A Study of the Dynamic Relationships between Depression, Treatment, and Work Behavior

by
M. Katherine Cloud

A dissertation submitted to the faculty of the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Economics.

Chapel Hill
2010

Approved by:

Donna B. Gilleskie, Advisor

David Guilkey, Reader

John S. Akin, Reader

David Blau, Reader

Marisa Domino, Reader

ABSTRACT

M. KATHERINE CLOUD: A Study of the Dynamic Relationships between Depression, Treatment, and Work Behavior
(Under the direction of Donna B. Gilleskie)

It has been shown in the literature that depression has a significant negative correlation with employment outcomes as measured by labor force participation, earnings, work attendance and job performance. I expand the understanding of this relationship by exploring the effect of depression on employment choices as well as treatment choices over time rather than simply examining correlations at a point in time. Other health related outcomes and the relationship between choices and mental health will be examined. My analysis follows initially depressed individuals for nine months and examines the dynamic relationship between health status and function, treatment decisions and employment outcomes. I consider a dynamic model of individual decisions over time where lagged endogenous behavior is allowed to influence current behavior or health outcomes. Results indicate that depression does have a significant effect on labor productivity. Individuals who were the most depressed at the baseline interview saw the largest improvements in productivity following treatment. However, the estimates imply that depression is not a significant determinant of the worker's attendance at work.

DEDICATION

This dissertation is dedicated to the loving memory of my grandfather Dr. Alvin Hugh Dempsey.

ACKNOWLEDGMENTS

I would like to thank Dr. Donna Gilleskie for all of her hard work and dedication to her students. Her advice and support has been invaluable. Additional thanks go to Eli Lilly for the use of the ARTIST trial data and Rebecca Robinson for her help with questions on these data.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

ARTIST	A Randomized Trial Investigating SSRI Treatment
CDS	Chronic Disease Score
CPS	Current Population Study
DALYs	Disability Adjusted Life Years
DDI	Depression Diagnosis Interview
DPCRC	Duke Primary Care Research Consortium
DSM-IV	Diagnostic and Statistical Manual, fourth edition
ECA	Epidemiological Catchment Area
FD	First Difference
FE	Fixed Effects
FOD	Forward Orthogonal Deviations
GBD	Global Burden of Disease
GMM	Generalized Method of Moments
HERO	Health Enhancement Research Organization
HSCL	Hopkins Symptoms Checklist
HSCL-20	Hopkins Symptoms Checklist-20
IID	Independent and identially distributed
IV	Instrumental Variables
LPT	Lost Productive Time
MOS	Medical Outcomes Study
NCS	National Comorbidity Study
NIMH	National Institutes for Mental Health
OLS	Ordinary Least Squares
PCN	Primary Care Network
PCP	Primary Care Provider

PCS-12 Physical Component Score

PHQ-9 9-item Patient Health Questionnaire

PSS Physical Symptoms Scale

QALY Quality Adjusted Life Year

SF-12 12-item Short-form Health Survey

SF-36 36-item Short-form Health Survey

SSRI Selective Serotonin Reuptake Inhibitors

USPS United States Postal Service

WG Within-Groups

WLQ Work Limitations Questionnaire

YLD Years Lived with Disability

ZCTA Zipcode Tabulation Areas

Chapter 1

Introduction

The impact of mental health on the labor market is a topic of increasing interest to employers, policy makers, and to both labor and health economists. Following the findings of the Harvard School of Public Health in its 1996 (updated in 2003) study "Global Burden of Disease (GBD)," the *Global Business and Economic Roundtable on Addiction and Mental Health* was formed in 1998. The Roundtable consists of business and health leaders who have proposed that mental health is a business and economic issue.[1] They highlight the fact that mental health problems are driving disability rates within the North American labor force representing significant social and business costs and leading to lower productivity. The global information economy is, by definition, dependent on mental performance. Therefore, mental health among the labor force is a significant determinant of output much like physical health was in the industrial economy. Mental health is also closely tied to a depressed individual's physical health, notably, heart disease and fatigue. In the public sector, congressional lawmakers and mental health experts are discussing ways to meet the goals laid out in *The President's New Freedom Commission on Mental Health.*[2]

The term mental illness includes a vast array of diseases causing varying levels of limitations. Major depressive disorder, also known as major depression, unipolar depression, clinical depression or simply depression, is a common mental disorder characterized by a pervasive low mood, loss of interest in a person's usual activities, diminished ability to experience

[1] They base this proposal on four facts drawn from the Harvard and other studies.

[2] The commission from the George W. Bush Administration, 2002, pinpoints a range of mental health challenges.

pleasure, feelings of guilt or low self-worth, disturbed sleep or appetite, low energy and poor concentration. It is the most widespread of all psychiatric disorders and one of the most common medical conditions in the United States. It has been estimated that depression affects 18.8 million Americans each year, afflicting between 10 and 25 percent of American women and between five and 12 percent of men in their lifetime (Kessler et al., 2005).

The Harvard Study reports that in 2000, depression was estimated to be the fourth leading contributor to the global burden of disease, and today, depression is already the second cause of Disability Adjusted Life Years (DALYs) in the age category 15-44 years for both sexes combined.[3] The same study estimated that by 2020, only ischemic heart disease will contribute a larger worldwide economic burden in DALYs calculated for all ages, both sexes. These projections show that with the aging of the world population and the conquest of infectious diseases, psychiatric and neurological conditions could increase their share of the total global disease burden by almost half, from 10.50% of the total burden to almost 15.00% in 2020.

This burden results from both direct health care costs and indirect costs including effects in the labor market, high correlation with poor physical health, suicide and disability costs. Greenberg et al. (2003) estimated that the United States spent $83.10 billion in 2000 for costs associated with depression, an increase from $77.40 billion in 1990. The direct treatment costs were estimated to account for 31.40% of the total costs for depression, with workplace costs making up 62.00% of the economic costs of depression (Greenberg et al., 2003). Multiple studies have shown people with depression also have higher total healthcare utilization rates when compared to individuals without depression. The Health Enhancement Research Organization (HERO) analyzed employee medical costs for 46,000 employed persons and found that of the health risk factors studied (smoking, sedentary lifestyle, high cholesterol levels, hypertension, poor diet, being overweight, excessive alcohol consumption, high blood glucose

[3] The Harvard Global Burden of Disease study developed a single measure to allow comparison of the burden of disease across many different disease conditions by including both death and disability. This measure was called Disability Adjusted Life Years (DALYs). DALYs measure lost years of healthy life regardless of whether the years were lost to premature death or disability. The disability component of this measure is weighted for severity of the disability. For example, disability caused by major depression was found to be equivalent to blindness or paraplegia whereas active psychosis seen in schizophrenia produces disability equal to quadriplegia.

levels, high stress and depression), depression predicted the largest increase in medical costs. The HERO study estimated that depression predicted a 70.00% increase in medical costs compared to medical costs for those without depression. Another study noted that depressed employees use, on average, more than $4000 per year in medical services compared to less than $1000 per year for those without depression indicating that depression is a significant element in rising healthcare costs (Sipkoff, 2006). Both direct and indirect costs will vary based on the severity of the depression.

In the labor market, depressed individuals are less likely to be employed, more likely to be absent if employed and have lower productivity while at work. Compounding the effect, many affected individuals are in the prime of their working lives, unlike other disabling conditions that often occur later in life. The GBD study estimated that major depression is the single most burdensome illness in the middle life years.

The costs related to depression often come in the form of lower productivity (due to reduced cognitive abilities or impaired concentration) while at work, called presenteeism, and a higher number of missed days of work, termed absenteeism. Data from the Epidemiological Catchment Area (ECA) and the National Comorbidity Study (NCS) provide estimates of the annual total cost of depression at $43.70 billion in 1990. Employment effects were substantial: $11.70 billion for work absenteeism and $12 billion for reduced productivity. A 1996 reestimate of the losses totaled $53 billion (Swindle et al., 2001). Other estimates have been as high as $51.50 billion a year in lost productivity (Lerner et al., 2004).

Many studies address issues concerning either mental health or depression and labor outcomes. These studies are primarily static, considering only one time period or using repeated cross sections, despite the fact that depression is a recurring disease for the individual. It is estimated that of individuals who have had a single major depressive episode, 50%-60% may develop a second episode within 10 years of the initial episode. About 90% of those who have had two episodes may have a third. The relapse rate is estimated to be 40% within 15 weeks among persons who have had at least three lifetime episodes and a relapse rate of 65% within the first year, if left untreated, for these individuals (Bloom, 2004).

A few studies use panel data but can only report an association between depression and work outcomes. They do not account for treatment decisions and efficacy or they do not address individual unobserved heterogeneity that may affect the individual's decisions and outcomes.

The ideal data set would detail the health, labor and demographic characteristics of a nationally representative sample of individuals followed over time. In reviewing the literature, there were not any studies that use data with all of these characteristics. I use the ARTIST (A Randomized Trial Investigating SSRI [4] Treatment) trial data which has the advantages of including many health measures (both physical and mental) and labor questions asked together with a substantial amount of geographic variation. It is in the form of a panel and follows 573 individuals for 9 months. Disadvantages include the fact that these data are based on self report, the data only cover a nine-month period and all of the individuals are depressed at baseline causing selection bias issues. The purpose of the trial was to investigate the efficacy of specific SSRI treatments on depressed primary care patients.

The contributions of this dissertation follow: This study extends the models estimated with panel data by allowing for potential dynamic relationships between depression and employment over time, considering the role of treatment in affecting both health transitions as well as observed employment outcomes, using the time series element of the data to address instrumentation and using multiple measures of depression and labor outcomes.

The remainder of this dissertation is organized in the following chapters. Chapter 2 contains a literature review. Chapter 3 describes the data. Chapter 4 presents a theoretical model. The empirical model is found in Chapter 5. Chapter 6 discusses results. Chapter 7 concludes, discussing implications for policy formulation and further research.

[4] Selective Serotonin Reuptake Inhibitors (SSRIs) are the most commonly prescribed class of antidepressants.

Chapter 2

Background

The literature confirms an association between depression and employment probabilities, presenteeism and absenteeism. It further suggests that there are substantial employer costs associated with the employment of depressed individuals.

Using data from the National Comorbidity Survey (NCS)[1] to identify the importance of depression in the labor force, Marcotte et al. (1999) find that rates of depression are similar in and out of the labor force. However, they report that within the labor force, depression is strongly associated with unemployment, with a particularly strong relationship for middle age workers. Using an instrumental variables approach, Ettner et al. (1997) find a decrease in the employment probability for both women and men with major depression. Mullahy and Sindelar (1990) show that mental health was a significant determinant of the probability of full-time work for New Haven adults. Broadhead et al. (1990) find similar results and report that persons with minor depression were more likely to be unemployed. In more recent work, Lerner et al. (2004) report that all forms of depression lead to significant increases in job turnover. Specifically focusing on employed individuals with depression, they report that at the six-month follow-up, unemployment rates had increased by 12% for the participants in the major depression group, compared with 2% for individuals in the control group and 3% for those affected by rheumatoid arthritis. The same study reports that "among participants who were still employed, those with depression had significantly more job turnover, presenteeism

[1] The NCS is a nationally representative sample. It is longitudinal and cross-sectional and only compares cohorts of individuals rather than a specific individual over the life cycle.

and absenteeism."

Depite higher unemployment rates for depressed individuals, it has been estimated that more than 70 percent of people diagnosed with depression are employed (Lerner et al., 2004). Estimates have varied with respect to whether depression impacts productivity costs predominantly through absenteeism or presenteeism.

Kessler et al. (2001) find that employed patients with psychiatric disorders (including depression) had significantly greater mean monthly number of work loss days (0.25) and work cutback days (1.09) compared to people with no disorders. Druss et al. (2001), using data from three major American companies, report that those who displayed chronic symptoms of depression were twice as likely to miss work due to health reasons, and seven times as likely to report missed work days at the time of the follow up survey. Bill Wilkerson[2] reports in his speech "Mental Health: The Ultimate Productivity Weapon in the Post- September 11th economy" that the average number of workdays lost to one case of depression is about forty or $10,000 per absent employee for wage replacements and the company's share of drug therapies with a group health benefit plan. In more recent studies, Kessler et al. (2006) studied work impairment due to chronic conditions including major depression. They report estimates that major depressive disorder was associated with 27.20 lost workdays per ill worker per year.

The indirect costs of depression continue to increase when considering lowered productivity (due to reduced cognitive abilities or impaired concentration) while at work, or presenteeism. Brouwer et al. (2002) report that presenteeism can occur before and after absence from work. Berndt et al. (1998) find that perceived work performance is inversely related to the severity of depression and that a reduction in depression severity is associated with a rapid improvement in perceived work performance. In their study to explain depression's effect on work productivity, they use panel data with three interviews per individual over a 12-week period. They do not address the issue of unobserved heterogeneity and explain much of the change in outcomes as regression to the mean. Von Korff et al. (1992) evaluated untreated depressed patients who frequently used health care in a Health Maintenance Organization

[2]He is the Co-Founder and CEO, *Global Business and Economic Roundtable on Addiction and Mental Health.*

over a period of 12 months. Respondents whose depression did not improve over the period reported very high levels of work impairment that did not change significantly. In comparison, respondents with depression rated as severe at baseline whose depression did improve over the follow-up reported a 36% reduction in work impairment days. The study does use panel data. However, these results do not prove that the reduced productivity is a consequence of the depression. Another possibility is that some other unmeasured variable (e.g., difficulty in getting along with a work supervisor) increases both depression and work impairment. The purpose of the Von Korff et al. (1992) study was primarily medical and the results are based on associations between depressive episode improvement and work impairment. Stewart et al. (2003) use data from the *American Productivity Audit* in 2001-2002. They find that depressed individuals had 5.6 hours per week more lost productive time than non-depressed workers. They conclude that 81% of the lost productive time was due to presenteeism. In this study, excess lost productive time (LPT) costs from depression were derived as the difference in LPT for depressed individuals minus the expected LPT for non-depressed individuals projected to the US labor force. Again, they do not account for individual differences or treatment decisions and outcomes. Using data from 16,651 employees, Burton et al. (2004) demonstrated that depression was highly associated with work limitations that predict a worker's productivity while on the job, including limitations in time management, interpersonal/mental functioning, and overall output.

