Employment Effects of Healthcare Policy: Evidence from the 2007 FDA Black Box Warning on Antidepressants

Aline Bütikofer*
Department of Economics, Norwegian School of Economics

Christopher Cronin†
Department of Economics, University of Notre Dame

Meghan Skira‡
Department of Economics, University of Georgia

October 18, 2018

Abstract: Public policies and regulatory decisions aimed at improving health may have indirect effects on outcomes such as education and employment. We study the labor market effects of the US Food and Drug Administration's 2007 expanded black box warning on antidepressants. Using nationally representative data from the National Survey on Drug Use and Health and a difference-in-differences strategy, we find that employment among women aged 35-49 with a history of depression decreased by 6.1 percent (4.4 percentage points) in response to the warning. We also present evidence that antidepressant use among women aged 35-49 decreased along both the extensive and intensive margins after the warning. We find no employment or antidepressant response among men or among women younger than 35.

Keywords: Mental Health, Antidepressants, Black Box Warnings, Employment.
JEL Classification: I18, D83, J22

*aline.butikofer@nhh.no
†ccronin1@nd.edu
‡skira@uga.edu

1 Introduction

Economists view health as a form of human capital (Grossman, 1972). Supporting this perspective is a large body of empirical research indicating that both physical and mental health have important effects on educational attainment, labor supply, and earnings (Currie and Madrian, 1999). Public policies and regulatory decisions aimed at improving health may therefore have indirect and unintended effects on scholastic and work-related outcomes. In order to understand and evaluate the full implications of such policies and regulations, it is important that these indirect effects are understood.

In recent years, policy makers have made numerous attempts to curb mental illness in the US; diagnosis rates have grown steadily since at least the mid-1990s (see Figure 1).[1] According to the National Association of Mental Illness (NAMI), approximately one in five adults in the US experiences mental illness in any given year, most commonly depression and anxiety disorders.[2] Depression and anxiety, in particular, have a disproportionate effect on middle-aged women; they experience depression at almost twice the rate of similarly aged men and 1.3 times the rate of women under 40 (CDC). Ettner et al. (1997), and more recently Cronin et al. (2018), show empirically that these disorders can have a substantial negative impact on employment and wages. For example, Ettner et al. (1997) estimate that individuals with a mental illness are roughly 13 percentage points less likely to work and that employed women with a mental illness earn 30 percent less annually than women without a mental illness. Prescription antidepressants are the most common treatment for depression and anxiety. Since the year 2000, antidepressant use in the US has risen by 65 percent (CDC). Currently, antidepressants are one of the three most commonly used drug classes in the US (CDC). Between 2011 and 2014, 21.2 percent of all middle-aged women, and 12.7 percent of all Americans, had consumed an antidepressant in the past month (CDC).

The extent to which antidepressants protect the employment and productivity of those with depression is an open question. Furthermore, it is unclear the extent to which government policy could or should impact labor market outcomes by promoting mental health treatment. We provide insight on these questions by analyzing the labor market effects of the US Food and Drug Administration's (FDA) 2007 expanded black box warning on antidepressants. A black box warning appears on a prescription drug's label (or package insert) and is meant to call attention to serious adverse or life-threatening side effects. It is the strongest warning required by the FDA. In October 2004, the FDA required that drug manufacturers of 36 antidepressants include a black box warning of increased risk of suicidal thinking and behavior (suicidality) for children and adolescents.

[1] Examples include state-level mental health parity laws passed throughout the 1990s and early 2000s; the (federal) Mental Health Parity Act of 1996 and Mental Health Parity and Addiction Equity Act of 2008; and the Patient Protection and Affordable Care Act of 2010, which made mental health one of 10 essential health benefits all individual and small-group insurance plans must cover.

[2] Depression and anxiety are often discussed as a single mental illness because they share common symptoms, treatments, and causes. In this paper, we do not distinguish between the two, but recognize there are detailed and nuanced clinical differences.

The warning was expanded on May 2, 2007 to include young adults aged 18-24. Furthermore, the expanded warning contained language recommending patients of all ages who are started on antidepressant therapy be monitored carefully.

While a number of studies have examined the impact of the 2004 warning on pediatric and adolescent antidepressant use, few have examined the adult response and even fewer have analyzed the prescription utilization response to the expanded 2007 warning. More importantly, there has been no research on the labor market effects of the 2007 warning, which is the primary focus of our paper. The expanded warning may have had important labor market implications for a variety of reasons. First, the warning targeted young adults aged 18-24, a group that is less likely to live in their parents' home and more likely to be in the workforce. Second, the limited research that does examine how adult antidepressant use responded to the 2004 warning that targeted children and adolescents finds a significant negative adult response; thus, it is possible adults older than 24 responded to the 2007 warning. Last, the 2007 expanded warning recommended that new patients of all ages receive enhanced monitoring.

We use cross-sectional data from the nationally representative National Survey on Drug Use and Health (NSDUH) to estimate the effects of the FDA's 2007 expanded black box warning on the employment of men and women aged 18-49. We use a difference-in-differences strategy that compares the employment of ever-depressed individuals, before and after the warning, to the employment of never-depressed individuals. We find that employment among ever-depressed women aged 35-49 decreased by roughly 6.1 percent (4.4 percentage points) in response to the expanded warning. The warning did not have a significant effect on employment for ever-depressed men or for ever-depressed women younger than 35. These results survive numerous robustness tests and are corroborated in analyses using another nationally representative dataset, the National Health Interview Survey (NHIS).

To further understand these findings, we explore several mechanisms through which the 2007 expanded black box warning may have affected employment. We study the most obvious mechanism, a decrease in antidepressant use, using the Medical Expenditure Panel Survey (MEPS). Consistent with the employment effects, we find that (i) depressed women aged 35-49 were 21 percent (8 percentage points) less likely to use antidepressants in the eight months following the warning than they were prior to the warning and (ii) antidepressant use among men and among women younger than 35 did not respond to the warning. Consistent with the warning's language regarding new antidepressant users, we find the extensive margin decrease in antidepressant use was most pronounced among new users. That said, we also find a significant decrease in the average antidepressant dose after the 2007 expanded warning, which could further explain the negative employment effects. Depressed women aged 35-49 who took antidepressants before and after the warning reduced their average dose by 36-43 percent after the warning, while men and women under the age of 35 experienced no change. We find no evidence that the 2007 expanded warning

had an effect on alcohol consumption, marijuana use, or the use of psychotherapy.

The fact that 35-49 year old women, who were not targeted by the FDA, responded to 2007 warning, while 18-24 year olds, who were targeted, did not respond is somewhat surprising. That said, neither finding is without precedent. Several papers find that in response to the 2004 black box warning on antidepressants, non-targeted adults reduced their utilization (Olfson et al., 2008; Libby et al., 2009; Parkinson et al., 2014). Given that the 2007 warning expanded the 2004 warning to an older age group (18-24 year olds) and contained language that applied to all new antidepressant users, there is reason to expect an even larger response among adults in 2007. Responses among non-targeted groups have been observed with other FDA advisories and warnings as well (see Dusetzina et al. 2012 for a summary). For example, Dorsey et al. (2010) find that the FDA's 2005 black box warning on antipsychotic medication, which targeted elderly individuals with dementia, led to statistically significant declines in atypical antipsychotic drug use among non-elderly individuals without dementia. Regarding non-response among targeted 18-24 year olds, many studies find no change in drug utilization among targeted groups in response to various FDA warnings (again, see Dusetzina et al. 2012 for a summary). Moreover, the collective literature on the impact of the 2004 warning on children and adolescents suggests declines in extensive margin antidepressant use of 20 percentage points or more (Gibbons et al., 2007; Nemeroff et al., 2007; Olfson et al., 2008; Libby et al., 2009; Busch et al., 2010, 2011; Parkinson et al., 2014). Given the large response in 2004 and the narrow age group that received "new information" about suicidality risk (i.e., those 19-21 at the time of the 2004 warning), it seems plausible that young adults fully updated their priors on the dangers of antidepressant use in 2004.

Among the papers in the economics literature that examine the black box warning on antidepressants, our focus on (i) the 2007 expanded warning and (ii) indirect employment effects is unique. In a related paper, Busch et al. (2014) study indirect effects of the 2004 warning. Using annual cross-sectional data from the NSDUH, the authors estimate the impact of the warning on the academic and behavioral outcomes of adolescents aged 12-17 with probable depression via a difference-in-differences design. They find that adolescents with recent probable depression experienced a drop of 0.14 points in grade point average after the 2004 warning relative to those without probable depression. Consistent with our finding that the employment declines are concentrated among females, the decline they find is entirely driven by girls. They also find increases in substance use and delinquency measures for girls. Our findings are further supported by studies from the psychiatry literature, which suggest women may be more responsive to antidepressants than men (Kornstein et al., 2000; Young et al., 2009; Sramek et al., 2016).

Our results underscore that FDA and other public health policies and regulatory actions can have significant and potentially unintended consequences beyond health and clinical considerations. Regarding the 2007 expanded black box warning in particular, employment effects seem to be limited to depressed women in their mid-30s to late-40s, a demographic with a relatively high

rate of antidepressant use. Given that these women are in their prime-working years, the overall economic impact is not trivial. In Section 6, we argue that, among other consequences, the warning decreased overall labor force participation by 0.23 percentage points and led to roughly $13 billion in lost wages. Given existing research showing that the 2004 black box warning actually *increased* suicide rates (Gibbons et al., 2007; Busch et al., 2014) and worsened academic and behavioral outcomes for adolescents (Busch et al., 2014), our findings echo the concerns of other researchers regarding the efficacy of the FDA's black box warnings on antidepressants.

Our paper also contributes to a growing literature that studies the labor market effects of medical treatments and innovations by exploiting treatment-related policies and information shocks. Two papers exploit the removal of Vioxx, a type of Cox-2 inhibitor (used for the treatment of pain and inflammation), from the worldwide market in 2004 due to concerns over negative side effects. Garthwaite (2012) shows that Vioxx's removal, and the resulting reduction in the use of all Cox-2 inhibitors, decreased the probability of working for individuals in the US with joint conditions by 22 percentage points. Bütikofer and Skira (2018) find Vioxx's market entry decreased sickness absence days among individuals with joint pain by 7-12 percent in Norway, and the withdrawal of Vioxx increased sickness days by 12-16 percent and increased the quarterly probability of receiving disability benefits by 6-15 percent. Daysal and Orsini (2012) study the employment effects of hormone replacement therapy (HRT) on middle-aged women by exploiting the release of findings from the Women's Health Initiative Study, which documented health risks associated with long-term HRT use. Their estimates imply HRT use increases the short-term employment of middle-aged women by 33 percentage points. Shapiro (2018) finds direct-to-consumer advertising of antidepressants significantly decreases missed days of work. In sum, our paper complements the existing literature which shows that regulatory policies and information shocks related to medical treatments can have substantial spillover effects on the labor supply of those using the treatments.[3]

The remainder of the paper is organized as follows: Section 2 details the history of the FDA's black box warning on antidepressants. In Section 3, we describe the data and our empirical strategy. Section 4 contains our estimates of the impact of the 2007 expanded black box warning on employment, as well as robustness checks. In Section 5, we explore mechanisms through which the warning may have altered employment. Finally, in Section 6, we discuss the implications of our findings for public policy moving forward.

[3]Our paper also relates to a literature that studies the labor market effects of medical treatments, but does not exploit regulatory policies or information shocks (e.g., Nichol (2001); Nichol et al. (2009); Thirumurthy et al. (2008); Papageorge (2016); Jeon and Pohl (2018); Harris et al. (2018)) as well as a literature on the effects of regulatory policies and information shocks on medical treatment use, which does not study impacts on labor market outcomes (e.g., Chintagunta et al. (2009); Collins et al. (2013); Balat et al. (2018)).

2 Background

Modern pharmaceutical antidepressants were first introduced in the 1950s and are currently one of the three most commonly used drug classes in the US (CDC). The most popular antidepressants today are selective serotonin reuptake inhibitors (SSRIs), the first of which, fluoxetine (marketed as Prozac), was introduced in 1987.[4] SSRIs grew in popularity, in part, because they have fewer side effects and are less toxic in overdose compared to earlier antidepressants, such as tricyclic antidepressants (TCAs). In addition to depression, SSRIs and other antidepressants are now commonly prescribed for a multitude of ailments, including generalized anxiety, panic disorder, obsessive compulsive disorder, post-traumatic stress disorder, eating and sleep disorders, pain, and migraines (Wong et al., 2017).

The introduction of SSRIs produced an abundance of clinical research. By the late 1990s and early 2000s, a growing body of evidence indicated that some younger patients, particularly those under age 18, experienced an increase in the incidence of suicidal ideation, and potentially an increase in suicide attempts, after beginning treatment with an SSRI.[5] The most commonly cited theory as to why SSRIs, in particular, might increase suicidality, is that they tend to give new patients energy before altering their mood, potentially aiding a suicidal thought/attempt (Ludwig et al., 2009). The FDA's first public recognition of this association came on June 19, 2003, when it released a statement saying Paroxetine (marketed as Paxil), an SSRI, should not be used to treat major depressive disorder for children under 18. After a series of public health advisories,[6] the FDA mandated on October 15, 2004 that a black box warning be added to all antidepressants describing increased risks of suicidality in children and adolescents. The specific wording of the warning can be seen in Appendix Figure A1.

Shortly after the FDA's 2004 decision, the agency's Division of Psychiatry Products was asked to expand their exploration of suicidality in antidepressant trials to the adult population. This was a major effort, involving 372 placebo-controlled antidepressant trials and almost 100,000 patients (Noel, 2015). On December 13, 2006 the FDA's Psychopharmacologic Drugs Advisory (PDA) Committee met to consider the results of these trials. The meeting convened with a vote of six-to-

[4] Seratonin is a neurotransmitter, a chemical that carries signals between brain cells. SSRIs block the brain's reabsorption of seratonin, making more available.

[5] This body of evidence was systematically reviewed when the FDA sponsored a meta-anlaysis of over 20 pediatric randomized control trials and found among those 18 or younger, SSRI doubled the risk of suicidality versus receiving a placebo (4 percent versus 2 percent). None of the suicide attempts documented in the trials were fatal (Hammad et al., 2006). Antidepressants were first linked to an increase in the risk of suicide in the 1950s, when TCAs were introduced (Ludwig et al., 2009).

[6] On October 27, 2003, the FDA announced it would convene an advisory committee to examine risk of suicidality associated with antidepressant use in pediatric and adolescent patients. This announcement was accompanied by a public health advisory, distributed via MedWatch, to all US physicians recommending close supervision of all high-risk patients. On March 22, 2004, the FDA issued a second public health advisory urging clinicians to "carefully monitor patients on antidepressants for possible worsening of depression or suicidality." Moreover, the FDA called on manufacturers of 10 antidepressants to add to their label a warning that all patients should be carefully monitored for suicidality.

two in favor of altering the black box warning in three ways: (i) expand the warning that the drug increases the risk of suicidality to include young adults *under the age of 24*; (ii) notify individuals 65 and older of a *reduced* risk of suicide while taking the drug; and (iii) state that patients of *all ages* who are started on antidepressant therapy should be monitored appropriately and observed closely for clinical worsening, suicidality, or unusual changes in behavior. The new warning was announced to the public on May 2, 2007. The exact wording of the expanded warning can be seen in Appendix Figure A2. Both the FDA's PDA committee meeting in December 2006 and the black box announcement in May 2007 generated significant media coverage.[7] Figure 2 contains a timeline of FDA warnings on the association between antidepressants and sucidality.

