

ESSAYS ON THE ECONOMICS OF INVESTMENTS IN HEALTH

BY

ALEXANDER N. SLADE

DISSERTATION

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Agricultural and Consumer Economics
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2012

Urbana, Illinois

Doctoral Committee:

Associate Professor Paul E. McNamara, Chair
Professor Andrea H. Beller
Professor Craig Gundersen
Associate Professor Elizabeth T. Powers

ABSTRACT

This dissertation uses three longitudinal data sets to investigate the short- and long-term effects of family and disease on individual health outcomes and behaviors. The first essay begins by exploring the role of a recent diagnosis of diabetes on the dynamics of several health investment behaviors in older Americans, including smoking, alcohol use, frequent exercise, and the outcome of overweight or obese status. Nine waves of longitudinal data from the Health and Retirement Study are used. The behavioral response of diagnosed diabetics is compared to a group of individuals who are statistically likely to have high blood sugar levels based on Hemoglobin A1c (HbA1c) levels. A population average probit model, estimating using the Generalized Estimating Equation (GEE) framework, is used to clarify this relationship. While individuals tend to respond initially by increasing exercise levels and decreasing weight, cigarette and alcohol consumption, the response decreases with time and recidivism is present after initial diagnosis, especially through reduced exercise levels and weight gain.

Essays two and three examine family-related (and more specifically, father-related) determinants of child and maternal health outcomes. The second essay uses panel data from the National Longitudinal Study of Adolescent Health (Add Health) to explore how childhood family structure through age 15 (as measured by biological father absence and other male entrance) affects physical and mental health outcomes (including self-reported health status, overweight or obesity status, and depression) smoking behavior during adolescence (Wave I), as well as subsequently into young adulthood (Waves II through IV). Static logit models are estimated to assess the role of family structure on adolescent health outcomes and smoking, whereas discrete-time hazard models are used to estimate changes in these outcomes (e.g.,

quitting or starting smoking) after adolescence. These models are estimated separately for boys and girls. Results suggest that while most of the effect of family structure occurs during adolescence, there are residual effects of paternal absence in later life that can discourage smoking and reporting favorable health outcomes.

Finally, the third essay uses panel data from the Fragile Families and Child Wellbeing Study (FFCWS) are used to test the hypothesis that increased father involvement as measured through any child support, informal support, in-kind support, and father-child contact (i.e., visitation) can have spillover effects on the mother's mental health, as measured by depression. Maternal depression is an important outcome to consider as it is highly prevalent (especially in families headed by a single mother). Pooled linear probability models and fixed-effects models are estimated to clarify the relationship between various forms of nonresident father involvement on maternal depression. The results suggest that in-kind support is most protective for maternal depression. However, father visitation, especially when done so infrequently, can raise maternal depression levels.

ACKNOWLEDGEMENTS

I want to express my gratitude for all the individuals who made this dissertation possible. My thanks go out to my committee members: Paul McNamara, Andrea Beller, Craig Gundersen, and Elizabeth Powers, who helped me improve my drafts, clarify ideas, and hone my writing skills. Thanks also goes out to Carl Nelson and Alex Winter-Nelson, who have provided advice and guidance throughout my doctoral career. I am also grateful for the generous support of the Henry Fellowship, without which much of this research would not be possible. Additional thanks goes to several people at the College of Medicine: Nora Few, Jim Hall, and Susan Kies, who have provided me with advice and counsel throughout my doctoral studies.

TABLE OF CONTENTS

Chapter 1	
Introduction.....	1
Chapter 2	
Health Investment Decisions in Response to Diabetes Information in Older Americans..	10
Chapter 3	
The Role of Family Structure in the Evolution of Health from Adolescence to Young Adulthood by Gender.....	65
Chapter 4	
The Relationship between Nonresident Father Involvement and Maternal Depression in Fragile Families	116
Chapter 5	
Conclusions.....	169

Chapter 1

Introduction

This dissertation uses three panel data sets to empirically examine the evolution of health outcomes and behaviors throughout the life course, and potential factors that may impact this trajectory. In all cases, the panel nature of the data sets is exploited using longitudinal, microeconomic techniques. Health capital (and health behaviors that can influence health) is an important form of human capital development and skill formation. All three essays deal with the production of health (or the demand for inputs in this production function), which was formalized by Grossman (1972, 2000). Grossman's model posits that individuals are endowed with a certain amount of health capital, which depreciates over time. Individuals can gain utility directly from good health and through increased available time. Importantly, health investments, such as increasing exercise or quitting smoking, can reduce the rate of depreciation and improve one's stock of good health. Individuals maximize their present discounted lifetime utility to invest in healthy behaviors (such as smoking cessation or exercise) until the returns on the health investments equal the opportunity costs of making these health investments.

Health can be thought of as a commodity that can be produced through time and purchased good inputs (Becker 1965). For instance, individuals may invest in their own health through the purchase of a gym membership and spending time to exercise. In the context of family economics, the family can be viewed as a producer of child health. The second two papers of this dissertation focus on relatively disadvantaged families: those with absent biological fathers. Mothers who head these families almost universally have lower incomes available, which constrains their budget for purchased inputs to the child's (and their own) health. These

mothers also tend to spend a larger fraction of their time in the labor market, which limits their ability to be involved in the production of their child's (or their own) health.

This dissertation investigates how individuals make investments in their own health, and how important family correlates and dynamics may affect the health of both parents and children. All three papers are united by Grossman's theory of health production, though they represent different applications of this theory. Further, each paper uses different econometric techniques are used to estimate the reduced form demand systems for health outcomes and behaviors.

The first paper examines the production of secondary and tertiary health preventative behaviors in response to a new diagnosis of type II diabetes. Type II diabetes is a disease that is characterized by an excess of glucose in the blood (hyperglycemia) that is often related to insulin resistance and sometimes insulin deficiency. Importantly, blood sugar levels are largely related to one's dietary and lifestyle choices. In some cases, diabetes can be completely controlled through lifestyle changes without the need for any medication. Much of the existing literature on preventative behaviors deals with primary preventive behaviors, which are behaviors which lower the probability of acquiring a particular disease (such as a flu vaccination). Secondary prevention deals with actions one can take to minimize the probability that the disease will cause illness or morbidity. Finally, tertiary prevention deals with actions taken to reduce or alleviate some of the complications associated with the disease. The behaviors that I focus on in this paper include exercise, weight loss, smoking cessation, and alcohol abstinence. While alcohol and exercise can directly affect blood sugar levels, smoking and weight loss have more indirect effects (American Diabetes Association 2008; Sigal et al. 2006). Weight loss has been shown to increase insulin sensitivity and consequently lower blood sugars (Wing et al. 1987). While

smoking is an independent risk factor for the development of diabetes, smoking has high costs for diabetics, especially in increasing the likelihood for experiencing vascular complications. These can include sequelae such as foot amputation, heart disease, hypertension, and strong.

The paper capitalizes on the panel aspects of the RAND-enhanced Health and Retirement Study (RAND HRS) to estimate a dynamic, population average model to predict the long-term behavioral trajectory in these health investment behaviors after a new diagnosis of diabetes, and up to 14 years subsequent to diagnosis. This paper seeks to ascertain the treatment effect of the “average treatment on the treated.” That is, the contrast between a newly-diagnosed diabetic with an undiagnosed diabetic as if they were newly diagnosed. In order to estimate this contrast, a comparison group is constructed from 2006 HRS Biomarker data to using propensity scores to represent individuals who are statistically likely to be diagnosed with diabetes.

Individuals who will ultimately get diagnosed with diabetes in the HRS are included in the comparison group as well. Individuals who are not either diagnosed with diabetes or in the comparison group are excluded from the analyses. An additional analytical issue in the model is the unobserved clinical latency period that can precede diagnosis for many years. As a result, diabetes can be detected at vastly different points of the clinical evolution of the disease. As such, an individual getting diagnosed “late” in the course of the disease may have a different trajectory than someone who caught the disease at a relatively early point. I use medication use at the point of diagnosis to make this distinction. Increased medication use correlates well with clinical severity of the disease (as measured by HbA1c) in the HRS. A dynamic population average probit model is used to estimate the trajectory of these health behaviors. Inverse Probability Weights (IPWs) are applied in the models to account for attrition (especially health-related attrition), and baseline selection propensity into the HRS (Kapteyn et al. 2006). Results

suggest that individuals have a strong response at the point of diagnosis in quitting smoking and drinking, losing weight, and increasing exercise. However, the effect diminishes as time endures after diagnosis. While smoking and alcohol use remain low after diagnosis, individuals also to reduce exercise levels and gain weight two or more years after diagnosis, suggesting that weight and exercise maintenance (which are in many ways more important for diabetes self-management than smoking or drinking) may be difficult tasks for many diabetics.

Essays 2 and 3 focus on family-related correlates on the production of child and maternal health, respectively. The prevalence of the traditional, intact, two-biological-parent family has been steadily declining in the United States (Kreider and Ellis 2011). This trend is largely precipitated by the fact that over 40% of all children born today are born to an unmarried mother (Hamilton, Martin and Ventura 2011). An extensive body of literature has documented that children living in non-intact families (e.g., single-parent or step-parent families) have adverse outcomes (especially in cognitive domains) (e.g., Hill, Yeung and Duncan 2001; Krein and Beller 1988). Children growing up in these family structures are also at risk for adverse health outcomes during adolescence, as well as an increased risk for participating in unhealthy behaviors (especially smoking) (Antecol and Bedard 2007; Francesconi, Jenkins and Siedler 2010; Stewart and Menning 2009). Literature documenting the health consequences of growing up in non-intact families has also been emerging. Adolescents spending time in single-parent or step-parent families are more likely to have lower self-reported health status, higher rates of accidental injury, asthma, and depression compared to their counterparts in two-biological-parent families (Bramlett and Blumberg 2007; Dawson 1991; Harknett 2009; Heard, Gorman and Kapinus 2008; Langton and Berger 2011). While some studies find that family structure has an effect on child outcomes *per se*, Ginther and Pollak (2004) find that much of the effect of family

structure on child educational attainment can be explained by socioeconomic factors, including income.

The second paper in this dissertation addresses several shortcomings in the literature on the relationship between family structure and child health to date. The majority of the existing literature use very crude measures of family structure, as binary variables of the contemporaneous family structure collected at the time of the survey interview. This can be problematic, as many surveys collect data on children over wide age range. In the case of the National Longitudinal Study of Adolescent Health (Add Health), Wave I surveys adolescents from age 12 to 21. Spending time in a given family structure (e.g., a single-mother family) at age 12 may have very different effects on child outcomes (including health) than at age 21. Studies that construct richer measures of family structure over time often use the mother's marital history to construct variables about family structure. This generally requires the assumption that the mother's spouse at the time of the child's birth is the child's biological father. This assumption requires that, by default, children born to unmarried mothers are classified as growing up without a father. While this assumption may have been appropriate in older cohorts, it will likely result in misclassification when applied to more recent cohorts where the out-of-wedlock birth rate is much higher. Furthermore, many studies which consider the effects of paternal absence on child outcomes do not consider the role of step- or cohabiting- fathers that may enter the household subsequent to the departure of the biological father. Our analyses using Add Health data suggest that over half of children who experienced the departure of their biological father experienced the entrance of at least one other man subsequent to the departure of the biological father. Finally, the persistence of the adverse health behaviors and outcomes that are related to paternal absence past adolescence has remained unexplored.

The second paper uses all four waves of the Add Health data to explore the short- and long- term effects of family structure on health, with a focus on paternal absence and other male entrance. Other males refer to men with whom the child’s biological mother shared a marriage or marriage-like¹ relationship, which represents an aggregation of step-fathers and men with whom the mother shared a marriage-like relationship. We consider four health outcomes or behaviors: self-reported health status, depression, overweight or obesity, and smoking. We paint the most accurate possible picture of family structure from birth through age 15 by combining the maternal marital history with questions specifically asking about the child’s biological father, which is often unavailable in other surveys. We then estimate static logit models to ascertain the role of paternal absence and other male entrance on these adverse health outcomes (and smoking). Discrete-time hazard models are estimated to assess how family structure affects the persistence of these outcomes subsequent to adolescence. Results suggest that paternal absence has a large role in raising the propensity for developing adverse health outcomes or for regularly smoking during adolescence, though the effects are somewhat stronger for girls. Results from discrete-time hazard models reveals that boys whose father left during the early part of their lives are less likely to quit smoking (if they had ever regularly smoked by adolescence) or report excellent or very good health status (if they had reported good or worse health during adolescence). Girls are marginally more likely to start smoking if their father ever left from birth to age 15. Other male entrance also tended to reduce the magnitude and significance of biological father absence in many cases.

While the second paper explored the potential effects of family structure on children’s health, the health of the mother (especially her mental health) is often overlooked consequence

¹ A marriage-like relationship is defined by Add Health as “living with someone as if you were married to him or her when you are not.”

of paternal absence. In the third paper of this dissertation, the effects of nonresident father involvement on maternal depression levels are explored. Depression rates among young mothers have been found to be very high (about 15%). Among single mothers, it has been reported to be almost twice as high. While depression can lead to morbidity and disability in this population, maternal depression can adversely affect the health and developmental trajectory of her children, as well.

As income is generally much lower in families without resident fathers, the child support system in the United States was designed to alleviate some of the financial consequences that these mothers may face. As child support is designed to primarily help children living in these families, there is an extensive literature documenting the positive effects of child support receipt on childrens' outcomes (especially involving their cognitive development) (Amato and Gilbreth 1999; Argys et al. 1998; Knox 1996). However, the indirect effects of nonresident father involvement on the mother have remained unexplored. Food insecurity, which is closely related to depression, has been found to be reduced with nonresident father involvement, including child support and father-child contact (i.e., visitation) (Garasky and Stewart 2007; Nepomnyaschy and Garfinkel 2011). Other measures of material hardship (e.g., eviction or inability to pay utility bills) have also shown to be sensitive to nonresident father involvement (Nepomnyaschy and Garfinkel 2011). In both cases, visitation had a stronger role in improving maternal welfare than material forms of support, including child support, informal support, and in-kind support.

The third paper uses panel data from the Fragile Families and Child Wellbeing Study (FFCWS) to the hypothesis that increased levels of nonresident father involvement can improve depression levels in mothers. The FFCWS is a longitudinal study focusing on children born out-of-wedlock in major metropolitan areas with populations exceeding 200,000. Pooled and fixed-

effects linear probability models are estimated. Nonresident father involvement can be disaggregated into the receipt of any child support, any informal support, any in-kind support, and days of father-child contact. Maternal material hardship and paternal incarceration are explored as mediating variables. Results suggest that while in-kind support has a protective effect on maternal depression, father visitation, especially if done so infrequently, can actually harm mothers by increasing their depression levels. These findings are robust to the inclusion of maternal fixed effects. This paper provides evidence that the relationship between nonresident father involvement, material hardship, and maternal depression is likely very complex. Furthermore, the paper underscores the importance of informal means of support, especially among disadvantaged families in urban settings.

1.1. References

- Amato, P.R. and J.G. Gilbreth. 1999. "Nonresident Fathers and Children's Well-Being: A Meta-Analysis." *Journal of Marriage and Family* 61(3):557-573.
- American Diabetes Association. 2008. "Nutrition Recommendations and Interventions for Diabetes." *Diabetes Care* 31(Supplement 1):S61-S78.
- Antecol, H. and K. Bedard. 2007. "Does Single Parenthood Increase the Probability of Teenage Promiscuity, Substance Use, and Crime?" *Journal of Population Economics* 20(1):55-71.
- Argys, L.M., H.E. Peters, J. Brooks-Gunn, and J.R. Smith. 1998. "The impact of child support on cognitive outcomes of young children." *Demography* 35(2):159-173.
- Becker, G.S. 1965. "A Theory of the Allocation of Time." *The Economic Journal* 75(299):493-517.
- Bramlett, M.D. and S.J. Blumberg. 2007. "Family Structure And Children's Physical And Mental Health." *Health Affairs* 26(2):549-558.
- Dawson, D.A. 1991. "Family Structure and Children's Health and Well-Being: Data from the 1988 National Health Interview Survey on Child Health " *Journal of Marriage and Family* 53(3):573-584.
- Francesconi, M., S.P. Jenkins, and T. Siedler. 2010. "The effect of lone motherhood on the smoking behavior of young adults." *Health Economics* 19(11):1377-1384.
- Garasky, S. and S. Stewart. 2007. "Evidence of the Effectiveness of Child Support and Visitation: Examining Food Insecurity among Children with Nonresident Fathers." *Journal of Family and Economic Issues* 28(1):105-121.

- Ginther, D. and R. Pollak. 2004. "Family structure and children's educational outcomes: Blended families, stylized facts, and descriptive regressions." *Demography* 41(4):671-696.
- Grossman, M. 1972. "On the Concept of Health Capital and the Demand for Health." *Journal of Political Economy* 80(2):223.
- . 2000. "The human capital model." in *The handbook of health economics*, edited by A. Culyer and J.P. Newhouse. Amsterdam: North Holland: Elsevier.
- Hamilton, B.E., J.A. Martin, and S.J. Ventura. 2011. "Births: Preliminary Data for 2010." in *National Vital Statistics Reports* 60, no. 2, edited by N.C.f.H. Statistics. Hayattsville, MD.
- Harknett, K. 2009. "Why are Children with Married Parents Healthier? The Case of Pediatric Asthma." *Population Research and Policy Review* 28(3):347-365.
- Heard, H., B. Gorman, and C. Kapinus. 2008. "Family Structure and Self-Rated Health in Adolescence and Young Adulthood." *Population Research and Policy Review* 27(6):773-797.
- Hill, M.S., W.-J.J. Yeung, and G.J. Duncan. 2001. "Childhood family structure and young adult behaviors." *Journal of Population Economics* 14(2):271-299.
- Kapteyn, A., P.-C. Michaud, J. Smith, and A. Van Soest. 2006. "Effects of Attrition and Non-Response in the Health and Retirement Study. ." *RAND Working Paper* WR-407.
- Knox, V.W. 1996. "The Effects of Child Support Payments on Developmental Outcomes for Elementary School-Age Children." *The Journal of Human Resources* 31(4):816-840.
- Kreider, R.M. and R. Ellis. 2011. "Living arrangements of children: 2009." *Current population reports*, P70-126.
- Krein, S.F. and A.H. Beller. 1988. "Educational Attainment of Children From Single-Parent Families: Differences by Exposure, Gender, and Race." *Demography* 25(2):221-234.
- Langton, C.E. and L.M. Berger. 2011. "Family Structure and Adolescent Physical Health, Behavior, and Emotional Well-Being." *Social Service Review* 85(3):323-357.
- Nepomnyaschy, L. and I. Garfinkel. 2011. "Fathers' Involvement with Their Nonresident Children and Material Hardship." *Soc Serv Rev* 85(1):3-38.
- Sigal, R.J., G.P. Kenny, D.H. Wasserman, C. Castaneda-Sceppa, and R.D. White. 2006. "Physical Activity/Exercise and Type 2 Diabetes." *Diabetes Care* 29(6):1433-1438.
- Stewart, S.D. and C.L. Menning. 2009. "Family Structure, Nonresident Father Involvement, and Adolescent Eating Patterns." *Journal of Adolescent Health* 45(2):193-201.
- Wing, R.R., R. Koeske, L.H. Epstein, M.P. Nowalk, W. Gooding, and D. Becker. 1987. "Long-term Effects of Modest Weight Loss in Type II Diabetic Patients." *Arch Intern Med* 147(10):1749-1753.

Chapter 2

Health Investment Decisions in Response to Diabetes Information in Older Americans

Diabetes is a very common and serious chronic disease, and one of the fastest growing disease burdens in the United States. Further, health behaviors, such as exercise, smoking, drinking, as well as weight status, are instrumental to diabetes management and the reduction of its medical consequences. Nine waves of the Health and Retirement Study are used to model the role of a recent diabetes diagnosis and medication on present and subsequent weight status, exercise, drinking and smoking activity. Several non-linear dynamic population average probit models are estimated. Results suggest that compared to non-diagnosed individuals at risk for high blood sugar, diagnosed diabetics respond initially in terms of increasing exercise, losing weight, and curbing smoking and drinking behavior, but the effect diminishes after diagnosis. Evidence of recidivism is also found in these outcomes, especially weight status and physical activity, suggesting that some behavioral responses to diabetes may be short-lived.

2.1. Introduction

Diabetes is a collection of diseases¹ that lead to elevated blood sugar levels (hyperglycemia). If untreated, diabetes can lead to complications including nerve damage, blindness, limb amputation, and kidney damage. Diabetics are also more than twice as likely to suffer from heart attacks or strokes. In some cases, diabetes is treatable through lifestyle modifications, without the need for medications. Other chronic diseases, such as high blood pressure, arthritis, and heart disease (among others), can be mitigated by lifestyle modifications

¹ There are three major types of diabetes: type I, type II, and gestational. Type I diabetes is an auto-immune condition by which the body attacks the pancreas, leading to an absolute deficiency of insulin, the major hormone that lowers blood sugar. Its onset is generally before age 30. Type II diabetes is the most common form of diabetes (~90-95% of all diabetes cases) that is characterized by insulin resistance and can often be mitigated through diet and lifestyle changes. Gestational diabetes is a transient condition that affects women during pregnancy, though it can raise the risk of developing type II diabetes. This paper focuses on Type II diabetes, since new cases of Type I or gestational diabetes are extremely unlikely to appear later in life.

as well. However, given the particular risk of long-term complications in the case of diabetes, the cost of a poor lifestyle for diabetics is often higher than for sufferers of these other chronic diseases. This characterization is even more true for older Americans, a population whose disease prevalence is higher (CDC, 2007), and whose ability to manage the disease through lifestyle changes (especially through physical activity) is often much more difficult (e.g. Sinclair, Conroy and Bayer 2008).

For the newly diagnosed diabetic, initial medical recommendations would commonly include lowering carbohydrate and fat intakes, increasing exercise, quitting smoking, losing weight, and decreasing alcohol consumption. Unfortunately, only about 45% of diabetics follow recommended care guidelines (McGlynn et al. 2003). The cost of compliance has declined over time due to technologies such as artificial sweeteners, and the benefits have risen over time due to medical advances that can subvert or delay complications. Nevertheless, there have been several explanations put forth as to why people might engage in behaviors that have substantial future health consequences (such as high carbohydrate consumption in diabetics).

Becker and Murphy (1988) suggest that people weigh the total costs and benefits of an addictive behavior and initiate behaviors in a forward-looking, time-consistent manner. A diabetic, for example, might indulge in a slice of cheesecake but commit to going to the gym later that day. Likewise, diabetic patients who anticipate increasing treatment availability and efficacy in the future might be less likely to be compliant with lifestyle recommendations. Thaler and Shefrin (1981) suggest that an individual faces a competition between far-sighted desire to obtain better health and a near-sighted desire for gratification. Cognitive functioning, especially in regard to executive functioning tasks, has been shown to be worse among older diabetics as compared to nondiabetic controls (Yeung, Fischer and Dixon 2009). Impairments in executive

functioning may obfuscate the costs of noncompliance with lifestyle habits, relative to the immediate benefits. More recent evidence indicates that male diabetics may have deficits associated with the prefrontal cortex that can lead to losses in inhibitory control (Ishizawa et al. 2010).