It is important to address the treatment decision in any study of depression. Between 80 and 90 percent of individuals suffering from a major depressive disorder can be treated successfully, yet the National Institute of Mental Health estimates that only one in three with the illness ever seek treatment. Among a depressed sample, absenteeism is reduced and productivity improves regardless of treatment choice[3] (Mintz et al. 1992, Simon et al. 1998). In 1992, Mintz et al. (1992) looked at different studies that evaluated the effects of antidepressants and psychotherapy on work impairment. Work outcomes were generally good when the treatment choice was effective. Relapse was found to be a consistent predictor of

[3]In this research, treatment choices are of short-term medications, psychotherapy lasting 10 to 16 weeks, and maintenance therapy over 6 to 9 months.

long-term work impairment. When considering treatment, Claxton et al. (1999) showed that absenteeism increased prior to antidepressant use, and decreased after treatment began.

Many aspects of the relationships between depression, labor outcomes, and treatment decisions have been addressed in the existing literature. Most studies are found in medical or policy literature and fail to control for many factors. Many separate their sample into two groups, a treatment and a control group. Or, they just look at the changes in outcome variables without controlling for any characteristics, either observable or unobservable, other than depression.

In the economics literature, the majority of studies are based on cross section data and do not investigate the role of permanent unobserved heterogeneity on these outcomes. Longitudinal studies are rare, but do exist. However, they rarely incorporate treatment decisions.

I will extend this previous research by focusing on the relationships between depressive disorders, treatment decisions and multiple labor outcomes. The specific goals of the study are to: (i) estimate the impact of depression on labor outcomes while accounting for permanent unobserved heterogeneity in order to get unbiased estimates of depression's effects on labor outcomes, (ii) to incorporate treatment decisions into a model that uses different labor outcomes and varying measures of depression (including continuous, binary and categorical scales) and (iii) to use The Generalized Methods of Moments estimator and lags of endogenous regressors as instruments since no time-varying instruments are readily available.

Chapter 3

Data Description

3.1 ARTIST trial

Much effort was made to find a data set with varying health and employment measures that followed individuals over time in order to capture the effect of a potentially changing (either improving or worsening) depressive state on work behavior. Through special agreement with the pharmaceutical manufacturer Eli Lilly, I have obtained access to the ARTIST (A Randomized Trial Investigating SSRI Treatment) trial data. These data were collected over a 9-month period with enrollment occurring between April and November 1999. The participants were recruited from 37 clinical practices in two primary care research networks: the Primary Care Network (PCN)[1] and the Duke Primary Care Research Consortium (DPCRC).[2] Overall, 77 practitioners participated in the ARTIST trial, 51 from the PCN and 26 from the DPCRC. To be eligible to participate in the study, individuals had to be at least 18 years of age, receive their primary care from a physician participating in the PCN or DPCRC, have access to a home telephone, and be diagnosed with a depressive disorder. There were also exclusionary restrictions. Restrictions included: being actively suicidal, cognitive impairment, terminal illness, nursing home residence, taking an SSRI currently or within the last two months, taking a non- SSRI antidepressant at more than low doses, history of bipolar disorder, active cocaine or opiate user, and pregnant or breast feeding. The purpose of the

[1]The PCN is a not-for-profit voluntary organization and nationwide network of primary care practitioners whose goal is to optimize the care they provide by continuing education.

[2]The DPCRC is an academic site management organization that operates within the Duke Health Care System.

trial was to investigate the efficacy of specific SSRI treatments on depression. As such, the participants were randomly assigned particular drug treatments and followed over time. The data set has the unique feature that it follows individuals over time. The decision to begin antidepressant therapy was based solely on the primary care provider's (PCP) judgment that the patient's depression warranted antidepressant therapy.[3] Once the individual qualified for the study, a touch tone telephone procedure was used to randomly assign the patient to a specific antidepressant.[4] Both patients and PCP's were aware of the SSRI assignment. Decisions regarding treatment changes, including SSRI type, dosage, and discontinuation were allowed and jointly made by the patient and the PCP.[5] Patients received a pharmacy benefits card covering the costs for the SSRI and any non-SSRI antidepressant that the PCP prescribed during the study.[6] Study patients also received payments as reimbursement for their time. Participants received $20 for each completed telephone interview, with additional payments of $20 for 4 completed interviews and $30 for 5 completed interviews. A participant of all 5 interviews received a maximum of $150.

Computer assisted telephone interviews were used to assess outcomes. Interviewers did not divert from the interview format and were required to use preselected response options. Certain follow-up interviews were not complete. In order to get complete data, in the cases where individuals were difficult to reach, a prioritized list of questions was used, with the main outcome being asked first. The participants were considered still participating even if the interview was incomplete. They may have also been included in a later wave even if they were missing from a previous wave.

[3] The trial was set up to resemble real world practice and the practitioners did not receive additional training that could potentially change typical treatment patterns.

[4] The possible antidepressants were either 20 of mg paroxetine (Paxil), 20 mg of fluoxetine hydrochloride (Prozac), or 50 mg of sertraline (Zoloft).

[5] It is important to note that the decision to begin antidepressant therapy was based solely on the PCP's judgement that the patient's depression warranted antidepressant therapy. Using HSCL categories, approximately 6.67% of the sample individuals are categorized in the 'depression in remission' group at baseline.

[6] Providing the card minimized differences in patient compliance that could be due to socioeconomic factors. Antidepressant drugs covered by the card included: Amitriptyline chlordiazepoxide, Amitriptyline hydrochloride, Amitriptyline hydrochloride/perphenazine, Amoxapine, Bupropion hydrochloride, Citalopram hydrobromide, Clomipramine hydrochloride, Desipramine hydrochloride, Doxepin hydrochloride, Fluoxetine hydrochloride, Fluvoxamine maleate, Imipramine hydrochloride, Imipramine pamoate, Maprotiline hydrochloride, Mirtazapine, Nefazodone hydrochloride, Nortriptyline hydrochloride, Paroxetine hydrochloride, Phenelzine sulfate, Protriptyline hydrochloride, Sertraline hydrochloride, Tranylcypromine sulfate, Trazodone hydrochloride, Trimipramine maleate, and Venlafaxine hydrochloride.

These data have the limitation of following only those initially diagnosed as depressed, making selection an issue. Additional limitations include the patient self-reporting on therapy compliance and the fact that the treating physicians were not privileged to outcome information from the telephone interviews.[7] 601 patients provided informed consent and were randomized to treatment. 573 of these individuals completed the baseline telephone assessment. The survey consists of a baseline interview and 4 follow-up interviews in varying time intervals. Follow-up phone interviews were successfully completed for 79% of the participants at 9 months. Table 3.1 below details participation and follow-up of the original sample.

Table 3.1: Number of Patients at each interview

Baseline Interview	573 patients
At least 1 follow up interview	546 patients
1 month interview	538 patients
3 month interview	504 patients
6 month interview	483 patients
9 month interview	455 patients

Table 3.2 details availability of data in each interview, where 10111, for example, indicates that the participant missed the second interview. Note that I drop individuals aged 65 and older due to retirement possibilities, those with ages 18-22 due to schooling possibilities and one individual with no age information from the sample. Other individuals with missing employment information were also dropped from the sample. Statistics reported from here on are based on the 315 individuals with complete information for all interviews. These data provide 1575 observations for analysis.

3.2 Health Variables and Scales

It is important to consider different issues in the measurement of mental illness. In the mental health and labor literature, different strategies are used. One common approach is the

[7] This was to resemble real work practice.

Table 3.2: Individual Participation Patterns over the 5 interviews

Pattern	Longest Spell	Number of Individuals
01110	3	1
01111	4	2
10000	1	24
10100	1	2
10101	1	1
10110	2	1
10111	3	3
11000	2	20
11010	2	5
11011	2	12
11100	3	15
11101	3	15
11110	4	34
11111	5	315
total		450

utilization-based approach (Frank and Gertler, 1991). With this method, the individuals are asked if they have ever been treated for a mental disorder. If they answer yes, they are asked to identify the diagnosed disorder. With another approach, individuals are asked varying questions about their mental health status. They characterize each element as excellent, good, fair or poor. A more recent approach utilizes symptoms and clinical algorithms to determine diagnoses. Individuals respond to questionnaires that identify clinical symptoms. These symptoms either directly indicate a mental health problem or are used in algorithms to determine a diagnosis. Specifically, the diagnosis of major depressive disorder is based on the patient's self-reported experiences, behavior reported by relatives or friends, and a mental status exam. There is no laboratory test for major depression.

The ARTIST trial consists of multiple scales that can be used to measure depression and/or physical health: The Depression Diagnosis Interview Questions (DDI), The Hopkins Symptom Checklist (HSCL), The Physical Symptoms Scale (PSS), The 9-item Patient Health Questionnaire (PHQ-9), The 36-item Short-Form Health Survey (SF-36) and The 12-item

Short-Form Health Survey (SF-12).[8] Each scale has specific benefits and shortcomings. There is an SF-12 score for each interview. SF-36 is only used at baseline and in months 3 and 9. The DDI questions are also only asked at baseline, and in months 3 and 9. Depression outcome is assessed with two measures of core depressive symptoms, the HSCL-20[9] and SF-12 mental health component scores. I use both the HSCL-20 mental health component score and SF-12 mental health component scores as measures of depression in this paper. Both are validated measures of depression severity.

Depression severity is often defined as mild, moderate or severe. The levels are described in terms of the extent to which the patient's everyday life is affected: mild depression causes only minor impairment of the patient's work, social life and relationships with others. Major depression, as defined in the DSM-IV, can be of mild severity. Moderate depression is associated with more obvious symptoms and is more likely to be noticeable to others. Severe depression is characterized by affecting the patient so badly that he or she may be unable to work or to relate socially to others. Full remission is defined as the absence of symptoms for at least two months. For partial remission, full criteria for a major depressive episode are no longer met, or there are no substantial symptoms, but two months have not yet passed. In this study, I do not distinguish between full and partial remission.

HCSL-20

The Hopkins Symptom Checklist (HSCL) is a self report symptom inventory originally comprised of 58 items representing symptoms commonly observed in outpatients. There are five underlying symptom dimensions: somatization, obsessive-compulsive, interpersonal sensitivity, anxiety and depression. The original 58-item symptom inventory expanded to incorporate 32 additional items in four symptom dimensions: hostility, phobic anxiety, paranoid ideation and psychoticism. The HSCL-20 is a 20-item modified subscale of the 90-item Hopkins Symptom Checklist. The scores range from 0-4 with a higher score indicating a more

[8] SF-36 and SF-12 are versions of the same scale.

[9] The HSCL-20 is a 20-item modified subscale of the 90-item Hopkins Symptom Checklist. The 20-item scale includes the full 13-item depression subscale of these longer instruments plus 7 additional items that allow for an assessment of all Diagnostic and Statistical Manual, fourth edition (DSM-IV) items. Using these scales, an HSCL-20 mean score is calculated for each individual in each time period.

serious depressive episode. This continuous variable will be used in all estimations. To serve as a frame of reference, the scores are divided into 4 categories, each representing a different degree of depression.

$$HSCL_t = \begin{cases} 1 & \text{depression in remission:} & 0.00 \leq HSCL_t < 0.75 \\ 2 & \text{mild depression:} & 0.75 \leq HSCL_t < 1.50 \\ 3 & \text{moderate depression:} & 1.50 \leq HSCL_t < 2.00 \\ 4 & \text{severe depression:} & 2.00 \leq HSCL_t \leq 4.00 \end{cases} \tag{3.1}$$

Table 3.3: Summary of Depression using HSCL-20 Score

Month	Obs	Mean	SD	Min	Max
all	1575	1.05	0.73	0	3.20
Baseline	315	1.68	0.66	0.2	3.15
1	315	1.07	0.64	0	3.20
3	315	0.90	0.66	0	3.20
6	315	0.84	0.66	0	3.20
9	315	0.76	0.63	0	2.85

As seen in Table 3.3, the HSCL-20 mean measure continues to decrease throughout the ARTIST trial, meaning that the patients' general depression levels continuously improve. It decreases at a decreasing rate until the last period with the largest improvement occurring in the first period with a 36.30% change.

The trial follows individuals over time, and it is important to consider the breakdown of the variation in the depression score into between-individual variation and within-individual variation. I report similar statistics for the remainder of the outcome variables. Refer to Table 3.4. Overall, 39.62% of the person-year observations are considered to be in remission and 13.90% are categorized as having severe depression. Taking patients individually, 71.11% of the individuals are in remission at some point, with 39.68% considered to be severely depressed at least once; thus some individuals are characterized as having severe depressive symptoms at one interview but not at others. Taking one individual at a time, if an individual is considered to have major depression at least once, 35.04% of her observations fall into the severe depression category. Similarly, if the patient is ever categorized to be in remission, 55.71% of her observations are considered to be in remission. If an individual's depressive status never varied over time in these data, the within individual variation percentages would all be equal to 100. With a continuous variable, the mean and decomposed standard deviation are reported. The standard deviation is decomposed into between and within components. If the within standard deviation is equal to 0, all of the variation is between individuals. The 0.50 for the within variation confirms that there is variation within the individual.

Table 3.4: Variation in HSCL-20 Mental Health Variable

Variable		Mean	SD	Min	Max	Obs
HSCL-20 Score						
	overall	1.05	0.73	00.00	3.20	N = 1575
	between		0.53	00.09	2.78	n = 315
	within		0.50	-0.54	3.11	T-bar = 5
		Overall(%)	Between (%)	Within(%)		
HSCL-20 category (n=315)						
	in remission	39.62	71.11	55.71		
	mild depression	32.57	78.73	41.37		
	moderate depression	13.90	46.35	30.00		
	severe depression	13.90	39.68	35.04		

SF-36 and SF-12

The 36-item Short-form Health Survey (SF-36) measures health-related quality of life in eight domains, including physical functioning, social functioning, mental health, general health perception, pain, vitality, and physical and emotional role functioning. Selected measures from the Medical Outcomes Study (MOS) are also used to assess social functioning, concentration, positive well-being, hopefulness, sleep, and sexual function. Anxiety and alcohol disorder items are taken from the PHQ-9. Additionally, questionnaires were used to evaluate quality of close relationships and disposition.

