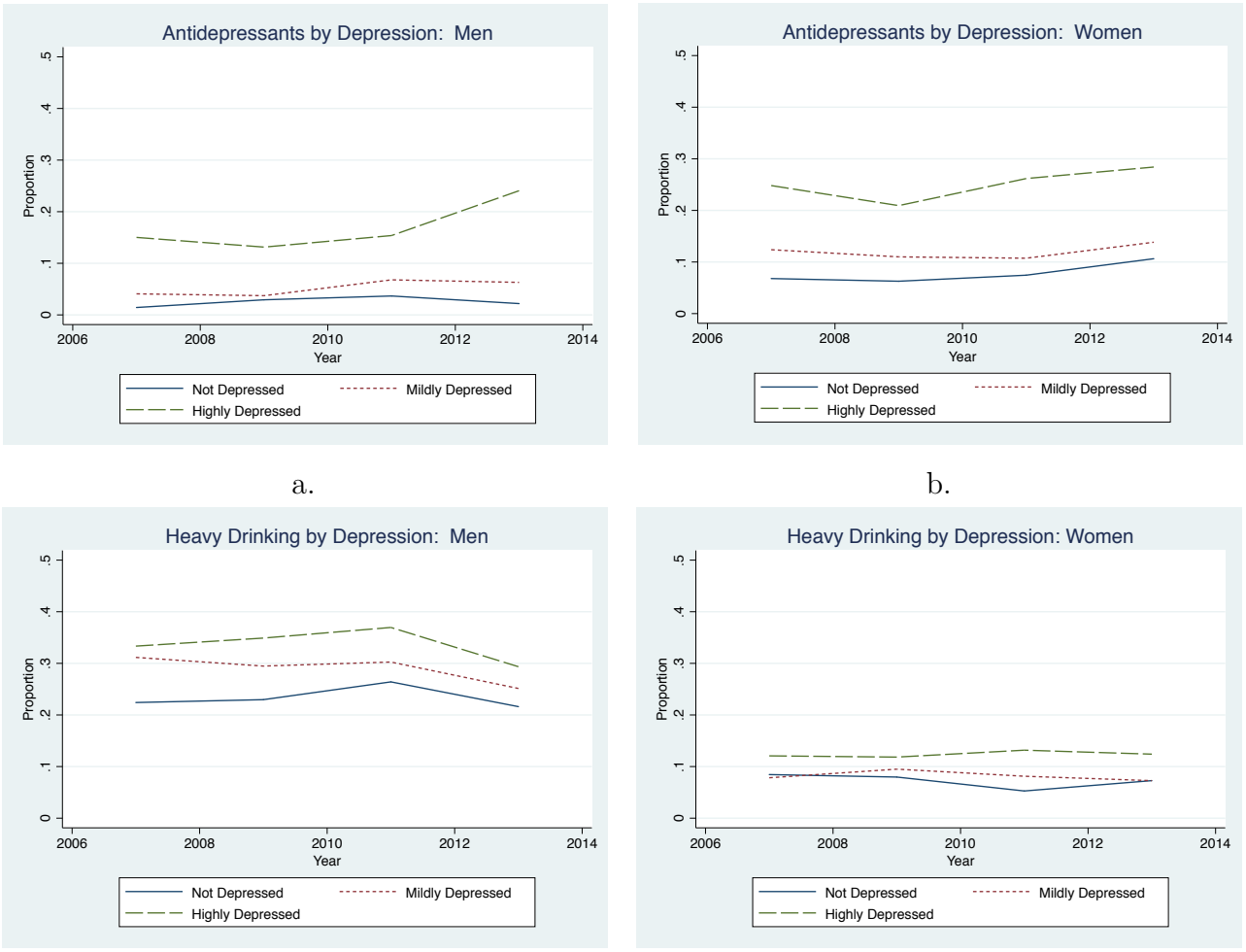
Table 6: Model Fit: Transitions.