Even though compliance with recommended guidelines is sometimes absent, there is strong cross-section and pooled longitudinal evidence to suggest that individuals respond (at least somewhat) to health changes. For example, Kahn (1999) employs National Health and Nutrition Evaluation Survey (NHANES) data to examine the differential effects of diabetes education and diagnosis in determining own-health investment behaviors such as exercise, dietary compliance, and overall calorie intake. Kahn (1999) finds that compared to people with diabetes who are unaware of their condition, diagnosed diabetics in general invest more in their own health by smoking less and adhering to dietary guidelines surrounding diabetes management. Education about self-management behaviors further encouraged positive investment outcomes (Kahn, 1998). Additionally, Keenan (2009) uses pooled Health and Retirement Survey (HRS) longitudinal data to estimate cross-section “change” regressions, where the difference in smoking status (i.e. quitting) among previous smokers and change in body mass index (BMI) among overweight individuals are outcome variables regressed on a new diagnosis of a health condition, such as lung disease, heart disease, and diabetes. Keenan finds that while a new diabetes diagnosis does increase a smoker’s propensity to quit, a new diabetes diagnosis was smaller in magnitude and significance compared to other chronic conditions. However, these studies do not take into account methodological features such as state-dependence or the dynamic or long-term nature of health behaviors underlying diabetes compliance, which are processes that influence individual decision-making in health and labor

domains (e.g. Contoyannis, Jones and Rice 2004b; Hernández-Quevedo, Jones and Rice 2008; Slade 1987)

State dependence can be thought of as the role of the previous period's behavior on current behavior. As an example from the Health and Retirement Study (HRS), the correlation between whether a respondent currently smokes any cigarettes in a two-year period t and the previous period, $t-1$, is 0.84. Similarly, whether an individual ever drinks any alcoholic beverages in a two-year period has a 0.70 correlation from one period to the next. Obesity, or having a body mass index (BMI, an indicator of height for weight) over 30 at a survey interview, has a correlation of 0.76 between one period and the next. While these correlations may appear small relative to other studies, the two-year gap between HRS survey periods allows processes that are not captured by the lagged outcome to influence the contemporaneous outcomes. Self-reported smoking or drinking behavior may also be measured with error. These relatively high correlations reflect not only the importance of previous behavior in explaining present behavior, but also the likely difficulty in changing these behaviors.

Similarly, recidivism – individuals falling back on poor health habits (e.g. smoking, alcohol use, and low levels of exercise) after making behavioral modifications, is an important issue from a policy and clinical standpoint. Adhering to scheduled provider appointments, which is often diabetes control is assessed, is also a problem among diabetic patients (Schectman, Schorling and Voss 2008). As a result, providers for diabetic patients often make it a priority to help patients not only to change lifestyle behaviors, but to maintain these changes, as well (Anderson and Funnell 2000).

As a result of the overwhelming medical evidence that exists about the mitigating effects of a healthy lifestyle on the consequences of diabetes, this paper aims to study four behavioral “flows” that can ameliorate the deleterious effects of diabetes and improve the prognosis of diabetics: smoking, alcoholic beverage consumption, frequent exercise, and body mass. These “flows” can affect health stocks relevant to diabetics, such as long-term blood sugar levels. The most important flows in the case of glycemic control in diabetes would include diet and physical activity, which can directly affect blood sugars. However, there are flows that have indirect benefits as well, such as curtailing smoking and drinking behavior. Current body weight, which is not directly a “flow” to glycemic control, is important to consider, as weight loss can improve insulin sensitivity and potentially mitigate reliance on medication. Body weight can also serve as proxy for energy intake versus expenditure, albeit crudely. However, given the extremely powerful biological ties between body weight and diabetes, I use overweight or obese status (corresponding to a $BMI \geq 25$) as one of the four outcome measures in this analysis.

In this study, reduced form demand equations are estimated for these health behaviors and weight status. I consider any smoking, frequent physical activity (at least 3 times per week), any alcoholic beverage consumption, and overweight or obese status as discrete outcomes when estimating their demands, because the risk for developing diabetic complications are often mitigated by achieving these particular threshold values. For instance, evidence from clinical trials suggests that exercise at least three times a week can significantly reduce long-term blood sugar levels and rates of microvascular complications (Boulé et al. 2001; Sigal et al. 2006). Similarly, achieving and maintaining weight goals is a major component of diabetes management (American Diabetes Association 2008). Evidence suggests that diabetics require at least a 5% reduction in body weight in order to achieve any significant

glycemic improvement among type II diabetics (Wing et al. 1987). Further, completely quitting smoking is particularly beneficial for decreasing the risk of vascular problems that can often accompany diabetes.

The costs of smoking in the context of diabetes are high. In addition to serving as an independent risk factor for the development of diabetes, smoking also tends to increase the risk of complications of diabetes, including heart and vascular disease. Drinking behavior is not as well characterized in the literature, though it is an important aspect of diabetes self-management. Because the effects of alcohol on blood sugar are often difficult to predict, glycemic control is often more difficult for diabetics who choose to drink. However, the clinical evidence on alcohol and diabetes is somewhat mixed. While the beneficial effects of some alcohol consumption on cardiovascular disease are well characterized even among diabetics (Howard, Arnsten and Gourevitch 2004), the American Diabetes Association currently recommends keeping alcohol consumption to a minimum, with an maximum daily consumption of one drink for women and two drinks per men (Wheeler, Franz and Froehlich 2004).

This study makes contributions on the role of diabetes information in health behavior modification and health outcomes. Specifically, this paper presents a dynamic framework for understanding the role of a recent diabetes diagnosis as a “shock” for people to initiate health behavior, and how diagnostic information persists subsequent to the initial diagnosis. I also consider the role of anti-diabetic medication in the evolution of these health outcomes, which can physiologically affect the outcomes themselves, as well as serve as a proxy for the current clinical stage of the diabetes.

2.2. Review of the Role of Information of Health Investment Behavior

Much of the existing literature on the effects of health shocks on behavior has focused on “primary preventative behaviors,” or actions that reduce the probability of contracting a particular disease, such as flu vaccinations (Mullahy 1999). This paper focuses on secondary and tertiary preventative behaviors, or behaviors that lower the health risk of a disease and reduce the likelihood of complications or disability resulting from the disease.

A newly diagnosed diabetic who maximizes his expected utility will face a tradeoff between immediate gratification and the benefits of good future health. Since this study focuses on elderly Americans, the question of whether future improved health is worth the costs of increased exercise and dietary compliance is particularly salient, since some individuals might feel like their longevity is limited.

Studies throughout various fields in economics have found a significant link between health information and behavior. Health information is generally collected by using respondents’ answers to a series of questions regarding medical facts of disease. Several studies have estimated demand systems conditional on health information. Carlson and Gould (1994) found that household meal planners had substantially different consumption of fats and oils given increased health awareness about cholesterol and fat. Similarly, Chern et al. (1995) used information from the Health and Diet Survey (HDS) to develop Bayesian-based measures of health information (specifically knowledge of the link between diet and disease). They estimated a demand model and demonstrated that increased information led to an increased consumption of corn, cottonseed, and soybean oils, but decreased consumption of lard and butter.

With regard to health services, Hsieh and Lin (1997) use national survey data of elderly individuals in Taiwan to estimate demand for preventive care among the elderly. They find that more and better health information increases the rate at which the elderly consume preventive care. Kenkel (1991) makes the important distinction between education and health information. He finds that after controlling for health information, the positive effects of schooling on health investment behavior (especially cigarette smoking) still remain.

Kahn (1998) reports that smoking quit rates are quite high among people recently diagnosed with diabetes. Among recently diagnosed diabetics with less than nine years of education, about 53% quit smoking within one year of diagnosis. Among college graduates, about 58% quit after a year of diagnosis. However, the effect of education on health investment behaviors such as smoking and diet is more pronounced among non-diabetic than among diabetic individuals.

Benítez-Silva and Ni (2008) use an empirical framework that assesses the role of changes in self-reported health and new disease diagnoses on subjective perceptions of longevity. They find that trends in subjective longevity follow patterns that would be expected of the Grossman (1972) model, in that a negative shock to health status in one period leads to updated beliefs that reduce the subjective probability of longevity (measured as the subjective probability of living until 75). They find this is true for a new diagnosis of most diseases, though diabetes has a positive effect on this probability. However, the authors suggest that access to treatment subsequent to receiving a diagnosis may substantially raise one's perception of longevity in the case of diabetes.

Sloan et al. (2009) use the 2003 Diabetes Supplement to the Health and Retirement Study (HRS) to discuss potential reasons behind differential adherence to recommended self-care guidelines. They find that perceived control over life events, as well as the subjective probability of being alive in ten years, are two important factors that contribute to one's self-management activities. Klick and Stratmann (2007) use Behavioral Risk Factor Surveillance System (BRFSS) data to investigate the health effects of diabetes treatment mandates in various states. They found that there exists a moral hazard by which treatment mandates can be viewed as substitutes for self-care (e.g. weight loss).

2.3. Data and Empirical Strategy

I use the 1992 – 2008 biennial waves of the RAND-enhanced Health and Retirement Study (RAND HRS), a survey that broadly covers mental and physical health, and labor outcomes of older Americans (over 50). In addition to the anthropometric measures of height and weight, which can be used to calculate body mass index, the HRS collects a wide array of information on preventative and risky behaviors: such as smoking, alcohol consumption, and physical activity². Each period, individuals are asked if they currently have diabetes, as well as any medications they are taking to treat the disease. Similar questions are asked for other chronic diseases such as hypertension, heart disease, lung disease, and cerebrovascular disease (i.e. stroke). Demographic information, such as marital status, gender, age and race, is also surveyed. Though spouses are included in the overall HRS, I omit spouses from my analyses. The HRS study design oversamples Black and Hispanic populations as well as Florida residents, and requires weighting in order to generalize the results to the population at large. Furthermore,

² Because of question differences across survey waves, I only include three waves of data from 2004-2008 in the exercise specifications.

survey respondents may drop out of the survey through death (i.e. passive attrition) or loss to follow-up (i.e. active attrition), which has particular relevance to the analysis conducted here.

In the following subsections, problem of misreported and undiagnosed diabetes that is common in older populations is discussed, along with the proposed reduced form econometric model. This is followed by a discussion of the construction of a comparison group to serve as an “at-risk nondiabetic” counterfactual to diagnosed diabetics. Finally, a weighting scheme to account for attrition and oversampling among HRS respondents, and the use of Average Marginal Effects in my analyses, is discussed.

2.3.1. Undiagnosed and Misreported Diabetes

Since type II diabetes often develops insidiously and “quietly” over time, there is a potential that many individuals are unaware of their elevated blood glucose levels. Indeed, it has been estimated that subclinical or undiagnosed diabetes can often precede diagnosis for more than 10 years, and that organ damage (such as retinopathy) can occur during the pre-diagnosis period (Harris et al. 1992). Furthermore, recent analyses using NHANES data estimates that of the approximately 30% of U.S. adults between the ages of 60 and 74 who have diabetes, more than 40% are undiagnosed, but have at least one laboratory assay that is suggestive of the disease (Cowie et al. 2009). Cowie et al. suggests that Hispanics and males have higher rates of undiagnosed diabetes. Older age also raises the propensity for undiagnosed diabetes. However, other demographic characteristics, such as education and income, were not found to be associated with diabetes diagnostic status (Wilder et al. 2005).

One of the most common laboratory tests used to measure an individuals’ control of their blood sugar is the Hemoglobin A1c (HbA1c). A HbA1c measurement is a 2-3 month average of

blood sugar that is an important prognostic indicator for future complications. While previously restricted to the diabetic population, this test has gained credence as a screening tool for diabetes (e.g. Bennett, Guo and Dharmage 2007; Rohlfing et al. 2000), and in 2009 was adopted as a diagnostic test by the American Diabetes Association and the International Expert Committee (2009)³.

Because patients with certain characteristics may have clinical diabetes longer before diagnosis compared to other patients, the effect of a new diabetes diagnosis may be somewhat different if a patient is diagnosed late in the clinical disease. For instance, at later points in the disease, the risk of complications might be much higher, making nonadherence to self-management behaviors very costly. In the analysis, since the clinical latency period is unobserved, and because it is difficult to ascertain how late an individual is diagnosed based on general health, demographic, and socioeconomic observable characteristics alone, medication status (oral medication, insulin use, or both) at the time of the diabetes diagnosis is used to gauge the approximate “timing” of the diagnosis relative to the clinical disease.

Diabetes treatment algorithms have remained fairly constant over the past two decades, and generally involve lifestyle (i.e. diet and exercise) changes, and oral antidiabetic medications like metformin (Nathan et al. 2006). Insulin is added if oral agents are insufficient to control blood sugar. If the hyperglycemia is severe enough at the time of diagnosis, insulin may be initially prescribed (though this is rare). Likewise, if there is very moderate hyperglycemia, no medication may be prescribed. In this way, medication status can be used as an indicator for the approximate severity of the disease at diagnosis. Since diabetes generally progresses slowly over time, increases in medication use at diagnosis can serve as a proxy for if the disease is diagnosed

³ The American Diabetes Association recommends a cutoff value of HbA1c at or above 6.5% to warrant a diagnosis of diabetes, which corresponds to an increased risk of retinopathy. Further, HbA1c values between 5.7% and 6.4% are often considered evidence for pre-diabetes. Rohlfing et al. (2000) recommends a cutoff HbA1c of 6.1%

at a later or earlier stage in its clinical evolution. Therefore, in my analyses, I treat diabetes as a spectrum of diseases; with individuals without any anti-diabetic medication (or insulin) being “mild” cases of diabetes, individuals on oral medication as “moderate” severity, and insulin users as the most severe cases. This assumption is reasonable given the individuals who had their HbA1c measured in the 2006 HRS Biomarker data, in which the average HbA1c among diagnosed diabetics not taking medication is 6.2%. Diabetics taking oral medication have an average HbA1c of 6.8%, while insulin users had an average HbA1c of 7.5%.

Finally, it is important to recognize that some individuals who have been diagnosed with diabetes may misreport their diabetes status on surveys (Goldman et al. 2003; Kriegsman et al. 1996; Mackenbach, Looman and van der Meer 1996; Martin et al. 2000; Shah and Manuel 2008). Diagnosed diabetics may misreport their diabetes status for a variety of reasons, including: stigma surrounding the disease, misunderstanding the terms used by the survey, or disagreement over the diagnosis, among others. Studies have demonstrated that higher levels of accuracy among self-reported diabetes status occurs in individuals with higher incomes, higher levels of education, and among white individuals. Individuals who visit the doctor more frequently, or who have co-morbid chronic conditions are also more likely to accurately self-report their diabetes status (Goldman et al. 2003; Mackenbach et al. 1996; Shah and Manuel 2008). It has also been suggested that diagnosed individuals who self-report their diagnoses tend to have better glycemic control, and are more likely to engage in self-management habits, such as glucose monitoring (Garay-Sevilla et al. 1999; Shah and Manuel 2008).

2.3.2. Econometric Model

At each wave of the HRS, an individual can be classified as overweight or obese, choose to engage in any smoking or drinking behavior, or engage in frequent (i.e. 3 or more times per week) physical activity (these can all be represented by a vector of outcomes, d_{it}). Though there are many such behaviors, I will be focusing on these four discrete outcomes because they have substantial potential to normalize blood sugar and reduce the rate of complications. For explanatory variables, there will be several variables indicating if there was a new diabetes diagnosis in the current period or in previous periods. Further, there will be a binary variable if the individual is currently diabetic (whether they are newly diagnosed or not). Current medication status is included as well. Finally, non-diabetes health status, insurance and demographic characteristics are included as covariates as well (discussed below). The full model takes the form shown in Equation 1.

$$(1) \quad d_{it} = \text{Diabetes}_{it}\beta + \phi\Delta\text{Diabetes}_{i,t,t-1} + \delta d_{i,t-1} + \gamma x_{it} + \varepsilon_{it}$$

In this context, the first term represents a vector of three variables capturing the clinical stages of diabetes as proxied by medication status for an individual at time t . These variables represent diabetics taking no medication, diabetics taking oral medication, and diabetics taking insulin. The second term represents if an individual acquired a new diagnosis of diabetes in period t . Individuals who are newly diagnosed are divided into two categories: those not taking any medication or insulin at the time of diagnosis (about 1/3 of new diagnoses), and those requiring oral medication or insulin at the time of diagnosis (about 2/3 of new diagnoses). Newly diagnosed diabetics requiring oral medication are grouped together with newly diagnosed insulin users because very few recently diagnosed diabetics report requiring insulin for treatment. Those not taking any anti-diabetic medication are considered as having been diagnosed early in the

clinical stage, while those requiring medication are considered as having been diagnosed at a later period in the clinical evolution of the disease. While this is not a perfect measure of the clinical stage, medication use does correlate well with an individual's glycemic control as measured by the HbA1c (see above). For both groups of newly diagnosed diabetics, lagged measures of diabetes diagnosis are added, up to 14 years, to capture both the long- and short-term effects of a new diagnosis. As discussed above, it is conceivable that some individuals who in fact have received a diagnosis may misreport their status. As a result of this misclassification, the estimates derived from this model represent a lower bound on the effect of a diabetes diagnosis on the outcomes of interest.

The third term incorporates dynamics into the model through the addition of a lagged dependent variable, to model the evolution of smoking, drinking and exercise behavior, along with body mass. Finally, the fourth term represents time-invariant and contemporaneous covariates such as age, race, marital status, gender, education, U.S. birth status, ethnicity, total income, total assets, number of doctor visits, and insurance status. Current and newly diagnosed cases of other chronic diseases, such as hypertension, heart disease, lung disease, and cerebrovascular disease (stroke) are included in x_{it} . Regional and time dummy variables are included in all models.

While the most efficient estimator to estimate the reduced form model specified in Equation 1 would be a nonlinear (e.g. probit) random-effects model incorporating individual-level unobserved heterogeneity, the use of such a model would make strong assumptions on the time-invariant error term that may not provide a good approximation to the data-generating process in a panel spanning almost 20 years, as is the case here. In this case, inconsistent estimates may be obtained. However, completely ignoring the correlation structure of the data, as

in a pooled model, would neglect important aspects of the data and provide inconsistent estimates as well. To incorporate a correlation structure among clustered data a population average probit model is estimated, which will provide consistent inference statistics (assuming the assumed correlation structure reflects the true structure). In addition to accounting for a correlation structure of the data, the use of a population average estimator offers several advantages in this context. A population average model more easily allows for the inclusion of sampling weights that address individual propensities of attrition and oversampling. Additionally, employing a population average model avoids the initial conditions problem found in dynamic random effects models, in which the initial unobserved values of the outcome are correlated with the individual heterogeneity of the error term.

The population average estimator is estimated using the Generalized Estimating Equation (GEE) method⁴, which is an extension of the Generalized Linear Model for clustered data (with panel data being a special case). For a detailed description of the GEE methodology, see Fitzmaurice, Laird, and Ware (2004) and Zeger, Liang and Albert (1988). In brief, GEE solves the adjusted score equation $\sum_{i=1}^N \mathbf{D}_i' \mathbf{V}_i^{-1}(\alpha) (\mathbf{Y}_i - \boldsymbol{\mu}_i) = 0$, in which \mathbf{D} is a matrix of derivatives of the conditional means $\boldsymbol{\mu}_i$ with respect to β . Here, $\mathbf{V}_i(\alpha) = \mathbf{A}_i^{1/2} \mathbf{R}_i(\alpha) \mathbf{A}_i^{1/2}$, where, \mathbf{A}_i is the diagonal variance matrices for these parameters, and $\mathbf{R}_i(\alpha)$ represents the user-specified working correlation matrix for repeated observations from the same individual. In this study, I assume a correlation structure of exchangeability, which implies that the correlation of all observations per individual is identical. This implies that the matrix $\mathbf{R}_i(\alpha)$ will have ones along the diagonal, and α (the correlation) everywhere else.

⁴ Estimations were performed using the xtprobit, pa command in Stata MP Version 12.0.

Because the Population Average/GEE framework does not explicitly model individual heterogeneity in the non-linear context, the interpretation of estimates derived from these models has a subtle difference compared to those derived from models directly incorporating individual heterogeneity. As a result, population average coefficients (and the Average Marginal Effects derived therefrom) are interpreted as effects for the population at large, rather than for a specific individual.

2.3.3. Construction of Comparison Group

The comparison group for diabetics (both newly diagnosed and otherwise) is not necessarily all non-diabetics, but rather those who could become diabetic during their lifetime. This includes individuals who ever receive a diagnosis in the HRS, as well as individuals who never receive a diagnosis, but whose blood sugar reaches a level that warrants a diabetes diagnosis or comes fairly close to this level (i.e. pre-diabetes).

In order to create a comparison group that satisfies these properties, I estimate a basic probit model for diabetes risk using the 2006 HRS Biomarker Data, which consists of blood measurements analyzed for a random subsample of the 2006 HRS respondents. These data include about 6,500 individuals who have reported laboratory values for the HbA1c assay. In this model, the outcome is having a $HbA1C \geq 5.7\%$, which corresponds to having glucose metabolism impaired enough to be classified as a “pre-diabetic.” Included in this model is gender, race, ethnicity, education, U.S. birth status, as well as 2006 measures of income, assets, body mass index, self-reported health status, and diagnoses of high blood pressure, stroke, heart disease, or lung disease. From the results of this model, each individual in the RAND HRS is assigned a predicted probability (propensity score) for high blood sugar. I make the assumption

that once an individual is statistically “at-risk” for developing diabetes or pre-diabetes, they are “at-risk” for their entire time in the HRS. Therefore, of the propensity scores observed for each individual, the maximum propensity score is used to classify individuals into the comparison group.

A cut-point value, such that individuals at or exceeding the cut-point would be classified as “at risk,” is obtained by maximizing the sum of sensitivity and specificity, minus one (also known as the Youden Index) (Fluss, Faraggi and Reiser 2005; Klotsche et al. 2009; Schisterman et al. 2005; Youden 1950). This value is optimal in the sense that it maximizes the discriminating ability of the propensity score when the errors associated with sensitivity and specificity are given equal weight. To calculate this value, I use kernel density smoothing to estimate the distribution of the propensity scores for individuals in the 2006 HRS Biomarker Data with HbA1c values considered normal and elevated. The propensity score that gave the Youden Index was 0.45⁵. Thus, an individual is considered “at-risk” if his maximum propensity score is 0.45 or above.

There is considerable overlap between the constructed group and diagnosed diabetics, with about 35% of those within the imputed “at-risk” group actually being diagnosed at some point with diabetes. The final analytic sample combines both groups: individuals who are ever diagnosed with diabetes in their time in the HRS, and individuals who are inferred to be at risk based on them having a high imputed probability of high blood sugar at some point in their evolution in the HRS. The final sample includes about 63% of the original HRS sample, but is

⁵ Among the individuals with reported HbA1c values in the 2006 HRS Biomarker Data, this propensity score scheme correctly classifies about 63% of individuals with elevated HbA1c values, with a sensitivity of 68% and specificity of 59%. The unobserved components in determining blood sugar, most notably diet and genetic factors, make obtaining more predictive values using propensity scores difficult.

consistent with the high risk for diabetes and pre-diabetes that has been noted in older individuals.

2.3.4. Sampling Design, Attrition, and Inverse Probability Weighting

Two important practical features of the HRS data, oversampling and attrition, necessitates weighting of observations in order to derive correct inference for the older American population at large. Furthermore, given that this study focuses on behavioral responses to diabetes, the analysis requires attention to the issue of non-random attrition, and health-related attrition in particular. Further, the HRS oversamples minority populations and residents of Florida. Accounting for these features of the data requires the construction of sampling weights to give more importance to individuals who were undersampled or who were more likely to attrite from the survey over time (Fitzgerald, Gottschalk and Moffitt 1998). The HRS provides sampling weights for each wave, which represents the inverse of the probability that an individual is in a given wave after controlling for four demographic characteristics: birth cohort, race, ethnicity, and gender. Additionally, this weight is adjusted using the Current Population Survey (CPS) during the year of the interview for those four characteristics. The HRS sampling weights therefore account for attrition and oversampling, assuming that attrition does not affect the CPS, and that attrition is random, conditional on the four demographic variables (known as the missing at random assumption, see Little and Rubin (1987)). However, these weights do not capture health-related attrition, which is of critical importance to the longitudinal analysis of health behaviors as a result of diabetes information. That is, the missing at random assumption can incorporate a more extensive set of observable baseline characteristics, including health-related measures.