The 36-item short form was constructed to survey health status in the Medical Outcomes Study (MOS), which was a part of the phone interviews at baseline, in month 3 and in month 9. The SF-36 was designed for use in clinical practice and research, health policy evaluations, and general population surveys. It includes one multi-item scale that assesses aspects of both the individual's physical and mental health: limitations in physical activities because of health problems, limitations in social activities because of physical or emotional problems, limitations in usual role activities because of physical health problems, bodily pain, general mental health (psychological distress and well being), limitations in usual role activity because of emotional problems, vitality (energy and fatigue), and general health perceptions. Since this study is primarily concerned with questions surrounding mental health issues, I will describe the questions that pertain to mental health. The 14 questions, out of 36 total questions, affecting mental health fall into one of four scales (Vitality, Social Functioning, Role- Emotional, Mental Health). The questions, scales and response alternatives that affect the SF-36 score are found in Appendix A. The scores range from 0 to 100, with higher scores indicating better mental health. As summarized in *The SF-36 physical and mental health summary scales: A user's manual* the best cut off for defining depression using the SF-36 mental health component score is at a score of 42.00 or below. This score minimizes the possibility of either a false negative or a false positive. The false positive rate using this SF-36 mental component scale cutpoint is 19.40% and the false negative rate is 26.30%.

The SF-12 was developed to be a much shorter, yet useful, alternative to the SF-36. Only one to two questions are asked for each of the four scales (Vitality, Role Emotional, Social Functioning, and Mental Health) rather than the three to four questions in the SF-36. The choice between the 36-item or 12-item versions is largely practical and each has advantages. The SF-12 is much shorter while reproducing the SF-36 summary scales very well. It has the advantage of maximizing participant retention. The SF-36 profile has the advantage of providing more information about the nature of differences in physical and mental health outcomes. The SF-12 also uses a composite score of 42.0 as the cutoff score for defining depression. The SF-12 interview is calculated in each month of the ARTIST trial. At baseline, and in months 3 and 9, there are SF-12 and SF-36 scores for each individual.

In months 1 and 6, only twelve questions are asked so there is only a SF-12 score for each individual interviewed in these months. The SF-12 scores will be used to divide the sample into nondepressed and depressed individuals.

Table 3.5: Summary of Depression using SF-12 Score

Month	Obs	Mean	SD	Min	Max
all	1575	44.04	12.06	4.45	65.83
Baseline	315	32.48	10.91	4.45	61.68
1	315	43.83	10.63	9.10	65.84
3	315	47.13	10.86	9.82	65.20
6	315	47.49	10.01	13.68	64.11
9	315	49.26	9.79	16.81	64.91

Similarly to when using the HSCL-20 component score, the SF-12 measure of mental health continuously improves throughout the ARTIST trial as reported in Table 3.5, with a higher score indicating better mental health. Patients experienced improvement each period, with the largest improvement occurring in the first period with a 34.94% change. In Table 3.6, it can be seen that there is within-individual variation for the SF-12 mental health measure variation (SD=8.98).

Table 3.6: Variation in SF-12 Mental Health Variable

Variable	Mean	SD	Min	Max	Obs
SF-12 Score					
overall	44.04	12.06	4.45	65.83	N = 1575
between		8.06	15.26	61.78	n = 315
within		8.98	9.60	72.64	T-bar = 5

The SF-12 also includes questions related to physical health. The Physical Symptoms Scale (PCS-12) is a scale consisting questions regarding the patient's physical health. The scores range from 0 to 100, with higher scores indicating better physical health. These questions are asked at each interview. See Appendix B for the list of questions.

It can be assumed that any chronic condition will affect all decisions and outcomes in the model. As a measure of medical comorbidity, I will use the Chronic Disease Score (CDS) which is calculated at baseline for each patient. The CDS score increases with the number of different chronic diseases as inferred from the subject's medication profile. At baseline, each patient is asked to list up to 14 prescription medications she is currently taking. With these responses, the CDS score is calculated and used as a comorbidity index.[10] Scores have been shown to predict mortality and health care resource utilization after controlling for demographics and previous resource utilization.[11] As a measure of chronic depression, at baseline, participants are asked if they had ever been treated for depression prior to the ARTIST study. 34.60% of the individuals had experienced a depressive episode prior to the ARTIST study. Baseline characteristics of these individuals are provided in Table 3.7.

Table 3.7: Baseline Health Characteristics of Sample

Baseline Characteristic	Mean	SD
Mental Health HSCL-20 score	1.68	0 .66
Mental Health HSCL-20 category (%)		
in remission	6.67	
mild depression	32.06	
moderate depression	25.71	
severe depression	35.56	
Mental Health SF-12 score	32.48	10.91
Physical Health SF-12 score	49.94	9.79
Chronic Disease Score	1.62	2.08
Past history of depression (%)	34.60	

3.3 Labor Variables

The indirect costs related to depression can come in the form of presenteeism or absenteeism. The probability of being employed is also affected. This study uses three categories of labor variables: participation variables (whether or not the individual is employed); absen-

[10]Specific medications are assigned to medication classes, which are then mapped to different chronic diseases.
[11]The CDS score employs empirical weights.

teeism variables (conditional on being employed, how often the worker is absent from work); and presenteeism variables (conditional on employment, the worker's productivity while at work).

Work Limitations Questionnaire-Presenteeism and Absenteeism Variables

The Work Limitations Questionnaire (WLQ) labor variables are only used if the individual answers yes to the following question: In the past two weeks did you work at anytime at a job or business not counting work around the house?[12]

The objective of the WLQ was to develop a psychometrically sound questionnaire for measuring the impact of chronic health problems and/or treatment on different job aspects of employed workers. The WLQ is a self-reported instrument for measuring the degree to which chronic health problems interfere with a workers' ability to perform job roles with four distinct dimensions: handling time, physical, mental-interpersonal, and output demands. With 25 items, and a 2-week reporting period, the Work Limitations Questionnaire demonstrates high reliability and validity. Unlike other questionnaires it addresses aspects of the job through a demand-level methodology. In comparison to other questionnaires, the WLQ Output Demands scale, which I use to measure presenteeism, has superior performance for predicting productivity. The Mental Interpersonal Demands and the SF-36 Role Limitation scales, which are not used in this study, have moderate validity.[13] A full listing of the questions can be found in Appendix C.

Tables 3.8 and 3.9 report between-individual and within-individual variation for the labor variables. In Table 3.8 it can be seen that there is within-individual variation in both the employment probability measure and the ability to meet the required number of work hours. Table 3.9 reports means and standard deviations of the productivity measure for the overall sample, between-individual and within-individual. The within statistics show the variation of the ability to meet output demands within person around the global mean 74.62. The within-individual SD shows that there is also within- individual variation in the productivity

[12] The wording of this question limits its use as an indicator of employment. An individual may still be employed if he/she has not worked in the last two weeks due to vacation, maternity leave, or an illness.

[13] For a full discussion of validity tests, see Lerner et al., 2002.

measure.

Table 3.8: Variation in Labor Variables

	Overall(%)	Between (%)	Within(%)
Employment Probability (n=315)			
does work	74.22	86.98	85.33
does not work	25.78	46.98	54.86
Ability to Meet Required Work Hours (n=274)			
All of the time	56.29	86.50	62.91
Most of the time	20.19	54.38	35.33
Some of the time	16.42	44.53	34.85
A slight bit of the time	6.07	19.34	30.74
None of the time	1.03	4.38	24.00

Table 3.9: Variation in Productivity Measure

Variable	Mean	SD	Min	Max	Obs
Productivity					
overall	74.62	19.46	5	100	N=1169
between		15.39	26	100	n=274
within		12.12	26.18	115.62	T=4.26

Table 3.10 presents descriptives statistics for each labor outcome by depressive state. As expected the less depressed samples perform better for all labor variables. Employment probabilities and depression levels show an inverse relationship, with 74.22% of the remission sample being employed and only 67.00% of the sample categorized as having severe depression working outside of the home. As depression levels decrease, the probability of working continuously increases. Less depressed workers are also better able to work the required number of work hours. Again we see that the largest difference can be found between the depression in remission sample and the mild depression group. The productivity measure also shows less desired outcomes for more depressed individuals. It can be seen that the difference in

productivity between the remission sample and those classified as having major depression is larger than when comparing the sample's employment probabilities and absenteeism.

Table 3.10: Summary of Labor Variables by HSCL-20 Mental Health Category

Variable	n	Mean	SD	Min	Max
Employment Probability					
full sample	1575	0.74	0.43	0	1
severe depression	219	0.67	0.47	0	1
moderate depression	219	0.72	0.44	0	1
mild depression	513	0.73	0.44	0	1
in remission	624	0.78	0.41	0	1
Ability to Meet Required Hours at Work					
full sample	1169	4.24	1.00	1	5
severe depression	147	3.26	1.14	1	5
moderate depression	159	3.74	1.05	1	5
mild depression	376	4.22	0.92	1	5
in remission	487	4.72	0.61	1	5
Productivity					
full sample	1169	74.62	19.46	5	100
severe depression	147	51.79	19.11	5	100
moderate depression	159	62.62	18.32	10	100
mild depression	376	73.11	16.25	5	100
in remission	487	86.60	11.73	35	100

3.4 Treatment Variables

At baseline, a touch tone telephone procedure was used to randomly assign the patient to a specific antidepressant: either Prozac, Zoloft or Paxil. A variable to indicate initial assignment is created for each individual. Prior to randomization, the PCP is asked what antidepressant he would prescribe in the absence of randomization. Responses included Celexa, Paxil, Prozac, Wellbutrin, Zoloft or NA.[14] These variables are summarized in Table 3.11.

A variable indicating if the same drug the PCP would have prescribed was randomized

[14] N/A indicates that the PCP's response was not Celexa, Paxil, Prozac, Wellbutrin or Zoloft.

to the patient at baseline is defined for each individual.[15] Both patients and PCP's were aware of the baseline SSRI assignment. Decisions regarding treatment changes, including SSRI type, dosage, and discontinuation were allowed and jointly made by the patient and the PCP. At each interview, the individual is asked about her treatment decisions. She may choose between continuing on the current drug, immediately switching to a new drug or discontinuing use of the current drug without switching to an alternate medication. If she chooses to change medication type, she is asked what type of SSRI she is currently taking. At the next interview, she is asked about her treatment decisions with respect to the new drug. 61% of the patients remain on the randomized drug throughout the trial. 20% of the sample eventually stops the randomly assigned medication and does not immediately switch to a new drug within the survey period, while 19% change SSRI medication and immediately switch to a different SSRI at least once.

Table 3.11: Treatment Variables

Treatment Variable		%
Randomized Drug		
	Prozac	32.38
	Paxil	32.38
	Zoloft	35.24
PCP Preferred Drug		
	Prozac	27.62
	Paxil	22.22
	Zoloft	26.98
	other	23.18
Same Drug Randomized & Prescribed		33.33

[15]If the PCP would have prescribed Celexa, Wellbutrin or another antidepressant, this value is automatically equal to 0. These common responses, Celexa and Wellbutrin, are each in another class of antidepressants (not SSRI).

3.5 Exogenous Variables

Exogenous variables include the individual's age, gender and race. Baseline characteristics of the sample are found in Table 3.12. Additional time-varying information that may affect outcomes exogenously are available from other data sources. Although participants report five digit zip code information at each interview, Zip Code Tabulation Areas (ZCTAs) are used for extracting Census 2000 and Current Population Study (CPS) information on a geographic area from the reported zip code.[16] Monthly state unemployment rates from the Current Population Study (CPS) are also used in the estimations. 21 states and Washington D.C. are represented in the study. Table 3.13 describes the sample by state distribution. Current local economic conditions facing the individual are likely to affect all employment decisions. Therefore, they will be included in all labor equations.

Table 3.12: Summary of Baseline Individual Exogenous Characteristics

Baseline Characteristic	
Age	43.84
	(9.89)
Age Range	23-64
Women %	80.95
Race %	
White	84.13
Black	13.65
Other	2.22

The month of the year is also included as a time-varying explanatory variable. The calendar month of the year may affect both labor and health outcomes. Individuals may experience more mental and physical health problems during the winter months. It is also possible that labor outcomes may differ during specific times of the year. For example, a worker may be more likely to take vacation during the summer months or during the winter holiday time. Productivity may also be affected during specific times of the year depending on the worker's industry. Individuals enrolled in the study and completed baseline interviews

[16] See Appendix E for discussion of ZCTAs.

Table 3.13: Represented States

State	Person-months	% of sample
North Carolina	648	41.14
Georgia	160	10.16
Pennsylvania	130	8.25
Missouri	82	5.21
California	70	4.44
Maryland	47	2.98
New York	65	4.13
Ohio	70	4.44
Texas	50	3.17
Florida	45	2.86
Arizona	45	2.86
Iowa	40	2.54
Michigan	45	2.86
Illinois	18	1.14
Indiana	15	0.95
Virginia	15	0.95
Tennessee	10	0.63
Oklahoma	5	0.32
Washington D.C.	5	0.32
South Carolina	5	0.32
Wisconsin	5	0.32

during different months. On the interview date I create indicator variables=1 for each of the twelve months.

Chapter 4

Theoretical Motivation

This section includes the theoretical background to the paper. In Section 5, I describe an empirical approximation to the theoretical model that follows.

4.1 Overview of Decision Making Process

The individual enters a decision making period t knowing her employment history, current health status and her mental and physical health history. Future mental and physical health are uncertain. With this knowledge, she decides whether or not to work in period t. By accepting employment, she contracts for a certain number of hours of required work hours (full time or part time). The required work hours are unobserved by the econometrician. Conditional on working, the individual then decides the number of these required hours to work and how productive to be for the duration of period t. The individual must also decide whether to continue her current medication, switch to another medication or discontinue medication treatment entirely. The individual then realizes her health status at the end of each period t. The per-period timing of decisions and realizations of health are depicted below.