	No Drinking=1		Light Drinking=1		Heavy Drinking=1		Antidepressants=1		Smoking=1	
	Data	Sim.	Data	Sim.	Data	Sim.	Data	Sim.	Data	Sim.
Lagged Behavior $t-1$										
No Drinking=1	0.672	0.724	0.229	0.263	0.009	0.013	0.084	0.113	0.138	0.139
Light Drinking=1	0.157	0.166	0.689	0.727	0.097	0.107	0.052	0.060	0.136	0.143
Heavy Drinking=1	0.047	0.051	0.300	0.331	0.594	0.618	0.054	0.059	0.230	0.213
Antidepressants=1	0.372	0.435	0.366	0.428	0.118	0.137	0.623	0.744	0.144	0.113
Smoking=1	0.281	0.278	0.427	0.484	0.224	0.238	0.062	0.055	0.689	0.669

Notes: Results are from models estimated on data in exams 2-9.

Main Figures

Figure 1: Depression, Anti-Depressants, and Alcohol: Evidence from NHANES

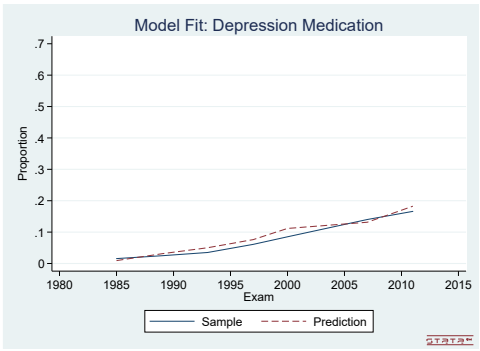


a. b.

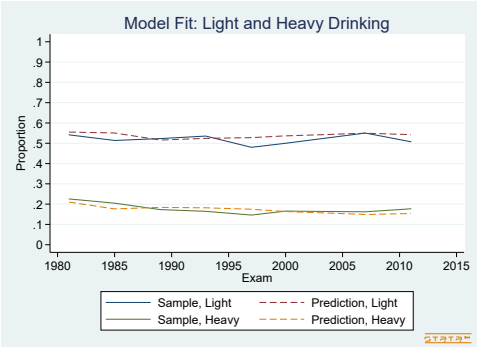
c. d.

Notes: Author's calculations from NHANES data from 2007-2013. Proportions are weighted by the NHANES full sample 2-year interview weight. Proportions are presented by tertiles of the Patient Health Questionnaire (PHQ-9) Depression Score. $n = 16,940$.

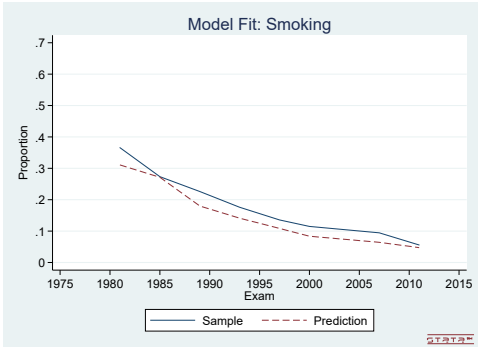
Figure 2: Model Fit



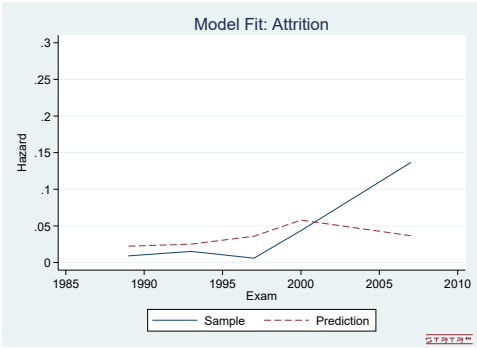
a.



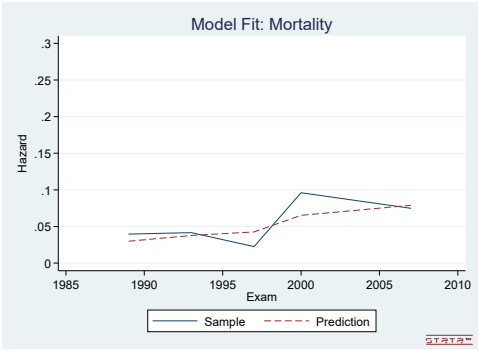
b.



c.



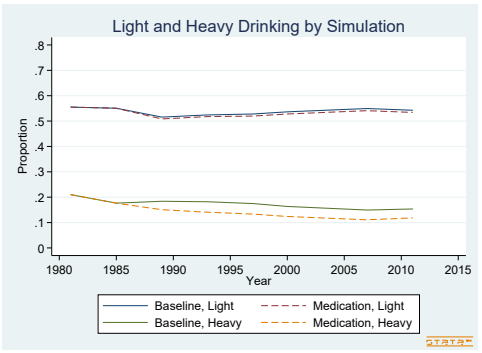
d.