To capture factors that are related to survey non-response (either through death or active non-response), for each period after an individual enters the HRS, a univariate probit model is estimated to predict response at a given wave (r_i) conditional on expanded observable characteristics observed when the respondent first entered the survey (x_{i1}). Active non-response and death are modeled identically in this context. While this scheme will account for attrition on observables, subject to the missing at random assumption, individuals also have differential baseline probabilities of selection into the HRS (i.e. minorities and Florida residents are oversampled). Following Kapteyn (2006), an inverse probability weight comparable to the HRS sampling weights can be constructed for each individual at each wave (except the first wave) by inverting the predicted probability of survey response from this model, and multiplying by the baseline HRS sampling weight, as in in Equation 2.

$$(2) \quad IPW: w_i(x) = \frac{w_1(q_i)}{p(r_i|x_{i1})}$$

The variables included in x_{i1} are largely the same as those used to predict hyperglycemia for the individuals status, self-reported whose HbA1c was measured in 2006, including time-invariant factors such as ethnicity, race, education, and U.S. birth status, and baseline values of age, body mass, asset level, household income, insurance health status, and chronic disease diagnoses. The baseline HRS sampling weight $w_1(q_i)$ adjusts for selection probability as well as attrition based on the four demographic characteristics listed above (q_i).

While observable characteristics observed at baseline do play a large role in predicting attrition in the HRS, there are variables that are observed after the baseline interview that can drive attrition, as well as unobserved factors that may lead to death or other non-response. With regard to the former issue, including lagged measures of time varying factors would confine the scope of the weighting to “monotone” attrition, in which respondents never re-enter the panel

subsequent to the period of non-response (Contoyannis, Jones and Rice 2004a). In the HRS, there is a substantial fraction of respondents who do not respond to a survey wave but respond in subsequent waves, making this strategy unattractive. With regard to unobservable characteristics driving attrition, weighting schemes accounting for unobserved heterogeneity generally require excluded variables that are related to attrition but unrelated to the outcome of interest. This is especially important in the case of attrition due to death, which may be influenced by unobserved characteristics of the outcome measures that are related to both death and a diabetes diagnosis. Such exclusion restrictions that meet these criteria are very difficult to identify, and in practice, many studies rely on an extensive vector of observable factors as well as lagged (or initial) measures of the dependent variable to correct models for attrition.

In this study, these weights are computed separately for each econometric model, with the initial (period 1) outcome variable for each model included in x_{i1} as appropriate. In the smoking, alcohol consumption, and weight status models, period 1 represents the first period that the individual is interviewed (which may not necessarily be 1992, or Wave 1). In the exercise specifications, period 1 represents Wave 7 (2004), since the measures of physical activity I use are only consistently measured from 2004-2008. In the descriptive analysis, I directly apply the HRS weights, since an IPW scheme is a model-specific correction. For econometric models, such as fixed-effects or population-averaged models, weights must be constant for all observations for each respondent. Following Mellor (2011), I employ the average sampling weight for each year they are in the sample for the specification of interest. In this way, longitudinal results can be adjusted for attrition and oversampling.

2.3.5. *Average Marginal Effects*

An Average Marginal Effect can be used to scale dichotomous coefficients in order to assess the magnitude of the effect of a new diabetes diagnosis on the outcome variables of interest (see, for example, Wooldridge (2005)). An Average Marginal Effect is computed by, for each observation, obtaining a predicted value that corresponds to the variable of interest (for instance, the new diabetes information) equaling 1, and obtaining the corresponding point on the normal distribution for this prediction. This value is subtracted from the point on the normal distribution that corresponds to the predicted value when the variable of interest equals zero. The Average Marginal Effects for each variable are computed by taking the mean of this difference across the distribution of the covariates and heterogeneity of the entire sample.

A “zero” for diabetes information reflects individuals who have not yet been diagnosed, or the constructed comparison group of people likely to be diabetic. The comparison between the diagnosis group and the “at risk” group is intuitive for contemporaneous variables of diabetes status. However, with regard to lagged measures of diabetes status (e.g. if someone received a diagnosis two years ago), the ideal comparison would be non-diabetic “at risk” individuals who would have been diagnosed two years ago. However, the chronic nature of diabetes makes it such that “at risk” individuals will likely become diabetic at some point. Therefore, this variable has an interpretation of the effect of being diagnosed two years ago versus non-diabetic “at risk” individuals who would eventually be diagnosed with diabetes or pre-diabetes.

Computing the Average Marginal Effect allows for a relatively simple interpretation in terms of the change in probability of alcohol use, smoking, exercise or weight based on the change in the explanatory variable (i.e., diabetes information) from 0 to 1. The values of this

marginal effect can be directly compared to linear coefficients as well. In the nonlinear models estimated, the contrasts of interest would be of a (diagnosed) diabetic with an individual who either has been diagnosed with diabetes at a future date, but not at the present time, or someone who is observationally likely to have high blood sugar measurements in their time in the HRS. I bootstrap the average marginal effects with 500 replications⁶, clustered at the individual level, to generate a standard error⁷.

2.4. Results and Discussion

2.4.1. Descriptive Analysis

Table 1 shows summary statistics of several demographic and health variables for all individuals in the sample in 1992, and for those who remained in the panel until 2008. Individuals at the start of the HRS survey in 1992 represent, by and large, typical older adults living in America. The rates of diseases that can be brought about by older age, such as heart problems, lung problems, cancer, and diabetes, all increased in prevalence over time, with the rate of heart disease more than doubling, and stroke tripling. Notably, the prevalence of diabetes in this group more than doubled as well, making almost 1 in 4 individuals who originally started the HRS in 1992 diabetic in 2008. The reliance on anti-diabetic treatments for diabetics (especially oral medication) also increased dramatically during this time, suggesting that the clinical course of diabetes was, on average, worse in 2008 than in 1992. Additionally, those who remained in the survey in 2008 were more likely to be female, white, have a higher household income, and greater household assets. Body mass index (BMI) remained fairly stable with only a

⁶ In some models, some bootstrapped parameters could not be estimated for all 500 replicates. In general, the number of replicates was above 490, and in no case was it below 374 total replicates.

⁷ In most models, compared to the bootstrapped standard error, the relative bias was small.

one unit increase, which may reflect an age-related decline, counteracted by an increasing secular trend, in body mass. Smoking and drinking behavior tended to decline substantially over time, as well. These statistics are consistent with some of the health consequences of aging.

Table 2 presents a simple probit model that identifies several risk factors for the development of hyperglycemia (as measured by the HbA1c) on a vector of observable characteristics. As expected, the most important risk factors for the development of higher blood sugar include age, body mass index, self-reported health, and the existence of comorbid chronic conditions like heart disease or hypertension. Race and ethnicity also had a strong role in predicting the risk for high blood sugar, with non-white individuals being much more likely to have elevated blood sugar than whites. While the simple model presented in Table 2 does identify important risk factors for the development of high blood sugar (and ultimately, diabetes), the ultimate goal of the model is to use the estimates to impute the probability of high blood sugar to generate an appropriate comparison group for newly diagnosed diabetics as described above, and shown in Table 3. It is important to note that these observable factors represent a very small fraction of the overall physiological determinants of blood sugar. Diet, physical activity (which is not consistently measured in the HRS), genetics, and medication compliance are other important drivers of blood sugar that are unobserved in the HRS. However, the existence of these strong observable risk factors allows for the identification of a sample of individuals who is likely to become diabetic (or pre-diabetic) at some point in their lives.

Table 3 shows descriptive statistics by various measures of diabetes status, including individuals never diagnosed with diabetes, people who ever received a diabetes diagnosis, newly (current period) diagnosed diabetics, all individuals currently diagnosed with diabetes, as well as non-diagnosed individuals who are likely to be at risk for diabetes, which is constructed as

described above. Average HbA1c results are reported at the bottom of each column for the 2006 respondents in the Biomarker data who fall into the respective category of diabetes status. The overall results displayed in Table 3 reflect the substantial heterogeneity of diabetics, though there are some common threads between the groups. In general, diabetics (and people who are at risk for diabetes) tend to have higher body mass and rates of obesity, tend to be older, non-white, have a lower level of education and income than those who have never been diagnosed with diabetes (either prior to or while they were in the HRS). Diagnosed diabetics (Column 5) also tend to have higher rates of co-morbid chronic diseases, especially hypertension, heart disease, and cerebrovascular disease.

Columns 3 and 4 of Table 3 show a comparison of individuals who were diagnosed while in the HRS before and after they received a diabetes diagnosis, respectively. Overall, post-diagnosis, individuals are likely to see the doctor more frequently, have health insurance, and become diagnosed with another chronic disease (especially hypertension). After a diagnosis, individuals are likely to smoke and drink less and exercise slightly more, though the exercise changes are relatively small. However, individuals do not seem to lose a substantial amount of weight after diagnosis, with the average BMI (and overweight or obese indicator) being higher post-diagnosis. Also of interest are the characteristics of individuals who were newly diagnosed with diabetes, which are depicted in Columns 6 and 7. As discussed above, the clinical manifestation of diabetes often takes an insidious course that results in individuals being diagnosed at differential stages of the disease. I use medication status at diagnosis to approximate whether an individual was diagnosed “early” (diagnosed without medication) or “later” (diagnosed with medication or insulin). Overall, those diagnosed without medication tend to be white, have a higher level of education and income, and see the doctor less frequently.

Individuals diagnosed with medication drank less, but weighed slightly more than those diagnosed without medication.

The last column of Table 3 provides descriptive statistics on individuals who are statistically likely to have elevated HbA1c values, but are not diagnosed. Overall, this group is more likely to be older, female, unmarried, have higher rates of overweight and obesity, and lower incomes than the non-diabetic group. These characterizations are largely similar to the diagnosed groups as well, with the exception of fewer doctor visits, which may justify why this group is undiagnosed despite being likely to have high blood sugar. It should be noted that this group of people is a subset of the whole population based on a propensity score; therefore, there are people in this sample who may not develop diabetes in their lifetime (in the HRS). However, the general characteristics of ethnicity (minority status is an independent risk factor for diabetes), elevated body weight, and existence of comorbid conditions combines some of the most important risk factors for diabetes. In the econometric models, this group, combined with individuals who have yet to be diagnosed (Column 3) serve as the comparison group to the diagnosed individuals.

Because diabetes is a chronic condition, individuals must adhere to treatments (including medication and lifestyle changes) over time to maintain proper glycemic control and to reduce the risk of complications. From a purely biological point of view, the longer high blood sugar concentrations are present in the bloodstream, the higher the probability of organ damage over time. The HRS data allows for the construction of the duration of time since an individual was diagnosed with diabetes, given that the diagnosis occurred any time in between 1992 and 2008. However, there is no information on whether the years with diabetes are well controlled or not. This is important because an additional year of diabetes is more meaningful for individuals with

poor control than for individuals with strong control. However, on average, it is safe to assume that each additional year of having diabetes marginally increases the risk for complications, suggesting that negative health behaviors should be declining as the duration with diabetes increases.

The behavioral response trajectories over time for individuals diagnosed with and without medication are shown Fig. 1 which provides unadjusted (except for sampling weights) averages of the four outcomes I consider: smoking (a), drinking (b), frequent exercise (c) and overweight or obese status (d). The averages for each outcome are grouped by medication status, in order to give a rough indication as to whether an individual was diagnosed during an early or later part of the clinical disease. For all four outcomes, people respond to a new diabetes diagnosis (year 0) by smoking and drinking less, losing weight, and exercising more compared to the prior period (year -2). With regard to smoking, the trajectory is similar for those diagnosed with and without medication, though those diagnosed without medication tend to smoke less than those diagnosed with medication. For all diagnosed diabetics, smoking levels decrease as time elapses since diagnosis, suggesting that diabetics tend to adhere to smoking guidelines over long periods of time. The continuous decline of any alcoholic beverage consumption for all diabetics is similar to smoking patterns seen in Fig. 1a, though individuals diagnosed with medication tend to consistently drink less than those diagnosed without medication. Exercise and weight status do not follow the same optimistic patterns as seen with smoking and drinking behavior. At the time of initial diagnosis, individuals tend to lose weight and increase their level of exercise. Subsequent to the initial diagnosis, however, exercise activity wanes, and individuals tend to gain weight. Some of the weight gain may be attributable to the physiological activities of medication itself. For instance, insulin is known to promote weight gain (e.g. Mäkimattila,

Nikkilä and Yki-Järvinen 1999), whereas certain oral medications, like metformin, can promote weight loss. However, from Table 3, Column 7, only about 5% of the diabetics diagnosed with medication were prescribed insulin therapy, making it unlikely that a substantial part of the difference in trajectory is due to the effects of medication alone.

2.4.2. The Role of Diabetes Information and Medication: Smoking and Drinking Behavior

I estimate several dynamic population average probit models that consider the trajectory of smoking or alcoholic beverage consumption in individuals after they are diagnosed with diabetes. Several comparison models, as described below, are presented along with the dynamic population average specifications. The analytic sample used includes individuals who have ever been diagnosed with diabetes, along with a group of individuals who were never diagnosed with diabetes, but are statistically likely to develop high blood sugar over their lifetimes. Therefore, individuals who have yet to be diagnosed, along with never-diagnosed but “at risk” individuals comprise the comparison group to diagnosed diabetics.

Table 4 presents Average Marginal Effects for a recent diabetes diagnosis with and without medication, and the long-term effects of such a diagnosis. This effect is considered for the entire sample of individuals “at risk” for diabetes and pre-diabetes, including all individuals who were ever diagnosed with diabetes, as reported in the HRS (Panel A), along with the male, non-white, and younger (<60 years old when they entered HRS) subsamples of this overall “at risk” group (Panels B, C, and D, respectively). In Panel A, there is a marginally significant average “shock effect” for individuals who were diagnosed without medication of about 1.7 percentage points. Receiving a new diagnosis of diabetes with medication lowers the probability of smoking by about 1.5 percentage points compared to individuals who are undiagnosed but “at

risk”, but this effect wanes in magnitude and significance as an individual is farther away from the original diagnosis. However, the insignificant effects over time are either very close to zero or negative, suggesting that there is no net increase in smoking propensity in years subsequent to the diagnosis.

Fig. 2 shows the predicted trajectory of smoking behavior among individuals who were diagnosed with and without medication. The probabilities of smoking were predicted from the dynamic, population average probit model reported in Table 4, Panel A, Column 1. The probabilities were then averaged among the sample of diabetics according to their time since diagnosis. Fig. 2 reports the unconditional probability of smoking (a), the quit probability (b), and the start probability (c)⁸. The means of these probabilities (i.e., the point estimates) are presented with 95% confidence intervals. In each figure, a horizontal line is drawn to represent the average of the respective predicted probabilities for individuals in the comparison group. It is valuable to consider behavioral changes in either direction due to the possibility that the changes that newly diagnosed diabetics make are not sufficient to mitigate the consequences of diabetes. The predicted probabilities in Fig. 2a generally parallel the descriptive means presented in Fig. 1a, with individuals tending to curb smoking behavior after a new diagnosis, and continuing to do so subsequent to the diagnosis. The quit probability is depicted in Fig. 2b, where there is a rise in the probability of quitting during the diagnosis period (year 0), which is higher than the comparison group’s quit probability. Similarly, diagnosed individuals who were previously nonsmoking are less likely to start smoking (Fig. 2c). However, an individual’s quit probability drops after the initial diagnosis period, where diagnosed individuals are no more likely to quit than “at-risk” individuals. In fact, in the period after diagnosis (i.e., two years after diagnosis),

⁸ The quit probability is the probability of not smoking in the current period conditional on smoking in the previous period. The start probability is the probability of smoking in the current period given not smoking in the previous period.

diabetics tend to start smoking at higher rates than comparison individuals. However, the absolute magnitude of this recidivism is relatively small; only about 2%. That is, almost all nonsmoking newly diagnosed diabetics (98%) remain nonsmoking after the diagnosis.

Table 4, Panel A, Column 2 shows estimates for the population average probit model without dynamics (i.e., with the lagged smoking variable removed). The removal of this variable changes the interpretation of the model from a dynamic model, focused on the changes in smoking behavior, to a model specifically considering the overall probability of smoking at a given point in time. As expected from the descriptive analysis, there is an overall negative effect over time, with nearly every diabetes coefficient having a negative sign. In this case, the initial diagnosis with and without medication are marginally significant, at slightly higher levels than in the dynamic specification. In both population average models, other variables behave similarly to other models estimated in the literature. Smoking propensity declines with increased education and age, as well as for individuals who are married. New diagnoses of other chronic diseases, especially heart disease, encourage people to stop smoking, as well.

The direction of this effect is consistent with Keenan (2009) who specifically looked at the “shock” factor in a pooled sample, as well as Kahn (1999) who estimated the effect through the difference between diabetics who were diagnosed and diabetics who were undiagnosed. Keenan (2009) reports an odds ratio of 1.69 for previously smoking, newly diagnosed diabetics to quit. When evaluated at the sample mean for quitting smoking (0.18), this translates to a crude marginal effect of a 7.7 percentage point increase of the quit probability for a new diabetes diagnosis. While this is a sizeable effect, Keenan’s results suggested that of the disease diagnoses considered (including cancer, stroke, heart disease, lung disease, and diabetes), a diabetes diagnosis played the smallest role, in magnitude and significance, in promoting quitting.

Similarly, Kahn uses NHANES data to predict that 16% of diagnosed diabetics smoke, as compared to 28% of undiagnosed diabetics, which is plausible given the results of these models. However, Kahn considers all cases of diagnosed diabetes regardless of time since diagnosis, whereas I focus on more behavior closer to the time of diagnosis. The two studies further differ in that I consider anti-diabetic medication use, which captures some of the effect of a new diagnosis, as well a different dependent variable (I consider the probability of smoking, regardless of previous smoking behavior). Finally, my use of a comparison group makes the interpretation of the effects of interest in my analysis different than in Keenan (2009), as Keenan considers the entire sample of non-diagnosed elderly and near elderly HRS respondents as a comparison group.

There are substantial variations in the response to a diabetes diagnosis according to various characteristics of the sample. Table 4 suggests that newly diagnosed male diabetics (Panel B) tend to be quite responsive to smoking, especially those who are diagnosed without medication (i.e. earlier, on average, in the clinical evolution of the disease). Notably, individuals who are not white (Panel C) tend to be less responsive to a diabetes diagnosis. This result is consistent with the observation that differentials in diabetes self-management are highly responsive to the socioeconomic gradient (Goldman and Smith 2002). Importantly, individuals entering the HRS under 60 tend to respond more to a diabetes shock than those in the full sample. This is consistent with the biological mechanism of diabetes complications, in which younger individuals face a higher risk of complications since they will presumably have the disease for a longer period of time than diabetics who acquire the disease at a later age. Additionally, this might reflect younger individuals being more physically and/or mentally able to manage diabetes through lifestyle changes.

The unscaled coefficients from the dynamic linear probability model (LPM) specification (Table 4, Column 3) are similar to the Average Marginal Effects obtained from the population average model. The LPM specification does not take into account the correlation of the observations within each individual (cluster), other than adjusting the standard errors to be robust to clustering. This similarity is likely due, in part, to the highly significant lagged smoking variable, which captures a substantial amount of heterogeneity in smoking behavior. The fixed effects LPM model omits the lagged smoking variable, as the inclusion of a lagged dependent variable in a fixed effects model can be correlated with the fixed effects error term. The fixed effects LPM model, reported in Column 4, transforms the dependent variable (smoking) into a deviant of contemporaneous smoking subtracted from an individual's history of smoking (their average smoking level across their time in the HRS). From Fig. 1, this deviant value increases in magnitude as time elapses since diagnosis for all individuals. However, Fig. 1 only depicts behavior near and after a diabetes diagnosis. For many diagnosed diabetics, smoking rates started declining many years before the diagnosis. For example, Table 1 shows a substantial decline in smoking rates for individuals in the HRS during 1992 (27%) to those same individuals 16 years later in 2008 (11%). Similar declines in smoking behavior occurred in non-white individuals and individuals who entered the HRS with diabetes across this time period, as well. Thus, the deviant value will get more negative over time for individuals with diabetes as they evolve through the HRS. As expected, the year fixed effects included in this model are highly negative and significant.

The FE estimates in all Panels suggest that individuals diagnosed with medication tend to substantially reduce their propensity to smoke relative to their overall history of smoking (in the HRS) compared to undiagnosed “at risk” individuals. Of particular note is the FE models in non-

whites (Panel C, Column 4), with very strong effects of a new diagnosis with medication on smoking behavior. This effect is driven primarily by a relatively high baseline rate of smoking among non-whites in this sample, followed by a rather precipitous drop in smoking rates 2-6 years prior to a diabetes diagnosis, with the decline continuing post-diagnosis, rendering the longitudinally averaged smoking rates much greater than contemporaneous smoking rates after individuals are diagnosed with diabetes. This difference persists after the inclusion of year fixed effects. Further, individuals diagnosed without medication do not seem to have a high propensity to consistently avoid smoking, though from Panel B, Column 5, newly diagnosed males without medication seem to quit at relatively high rates. The FE estimates are higher than the models without dynamics presented in Column 2 of Table 4, because while the estimates in Column 2 consider the pre- and post- diagnosis contrast, the FE estimates consider this contrast on the overall decline in smoking prevalence, which in many cases began years before a new diabetes diagnosis.

The overall results for smoking suggest that diabetes serves as an impetus for quitting. However, individuals tend to respond over a long period of time, and start lower smoking rates prior to the receipt of a diabetes diagnosis, and in many cases continue doing so after a diagnosis. I find that diabetics are generally compliant with smoking abstinence, having an overall lower rate of smoking, as well as having a higher cumulative “quit rate” over their evolution in the HRS than “at risk” individuals. Further, individuals diagnosed without medication seem to follow a similar trajectory to those diagnosed with medication.

Drinking behavior is not as well characterized in the literature, though drinking is an important aspect of diabetes self-management. Because the effects of alcohol on blood sugar are often difficult to predict, glycemic control is often more difficult for diabetics who choose to

drink. Though alcohol consumption is not as unequivocally harmful for diabetics as smoking, newly diagnosed diabetics tend to respond more strongly to drinking than smoking. Notably, the results from the dynamic population average probit model presented in Table 5 suggest there is a decrease in the demand for drinking by about 8.5 percentage points for individuals newly diagnosed with oral medication, but the effect decreases as diabetes persists. Unlike smoking, however, drinking dynamics are influenced by medication status; individuals taking oral medication or insulin are likely to transition away from drinking (i.e. stop drinking), with an Average Marginal Effect of about 4 percentage points for oral medication and about 5 percentage points for insulin.

These predicted trends are presented in Fig. 3a, which shows that while newly diagnosed diabetics stop drinking, the effect is much greater for individuals diagnosed with medication compared to those diagnosed without. In fact, the unconditional predicted probability of drinking falls by about 10 percentage points for individuals diagnosed with medication. These individuals tend to quit drinking at high rates (Fig. 3b), and start drinking at relatively low rates (Fig. 3c). While individuals diagnosed without medication tend to continue lowering drinking propensity after diagnosis, those who are diagnosed with medication tend to remain non-drinking throughout their evolution in the HRS. Part of the subsequent decline in the non-medicated group may be due to the fact that many individuals diagnosed without medication will ultimately need oral medication (of those diagnosed without medication, 36% will require oral medication two years after diagnosis, and 56% will require oral medication four years after diagnosis), so in that regard, individuals diagnosed “early” may in fact respond more strongly if they subsequently take medication.

In considering comparison models reported in Table 5, removing the lagged drinking variable (Columns 2 and 6), current diabetes status (regardless of time of diagnosis) including medication explains much of the effect of diabetes in the overall probability of drinking, suggesting that diagnosed diabetics, overall, have much lower rates of alcohol consumption than their observationally similar non-diabetic counterparts. Fixed effects models for drinking generally mirror the fixed effects models for smoking, in that individuals face a decline in drinking propensity over time, and this decline is higher in individuals who ultimately get diagnosed with diabetes.

In general, diabetes provides a “shock” to induce individuals to curb drinking behavior, with drinking activity generally remaining low after the period of initial diagnosis. Additionally, individuals diagnosed with medication tend to drink less and stop drinking faster than those diagnosed without medication. These trends are somewhat surprising given that in most cases, alcohol is not as harmful to the prognosis of diabetes as smoking is. However, it is consistent with Kahn’s (1999) finding that diagnosed diabetics tend to respond more to a diagnosis through dietary choices (including drinking) than through other behavioral changes, such as smoking or exercise.