4.2 Decision Variables

Each period, individuals make optimal decisions with respect to employment and mental health treatment. At the beginning of time period t, the worker decides whether or not to

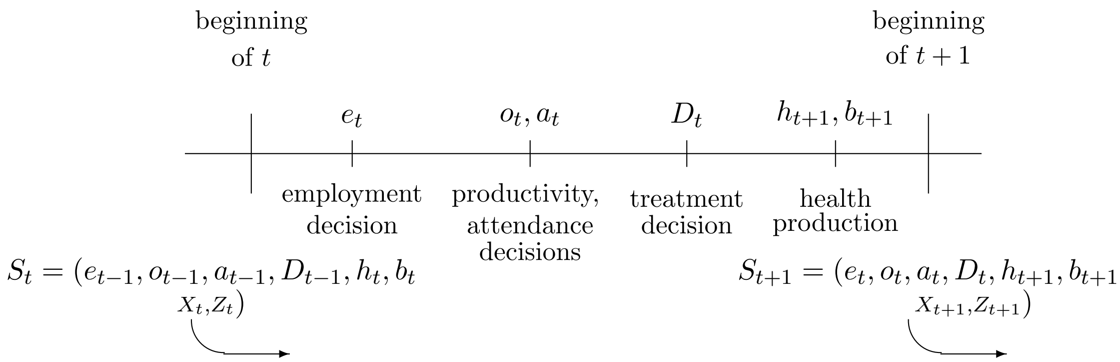

Figure 4.1: Timing of per-period Employment and Treatment Decisions and Health Production

work outside of the home (e_t) where

$$e_t = \begin{cases} 0 & \text{if individual does not work} \\ 1 & \text{if individual works.} \end{cases}$$

If she decides to accept outside employment, she chooses her level of productivity for the remainder of the period (o_t) and attendance (a_t). The productivity measure (o_t), a self reported ability to meet output demands, takes the value 0 to 100, with 100 indicating the highest productivity. With regard to absenteeism, the respondents answer the following question: "Conditional on working, how much of the time during the past two weeks did your physical health or emotional problems make it difficult for you to work the required number of hours?" Attendance (a_t) takes on the values (with a higher score indicating less limitation):

$$a_t = \begin{cases} 1 & \text{All of the time} \\ 2 & \text{Most of the time} \\ 3 & \text{Some of the time} \\ 4 & \text{A slight bit of the time} \\ 5 & \text{None of the time} \end{cases}$$

The individual makes the medication treatment decision (D_t). Treatment options include:

$$D_t = \begin{cases} 0 & \text{if patient continues on the same drug} \\ 1 & \text{if patient immediately switches from the current drug to another drug} \\ 2 & \text{if patient stops current drug, does not immediately switch to new drug} \end{cases}$$

When making decisions in period t, the individual may consider her employment status (e_{t-1}), her productivity (o_{t-1}), and her attendance at work (a_{t-1}) in the previous period. She also considers her wage offer which depends on the worker's employment history and past performance. The wage offer is unobserved by the researcher.

Decisions may also be influenced by the individual's health history including her treatment decisions in the previous period (D_{t-1}), mental health state entering the current period (h_t)[1] where

$$h_t = \begin{cases} 1 & \text{if depression in remission} \\ 2 & \text{if mildly depressed} \\ 3 & \text{if moderately depressed} \\ 4 & \text{if severe depression} \end{cases}$$

and physical health entering the current period (b_t) where b_t scores range from 0 to 100 with higher scores indicating better physical health.

Current economic conditions such as median income by zipcode and monthly state unemployment rates may also influence labor decisions and are denoted by the vector Z_t. All decisions are influenced by observed personal characteristics, X_t, including gender, race, and age. The individual's information, or state vector(S_t), entering period t is:

$$S_t = (e_{t-1}, o_{t-1}, a_{t-1}, D_{t-1}, h_t, b_t, Z_t, X_t). \tag{4.1}$$

[1] In the empirical estimation different linear and categorical scales for mental health are used.

4.3 Health Production

An important part of this study is understanding how mental health evolves. The mental health production function, or probability of a health transition from $h_t = h$ to $h_{t+1} = h'$, is modeled as

$$\begin{aligned} \pi_{t+1}^{hh'} &= Pr(h_{t+1} = h' \mid h_t = h) \;\; \forall t \\ &= \frac{\exp\{\pi_{0h'} + \pi_{1h'} h_t + \pi_{2h'} S_{t+1}\}}{\sum_{h''=1}^{4} \exp\{\pi_{0h''} + \pi_{1h''} h_t + \pi_{2h''} S_{t+1}\}} \end{aligned} . \tag{4.2}$$

Particular components of S_{t+1}, namely $(e_t, o_t, a_t, D_t, b_{t+1}, X_{t+1})$, may influence depression. Physical health (b_t) also develops based on decisions and observed characteristics where

$$b_{t+1} = b(b_t, h_t, S_{t+1}). \tag{4.3}$$

4.4 Optimization

An individual's per period utility is a function of consumption (C_t), mental health status (h_t), labor choices $(e_t$, a_t, $o_t)$ and treatment decisions (D_t). Let J equal the set of combinations of the e_t, a_t, o_t and D_t alternatives available to an individual. Each unique combination is denoted j, j=1,2,3...J. Health status is assumed to have both direct effects (i.e., healthier individuals derive more utility in all activities) and indirect effects (through its effect on labor and treatment decisions) on utility. Treatment decisions affect utility through the medication's effects on physical health (side effects) and have productive effects on mental health status. The individual's per-period alternative-specific utility function is given by

$$U_t^j = U(j, C_t, h_t, b_t, X_t, \varepsilon_t^j) \tag{4.4}$$

where ε_t^j represents alternate-specific and time-specific unobserved utility preferences and consumption (C_t) is constrained by her budget. Sources of income include employment income

(Y_t) if employed $(e_t = 1)$ and unearned income (W_t) such that

$$C_t = e_t * Y_t + W_t. \tag{4.5}$$

Employment income, $Y_t = Y(X_t, e_{t-1}, a_{t-1}, o_{t-1})$, is a function of personal characteristics, and her employment history, days present at work and productivity in period $t-1$. If $e_t = 0$, consumption is constrained by unearned income. Treatment for depression is costless in these data in pecuniary terms but may cause disutility through side effects, discomfort or stigma.

An individual is assumed to make decisions that will maximize her expected lifetime utility. This includes the current utility of each alternative plus the discounted present value of expected future utility where health and future utility shocks are uncertain. Her lifetime value of each employment and treatment alternative is:

$$V_j(S_t, \epsilon_t) = U_t^j + \sum_{h'=1}^{4} \pi_{t+1}^{hh'} E_t \left[\max_{k \in J} V_k(S_{t+1}, \epsilon_{t+1}^k) \mid j \right] \Big\} \;\; \forall j \;\; \forall t \tag{4.6}$$

where β is the discount factor. Since this research focuses on mental health and its effects on labor market outcomes, physical health transitions are not modeled explicitly in the value function. While physical health develops overtime, the value function assumes as written that physical health is exogenous and known with perfect foresight.

Chapter 5

Empirical Model

In measuring the effect of mental health on labor outcomes, many studies use the human capital framework (Becker 1967, Grossman 1972) and assume people value good health for both its direct effect on utility and its effect on human capital formulation. This paper focuses on the relationship between depression and different labor market outcomes. Underlying the analysis is the idea that for a group of depressed individuals, changes in depression levels lead to changes in labor market outcomes. Additionally, treatment decisions affect health status. In measuring the effects of depression on productivity and attendance for a group of initially depressed individuals, there are multiple sources of bias that must be addressed in order to obtain consistent estimates.

Two sources of endogeneity can lead to biased estimates of depression's effect on labor market outcomes. Structural endogeneity results from possible reverse causality. In this case, the onset of depression may occur due to labor market outcomes. Statistical endogeneity comes from potential unobserved factors causing depression that may also affect labor market outcomes independent of depression. For example, unobserved personal characteristics, such as excessive motivation or employment in high stress occupations may increase the risk of depression and can lead to different labor outcomes. Including treatment decisions variables explaining mental health transitions introduces additional potential bias in the measured effect. Potential endogeneity results from permanent unobserved characteristics that affect both the individual's mental health status and treatment decisions.

In this dynamic model, the labor outcomes L_t are the individual's log productivity at work o_t and attendance a_t, ($L_{it} = [o_{it}, a_{it}]$) if employed. The relationship is dynamic because an individual's labor outcomes depend on lagged endogenous variables. Let

$$e_{it} = \begin{cases} 1 & \text{if individual works in period } t \\ 0 & \text{if individual does not work in period } t. \end{cases}$$

The work productivity (o_{it}) and work attendance (a_{it}) measures are only observed if the individual works in period t ($e_{it} = 1$). This dynamic model of labor outcomes includes a lagged dependent variable as a regressor as follows:

$$L_{it} = \delta L_{i,t-1} + \alpha T_i + \beta X_{it} + \varepsilon_{it} \text{ if } e_{it} = 1. \tag{5.1}$$

An individual's labor outcome L_{it} is related to time-invariant personal characteristics T_i[1], a vector of time-varying observables for the individual, X_{it}, and a disturbance term, ε_{it}. The error component ε_{it} is decomposed as

$$\varepsilon_{it} = \mu_i + \nu_{it} \tag{5.2}$$

where μ_i represents unobserved time invariant personal characteristics that will affect labor outcomes (such as genetics) and ν_{it} accounts for random individual disturbances (such as time-varying work demands). The vector of disturbances ν_{it} is assumed to be independent and identically distributed (IID) $IID(0, \sigma_\nu^2)$ and μ_i is distributed $IID(0, \sigma_\mu^2)$. Substituting equation (5.2) into equation (5.1),

$$L_{it} = \delta L_{i,t-1} + \alpha T_i + \beta X_{it} + \mu_i + \nu_{it} \text{ if } e_{it} = 1. \tag{5.3}$$

The binary choice selection model explaining whether an individual is working is described by

$$e_{it}^* = \omega \Gamma_{it} + \eta_i + \zeta_{it}. \tag{5.4}$$

[1] T_i is not included in the theoretical model in the previous section. I separate this to clearly show the effects of purging the model of the time-invariant individual effect.

where

$$e_{it} = \begin{cases} 1 & \text{if } e_{it}^* > 0 \\ 0 & \text{if } e_{it}^* \leq 0 \end{cases}$$

Γ_{it} is a vector of explanatory variables that may share elements with X_{it}, η_i is an unobservable time-invariant individual-specific effect and ζ_{it} represents unobserved individual time-specific disturbances. It is assumed that the error components in the two equations have a joint normal distribution. This selection equation must be estimated jointly with the labor outcomes in order to eliminate selection bias.

Dynamic panel bias and selection bias are additional sources of potential bias. Introducing a lagged dependent variable complicates estimation when compared to a static model because labor outcomes in past periods may also be correlated with unobserved permanent individual heterogeneity, $E(L_{i,t-1}, \mu_i) \neq 0$. Sample selection bias refers to problems where the dependent variable is observed only for a restricted, nonrandom sample. In this case, the individuals productivity in period t is only observed if the individual is employed in that period. Similarly to the productivity equation (the equation of interest), the employment equation (the selection equation) may contain individual specific effects that are correlated with the explanatory variables. To obtain consistent estimates, all potential sources of bias must be addressed.

In this study, the employment of an individual depends on unobserved time-invariant characteristics of the individual such as motivation, ability or education level, $E(e_{it}\mu_i) \neq 0$. These characteristics are correlated with the unobserved individual effects from the productivity and attendance equations.

If the potential sources of bias are due to time-invariant characteristics, a researcher can time difference the data then use instrumental variables (IV) in a Generalized Method of Moments (GMM) framework. The First-Differenced (FD) equation is:

$$(L_{it} - L_{i,t-1}) = \delta(L_{i,t-1} - L_{i,t-2}) + \beta(X_{it} - X_{i,t-1}) + \nu_{it} - \nu_{i,t-1}. \tag{5.5}$$

This is equivalent to

$$\Delta L_{it} = \delta \Delta L_{i,t-1} + \beta \Delta X_{it} + \Delta \nu_{it}. \tag{5.6}$$

If the selection process is time constant, removing the fixed effect from productivity or attendance equations eliminates any potential selection problem operating through μ_i. Now, selection into employment is random when unrelated to the idiosyncratic errors ν_{it}, (i.e., $E(e_{it}\nu_{it}) = 0$). Since $T = 5$ and the entire panel covers 9 months, this is a realistic assumption. For consistency, it is required that

$$E[\Delta \nu_{it} | X_{it}, X_{is}, e_{it} = e_{is} = 1] = 0 \;\; \forall \;\; s < t. \tag{5.7}$$

First-differencing has the advantage that it eliminates the individual specific effects that are correlated with both the explanatory variables in the labor equation and the selection equation. However, it does have an important weakness when an individual is not employed in period t, $e_{it} \neq 1 \;\; \forall \;\; t$, since there is a gap in the panel.[2] In these instances, there are no productivity measures for these observations and with $T = 5$, many needed observations are lost. A second transformation that will eliminate fixed effects without losing as many observations is called forward orthogonal deviations (FOD) (Arellano and Bover, 1995). Rather than taking first-differences as in equation (5.5), it subtracts the average of all future available observations of a variable. Therefore, despite the number of gaps, it can be computed for all observations of L_{it} and explanatory variables except for $t = 5$. The orthogonal deviations transform for any variable w is:

$$w_{i,t+1} \equiv c_{it}\left(w_{it} - \frac{1}{T_{it}} \sum_{s>t} w_{is} \right) \tag{5.8}$$

where T_{it} is the number of all future available observations. The scale factor to equalize the variance of the transformed error, c_{it}, equals $\sqrt{T_{it}/(T_{it}+1)}$. With this transformation, the disturbance term becomes

$$\nu^*_{i,t+1} \equiv c_{it}\left(\nu_{it} - \frac{1}{T_{it}} \sum_{s>t} \nu_{is} \right). \tag{5.9}$$

[2] When some value of L_{it} is missing, then both $\Delta L_{i,t+1}$ and ΔL_{it} are missing when first-differenced.