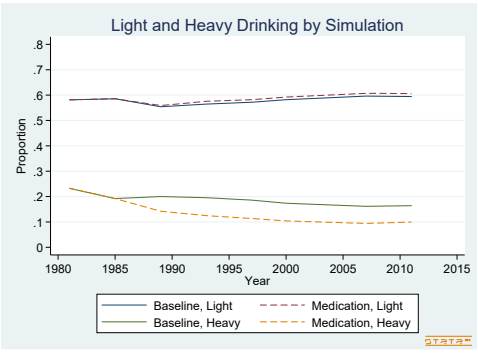
e.

Notes: Each figure presents results from the baseline simulation of our estimated dynamic model.

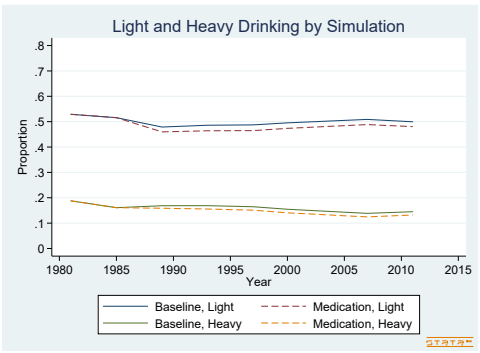
Figure 3: Comprehensive Antidepressants vs. Baseline: Alcohol Consumption



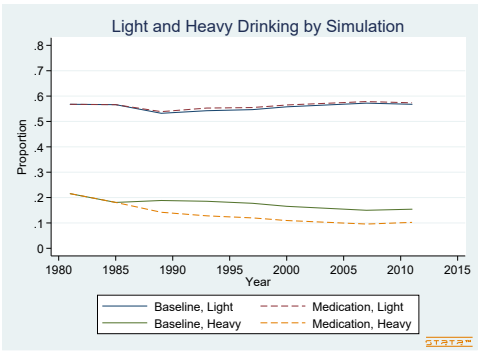
a.



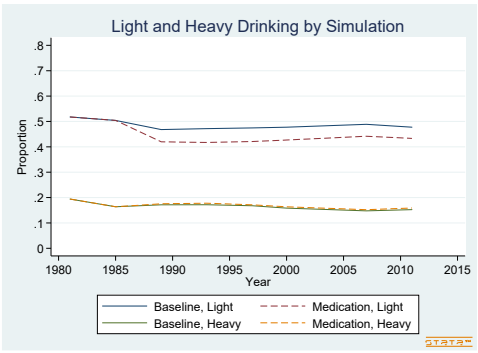
b.



c.



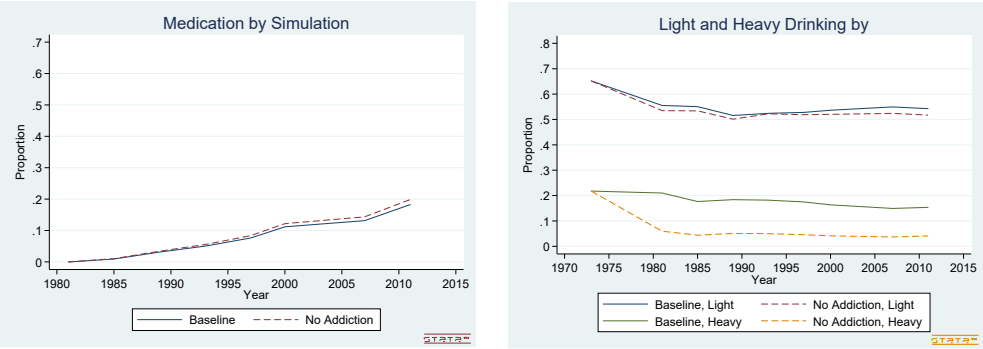
d.



e.

Notes: Each figure presents baseline simulated trends in light and heavy drinking as well as those behaviors when we impose that all individuals take an antidepressant from exam 4 onwards. Figure 3a presents the simulations for the entire sample. Figures 3b and 3c present results separately for men and women. Figures 3d and 3e present results for those never simulated to be depressed and ever simulated to be depressed, respectively.