2.4.3. The Role of Diabetes Information and Medication: Exercise and Weight Status

Increased physical activity and weight management are two of the most important measures that a diabetic can take to improve glycemic control and slow the progress of the disease. Significant weight loss is also often a challenge for diabetics because many newly diagnosed diabetics are overweight or obese, and dietary change and exercise often difficult for diabetic patients. Frequent exercise can benefit diabetics not only by promoting weight loss, but

by physiologically decreasing blood sugar in the process⁹. Further, some of the anti-diabetic medications that are prescribed, especially insulin, often promote weight gain, making it even harder for diabetics to lose weight. I focus on a discrete weight category for diabetics: overweight or obese ($BMI \geq 25$). Using this threshold, while somewhat ambitious, is appropriate because it often takes weight loss of more than 5% of body weight to attain glycemic improvement, and overweight or obese status are commonly used as risk factors for the development of diabetes and other chronic diseases.

Table 6 presents results for frequent physical activity (three or more times a week of physical activity of any intensity level). Exercise is somewhat persistent, but not nearly as much as smoking or drinking behavior. Overall, population average dynamic results reported in Column 1 suggest that individuals diagnosed with medication tend to respond the most, increasing their likelihood of starting frequent exercise by about 4.2 percentage points, which reflects the positive effect of the “shock” (6.8 percentage points) combined with a negative effect of medication status (2.6 percentage points) on this transition. This effect is particularly powerful for individuals who entered the HRS at a younger age (Panel D, Column 5), where newly diagnosed diabetics diagnosed with oral medication are 11.4 percentage points more likely to start exercising as compared to undiagnosed “at risk” individuals. Recidivism, or stopping frequent exercise, however, is apparent in these models, where individuals taking oral medication are likely to transition away from exercise, with the approximately 2.6% average marginal effect of oral medication being unopposed by recent diabetes information at all points subsequent to the initial diagnosis. Of the other covariates, being female raise the propensity for starting exercise, while frequent visits to the doctor and being Black lowered this propensity. In

⁹ Most forms of mild to moderate exercise can lower blood sugar levels in diabetics. However, strenuous or extreme forms of exercise can raise blood sugar.

models without dynamics (Columns 2 and 6), oral medication remains significant in lowering the propensity for exercise. Additionally, individuals who received a diabetes diagnosis without medication 2 years prior faced a much lower propensity for engaging in exercise. Individuals who were prescribed oral medication in that period were even less likely to exercise. This effect appears to be driven mostly by younger males in the sample (Panels B and D, Column 6).

Fig. 4 shows the predicted trajectory for exercise from the dynamic, population average probit model. Fig. 4a suggests that the unconditional probability of frequent exercise rises to approximately 20% for newly diagnosed diabetics, but quickly drops to around 16% for individuals diagnosed with medication and 13% for those diagnosed without. This increase in frequent exercise is precipitated largely by an increase in the probability of starting exercise in both groups (Fig. 4c), and the maintenance of a (relatively) low quit probability (Fig. 4b). Compared to individuals who are at-risk, but never diagnosed, individuals who become diagnosed tend to exercise at higher levels (and shift their behavior towards frequent exercise) in the period prior to diagnosis (year -2). This may be due to the awareness of impaired glucose metabolism or pre-diabetes in the absence of a diabetes diagnosis, which may prompt individuals to increase their level of exercise. Nevertheless, these individuals increase their exercise levels even further when they ultimately get diagnosed.

Table 7 presents similar results for weight status. In general, newly diagnosed individuals are about 4 percentage points less likely to be overweight or obese, though the effect is driven by overall diabetes status for individuals diagnosed without medication, whereas there is an explicit diagnosis effect for those diagnosed with medication. In Panel A, the net magnitude of a new diabetes diagnosis for those diagnosed with or without medication is very similar at around a 4 percentage point reduction in the probability of being overweight or obese compared to non-

diabetic “at risk” individuals. Importantly, both individuals diagnosed late and early face some recidivism with regard to weight status. In individuals diagnosed without medication, many will subsequently require oral medication, which will increase the probability of weight gain.

Furthermore, this suggests that individuals diagnosed with medication who are able to wean themselves off medication (through diet and exercise, for instance), are likely to lose weight as well. Recidivism is apparent in this specification as well, with both groups having significant positive coefficients for weight gain subsequent to the diagnosis.

Columns 2 and 6 of Table 7 show population average probit results without dynamics. Most notably, individuals diagnosed with medication face a decline in the probability for being overweight or obese, which persists after diagnosis. The use of the population average framework, which takes into account the (high) correlation between the observations in this model, makes these results plausible, as ignoring this correlation would produce positive, significant, coefficients for current and recent diabetes status (results not shown). That is, simply comparing diabetics to non-diabetic “at risk” individuals would suggest that diagnosed diabetics would be more likely to be overweight or obese, which follows from Table 3. The linear fixed effects specifications (Column 4) are similar to the specification without dynamics. Like with drinking and smoking, after a diabetes diagnosis, individuals weigh less (on average) than they previously did before the diagnosis. However, from Table 3, we see that the average of body weight is higher for diabetic individuals than for non-diabetic comparisons, which will magnify any weight loss in the fixed effects model. Furthermore, individuals who have been diagnosed for a longer time tend to have a higher average body mass, as well.

It should be noted that overweight and obesity results are driven primarily by individuals around the threshold BMI (25). The impact of diabetes is somewhat different at other weight

thresholds, or when BMI is considered as a continuous variable, though the general trends are similar. Other measures, especially medication status, tend to have a larger effect on weight status for those at higher BMI categories (e.g, over 30). For instance, the clinical observation that insulin tends to raise BMI is not reflected in the dynamic models since most of these individuals tend to start and stay in the overweight or obese category.

Fig. 5 displays the predicted probability of overweight or obese status from the dynamic, population average models. Unfortunately, the persistence for remaining classified as overweight or obese is quite high. Additionally, individuals who are diagnosed with diabetes tend to consistently weigh more than individuals in the comparison group. However, individuals are less likely to be overweight or obese after a diabetes diagnosis, where the predicted probability drops from about 85% to 82% among individuals diagnosed with medication, and from 82% to 79% for individuals diagnosed without medication. Likewise, the predicted probability of transitioning out of overweight or obese status to normal weight status increases from about 5% to 8% after diagnosis. Likewise, in the period of diagnosis, newly diagnosed individuals tend to be less likely to transition into overweight or obese status than individuals in the comparison group. However, similar to the trends observed for exercise, recidivism remains fairly high, with diabetics more likely to transition into overweight or obese status, and less likely to transition out of overweight or obese status post-diagnosis, compared to the comparison group. Put another way, diabetics tend to gain more weight than their at-risk counterparts two or more years subsequent to the initial diagnosis, which suggests that while diabetics are capable of losing weight, the maintenance of this weight loss can be a challenging task.

Overall, individuals respond to diabetes by losing weight and initiating exercise at the time near diagnosis, but this effect wanes subsequent to diagnosis. In addition, those diagnosed

with medication tend to exercise at the time of diagnosis compared to those diagnosed without medication, whereas there is a smaller difference in these trajectories in considering overweight or obese status. However, overall rates of exercise remain relatively low, while rates of overweight or obesity status remain high, suggesting that long-term exercise compliance and weight management may be more difficult than curbing smoking or drinking behavior.

2.5. Conclusions

Diabetes is a highly prevalent and deleterious disease in America, especially in older Americans. Medical evidence suggests that health behaviors significantly mitigate the disease, however, compliance with these health investment behaviors is relatively low. This paper analyzes the role of recent (including new) diabetes information and medication on smoking, drinking, exercise and weight status outcomes using population average probit models and linear comparison models. A comparison group of individuals who have not yet been diagnosed with diabetes, but are likely to suffer from high blood sugar, was used. Further, medication status at diagnosis was used to gauge the relative clinical stage at which a diabetic was diagnosed: those diagnosed without medication were considered to have been diagnosed somewhat earlier than those diagnosed with oral medication or insulin.

From the descriptive analysis, individuals who are diagnosed with diabetes tend to weigh more, but smoke and drink less than individuals in the comparison group. Considering the trajectory of these behaviors over time, diagnosed diabetics tended to substantially curb their smoking or drinking behavior both after the diagnosis and in subsequent years. However, exercise and overweight or obese status showed a short “shock” behavior, in which diabetics

initially responded to the diagnosis by losing weight and increasing exercise, but quickly returned to their original habits subsequent to their diagnosis.

The dynamic population average models confirmed these trends, with individuals tending to curb smoking and drinking behavior, lose weight, and increase exercise during the diagnosis period. Overall, individuals diagnosed with medication tended to respond slightly more to a new diabetes diagnosis, especially in reducing drinking activity and increasing exercise activity. In other cases, the trajectory of these behaviors was relatively similar between those diagnosed with and without medication. Two years or longer subsequent to diagnosis, individuals reduced smoking, drinking, and exercise levels, and gained weight. Linear fixed effects models reflected a much lower propensity for smoking and drinking behavior post-diagnosis compared to an individual's entire longitudinal history of these activities. Similar results were found in the body weight specifications.

Given that individuals tend to respond most strongly to a diabetes diagnosis closer to the time of diagnosis, especially with regard to exercise and weight loss, the importance of an accurate and fast diagnosis of diabetes is all the more important, especially in older individuals. Since there are many diabetics who do not know they have it, a faster diagnosis can yield lower rates of complications and better long-term outcomes.

2.6. References

- American Diabetes Association. 2008. "Nutrition Recommendations and Interventions for Diabetes." *Diabetes Care* 31(Supplement 1):S61-S78.
- Anderson, R.M. and M.M. Funnell. 2000. "Compliance and Adherence are Dysfunctional Concepts in Diabetes Care." *The Diabetes Educator* 26(4):597-604.
- Becker, G.S. and K.M. Murphy. 1988. "A Theory of Rational Addiction." *The Journal of Political Economy* 96(4):675-700.

- Benítez-Silva, H. and H. Ni. 2008. "Health status and health dynamics in an empirical model of expected longevity." *Journal of Health Economics* 27(3):564-584.
- Bennett, C.M., M. Guo, and S.C. Dharmage. 2007. "HbA1c as a screening tool for detection of Type 2 diabetes: a systematic review." *Diabetic Medicine* 24(4):333-343.
- Boulé, N.G., E. Haddad, G.P. Kenny, G.A. Wells, and R.J. Sigal. 2001. "Effects of Exercise on Glycemic Control and Body Mass in Type 2 Diabetes Mellitus." *JAMA: The Journal of the American Medical Association* 286(10):1218-1227.
- Carlson, K.A. and B.W. Gould. 1994. "The Role of Health Knowledge in Determining Dietary Fat Intake." *Review of Agricultural Economics* 16(3):373-386.
- Centers for Disease Control and Prevention. 2007. "National diabetes fact sheet."
- Chern, W.S., E.T. Loehman, and S.T. Yen. 1995. "Information, Health Risk Beliefs, and the Demand for Fats and Oils." *The Review of Economics and Statistics* 77(3):555-564.
- Contoyannis, P., A.M. Jones, and N. Rice. 2004a. "The dynamics of health in the British Household Panel Survey." *Journal of Applied Econometrics* 19(4):473-503.
- . 2004b. "Simulation-based inference in dynamic panel probit models: An application to health." *Empirical Economics* 29(1):49-77.
- Cowie, C.C., K.F. Rust, E.S. Ford, M.S. Eberhardt, D.D. Byrd-Holt, C. Li, D.E. Williams, E.W. Gregg, K.E. Bainbridge, S.H. Saydah, and L.S. Geiss. 2009. "Full Accounting of Diabetes and Pre-Diabetes in the U.S. Population in 1988–1994 and 2005–2006." *Diabetes Care* 32(2):287-294.
- Fitzgerald, J., P. Gottschalk, and R. Moffitt. 1998. "An Analysis of the Impact of Sample Attrition on the Second Generation of Respondents in the Michigan Panel Study of Income Dynamics." *The Journal of Human Resources* 33(2):300-344.
- Fitzmaurice, G.M., N.M. Laird, and J.H. Ware. 2004. *Applied Longitudinal Analysis*. Hoboken, NJ: John Wiley & Sons.
- Fluss, R., D. Faraggi, and B. Reiser. 2005. "Estimation of the Youden Index and its Associated Cutoff Point." *Biometrical Journal* 47(4):458-472.
- Garay-Sevilla, M.E., J.M. Malacara, A. Gutierrez-Roa, and E. Gonzalez. 1999. "Denial of disease in Type 2 diabetes mellitus: its influence on metabolic control and associated factors." *Diabet Med* 16(3):238-244.
- Goldman, D.P. and J.P. Smith. 2002. "Can patient self-management help explain the SES health gradient?" *Proceedings of the National Academy of Sciences of the United States of America* 99(16):10929-10934.
- Goldman, N., I.F. Lin, M. Weinstein, and Y.-H. Lin. 2003. "Evaluating the quality of self-reports of hypertension and diabetes." *Journal of Clinical Epidemiology* 56(2):148-154.
- Grossman, M. 1972. "On the Concept of Health Capital and the Demand for Health." *The Journal of Political Economy* 80(2):223-255.
- Harris, M.I., R. Klein, T.A. Welborn, and M.W. Knuiman. 1992. "Onset of NIDDM occurs at least 4-7 yr before clinical diagnosis." *Diabetes Care* 15(7):815-819.
- Hernández-Quevedo, C., A.M. Jones, and N. Rice. 2008. "Persistence in health limitations: A European comparative analysis." *Journal of Health Economics* 27(6):1472-1488.
- Howard, A.A., J.H. Arnsten, and M.N. Gourevitch. 2004. "Effect of Alcohol Consumption on Diabetes Mellitus." *Annals of Internal Medicine* 140(3):211-219.
- Hsieh, C.-R. and S.-J. Lin. 1997. "Health Information and the Demand for Preventive Care among the Elderly in Taiwan." *The Journal of Human Resources* 32(2):308-333.

- International Expert Committee. 2009. "International Expert Committee Report on the Role of the A1C Assay in the Diagnosis of Diabetes." *Diabetes Care* 32(7):1327-1334.
- Ishizawa, K.T., H. Kumano, A. Sato, H. Sakura, and Y. Iwamoto. 2010. "Decreased response inhibition in middle-aged male patients with type 2 diabetes." *Biopsychosoc Med* 4(1):1.
- Kahn, M.E. 1998. "Education's role in explaining diabetic health investment differentials." *Economics of Education Review* 17(3):257-266.
- . 1999. "Diabetic Risk Taking: The Role of Information, Education and Medication." *Journal of Risk and Uncertainty* 18(2):147-164.
- Kapteyn, A., P.-C. Michaud, J. Smith, and A. Van Soest. 2006. "Effects of Attrition and Non-Response in the Health and Retirement Study. ." *RAND Working Paper* WR-407.
- Keenan, P.S. 2009. "Smoking and Weight Change After New Health Diagnoses in Older Adults." *Arch Intern Med* 169(3):237-242.
- Kenkel, D.S. 1991. "Health Behavior, Health Knowledge, and Schooling." *The Journal of Political Economy* 99(2):287-305.
- Klick, J. and T. Stratmann. 2007. "Diabetes Treatments and Moral Hazard." *The Journal of Law and Economics* 50(3):519-538.
- Klotsche, J., D. Ferger, L. Pieper, J. Rehm, and H.-U. Wittchen. 2009. "A novel nonparametric approach for estimating cut-offs in continuous risk indicators with application to diabetes epidemiology." *BMC Medical Research Methodology* 9(1):63.
- Kriegsman, D.M.W., B.W.J.H. Penninx, J.T.M. Van Eijk, A.J.P. Boeke, and D.J.H. Deeg. 1996. "Self-reports and general practitioner information on the presence of chronic diseases in community dwelling elderly: A study on the accuracy of patients' self-reports and on determinants of inaccuracy." *Journal of Clinical Epidemiology* 49(12):1407-1417.
- Little, R.J. and D.B. Rubin. 1987. *Statistical analysis with missing data*. New York: Wiley.
- Mackenbach, J.P., C.W. Looman, and J.B. van der Meer. 1996. "Differences in the misreporting of chronic conditions, by level of education: the effect on inequalities in prevalence rates." *Am J Public Health* 86(5):706-711.
- Mäkimattila, S., K. Nikkilä, and H. Yki-Järvinen. 1999. "Causes of weight gain during insulin therapy with and without metformin in patients with Type II diabetes mellitus." *Diabetologia* 42(4):406-412.
- Martin, L.M., M. Leff, N. Calonge, C. Garrett, and D.E. Nelson. 2000. "Validation of self-reported chronic conditions and health services in a managed care population." *American Journal of Preventive Medicine* 18(3):215-218.
- McGlynn, E.A., S.M. Asch, J. Adams, J. Keesey, J. Hicks, A. DeCristofaro, and E.A. Kerr. 2003. "The quality of health care delivered to adults in the United States." *N Engl J Med* 348(26):2635-2645.
- Mellor, J.M. 2011. "Do cigarette taxes affect children's body mass index? The effect of household environment on health." *Health Economics* 20(4):417-431.
- Mullahy, J. 1999. "It'll only hurt a second? Microeconomic determinants of who gets flu shots." *Health Economics* 8(1):9-24.
- Nathan, D.M., J.B. Buse, M.B. Davidson, R.J. Heine, R.R. Holman, R. Sherwin, and B. Zinman. 2006. "Management of Hyperglycemia in Type 2 Diabetes: A Consensus Algorithm for the Initiation and Adjustment of Therapy." *Diabetes Care* 29(8):1963-1972.
- Rohlfing, C.L., R.R. Little, H.M. Wiedmeyer, J.D. England, R. Madsen, M.I. Harris, K.M. Flegal, M.S. Eberhardt, and D.E. Goldstein. 2000. "Use of GHb (HbA1c) in screening for undiagnosed diabetes in the U.S. population." *Diabetes Care* 23(2):187-191.

- Schectman, J., J. Schorling, and J. Voss. 2008. "Appointment Adherence and Disparities in Outcomes Among Patients with Diabetes." *Journal of General Internal Medicine* 23(10):1685-1687.
- Schisterman, E.F., N.J. Perkins, A. Liu, and H. Bondell. 2005. "Optimal Cut-point and Its Corresponding Youden Index to Discriminate Individuals Using Pooled Blood Samples." *Epidemiology* 16(1):73-81 10.1097/1001.ede.0000147512.0000181966.ba.
- Shah, B. and D. Manuel. 2008. "Self-reported diabetes is associated with self-management behaviour: a cohort study." *BMC Health Services Research* 8(1):142.
- Sigal, R.J., G.P. Kenny, D.H. Wasserman, C. Castaneda-Sceppa, and R.D. White. 2006. "Physical Activity/Exercise and Type 2 Diabetes." *Diabetes Care* 29(6):1433-1438.
- Sinclair, A.J., S.P. Conroy, and A.J. Bayer. 2008. "Impact of Diabetes on Physical Function in Older People." *Diabetes Care* 31(2):233-235.
- Slade, F.P. 1987. "Retirement Status and State Dependence: A Longitudinal Study of Older Men." *Journal of Labor Economics* 5(1):90-105.
- Sloan, F.A., N.A. Padron, and A.C. Platt. 2009. "Preferences, beliefs, and self-management of diabetes." *Health Serv Res* 44(3):1068-1087.
- Thaler, R.H. and H.M. Shefrin. 1981. "An Economic Theory of Self-Control." *The Journal of Political Economy* 89(2):392-406.
- Wheeler, M.L., M.J. Franz, and J.C. Froehlich. 2004. "Alcohol Consumption and Type 2 Diabetes." *DOC News* 1(2):7.
- Wilder, R.P., S.R. Majumdar, S.W. Klarenbach, and P. Jacobs. 2005. "Socio-economic status and undiagnosed diabetes." *Diabetes Research and Clinical Practice* 70(1):26-30.
- Wing, R.R., R. Koeske, L.H. Epstein, M.P. Nowalk, W. Gooding, and D. Becker. 1987. "Long-term Effects of Modest Weight Loss in Type II Diabetic Patients." *Arch Intern Med* 147(10):1749-1753.
- Wooldridge, J.M. 2005. "Simple solutions to the initial conditions problem in dynamic, nonlinear panel data models with unobserved heterogeneity." *Journal of Applied Econometrics* 20(1):39-54.
- Yeung, S.E., A.L. Fischer, and R.A. Dixon. 2009. "Exploring effects of type 2 diabetes on cognitive functioning in older adults." *Neuropsychology* 23(1):1-9.
- Youden, W.J. 1950. "Index for rating diagnostic tests." *Cancer* 3(1):32-35.
- Zeger, S.L., K.-Y. Liang, and P.S. Albert. 1988. "Models for Longitudinal Data: A Generalized Estimating Equation Approach." *Biometrics* 44(4):1049-1060.

2.7. Tables and Figures

Table 2.1. Demographic and health characteristics of respondents.

	1992	2008
Female	0.521	0.563
Age	56.013 (3.222)	70.993 (5.205)
Black	0.102	0.075
Hispanic	0.064	0.058
Other race	0.035	0.033
Highest grade- completed high school	0.338	0.340
Highest grade- completed some	0.198	0.211
Highest grade- completed college	0.186	0.211
Total income (tens of thousands)	5.007 (5.513)	7.925 (96.644)
Total assets (hundreds of thousands)	2.36 (4.764)	6.027 (15.43)
Married	0.739	0.650
Number of doctor visits ⁺	4.261 (8.051)	10.728 (15.377)
Employer or government insurance	0.580	0.943
Diagnosed with diabetes	0.098	0.229
Diabetes: not taking any medication	0.037	0.040
Diabetes: taking oral medication	0.042	0.169
Diabetes: taking insulin	0.023	0.045
Heart disease	0.126	0.277
Lung disease	0.083	0.144
Cerebrovascular disease (stroke)	0.025	0.061
Hypertension	0.380	0.629
Currently smokes	0.269	0.119
Currently drinks any alcohol	0.636	0.516
Body mass index (BMI)	26.985 (5.003)	28.031 (5.546)
Overweight or obese status	0.636	0.698
Frequent exercise*		0.200
Number of individuals	9671	6863

Note: HRS sampling weights used in all analyses. The 1992 column represents all respondents in that period; the 2008 column represents those respondents still in the panel in 2008.

Standard deviation in parentheses for continuous variables.

⁺In 1992, number of doctor visits was reported for the past 12 months, while in subsequent waves, was reported for the past two years.

*Frequent exercise (3 or more times per week) is only measured in three waves from 2004-2008.

Overweight or obese status represents individuals whose BMI ≥ 25 .

Table 2.2. Probit model predicting persistent hyperglycemia in HRS respondents.

	Coef.	Std. Error.
Age 40-50	-0.027	(0.369)
Age 50-60	0.280	(0.355)
Age 60-70	0.537	(0.355)
Age 70-80	0.673*	(0.357)
Age 80-90	0.752**	(0.359)
Age >90	0.634*	(0.382)
Female	0.030	(0.035)
U.S. Born	0.081	(0.068)
Highest grade- completed high school	-0.091*	(0.048)
Highest grade- completed some college	-0.074	(0.053)
Highest grade- completed college	-0.103*	(0.056)
Black	0.370***	(0.053)
Hispanic	0.354***	(0.075)
Other race	0.222**	(0.089)
Body mass index (BMI)	0.076***	(0.017)
Self-reported health status: fair	0.341***	(0.068)
Self-reported health status: good	0.271***	(0.060)
Self-reported health status: very good	0.084	(0.059)
Currently diagnosed with high blood pressure	0.169***	(0.037)
Currently diagnosed with heart disease	0.110**	(0.043)
Currently diagnosed with cerebrovascular disease	0.073	(0.078)
Currently diagnosed with lung disease	-0.045	(0.040)
Constant	-2.793***	(0.456)
Log likelihood	-3739.703	
Number of observations	5988	
Pseudo-R ²	0.08	

* $p \leq 0.10$ ** $p \leq 0.05$ *** $p \leq 0.01$

Probit coefficients reported. Standard errors in parentheses.

A sub-sample of 2006 HRS respondents consented to a blood test.

A HbA1c greater than 5.7 represents an abnormally high blood sugar level over a 2-3 month period.

Table 2.3. Descriptive means by diabetes diagnostic status.