Again, individual specific effects are removed. Additionally, Arellano and Honore (2001) note that using orthogonal deviations does not introduce a moving average process into the disturbance. The time-invariant individual effect is purged, but $\nu_{i,t-1}$ is now correlated with $L_{i,t-1}$, $E(L_{i,t-1}, \nu_{i,t-1}) \neq 0$ and any other lagged endogenous variables. Thus, Ordinary Least Squares (OLS) produces inconsistent estimates. However, instruments for $\Delta L_{i,t-1}$ and other lagged endogenous covariates can be used since they are uncorrelated with the error term.

Estimation

For comparison purposes, I first estimate the equation of interest using ordinary least squares. I pool the data and estimate an OLS regression,

$$L_{it} = \delta L_{i,t-1} + \alpha T_i + \beta X_{it} + \varepsilon_{it} \text{ if } e_{it} = 1. \tag{5.10}$$

OLS estimates are biased due to dynamic panel bias, sample selection bias and unobserved heterogeneity. Standard results indicate that the OLS levels estimator is biased upwards by attributing unobserved characteristics in the individual's fixed effect to endogenous covariates.

In the model, it is probable that both the time-varying and time-invariant explanatory variables are correlated with an individual's unobserved characteristics. It is realistic to suspect that mental health, the variable of primary interest, is endogenous, $E(X_{it}\mu_i) \neq 0$. A time-invariant observed characteristic, such as whether or not the individual had ever been treated for depression before the study, is likely endogenous as well, $E(T_i\mu_i \neq 0)$.

The fixed effects (FE) approach allows for the fact that an error term for a specific individual will be correlated over time. Fixed effects estimates are also included for comparison purposes. As μ_i is constant over time within-individuals, differing only between units, this approach eliminates μ_i from the equation by measuring the deviation of each variable from the within-individual mean of the variables. Time-constant unobserved heterogeneity is no longer a problem. Averaging equation (5.3) over time gives

$$\bar{L}_i = \delta \bar{L}_{i,t-1} + \alpha T_i + \beta \bar{X}_i + \mu_i + \bar{\nu}_i. \tag{5.11}$$

Subtracting equation (5.11) from equation (5.3), μ_i, the individual error component, and T_i, the vector of constant explanatory variables, are purged from the equation leaving

$$(L_{it} - \bar{L}_i) = \delta(L_{i,t-1} - \bar{L}_i) + \beta(X_{it} - \bar{X}_i) + \nu_{it} - \bar{\nu}_i. \tag{5.12}$$

After removing the individual effect from the equation, applying the standard least squares estimator to the deviations gives the covariance estimator of β. Intuitively, fixed effects estimators are assuming that each individual has her own intercept (individuals have different time-invariant characteristics) but each individual's regression between labor outcome and regressor will have the same slope. In this study, it is assuming that each person is different, but a change in mental health status will affect each individual's labor outcomes the same. The fixed effects estimator removes the bias resulting from unobserved heterogeneity. However, it does not correct for dynamic panel bias. The estimator proves to be biased downwards in standard results. Bond (2002) points out that good estimates of δ should lie in the range between the OLS and the FE estimates, thus providing a useful check on results.

The lagged labor variable $L_{i,t-1}$ and other variables in X_{it} may be endogenous. To correct for the dynamic panel bias, instruments for $\Delta L_{i,t-1}$ and other lagged endogenous covariates must be used since they are uncorrelated with the error term. If no exogenous time-varying instruments are readily available, lags of endogenous regressors can be used to instrument $\Delta L_{i,t-1}$ and ΔX_{it} in equation (5.6). Anderson and Hsiao (1981, 1982) note that $\Delta L_{i,t-2}$ and $L_{i,t-2}$ are correlated with $\Delta L_{i,t-1}$, but not with $\Delta\nu_{it}$, making it a valid instrument for $\Delta L_{i,t-1}$. They propose two instrumental variables estimators of δ that are consistent:

$$\hat{\delta}_{IV1} = \frac{\sum_{i=1}^{N}\sum_{t=1}^{T}(L_{it} - L_{i,t-1})(L_{i,t-2} - L_{i,t-3})}{\sum_{i=1}^{N}\sum_{t=1}^{T}(L_{i,t-1} - L_{i,t-2})(L_{i,t-2} - L_{i,t-3})} \tag{5.13}$$

and

$$\hat{\delta}_{IV2} = \frac{\sum_{i=1}^{N}\sum_{t=1}^{T}(L_{it} - L_{i,t-1})L_{i,t-2}}{\sum_{i=1}^{N}\sum_{t=1}^{T}(L_{i,t-1} - L_{i,t-2})L_{i,t-2}}. \tag{5.14}$$

The first (equation 5.13) uses $\Delta L_{i,t-2}$ as the instrument and the second (equation 5.14) uses $L_{i,t-2}$. Similarly, $\Delta X_{i,t-2}$ or $X_{i,t-2}$ can be instruments for any additional endogenous explanatory variables since $\Delta X_{i,t-2}$ and $X_{i,t-2}$ are correlated with $\Delta X_{i,t-1}$, but not with $\Delta\nu_{it}$.

Arellano and Bond (1991) report that Anderson-Hsiao estimators are inefficient because they do not utilize all available instruments. The Generalized Method of Moments (GMM) estimator uses additional moment restrictions, based on the orthogonality between lagged values of the dependent variable L_{it} and the errors ν_{it}, thus enlarging the set of instruments. It is a general estimator designed for dynamic models with the following characteristics:

1. Many individuals and few time periods ("small T, large N"). If T is large, dynamic panel bias is no longer a significant source of bias.
2. a linear functional relationship
3. a dependent variable that is dynamic
4. independent variables that are not strictly exogenous
5. fixed individual effects
6. heteroskedasticity and autocorrelation within individuals, but not across them.

Arellano and Bond estimators are based on the assumptions that the disturbance term is mean zero (independent across individuals, and not serially correlated, $E(\nu_{it}) = 0$ and $E(\nu_{it}\nu_{is}) = 0$). Under these assumptions, the transformed residuals from equation (5.6) are no longer correlated, for all L_{it} and X_{it} dated $t-2$ and earlier. Now, more instruments are available to rid the model of the correlation between $\Delta L_{i,t-1}$, ΔX_{it} and $\Delta\nu_{it}$ in equation (5.6). With $T = 5$, the transformed residuals satisfy the following moment conditions (with Z indicating the instruments):

$$E(Z'_{it}\Delta\nu_{it}) = 0 \;\; \forall\, t = 2, 3, 4, 5. \tag{5.15}$$

For each individual, a system of five equations is denoted

$$L_i = X_i\Theta + \nu_i$$

where

$$L_i = \begin{bmatrix} \Delta L_{i3} \\ \Delta L_{i4} \\ \Delta L_{i5} \end{bmatrix}, \quad X_i = \begin{bmatrix} \Delta L_{i2} & \Delta X_{i3} \\ \Delta L_{i3} & \Delta X_{i4} \\ \Delta L_{i4} & \Delta X_{i5} \end{bmatrix}, \quad \text{and } \nu_i = \begin{bmatrix} \Delta \nu_{i3} \\ \Delta \nu_{i4} \\ \Delta \nu_{i5} \end{bmatrix}.$$

The set of instruments is

$$Z_i = \begin{pmatrix} Z_{i3} & 0 & 0 \\ 0 & Z_{i4} & 0 \\ 0 & 0 & Z_{i5} \end{pmatrix}$$

where $Z_{it} = (L_{i,t-2}, X_{i,t-2}, L_{i,t-3}, X_{i,t-3}, ..., L_{i1}, X_{i1})'$ at period t. Rows of the instrument matrix correspond to the transformed (either by first-differencing or using forward orthogonal transformation) equations for periods $t = 3, 4, 5$ for individual i with moment conditions in equation (5.15). It can be seen that the number of instruments increases as t increases.[3]

The efficient instrument martrix Z_{it} includes different moment conditions depending upon whether the explanatory variables are correlated with the fixed effects or not, and whether they are endogenous, predetermined or exogenous.

X_{it} may include variables that are considered endogenous in the sense that X_{it} are correlated with the error term ν_{it} and earlier shocks but uncorrelated with $\nu_{i,t+1}$ and subsequent shocks. Endogenous variables are treated symmetrically with the dependent variable and lagged values $X_{i,t-2}$ and longer lags are valid instruments for the transformed equations. If the X_{it} are predetermined, they are also uncorrelated with the current disturbances but future values of these variables are correlated with the error term, $E(X_{it}\nu_{is}) \neq 0$ for $s < t$ and 0 otherwise. Therefore, $X_{i,t-1}$ is also a valid instrument because the differenced error in equation (5.6) includes $\nu_{i,s-1}$. For strictly exogenous explanatory variables, $E(X_{it}\nu_{is}\nu_{i,s-1}) = 0$, all x_{is} are valid instruments.

The validity of GMM estimates depends on crucial assumptions. The consistency of the estimators depends on the assumption that there is no second order correlation in the error term of the transformed equation, $E(\Delta\nu_{it}\Delta\nu_{i,t-2}) = 0$. If the ν_{it} are correlated, many of the instruments will not be valid as the crucial moment conditions in equation (5.15) will no

[3]See Appendix E for more information about GMM.

longer hold. I use the Arellano and Bond (1991) test for serial correlation in transformed errors to test the validity of the instruments and the moment restrictions. I also use Hansen's J-test of overidentifying restrictions to test the assumptions on the moment conditions and assess the model specification.

Chapter 6

Results

This section presents the results of my analysis. The following is a summary of findings. Using the HSCL-20 score for mental health, improvements in mental health do have a positive and significant effect on changes in labor productivity as expected. This is found to be true using the SF-12 mental health scale as well. When the sample is divided into two samples by severity of depression at baseline, mental health has a significant effect on productivity for those categorized as having severe or moderate depression at baseline. For the less depressed individuals at baseline, improvements in mental health do not significantly improve the worker's productivity.

Improvements in mental health do not have a significant effect on a worker's attendance at work except for with the more depressed sub-sample at baseline. Physical health seems to be a better predictor of an individual's attendance at work but this result could be biased due to endogeneity.

For both presenteeism and attendance equations, I present estimates using OLS, Within-Groups (WG) estimates and two versions of GMM estimation (one using first differencing to remove the fixed effect and the second using forward orthogonal deviations). Ordinary least squares violates many of the assumptions necessary for consistent estimates. However, these estimates can be used for comparison as it inflates the coefficient for the lagged endogenous variable. Particularly, OLS inflates the coefficient for the lagged employment outcome by attributing explanatory power to it rather than an individual's fixed characteristic such as genetics. The WG estimates of the lagged dependent variable are biased downwards as this method does not account for dynamic panel bias. They are included for comparison purposes also. Bond (2002) points out that unbiased estimates of the true value of the lagged dependent variable should be less than the OLS estimate and greater than the WG estimate. For each specification, I compare the estimated coefficient to the OLS and WG estimates.

Since the unobserved individual effects (that are correlated with both the regressors in the labor equation and the employment equation) are removed through differencing, it is not possible to estimate the effects of any time-invariant regressors, either exogenous or endogenous. For this reason a quadratic in age and an interaction between age and the female indicator are included. Month dummies are used in all specifications. Summary statistics for first-differenced variables are found in Appendix G.

All GMM estimations utilize all available moment conditions. The lagged independent labor outcome, depression measure, physical health measure, and treatment decisions are considered predetermined. The strictly exogenous time-varying regressors include the unemployment rate, age-squared, the age-female interaction variable and the calendar month.

6.1 GMM Results

6.1.1 Productivity

The log of the presenteeism measure is used as the dependent variable in all presenteeism specifications. Table 6.1 includes both OLS in levels and WG estimates for the presenteeism specification using the HSCL-20 score as the measure for mental health. Mental health is a significant and negative determinant of work productivity using both methods. The OLS estimate indicates that a one point increase in mental health score (which indicates worse health) leads to a 3.50% decrease in the WLQ productivity measure, without accounting for the fact that there are multiple observations per individual. The within-groups estimate accounts for the correlation in error terms for a specific individual. The coefficient indicates that after controlling for the fact that there are multiple observations per individual a one point increase in depression score results in a 4.70% decrease in the work productivity measure.

For the GMM results, specification tests are discussed before the estimated coefficients.

The Hansen statistic tests the null hypothesis that the instruments are valid. Table 6.2 includes estimates of the presenteeism equation with first differences and with forward orthogonal deviations to remove the fixed effect. The specification statistics using FD are found in the left column of the lower panel of Table 6.2. The Hansen test p-value is 0.586, indicating that the null can not be rejected at the 5% significance level. The Arellano and Bond AR(2) statistics test the null that there is no second-order serial correlation present, $E(\Delta\nu_{it}\Delta\nu_{i,t-2}) = 0$. Again, the p-value (m2=0.072) is above the 5% significance level critical value, thus implying that there is no second-order serial correlation present and that the estimates are consistent. As expected, first order correlation is present due to first differencing (m1=0.000). There are 47 total instruments, which is less than the number of observations. Additionally, the estimate of the parameter on lagged productivity does fall within the desired range determined from the OLS and WG estimates. In the right column of Table 6.2, the fixed effect is removed using forward orthogonal deviations rather than first differencing. Again, no second serial correlation is present (m2=0.051) and the instruments are determined to be exogenous (Hansen=0.575).

Table 6.1: Estimation Results for Productivity using OLS and WG and the HSCL-20 mental health measure

Variable	OLS levels			Within groups		
Productivity $t-1$	0.405	(0.027)	***	-0.085	(0.039)	**
Attendance $t-1$	-0.023	(0.009)	***	0.029	(0.011)	***
HSCL-20 depression measure t	-0.035	(0.013)	***	-0.047	(0.018)	***
SF-12 physical health measure t	0.003	(0.001)	***	0.001	(0.002)	
Continue current medication $t-1$	0.003	(0.028)		0.037	(-0.950)	
Switch medications $t-1$	-0.011	(0.042)		0.049	(-0.800)	
Unemployment rate by zipcode t	-0.001	(0.002)		0.068	(0.039)	*
Age squared t	0.000	(0.000)		0.000	(0.000)	
Age female interaction t	0.000	(0.000)		-0.029	(0.041)	

Note: 254 individuals, 839 observations, 5 available interviews. Month dummies included in all specifications (coefficients are not reported). *significant at 10%; **significant at 5%, ***significant at 1%.