3 Empirical Approach

3.1 Data

The data we use for our main analyses come from the National Survey on Drug Use and Health (NSDUH). The NSDUH is a nationwide survey sponsored by the Substance Abuse and Mental Health Services Administration (SAMHSA) that provides information on demographics, substance use, mental health, other-health related issues, and labor market status. Every year the NSDUH surveys about 70,000 individuals age 12 and older.[8]

In 2004, the NSDUH began classifying individuals as having had a major depressive episode (MDE) in their lifetime if for a period of two weeks or longer, they (i) experienced at least 5 out of 9 symptoms commonly associated with depression and (ii) one of the symptoms was a depressed mood or loss of interest or pleasure in daily activities.[9] The other symptoms reflect changes in functioning, such as problems with sleep, eating, energy, concentration and decisiveness, self-image and worth, or recurrent thoughts of death or suicide. We define our treatment group as those who have ever had an MDE in their lifetime and our control group as those who have never experienced an MDE

[7] A LexisNexis search of "antidepressant warning (warning*s*)" for May 2007 reveals 807 (933) news articles, which is more than twice as many articles found in April or June of 2007. A search of those same terms for December 2006 reveals 672 (808) articles, also more than twice the articles found in November 2006 or January 2007.

[8] Our identification strategy requires that individuals be categorized as depressed (or not) and employed (or not) consistently during our time period of interest. NSDUH and NHIS (∼90,000 annual participants), the latter of which we analyze in Section 4.3, are the largest US datasets that meet these minimal criteria. Alternatives, such as NHANES (∼5,000 annual participants), NLSY97 (∼7,500 annual participants), Add Health (∼16,000 annual participants), PSID (∼24,000 annual participants), and MEPS (∼30,000 annual participants) contain employment and mental health information, but offer substantially smaller samples, especially when we consider heterogeneity by coarse age groups. Furthermore, survey waves during our period of interest in Add Health only occur in 2001-2002 and 2008 and provide us a narrow age range to study (24-32 year olds in 2008). Likewise, NLSY97 provides only a small age range to consider as individuals were aged 22-28 in the 2007 wave. Both the PSID and NLSY79 were conducted bienially during our period of interest, making them unattractive for our empirical design. BRFSS (∼430,000 annual participants) contains a mental health/stigma module as well as an anxiety and depression module, but these modules are not asked consistently each year of the survey and different subsets of states ask them in different years.

[9] The NSDUH's definition of MDE is based mainly on the 4th edition of the Diagnostic and Statistical Manual of Mental Disorders (DSM-IV). Unlike the definition in the DSM-IV, no exclusions were made for an MDE caused by medical illness, bereavement, or substance use disorders in NSDUH.

in their lifetime.[10] Note that assignment to the treatment and control groups using lifetime MDE is *not* based on a medical diagnosis or an individual's decision to seek treatment for depression. Assignment based on diagnosis and/or mental health treatment receipt could be problematic if the warning altered individual healthcare behaviors, potentially changing the composition of the treatment and control groups in a way that biases the difference-in-differences estimates.[11] We revisit this point in Section 4.2.

Our main outcome of interest is an indicator for whether the individual was employed in the past week, which includes those working full-time, those working part-time, and those who report having a job or a business but did not work in the past week. All others are defined as not employed, which includes the unemployed, the disabled, homemakers, those in school or training, the retired, and those without a job for some other reason.

The NSDUH contains rich information on demographics, such as the individual's gender, race, ethnicity, marital status, educational attainment, and whether he/she lives in a large metro, small metro, or non-metro. For those aged 12-21, age is known precisely in years. In the public-use data, an individual's precise age in years is not available if he/she is over the age of 21. Instead, the NSDUH bins the following ages together: 22-23, 24-25, 26-29, 30-34, 35-49, 50-64, 65 and older. Our sample consists of those aged 18-49; individuals in their prime-working years. We also estimate our models on subsamples of coarse age groups, including those aged 18-25, 26-34, and 35-49. We consider these groups because 18-25 year olds were directly targeted by the 2007 warning and 26-34 year olds were just outside the targeted age range. Studying those aged 35-49 allows us to examine whether the warning affected individuals well outside the targeted group. Moreover, the warning also applied to "patients of all ages who are started on antidepressant therapy," providing another reason to examine the response among older individuals. We have chosen not to study individuals over the age of 50 to avoid distinguishing between unemployment and retirement. By excluding those over 50, we also avoid considering how conditions like Alzheimer's and menopause interact with depression, which would lead to a more varied and complicated understanding of depression and employment across age groups.

We use data from 2005 to 2008 inclusive. Our sample starts in 2005, which was the first full year the MDE-related questions were asked to all adult respondents. Beginning in 2005 is also advantageous because it allows us to avoid capturing responses to the 2004 black box warning and FDA advisories that preceded the initial warning.[12] We do not consider years beyond 2008 to avoid

[10] In referencing the treatment group, we will often describe individuals as having experienced "a lifetime MDE."

[11] In comparison, the treatment group in Busch et al. (2014) consists of those with probable depression, defined as receiving any treatment or counseling services for emotional or behavioral problems from a range of sources over the past 12 months. The authors conduct, and pass, a number of specification tests to defend their identifying assumption that the composition of their treatment and control groups is stable; however, none of these tests can guarantee that unobserved characteristics of the groups does not change. They use the "probable depression" definition because the validated diagnostic instrument for MDE that we use was not asked until 2004.

[12] To address concerns that the 2004 warning impacted antidepressant use trends, and subsequently employment, in 2005, we conduct robustness tests in Section 4.3 that drop the first and second half of 2005 from the sample.

the worst of the Great Recession and its potential effects on employment and healthcare. NSDUH interviews are conducted quarterly. To improve statistical power, we consider 6-month periods (i.e., in a given calendar year, quarters 1 and 2 make up one period and quarters 3 and 4 make up another period).[13] We consider all eight 6-month periods from 2005 to 2008. Our sample consists of 59,048 male observations and 67,468 female observations.

In Tables 1 and 2 we present summary statistics for men and women, respectively, by whether or not they experienced a lifetime MDE and by coarse age groups from 2005-2008. We also show the pre- and post-warning employment means for each group. Women are almost two times more likely to have experienced a lifetime MDE than men. Men of all ages with a lifetime MDE are less likely to be employed. For women younger than 35, employment rates are similar for those with and without a lifetime MDE. However, women aged 35-49 with a lifetime MDE are less likely to be employed than those without. On average, both men and women with a lifetime MDE are more highly educated, are less likely to be married, and are more likely to be white than those who have not experienced an MDE in their lifetime. We control for these characteristics in our regressions.

3.2 Empirical Strategy

We estimate the following difference-in-differences model using OLS:

$$Emp_{it} = \alpha_0 + \alpha_1 MDE_{it} + \alpha_2 MDE_{it} \times Post_t + \alpha_3 X_{it} + \tau_t + \varepsilon_{it} \tag{1}$$

where Emp_{it} is an indicator for whether person i is employed at time t (where time is measured in 6-month intervals). MDE_{it} is an indicator for whether person i has had an MDE in their lifetime. $Post_t$ is an indicator equal to 1 starting in 2007. We define the first half of 2007 as "treated" because the FDA advisory committee agreed that antidepressant labeling changes were needed during their meeting in mid-December 2006 and the expanded black box warning was announced by the FDA on May 2, 2007. Given the media coverage of the December meeting (see footnote 7), healthcare providers and patients may have responded to that meeting and the proposed labeling changes before the black box warning was formally announced. If providers and patients responded only to the May 2007 announcement, then including the first half of 2007 in the post-warning period will cause our estimates to be conservative.[14] X_{it} are individual controls, including age, education, race, marital status, and metro status, and τ_t are time fixed effects. In all our models, we use NSDUH sampling weights, which are representative of the US population and account for survey nonresponse, and we report heteroskedastic-robust standard errors.

The coefficient of interest is α_2, which represents the change in the probability of employment of

[13] We have also estimated our models with time measured in quarters (rather than 6-month periods) and results are qualitatively similar.

[14] We examine the sensitivity of our results to the definition of $Post_t$ as well as the exclusion of the first half of 2007 in Section 4.3.

individuals with a lifetime MDE following the expanded black box warning in 2007, compared to the change in the probability of employment of individuals without a lifetime MDE. Our identification strategy, like that of Busch et al. (2014), relies on two key assumptions: (i) the composition of the treatment and control groups did not change in response to the warning and (ii) employment trends for the treatment and control groups would have been the same in the absence of the warning (i.e., the parallel trends assumption). We provide evidence supporting these assumptions in Section 4.2.

4 Results

4.1 Main Findings

The difference-in-difference estimates are displayed in Table 3. We present results separately by gender, for the full sample (i.e., those aged 18-49) and for our three coarse age groups (18-25, 26-34, 35-49 year olds). We find no significant effect of the 2007 black box warning on the employment of men who have experienced a lifetime MDE. Women aged 35-49 with a lifetime MDE experienced a 4.4 percentage point decline in employment in response to the expanded warning, a 6.1 percent decline relative to the pre-warning mean for this group of women. There was no significant impact of the warning on the employment of women under the age of 35.

We then reestimate Equation 1 allowing the effects of the warning to vary over time. That is, we replace $MDE_{it} \times Post_t$ with separate interactions between MDE_{it} and indicators for each post-warning period. Results are shown in Table 4. Again, we find that the warning had a significant effect on the employment of 35-49 year old women, but no other subgroup. The negative employment effect for this group in the baseline difference-in-differences model is driven by an 8 percentage point decline in employment in the second half of 2007, with no significant decline in the first half of 2007. The point estimates in the first and second halves of 2008 are negative and smaller in magnitude (a 3-4 percentage point decline), but are not precisely estimated.

These results suggest that the 2007 expanded black box warning did not affect the employment of the targeted age group of 18-24 year olds. Instead, the warning led to a decline in employment among women aged 35-49, with the effect concentrated shortly after the warning was announced. While these results may seem surprising, a response by adults, particularly female adults, can be rationalized. First, while the 2007 warning explicitly targeted 18-24 year olds, language was added that applied to all new antidepressant users. Second, Tables 1 and 2 show that females experience higher rates of depression than men, and Blanchflower and Oswald (2016) as well as the CDC show that among the subsamples we study, antidepressant use is highest for women aged 35-49. Thus, it could be that statistical power is larger in our 35-49 year old female sample as there are more individuals who consider antidepressant therapy. Third, several studies from the psychiatry literature argue that women may be more responsive to antidepressants than men (Kornstein et al., 2000; Young et al., 2009; Sramek et al., 2016). Fourth, Olfson et al. (2008), Libby et al. (2009),

and Parkinson et al. (2014) show that adults responded to the 2004 black box warning by reducing their antidepressant use, despite the warning having no stated relevance for adults. Last, Busch et al. (2014) only find significant indirect (adverse) effects of the 2004 warning for adolescent girls.

4.2 Support for Identifying Assumptions

We next provide support for the two main identifying assumptions underlying the difference-in-differences model. First, we address concerns related to changes in the composition of our treatment group. As previously stated, because the individuals comprising our treatment group have experienced an MDE in their lifetime, we think these concerns are muted relative to strategies that define the treated group based on diagnosis or the receipt of mental healthcare. Nevertheless, one might worry about individuals who never experienced a lifetime MDE prior to the warning and develop depressive symptoms after the warning. If in the absence of the warning some of these individuals would seek out treatment (that prevents their symptoms from intensifying), but do not do so in the presence of the warning and subsequently experience a major depressive episode, then this may introduce a different set of depressed individuals into the treatment group.[15] To address these concerns, in Figures 3 and 4, we plot the share of men and women, respectively, who ever experienced MDE in their lifetime throughout our sample period. There are no abrupt or sharp changes in the proportion of individuals (overall or by age group) with a lifetime MDE after the warning.[16]

In the spirit of Busch et al. (2014), we also estimate versions of Equation 1, where we use each covariate in X_{it} (i.e., age, education, race, marital status, and metro status) as the outcome variable and test whether the coefficient on $MDE_{it} \times Post_t$ is significant. While all tested characteristics are included as controls in the difference-in-differences model (where employment is the outcome of interest), evidence of significant changes in these characteristics would raise concerns regarding *unobserved* compositional changes, which would threaten the validity of our identification strategy. For men aged 18-25, 4 out of 17 coefficients were significant at the 5 percent level and for men aged 26 to 34, 2 out of 17 coefficients were significant at the 5 percent level. There were no coefficients significant at $p = 0.05$ for women in the three age groups we consider or for men aged 35-49. Thus, out of 102 coefficients, only 5.9 percent were significant at the 5 percent level. Half of the

[15] Similarly, one might worry about those who never experienced an MDE prior to the warning because they took antidepressants to keep mild or moderate depression under control. The warning could lead some of these individuals to stop taking antidepressants, triggering an MDE.

[16] It has been suggested that the group of individuals who have experienced MDE in their lifetime may be largely comprised of those with a history of depression, meaning this particular depression measure may not change much over time and may mask small changes in the proportion of individuals who recently experienced an MDE. In Appendix Figures A3 and A4, we present plots of the share of individuals who experienced MDE in the past year. We find these proportions are quite stable. We prefer using MDE in one's lifetime rather than in the past year to define the treatment group because it allows us to include more individuals who may be affected by the warning (i.e., individuals who may use or consider using antidepressant therapy), such as those with a history of depression (not just recent depression) as well as those who currently suffer from mild or moderate depression that was more severe in the past.

significant male coefficients relate to small changes in racial or ethnic groups that make up less than 1 percent of the sample (e.g., Hawaiian/Pacific Islander), which we suspect are unlikely to impact our estimates.[17] Thus, we interpret the results of this exercise as evidence that compositional changes in the treatment group are not a large concern, particularly for women aged 35-49, the group for whom we find declines in employment due to the warning.

Our second identifying assumption is that pre-warning employment trends for ever- and never-depressed groups are parallel. We test this assumption using an event-study design. Specifically, we estimate a generalized difference-in-differences model of the form:

$$Emp_{it} = \alpha_0 + \alpha_1 MDE_{it} + \sum_{\ell=2005Q1/2}^{2008Q3/4} \pi_\ell \mathbb{1}_{(t=\ell)} \times MDE_{it} + \alpha_3 X_{it} + \tau_t + \varepsilon_{it}. \quad (2)$$

The indicator for the second half of 2006 is omitted; thus, the π_ℓ coefficients describe the evolution of employment before and after the black box warning relative to 2006Q3/Q4. This specification provides both a visual and statistical summary of the differences in pre-warning employment trends for the treatment and control groups across subsamples. Figures 5 and 6 show plots of the π_ℓ coefficients from Equation 2 along with 95 percent confidence intervals for men and women, respectively. The figures show no differences in the pre-warning employment trends of the treatment and control groups. In fact, relative to the base period, we find no statistically significant employment differences between treatment and control groups in any of the pre-warning periods for any of the subsamples analyzed.

4.3 Robustness Checks

We perform a number of robustness exercises. First, we estimate our model under alternative timing assumptions. Second, we address concerns related to the Great Recession. Third, we estimate our model using two alternative definitions of depression. Fourth, we provide evidence from the National Health Interview Survey (NHIS) that corroborates our findings from the NSDUH. For brevity, for most of these robustness checks, we show results only for women aged 35-49, the group for which we find significant employment effects, and we briefly discuss the results for the other subgroups. Parameter estimates for the other groups are available by request.