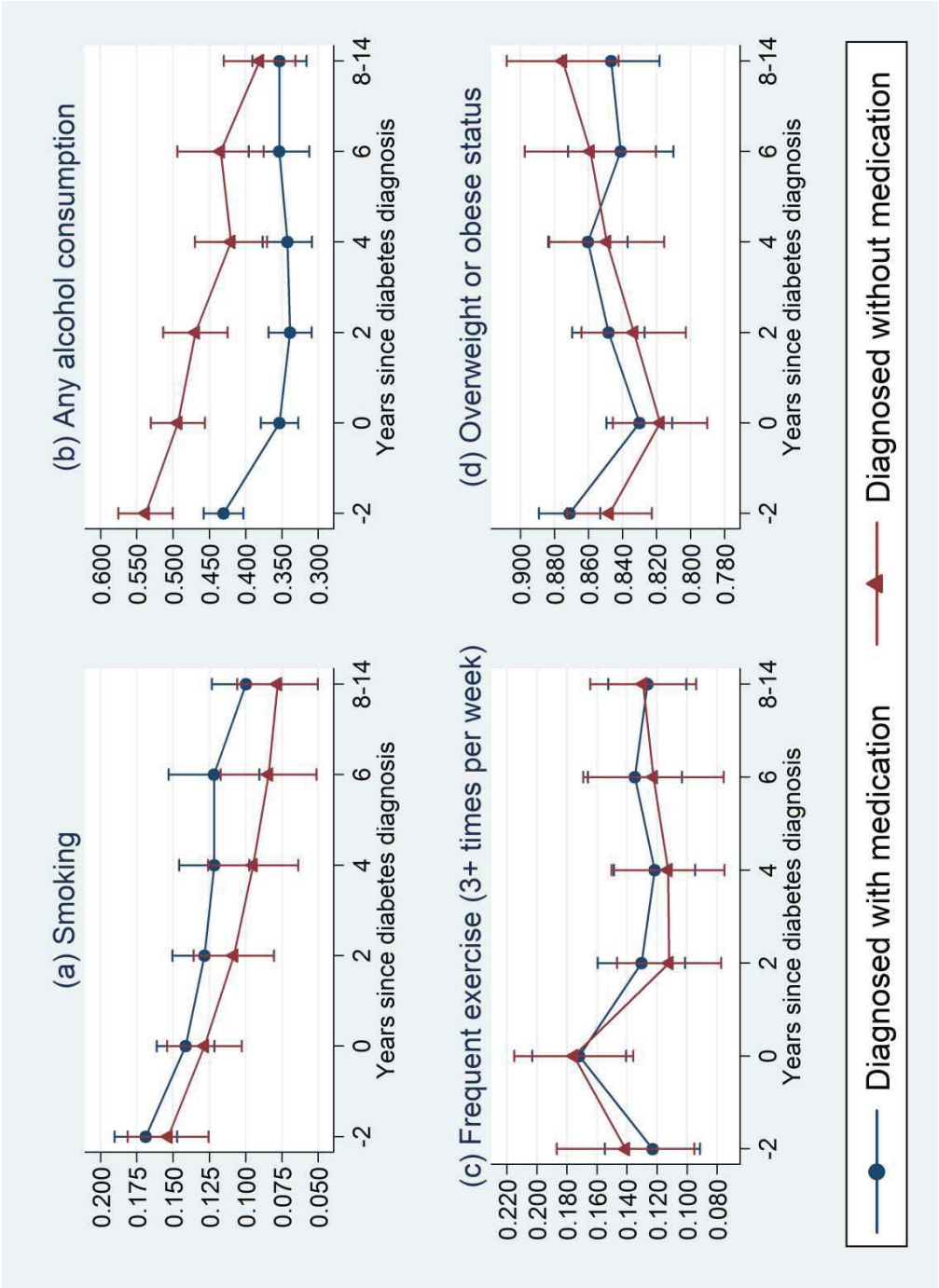
	(1) Entire HRS	(2) Never diagnosed	(3) Diagnosed during Pre-diagnosis	(4) HRS Post-diagnosis	(5) Currently diagnosed	(6) Newly diagnosed without medication	(7) Newly diagnosed with medication	(8) At risk, never diagnosed
Female	0.542	0.553	0.515	0.501	0.499	0.504	0.490	0.570
Age	66.015	66.116	63.826	68.085	66.651	66.225	66.263	69.795
Black	0.089	0.076	0.112	0.116	0.149	0.080	0.137	0.146
Hispanic	0.062	0.054	0.086	0.093	0.096	0.080	0.097	0.099
Other race	0.039	0.035	0.052	0.055	0.060	0.065	0.054	0.050
Highest grade- completed high school	0.312	0.315	0.308	0.313	0.300	0.316	0.306	0.310
Highest grade- completed some college	0.217	0.222	0.204	0.194	0.199	0.239	0.193	0.194
Highest grade- completed college	0.216	0.229	0.170	0.167	0.166	0.192	0.159	0.123
Total income (tens of thousands)	6.293	6.581	5.164	6.181	5.301	6.324	4.940	4.107
Total assets (hundreds of thousands)	4.156	4.481	3.088	3.245	2.927	3.687	2.855	2.669
Married	0.633	0.635	0.658	0.623	0.605	0.632	0.632	0.570
Number of doctor visits in past two years	9.026	8.075	8.814	13.433	14.447	12.251	14.338	10.243
Employer or government insurance	0.793	0.789	0.760	0.852	0.835	0.845	0.809	0.858
Diagnosed with diabetes	0.140	0.000	0.000	1.000	1.000	1.000	1.000	0.000
Diabetes: not taking any medication	0.029	0.000	0.000	0.251	0.204	1.000	0.000	0.000
Diabetes: taking oral medication	0.095	0.000	0.000	0.719	0.678	0.000	0.975	0.000
Diabetes: taking insulin	0.031	0.000	0.000	0.074	0.218	0.000	0.052	0.000
Heart disease	0.212	0.188	0.220	0.328	0.340	0.251	0.306	0.296
Lung disease	0.114	0.107	0.130	0.159	0.148	0.141	0.175	0.135
Cerebrovascular disease (stroke)	0.052	0.044	0.043	0.086	0.102	0.054	0.079	0.076
Hypertension	0.482	0.428	0.565	0.750	0.742	0.712	0.728	0.630
Currently smokes	0.161	0.165	0.184	0.118	0.131	0.125	0.140	0.138
Currently drinks any alcohol	0.533	0.564	0.494	0.393	0.378	0.514	0.362	0.462
Body mass index (BMI)	27.264	26.481	29.541	30.427	30.384	29.794	30.622	28.465
Overweight or obese	0.645	0.595	0.828	0.839	0.820	0.821	0.837	0.750
Frequent exercise*	0.153	0.161	0.121	0.137	0.128	0.175	0.172	0.136
Hemoglobin A1c ⁺	5.801	5.558	6.144	6.475	6.750	5.995	6.371	5.661
Number of individuals	28011	21574	3520	3031	6081	858	1722	10565
Number of observations	129101	98766	11208	8635	19127	858	1722	50995

HRS sampling weights used in all analyses.

*Frequent exercise (3 or more times per week) is only measured in three waves from 2004-2008. Means are reported for all values for which exercise is observed.

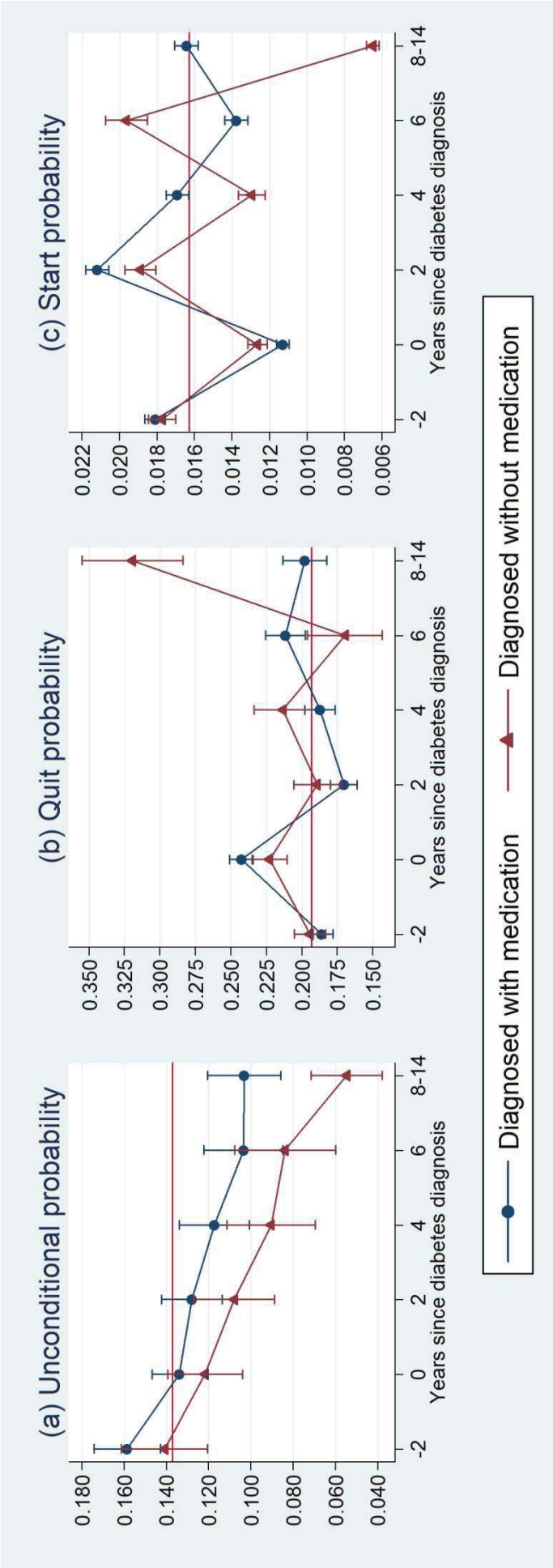
⁺Hemoglobin A1c is a test that measures long-term blood sugar levels. Completely normal values are below 5.7%. HbA1c is only reported for respondents in the 2006 Biomarker Data.

Figure 2.1.1. Health habit trajectory after a diabetes diagnosis.



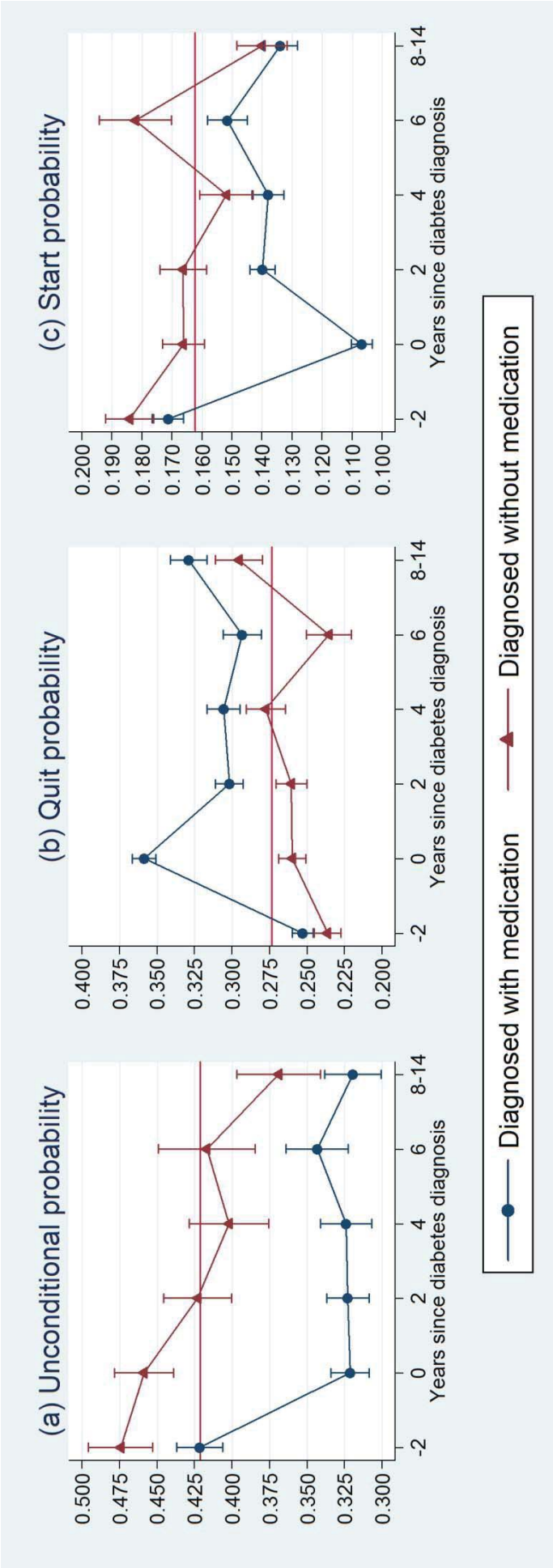
Note: Vertical bars represent 95% confidence intervals of the mean.

Figure 2.2. Predicted smoking trajectory after a diabetes diagnosis based on dynamic, population



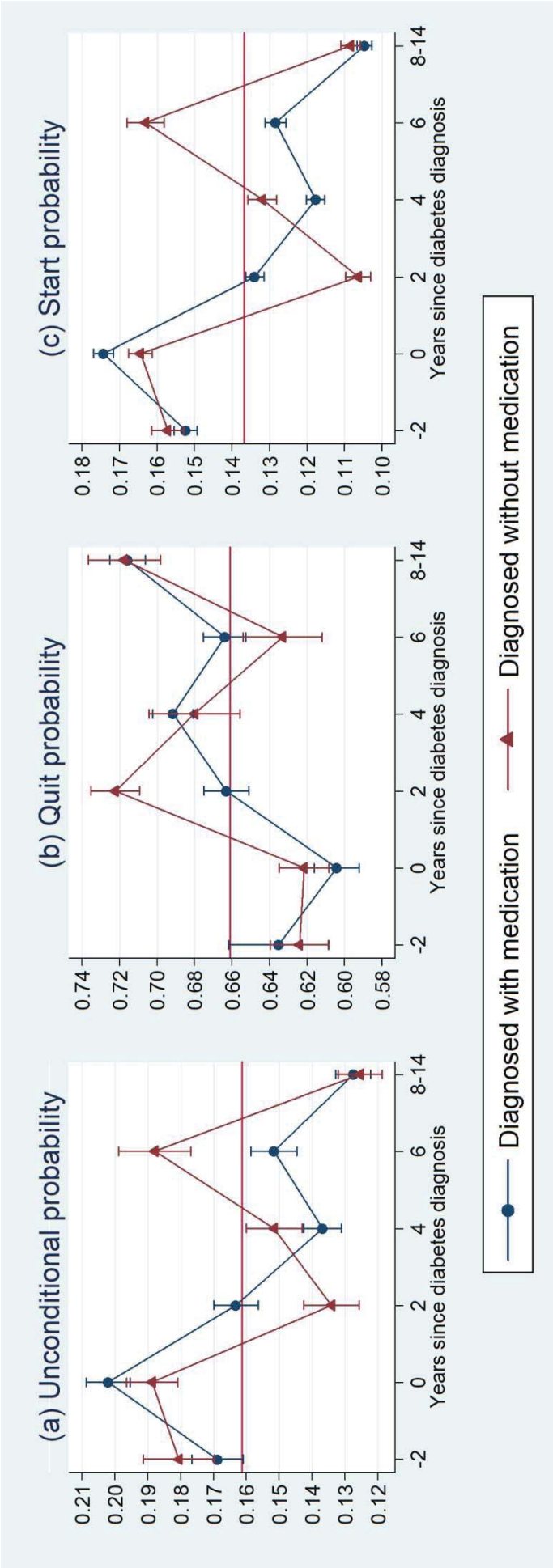
Note: Smoking reflects if an individual ever smoked cigarettes in a two-year period. Vertical bars represent 95% confidence intervals of the mean predicted probabilities. The horizontal line represents the mean predicted probability for individuals in the comparison group, including individuals who have yet to be diagnosed with diabetes, and undiagnosed individuals who are “at-risk” for developing diabetes or pre-diabetes.

Figure 2.3. Predicted drinking trajectory after a diabetes diagnosis based on dynamic, population average model.



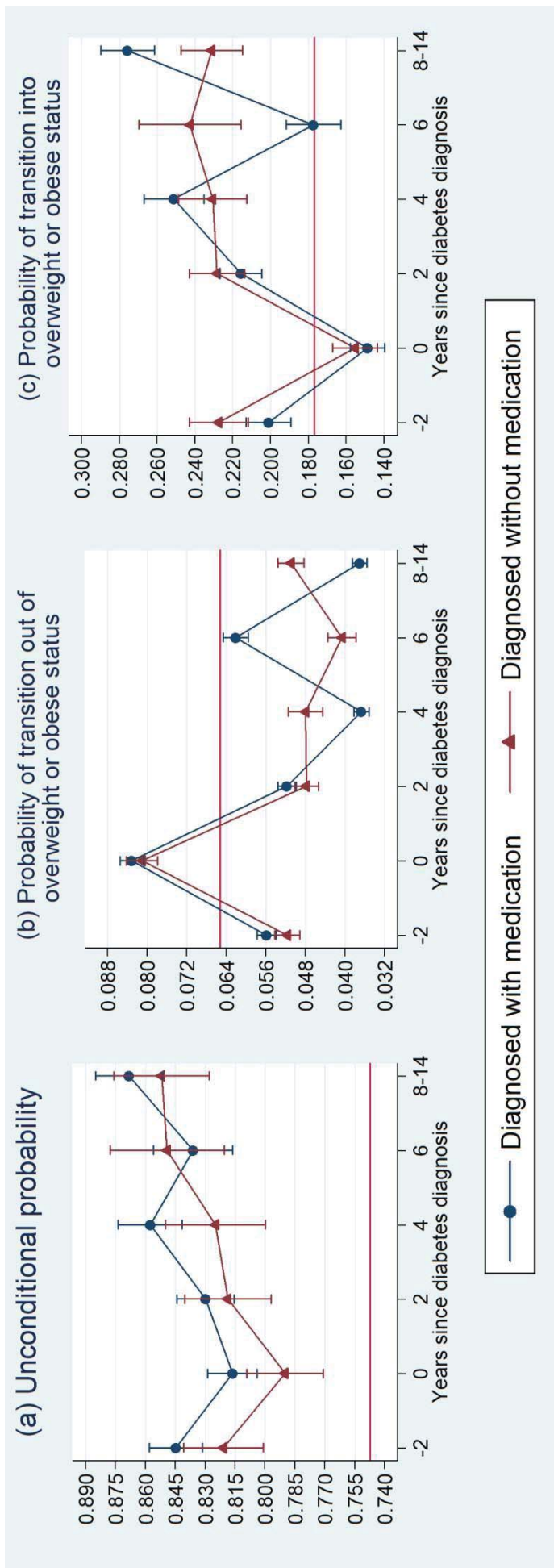
Note: Drinking reflects if an individual ever consumed alcoholic beverages in a two-year period. Vertical bars represent 95% confidence intervals of the mean predicted probabilities. The horizontal line represents the mean predicted probability for individuals in the comparison group, including individuals who have yet to be diagnosed with diabetes, and undiagnosed individuals who are “at-risk” for developing diabetes or pre-diabetes.

Figure 2.4. Predicted frequent exercise trajectory after a diabetes diagnosis based on dynamic, population average model.



Note: Frequent exercise reflects if an individual engaged in physical activity three or more times per week. Vertical bars represent 95% confidence intervals of the mean predicted probabilities. The horizontal line represents the mean predicted probability for individuals in the comparison group, including individuals who have yet to be diagnosed with diabetes, and undiagnosed individuals who are “at-risk” for developing diabetes or pre-diabetes.

Figure 2.5. Predicted overweight or obese status trajectory after a diabetes diagnosis based on dynamic, population average model.



Note: Overweight or obese status reflects if an individual's body mass index was 25 or above. Vertical bars represent 95% confidence intervals of the mean predicted probabilities. The horizontal line represents the mean predicted probability for individuals in the comparison group, including individuals who have yet to be diagnosed with diabetes, and undiagnosed individuals who are "at-risk" for developing diabetes or pre-diabetes.

Table 2.4. Average effects for smoking activity by diabetes status and medication.

	(1)		(2)		(3)		(4)		(5)		(6)		(7)		(8)	
	PA Probit	AME	PA Probit	AME	LPM	Coef.	LPM Fixed Effects	Std.Error	PA Probit	AME	PA Probit	AME	LPM	Coef.	LPM Fixed Effects	Std.Error
Panel A: Entire sample																
Newly diagnosed without medication	-0.017	(0.010)	-0.024	(0.014)	-0.020	(0.012)	-0.028	(0.019)	-0.036	(0.014)	-0.030	(0.022)	-0.046	(0.019)	-0.056	(0.031)
Diagnosed 2 years ago without medication	-0.003	(0.011)	-0.011	(0.014)	-0.005	(0.011)	-0.014	(0.018)	-0.021	(0.014)	-0.024	(0.023)	-0.021	(0.014)	-0.039	(0.029)
Diagnosed 4 years ago without medication	-0.011	(0.012)	-0.016	(0.015)	-0.011	(0.012)	-0.019	(0.019)	-0.036	(0.015)	-0.037	(0.023)	-0.032	(0.013)	-0.054	(0.030)
Diagnosed 6 years ago without medication	0.002	(0.016)	-0.011	(0.020)	-0.000	(0.014)	-0.011	(0.022)	0.009	(0.025)	-0.014	(0.033)	0.006	(0.021)	-0.034	(0.037)
Diagnosed 8-14 years ago without	-0.031	(0.015)	-0.014	(0.025)	-0.030	(0.013)	-0.008	(0.025)	-0.041	(0.027)	-0.014	(0.038)	-0.035	(0.020)	-0.025	(0.039)
Newly diagnosed with medication	-0.015	(0.007)	-0.019	(0.009)	-0.017	(0.009)	-0.047	(0.014)	-0.010	(0.011)	-0.006	(0.017)	-0.009	(0.014)	-0.058	(0.026)
Diagnosed 2 years ago with medication	0.003	(0.008)	-0.019	(0.010)	0.002	(0.009)	-0.042	(0.015)	-0.012	(0.011)	-0.021	(0.019)	-0.010	(0.013)	-0.068	(0.028)
Diagnosed 4 years ago with medication	-0.002	(0.007)	-0.010	(0.011)	-0.000	(0.007)	-0.032	(0.015)	-0.003	(0.011)	-0.013	(0.020)	-0.001	(0.012)	-0.056	(0.027)
Diagnosed 6 years ago with medication	-0.009	(0.008)	-0.005	(0.013)	-0.005	(0.007)	-0.022	(0.016)	-0.007	(0.012)	-0.011	(0.022)	-0.003	(0.012)	-0.053	(0.029)
Diagnosed 8-14 years ago with medication	-0.004	(0.008)	-0.005	(0.015)	-0.005	(0.007)	-0.021	(0.018)	0.003	(0.011)	-0.007	(0.025)	0.003	(0.012)	-0.049	(0.031)
Currently diabetic: not taking medication	0.008	(0.006)	0.007	(0.010)	0.010	(0.007)	0.014	(0.015)	0.021	(0.010)	0.007	(0.017)	0.023	(0.011)	0.032	(0.024)
Currently diabetic: taking oral medication	0.002	(0.004)	-0.002	(0.008)	0.001	(0.004)	0.023	(0.013)	0.001	(0.006)	-0.010	(0.015)	-0.002	(0.007)	0.035	(0.024)
Currently diabetic: taking insulin	0.005	(0.004)	-0.015	(0.008)	0.005	(0.005)	-0.019	(0.011)	0.003	(0.007)	-0.034	(0.013)	-0.000	(0.008)	-0.033	(0.018)
Smoked, t-1	0.757	(0.016)			0.780	(0.007)			0.779	(0.022)			0.778	(0.010)		
Number of individuals (n)	12878		13019		12878		13019		5740		5817		5740		5817	
Number of observations (nT)	56937		58614		56937		58614		25046		25847		25046		25847	
Panel C: Non-whites																
Panel D: Individuals entering the HRS before age 60																
Newly diagnosed without medication	0.005	(0.039)	-0.030	(0.022)	0.002	(0.040)	0.020	(0.053)	-0.020	(0.013)	0.030	(0.050)	-0.032	(0.016)	-0.044	(0.027)
Diagnosed 2 years ago without medication	0.009	(0.033)	-0.024	(0.023)	0.002	(0.036)	0.020	(0.045)	-0.011	(0.013)	0.043	(0.042)	-0.017	(0.014)	-0.031	(0.026)
Diagnosed 4 years ago without medication	0.039	(0.056)	-0.037	(0.023)	0.037	(0.056)	0.060	(0.059)	-0.008	(0.017)	0.072	(0.059)	-0.011	(0.016)	-0.027	(0.028)
Diagnosed 6 years ago without medication	-0.020	(0.029)	-0.014	(0.033)	-0.012	(0.012)	0.047	(0.050)	-0.004	(0.024)	0.055	(0.049)	-0.010	(0.020)	-0.019	(0.032)
Diagnosed 8-14 years ago without	-0.076	(0.050)	-0.014	(0.038)	-0.051	(0.038)	0.021	(0.076)	-0.035	(0.023)	0.007	(0.071)	-0.040	(0.018)	-0.023	(0.035)
Newly diagnosed with medication	-0.006	(0.017)	-0.006	(0.017)	-0.010	(0.018)	-0.093	(0.034)	-0.017	(0.010)	-0.032	(0.022)	-0.024	(0.012)	-0.064	(0.023)
Diagnosed 2 years ago with medication	0.019	(0.019)	-0.021	(0.019)	0.020	(0.019)	-0.080	(0.036)	0.000	(0.010)	-0.019	(0.025)	-0.002	(0.012)	-0.061	(0.024)
Diagnosed 4 years ago with medication	-0.008	(0.018)	-0.013	(0.020)	-0.007	(0.019)	-0.082	(0.038)	0.001	(0.010)	-0.027	(0.026)	0.001	(0.011)	-0.042	(0.024)
Diagnosed 6 years ago with medication	0.017	(0.018)	-0.011	(0.022)	0.021	(0.016)	-0.062	(0.039)	-0.012	(0.010)	-0.006	(0.031)	-0.012	(0.010)	-0.036	(0.024)
Diagnosed 8-14 years ago with medication	-0.006	(0.016)	-0.007	(0.025)	-0.005	(0.017)	-0.096	(0.044)	-0.008	(0.011)	-0.020	(0.033)	-0.011	(0.010)	-0.034	(0.027)
Currently diabetic: not taking medication	0.001	(0.015)	0.007	(0.017)	0.003	(0.017)	0.004	(0.036)	0.013	(0.009)	-0.014	(0.025)	0.019	(0.010)	0.036	(0.023)
Currently diabetic: taking oral medication	0.000	(0.009)	-0.010	(0.015)	0.001	(0.010)	0.055	(0.029)	0.005	(0.006)	0.007	(0.018)	0.006	(0.007)	0.036	(0.021)
Currently diabetic: taking insulin	-0.005	(0.009)	-0.034	(0.013)	-0.005	(0.009)	-0.026	(0.024)	0.007	(0.006)	-0.033	(0.018)	0.008	(0.007)	-0.019	(0.014)
Smoked, t-1	0.719	(0.037)			0.761	(0.013)			0.760	(0.02)			0.790	(0.008)		
Number of individuals (n)	3482		3540		3482		3540		6537		6632		6537		6632	
Number of observations (nT)	14205		14785		14205		14785		32187		33334		32187		33334	

Notes: PA, population average (GEE); AME, average marginal effect; LPM, linear probability model. Bootstrapped standard errors, clustered at the individual level, in parentheses for population average models. LPM standard errors are robust to clustering.