The estimated coefficients using GMM are presented in the top panel of Table 6.2. Both methods result in similar estimates indicating that improvements in the HSCL-20 mental health scores do have a significant and positive effect on the worker's productivity score.

Table 6.2: Estimation Results for Productivity using GMM and the HSCL-20 mental health measure

Variable	First Differences			Forward Orthogonal Deviations		
Δ Productivity $t-1$	0.090	(0.101)		0.062	(0.100)	
Δ Attendance $t-1$	0.009	(0.020)		0.005	(0.021)	
Δ HSCL-20 depression measure	-0.075	(0.026)	***	-0.060	(0.023)	***
Δ SF-12 physical health measure	0.003	(0.010)		0.003	(0.011)	
Δ Continue current medication	0.007	(0.145)		0.079	(0.125)	
Δ Switch medications	-0.061	(0.121)		-0.126	(0.105)	
Δ Unemployment rate by zipcode	0.128	(0.047)	***	0.140	(0.035)	***
Δ Age squared	0.000	(0.000)		0.000	(0.000)	
Δ Age female interaction	-0.028	(0.066)		-0.029	(0.039)	
Serial Correlation (m1) p-value	0.00			0.00		
Serial Correlation (m2) p-value	0.072			0.051		
Hansen Statistic	0.586			0.575		
Instruments (Full)	47			47		

Note: 221 individuals, 575 observations, 5 available interviews. Month dummies included in all specifications (coefficients are not reported). *significant at 10%, **significant at 5%, ***significant at 1%. The reported results are two-step estimates with heteroskedasticity consistent errors and adjusted using the Windmeijer (2005) finite sample correction.

Using FD, a one point improvement in HSCL-20 mental health score (a decrease in HSCL-20 score) results in a 7.50% improvement in the productivity measure. Using the full sample of individuals, with a mean HSCL-20 score equal to 1.68 at baseline, a 1 point improvement (decrease in score) would indicate that the individual's mental health had improved and that the individual would be categorized as having mild depression rather than moderate depression. The WLQ scale score ranges from 0 (limited all of the time) to 100 (limited none of the time) and represents the amount of time that the worker was limited on the job in the previous two weeks. Using an algorithm, Lerner et al (2001) translated WLQ scores into an estimate of productivity loss when compared to a healthy sample.[1] Using these estimates and the full sample at baseline with a mean productivity score equal to 74.62, the significant 7.50% change in productivity (to a productivity index score of 80.21) indicates an improvement in productivity of approximately 4 percent (from a 22.10% relative decrease to an 18.10% relative decrease) when compared to a healthy worker. This translates into a 6%

[1] See Appendix E.

decrease in work hours to compensate for the loss in productivity due to depressive symptoms. The unemployment rate is significant with the expected sign indicating that when there is a one point change in the unemployment rate in an area the worker's productivity increases by 13.00%. The lagged productivity measure is not significant. Similarly to when using first differencing, using FOD indicates that the mental health coefficient is negative and significant with a similar value (coefficient=-0.06). This indicates that a one point decrease in HSCL-20 score results in a six percent change (improvement) in the work productivity measure. Again, deviations in unemployment rates do have a significant effect on productivity. Though the lagged productivity measure is not significant, its coefficient does fall within the desired upper and lower bounds given by the OLS and WG estimates.

Tables 6.3 and 6.4 report the same specification as found in Tables 6.1 and 6.2 using the SF-12 mental health scale rather than the HSCL-20 score to measure depression outcome. Unlike the HSCL-20 scale, the SF-12 is positively scored with higher scores indicating better mental health. Again, both sets of estimates fulfill all specification requirements. The same variables are determined to be significant with the same expected signs. Estimates indicate that a one point increase in the SF-12 mental health score leads to improvements in the productivity measure of less than one percent. With a mean SF-12 score equal to 32.48 at baseline, a 1 point improvement to 33.48 would indicate that the individual's mental health had improved but, using the preferred cut-off score of 42, the individual would still meet the National Institutes for Mental Health (NIMH) criteria for major depression. Using these estimates and the full sample at baseline with a mean productivity score equal to 74.62, the significant change in productivity is negligible when compared to a healthy worker. Across all estimations methods, there is strong evidence that improvements in mental health do lead to improvements in productivity.

To compare more similar groups, I divide the sample into two groups- those with either severe depression or moderate depression at baseline and those with mild or remissive depression at baseline. This is to determine if the mental health improvements have similar effects for both groups of workers. GMM (using the HSCL-20) results are presented in Tables

Table 6.3: Estimation Results for Productivity using OLS and WG and the SF-12 mental health measure

Variable	OLS levels			Within groups		
Productivity $t-1$	0.408	(0.027)	***	-0.067	(0.038)	***
Attendance $t-1$	-0.023	(0.009)	***	0.032	(0.011)	***
SF-12 depression measure t	0.002	(0.000)	***	0.002	(0.001)	*
SF-12 physical health measure t	0.004	(0.001)	***	0.001	(0.002)	
Continue current medication $t-1$	0.008	(0.028)	*	-0.036	(0.037)	
Switch medications $t-1$	-0.005	(0.042)		-0.035	(0.049)	
Unemployment rate by zipcode t	-0.001	(0.002)		0.073	(0.039)	
Age squared t	0.000	(0.000)		0.000	(0.000)	
Age female interaction t	0.000	(0.000)		-0.034	(0.042)	

Note: 254 individuals, 839 observations, 5 available interviews. Month dummies included in all specifications (coefficients are not reported). *significant at 10%, **significant at 5%, ***significant at 1%.

Table 6.4: Estimation Results for Productivity using GMM and the SF-12 mental health measure

Variable	First Differences			Forward Orthogonal Deviations		
Δ Productivity $t-1$	0.026	(0.096)		0.060	(0.100)	
Δ Attendance $t-1$	0.011	(0.020)		0.007	(0.019)	
Δ SF-12 depression measure	0.003	(0.001)	**	0.002	(0.001)	**
Δ SF-12 physical health measure	0.003	(0.009)	***	0.005	(0.010)	
Δ Continue current medication	-0.083	(0.159)		-0.145	(0.119)	
Δ Switch medications	-0.108	(0.123)		-0.146	(0.101)	
Δ Unemployment rate by zipcode	0.133	(0.047)	***	0.158	(0.036)	***
Δ Age squared	0.000	(0.000)		0.000	(0.000)	
Δ Age female interaction	-0.039	(0.064)		-0.036	(0.037)	
Serial Correlation (m1) p-value	0.001			0.001		
Serial Correlation (m2) p-value	0.079			0.069		
Hansen	0.801			0.819		
Instruments (Full)	47			47		

Note: 221 individuals, 575 observations, 5 available interviews. Month dummies included in all specifications (coefficients are not reported). *significant at 10%, **significant at 5%, ***significant at 1%. The reported results are two-step estimates with heteroskedasticity consistent errors and adjusted using the Windmeijer (2005) finite sample correction.

6.5-6.6.[2] For the more depressed sample, improvements in mental health have a significant and positive effect on productivity (t=2.82). When comparing results from Table 6.2 and Table 6.5, estimated coefficients and standard deviations for the relatively more depressed sample are similar to those for the whole sample. This is expected since 61.20% of the full sample are characterized as either having moderate or severe depression at baseline versus only 38.73% of the sample determined to be as either having mild depression or being in remission initially. For those who were less depressed at baseline the change in mental health is no longer a significant determinant of productivity outcomes. The standard deviation for the least depressed group is much larger than for the full sample or the more depressed sub-sample. However, physical health becomes a significant determinant of productivity at the 10% level when using FOD. Unexpectedly, the coefficient is negative indicating that an improvement in physical health leads to a decrease in productivity. This could be due to the self report of health indicators.

The results indicate that accounting for the endogeneity of different variables, dynamic panel bias and the selection bias in the model gives different results of effects of depression changes on productivity. OLS coefficients underestimate effects of depression on productivity. When accounting for the effect of unobserved heterogeneity, the WG coefficients increase for the HSCL-20 measure and stay the same using the SF-12 measure for depression.

6.1.2 Absenteeism

The categorical attendance scale is treated as a linear variable in all specifications. Attendance estimates are found in Tables 6.7-6.12. Regardless of the differencing method or the mental health measure, improvements in mental health do not have a significant effect on a worker's attendance at work except for with the most depressed at baseline sub-sample.

As with the presenteeism equations, specification tests are discussed before the estimated coefficients. Table 6.7 includes both OLS in levels and WG estimates for the absenteeism specification using the HSCL-20 score as the measure for mental health. Mental health is not a significant and determinant of a worker's ability to meet the number of required work

[2] I only present HSCL-20 GMM estimates as SF-12 results are similar.

Table 6.5: Estimation Results for Productivity using GMM and the HSCL-20 mental health measure for the most depressed sample at baseline

Variable	First Differences			Forward Orthogonal Deviations		
Δ Productivity $t-1$	0.046	(0.118)	***	0.073	(0.117)	
Δ Attendance $t-1$	-0.004	(0.026)	***	-0.010	(0.024)	
Δ HSCL-20 depression measure	-0.074	(0.028)	***	-0.072	(0.025)	***
Δ SF-12 physical health measure	0.009	(0.009)		0.010	(0.009)	
Δ Continue current medication	0.116	(0.190)		0.079	(0.188)	
Δ Switch medications	-0.040	(0.176)		-0.065	(0.159)	
Δ Unemployment rate by zipcode	0.047	(0.031)		0.095	(0.033)	***
Δ Age squared	0.001	(0.000)		0.001	(0.000)	
Δ Age female interaction	-0.015	(0.091)		-0.053	(0.061)	
Serial Correlation (m1) p-value	0.007			0.004		
Serial Correlation (m2) p-value	0.410			0.552		
Hansen Statistic	0.563			0.621		
Instruments (Full)	47			47		

Note: 130 individuals, 335 observations, 5 available interviews. Month dummies included in all specifications (coefficients are not reported). *significant at 10%, **significant at 5%, ***significant at 1%. The reported results are two-step estimates with heteroskedasticity consistent errors and adjusted using the Windmeijer (2005) finite sample correction.

Table 6.6: Estimation Results for Productivity using GMM and the HSCL-20 mental health measure for the least depressed sample at baseline

Variable	First Differences			Forward Orthogonal Deviations		
Δ Productivity $t-1$	0.105	(0.113)		0.142	(0.117)	
Δ Attendance $t-1$	-0.007	(0.028)		-0.002	(0.024)	
Δ HSCL-20 depression measure	-0.061	(0.055)		-0.059	(0.051)	
Δ SF-12 physical health measure	-0.009	(0.006)		-0.010	(0.006)	*
Δ Continue current medication	0.055	(0.132)		0.052	(0.113)	
Δ Switch medications	0.086	(0.128)		0.076	(0.100)	
Δ Unemployment rate by zipcode	0.317	(0.380)		0.123	(0.290)	
Δ Age squared	0.000	(0.000)		0.000	(0.000)	
Δ Age female interaction	0.000	(0.000)		-0.071	(0.055)	
Serial Correlation (m1) p-value	0.022			0.012		
Serial Correlation (m2) p-value	0.555			0.535		
Hansen Statistic	0.430			0.354		
Instruments (Full)	47			47		

Note: 91 individuals, 240 observations, 5 available interviews. Month dummies included in all specifications (coefficients are not reported). *significant at 10%, **significant at 5%, ***significant at 1%. The reported results are two-step estimates with heteroskedasticity consistent errors and adjusted using the Windmeijer (2005) finite sample correction.

hours using both methods. In Table 6.8, the Hansen test p-value is 0.412, indicating that the null can not be rejected at the 5% significance level. The p-value testing for second order serial correlation is above the 5% significance level critical value (m2=0.886), thus implying that there is no second-order serial correlation present and that the estimates are consistent. As expected, first order correlation is present due to first differencing (m1=0.00). There are 47 total instruments, which is less than the number of observations. Additionally, the estimate of the parameter on lagged attendance measure does fall within the desired range determined from the OLS and WG estimates. The estimated coefficients in Tables 6.8-6.10 indicate that improvements in mental health do not significantly affect the worker's work attendance regardless of the transformation method or depression score that is used.

Results indicate that lagged absenteeism is significant using either FD or FOD and the HSCL-20 or SF-12 scores. Physical health changes significantly affect attendance when using FOD to remove the unobserved heterogeneity and the HSCL-20 score (t=1.95).

Results differ when divided into the two groups- the more depressed individuals at baseline and those less depressed at baseline. The results for the most depressed individuals are found in Table 6.11 and indicate that mental health changes do affect the worker's attendance at the 90% confidence level. Changes in depressive levels are not significant determinants of attendance for the less depressed group as seen in Table 6.12.

Table 6.7: Estimation Results for Attendance using OLS and WG and the HSCL-20 mental health measure

Variable	OLS levels			Within groups		
Attendance $t-1$	0.201	(0.032)	***	-0.193	(0.003)	***
Productivity $t-1$	-0.006	(0.002)	***	0.007	(0.003)	***
HSCL-20 depression measure t	-0.044	(0.047)		0.006	(0.070)	
SF-12 physical health measure t	0.023	(0.003)	***	0.025	(0.006)	***
Continue current medication $t-1$	0.065	(0.098)		0.067	(0.131)	
Switch medications $t-1$	0.049	(0.147)		0.090	(0.173)	
Unemployment rate by zipcode t	-0.004	(0.009)		-0.047	(0.137)	
Age squared t	0.000	(0.000)		.000	(0.001)	
Age female interaction t	-0.001	(0.001)		-0.011	(0.146)	

Note: 254 individuals, 839 observations, 5 available interviews. Month dummies included in all specifications (coefficients are not reported). *significant at 10%, **significant at 5%, ***significant at 1%.