4.3.1 Alternative Timing Assumptions

In our baseline specification, we classify the first half of 2007 as part of the post-warning period because the FDA advisory meeting where labeling changes were discussed took place in December 2006 and the black box warning was announced in May 2007 (see Figure 2). If individuals did not respond to the FDA meeting but only to the black box announcement, then the first half

[17] Results from this exercise are available upon request.

of 2007 is only partially treated. We reestimate Equation 1 dropping the first half of 2007 from the sample. Results are presented in column 1 of Table 5. Relative to our baseline estimates, the decline in employment for women aged 35-49 increases in magnitude to 5.1 percentage points, which is consistent with the first half of 2007 being an intermediate, partially-treated period that makes our baseline estimates conservative. We then reestimate Equation 1 including the first half of 2007 but defining $Post_t$ to be equal to 1 starting in the second half of 2007, such that only 6-month periods that are fully treated after the black box warning announcement are considered to be in the post-warning period. Results are presented in column 2 of Table 5. We find that women aged 35-49 experienced a 4.6 percentage point decline in employment. For both these specification checks, we find no significant impact of the warning for men or for women under the age of 35. Thus, our results are not sensitive to alternative timing assumptions regarding the warning.

Our baseline specification includes the first and second halves of 2005 in the pre-warning period. Including these periods could be problematic if the 2004 warning continued to have an effect on antidepressant use trends in 2005. Specifically, the 2004 warning could cause a declining employment trend for the (lifetime MDE) treatment group that is not mimicked by the control group, leading to a violation of the parallel trends assumption and an overstatement of the true treatment effects. The event study results do not indicate that this is a concern; nevertheless, we reestimated our model, first excluding the first half of 2005 from the sample, then excluding the second half of 2005 as well. The results are presented in columns 3 and 4 of Table 5. When we drop the first half of 2005, we find a 5.9 percentage point decline in employment for women aged 35-49. When we drop all of 2005, we find a 6.9 percentage point decline in employment for these women. We find no significant effects for the other subgroups. These estimates are similar in magnitude to our main findings; moreover, our treatment effects are larger when 2005 is removed, which runs counter to the narrative above. Existing literature on the antidepressant response to the 2004 black box warning further supports including data from 2005 in our analysis, as most studies find that the primary antidepressant response took place after the various FDA hearings and advisories in late 2003 and early 2004 that eventually led to the black box announcement in October (Nemeroff et al., 2007; Olfson et al., 2008; Busch et al., 2014). In fact, in their examination of the indirect effects of the 2004 warning on adolescents, Busch et al. (2014) treat the first quarter of 2004 as the first treated period, despite the fact that the warning was not announced until October of that year.

4.3.2 The Great Recession

Another potential concern is that our sample period, particularly the post-warning period, overlaps with the Great Recession (which started in December 2007 and ended in June 2009 according to the National Bureau of Economic Research), which may impact our estimates. In particular, if the recession had a greater employment impact on depressed individuals than non-depressed individuals, then our estimates would confound the effects of the black box warning and the recession.

We think this is unlikely for two reasons. First, the time-varying effects of the warning suggest the employment response by women occurred in the second half of 2007 and eroded in magnitude after. If the Great Recession influenced our results, we would expect to find larger effects towards the end of the sample period when the recession worsened. Second, we find significant effects of the warning for women and no significant effects for men, while the literature suggests men experienced more employment declines and unemployment increases than women during the Great Recession (Hoynes et al., 2012).

Nevertheless, it is possible that those in poor health were impacted earlier and more significantly by the recession, perhaps because they were less likely to be hired and/or more likely to be laid off. If this is the case, we would also expect to find declines in employment for individuals with chronic health conditions, such as asthma, diabetes, or high blood pressure. Moreover, given that body weight is (i) correlated with diabetes and high blood pressure and (ii) more easily observable to employers than mental health problems, we might expect an even stronger employment effect for individuals suffering from those conditions. To test this idea, we reestimate Equation 1 but define the treatment group as individuals who have ever been diagnosed with asthma, diabetes, or high blood pressure. Results for all groups are presented in Appendix Tables A1 and A2. We find no significant effect of the warning (i.e., post-2007) on individuals with these conditions, giving us confidence that the employment declines we find for women are not driven by the recession impacting those with poor physical or mental health. Furthermore, we view this exercise as a more general placebo check that provides evidence that our estimated effect of the warning on women does not simply reflect some other change (perhaps in healthcare policy) that impacted individuals with poor health during this time.

4.3.3 Alternative Depression Definitions

As explained in Section 4.2, our preference is to define the treatment and control groups using lifetime MDE status because the warning is unlikely to alter the composition of these groups. That said, one concern is that the NSDUH definition of MDE is almost the exact diagnostic indicator for Major Depressive Disorder outlined by the DSM-IV (see footnote 9), meaning some individuals suffering from minor forms of depression could be in our control group. In this section, we reestimate Equation 1 using two alternative definitions of depression to categorize individuals into treatment and control groups. First, we define the treatment group as those who have ever been told by a doctor or medical professional they have depression or anxiety. While this definition allows for individuals with more minor forms of depression to be categorized as depressed, entry into the treatment group depends on a diagnosis (i.e., seeking out treatment); therefore, the warning is more likely to lead to compositional changes in the treatment and control groups than our preferred definition. Given that our objective is to expand the original treatment group to include individuals with minor depression, our second new definition categorizes as depressed anyone with a lifetime

MDE *or* who has ever been told by a doctor they have depression or anxiety.

The results using these alternative depression definitions can be found in columns 5 and 6 of Table 5. Using the first definition, we find that for women aged 35-49, the warning reduced employment by 3.9 percentage points, and we find a 3.6 percentage point decline for this group using the second definition. We again find no significant effects for males or for women under the age of 35.

4.3.4 National Health Interview Survey

We show in the above subsections that our main finding – that the FDA's 2007 expanded black box warning on antidepressants led to a significant reduction in employment for ever-depressed women aged 35-49 – is robust to several alternative specifications and assumptions. Moreover, we argued in Section 4.1 that there are numerous reasons why this result should be viewed as unsurprising. Despite this evidence, because the largest employment effects are found among a non-targeted group, there may be concerns that our results are a statistical anomaly or are in some way driven by the survey data used. In response to these concerns, in this section we repeat the above analysis using the National Health Interview Survey (NHIS), an annual household survey conducted by the National Center for Health Statistics. NHIS is a large cross-sectional survey that is representative of the non-institutionalized US population, and it contains information on a broad range of health topics, as well as demographics and labor market status.

Questions about experiencing a major depressive episode are not available in the NHIS for our time period of interest. Instead, we use the Kessler-6 (K6) scale to define our treatment and control groups.[18] The K6 is commonly used in the mental health literature. The scale is calculated from responses to six questions of the form: "During the past 30 days, how often did you feel . . . [nervous, hopeless, restless or fidgety, so depressed that nothing could cheer you up, that everything was an effort, worthless]?" For each question, a value of 0, 1, 2, 3, or 4 is assigned to the answers "none of the time," "a little of the time," "some of the time," "most of the time," or "all of the time," respectively. The K6 is calculated by summing the scores from each of the six questions, generating a 0-24 scale, with higher scores indicating a greater tendency towards mental illness. Scores greater than 13 are associated with severe mental illness, while scores greater than 4 and less than 13 are associated with moderate mental distress. A cutoff of 4 has been identified as optimal in identifying respondents with mental health treatment needs (Prochaska et al., 2012). Thus, our treatment

[18] The only other depression measure consistently recorded by NHIS requires that an individual reports an activity limitation. If an individual reports being unable to work, perform personal care activities, like bathing and dressing, or perform routine errands and chores, then the survey design allows them to attribute the limitation to depression (among several other conditions). Thus, for individuals who do not experience activity limitations, their depression status is unknown. During our period of interest (2005-2008), the NHIS included some specific questions about depression and anxiety, but did not do so consistently across years. For example, only in 2008 did the survey ask individuals whether they ever had depression or anxiety, and only in 2007 did it ask whether the respondent had been frequently depressed or anxious in the past year.

group consists of individuals with a K6 score greater than 4, which describes roughly 20 percent of women and 15 percent of men aged 18-49 in our sample. Because the K6 score reflects recent mental distress, it likely does not capture individuals who have suffered from depression in the past, but do not currently experience depressive symptoms, perhaps because they receive mental health treatment like antidepressants. If those individuals responded to the black box warning, our estimates will be biased toward zero since individuals in the control group responded to the treatment. This is also the reason the NHIS is not our preferred sample.

We estimate Equation 1 using the NHIS, replacing MDE_{it} with an indicator for whether an individual has a K6 score greater than 4. Our main outcome of interest is again an indicator for whether the individual was employed in the past week, which includes individuals who worked for pay in the past week, did not work in the past week but have a job, and those who worked but not for pay in the past week.[19] Like the NSDUH, the NHIS data contains the quarter the interview was conducted. Again, to increase statistical power, we group the data into 6-month periods. The NHIS also has detailed demographic information on race, ethnicity, education, marital status, and Census region, which we control for in the regressions. An advantage of the NHIS is that it contains precise age in years. We, therefore, control for age and age squared.[20] We again estimate our model on individuals 18-49 and the three coarse age groups (18-25, 26-34, 35-49 year olds). Descriptive statistics for our NHIS sample are presented in Appendix Tables A3 and A4.

The same identifying assumptions discussed in Section 4.2 must be met when using the NHIS. In Appendix Figures A5 and A6, we show the proportion of individuals with a K6 score greater than 4 over the sample period.[21] For men aged 26-34 and 35-49 as well as women aged 35-49, the proportion of individuals with a K6 greater than 4 is quite stable over the sample period. For other groups, there is about a 5 percentage point increase in depression towards the end of the sample period. We also present event study estimates in Figures A7 and A8. We generally do not find evidence of strong pre-trends. In sum, we view our identification strategy as imperfect for men aged 18-25 and women under 35 when using NHIS data; however, our main objective in turning to the NHIS is to determine whether the significant negative employment effects found for women aged 35-49 in the NSDUH are robust. For this subsample, the assumptions required for our identification strategy seem to be satisfied.

The difference-in-differences estimates using the NHIS are presented in Table 6. We find no significant impact of the warning on men or women. For women aged 35-49, there is a decline in employment after the warning, but it is not precisely estimated. Given the robust decline in employment for women aged 35-49 with a lifetime MDE in the NSDUH and the fact that age in years is available in the NHIS, we take advantage of knowing age precisely and explore whether there are

[19] Our results are not sensitive to including those who work but not for pay as they make up a tiny part of the sample.

[20] Our results using the NHIS are not at all sensitive to the inclusion of age and age squared, which alleviates concerns that not controlling for age in years when using the NSDUH affects our estimates.

[21] For brevity, we often refer to those with a K6 score greater than 4 as "depressed."

effects for women within the 35-49 age range. We experimented with numerous smaller age bands and found several significant declines in employment among those with a K6 score greater than 4. The widest age band where we find significant decreases is among 36-44 year old women, who experience a 5.5 percentage point (p=0.055) decline in employment, an 8.4 percent decline relative to the pre-warning mean for this group of women. Results are shown in column 1 of Table 7.[22]

Another advantage of the NHIS is it contains additional outcomes related to employment and health that could be impacted by the 2007 warning. In particular, respondents are asked to report the number of days during the past year that an illness or injury kept them in bed for more than half the day (henceforth "bed days"). It also contains information on whether individuals are currently unable to work due to a physical, mental, or emotional problem and whether individuals are limited in the kind or amount of work they can do because of such a health problem. We consider an indicator that takes on a value of 1 for those who are unable to work or are limited in the work they can do due to a health problem. Finally, the survey asks individuals if they have any "activity limitations" related to work, personal care (e.g., eating, bathing, dressing, etc.), and routine needs (e.g., household chores, shopping, etc.). If any limitations are identified, the respondent is asked to specify the health condition(s) causing the limitation(s), with one of the conditions being "depression/anxiety/emotional problem." We create an indicator that takes on a value of 1 for those who have a limitation due to depression, anxiety, or emotional problem, and 0 otherwise (including both those without any limitations and those with limitations that are not due to depression).

We examine whether the 2007 warning impacted these additional outcomes for women aged 36-44 using our difference-in-differences strategy, with results presented in columns 2-4 of Table 7.[23] We find no significant change in bed days. The probability of being unable to work or having a work limitation due to a health problem increases by 4.3 percentage points among those with depression, a 19 percent increase relative to the pre-warning mean for this group. We find a 4.7 percentage point increase in the probability of having an activity limitation due to depression (anxiety or emotional problem) among depressed women, a 66 percent increase relative to the pre-warning mean. Thus, this analysis corroborates the declines in employment we find among depressed women. For these same women, work limitations increased after the warning as did the probability of having activity

[22] The proportion of women aged 36-44 with a K6 score greater than 4 is stable across the sample period, and we do not find evidence of differential pre-warning trends among this subgroup. Within this 36-44 age range, effects on employment are largest and most precisely estimated when considering 38-44 year old women (7 percentage point decline, p=0.030). When we extend the 36-44 age band by 1 year on either side, we find qualitatively similar declines in employment (of about 4 percentage points), but with p-values between 0.10 and 0.20. We also experimented with different age bands for men and for younger women. None of that experimentation generated employment effects that were significantly different from zero.

[23] Estimates of the impact of the warning on these additional outcomes for our standard subgroups of men and women are available by request. We find increases in the probabilities of having a work limitation due to a health problem as well as an activity limitation due to depression for men aged 18-25 with depression. As mentioned above, in Figure A5, there is a noticeable up-tick in depression among this group of men, which may be driving the effects we find on these other outcomes. We find no significant effects for the other subgroups.

limitations due to depression, suggesting depression became a barrier to employment for some women after the expanded warning.[24]

5 Mechanisms

The above results suggest that employment fell for women aged 35-49 in response to the 2007 expanded black box warning on antidepressants. The most likely reason for this effect is that the warning led to a decline in antidepressant use, worsening mental health, leading to lower rates of employment. In this section, we first provide evidence that antidepressant use did in fact decline for women aged 35-49. We then explore a number of other potential mechanisms.

5.1 Impact of 2007 Expanded Warning on Antidepressant Use

We use the Medical Expenditure Panel Survey (MEPS) data and a difference-in-differences strategy similar to what was described above to study the impact of the 2007 expanded black box warning on antidepressant use.[25] The MEPS, which is collected by the Agency for Healthcare Research and Quality (AHRQ), has been administered continuously since 1996. The MEPS contains detailed health, illness, medical expenditure, health insurance, and demographic information for a nationally representative sample of households in the US. New participants are added annually, drawn randomly from the previous year's NHIS sample. Each cohort is interviewed five times over the two years that follow January 1st of the cohort year.

The MEPS data have several characteristics that make it well-suited for this analysis. First, survey participants report all prescription drugs that they take by name, allowing us to determine whether an individual is taking an antidepressant known to carry a black box warning.[26] Second,

[24] As with the NSDUH data, we conduct placebo tests on the 4 outcomes examined in Table 7, replacing the previous treatment group definition (i.e., K6 score greater than 4 in the past 30 days) with "ever diagnosed with asthma, diabetes, or high blood pressure." Using these alternative placebo treatment groups, we find no significant impact of the post-warning period on the 4 outcomes for 35-49 or 36-44 year old women. Again, these findings suggest that the estimated effect of the black box warning on women aged 35-49 does not reflect the effect of some alternative healthcare-related shock that affected individuals with poor health.