The overall sample is restricted to all individuals who ultimately receive a diabetes diagnosis, as well as non-diabetics who are likely to have high blood sugar (described in Table 3, Column 8).

All models include inverse probability weights (IPW), which account for attrition and differential baseline selection probability, where the weight represents the means of these IPW for each period the individual is in the model.

All models include gender, age, age², race, ethnicity, education, income, assets, employer or government insurance, U.S. birth status, and current or new diagnoses of heart disease, lung disease, stroke, or hypertension.

Bolded coefficients represent significance at the 5% level; italics represents significance at the 10% level.

Table 2.5. Average effects for drinking activity by diabetes status and medication.

	(1)		(2)		(3)		(4)		(5)		(6)		(7)		(8)	
	PA Probit	AME	PA Probit	AME	LPM	Coef.	LPM Fixed Effects	Std.Error	PA Probit	AME	PA Probit	Std. Error	LPM	Coef.	LPM Fixed Effects	Std. Error
Panel A: Entire sample																
Panel B: Males																
Newly diagnosed without medication	0.005	(0.022)	0.019	(0.020)	0.001	(0.020)	-0.032	(0.024)	0.023	(0.034)	0.044	(0.029)	0.018	(0.030)	-0.003	(0.034)
Diagnosed 2 years ago without medication	0.010	(0.020)	0.005	(0.020)	0.007	(0.019)	-0.047	(0.024)	-0.001	(0.028)	0.013	(0.027)	-0.004	(0.025)	-0.027	(0.034)
Diagnosed 4 years ago without medication	-0.001	(0.022)	-0.010	(0.022)	0.000	(0.020)	-0.061	(0.028)	0.026	(0.036)	0.021	(0.034)	0.026	(0.030)	-0.028	(0.040)
Diagnosed 6 years ago without medication	0.036	(0.029)	0.018	(0.027)	0.036	(0.026)	-0.033	(0.031)	0.035	(0.040)	0.033	(0.042)	0.039	(0.038)	-0.018	(0.049)
Diagnosed 8-14 years ago without medication	-0.015	(0.026)	-0.014	(0.030)	-0.019	(0.023)	-0.071	(0.035)	-0.005	(0.038)	0.002	(0.041)	-0.010	(0.034)	-0.051	(0.047)
Newly diagnosed with medication	-0.043	(0.015)	-0.026	(0.017)	-0.044	(0.014)	-0.075	(0.021)	-0.067	(0.022)	-0.055	(0.028)	-0.067	(0.021)	-0.108	(0.034)
Diagnosed 2 years ago with medication	0.002	(0.017)	-0.023	(0.018)	0.003	(0.015)	-0.069	(0.021)	-0.002	(0.025)	-0.045	(0.028)	-0.001	(0.023)	-0.100	(0.035)
Diagnosed 4 years ago with medication	0.002	(0.019)	-0.013	(0.019)	0.001	(0.016)	-0.060	(0.023)	-0.029	(0.026)	-0.054	(0.030)	-0.026	(0.024)	-0.109	(0.037)
Diagnosed 6 years ago with medication	0.019	(0.024)	0.005	(0.022)	0.018	(0.020)	-0.044	(0.027)	0.006	(0.034)	-0.027	(0.039)	0.009	(0.031)	-0.084	(0.045)
Diagnosed 8-14 years ago with medication	0.000	(0.018)	0.003	(0.024)	0.000	(0.015)	-0.048	(0.028)	0.025	(0.025)	0.018	(0.039)	0.018	(0.023)	-0.044	(0.045)
Currently diabetic; not taking medication	-0.023	(0.014)	-0.025	(0.015)	-0.017	(0.011)	0.023	(0.019)	-0.032	(0.021)	-0.034	(0.022)	-0.026	(0.017)	0.012	(0.028)
Currently diabetic; taking oral medication	-0.042	(0.009)	-0.050	(0.011)	-0.031	(0.007)	0.013	(0.017)	-0.021	(0.013)	-0.029	(0.020)	-0.017	(0.010)	0.030	(0.028)
Currently diabetic; taking insulin	-0.047	(0.012)	-0.062	(0.014)	-0.032	(0.009)	-0.002	(0.019)	-0.051	(0.020)	-0.074	(0.021)	-0.039	(0.014)	-0.024	(0.025)
Any alcohol consumption, t-1	0.506	(0.044)			0.636	(0.006)			0.545	(0.071)			0.654	(0.008)		
Number of individuals (n)	12899		13032		12899		13032		5749		5822		5749		5822	
Number of observations (nT)	57394		58924		57394		58924		25244		25986		25244		25986	
Panel C: Non-whites																
Panel D: Individuals entering the HRS before age 60																
Newly diagnosed without medication	0.011	(0.049)	0.009	(0.049)	-0.003	(0.047)	-0.066	(0.060)	0.031	(0.032)	0.025	(0.027)	0.022	(0.026)	-0.046	(0.031)
Diagnosed 2 years ago without medication	-0.078	(0.039)	-0.072	(0.048)	-0.065	(0.035)	-0.142	(0.061)	0.020	(0.027)	0.009	(0.028)	0.013	(0.024)	-0.065	(0.032)
Diagnosed 4 years ago without medication	-0.044	(0.042)	-0.051	(0.050)	-0.027	(0.034)	-0.115	(0.059)	-0.015	(0.030)	-0.021	(0.031)	-0.013	(0.025)	-0.089	(0.036)
Diagnosed 6 years ago without medication	-0.031	(0.057)	-0.026	(0.051)	-0.015	(0.043)	-0.101	(0.056)	0.025	(0.038)	0.011	(0.037)	0.025	(0.034)	-0.059	(0.040)
Diagnosed 8-14 years ago without medication	-0.060	(0.047)	-0.063	(0.056)	-0.037	(0.035)	-0.123	(0.065)	-0.002	(0.030)	-0.014	(0.037)	-0.013	(0.024)	-0.088	(0.043)
Newly diagnosed with medication	0.000	(0.027)	-0.011	(0.031)	-0.006	(0.023)	-0.108	(0.041)	-0.042	(0.022)	-0.035	(0.022)	-0.048	(0.018)	-0.103	(0.029)
Diagnosed 2 years ago with medication	0.013	(0.037)	-0.024	(0.035)	0.011	(0.032)	-0.118	(0.045)	-0.012	(0.024)	-0.041	(0.023)	-0.004	(0.019)	-0.105	(0.029)
Diagnosed 4 years ago with medication	0.018	(0.042)	-0.007	(0.044)	0.008	(0.039)	-0.106	(0.050)	-0.011	(0.024)	-0.030	(0.024)	-0.005	(0.020)	-0.098	(0.030)
Diagnosed 6 years ago with medication	0.056	(0.044)	0.019	(0.048)	0.042	(0.042)	-0.085	(0.056)	0.045	(0.031)	0.017	(0.029)	0.044	(0.025)	-0.056	(0.034)
Diagnosed 8-14 years ago with medication	0.057	(0.031)	0.033	(0.040)	0.044	(0.026)	-0.082	(0.051)	0.000	(0.023)	-0.003	(0.028)	-0.002	(0.017)	-0.077	(0.035)
Currently diabetic; not taking medication	-0.036	(0.028)	-0.038	(0.031)	-0.028	(0.027)	0.041	(0.038)	-0.035	(0.019)	-0.020	(0.020)	-0.030	(0.015)	0.045	(0.025)
Currently diabetic; taking oral medication	-0.062	(0.017)	-0.057	(0.025)	-0.043	(0.013)	0.050	(0.034)	-0.050	(0.015)	-0.045	(0.016)	-0.038	(0.009)	0.032	(0.024)
Currently diabetic; taking insulin	-0.020	(0.019)	-0.042	(0.022)	-0.009	(0.015)	0.014	(0.026)	-0.053	(0.015)	-0.061	(0.018)	-0.031	(0.011)	-0.003	(0.022)
Any alcohol consumption, t-1	0.488	(0.057)			0.603	(0.012)			0.384	(0.040)			0.640	(0.008)		
Number of individuals (n)	3492		3546		3492		3546		6553		6642		6553		6642	
Number of observations (nT)	14321		14871		14321		14871		32515		33531		32515		33531	

Notes: PA, population average (GEE); AME, average marginal effect; LPM, linear probability model. Bootstrapped standard errors, clustered at the individual level, in parentheses for population average models. LPM standard errors are robust to clustering.

The overall sample is restricted to all individuals who ultimately receive a diabetes diagnosis, as well as non-diabetics who are likely to have high blood sugar (described in Table 3, Column 8).

All models include inverse probability weights (IPW), which account for attrition and differential baseline selection probability, where the weight represents the means of these IPW for each period the individual is in the model.

All models include gender, age, age², race, ethnicity, education, income, assets, employer or government insurance, U.S. birth status, and current or new diagnoses of heart disease, lung disease, stroke, or hypertension.

Bolded coefficients represent significance at the 5% level; italics represents significance at the 10% level.

Table 2.6. Average effects for frequent physical activity by diabetes status and medication.

	(1)		(2)		(3)		(4)		(5)		(6)		(7)		(8)	
	PA Probit	AME	PA Probit	AME	LPM	Std.Error	LPM	Coef.	PA Probit	AME	PA Probit	AME	LPM	Coef.	LPM	Std.Error
Panel A: Entire sample																
Newly diagnosed without medication	0.013	(0.030)	-0.007	(0.024)	0.009	(0.033)	-0.008	(0.044)	0.060	(0.054)	0.020	(0.042)	0.053	(0.050)	0.065	(0.056)
Diagnosed 2 years ago without medication	-0.038	(0.024)	-0.044	(0.017)	-0.040	(0.027)	-0.040	(0.046)	-0.051	(0.033)	-0.056	(0.024)	-0.052	(0.037)	-0.047	(0.060)
Diagnosed 4 years ago without medication	-0.001	(0.031)	-0.023	(0.021)	-0.005	(0.029)	0.001	(0.051)	0.060	(0.051)	0.000	(0.034)	0.052	(0.049)	0.046	(0.069)
Diagnosed 6 years ago without medication	0.040	(0.038)	0.015	(0.029)	0.033	(0.039)	0.054	(0.064)	0.053	(0.065)	0.013	(0.047)	0.044	(0.062)	0.067	(0.094)
Diagnosed 8-14 years ago without	-0.010	(0.026)	-0.016	(0.020)	-0.011	(0.023)	-0.001	(0.066)	0.027	(0.037)	0.006	(0.028)	0.022	(0.033)	0.047	(0.087)
Newly diagnosed with medication	0.068	(0.025)	0.048	(0.020)	0.061	(0.024)	0.036	(0.037)	<i>0.071</i>	<i>(0.037)</i>	0.044	(0.028)	<i>0.062</i>	<i>(0.034)</i>	0.014	(0.051)
Diagnosed 2 years ago with medication	0.021	(0.022)	0.012	(0.018)	0.021	(0.022)	0.014	(0.039)	0.030	(0.037)	0.018	(0.027)	0.029	(0.034)	-0.002	(0.054)
Diagnosed 4 years ago with medication	0.004	(0.023)	0.005	(0.019)	0.002	(0.022)	0.020	(0.045)	-0.015	(0.033)	-0.016	(0.024)	-0.015	(0.031)	-0.028	(0.065)
Diagnosed 6 years ago with medication	0.020	(0.025)	0.023	(0.020)	0.019	(0.023)	0.042	(0.049)	-0.009	(0.031)	-0.003	(0.028)	-0.007	(0.030)	-0.009	(0.067)
Diagnosed 8-14 years ago with medication	-0.006	(0.020)	-0.009	(0.018)	-0.005	(0.018)	-0.001	(0.054)	-0.050	(0.023)	-0.037	(0.022)	<i>-0.041</i>	<i>(0.023)</i>	-0.074	(0.074)
Currently diabetic: not taking medication	0.005	(0.019)	0.021	(0.017)	0.007	(0.020)	0.036	(0.030)	-0.021	(0.026)	0.011	(0.027)	-0.017	(0.031)	-0.016	(0.041)
Currently diabetic: taking oral medication	-0.026	(0.011)	-0.018	(0.008)	-0.024	(0.010)	-0.000	(0.024)	-0.021	(0.015)	-0.010	(0.013)	-0.019	(0.015)	0.011	(0.036)
Currently diabetic: taking insulin	-0.016	(0.013)	-0.011	(0.011)	-0.015	(0.012)	-0.012	(0.024)	-0.015	(0.019)	-0.009	(0.017)	-0.015	(0.018)	0.007	(0.032)
Frequent physical activity, t-1	0.204	(0.023)			0.165	(0.013)			0.182	(0.034)			0.160	(0.020)		
Number of individuals (n)	9326		9502		9326		9502		4043		4125		4043		4125	
Number of observations (nT)	16371		24856		16371		24856		7120		10797		7120		10797	
Panel C: Non-whites																
Panel D: Individuals entering the HRS before age 60																
Newly diagnosed without medication	0.001	(0.065)	-0.043	(0.042)	0.002	(0.065)	0.070	(0.085)	-0.027	(0.048)	-0.057	(0.038)	-0.030	(0.059)	-0.060	(0.102)
Diagnosed 2 years ago without medication	-0.026	(0.048)	-0.040	(0.036)	-0.028	(0.049)	0.021	(0.083)	-0.072	(0.042)	-0.090	(0.034)	-0.067	(0.054)	-0.070	(0.104)
Diagnosed 4 years ago without medication	0.015	(0.072)	0.005	(0.053)	0.006	(0.072)	0.057	(0.103)	-0.060	(0.060)	-0.074	(0.042)	-0.046	(0.059)	-0.038	(0.115)
Diagnosed 6 years ago without medication	0.184	(0.144)	0.110	(0.104)	0.176	(0.145)	0.180	(0.165)	0.035	(0.094)	0.011	(0.078)	0.025	(0.097)	0.034	(0.145)
Diagnosed 8-14 years ago without	-0.065	(0.044)	-0.070	(0.032)	-0.060	(0.045)	-0.039	(0.132)	-0.055	(0.087)	-0.065	(0.070)	-0.040	(0.074)	-0.025	(0.176)
Newly diagnosed with medication	0.026	(0.042)	0.028	(0.031)	0.023	(0.042)	0.020	(0.077)	0.151	(0.049)	0.121	(0.040)	0.133	(0.043)	0.102	(0.092)
Diagnosed 2 years ago with medication	-0.014	(0.039)	-0.018	(0.030)	-0.012	(0.040)	-0.040	(0.086)	0.030	(0.047)	0.046	(0.042)	0.021	(0.046)	0.056	(0.098)
Diagnosed 4 years ago with medication	-0.111	(0.021)	<i>-0.047</i>	<i>(0.028)</i>	-0.092	(0.025)	-0.078	(0.091)	-0.024	(0.058)	0.021	(0.053)	-0.026	(0.055)	0.016	(0.115)
Diagnosed 6 years ago with medication	-0.015	(0.043)	0.023	(0.041)	-0.016	(0.043)	-0.033	(0.110)	-0.019	(0.062)	-0.022	(0.053)	-0.018	(0.057)	-0.024	(0.122)
Diagnosed 8-14 years ago with medication	-0.025	(0.033)	-0.026	(0.031)	-0.017	(0.033)	-0.120	(0.115)	-0.115	(0.045)	-0.095	(0.040)	-0.085	(0.044)	-0.153	(0.129)
Currently diabetic: not taking medication	-0.013	(0.033)	0.018	(0.035)	-0.007	(0.035)	-0.049	(0.065)	0.048	(0.035)	0.080	(0.036)	0.046	(0.037)	0.115	(0.080)
Currently diabetic: taking oral medication	-0.010	(0.018)	-0.007	(0.015)	-0.010	(0.018)	-0.010	(0.049)	<i>-0.037</i>	<i>(0.021)</i>	-0.027	(0.017)	-0.029	(0.020)	0.043	(0.069)
Currently diabetic: taking insulin	-0.017	(0.021)	-0.029	(0.019)	-0.014	(0.022)	0.017	(0.061)	-0.010	(0.029)	0.003	(0.025)	-0.009	(0.026)	0.016	(0.065)
Frequent physical activity, t-1	0.184	(0.042)			0.164	(0.027)			0.115	(0.055)			0.134	(0.028)		
Number of individuals (n)	2489		2569		2489		2569		1698		1740		1698		1740	
Number of observations (nT)	4359		6452		4359		6452		3096		4006		3096		4006	

Notes: PA, population average (GEE); AME, average marginal effect; LPM, linear probability model. Bootstrapped standard errors, clustered at the individual level, in parentheses for population average models. LPM standard errors are robust to clustering.

The overall sample is restricted to all individuals who ultimately receive a diabetes diagnosis, as well as non-diabetics who are likely to have high blood sugar (described in Table 3, Column 8).

All models include inverse probability weights (IPW), which account for attrition and differential baseline selection probability, where the weight represents the means of these IPW for each period the individual is in the model.

All models include gender, age, age², race, ethnicity, education, income, assets, employer or government insurance, U.S. birth status, and current or new diagnoses of heart disease, lung disease, stroke, or hypertension.

Bolded coefficients represent significance at the 5% level; italics represents significance at the 10% level.

Table 2.7. Average effects for Overweight or obese status by diabetes status and medication.