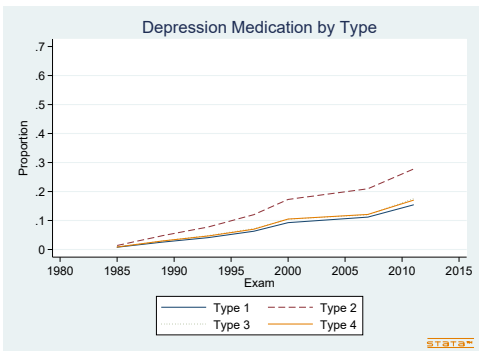
Figure 4: No Dependence on Past Alcohol vs. Baseline



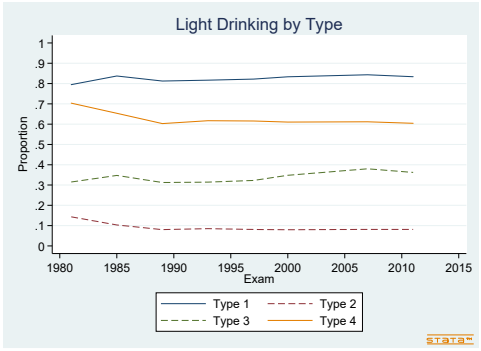
a. b.

Notes: Each figure presents baseline trends relative to a counterfactual simulation in which the dependence on past alcohol consumption is removed in all behavioral equations. Figure Figure 4a presents trends in antidepressants by simulations, and Figure 4b presents trends in alcohol consumption by simulation.

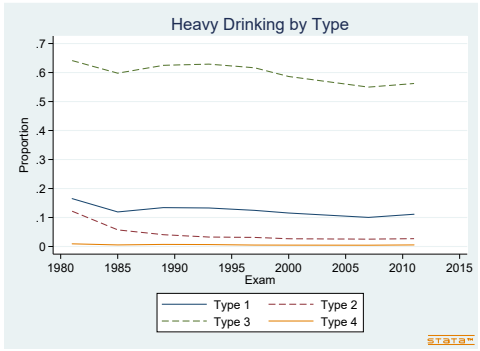
Figure 5: Behaviors and Outcomes by Unobserved Type



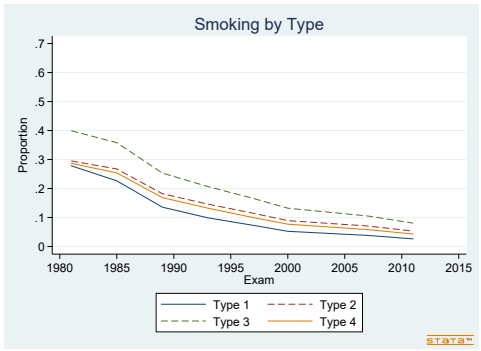
a.



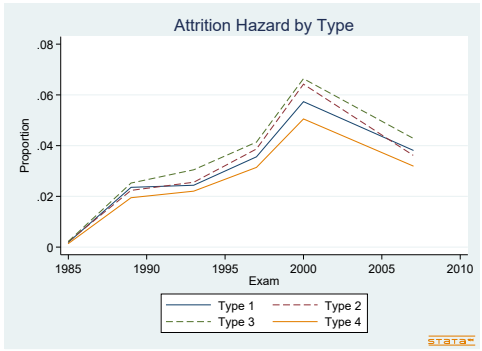
b.

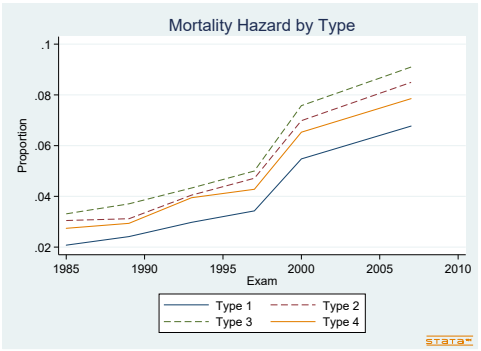


c.



d.





e.

Notes: Each figure presents results from the baseline simulation of our estimated dynamic model by each of the four unobserved types.