Table 6.8: Estimation Results for Attendance using GMM and the HSCL-20 mental health measure

Variable	First Differences			Forward Orthogonal Deviations		
Δ Attendance $t-1$	0.167	(0.083)	**	0.186	(0.076)	***
Δ Productivity $t-1$	-0.010	(0.005)		-0.002	(0.005)	
Δ HSCL-20 depression measure	-0.076	(0.139)		-0.007	(0.129)	
Δ SF-12 physical health measure	0.065	(0.040)		0.082	(0.042)	*
Δ Continue current medication	0.185	(0.603)		0.017	(0.587)	
Δ Switch medications	0.066	(0.547)		-0.012	(0.525)	
Δ Unemployment rate by zipcode	0.080	(0.139)		0.063	(0.115)	
Δ Age squared	0.002	(0.002)		0.001	(0.002)	
Δ Age female interaction	-0.278	(0.222)		-0.132	(0.181)	
Serial Correlation (m1) p-value	0.000			0.000		
Serial Correlation (m2) p-value	0.886			0.989		
Hansen Statistic	0.412			0.426		
Instruments (Full)	47			47		

Note: 221 individuals, 575 observations, 5 available interviews. Month dummies included in all specifications (coefficients are not reported). *significant at 10%, **significant at 5%, ***significant at 1%. The reported results are two-step estimates with heteroskedasticity consistent errors and adjusted using the Windmeijer (2005) finite sample correction.

Table 6.9: Estimation Results for Attendance using OLS, WG and the SF-12 mental health measure

Variable	OLS levels			Within groups		
Attendance $t-1$	0.203	(0.032)	***	-0.188	(0.040)	***
Productivity $t-1$	-0.006	(0.002)	***	0.009	(0.003)	***
HSCL-20 depression measure t	0.002	(0.003)		-0.004	(0.004)	
SF-12 physical health measure t	0.024	(0.003)	***	-0.025	(0.006)	***
Continue current medication $t-1$	0.069	(0.098)		0.056	(0.131)	
Switch medications $t-1$	-0.053	(0.147)		0.080	(0.173)	
Unemployment rate by zipcode t	-0.005	(0.009)		-0.050	(0.137)	
Age squared t	0.000	(0.000)		0.000	(0.001)	
Age female interaction variable t	-0.001	(0.001)		-0.006	(0.146)	

Note: 254 individuals, 839 observations, 5 available interviews. Month dummies included in all specifications (coefficients are not reported). *significant at 10%, **significant at 5%, ***significant at 1%.

Table 6.10: Estimation Results for Attendance using GMM and the SF-12 mental health measure

Variable	First Differences			Forward Orthogonal Deviations		
Δ Attendance $t-1$	0.145	(0.075)	*	0.142	(0.071)	**
Δ Productivity $t-1$	-0.004	(0.005)		-0.005	(0.005)	
Δ HSCL-20 depression measure	-0.005	(0.006)		-0.004	(0.006)	
Δ SF-12 physical health measure	0.027	(0.032)		0.040	(0.032)	
Δ Continue current medication	-0.373	(0.542)		-0.413	(0.479)	
Δ Switch medications	-0.357	(0.581)		-0.371	(0.523)	
Δ Unemployment rate by zipcode	-0.036	(0.115)		-0.027	(0.130)	*
Δ Age squared	0.002	(0.002)		0.001	(0.002)	
Δ Age female interaction	-0.190	(0.221)		-0.089	(0.173)	
Serial Correlation (m1) p-value	0.000			0.000		
Serial Correlation (m2) p-value	0.504			0.562		
Hansen Statistic	0.327			0.310		
Instruments (Full)	47			47		

Note: 221 individuals, 575 observations, 5 available interviews. Month dummies included in all specifications (coefficients are not reported). *significant at 10%, **significant at 5%, ***significant at 1%. The reported results are two-step estimates with heteroskedasticity consistent errors and adjusted using the Windmeijer (2005) finite sample correction.

Table 6.11: Estimation Results for Attendance using GMM and the SF-12 mental health measure for the most depressed sample at baseline

Variable	First Differences			Forward Orthogonal Deviations		
Δ Attendance t	0.156	(0.100)		0.115	(0.094)	
Δ Productivity t	-0.004	(0.009)		-0.007	(0.007)	
Δ SF-12 depression measure	-0.017	(0.010)	*	-0.018	(0.010)	*
Δ SF-12 physical health measure	-0.035	(0.045)		-0.041	(0.048)	
Δ Continue current medication	-0.435	(0.956)		-0.264	(1.055)	
Δ Switch medications	-0.107	(1.006)		0.053	(1.080)	
Δ Unemployment rate by zipcode	0.029	(0.108)		-0.219	(0.251)	
Δ Age squared	0.001	(0.003)		-0.001	(0.003)	
Δ Age female interaction	0.026	(0.308)		0.081	(0.295)	
Serial Correlation (m1) p-value	0.002			0.001		
Serial Correlation (m2) p-value	0.162			0.236		
Hansen Statistic	0.142			0.251		
Instruments (Full)	47			47		

Note: 130 individuals, 335 observations, 5 available interviews. Month dummies included in all specifications (coefficients are not reported). *significant at 10%, **significant at 5%, ***significant at 1%. The reported results are two-step estimates with heteroskedasticity consistent errors and adjusted using the Windmeijer (2005) finite sample correction.

Table 6.12: GMM estimates for Attendance and Mental Health using the SF-12 mental health measure for the least depressed sample at baseline

Variable	First Differences			Forward Orthogonal Deviations		
Δ Attendance $t-1$	-0.093	0.159		-0.059	0.167	
Δ Productivity $t-1$	-0.128	0.007	***	-0.013	0.008	
Δ SF-12 depression measure	0.011	0.009		0.008	0.008	
Δ SF-12 physical health measure	0.047	0.028	*	0.052	0.029	*
Δ Continue current medication	- 0.249	0.393		-0.316	0.375	
Δ Switch medications	-0.305	0.520		-0.384	-0.496	
Δ Unemployment rate by zipcode	-0.957	0.811		0.384	0.534	
Δ Age squared	-0.000	0.002		.000	.002	
Δ Age female interaction	0.342	0.000		-0.104	0.272	
Serial Correlation (m1) p-value	0.037			0.030		
Serial Correlation (m2) p-value	0.476			0.463		
Hansen Statistic	0.334			0.536		
Instruments (Full)	47			47		

Note: 91 individuals, 240 observations, 5 available interviews. Month dummies included in all specifications (coefficients are not reported). *significant at 10%, **significant at 5%, ***significant at 1%. The reported results are two-step estimates with heteroskedasticity consistent errors and adjusted using the Windmeijer (2005) finite sample correction.

Chapter 7

Conclusion

Using different estimation techniques and measures of mental health status, this research provides a better understanding of the relationships between depression, treatment decisions and employment outcomes for a sample of initially depressed individuals. After using several methods to account for the endogeneity of different variables, the results consistently indicate that depression severity has a significant effect on a worker's productivity if employed. However, it is seen that depression does not significantly affect the worker's attendance. One possible explanation is that the worker attends work but performs at a low level of productivity in order to maintain employment for income or other benefits including health insurance. Additionally, results are consistent with the literature that the most depressed sub sample experiences the largest change in performance.

The results are consistent with the conclusion that it is important for a firm to consider productivity and not just attendance at work. If the workers are still coming to work, but are not efficient, it can be a major yet often overlooked cost to the employer.

It is always important to address possible problems with the results. I do not model attrition in the ARTIST trial or in the individual's decisions to continue, switch or stop treatment. It is also important to note that these individuals sought depressive treatment and participated in the treatment trial. This could lead to additional selection issues. A benefit is that people act as their own control and there will be less baised based on recall of health choices, outcomes and employment outcomes.

Appendix A

Sample Characteristics by Participation Pattern

Sample attrition is a potential source of selection bias in this study. It arises as certain participants drop out of the study before completing each of the five interviews. Here, the researcher does not know why the individual drops out of the sample. Possible explanations that are systematically related to the response variable include no change in mental health, or a recovery from depression. Death is also a possibility.

Table A.1: Health Characteristics by Participation Pattern

	All	11111	11110	11100	11000	11101
n	399	315	34	15	20	15
Age	43.09 (10.24)	43.84 (9.90)	40.26 (11.09)	42.20 (11.06)	38.35 (12.67)	41.13 (9.17)
Female	79.94	80.95	82.35	80.00	63.64	80.00
Race						
White	82.96	84.13	85.29	80.00	70.00	73.33
Black	14.29	13.65	11.76	13.33	30.00	13.33
Other	02.75	02.22	02.94	06.67	00.00	13.33
At Baseline						
HSCL-20	1.69 (0.68)	1.69 (0.66)	1.75 (0.65)	1.71 (0.64)	1.40 (0.85)	2.21 (0.63)
SF-12	31.79 (10.98)	32.48 (10.91)	30.06 (11.40)	30.07 (10.50)	31.02 (11.66)	23.25 (7.55)
PCS-12	50.36 (9.82)	49.95 (9.79)	51.35 (10.61)	50.32 (11.07)	54.58 (7.30)	51.30 (10.27)
HSCL-20						
All	1.08 (0.74)	1.05 (0.73)	1.16 (0.73)	1.34 (0.76)	1.09 (0.78)	1.49 (0.83)
Baseline	1.70 (0.68)	1.68 (0.66)	1.75 (0.65)	1.71 (0.64)	1.48 (0.85)	2.21 (0.63)
month 1	1.08 (0.64)	1.07 (0.63)	1.11 (0.63)	1.18 (0.66)	0.78 (0.56)	1.44 (0.74)
month 3	0.92 (0.67)	0.90 (0.66)	0.92 (0.66)	1.13 (0.86)	-	1.25 (0.76)
month 6	0.83 (0.65)	0.83 (0.66)	0.85 (0.64)	-	-	-
month 9	0.77 (0.64)	0.76 (0.63)	-	-	-	1.07 (0.77)
SF-12						
All	43.29 (12.21)	44.04 (12.06)	40.41 (11.91)	38.20 (13.19)	38.27 (13.01)	36.82 (11.41)
Baseline	31.79 (10.98)	32.48 (10.91)	30.06 (11.40)	30.07 (10.50)	31.02 (11.66)	23.25 (7.55)
month 1	43.35 (10.54)	43.83 (10.64)	41.17 (8.98)	39.78 (12.30)	45.52 (10.05)	38.77 (9.21)
month 3	46.57 (10.91)	47.13 (10.86)	44.96 (11.08)	44.32 (13.06)	-	40.89 (7.28)
month 6	47.29 (9.96)	47.49 (10.01)	45.44 (9.42)	-	-	-
month 9	49.00 (9.85)	49.26 (9.79)	-	-	-	43.49 (9.77)

Refer to Table 3.2.

Table A.2: Labor Characteristics by Participation Pattern

	All	11111	11110	11100	11000	11101
n	399	315	34	15	20	15
Employment %						
All	74.73 (43.46)	74.22 (43.75)	82.35 (38.26)	60.00 (49.54)	75.00 (43.85)	81.67 (39.02)
Baseline	75.94 (42.79)	75.23 (43.23)	85.29 (35.94)	60.00 (50.71)	70.00 (47.01)	93.30 (25.80)
month 1	74.18 (43.81)	72.69 (44.62)	85.29 (35.94)	66.66 (48.79)	80.00 (41.04)	80.00 (41.40)
month 3	74.67 (43.54)	74.60 (43.59)	79.41 (41.04)	53.33 (51.64)	-	86.67 (35.18)
month 6	74.78 (43.48)	74.28 (43.77)	79.41 (41.04)	-	-	-
month 9	73.94 (43.96)	74.28 (43.77)	-	-	-	66.67 (48.79)
Work Hours						
All	4.17 (1.05)	4.24 (1.00)	3.90 (1.16)	3.63 (1.30)	3.93 (1.22)	3.55 (1.20)
Baseline	3.72 (1.13)	3.83 (1.09)	3.31 (1.16)	3.11 (1.53)	3.78 (1.05)	3.07 (1.27)
month 1	4.19 (1.02)	4.27 (0.94)	3.93 (1.25)	3.80 (1.13)	4.06 (1.39)	3.58 (1.16)
month 3	4.36 (0.97)	4.43 (0.94)	4.14 (0.99)	4.00 (1.19)	-	3.77 (1.24)
month 6	4.31 (0.99)	4.31 (0.98)	4.26 (1.06)	-	-	-
month 9	4.36 (0.93)	4.38 (0.92)	-	-	-	3.90 (1.10)
Productivity						
All	73.41 (20.02)	74.62 (19.46)	70.23 (18.98)	67.03 (25.73)	71.00 (21.63)	56.73 (22.76)
Baseline	62.89 (20.76)	64.60 (20.61)	57.80 (17.29)	59.44 (26.97)	64.28 (21.91)	45.35 (16.57)
month 1	73.43 (19.70)	74.07 (19.07)	72.24 (19.89)	71.50 (25.06)	76.87 (20.23)	60.83 (24.29)
month 3	75.84 (19.49)	77.11 (18.99)	74.26 (17.19)	70.00 (26.59)	-	59.61 (22.21)
month 6	78.90 (17.11)	79.07 (17.28)	77.40 (15.71)	-	-	-
month 9	77.77 (18.16)	78.36 (17.57)	-	-	-	64.00 (26.33)

Refer to Table 3.2.

Appendix B

Questions from the 36-item Short-form Health Survey (SF-36)

1. During the past 4 weeks have you had any of the following problems with your work or other regular daily activities as a result of any emotional problems (such as feeling depressed or anxious)?