	(1)		(2)		(3)		(4)		(5)		(6)		(7)		(8)	
	PA Probit	AME	PA Probit	AME	LPM	Coef.	LPM Fixed Effects	Std.Error	PA Probit	AME	PA Probit	AME	LPM	Coef.	LPM Fixed Effects	Std.Error
Panel A: Entire sample																
Panel B: Males																
Newly diagnosed without medication	-0.003	(0.016)	-0.016	(0.019)	-0.001	(0.017)	-0.046	(0.019)	-0.014	(0.026)	-0.023	(0.027)	-0.014	(0.030)	-0.055	(0.031)
Diagnosed 2 years ago without medication	0.030	(0.015)	-0.022	(0.020)	0.024	(0.014)	-0.052	(0.020)	0.073	(0.019)	-0.001	(0.026)	0.048	(0.019)	-0.039	(0.030)
Diagnosed 4 years ago without medication	0.027	(0.018)	-0.013	(0.024)	0.020	(0.016)	-0.045	(0.023)	0.005	(0.027)	-0.023	(0.034)	0.001	(0.024)	-0.055	(0.033)
Diagnosed 6 years ago without medication	0.032	(0.021)	-0.012	(0.025)	0.027	(0.020)	-0.042	(0.025)	0.043	(0.031)	-0.020	(0.035)	0.028	(0.029)	-0.046	(0.033)
Diagnosed 8-14 years ago without	0.006	(0.021)	-0.033	(0.031)	0.005	(0.020)	-0.060	(0.034)	0.027	(0.028)	-0.014	(0.042)	0.019	(0.028)	-0.026	(0.047)
Newly diagnosed with medication	-0.044	(0.012)	-0.048	(0.016)	-0.036	(0.012)	-0.066	(0.017)	-0.046	(0.019)	-0.051	(0.021)	-0.035	(0.017)	-0.064	(0.026)
Diagnosed 2 years ago with medication	0.005	(0.012)	-0.036	(0.017)	0.004	(0.011)	-0.065	(0.018)	-0.001	(0.015)	-0.040	(0.023)	-0.003	(0.014)	-0.059	(0.028)
Diagnosed 4 years ago with medication	0.033	(0.014)	-0.017	(0.018)	0.024	(0.013)	-0.053	(0.019)	0.033	(0.023)	-0.013	(0.025)	0.022	(0.023)	-0.038	(0.029)
Diagnosed 6 years ago with medication	-0.013	(0.016)	-0.048	(0.023)	-0.008	(0.014)	-0.083	(0.023)	-0.015	(0.021)	-0.034	(0.030)	-0.008	(0.018)	-0.060	(0.032)
Diagnosed 8-14 years ago with medication	0.033	(0.013)	-0.052	(0.023)	0.029	(0.011)	-0.094	(0.024)	0.045	(0.020)	-0.027	(0.032)	0.037	(0.018)	-0.059	(0.034)
Currently diabetic: not taking medication	-0.039	(0.009)	-0.023	(0.014)	-0.033	(0.008)	0.002	(0.015)	-0.048	(0.014)	-0.043	(0.019)	-0.034	(0.013)	-0.009	(0.024)
Currently diabetic: taking oral medication	0.008	(0.006)	0.023	(0.010)	0.006	(0.005)	0.022	(0.014)	0.009	(0.008)	0.015	(0.013)	0.006	(0.007)	0.014	(0.020)
Currently diabetic: taking insulin	0.001	(0.007)	0.005	(0.013)	0.003	(0.007)	0.021	(0.013)	-0.014	(0.012)	-0.013	(0.018)	-0.005	(0.010)	0.013	(0.017)
Overweight or obese status, t-1	0.696	(0.015)			0.705	(0.006)			0.647	(0.022)			0.696	(0.010)		
Number of individuals (n)	12841		13008		12841		13008		5742		5819		5742		5819	
Number of observations (nT)	56418		58280		56418		58280		25126		25913		25126		25913	
Panel C: Non-whites																
Panel D: Individuals entering the HRS before age 60																
Newly diagnosed without medication	0.026	(0.044)	0.015	(0.043)	0.021	(0.048)	-0.030	(0.047)	-0.008	(0.016)	-0.008	(0.019)	0.003	(0.017)	-0.044	(0.018)
Diagnosed 2 years ago without medication	0.005	(0.043)	-0.022	(0.046)	0.003	(0.037)	-0.043	(0.045)	0.028	(0.016)	-0.011	(0.019)	0.024	(0.016)	-0.047	(0.019)
Diagnosed 4 years ago without medication	0.035	(0.031)	0.002	(0.037)	0.018	(0.021)	-0.031	(0.039)	0.055	(0.014)	0.008	(0.021)	0.032	(0.012)	-0.036	(0.022)
Diagnosed 6 years ago without medication	0.030	(0.033)	0.043	(0.059)	0.018	(0.024)	0.031	(0.076)	0.017	(0.021)	-0.018	(0.026)	0.015	(0.019)	-0.057	(0.024)
Diagnosed 8-14 years ago without	0.007	(0.032)	0.025	(0.047)	0.008	(0.028)	0.012	(0.055)	0.004	(0.015)	-0.033	(0.026)	0.004	(0.015)	-0.070	(0.027)
Newly diagnosed with medication	-0.025	(0.025)	-0.034	(0.029)	-0.016	(0.023)	-0.043	(0.034)	-0.022	(0.013)	-0.033	(0.017)	-0.012	(0.011)	-0.057	(0.017)
Diagnosed 2 years ago with medication	-0.008	(0.027)	-0.032	(0.032)	-0.005	(0.024)	-0.049	(0.035)	0.023	(0.013)	-0.015	(0.015)	0.017	(0.011)	-0.050	(0.016)
Diagnosed 4 years ago with medication	0.028	(0.023)	0.003	(0.033)	0.023	(0.023)	-0.023	(0.036)	0.030	(0.012)	-0.011	(0.017)	0.017	(0.010)	-0.057	(0.017)
Diagnosed 6 years ago with medication	-0.070	(0.043)	-0.067	(0.052)	-0.055	(0.039)	-0.083	(0.052)	-0.024	(0.017)	-0.055	(0.023)	-0.013	(0.016)	-0.091	(0.022)
Diagnosed 8-14 years ago with medication	0.023	(0.026)	-0.033	(0.039)	0.028	(0.027)	-0.054	(0.042)	0.044	(0.013)	-0.041	(0.024)	0.034	(0.012)	-0.092	(0.023)
Currently diabetic: not taking medication	-0.022	(0.020)	-0.003	(0.028)	-0.022	(0.019)	0.038	(0.033)	-0.046	(0.010)	-0.046	(0.013)	-0.035	(0.010)	-0.002	(0.014)
Currently diabetic: taking oral medication	0.021	(0.013)	0.032	(0.019)	0.015	(0.011)	0.035	(0.027)	-0.006	(0.007)	0.007	(0.010)	-0.006	(0.006)	0.019	(0.013)
Currently diabetic: taking insulin	0.013	(0.012)	0.034	(0.021)	0.009	(0.011)	0.037	(0.028)	-0.007	(0.009)	-0.012	(0.014)	-0.001	(0.008)	0.017	(0.013)
Overweight or obese status, t-1	0.683	(0.044)			0.688	(0.015)			0.674	(0.029)			0.686	(0.010)		
Number of individuals (n)	3475		3540		3475		3540		6521		6631		6521		6631	
Number of observations (nT)	14120		14736		14120		14736		31931		33151		31931		33151	

Notes: PA, population average (GEE); AME, average marginal effect; LPM, linear probability model. Bootstrapped standard errors, clustered at the individual level, in parentheses for population average models. LPM standard errors are robust to clustering.

The overall sample is restricted to all individuals who ultimately receive a diabetes diagnosis, as well as non-diabetics who are likely to have high blood sugar (described in Table 3, Column 8).

All models include inverse probability weights (IPW), which account for attrition and differential baseline selection probability, where the weight represents the means of these IPW for each period the individual is in the model.

All models include gender, age, age², race, ethnicity, education, income, assets, employer or government insurance, U.S. birth status, and current or new diagnoses of heart disease, lung disease, stroke, or hypertension.

Bolded coefficients represent significance at the 5% level; italics represents significance at the 10% level.

Chapter 3

The Role of Family Structure in the Evolution of Health from Adolescence to Young Adulthood by Gender

The incidence of the intact two-biological-parent family has been steadily decreasing since the 1960s. Parents heading non-intact families often face reduced income and less available time, which can lower investments in children's human capital. While the majority of studies examine the impact of family structure on educational attainment, it may also affect children's health. We improve upon previous studies of this by measuring family structure by the timing of maternal relationships, examining effects by gender, and considering persistence in or entry into adverse health states after adolescence. Using four waves of panel data from the National Longitudinal Study of Adolescent Health (Add Health), we estimate both static logit models and discrete-time hazard models to test the hypotheses that growing up without a biological father leads to an increased likelihood of smoking and adverse physical and mental health outcomes during young adulthood, as commonly measured by self-reported health status, obesity, and depression. We find that boys whose biological father was absent during early childhood are more likely than other boys to continue smoking and remain in adverse physical health after adolescence. While adolescent health outcomes are more sensitive to childhood family structure in girls than in boys, many of the adverse effects tend to be limited to adolescence for girls, but to last through young adulthood for boys. We also find that the entrance of step-fathers and cohabiting males diminishes the effects of biological-father absence. Our findings suggest that spending time in non-intact families during childhood may have negative consequences after adolescence, but entry of other males can mitigate some of them

3.1. Introduction

Since 1960, the prevalence of the traditional, two-biological-parent family has been declining in the United States. In particular, the percentage of children living in a single-parent family increased from about 8% in 1960 to over 20% in 1984 (Norton and Glick 1986). This number had jumped to 27% by 1995 (U.S. Census Bureau 1997). In 2004, about 26% of children lived in one-parent households, most of them led by women (Kreider 2008). On average, single-parent households suffer from much lower income than two-parent households. In addition, the time of the single parent, usually the mother, can be spread thin among work, child care and other household production activities.

An extensive body of literature documents the effects of childhood family structure on well-being and human capital outcomes, primarily educational attainment. One of the first, Krein and Beller (1988), document negative effects of living in a single-parent family headed by a woman on young adult educational attainment that varied by length of exposure (duration), period of childhood, and child's gender, with larger negative effects for boys than for girls. In terms of other outcomes, adolescents and older children living in single-parent families tend to be more likely to engage in drug use (Hoffmann and Johnson 1998), sexual intercourse (Lammers et al. 2000) and to have a premarital birth (Hill, Yeung and Duncan 2001), thereby creating a single-parent family in the next generation.

There is a growing literature that suggests that family structure can affect children's health as well. Young children living in single-mother homes have been found to have worse physical, mental, and dental health outcomes than their counterparts in two-biological-parent homes (Angel and Worobey 1988; Bramlett and Blumberg 2007; Dawson 1991; Harknett 2009; Heard, Gorman and Kapinus 2008; Langton and Berger 2011; Montgomery, Kiely and Pappas

1996). Childrens' access to health services, such as physician visits or meeting healthcare needs, can also be compromised in single-parent families (Chen and Escarce 2006; Fairbrother et al. 2005; Leininger and Ziol-Guest 2008). Children who experience transitions away from a two-parent family structure are also more likely to face a myriad of mental and physical health problems, including depression (Brown 2006; Bzostek and Beck 2011; Harknett 2009; Langton and Berger 2011; Mauldon 1990; Spruijt and de Goede 1997).¹ Health problems that originate in childhood and adolescence can often persist and become exacerbated into adulthood (e.g. Case, Fertig and Paxson 2005; Fletcher, Green and Neidell 2010), making it all the more important to understand the correlates for the development of poor health outcomes in adolescents.

In addition to affecting health outcomes themselves, family structure also has been found to be associated with unhealthy behaviors during adolescence and young adulthood, such as drinking, substance use, and a poor diet, which may directly or indirectly affect contemporaneous physical and mental health. Spending time in single-mother families has been shown to increase the propensity of young adults to engage in risky behaviors, especially smoking (Antecol and Bedard 2007; Bjarnason et al. 2003; Fletcher and Sindelar 2012; Francesconi, Jenkins and Siedler 2010; Griesbach, Amos and Currie 2003). Antecol and Bedard (2007) find that paternal absence significantly increases the risk of adolescents engaging in sexual activity, marijuana use, drinking, or smoking before the age of 15. There is also evidence to suggest that children living in single- and step-parent families have poorer diets compared with those living in an intact, two biological-parent household. Johnson-Down et al. (1997) found that children living in single-parent families consume a higher percentage of their calories

¹ While the vast majority of the literature considers the most common non-intact family types (i.e., single-mother families and step-parent families), there is evidence that less common family structures (e.g., single-father or grandparent-headed families) may also have negative consequences for children's physical and emotional well-being (Conway and Li 2011).

from fat as compared to two-parent families. Furthermore, Stewart and Menning (2009) use National Longitudinal Study of Adolescent Health (Add Health) data to find that living in a single-mother family during Waves I and II (1994-1995 and 1996, respectively) decreased the frequency of vegetable and variety of simple sugar consumption, and raised the likelihood of skipping breakfast.

The majority of studies considering the association between family structure and child outcomes use a static measure of family structure only as collected at the time of the survey. The limitation of such a “snapshot” is that it does not capture variation in how much time children spend without their fathers, which varies greatly, especially between such households created by out-of-wedlock births and those resulting from a divorce, nor does it capture the period of childhood during which paternal absence occurs, both shown to make a difference (Krein and Beller 1988). Since many household-level surveys (including Add Health and NLSY) collect information about children of varying ages, the potential exists to observe children living in a single-parent family over a wide range of ages². Some studies are able to obtain richer measures of family structure by using the marital history of the mother (e.g., Antecol and Bedard 2007; Francesconi et al. 2010), or by exploiting the longitudinal nature of some surveys to examine cross-wave differences in family structure (e.g., Bzostek and Beck 2011; Langton and Berger 2011). These studies generally find that family structure can have differential effects on child outcomes depending on the period of life when the child faced the family structure. Likewise, family structure transitions can have effects on children above and beyond the effects of the family structure alone.

² This is important especially in studies using the full sample of Add Health, since the Wave I sample includes children from age 11 to 21. Living in a single- or step- parent family at age 12, for instance, can have different consequences than living in those family structures at age 18.

Despite the strong associations found between family structure and child outcomes, there is substantial heterogeneity in this relationship across countries and studies. For instance, Griesbach et al. (2003) found that living in a step-father household played a larger role in promoting smoking compared to living in a single-mother family in a sample of 15-year old adolescents. In contrast, Antecol and Bedard (2007) found that the presence of a step-father made no difference in adolescents' decisions to try smoking before age 15 compared to remaining continuously in an intact, two-biological-parent family. Furthermore, most studies examining the timing of paternal exit from the household find that departure during early childhood (i.e., before five years old) tends to have the most harmful effects on children's outcomes, including educational attainment and smoking (e.g. Antecol and Bedard 2007; Krein and Beller 1988). However, Francesconi et al. (2010) found that father exit during the later part of childhood played a larger role in explaining smoking behavior. Still others find very little association between family structure and child outcomes. Ginther and Pollak (2004) find that living in a single-parent household plays very little role in explaining educational outcomes after controlling for socioeconomic characteristics. Lang and Zagorsky (2001) also find little negative impact of paternal absence on subsequent child outcomes.

Additionally, little is known about the role of family structure on the trajectory of health and health behaviors subsequent to adolescence, as the majority of the aforementioned studies consider health only as measured at the time of the survey interview. Since health can change over time, it is important to understand both the short- and long- term effects of family structure on child health. Studies using the Add Health data that do consider health outcomes at later periods use binary measures of family structure during adolescence (Wave I), and found little impact on overweight or obesity status (Crossman, Sullivan and Benin 2006) or self-reported

health status (Heard et al. 2008) during young adulthood (Wave III). However, using German Socioeconomic Panel Data, Francesconi et al. (2010) finds that the presence of a lone mother up to the age of 16 raises the hazard for starting smoking by age 21.

No studies, to our knowledge, use dynamic measures to consider the role of family structure (or specifically, paternal absence or the entrance of other males) on the persistence of and risk for negative health outcomes after adolescence. Understanding the long-term consequences of living without a father can have important policy implications. For instance, if living in a single-parent family increases one's likelihood of smoking during adolescence, but the adolescents are more likely to quit shortly thereafter, this would have different implications than if they remained smokers as adults.

This study adds to the literature in several ways. First, we consider health and health behaviors as dynamic outcomes, by considering waves of data spanning adolescence to young adulthood (ages 15-32). Second, we combine the mothers' reports of presence or absence of the child's biological father with her marital history to get the most accurate picture possible of the family structure of the child from birth to age 15, rather than using a "snapshot" measure of family structure at the time of the survey interview. Third, this study examines the effects of family structure by gender. As noted above, Krein and Beller showed more detrimental effects of living in a single-parent family on the educational attainment of young men than of young women. Finally, we consider several measures of physical and mental health status, including self-reported health status, weight status, and depression. Smoking, an important health behavior, is also considered in this study.

3.2. Conceptual Framework

Grossman (1972) originally formulated a model for the production of health, whereby each individual was endowed with a “stock” of health capital that depreciates over their lifetime. Individuals gain utility from good health, which enables them to increase their income through increased wages as well as more productive time in both the labor market and the home. Importantly, investments in health enable a person’s health stock to be maintained or increased. Similarly, negative investments have long-term ramifications for future health. A single parent’s investment in their child’s health may be compromised as a result of financial or time constraints, or a lower average level of education (Haveman and Wolfe 1994). As time and purchased goods may be combined to produce commodities (such as child health) within the household (Becker 1965), child’s health may be negatively influenced by a single-parent family structure.

More recently, the theories of health production and time allocation have been unified by more modern and multi-dimensional concepts such as genetics and neurobiology, as described by Heckman (2007). Heckman stresses the notion that epigenetics (the role of environment on the expression of genetics) underlies much of capacity formation, including the ability to create and maintain one’s stock of health capital.

In order to incorporate the theories of Becker and Grossman, we develop a conceptual framework for the production of health following models put forth by Blau et al. (1996), Ruhm (2004), and Antecol and Bedard (2007). We can express the evolution of health H_t as:

$$(1) \quad H_t = H(F_1, N_1, \mu, S_t)$$

In this model, health (both mental and physical) at time t evolves according to a function which is dependent on parental inputs during adolescence (period 1), including fixed inputs such

as medical care or healthy food (F_1), and parental leisure time during adolescence (N_1). A genetic health endowment, which is determined at conception (μ), and exogenous (or production) shocks (S_t) also enter into the health production function. In this framework, health is increasing in N and F , as increased parental leisure time has health benefits for children directly through time investments, as well as indirectly through increases in household production efficiency. Parents face both budget and time constraints that limit the amount of these positive inputs to the health of their children.

As a result of the reduced time and money resources that can accompany single-parent families, the primary hypothesis of the model is that the less time a child spends in an intact, two-parent household, the lower the probability of having good physical or mental health. Further, we hypothesize that the longer a child spends without his or her biological father, the more likely they will be to experience a decline in their health. The model additionally hypothesizes that health is increasing in household income during adolescence, in that higher financial resources during adolescence will ultimately improve a child's health trajectory. Given the limitations of the data, we do not attempt to estimate Equation 1 structurally. Instead, we estimate several reduced-form models to consider both the short- and long- term implications of family structure on child health outcomes.

3.3. Data and Empirical Strategy

Our analyses are carried out with the National Longitudinal Study of Adolescent Health (Add Health) data³, which broadly surveys health, health behaviors and their contexts throughout adolescence and young adulthood. The first Wave of data (Wave I) consisted of approximately

³ We use the restricted version of Add Health, which includes the full sample of interviewed respondents.

90,000 adolescents in grades 7 through 12 collected in 1994-1995 from primary sampling units of high-schools and “feeder schools,” whose enrollees were expected to attend the high school. These students were ages 12-21 at the time of the survey. A subsample of 20,000 students participated in an in-home questionnaire. The adolescents were interviewed in follow-up waves in 1996, 2001-2002, and 2007-2009 (Waves II, III and IV, respectively).

We use all four waves for our analyses, though key household variables, such as parental relationship history and household income (during adolescence), are only collected from the parent in Wave I. In our study, we include only those adolescents whose biological mothers completed the parent survey (about 85% of parental respondents). We include only those mothers with non-missing answers to a questionnaire about the presence (or absence) of the child’s biological father in the home⁴, and a relationship history for their most recent three marriage or marriage-like relationships⁵ (to ascertain the presence of other males if the child’s biological father was ever absent). We impute income for observations for which it is missing using marital status of the mother in 1994-1995 (e.g. married, divorced, widowed) as well as her occupation. We also restrict our analysis to those adolescents who are between the ages of 15 and 18, who by Wave IV (2007-2009) had become adults ages 27 to 32⁶. We chose to bound the upper age limit during Wave I at 18 to have a sample of adolescents in secondary school and still

⁴ There are three questions regarding the presence of the child’s biological father that are asked of the mother at the Wave I interview. The mother is asked if the child’s biological father currently lives in the household. If the mother responds affirmatively, the question series ends. If the mother reports that the child’s biological father does not live in the household, she is asked if the child ever lived with his/her biological father. The question series ends if the mother reports that the child never lived with the biological father. However, if the mother reports that the child ever lived with the biological father, she is asked what year the child most recently lived with the biological father. We make the assumption that for children without their biological father in the household at Wave I, the last experience living with the biological father included the biological mother as well. We believe the misclassification due to this assumption is likely to be small given that the child lives with the biological mother at the Wave I interview.

⁵ A marriage-like relationship is defined by Add Health as “living with someone as if you were married to him or her when you are not.”

⁶ The adolescents in our sample were born from 1977 to 1980. The Wave I interview took place in 1994-1995 (though only two adolescents in our sample were interviewed in 1994). The Wave IV interview took place in 2007-2009. While there are some individuals who were 27 and 32 at the time of the Wave IV interview (22 and 12 individuals, respectively), the vast majority of respondents were aged 28-31.

living at home. Further, the mothers' relationship status was not available before 1977, which makes it impossible to glean family structure during the child's earliest years for adolescents older than 18. The lower age bound of 15 was chosen to ensure an adequate picture of family structure through the child's adolescent years. We use Wave I sampling weights provided by Add Health to correct for the complex survey design.

Though some other studies (e.g., Crossman et al. 2006) combine single-mother and single-father households together as "single-parent" households, we contend that there are fundamental differences between these two family types. For example, in contrast with fathers, mothers tend to make most of the investments in children's health (Case and Paxson 2001), and single-father families have significantly higher incomes and are less likely to receive child support than single-mother families (Beller and Graham 2003). Additionally, it would be difficult to draw inferences on the role of single-father families on health, since there are relatively few such families in these data.

3.3.1. Variable Definitions

Our independent variables of interest consist of a set of dummy variables capturing the timing of family structure changes, which are calculated from two series of questions involving (a) the history of the father's absence if the father is not living with the mother at the time of the Wave I interview, and (b) the history of the mother's most recent three marriage or marriage-like relationships (including the present such relationship). As indicated above, the Wave I interview took place in 1994-1995, and the mother was asked about her relationship status in each year from 1977 through 1995. Though some surveys do not differentiate between the mother's spouse and the child's biological father, Add Health does. The knowledge is beneficial compared to

having a marital history alone, primarily because the use of the marital history alone requires the assumption that the mother's spouse at the time of the child's birth was the child's biological father. This assumption may be problematic, since children of never-married mothers would by default be classified as having their father never present. That is, biological fathers who are not reported to be in a marriage or marriage-like relationship with the mother would be overlooked. The reliance solely on the mother's marital history to glean the child's living arrangements may further misclassify children if the father's departure from the household precedes the reported end of the relationship.

We employ similar measures of paternal absence as Antecol and Bedard (2007), who use data from the NLSY, though we generate those measures somewhat differently due to the availability of specific information about the biological father's presence in the household (discussed above). If the biological father is living with the mother at the time of the Wave I interview, we consider the adolescent to have always lived in an intact two-parent household⁷. For children whose father was not present at the time of the interview, the mother is asked when the father stopped living with the child. From the series of questions about the presence of the biological father in the household, we generate a series of dummy variables to denote when (and if) the mother ends the relationship with the father. We place this transition into one of four mutually exclusive categories: the father was never present, he left when the child was aged 0-5, he left when the child was 6-10, or he left when the child was aged 11-15. The entry of other males, which we consider to be men whom the mother shared a marriage (i.e., step-fathers) or

⁷ An assumption of this measure is that if the child's biological father is living with the mother at the time of the Wave I interview (1994-1995), he has done so continuously since the child was born. This may misclassify some fathers who left the household and subsequently returned prior to the Wave I interview.

marriage-like relationship with after the biological father left the household⁸, is included in a similar fashion, since many mothers re-marry or cohabit with other men subsequent to the father's departure. It is important to note that most men who enter subsequent to the absence of the biological father tend to stay for a prolonged period of time⁹. To ascertain the entrance of these other males into the household, we use the marital history of the mother, and develop similar measures of the entrance of another male after the departure of the biological father. Since the survey asks mothers about their relationship status by year (rather than by date), we assume that this status persisted for the entire year. Using these mutually exclusive categories for paternal absence, as opposed to a continuous measure of years in a single-parent family, is most useful because it captures the non-linearity of the relationship, by the period of childhood, between family structure and health.

We restrict the analysis to adolescents above the age of 15 as a balance of maintaining sufficient observations and getting a fairly complete picture of the father's presence and mother's marital history across the life of the adolescent (complete histories are only available for those adolescents who were 18 years old in 1995). This is in contrast to some studies that consider family structure up to older ages, such as 16 or 18 (e.g., Francesconi et al., 2010; Krein and Beller, 1988). We explore the potential misclassification from this classification scheme in Table 1. Panel A of this Table shows the patterns of paternal absence and other male (i.e., step-father or

⁸ We aggregate all men who report being in a marriage or marriage-like relationship with the mother as "other males" in our analyses. Step-fathers are traditionally defined as men who enter the household through marriage, while cohabiting fathers imply the entrance of men with more informal relationships with the mother. However, given the way marriage-like relationship is defined by Add Health, it is difficult to meaningfully disaggregate men who enter the household through a marriage, compared to a marriage-like relationship. A limitation of the aggregation of "other males" in this context is that to the extent that marriage is a signal of an increased commitment to the mother and her child, some men included as "other males" may actually have little interest in having a strong role in the child's life. Disaggregated results are similar to the results presented here, and are available from the authors upon request.

⁹ Of all the other males who enter the household during a child's life, 79% of them remain in the household until the child is at least 15.

male who shared a marriage-like relationship with the mother) entrance by age at 1995 for the full sample of adolescents. Overall, relatively few biological fathers leave after the child reaches age 15; about 3% of 18-year-old respondents report having their biological fathers leave from 16-18¹⁰. If we assume that the reports of the 18-year-olds in Add Health reflect the complete parental residential history for 18-year-olds nationally in 1995, the misclassification of family structure due to considering it only up to age 15, rather than to age 18, is likely to be small. Thus, our measures of family structure are not likely to be substantially different from measures that capture family structure up to age 18.