[25] The NSDUH includes some information that could be interpreted as antidepressant use. If a respondent reports (i) thoughts or plans of suicide in their lifetime, or (ii) experiencing at least 5 out of 9 MDE symptoms in the same 2-week period in their lifetime, then they are asked whether they are "currently taking prescription medications prescribed for mood." However, these data are less than ideal for our purposes, not just because of the MEPS advantages listed below, but because the vast majority of those categorized as not having a lifetime MDE are never asked about "prescription medications prescribed for mood." Within the context of our difference-in-differences setting, the non-depressed control group has missing antidepressant use by design. Moreover, this measure of prescription use does not capture well antidepressant use for those with a history of mild or moderate depression.

[26] The recorded name of the drug is important because roughly 38 percent of the prescription drug fills for depression and anxiety (according to reported ICD9-CM codes) were not affected by the 2007 expanded black box warning on antidepressants. The most common class of alternatives is benzodiazepines (e.g., Xanax, Valium, etc.). In addition to the name of the drug, MEPS participants report the pharmacy the drugs were obtained from. If the survey participant provides written permission for the pharmacy to release their records, the pharmacy is contacted by AHRQ. Participating pharmacies provide the following information on each reported prescription: date filled, national drug code, medication name, strength of medicine (amount and unit), quantity, total charge, and payments by source.

depression and anxiety, which can be reported in each interview period as diagnosed or inferred from interview responses, are coded using ICD9-CM codes.[27] Measuring depression and anxiety via ICD9-CM codes has advantages over the MDE measure from the NSDUH and the K6 measure from the NHIS, as it is easier to interpret, includes both those with major and minor depressive disorders, and is less likely (than the K6) to be influenced by the use of antidepressants. Third, the panel nature of the data allows us to examine prescription drug behavior for the same individuals, before and after the warning. As such, with the MEPS, we can ensure that the ever-depressed and never-depressed groups are fixed over time.

We use the 2006 MEPS cohort, which is comprised of individuals who complete their first interview in the first half of 2006 and their last interview in December of 2007, so they are observed before and after the warning. Summary statistics for this cohort by sex and age group (18-25, 26-34, 35-49, and 36-44 year olds) are presented in Tables 8 and 9. Individuals are categorized as depressed if over the two years that they are interviewed, they ever report an illness with any of the following ICD9-CM codes: 296, Episodic mood disorders (affective psychoses); 300, Anxiety, dissociative and somatoform disorders (neurotic disorders); 309, Adjustment disorder with depressed mood; or 311, Depressive disorder (not classified elsewhere). The relationship between depression and observables is similar to that seen in the NSDUH and NHIS. The relationship between depression and antidepressant use has several notable features. First, never-depressed individuals are very unlikely to take antidepressants.[28] Second, for most individuals, the 2007 black box warning occurs early in the fourth interview round, so rounds 4 and 5 represent the post-warning period.[29] For 35-49 year old men and women, where our sample size is largest, there is a clear positive trend in antidepressant use in the first three rounds, followed by what appears to be a break in that trend. Because this pattern is not mimicked by the non-depressed group, it will be important to account for this trend in our empirical specification.

For non-participating individuals and pharmacies, these data fields are imputed by AHRQ. Unfortunately, the MEPS data files do not distinguish the reason for imputation (non-participating individuals versus pharmacies); however, imputed observations are flagged. In the 2006 (2007) data files, the pharmacy details of roughly 52 (49) percent of the reported prescription fills are imputed.

[27] In addition to reporting medical care consumption in each interview, individuals are asked to report all "health problems (experienced during the current interview period) including physical conditions, accidents, or injuries that affect any part of the body as well as mental or emotional health conditions, such as feeling sad, blue, or anxious about something." Participants are told explicitly to include ailments even if they did not seek professional medical care. An individual's description of the illness is recorded as verbatim text, which is later coded to 5-digit ICD9-CM codes by professional coders.

[28] As mentioned above, antidepressants are used in the treatment of illnesses besides depression and anxiety, including pain, insomnia, and migraines (Wong et al., 2017). While a non-trivial fraction of antidepressant users may not suffer from depression or anxiety, the table suggests they represent a very small fraction of the total non-depressed population.

[29] May 1, 2007 occurs in the fourth interview period for 87 percent of the sample; it occurs in the third interview period for the remainder.

Our empirical model is as follows:

$$AD_{it} = \alpha_0 + \alpha_1 Dep_i + \alpha_2 Post_t + \alpha_3 Dep_i \times Post_t + \alpha_4 t + \alpha_5 Dep_i \times t + \alpha_6 Post_t \times t + \alpha_7 X_{it} + \varepsilon_{it} \quad (3)$$

where AD_{it} is an indicator for whether individual i uses any antidepressants in interview round t and $Post_t$ is an indicator for whether round t ended after May 2007. Dep_i is an indicator for whether individual i ever reports an ICD9 code associated with depression or anxiety; X_{it} are individual controls (education, race, marital status, etc.); and ε_{it} is the econometric error. By including a linear time trend, as well as interactions, we allow for different trends (i) before and after the warning and (ii) for depressed and non-depressed individuals. This structure accounts for the difference in pre-trends in antidepressant utilization between depressed and non-depressed individuals seen in Tables 8 and 9.[30]

Regarding the definition of $Post_t$, the MEPS data records the round in which prescriptions are filled, but not the exact fill date or whether/when drugs are taken. As such, each individual has one *transition* round that contains May 1, 2007 and we do not know whether prescriptions filled in this round occurred before or after the warning. Our baseline specification puts the transition round in the post-warning period, which is the most conservative approach, as some (re)fills early in this period occur before May 1, making it more difficult to find a significant decline in antidepressant drug use. We later show that our results are robust to dropping the transition round entirely.

We estimate the model via OLS, both with and without group-specific linear trends. All reported standard errors are robust to possible violations of the homoskedasticity assumption and are clustered at the individual level. MEPS longitudinal sample weights are used. Results without differential trends are presented in Table 10.[31] None of the six subgroups appear to respond to the 2007 expanded warning with a decrease in their antidepressant use. In fact, antidepressant use *increases* among women aged 26-34. This finding is unsurprising, given the positive trend in antidepressant use leading up to the warning observed in Tables 8 and 9.

Our preferred results, which allow for differential linear trends as in Equation 3, are presented in Table 11.[32] Consistent with the employment effects reported in Section 4, we find that an-

[30] By controlling for the antidepressant use of non-depressed individuals, we also capture other changes that could have occurred in the overall drug market *at the same time* as the warning. In other words, our specification measures a break in the antidepressant use trend among depressed individuals, accounting for other market-level changes that may have occurred at the same time as the warning.

[31] This specification allows the interview round to enter linearly via a trend, t. We have estimated an alternative, slightly more flexible model that replaces both the linear time trend, t, and post-warning trend, $Post \times t$, with interview round fixed effects, which yields almost identical results.

[32] Using NSDUH, we reestimated Equation 1 without time fixed effects, instead including t, $MDE_{it} \times t$, and $Post_t \times t$ to be more consistent with our MEPS specification. Busch et al. (2014) use a similar specification that allows for differential linear trends. We find employment among women aged 35-49 with a lifetime MDE declines by 8.9 percentage points (p=0.044) following the warning and no significant effects for any other group. Given the longer sample period when we use the NSDUH (or NHIS), we prefer the specification with time fixed effects because it allows

tidepressant use among women aged 35-49 decreased by 8 percentage points (a 20 percent decline relative to the pre-warning mean for this group) and that this decrease is significantly different from zero at the 5 percent level. When we further restrict this group of women to those aged 36-44, the group for whom significant employment effects were discovered in the NHIS, we find an even larger 14.3 percentage point decrease in antidepressant use (a 37 percent decline relative to the pre-warning mean for this group). The fact that α_5 is positive and significant for both groups of women highlights the importance of the differential linear trend. This parameter captures the fact that antidepressant use increases among these groups prior to the 2007 expanded warning, allowing α_3 to be identified by a break in this trend. Also consistent with our employment effects, the 2007 expanded warning does not significantly impact the antidepressant use of women younger than 35, men aged 18-25, or men aged 35-49. We do find (weakly) significant *positive* effects for 26-34 year old men; however, this subsample contains only 66 depressed men and fails every robustness test examined below, including the inclusion of individual-specific fixed effects, removal of the transition round, and use of an alternative definition of depression.[33]

These findings are consistent with previously estimated effects of the 2004 black box warning on adult antidepressant use. Libby et al. (2009) found that SSRI use within 30 days of a new depression diagnosis fell 15 percent for adults aged 25-89 after the 2004 warning was introduced. Using MEPS data, Parkinson et al. (2014) found that new antidepressant use fell 16 percent for adults aged 25-64 after the 2004 warning. We perform a similar analysis and categorize observed antidepressant use as new or continuing.[34] In the first column of Table 12, we show estimates from Equation 3 where (i) AD_{it} equals 1 if individual i fills a new antidepressant prescription in round t and zero otherwise and (ii) the sample includes all individuals aged 25-64. We find the probability an ever-depressed individual is a new antidepressant user declines by 3 percentage points (a 35 percent decline) after the 2007 warning. That the 2007 warning yields a larger response among adults than the 2004 warning is expected, given that the 2007 warning (i) targeted older antidepressant users and (ii) added language that applied to new antidepressant users of all ages.

Though neither Libby et al. (2009) nor Parkinson et al. (2014) show how the 2004 adult new user effect differs for men and women, we show in columns 2 and 3 of Table 12 that the decline in new antidepressant use in 2007 is primarily driven by women. In columns 5 and 6, we examine the impact of the 2007 warning on new and continuing antidepressant use for 35-49 year old ever-depressed women. In this subsample, the probability of being a new antidepressant user fell by 5.9

for more flexible effects of time on employment, which may be especially important during the Great Recession.

[33]The data are not well-suited for allowing time-varying (dynamic) treatment effects. Technically, three post-warning period effects can be estimated; however, the first post-warning period for every individual is the transition round and the third post-warning period is only identified by individuals for whom May 1 occurs in the third interview round, which is only 13 percent of the sample.

[34]For each drug taken during a MEPS survey period, the data contain the date that a particular drug (e.g., Prozac) was first prescribed to the individual, even if the first prescription occurred prior to the first MEPS interview. We cannot, however, observe when an individual was first prescribed a particular *class* of drugs (e.g., SSRIs) because first prescription use is not recorded for drugs that are never taken during the MEPS survey.

percentage points (a 56 percent decline), while the probability of being a continuing antidepressant user fell by just 2.1 percentage points (a 7 percent decline), though the latter effect is not precisely estimated. This result is compelling for two reasons. First, among 35-49 year old women, the 2007 warning provided relevant information only for new users; thus, it is sensible that the overall decline in antidepressant use among 35-49 year olds is driven by new users. Second, though not shown, we find no significant impact of the 2007 warning on new antidepressant use for men or for women under 35; thus, the 35 percent decline in new antidepressant use among ever-depressed adults reported above is driven almost entirely by women aged 35-49.

Our main findings are robust to a variety of alternative model specifications and assumptions. We present the results of these robustness checks for women aged 35-49 and 36-44 in Table 13. Column 1 contains estimates of Equation 3 where the transition round is dropped for each individual. Recall, the 2007 expanded black box warning was announced during the transition round, so that round contains both pre- and post-warning antidepressant use, likely leading our estimates to be biased towards zero. As expected, the estimated effect of the warning grows for women aged 35-49 and 36-44 when the transition round is dropped. The true impact of the warning on these women likely lies somewhere between our baseline estimates and these. Column 2 contains estimates of Equation 3 where individual-level fixed effects are included. Individual-level fixed effects purge the model of individual time-invariant unobservables, which may be correlated with observables of interest, causing bias. In our setting, the key observable of interest is $Dep_i \times Post_t$, which cannot be correlated with permanent unobservables, as (i) its permanent component, Dep_i, is controlled for explicitly and (ii) all included individuals are observed before and after the warning. As such, individual fixed effects can only impact estimation of the treatment effect through their correlation with other covariates. In light of this, it is unsurprising that point estimates for both subsamples are similar to our baseline results; that said, standard errors in the fixed effects specification are notably larger due to the substantial degrees of freedom loss. Finally, column 3 contains estimates where individuals are categorized as depressed if they ever register a K6 score greater than 4 during the survey. Again, these estimates are quantitatively similar to our preferred estimates.[35]

One last concern we address is our reliance on differential linear trends for identification in our main specification. This assumption may be undesirable for two reasons. First, if the growth in antidepressant use over time is truly quadratic, for example, as opposed to linear, then the negative results we estimate for 35-49 and 36-44 year old women could be driven by model misspecification. Second, with time, survey participants may learn that their interview will last longer if they report medical care consumption. As such, the decrease in antidepressant use that we observe post-warning

[35] Results for other subsamples are available upon request. Across all alternative specifications, we find no significant effects of the warning on antidepressant use for men or for women aged 26-34. Note that in the main specification, we find that the warning has a *positive*, though insignificant, effect on antidepressant use for women aged 18-25, which is unexpected. This positive insignificant effect persists when the transition round is dropped; however, the effect becomes significant when fixed effects are added (p=0.099) and when the alternative depression measure is used (p=0.051). Note that this group contains just 64 depressed women.

could simply result from under-reporting by survey participants in later rounds of the survey. Both of these concerns suggest that a negative treatment effect would be observed independent of the black box warning.

To address these concerns, we conduct a series of placebo tests, where we assume that a policy (or fake warning) occurred on May 1, every year between 2003 and 2009. For each placebo warning occurring in year y, the MEPS cohort from year $y-1$ is used, just as in our main analysis. If the claims discussed above are valid, we should consistently find a negative "May 1" effect across years. Results from this placebo analysis for women older than 35 can be seen in Table 14. First, note that independent of statistical significance, the results are not overwhelmingly negative – excluding 2007, only a quarter of the effects are negative. Second, significant treatment effects are not estimated in any of the placebo years for women over the age of 35. Though not shown, results for the other subsamples are similar. Of the 60 placebo treatment effects estimated (6 years and 10 total subsamples), 32 are negative and just 4 are significantly different from zero at the 10 percent level. Two of these significant effects are from 2003; likely a result of the 2004 black box warning and the preceding advisories. With a purely random treatment, one would expect 30 negative parameters and 6 significant effects.

5.2 Alternative Mechanisms

The most obvious channel underlying the employment decline generated by the FDA's 2007 black box warning is a reduction in extensive margin antidepressant use. That said, other individual responses could explain the observed drop in employment. For example, there may have been changes in the treatment patterns of those who continued to take antidepressants before and after the warning. To explore this, we estimate the following model using MEPS participants that (i) report depression or anxiety at some point during the survey and (ii) take antidepressants in at least one pre-warning round *as well as* one post-warning round:

$$ln(AD_{it}) = \alpha_0 + \alpha_1 Post_t + \alpha_2 t + \alpha_3 X_{it} + \varepsilon_{it}.$$

In our first specification, AD_{it} measures the total number of antidepressant pills taken in an interview round, regardless of strength. In our second specification, AD_{it} measures the average milligrams (MG) per pill taken in an interview round. We take the log of both measures to reduce the impact that right tail outliers have on our estimates. Other variables are defined as before.