   (a) Cut down the amount of time you spent on work or other activities (Role Emotional; Yes, No)

   (b) Accomplished less than you would like (Role Emotional; Yes, No)

   (c) Didn't do work or other activities as carefully as usual (Role Emotional; Yes, No)

2. During the past 4 weeks, to what extent has your physical or emotional problems interfered with your normal social activities with family, friends, neighbors or groups? (Social Functioning; Not at all, Slightly, Moderately, Quite a bit, Extremely)

3. These questions are about how you feel and how things have been with you during the past 4 weeks. For each question, please give the one answer that comes closest to the way you have been feeling. How much of the time during the past 4 weeks

   (a) Did you feel full of pep? (Vitality; All of the time, Most of the time, A good bit of the time, Some of the time, A little of the time, None of the time)

   (b) Have you been a very nervous person? (Mental Health; All of the time, Most of the time, A good bit of the time, Some of the time, A little of the time, None of the time)

(c) Have you felt so down in the dumps that nothing could cheer you up? (Mental Health; All of the time, Most of the time, A good bit of the time, Some of the time, A little of the time, None of the time)

(d) Have you felt calm and peaceful? (Mental Health; All of the time, Most of the time, A good bit of the time, Some of the time, A little of the time, None of the time)

(e) Did you have a lot of energy? (Vitality; All of the time, Most of the time, A good bit of the time, Some of the time, A little of the time, None of the time)

(f) Have you felt downhearted and blue? (Mental Health; All of the time, Most of the time, A good bit of the time, Some of the time, A little of the time, None of the time)

(g) Did you feel worn out? (Vitality; All of the time, Most of the time, A good bit of the time, Some of the time, A little of the time, None of the time)

(h) Have you been a happy person? (Mental Health; All of the time, Most of the time, A good bit of the time, Some of the time, A little of the time, None of the time)

(i) Did you feel tired? (Vitality; All of the time, Most of the time, A good bit of the time, Some of the time, A little of the time, None of the time)

4. During the past 4 weeks, how much of the time has your physical health or emotional problems interfered with your social activities (like visiting with friends, relatives etc.)? (Social Functioning; All of the time, Most of the time, A good bit of the time, Some of the time, A little of the time, None of the time)

Each item is used in scoring only one scale. The scores from each of the four scales are then combined to attain the mental component summary score.

Appendix C

Productivity Questions from the Work Limitations Questionnaire

A productivity scale is created from five questionnaire items. Conditional on working, a negatively scored scale is used with a higher score indicating more limitation. The response choices are: All of the time, Most of the time, Some of the time, A slight bit of the time,and None of the time.[1] The questionnaire items include:

1. How much of the time during the past two weeks did you finish work on time?
2. How much of the time during the past two weeks did you do your work without making mistakes?
3. How much of the time during the past two weeks did you feel you've done what you are capable of doing?
4. How much of the time during the past two weeks did your physical health or emotional problems make it difficult for you to handle the workload?
5. How much of the time during the past two weeks did your physical health or emotional problems make it difficult for you to work fast enough?

[1] See WLQ for scoring algorithm.

Appendix D

Zipcode Tabulation Areas

Zip code information is included in the ARTIST trial data set at each interview. Zip code information is included in this study; however, Zip Code Tabulation Areas are used for extracting information on a geographic area from the given zip code. The United States Census Bureau describes them,

> ZIP Code Tabulation Areas (ZCTAs) are a new statistical entity developed by the U.S. Census Bureau for tabulating summary statistics from Census 2000. This new entity was developed to overcome the difficulties in precisely defining the land area covered by each ZIP Code. Defining the extent of an area is necessary in order to accurately tabulate census data for that area. ZCTAs are generalized area representations of U.S. Postal Service (USPS) ZIP Code service areas. Simply put, each one is built by aggregating the Census 2000 blocks, whose addresses use a given ZIP Code, into a ZCTA which gets that ZIP Code assigned as its ZCTA code. They represent the majority USPS five-digit ZIP Code found in a given area. For those areas where it is difficult to determine the prevailing five-digit ZIP Code, the higher-level three-digit ZIP Code is used for the ZCTA code. It is important to note the following: In most instances the ZCTA code equals the ZIP Code for an area In creating ZCTAs, the Census Bureau took the ZIP Code used by the majority of addresses in a area for the ZCTA code, some addresses will end up with a ZCTA code different from their ZIP Code. Some ZIP Codes represent very few addresses (sometimes only one) and therefore will not appear in the ZCTA universe. The term ZCTA was created to differentiate between this entity and true USPS ZIP Codes. Information on the Census Bureau's position

regarding ZIP Code data. ZCTA is a trademark of the U.S. Census Bureau; ZIP Code is a registered trademark of the U.S. Postal Service.

Appendix E

The Work Limitations Questionnaire: Estimated Productivity Impact of Health-Related Work Limitations Based on WLQ Index Score

Table E.1: The Work Limitations Questionnaire: Estimated Productivity Impact of Health-Related Work Limitations Based on WLQ Index Score

WLQ Index Score	% decrease in productivity (compared to healthy)	% increase in work hours to compensate for productivity loss
100	-	-
95	4.9	5.1
90	9.5	10.5
85	14.1	16.2
80	18.1	22.1
75	22.1	28.4
70	25.9	34.9
65	29.5	41.9
60	32.9	49.2
55	36.2	56.8
50	39.4	64.9

I report WLQ Index Scores on a positive scale with a higher score indicating better labor outcomes.

Appendix F

The GMM Estimator

The GMM estimator based on these moment conditions minimises the limiting variance of the sample moments,

$$V = \left(\frac{1}{N}\sum_{i=1}^{N}\Delta\nu_i' Z_i\right)\left(\frac{1}{N}\sum_{i=1}^{N} Z_i'\Delta\nu_i\right). \tag{F.1}$$

Let

$$L = \begin{bmatrix} L_1 \\ \vdots \\ L_N \end{bmatrix}, \quad X = \begin{bmatrix} X_1 \\ \vdots \\ X_N \end{bmatrix}, \quad \text{and } \nu_i = \begin{bmatrix} \nu_1 \\ \vdots \\ \nu_N \end{bmatrix}.$$

The optimal GMM estimator is then

$$\hat{\Theta} = (X'Z'\hat{V}^{-1}Z'X)^{-1}X'Z\hat{V}^{-1}Z'L \tag{F.2}$$

where $\hat{V}$ is a consistent estimate of V. Generally, the optimal $\hat{V}$ is

$$\hat{V} = \frac{1}{N}\sum_{i=1}^{N} Z_i'\hat{\nu}_i\hat{\nu}_i{}'Z_i \tag{F.3}$$

where $\hat{\nu}_i$ is an estimate of the vector of residuals ν_{it}. This vector of residuals is obtained from an initial consistent estimator. For this initial consistent estimator, Arellano and Bond (1991) use

$$\hat{V}_c = \frac{1}{N}\sum_{i=1}^{N} Z_i' H Z_i, \tag{F.4}$$

where for $T = 5$,

$$H = \begin{pmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{pmatrix}.$$

With T fixed at 5 and $N \to \infty$, the resulting $\hat{\Theta} \sim N(\Theta, \sum)$ is consistent. Generally, the GMM estimator based on these moment conditions minimises

$$J_N = \left(\frac{1}{N}\sum_{i=1}^{N} \Delta\mu_i' Z_i\right) W_N \left(\frac{1}{N}\sum_{i=1}^{N} Z_i'\Delta\mu_i\right) \tag{F.5}$$

using the weight matrix

$$W_N = \left[\frac{1}{N}\sum_{i=1}^{N} \left(Z_i'\widehat{\Delta\mu_i}\widehat{\Delta\mu_i}' Z_i\right)\right]^{-1}. \tag{F.6}$$

Standard errors for the transformed estimates are obtained by taking the square root of the diagonal of sigma. With heteroskedastic disturbances, the two-step estimator is more efficient.

Appendix G

Additional Information on Variables

Table G.1: Constant Explanatory Variables- Definitions

Variable	Definition
Female	Dummy variable=1 if female, =0 if male
Race	
White	=1 if white, =0 otherwise
Black	=1 if black, =0 otherwise
Asian	=1 if asian, =0 otherwise
Other	=1 if other race, =0 otherwise
Evertreated	=1 if evertreated for depression prior to study, =0 otherwise
Randomized Drug	SSRI randomized
Paxil	=1 if Paxil randomized, =0 otherwise
Prozac	=1 if Prozac randomized, =0 otherwise
Zoloft	=1 if Zoloft randomized, =0 otherwise
Same Drug	=1 if randomized drug= prescribed drug, =0 otherwise

Table G.2: Summary of First-Differenced Characteristics of Sample

Δ Characteristic	Mean	SD
Δ Mental Health HSCL-20 score	-0.230	0 .559
Δ Mental Health SF-12 score	4.195	10.352
Δ Physical Health SF-12 score	0.443	7.053
Δ Productivity	3.560	16.347
Δ Attendance	0.132	0.379

REFERENCES

Anderson, T. and C. Hsiao (1981). Estimation of dynamic models with error components. *Journal of the American Statistical Association 76*(375), 598–606.

Anderson, T. and C. Hsiao (1982). Formulation and estimation of dynamic models using panel data. *Journal of econometrics 18*(1), 47–82.

Arellano, M. and S. Bond (1991). Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations. *The Review of Economic Studies 58*(2), 277–297.

Arellano, M. and O. Bover (1995). Another look at the instrumental variable estimation of error-components models* 1. *Journal of econometrics 68*(1), 29–51.

Arellano, M. and B. Honore (2001). Panel data. *Handbook of Econometrics 5*.

Becker, G. (1967). *Human capital and the personal distribution of income: An analytical approach*. Institute of Public Administration.

Berndt, E., S. Finkelstein, P. Greenberg, R. Howland, A. Keith, A. Rush, J. Russell, and M. Keller (1998). Workplace performance effects from chronic depression and its treatment. *Journal of Health Economics 17*(5), 511–535.

Bloom, B. (2004). Prevalence and economic effects of depression. *Managed care (Langhorne, Pa.) 13*(6 Suppl Depression), 9.

Bond, S. (2002). Dynamic panel data models: a guide to micro data methods and practice. *Portuguese Economic Journal 1*(2), 141–162.

Broadhead, W., D. Blazer, L. George, and C. Tse (1990). Depression, disability days, and days lost from work in a prospective epidemiologic survey. *Journal of the American Medical Association 264*(19), 2524–2528.

Brouwer, W., N. Van Exel, M. Koopmanschap, and F. Rutten (2002). Productivity costs before and after absence from work: as important as common? *Health Policy 61*(2), 173–187.

Burton, W., G. Pransky, D. Conti, C. Chen, and D. Edington (2004). The association of medical conditions and presenteeism. *Journal of Occupational and Environmental Medicine 46*(6), S38.

Claxton, A., A. Chawla, and S. Kennedy (1999). Absenteeism among employees treated for depression. *Journal of occupational and environmental medicine 41*(7), 605.

Druss, B., M. Schlesinger, and H. Allen Jr (2001). Depressive symptoms, satisfaction with health care, and 2-year work outcomes in an employed population. *American Journal of Psychiatry 158*(5), 731.

Ettner, S. L., R. G. Frank, and R. C. Kessler (1997, October). The impact of psychiatric disorders on labor market outcomes. *Industrial and Labor Relations Review 51*(1), 64–81.

Frank, R. and P. Gertler (1991). An assessment of measurement error bias for estimating the effect of mental distress on income. *Journal of Human Resources 26*(1), 154–164.

Greenberg, P., R. Kessler, H. Birnbaum, S. Leong, S. Lowe, P. Berglund, and P. Corey-Lisle (2003). The economic burden of depression in the United States: How did it change between 1990 and 2000?. *Journal of Clinical Psychiatry 64*(12), 1465–1475.

Grossman, M. (1972). On the concept of health capital and the demand for health. *Journal of Political economy 80*(2).

Kessler, R., H. Akiskal, M. Ames, H. Birnbaum, P. Greenberg, et al. (2006). Prevalence and effects of mood disorders on work performance in a nationally representative sample of US workers. *American Journal of Psychiatry 163*(9), 1561.

Kessler, R., W. Chiu, O. Demler, and E. Walters (2005). Prevalence, severity, and comorbidity of 12-month DSM-IV disorders in the National Comorbidity Survey Replication. *Archives of general psychiatry 62*(6), 617.

Kessler, R., J. Soukup, R. Davis, D. Foster, S. Wilkey, M. Van Rompay, and D. Eisenberg (2001). The use of complementary and alternative therapies to treat anxiety and depression in the United States. *American Journal of Psychiatry 158*(2), 289.

Lerner, D., D. Adler, H. Chang, L. Lapitsky, M. Hood, C. Perissinotto, J. Reed, T. McLaughlin, E. Berndt, and W. Rogers (2004). Unemployment, job retention, and productivity loss among employees with depression. *Psychiatric Services 55*(12), 1371.

Marcotte, D. E., V. Wilcox-Gok, and P. Redmon (1999). Prevalence and patterns of major depressive disorder in the united states labor force. *Journal of Mental Health Policy and Economics 2*, 123–131.

Mintz, J., L. Mintz, M. Arruda, and S. Hwang (1992). Treatments of depression and the functional capacity to work. *Archives of General Psychiatry 49*(10), 761–768.

Mullahy, J. and J. Sindelar (1990). *Research in Human Capital and Development*, Volume 6 of *Gender Differences in the Effects of Mental Health on Labor Force Participation.* Greenwich JAI Press.

Simon, G., W. Katon, C. Rutter, M. VonKorff, E. Lin, P. Robinson, T. Bush, E. Walker, E. Ludman, and J. Russo (1998). Impact of improved depression treatment in primary care on daily functioning and disability. *Psychological Medicine 28*(03), 693–701.

Sipkoff, M. (2006). Depression is prevalent and pernicious, costing employers billions each year. *Depression in the Workplace 1*(1), 4–8.

Stewart, W., J. Ricci, E. Chee, and D. Morganstein (2003). Lost productive work time costs from health conditions in the United States: results from the American Productivity Audit. *Journal of Occupational and Environmental Medicine 45*(12), 1234–1246.

Swindle, R., K. Kroenke, and L. A. Braun (2001). *Research in Human Capital and Development*, Volume 14 of *Energy and Improved Workplace Productivity in Depression.*

Von Korff, M., J. Ormel, W. Katon, and E. Lin (1992). Disability and depression among high utilizers of health care: a longitudinal analysis. *Archives of general psychiatry 49*(2), 91.

Windmeijer, F. (2005). A finite sample correction for the variance of linear efficient two-step GMM estimators. *Journal of econometrics 126*(1), 25–51.