In Table 15 we present our estimate of α_1 across samples and outcomes, using models that both exclude and include individual fixed effects. A full set of parameter results is available upon request. Given the control for time, t, α_1 captures any break from trend that occurs at the time of the warning. For completeness, we present results for all age-sex samples previously considered; however, 3 subsamples contain fewer than 15 people that meet the sample criteria, so little weight

should be placed on those results.

Our most notable findings are for women aged 35-49, our largest subsample and the only subsample (along with women aged 36-44) for which we find both employment and extensive margin antidepressant effects. The estimates in columns 1 and 2 of Panel B in Table 15 show that these women (who took antidepressants both before and after the warning) did not respond to the warning by reducing the average number of pills they took in a round, suggesting that neither the number of pills per refill nor the number of refills was reduced in response to the warning. However, estimates in columns 3 and 4 suggest a significant decline in the strength of the average pill taken – post-warning doses per pill (measured in MG) are 65-70 percent of pre-warning doses.[36] In addition to the extensive margin decline in antidepressant use for women aged 35-49, this intensive margin decline could further explain the drop in employment detailed in Section 4.1. Further, while the extensive margin response documented in the previous section could be driven by patients and/or healthcare providers, the intensive margin response requires provider involvement, even if the provider response was caused by patient requests.

We examine additional mechanisms that may be especially relevant for those who changed their antidepressant use. We first look at possible medical substitutes. Specifically, using the MEPS data, we reestimate Equation 3, using two new dependent variables: (i) an indicator for whether the individual went to psychotherapy during an interview round, and (ii) an indicator for whether the individual took non-black box drugs for depression or anxiety during an interview round. Both the 2004 and 2007 warnings applied to all antidepressant drugs on the market; however, the warning did not apply to several classes of drugs more targeted toward the treatment of anxiety, such as Alprazolam (marketed as Xanax) and Diazepam (marketed as Valium). These drugs, which are both benzodiazepines, are sedatives and could serve as a substitute for antidepressants in certain cases.[37] Finding that the warning caused a shift from antidepressants to benzodiazepines could potentially explain some of the observed drop in employment for women, as benzodiazepines reduce energy levels and are commonly abused.[38]

In Appendix Table A5, we present estimates of α_3 (i.e., the difference-in-differences interaction term) for women aged 35-49 and 36-44. A full set of estimation results, for these subsamples and others, is available upon request. The results in columns 1 and 2 show a (weakly) significant, negative effect of the 2007 warning on therapy use among women aged 35-49, implying therapy and antidepressants are not substitutes, but rather complements, for this subgroup. The estimates in columns 3 and 4 suggest that the warning did not lead to a significant change in the use of

[36] These results are robust to dropping the transition round. Results are available upon request.

[37] In the MEPS data for the 2006 cohort, 62 percent of the prescription drugs used to treat depression and anxiety are antidepressants carrying a black box warning, roughly 10 percent are benzodiazepines, and 28 percent are other alternatives.

[38] According to the SAMHSA, benzodiazepine abuse treatment admissions tripled from 1998 to 2008. In August 2016, the FDA added a black box warning to all benzodiazepines warning that the drugs can be fatal if taken with opioids, another commonly abused drug class.

non-black box drugs for women over 35.[39] Thus, while the warning had effects on a non-targeted population (i.e., women aged 35-49), it did not affect non-targeted prescription drug classes also used for treating depression and anxiety. Such spillover effects on non-targeted drug classes were found in the US in response to the removal of Vioxx, a type of Cox-2 inhibitor (painkiller), from the pharmaceutical market due to adverse cardiac side effects (Collins et al., 2013).

In addition to these medical substitutes, we use the NSDUH to examine whether individuals responded to the 2007 warning by using non-medical antidepressant substitutes, including alcohol and marijuana. In particular, we consider as outcomes indicators for: (i) any marijuana use in the past month; (ii) using marijuana 20 or more days in the past month; (iii) using alcohol (in any amount) 20 or more days in the past month; (iv) binging alcohol in the past month (defined as consuming 5 or more drinks on the same occasion at least once in the past 30 days); and (v) heavy alcohol use (defined as consuming 5 or more drinks on the same occasion 5 or more days in the past 30 days). For brevity, we present the results only for women aged 35-49 in Appendix Table A6, with the results for the other groups available by request. Generally, we find no effect of the warning on any of these substitutes for males or females, and notably no effects for women aged 35-49. Thus, it does not appear that women aged 35-49 substituted toward risky behaviors such as alcohol or marijuana consumption.

6 Discussion

We study the employment effects of the 2007 expanded black box warning on antidepressants. We find that the warning led to a significant decline in the employment of depressed women aged 35-49, but had no impact on the employment of depressed men or on depressed women under the age of 35. We show that this result exists in two nationally representative datasets and survives a number of robustness tests. Moreover, we find that depressed women aged 36-44 experienced an increase in self-reported work and activity limitations due to depression and anxiety after the announcement of the warning. We explore several mechanisms through which the warning may have affected employment, ultimately finding two channels. For depressed women aged 35-49, the same age group experiencing a decline in employment, we find (i) a significant decline in antidepressant drug use at the extensive margin, driven especially by new users, which is consistent with language used in the warnings, and (ii) continuing antidepressant users take notably smaller doses after the warning. Each of these responses likely contributed to the observed decline in employment.

Our findings provide an important example of how health policy can impact human capital development and the economy at large, which may not be considered by policy architects. In the case of the 2007 expanded black box warning on antidepressants, our estimates suggest that

[39] These results are generally robust to including individual fixed effects and dropping the transition round. Parameter magnitudes and significance differ slightly, but lead to qualitatively similar conclusions. The therapy and non-black box drug responses among all other subsamples are not significantly different from zero.

the economic impact was significant. According to the Bureau of Labor Statistics, women aged 35-49 comprised approximately 16.4 percent of the labor force in the US in 2006. Data from the 2005 MEPS cohort suggest that 23.3 percent of women ever working over the 2-year interview period also experienced some form of depression or anxiety in that time frame. Thus, prior to the announcement of the expanded black box warning, depressed women aged 35-49 comprised about 3.8 percent of the US labor force. Our estimates imply that employment among this group fell in response to the warning by about 6.1 percent, representing a 0.23 percentage point reduction in the size of the US labor force, or 352,042 fewer individuals. Again from the 2005 MEPS cohort, average annual income from wages for this demographic (i.e., ever-depressed employed women aged 35-49) was $37,425. As such, the warning led to roughly $13 billion in lost wages.

Reduced wages only capture a fraction of the total impact of the warning. The expansion of the black box warning and ensuing reduction in antidepressant use presumably increased pain and suffering for patients who stopped using or reduced their medications. There were also implications for suppliers of antidepressants. Maybe most importantly, a complete analysis of the costs and benefits of the warning would consider the impact on suicide rates, which speaks to the original objective of the warning. Recall, the 2004 and 2007 antidepressant black box warnings were the result of an extensive research effort by the FDA, which included meta-analyses of nearly 400 randomized control trials (RCTs) and over 100,000 patients. In short, these studies suggested that antidepressant use among children, adolescents, and young adults increased their risk of suicidality; the black box warning was added, and then expanded, to lower the incidence of suicide.

Several studies have shown that suicide rates among children and adolescents *increased* after the 2004 black box warning was added (Gibbons et al., 2007; Busch et al., 2014). Ludwig et al. (2009) provide a thorough rationalization of this unexpected outcome. The authors describe three problems with the RCTs supporting the decision to add the warnings. First, the RCTs were not externally valid as the most severely depressed patients were almost always excluded from the trials. Second, the RCTs were severely under-powered. As the authors explain, at typical suicide rates, to find a 20 percent effect of SSRIs on suicide, one would need 1.9 million subjects. Finding a 5 percent effect would require 30 million patients, which is about twice the number of Americans who suffer from major depressive disorder in any given year. As a result, the researchers conducting these RCTs examine the impact of antidepressants not on suicides, but on suicidal thoughts and attempts. The decision to examine these outcomes leads to the third problem, which is that considering and/or attempting suicide is very different from death via suicide. A small fraction of those who consider suicide attempt it and few attempted suicides are fatal (Cutler et al., 2001; Baldessarini et al., 2006). Moreover, much of the data on suicidal thoughts and attempts suffers from ascertainment bias (i.e., side effects are more common with any active drug than with placebo; thus, those receiving antidepressants complete more doctor visits, meaning more opportunities to report suicidal thoughts and attempts). Ludwig et al. (2009) go on to estimate the effect of SSRI

sales on suicide across 26 countries by exploiting institutional differences in how all drugs are priced, regulated, and distributed. They find that one SSRI pill per capita *reduces* suicide by 5 percent.

In addition to the research discussed above, which suggests that the warnings actually increased suicide rates, Busch et al. (2014) find that the warnings generated perverse indirect effects as well. Using a similar empirical strategy to the one employed in this paper, Busch et al. (2014) find that the 2004 warning lowered GPAs, increased delinquency, and increased tobacco and illicit drug use among depressed adolescent girls. In this paper, we show yet another unintended consequence of the black box warnings on antidepressants; namely, a reduction in employment among depressed women. The sum of these findings, paired with the apparent lack of evidence pointing to any ex-post benefits of the warnings, calls to question the efficacy of the black box warning that is still attached to all antidepressant drugs sold in the US and in many other countries around the world.

References

Balat, J. F., N. W. Papageorge, and S. Qayyum (2018): "Positively Aware? Conflicting Expert Reviews and Demand for Medical Treatment," Working Paper No. 24820, National Bureau of Economic Research.

Baldessarini, R. J., M. Pompili, and L. Tondo (2006): "Suicidal Risk in Antidepressant Drug Trials," *Archives of General Psychiatry*, 63, 246–248.

Blanchflower, D. G. and A. J. Oswald (2016): "Antidepressants and Age: A New Form of Evidence for U-Shaped Well-Being Through Life," *Journal of Economic Behavior & Organization*, 127, 46–58.

Busch, S. H., R. G. Frank, D. L. Leslie, A. Martin, R. A. Rosenheck, E. G. Martin, and C. L. Barry (2010): "Antidepressants and Suicide Risk: How Did Specific Information in FDA Safety Warnings Affect Treatment Patterns?" *Psychiatric Services*, 61, 11–16.

Busch, S. H., R. G. Frank, A. Martin, and C. L. Barry (2011): "Characterizing Declines in Pediatric Antidepressant Use After New Risk Disclosures," *Medical Care Research and Review*, 68, 96–111.

Busch, S. H., E. Golberstein, and E. Meara (2014): "The FDA and ABCs: Unintended Consequences of Antidepressant Warnings on Human Capital," *Journal of Human Resources*, 49, 540–571.

Bütikofer, A. and M. Skira (2018): "Missing Work is a Pain: The Effect of Cox-2 Inhibitors on Sickness Absence and Disability Pension Receipt," *Journal of Human Resources*, 53, 71–122.

Chintagunta, P. K., R. Jiang, and G. Z. Jin (2009): "Information, Learning, and Drug Diffusion: The Case of Cox-2 Inhibitors," *Quantitative Marketing and Economics*, 7, 399–443.

Collins, J. M., K. I. Simon, and S. Tennyson (2013): "Drug Withdrawals and the Utilization of Therapeutic Substitutes: The Case of Vioxx," *Journal of Economic Behavior & Organization*, 86, 148–168.

Cronin, C. J., M. P. Forsstrom, and N. W. Papageorge (2018): "Mental Health, Human Capital, and Labor Market Outcomes," Working Paper.

Currie, J. and B. Madrian (1999): *Health, Health Insurance and the Labor Market*, Elsevier, vol. 3 of *Handbook of Labour Economics*, chap. 50, 3309–3416.

Cutler, D. M., E. L. Glaeser, and K. E. Norberg (2001): "Explaining the Rise in Youth Suicide," in *Risky Behavior Among Youths: An Economic Analysis*, University of Chicago Press, 219–270.

Daysal, N. M. and C. Orsini (2012): "The Miracle Drug: Hormone Replacement Therapy and Labor Market Behavior of Middle-Aged Women," Working Paper.

Dorsey, E. R., A. Rabbani, S. A. Gallagher, R. M. Conti, and G. C. Alexander (2010): "Impact of FDA Black Box Advisory on Antipsychotic Medication Use," *Archives of Internal Medicine*, 170, 96–103.

Dusetzina, S. B., A. S. Higashi, E. R. Dorsey, R. Conti, H. A. Huskamp, S. Zhu, C. F. Garfield, and G. C. Alexander (2012): "Impact of FDA Drug Risk Communications on Health Care Utilization and Health Behaviors: A Systematic Review," *Medical Care*, 50, 466.

Ettner, S. L., R. G. Frank, and R. C. Kessler (1997): "The Impact of Psychiatric Disorders on Labor Market Outcomes," *Industrial and Labor Relations Review*, 51, 64–81.

Garthwaite, C. L. (2012): "The Economic Benefits of Pharmaceutical Innovations: The Case of Cox-2 Inhibitors," *American Economic Journal: Applied Economics*, 4, 116–137.

Gibbons, R. D., C. H. Brown, K. Hur, S. M. Marcus, D. K. Bhaumik, J. A. Erkens, R. M. C. Herings, and J. J. Mann (2007): "Early Evidence on the Effects of Regulators' Suicidality Warnings on SSRI Prescriptions and Suicide in Children and Adolescents," *American Journal of Psychiatry*, 164, 1356–1363.

Grossman, M. (1972): "On the Concept of Health Capital and the Demand for Health," *Journal of Political Economy*, 80, 223.

Hammad, T. A., T. Laughren, and J. Racoosin (2006): "Suicidality in Pediatric Patients Treated with Antidepressant Drugs," *Archives of General Psychiatry*, 63, 332–339.

Harris, M. C., L. M. Kessler, M. N. Murray, and E. Glenn (2018): "Prescription Opioids and Labor Market Pains," Working Paper.

Hoynes, H., D. L. Miller, and J. Schaller (2012): "Who Suffers During Recessions?" *Journal of Economic Perspectives*, 26, 27–48.

Jeon, S.-h. and R. V. Pohl (2018): "Labor Market Effects of Medical Innovation: The Case of Breast and Prostate Cancer," Working Paper.

Kornstein, S. G., A. F. Schatzberg, M. E. Thase, K. A. Yonkers, J. P. McCullough, G. I. Keitner, A. J. Gelenberg, S. M. Davis, W. M. Harrison, and M. B. Keller (2000): "Gender Differences in Treatment Response to Sertraline Versus Imipramine in Chronic Depression," *American Journal of Psychiatry*, 157, 1445–1452.

Libby, A. M., H. D. Orton, and R. J. Valuck (2009): "Persisting Decline in Depression Treatment After FDA Warnings," *Archives of General Psychiatry*, 66, 7.

Ludwig, J., D. E. Marcotte, and K. Norberg (2009): "Anti-Depressants and Suicide," *Journal of Health Economics*, 28, 659–676.

Nemeroff, C. B., A. Kalali, M. B. Keller, D. S. Charney, S. E. Lenderts, E. F. Cascade, H. Stephenson, and A. F. Schatzberg (2007): "Impact of Publicity Concerning Pediatric Suicidality Data on Physician Practice Patterns in the United States," *Archives of General Psychiatry*, 64, 466–472.

Nichol, K. L. (2001): "Cost-Benefit Analysis of a Strategy to Vaccinate Healthy Working Adults Against Influenza," *Archives of Internal Medicine*, 161, 749–759.

Nichol, K. L., S. J. D'Heilly, M. E. Greenberg, and E. Ehlinger (2009): "Burden of Influenza-Like Illness and Effectiveness of Influenza Vaccination Among Working Adults Aged 50-64 Years," *Clinical Infectious Diseases*, 48, 292–298.

Noel, C. (2015): "Antidepressants and Suicidality: History, the Black-Box Warning, Consequences, and Current Evidence," *Mental Health Clinician*, 5, 202–211.

Olfson, M., S. C. Marcus, and B. G. Druss (2008): "Effects of Food and Drug Administration Warnings on Antidepressant Use in a National Sample," *Archives of General Psychiatry*, 65, 94–101.

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

Parkinson, K., J. Price, K. I. Simon, and S. Tennyson (2014): "The Influence of FDA Advisory Information and Black Box Warnings on Individual Use of Prescription Antidepressants," *Review of Economics of the Household*, 12, 771–790.

Prochaska, J. J., H.-Y. Sung, W. Max, Y. Shi, and M. Ong (2012): "Validity Study of the K6 Scale as a Measure of Moderate Mental Distress Based on Mental Health Treatment Need and Utilization," *International Journal of Methods in Psychiatric Research*, 21, 88–97.

Shapiro, B. (2018): "Promoting Wellness or Waste? Evidence from Antidepressant Advertising," *SSRN Electronic Journal*.

Sramek, J. J., M. F. Murphy, and N. R. Cutler (2016): "Sex Differences in the Psychopharmacological Treatment of Depression," *Dialogues in Clinical Neuroscience*, 18, 447.

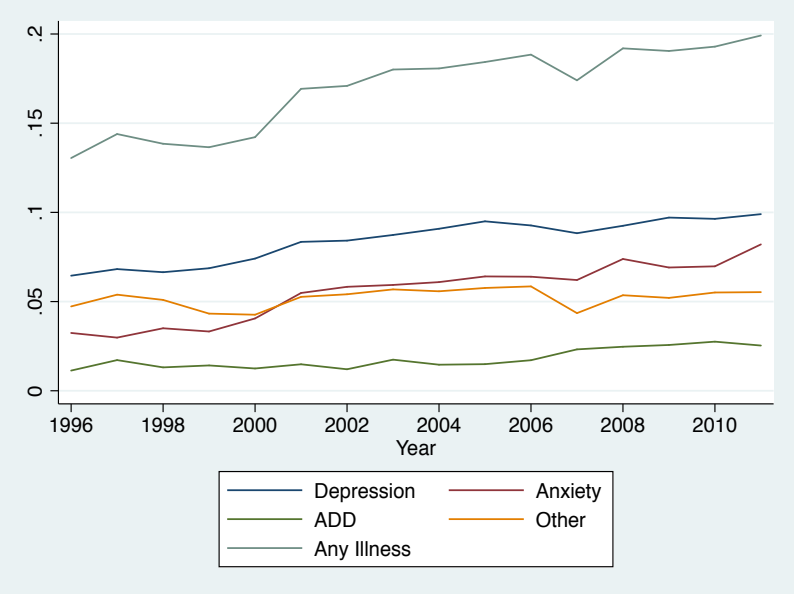
Thirumurthy, H., J. G. Zivin, and M. Goldstein (2008): "The Economic Impact of AIDS Treatment: Labor Supply in Western Kenya," *Journal of Human Resources*, 43, 511–552.

Wong, J., A. Motulsky, M. Abrahamowicz, T. Eguale, D. L. Buckeridge, and R. Tamblyn (2017): "Off-label Indications for Antidepressants in Primary Care: Descriptive Study of Prescriptions from an Indication Based Electronic Prescribing System," *British Medical Journal*, 356, j603.

Young, E. A., S. G. Kornstein, S. M. Marcus, A. T. Harvey, D. Warden, S. R. Wisniewski, G. Balasubramani, M. Fava, M. H. Trivedi, and A. J. Rush (2009): "Sex Differences in Response to Citalopram: A STAR* D Report," *Journal of Psychiatric Research*, 43, 503–511.

Tables and Figures

Figure 1: Mental Health Conditions, 1996 - 2011 MEPS Cohorts



Notes: Illness proportions are calculated using population weighted MEPS data found in the Medical Condition Files. Individuals are grouped by cohort and categorized according to whether they report the illness (as indicated by ICD9 code) over the two year interview period. Any (mental) illness corresponds to ICD9 codes 290 - 319.

Figure 2: Timeline of FDA Rulings and Announcements on Antidepressants and Suicidality

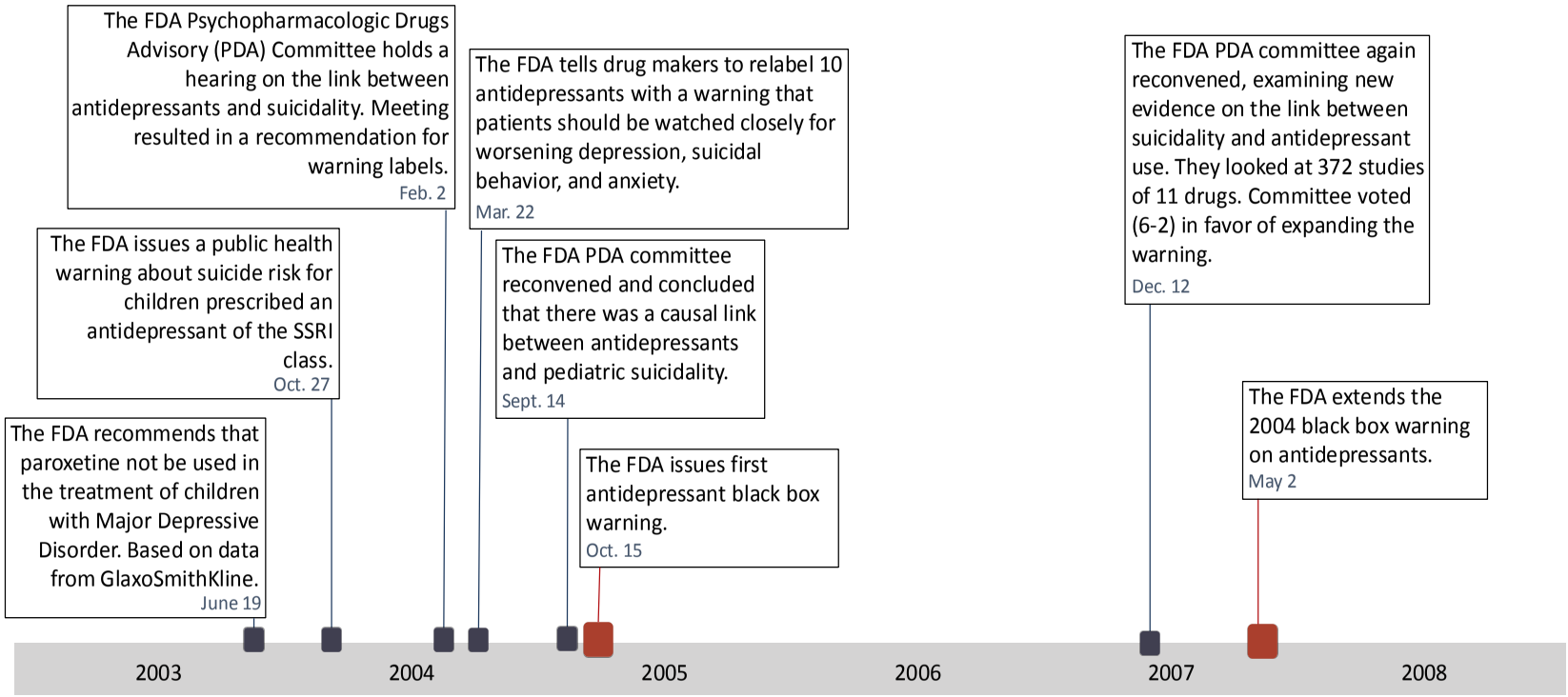


Figure 3: Proportion of Males with a Lifetime MDE Over Time (NSDUH)

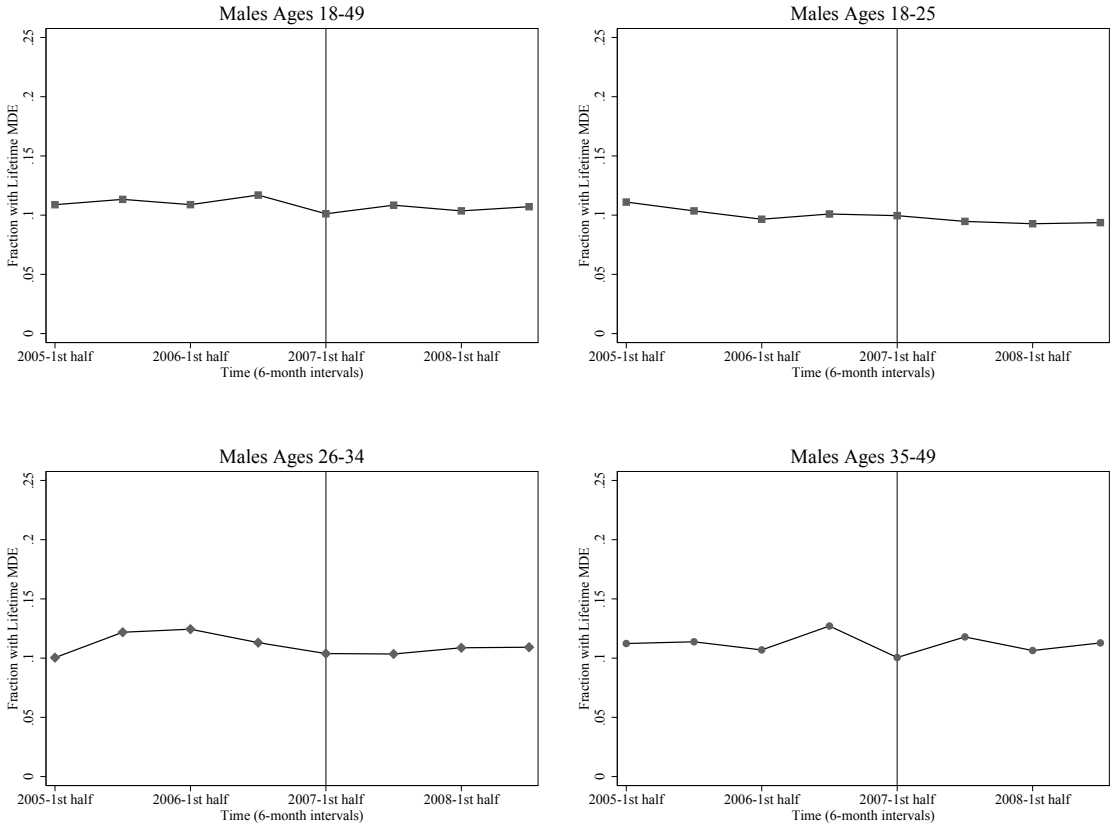


Figure 4: Proportion of Females with a Lifetime MDE Over Time (NSDUH)

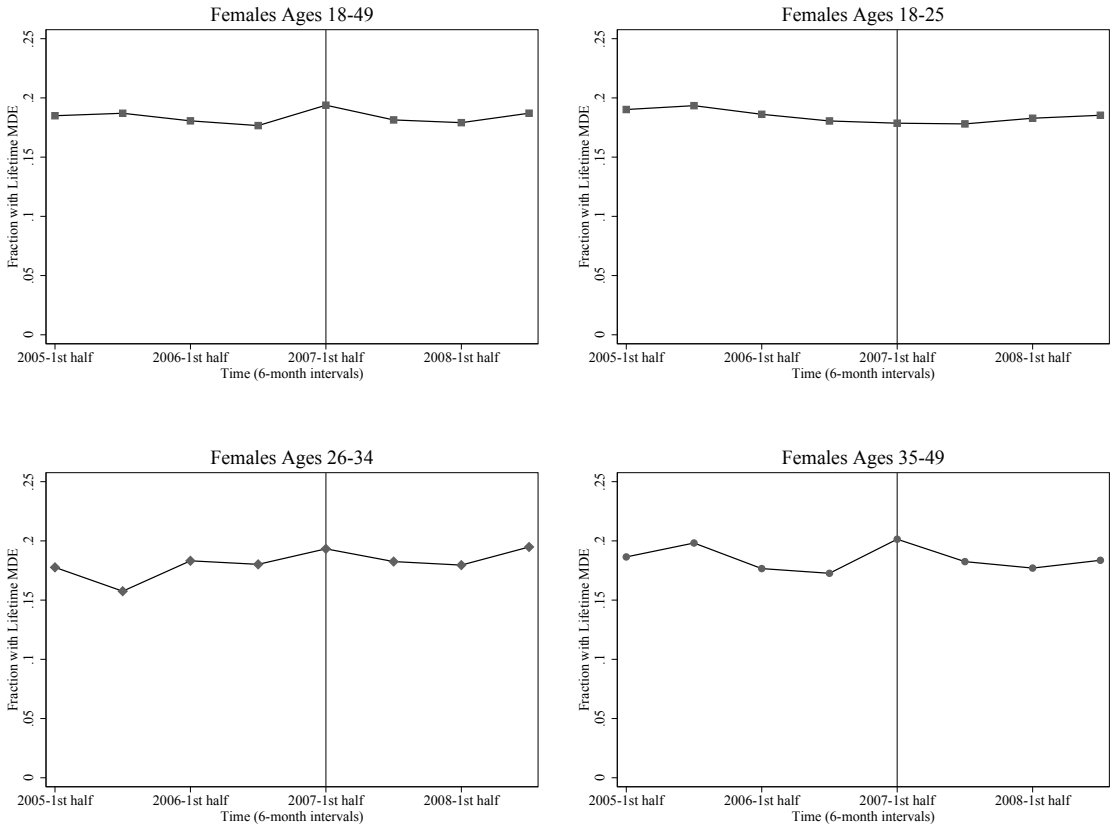
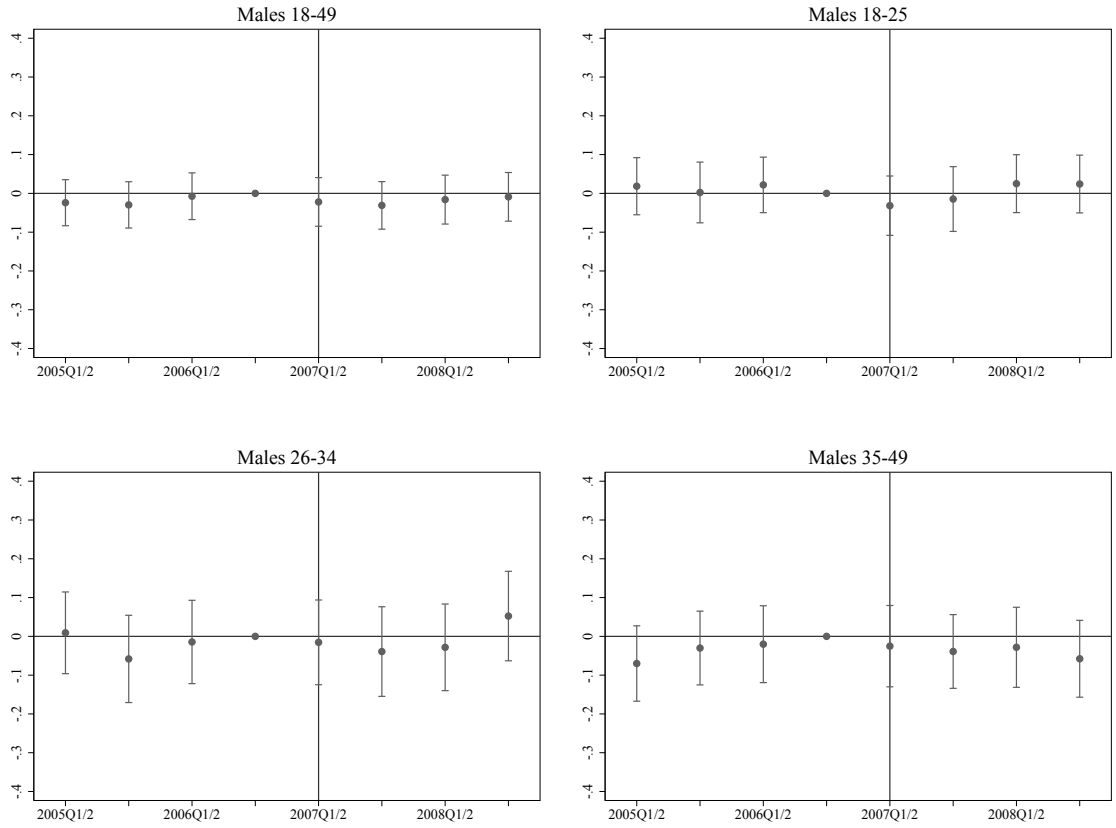
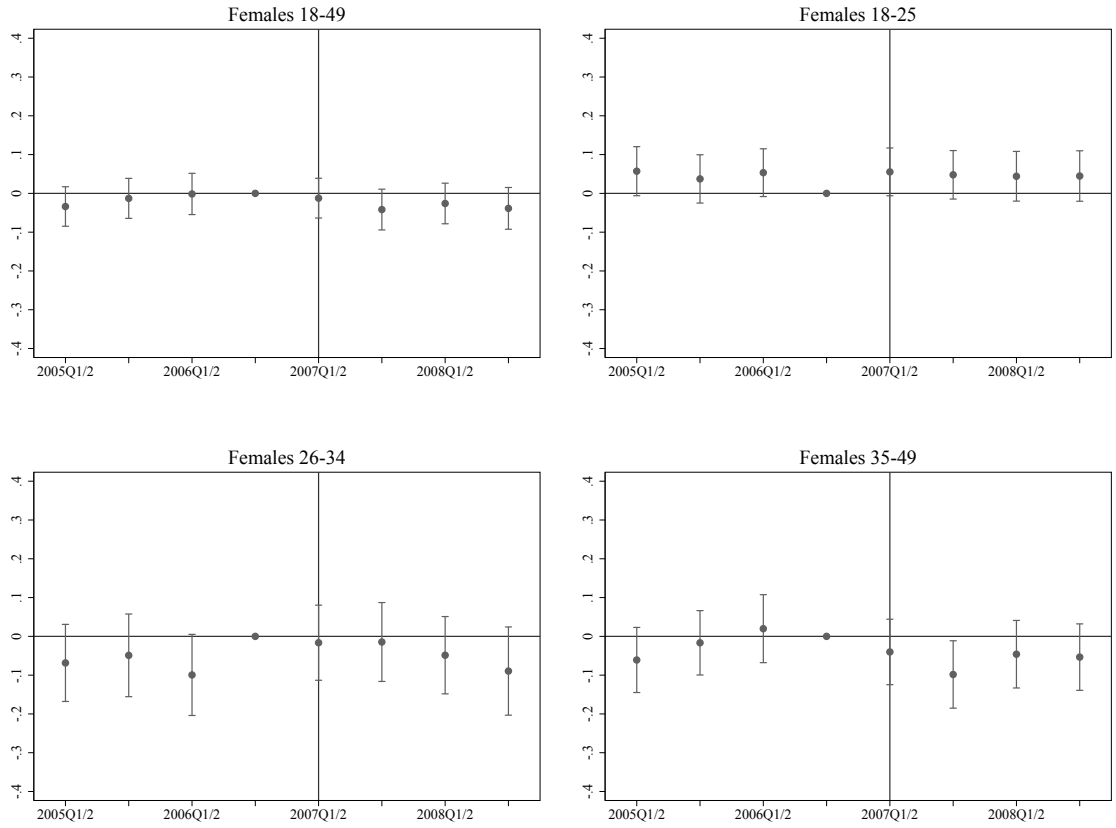


Figure 5: Event Study Estimates for Males (NSDUH)



Notes: Each panel contains plots of the estimates of π_ℓ from Equation 2 with 95 percent confidence interval bars where the probability of employment is the outcome.

Figure 6: Event Study Estimates for Females (NSDUH)



Notes: Each panel contains plots of the estimates of π_ℓ from Equation 2 with 95 percent confidence interval bars where the probability of employment is the outcome.

Table 1: Descriptive Statistics for Males by Age (NSDUH)

	All		18-25		26-34		35-49	
	No MDE	MDE	No MDE	MDE	No MDE	MDE	No MDE	MDE
Employed								
Pre-warning	88.00	79.05	77.11	75.37	91.13	81.50	91.95	79.40
Post-warning	86.48	77.01	74.45	71.98	90.18	81.36	90.73	76.92
Education								
Less than high school	17.11	13.74	22.28	17.78	16.28	12.58	14.86	12.54
High school	31.67	30.34	35.95	34.84	28.81	27.77	31.01	29.68
Some college	25.34	29.39	30.23	35.66	24.78	29.94	23.09	26.24
College graduate	25.89	26.53	11.54	11.73	30.14	29.71	31.04	31.54
Marital Status								
Married	49.42	38.30	9.63	8.72	50.46	37.25	69.62	52.30
Previously married	9.17	16.96	1.06	2.26	7.10	13.54	14.54	25.48
Never married	41.41	44.74	89.31	89.03	42.44	49.21	15.84	22.22
Race/Ethnicity								
White	62.99	73.81	61.58	68.13	57.67	68.71	66.66	79.15
Black	11.88	8.88	13.22	9.42	11.75	11.16	11.25	7.39
Native American	0.48	0.41	0.53	0.62	0.44	0.23	0.48	0.41
Hawaiian/Pacific Islander	0.43	0.32	0.48	0.50	0.42	0.26	0.41	0.28
Asian	4.98	3.31	4.45	4.12	5.84	3.23	4.78	2.98
Multiple	0.92	1.35	1.10	1.17	0.97	1.28	0.80	1.47
Hispanic	18.32	11.93	18.64	16.04	22.91	15.12	15.62	8.33
Metro Status								
Large metro	55.89	55.40	53.42	53.31	56.94	56.98	56.60	55.50
Small metro	28.92	29.23	30.93	31.14	29.21	26.50	27.71	29.84
Non-metro	15.19	15.37	15.65	15.55	13.85	16.52	15.68	14.66
N	52,815	6,233	31,146	3,406	9,107	1,195	12,562	1,632
Share	89.15%	10.85%	90.09%	9.91%	88.94%	11.06%	88.77%	11.23%

Notes: All means and proportions are estimated using the NSDUH sampling weights and include all observations in our sample from 2005-2008.

Table 2: Descriptive Statistics for Females by Age (NSDUH)

	All		18-25		26-34		35-49	
	No MDE	MDE	No MDE	MDE	No MDE	MDE	No MDE	MDE
Employed								
Pre-warning	73.93	71.92	69.25	70.32	72.29	71.93	77.06	72.70
Post-warning	74.59	71.10	69.43	71.23	73.28	74.22	77.86	69.34
Education								
Less than high school	13.29	11.98	16.72	14.32	12.93	10.86	11.81	11.42
High school	29.49	27.22	33.35	33.14	25.34	23.67	29.82	26.19
Some college	28.87	33.35	34.79	37.18	27.16	32.15	26.88	32.11
College graduate	28.35	27.44	15.15	15.37	34.56	33.32	31.50	30.27
Marital Status								
Married	53.55	44.37	17.31	14.62	58.47	47.73	68.70	57.16
Previously married	12.41	19.78	2.45	4.09	10.26	16.33	18.44	29.26
Never married	34.04	35.85	80.24	81.30	31.27	35.94	12.86	13.58
Race/Ethnicity								
White	61.36	74.66	60.04	68.49	57.67	71.96	63.98	79.10
Black	14.48	9.64	15.38	11.78	14.39	9.71	14.09	8.55
Native American	0.56	0.60	0.62	0.65	0.61	0.81	0.50	0.47
Hawaiian/Pacific Islander	0.45	0.28	0.34	0.53	0.50	0.22	0.48	0.19
Asian	5.66	2.29	4.89	3.46	6.36	2.49	5.67	1.62
Multiple	0.88	1.58	1.11	1.90	1.01	2.07	0.71	1.17
Hispanic	16.61	10.94	17.63	13.18	19.47	12.74	14.58	8.90
Metro Status								
Large metro	56.51	50.13	53.50	52.46	57.43	49.14	57.50	49.50
Small metro	28.57	32.35	31.37	32.05	28.86	34.52	27.04	31.37
Non-metro	14.92	17.52	15.14	15.49	13.70	16.34	15.46	19.13
N	54,659	12,809	31,150	7,122	9,542	2,311	13,967	3,376
Share	81.61%	18.39%	81.56%	18.44%	81.87%	18.13%	81.50%	18.50%

Notes: All means and proportions are estimated using the NSDUH sampling weights and include all observations in our sample from 2005-2008.

Table 3: Difference-in-Differences Estimates of the Effect of the Warning on Employment (NSDUH)

	All	18-25	26-34	35-49
		Panel A: Males		
MDE	-0.0894***	-0.0319**	-0.0909***	-0.116***
	(0.0107)	(0.0137)	(0.0198)	(0.0175)
MDE × Post	-0.00440	-0.0106	0.00956	-0.00890
	(0.0157)	(0.0202)	(0.0289)	(0.0256)
R^2	0.0952	0.0783	0.0419	0.0823
N	59048	34552	10302	14194
		Panel B: Females		
MDE	-0.0363***	-0.00992	-0.0252	-0.0532***
	(0.00931)	(0.0109)	(0.0189)	(0.0151)
MDE × Post	-0.0172	0.0105	0.0118	-0.0441**
	(0.0133)	(0.0154)	(0.0268)	(0.0214)
R^2	0.0636	0.0805	0.0890	0.0456
N	67468	38272	11853	17343

Notes: All models estimated by OLS with heteroskedasticty-robust standard errors (in parentheses). Unreported covariates include indicator variables for age (when possible), education, marital status, race and ethnicity, metro type, and time (measured in 6-month periods).

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 4: Time-Varying Effects of the Warning on Employment (NSDUH)

	All	18-25	26-34	35-49
		Panel A: Males		
MDE	-0.0894***	-0.0319**	-0.0910***	-0.116***
	(0.0107)	(0.0137)	(0.0198)	(0.0176)
MDE × 2007Q1/2	-0.00692	-0.0424	0.00155	0.00397
	(0.0256)	(0.0320)	(0.0442)	(0.0441)
MDE × 2007Q3/4	-0.0159	-0.0253	-0.0222	-0.00980
	(0.0247)	(0.0363)	(0.0482)	(0.0379)
MDE × 2008Q1/2	-0.000913	0.0143	-0.0112	0.00101
	(0.0258)	(0.0309)	(0.0458)	(0.0431)
MDE × 2008Q3/4	0.00619	0.0134	0.0693	-0.0285
	(0.0256)	(0.0308)	(0.0482)	(0.0406)
R^2	0.0952	0.0784	0.0426	0.0824
N	59048	34552	10302	14194
		Panel B: Females		
MDE	-0.0363***	-0.00992	-0.0250	-0.0531***
	(0.00931)	(0.0109)	(0.0189)	(0.0151)
MDE × 2007Q1/2	-0.0000384	0.0179	0.0381	-0.0253
	(0.0203)	(0.0232)	(0.0382)	(0.0329)
MDE × 2007Q3/4	-0.0295	0.0103	0.0400	-0.0834**
	(0.0211)	(0.0239)	(0.0413)	(0.0345)
MDE × 2008Q1/2	-0.0138	0.00657	0.00588	-0.0311
	(0.0210)	(0.0249)	(0.0400)	(0.0347)
MDE × 2008Q3/4	-0.0264	0.00724	-0.0350	-0.0386
	(0.0219)	(0.0256)	(0.0486)	(0.0337)
R^2	0.0637	0.0805	0.0893	0.0458
N	67468	38272	11853	17343

Notes: All models estimated by OLS with heteroskedasticty-robust standard errors (in parentheses). Unreported covariates include indicator variables for age (when possible), education, marital status, race and ethnicity, metro type, and time (measured in 6-month periods).

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 5: Robustness Checks of the Effect of the Warning on Employment of Females Aged 35-49 (NSDUH)

	(1) Drop 2007-1st Half	(2) Post=0 in 2007-1st Half	(3) Drop 2005-1st Half	(4) Drop 2005	(5) Alt. Dep. Defn. 1	(6) Alt. Dep. Defn. 2
MDE	-0.0540***	-0.0586***	-0.0378**	-0.0286		
	(0.0151)	(0.0134)	(0.0176)	(0.0224)		
MDE × Post	-0.0513**	-0.0457**	-0.0594**	-0.0694**		
	(0.0233)	(0.0223)	(0.0233)	(0.0271)		
Diagnosed anxiety/depression					-0.0487***	
					(0.0137)	
Diagnosed anxiety/depression × Post					-0.0391**	
					(0.0192)	
MDE or diagnosed anxiety/depression						-0.0530***
						(0.0126)
MDE or diagnosed anxiety/depression × Post						-0.0356**
						(0.0177)
R^2	0.0476	0.0456	0.0453	0.0463	0.0445	0.0460
N	15081	17343	15176	12866	17310	17260

Notes: All models estimated by OLS with heteroskedasticty-robust standard errors (in parentheses). Unreported covariates include indicator variables for age (when possible), education, marital status, race and ethnicity, metro type, and time (measured in 6-month periods).

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 6: Difference-in-Differences Estimates of the Effect of the Warning on Employment (NHIS)

	All	18-25	26-34	35-49
		Panel A: Males		
K6 > 4	-0.146***	-0.0795***	-0.122***	-0.206***
	(0.0122)	(0.0265)	(0.0219)	(0.0165)
K6 > 4 × Post	-0.000735	-0.0480	-0.000896	0.0297
	(0.0185)	(0.0432)	(0.0320)	(0.0239)
R^2	0.135	0.114	0.0721	0.122
N	25146	5477	7054	12615
		Panel B: Females		
K6 > 4	-0.114***	-0.0622***	-0.0728***	-0.164***
	(0.0108)	(0.0236)	(0.0194)	(0.0149)
K6 > 4 × Post	0.00346	0.0329	-0.00738	-0.0108
	(0.0161)	(0.0348)	(0.0291)	(0.0224)
R^2	0.0868	0.103	0.102	0.0752
N	30653	6521	8928	15204

Notes: All models estimated by OLS with heteroskedasticty-robust standard errors (in parentheses). Unreported covariates include age and age squared and indicator variables for education, marital status, race and ethnicity, Census region, and time (measured in 6-month periods).
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 7: Difference-in-Differences Estimates of the Effect of the Warning on Outcomes of Females Aged 36-44 (NHIS)

	Employed	# of Bed Days	Work Limitation Due to Health	Activity Limitation Due to Depression
K6 > 4	-0.129***	12.96***	0.171***	0.0627***
	(0.0189)	(1.903)	(0.0151)	(0.00955)
K6 > 4 × Post	-0.0553*	1.650	0.0431*	0.0468***
	(0.0288)	(2.816)	(0.0241)	(0.0162)
R^2	0.0732	0.0601	0.131	0.0819
N	9102	9020	9110	9101

Notes: All models estimated by OLS with heteroskedasticty-robust standard errors (in parentheses). Unreported covariates include age and age squared and indicator variables for education, marital status, race and ethnicity, Census region, and time (measured in 6-month periods).
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 8: Descriptive Statistics for Males by Age (2006 MEPS Cohort)

	All		18-25		26-34		35-49		36-44	
	No Dep	Dep	No Dep	Dep	No Dep	Dep	No Dep	Dep	No Dep	Dep
AD Use in Round										
1	0.005	0.281	0.000	0.326	0.002	0.235	0.007	0.282	0.013	0.194
2	0.004	0.285	0.000	0.266	0.001	0.271	0.008	0.291	0.010	0.183
3	0.005	0.304	0.000	0.190	0.000	0.257	0.011	0.363	0.017	0.260
4	0.006	0.325	0.000	0.299	0.003	0.260	0.010	0.369	0.012	0.271
5	0.004	0.248	0.003	0.175	0.003	0.183	0.005	0.307	0.005	0.181
Education										
Less than high school	0.167	0.135	0.289	0.190	0.149	0.065	0.126	0.153	0.148	0.144
High school	0.560	0.603	0.650	0.810	0.503	0.472	0.559	0.605	0.521	0.616
College graduate	0.268	0.261	0.055	0.000	0.345	0.463	0.310	0.239	0.328	0.234
Marital Staus										
Married	0.488	0.406	0.052	0.063	0.506	0.351	0.664	0.531	0.670	0.522
Previously married	0.111	0.190	0.022	0.033	0.102	0.154	0.157	0.243	0.140	0.247
Never married	0.401	0.404	0.926	0.904	0.392	0.496	0.179	0.226	0.191	0.231
Parent	0.394	0.284	0.040	0.008	0.354	0.210	0.566	0.400	0.611	0.397
Non-White	0.186	0.145	0.229	0.134	0.165	0.095	0.179	0.131	0.203	0.176
MSA	0.854	0.850	0.855	0.812	0.853	0.907	0.851	0.834	0.854	0.833
Employed	0.925	0.827	0.887	0.789	0.938	0.880	0.932	0.810	0.931	0.823
N	2554		491		601		1315		757	
% Depressed	10.4		7.6		11.1		11.3		10.1	

Notes: All means and proportions are estimated using the MEPS longitudinal sampling weights and include all observations in our sample from the 2006 MEPS cohort. We report the raw number of individuals in each subsample; however, depression shares are weighted.

Table 9: Descriptive Statistics for Females by Age (2006 MEPS Cohort)

	All		18-25		26-34		35-49		36-44	
	No Dep	Dep	No Dep	Dep	No Dep	Dep	No Dep	Dep	No Dep	Dep
AD Use in Round										
1	0.019	0.301	0.013	0.216	0.016	0.227	0.024	0.351	0.024	0.314
2	0.020	0.346	0.007	0.075	0.016	0.263	0.022	0.420	0.026	0.396
3	0.020	0.377	0.008	0.256	0.013	0.304	0.028	0.438	0.030	0.437
4	0.017	0.373	0.003	0.215	0.015	0.324	0.022	0.439	0.024	0.398
5	0.024	0.386	0.008	0.197	0.023	0.312	0.031	0.466	0.037	0.414
Education										
Less than high school	0.111	0.126	0.190	0.179	0.101	0.084	0.088	0.122	0.102	0.161
High school	0.571	0.610	0.670	0.744	0.537	0.634	0.550	0.575	0.532	0.552
College graduate	0.315	0.262	0.136	0.066	0.362	0.280	0.356	0.303	0.366	0.286
Marital Staus										
Married	0.545	0.504	0.147	0.201	0.564	0.429	0.679	0.591	0.683	0.488
Previously married	0.136	0.216	0.031	0.032	0.101	0.178	0.189	0.272	0.180	0.303
Never married	0.319	0.280	0.822	0.767	0.335	0.393	0.132	0.137	0.137	0.209
Parent	0.550	0.550	0.176	0.207	0.511	0.584	0.686	0.597	0.751	0.659
Non-White	0.231	0.179	0.256	0.209	0.222	0.230	0.226	0.145	0.246	0.184
MSA	0.865	0.828	0.867	0.855	0.887	0.828	0.848	0.825	0.835	0.800
Employed	0.813	0.737	0.853	0.837	0.785	0.795	0.812	0.695	0.797	0.692
N	3081		476		793		1638		890	
% Depressed	20.0		13.5		18.4		22.5		20.4	

Notes: All means and proportions are estimated using the MEPS longitudinal sampling weights and include all observations in our sample from the 2006 MEPS cohort. We report the raw number of individuals in each subsample; however, depression shares are weighted.

Table 10: Difference-in-Differences Estimates of the Effect of the Warning on Extensive Margin Antidepressant Use Excluding Differential Linear Trends (2006 MEPS Cohort)

	All	18-25	26-34	35-49	36-44
			Panel A: Males		
Dep	0.278***	0.377***	0.233***	0.303***	0.211***
	(0.028)	(0.100)	(0.060)	(0.057)	(0.044)
Post	0.010**	-0.016	0.020*	0.003	0.001
	(0.005)	(0.039)	(0.011)	(0.007)	(0.009)
Dep × Post	0.013	0.102	0.014	0.037	0.014
	(0.024)	(0.108)	(0.047)	(0.032)	(0.046)
R^2	0.242	0.245	0.222	0.259	0.149
N	2554	491	601	1315	757
% Depressed	10.4	7.6	11.1	11.3	10.1
			Panel B: Females		
Dep	0.315***	0.166***	0.237***	0.375***	0.361***
	(0.019)	(0.049)	(0.037)	(0.026)	(0.036)
Post	-0.019**	0.017	-0.004	-0.033***	-0.043***
	(0.008)	(0.014)	(0.018)	(0.010)	(0.014)
Dep × Post	0.035**	0.048	0.068*	0.036	0.005
	(0.017)	(0.050)	(0.037)	(0.022)	(0.030)
R^2	0.246	0.142	0.197	0.287	0.259
N	3081	476	793	1638	890
% Depressed	20.0	13.5	18.4	22.5	20.4

Notes: All models are estimated using the MEPS longitudinal sampling weights. Heteroskedasticity robust standard errors are clustered at the individual level. Individual-level controls are included in all models. We report the raw number of individuals in each subsample; however, depression shares are weighted.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 11: Difference-in-Differences Estimates of the Effect of the Warning on Extensive Margin Antidepressant Use Including Differential Linear Trends (2006 MEPS Cohort)

	All	18-25	26-34	35-49	36-44
			Panel A: Males		
Dep	0.323***	0.377***	0.351***	0.284***	0.182***
	(0.039)	(0.100)	(0.090)	(0.047)	(0.053)
Post	0.020	-0.016	0.027	0.033	0.074*
	(0.018)	(0.039)	(0.030)	(0.028)	(0.043)
Dep × Post	0.070	0.102	0.163*	0.013	-0.024
	(0.043)	(0.108)	(0.091)	(0.056)	(0.069)
t	-0.000	-0.001	0.001	-0.000	0.003
	(0.002)	(0.004)	(0.003)	(0.003)	(0.005)
Dep × t	-0.023	-0.058*	-0.061*	0.010	0.015
	(0.015)	(0.035)	(0..32)	(0.019)	(0.020)
Post × t	-0.005	0.004	-0.006	-0.008	-0.019*
	(0.005)	(0.011)	(0.008)	(0.008)	(0.012)
R^2	0.243	0.251	0.229	0.259	0.150
N	2554	491	601	1315	757
% Depressed	10.4	7.6	11.1	11.3	10.1
			Panel B: Females		
Dep	0.252***	0.186**	0.224***	0.283***	0.244***
	(0.031)	(0.079)	(0.077)	(0.037)	(0.050)
Post	-0.014	0.047	-0.014	-0.029	-0.026
	(0.021)	(0.051)	(0.042)	(0.030)	(0.043)
Dep × Post	-0.044	0.073	0.052	-0.080**	-0.143***
	(0.034)	(0.105)	(0.082)	(0.038)	(0.056)
t	-0.001	-0.004	-0.001	0.001	0.007
	(0.002)	(0.005)	(0.004)	(0.004)	(0.006)
Dep × t	0.032***	-0.010	0.006	0.047***	0.061***
	(0.012)	(0.036)	(0.031)	(0.014)	(0.019)
Post × t	0.003	-0.009	0.004	0.006	0.004
	(0.006)	(0.013)	(0.011)	(0.008)	(0.012)
R^2	0.246	0.142	0.197	0.289	0.262
N	3081	476	793	1638	890
% Depressed	20.0	13.5	18.4	22.5	20.4

Notes: All models are estimated using the MEPS longitudinal sampling weights. Heteroskedasticity robust standard errors are clustered at the individual level. Individual-level controls are included in all models. We report the raw number of individuals in each subsample; however, depression shares are weighted.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 12: Difference-in-Differences Estimates of the Effect of the Warning on New and Continuing Antidepressant Use (2006 MEPS Cohort)

	New AD Use				Cont. AD Use
	Adults 25-64	Men 25-64	Women 25-64	Women 35-49	Women 35-49
Dep	0.066***	0.068***	0.066***	0.044**	0.240***
	(0.011)	(0.017)	(0.015)	(0.022)	(0.035)
Post	0.014	0.023*	0.004	-0.002	-0.027
	(0.011)	(0.014)	(0.017)	(0.029)	(0.030)
Dep × Post	-0.030*	-0.011	-0.039*	-0.059*	-0.021
	(0.017)	(0.027)	(0.021)	(0.034)	(0.038)
t	0.001	0.002	0.001	0.001	-0.000
	(0.001)	(0.001)	(0.002)	(0.004)	(0.003)
Dep × t	0.009*	0.001	0.014*	0.025**	0.022
	(0.005)	(0.009)	(0.007)	(0.011)	(0.013)
Post × t	-0.004	-0.006	-0.002	-0.001	0.007
	(0.003)	(0.004)	(0.005)	(0.008)	(0.008)
N	7262	3284	4018	1632	1638
% Depressed	18.0	12.7	23.0	22.8	22.8

Notes: All models are estimated using the MEPS longitudinal sampling weights. Heteroskedasticity robust standard errors are clustered at the individual level. Individual-level controls are included in all models. We report the raw number of individuals in each subsample; however, depression shares are weighted.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 13: Robustness Checks of the Effect of the Warning on Extensive Margin Antidepressant Use of Females Aged 35-49 (2006 MEPS Cohort)

	(1)		(2)		(3)	
	Drop Transition Round		Individual FE		Alt. Dep. Defn.	
	35-49	36-44	35-49	36-44	35-49	36-44
Dep	0.281***	0.228***			0.099***	0.071***
	(0.039)	(0.053)			(0.023)	(0.030)
Post	-0.004	0.051	-0.015	-0.003	-0.035	-0.014
	(0.084)	(0.127)	(0.029)	(0.041)	(0.035)	(0.050)
Dep × Post	-0.093*	-0.190**	-0.050	-0.103**	-0.073***	-0.096***
	(0.056)	(0.079)	(0.033)	(0.049)	(0.025)	(0.035)
t	-0.001	0.004	-0.001	0.007	0.000	0.008
	(0.004)	(0.007)	(0.005)	(0.007)	(0.005)	(0.008)
Dep × t	0.048***	0.069***	0.039***	0.049***	0.031***	0.035***
	(0.016)	(0.022)	(0.012)	(0.017)	(0.009)	(0.012)
Post × t	0.003	-0.008	0.003	-0.003	0.009	0.002
	(0.017)	(0.026)	(0.008)	(0.087)	(0.009)	(0.012)
N	1638	890	1638	890	1638	890
% Depressed	22.5	20.4	22.5	20.4	22.5	20.4

Notes: All models are estimated using the MEPS longitudinal sampling weights. Heteroskedasticity robust standard errors are clustered at the individual level. Individual-level controls are included in all models. We report the raw number of individuals in each subsample; however, depression shares are weighted. Specification (1) drops the transition round for each individual. Specification (2) includes individual-specific fixed effects. Specification (3) uses an alternative definition of depression where $Dep_i = 1$ if $K6 > 4$ in any interview round.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 14: Effect of "May 1" Placebo Warning on Extensive Margin Antidepressant Use of Females Aged 35-49 in Other Years (MEPS)

	35-49	36-44
Dep × $\text{Post}_{1/5/2003}$	0.022	0.074
	(0.046)	(0.060)
	[329]	[184]
Dep × $\text{Post}_{1/5/2004}$	0.031	0.046
	(0.046)	(0.059)
	[364]	[213]
Dep × $\text{Post}_{1/5/2005}$	0.060	0.047
	(0.041)	(0.051)
	[347]	[193]
Dep × $\text{Post}_{1/5/2006}$	-0.028	-0.061
	(0.043)	(0.059)
	[359]	[187]
Dep × $\text{Post}_{1/5/2007}$	-0.080**	-0.143***
	(0.038)	(0.056)
	[364]	[182]
Dep × $\text{Post}_{1/5/2008}$	0.014	0.013
	(0.045)	(0.056)
	[230]	[130]
Dep × $\text{Post}_{1/5/2009}$	-0.000	0.069
	(0.039)	(0.046)
	[361]	[210]

Notes: All models are estimated using the MEPS longitudinal sampling weights. Heteroskedasticity robust standard errors (in parentheses) are clustered at the individual level. Individual-level controls are included in all models. We report the raw number of depressed individuals in each subsample in brackets.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 15: Difference-in-Differences Estimates of the Effect of the Warning on Intensive Margin Antidepressant Use (2006 MEPS Cohort)

	Pills		MG Per Pill	
	(1)	(2)	(3)	(4)
		Panel A: Males		
All [N=76]	0.046	0.127	0.107	0.315
	(0.373)	(0.349)	(0.266)	(0.216)
18-25 [N=9]	-0.584	-0.543	-0.536	-0.430
	(1.126)	(1.090)	(0.926)	(0.859)
26-34 [N=13]	-1.978**	-1.923**	-0.269	0.119
	(0.700)	(0.798)	(0.617)	(0.793)
35-49 [N=50]	0.561	0.574	0.307	0.350
	(0.423)	(0.453)	(0.264)	(0.289)
36-44 [N=17]	-0.963	-1.109	0.009	0.155
	(0.724)	(0.647)	(0.408)	(0.435)
		Panel B: Females		
All [N=195]	0.171	0.086	-0.097	-0.225
	(0.212)	(0.210)	(0.187)	(0.180)
18-25 [N=6]	0.669	0.361	0.799	0.761
	(0.967)	(0.965)	(1.486)	(1.238)
26-34 [N=39]	0.504	0.369	0.506	0.293
	(0.352)	(0.322)	(0.343)	(0.355)
35-49 [N=141]	0.113	0.067	-0.347*	-0.436**
	(0.267)	(0.260)	(0.190)	(0.185)
36-44 [N=64]	-0.014	-0.153	-0.371	-0.470*
	(0.433)	(0.380)	(0.270)	(0.255)
FE	N	Y	N	Y

Notes: All models are estimated using the MEPS longitudinal sampling weights. Heteroskedasticity robust standard errors (in parentheses) are clustered at the individual level. Individual-level controls are included in all models. We report the raw number of individuals in each subsample in brackets.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$