Table 1 also suggests that the majority of fathers who ultimately leave the household do so towards the beginning of the child's life. This characterization of the timing of paternal absence is similar to that suggested by Antecol and Bedard (2007) using NLSY data. However, in comparing the trajectory of paternal absence found in our data to that reported by Antecol and Bedard, we report a smaller number of fathers who are never present, and a substantially smaller fraction of fathers who are ultimately absent. The contrast between our approaches is highlighted in Panel B, where we generate paternal absence measures using Antecol and Bedard's approach. This discrepancy is likely due to a large fraction of unmarried mothers who do not report ever being married (or in a marriage-like relationship), but who, by contrast, do report living with the biological father at the time of the Wave I interview. As a result, a higher number of fathers are reported as never present in Panel B compared with Panel A¹¹. An additional source of this discrepancy comes from some mothers who report never being in a marriage or marriage-like

¹⁰ Of the approximately 2000 eighteen-year-old boys and girls (combined), only about 60 of these individuals experienced the departure of their biological father between ages 16 and 18. This proportion is approximately equal for boys and girls.

¹¹ Antecol and Bedard (2007) also state that there is a relatively high proportion of black women in the NLSY sample they employ, who are more likely to have out-of-wedlock births.

relationship by the time they were interviewed, but nevertheless report that the child's biological father lives with her in the household at the interview¹².

Although we can measure family structure over the course of the child's life, we can only measure mediating socioeconomic variables, like income, at the time of the Wave I interview in 1995. Income of the household is not reported retrospectively over the child's life when they were potentially in a single-parent family, which may misrepresent the "true" income across the evolution of the household (e.g. Wolfe et al. 1996). However, this variable is valuable as a "snapshot," albeit noisy measure of financial resources during adolescence.

Other than income, several household, socioeconomic, and demographic characteristics are collected at Wave I. These include variables collected from the child directly, such as their race, age, birth order and national origin. Most variables, however, are collected from the child's mother, including her education level, U.S. birth status, age, religious attendance, and self-reported health status. Further, maternal smoking is included in the smoking specifications. These observed variables can account for some of the heterogeneity at the household level. Given that family structure may affect children in complex and multi-dimensional ways, we consider several measures of mental and physical health, as well as smoking. These measures parallel those used by other studies in considering the relationship between family structure and child health. Mental health is measured as depressive symptomatology from the Center for Epidemiological Scale for Depression, as well as diagnoses of major depressive disorder in young adulthood. To consider physical health, we utilize the adolescents' self-reported general health scale, which ranges from excellent to poor. Overweight or obese status is assessed based on body mass index thresholds set by the Centers for Disease Control and Prevention. Finally,

¹² There were 153 mothers who reported living with the child's biological father at the Wave I interview, but were never married (or in a marriage-like relationship). The inclusion of a dichotomous variable indicating these mothers as a separate category did not change the results. These results are available from the authors upon request.

smoking is based on self-reported cigarette smoking questionnaires throughout adolescence and young adulthood.

Evidence for depressive symptoms is derived from questions from the Center for Epidemiologic Studies Scale for Depression (CES-D)¹³. The Add Health data contains modified versions of the CES-D scale. In the survey, individuals respond to each depressive symptom in one of four ways: 0 (never/rarely), 1 (sometimes), 2 (a lot of the time), or 3 (most/all of the time). While the regular CES-D scale asks 20 such questions to assess depressive symptomatology, Wave I of Add Health asks only 18 such questions, with the CES-D questions becoming more sparse in subsequent Waves such that by Wave IV, only five such questions are asked. Thus, we would be constrained to five consistently asked questions if we used the CES-D scale as the sole measure of depression for Add Health respondents. However, individuals are asked if they have been diagnosed with depression during both the Wave III and Wave IV interviews. Given that a diagnosis of major depressive disorder is a more relevant endpoint than responses to five CES-D items, we deal with depression in the following way. For Wave I depression, the available answers to the 18 items are averaged together and rescaled to have a theoretical maximum score of 60. Following Roberts, Lewinsohn and Seeley (1991), boys with scores 22 or above, along with girls with scores of 24 or above are classified as depressed. In Waves III and IV, a diagnosis of depression is used. As Wave II does not contain information about depression diagnoses, it is omitted from the longitudinal analyses of depression.

Self-reported health status is measured by an ordinal, five-point categorical scale that the adolescent answers during each survey wave, which can take values of excellent (1), very good (2), good (3), fair (4), or poor (5) health. This variable is commonly used in the literature to represent physical health. Though there are some limitations to this measure, the most serious of

¹³ See Radloff (1977) for a complete description of the CES-D scale.

which includes non-random measurement error (Crossley and Kennedy 2002), the measure is widely agreed upon to be a strong predictor of subsequent mortality (in adults) and health care utilization¹⁴. Following several studies, we condense self-reported health status into variables containing more than one category (e.g. Contoyannis and Li 2011; Polsky et al. 2009).

Specifically, we group very good or excellent health together as a single variable. The good, fair, and poor health categories are similarly grouped as a single, adverse health variable. This variable is dichotomized in this fashion in order to allow for implementation into a logit framework¹⁵.

We define an individual to be overweight or obese based on their body mass index (BMI)¹⁶, a commonly used indicator of weight-to-height status. We classify adolescents (Wave I) as obese based on the Centers for Disease Control and Prevention (CDC)'s age- and gender-based percentiles for children ages 2-20 (Kuczmarski et al. 2002). We define an individual who is less than 18 to be overweight or obese if their BMI exceeds the 85th percentile for their age and gender, respectively. Under this classification, we find that about 25% of the pooled sample is overweight or obese at the time of the Wave I interview, which is consistent with other estimates of the prevalence of overweight status or obesity in adolescents. As the individual transitions into young adulthood (i.e., after age 18), we use the CDC's definition of $BMI \geq 25$ to define if an individual is overweight or obese. Almost 46% of the sample meets this threshold by Wave III. By Wave IV (corresponding to when the adolescents are from ages 27 to 32), approximately 66% meet the classification for overweight status or obese status. This trend is consistent with

¹⁴ See Contoyannis and Li (2011) for a more detailed discussion of the strengths and limitations of self-assessed health measure.

¹⁵ Similar results were obtained using the full five-point scale in an ordered logit framework (for static analyses). These estimates are available from the authors upon request.

¹⁶ BMI is calculated as weight (in kilograms) divided by height (in meters), squared. Height and weight are self-reported by the adolescent in Wave I, and measured directly by Add Health in Wave II. We correct potential measurement bias in Wave I using a linear prediction model by using Wave II data to compare differences in self-reported and measured BMI.

the recent trends in the United States regarding these outcomes¹⁷.

While many lifestyle variables, including smoking, drinking, breakfast consumption, and physical activity, are important to the evolution of health (and mortality, in adults), we focus on smoking for two reasons. First, it is one of the few lifestyle variables that are consistently and comprehensively defined throughout the entire course of Add Health. Second, some lifestyle variables (e.g., drinking or sexual activity), unlike smoking cigarettes, are not unequivocally detrimental to health as children grow into adulthood. Further, the relevance of some of these activities as harmful behaviors wanes as individuals grow from adolescence into adulthood.

3.3.2. Empirical Strategy

As discussed in Section 2, health behaviors and outcomes can be expressed as functions of individual, household, and unobserved factors. Our outcomes of interest are self-reported health (dichotomized as two variables), depressive symptomatology, overweight or obese status and cigarette smoking. As previously mentioned, we segment our analyses by gender to reflect fundamental differences in their responses to social disadvantage and paternal absences, as well as biological differences in health outcomes themselves.

To be comparable with the existing literature examining family structure and health, we begin our analyses by considering the direct effect of family structure on measures of health during Wave I (adolescence), after controlling for total income in 1995:

¹⁷ These trends are very similar if the more conservative measure of obese status is used (corresponding to BMI-for-age percentile of 95 or above for children below age 18, or a BMI ≥ 30 for adults). For the pooled sample (including both boys and girls), about 8.5% of adolescents are classified as obese in Wave I. The prevalence of obesity increases to approximately 23% in Wave III, and 37% in Wave IV. The use of obesity alone as a dependent variable does not change the nature of the results. The prevalence of overweight and obesity documented here are consistent with national prevalence estimates obtained from the 2007-2008 National Health and Nutrition Evaluation Survey, in which approximately 68% of adults were found to be overweight (including obese), and 34% of adults found to be obese (Ogden and Carroll 2010).

$$(2) \quad H_i^* = \alpha + \beta FS_i + \phi Income_i + \gamma Z_i + u_i.$$

In equation 2, FS is a time-invariant vector of mutually exclusive categories that represent the timing of paternal absence from the household through age 15. The variable $Income$ represents the income of the household during Wave I in 1995, with a variable that denotes if the individual's income observation was imputed¹⁸. The vector Z represents demographic and other socioeconomic variables that may influence smoking, mental and physical health during adolescence¹⁹.

We then consider the role of family structure in the probability of entry into, or exits from, “bad” health states after Wave I. In the example of smoking, this would be the probability of a nonsmoking adolescent choosing to smoke subsequently, along with the subsequent quitting probability of smoking adolescents. Using Waves II²⁰ through IV of Add Health, we estimate several discrete-time hazard models to clarify the relationship between family structure and long-term transitions in health outcomes and smoking. A discrete-time hazard approach has two major advantages in this context. First, family structure may have different effects in promoting upward health transitions (i.e., to better health states) versus downward transitions (i.e., to poorer health states). Second, in adolescents and young adults, a fraction of respondents may never report having an unfavorable health outcome (i.e., good or lower health status) or start smoking during their course in Add Health, resulting in right censoring of the data.

¹⁸ Household income during Wave I is imputed for missing observations, as approximately 10% of the mothers did not report income in the survey.

¹⁹ The incorporation of sibling fixed effects is a potential way to address some of the unobserved heterogeneity. However, the variation in age of siblings is relatively small; in Add Health, the average age difference between full siblings is only slightly over two years (e.g. Jacobsen and Rowe 1999). Therefore, the limited inter-sibling variation in family structure makes it difficult to use as an identification strategy.

²⁰ For the quitting specifications, individuals could have started smoking before the Wave I interview and had reported quitting smoking by the time they were interviewed in Wave I. This is in contrast to other health behaviors in outcomes which are not observed prior to the time of the Wave I interview. Therefore, the time variable starts at Wave I in the quitting specifications.

To estimate a hazard model, the data is reconstructed to reflect a “person-period” data set, in which each individual in the sample has multiple records, corresponding to each discrete period (i.e., survey wave) in which the individual is observed. The time variable started after Wave I, which represents the time when adolescents’ health and health behaviors can potentially change. A logistic model is then fitted to the transformed data, in the form of Equation 3:

$$(3) \quad Y_{it} = \delta + \varphi FS_{i1} + \tau Income_{i1} + \theta Z_{i1} + e_{it} \quad (i = 1, \dots, N; t = 2, \dots, T).$$

For adolescents reporting a “good” health state in Wave I, the outcome variable Y represents an entrance, in period t , into a “bad” health state. An additional model is also estimated for individuals reporting “bad” health states in Wave I. In this case, Y represents an entrance, in period t , into a “good” health state. In the example of smoking, Y would represent if an adolescent who regularly smoked by Wave I had quit by time t . In the case of adolescents who did not regularly smoke by Wave I, Y would represent if that individual started regularly smoking by time t . Dichotomous variables are included to indicate survey periods.

While the most efficient estimator for the model shown in Equation 3 would incorporate individual-level unobserved heterogeneity (i.e., “frailty”), the use of such an estimator would make strong assumptions about the intertemporal nature of the data. Specifically, the intra-class correlation may not be a good approximation to the data generating process, especially since the data spans almost 13 years, with long gaps in between Waves II, III, and IV. In this case, inconsistent inference statistics may be obtained. Additionally, it is difficult to incorporate sampling weights in random-effects estimators, which makes the results difficult to generalize to the U.S. population at large. Therefore, we estimate the models without individual unobserved

heterogeneity, but employ a rich set of maternal and household characteristics to obtain plausible estimates for the effects of paternal absence and other male entrance on the outcomes of interest.

As with the descriptive statistics and static models, the hazard models are considered separately for boys and girls. The same vector of time-invariant covariates (Z) included in Equation 2 is included in Equation 3. In all models, marginal effects, evaluated at the sample means, are reported.

3.4. Results and Discussion

3.4.1. Descriptive Analysis

Tables 2 and 3 show descriptive statistics for measures of health status, health behaviors, family structure, and demographic and socioeconomic characteristics by father presence or absence and gender. About 30 percent of children spent some time without the biological father in the home during the first 15 years of their lives, though girls are more likely (32%) to have their fathers leave (before age 16) compared with boys (29%)²¹. The mothers of both boys and girls spending some time without their biological fathers tend to work more (i.e., full-time), but to have lower household incomes and educational attainment. Children whose biological father ever left were more than twice as likely to be black. Mothers living without the child's father at the time of the Wave I interview also reported higher rates of tobacco consumption and tended to rate their physical health more poorly. These differences are all statistically significant at the 1% level.

With respect to the timing of paternal absence, fathers who were not present at the Wave I interview tended either to have never been present, or to have left during the earlier part of the

²¹ This difference is significant at the 5% level. Note that the rates of paternal absence we find here are lower than those obtained from other data sources (Antecol and Bedard report a rate closer to 50%).

child's life. Of girls who spent any time without their fathers, about 33% of them had fathers who were never present in the household at all throughout their life. Another 31% had fathers who left before the girl reached 6 years old, when she usually enters first grade. A smaller percentage, left between the ages of 6 and 10, when they were in elementary to middle school grades, and from ages 11-16, as the adolescent transitioned from middle school into high school. Over half of all boys and girls aged 15 to 18, whose father ever left, will experience the arrival of at least one other male, whereas about one in ten, will experience the entry of more than one man. Other males tend to arrive somewhat uniformly across the child's life up to age 15, with the largest proportion arriving when the child is between 6 and 10 years old. Boys face a similar pattern of paternal absence, though fathers are more likely to be present for a small portion at the beginning of the boy's life, as opposed to being absent for the entirety of his life.

Concerning health outcomes of adolescents, individuals spending time without their fathers tend to report their health as being less well in every outcome except for overweight or obese status. Both boys and girls spending time without their biological fathers report higher rates of smoking, as well. Additionally, there is a larger disparity in the reporting of physical health for girls than boys by family structure. For instance, in adolescent girls spending time without their father, 58% reported excellent or very good health during adolescence, compared with about 69% of girls whose father never left before age 16, a more than ten percentage point differential. In adolescent boys, there was approximately a six percentage point differential, which is still considerable.

Depressive symptomatology, as measured by CES-D scores for depression propensity, has similar differentials by gender and family structure. Adolescent girls tend to report higher

rates of depressive symptomatology than boys do²². Further, there are differences between boys and girls who spend time without their biological fathers compared to those who do not. In adolescence, about 82% of girls age 15-18 expressed some depressive symptoms (a CES-D score greater than 0, not reported), and about 10% meet the criteria for depression, with a large differential between those spending time without their biological fathers and those who do not. Using the CES-D cutoff scores, we estimate that about 8% of the pooled Wave I adolescent sample age 15 or above (including both boys and girls) is depressed, which is consistent with previous estimates of depression prevalence in these data (e.g., Fletcher 2009; Goodman and Whitaker 2002).

Girls and boys spending time without their fathers are more likely to report regularly smoking (i.e., smoking at least one cigarette per day for 30 days), with girls' smoking habits being more sensitive to paternal absence than boys. Overweight and obesity outcomes do not appear to be significantly affected by paternal absence in either boys or girls.

We now turn to the persistence and evolution of these outcomes as individuals transition out of adolescence, which is depicted in Figures 1 and 2. In Figure 1, the quitting trajectory (Panel A) is depicted for adolescents who had ever regularly smoked (i.e., at least one cigarette per day for 30 days) at the time of the Wave I interview, along with the “start” smoking trajectory (Panel B) for adolescents who had never regularly smoked at the Wave I interview. A similar analysis is performed for the overweight or obesity outcomes (Panels C and D)²³.

The trajectory of quitting for adolescents who smoked varies by paternal absence (Panel A), in that girls without a father are consistently less likely to quit smoking subsequent to

²² This difference is significant at the 1% level.

²³ While we only focus on the right tail of the BMI distribution that is associated with overweight or obesity status, supplementary analyses revealed that family structure is not associated with occurrences on the left tail of the BMI distribution, such as underweight status (these results are available from the authors upon request).

adolescence as compared with girls whose fathers never left. Boys who were smokers during Wave I tend to quit less frequently during Waves III and IV if they had absent fathers. Likewise, girls who do not report regularly smoking during adolescence (Panel B) tend to start smoking at higher rates starting at Wave III if they spent any time without their father growing up. Weight outcomes do not appear to be substantially sensitive to living without one's father, but appear to be highly persistent over time. Panel C shows that very few ($< 5\%$) of adolescents who were classified as overweight or obese during adolescence become normal weight during adulthood at Wave IV. Panel D suggests that there is also a considerable upward trend in weight status for normal-weight adolescents over time, as 60% of these individuals will be classified as overweight or obese as adults.

Figure 2 displays similar trajectories for other health outcomes, including depressive symptomatology, and self-reported health status²⁴. Similar to the case of smoking, adolescents living without their fathers are more likely to stay in good or lower self-reported health states and less likely to stay in very good or excellent health states, with a much larger family structure differential for girls than for boys. These descriptive analyses suggest that while paternal absence can affect health outcomes and smoking substantially during adolescence (with the exception of weight status), there may be long-term consequences of living without a father that extend into adulthood, as well.

3.4.2. Estimation Results: Baseline Specifications

In order to understand the basis by which family structure influences baseline health outcomes during adolescence, we estimate a series of static logit models in which measures of

²⁴ Given that depression diagnostic information is not available prior to Wave III, we do not consider Wave I or Wave II in the longitudinal analysis of depression.

physical and mental health are regressed on paternal absence, and a vector of household covariates. Table 4 presents specifications that represent different sets of measures of family structure over the child's life from birth to age 15. The first (Panel A) includes a binary measure denoting if the father was ever absent during the child's life up to age 15. Panel B adds an additional measure indicating if another male entered the household up to age 15. Panel C disaggregates the binary measure in Panel A, to include measures of paternal absence according to the period of the child's life during which the absence occurred. The specifications reported in Panel D employ these same measures, but add several disaggregated variables related to the entrance of a other males that is similarly stratified by the period of childhood during which the other male's entrance occurred. We include all specifications for two reasons. First, many studies considering the effects of single-parent families and/or paternal absences often do not include any measure of other males, or include a crude measure (i.e., single dichotomous variable) if they do so at all. The results reported in Panels A and C are comparable to those studies. Second, a very large fraction (>50%) of children whose biological father ever left ultimately see the entrance of another male. The addition of these "other male" measures may therefore substantially mediate the effects of living without the biological father. Panels A and C therefore represent the full effect of living without the biological father, whereas Panels B and D reflect the effect adjusted for the entrance of other males.

As suggested in Tables 2 and 3, children growing up without a father are much more likely to have their mother smoke compared to those growing up in intact families. Further, several studies have suggested parental smoking is strongly related to the risk for smoking initiation in adolescents (Francesconi et al. 2010; Gilman et al. 2009; Göhlmann, Schmidt and Tauchmann 2010). Therefore, we explicitly examine maternal smoking as an additional variable

that may be important in the relationship between family structure and smoking habits of affected children. While Antecol and Bedard (2007) found very large marginal effects for paternal absence on smoking (almost 20 percentage points), they did not include a measure of maternal smoking. Other studies (e.g., Francesconi et al. 2010) include such measures and find somewhat lower effects. As a result, we consider regressions with and without the inclusion of the maternal smoking variable²⁵.

The findings in Panels A and C of Table 4 suggest that the overall effect of paternal absence serves to increase the probability of ever smoking regularly by the time of the Wave I interview, and of reporting symptoms suggestive of depression. Furthermore, living without a father decreases the likelihood of reporting very good or excellent health status. As expected from the descriptive analysis, paternal absence has little effect on the probability of being classified as overweight or obese as an adolescent, and it actually lowers this likelihood in boys. The largest effects are seen for females, especially in regards to the smoking outcome. These general findings are consistent with previous literature that suggests that spending time without the father can promote adolescent smoking and be harmful to physical and mental health during this time. The effect of paternal absence, especially during the early part of the child's life is slightly diminished with the inclusion of the maternal smoking variable. However, its inclusion does not strongly affect the magnitude or significance of the coefficients on paternal absence for girls (columns 1 and 2), but does so somewhat for boys (columns 6 and 7).

Turning to the inclusion of other males (Panels B and D), the effect of paternal absence is diminished in magnitude and significance, particularly for the smoking measures. When other

²⁵ The inclusion of the maternal smoking variable in other models did not substantially affect the magnitudes or significance of the family structure variables.

males are included (Panels B and D), the original effect of paternal absence diminishes substantially in both magnitude and significance in the smoking specifications for both genders (columns 1, 2, and 6), though the effect of biological father absence remains marginally significant. Other male entrance was insignificant in the aggregate (Panel B), but was significant in the disaggregated form in girls (Panel D, columns 1 and 2), if the entrance took place during the early part of the girl's life. By contrast, in the self-reported health, obesity and depression specifications in girls, the entrance of another male has a relatively small effect on health, and does not substantially alter the effect of biological father absence. However, in boys, other male entrance diminished the effect of biological father absence in all specifications. For the depression and self-rated health models in boys, the inclusion of other male variables rendered the biological father absence variables insignificant.

In terms of other demographic factors which are included in these models, but not reported in Table 4, increasing age increases the propensity to either try smoking ever or regularly. Non-white individuals, including Hispanics, Blacks and those of other races, tend to report smoking less. Other influential variables included maternal education, which lowered the probability of smoking and depression, and increased the probability of reporting very good or excellent health, though only in girls. Maternal health status also affected physical and mental health outcomes of the adolescents, as well. Again, girls seemed to be more affected than boys by these mother-reported measures.

3.4.3. Estimation Results: Discrete-Time Hazard Models

While the baseline models considered the role of family structure in health outcomes and behaviors as measured at the time of the Wave I interview (i.e., during adolescence), it is

conceivable that the relationship between family structure and these outcomes may be more long-term. For instance, adolescents who reported smoking during Wave I may have differential probabilities of quitting depending on the absence of their biological father or the presence of another male. Family structure may also play a role in the probability for non-smoking adolescents to start smoking after Wave I. Similar transitions may occur in other health outcomes, as well. To investigate these possibilities, we estimate a series of discrete-time hazard models.

We begin by considering the dynamics of smoking after adolescence. Smoking is measured uniquely in the Add Health in that respondents are asked if they ever regularly smoked prior to the Wave I interview. Likewise, they are asked if they ever quit smoking prior to Wave I. Because there are adolescents who already started and stopped smoking by the time they were interviewed in Wave I, we are able to use the full Add Health panel (i.e., Waves I through IV) to analyze quitting behavior. The descriptive analysis suggested that paternal absence may affect quitting or starting behavior during young adulthood. From Figure 1, both boys and girls who smoked during Wave I were less likely to quit during young adulthood (Waves III or IV) if their biological father was ever absent. Likewise, adolescents who did not smoke during Wave I were more likely to subsequently start smoking during Waves III or IV if their biological father was ever absent.

The discrete-time hazard results for smoking are presented in Table 5, following the same four-panel structure presented in Table 4. In contrast to the very strong associations between paternal absence and adolescent smoking reported in Table 3, the results reported in Table 5 show a more limited role for paternal absence on smoking dynamics after Wave I. However, there are some important effects to be noted. In considering quitting behavior, the relevant

sample is adolescents (aged 15 – 18) who report having ever regularly smoked by the time they were interviewed in Wave I (i.e., the sample at-risk for quitting). Similarly, starting behavior is analyzed for the sample of adolescents who did not report having ever regularly smoked at the Wave I interview, representing the sample at-risk for starting to regularly smoke.

The results presented in Panels A and B (columns 1, 2, 5 and 6) of Table 5 suggest that paternal absence or other male entrance has no significant net effect in promoting quitting behavior for adolescent girls or boys who ever regularly smoked at Wave I. However upon disaggregation by child's age of paternal absence (Panel C), the results in column 1 suggest that girls who ever smoked regularly during or prior to adolescence had an approximately 4.7 percentage point lower hazard of quitting smoking if their father left between ages 11 and 15. This effect is robust to the inclusion of other male in Panel D. The entrance of other males influence girls' hazard of quitting in a positive direction (though not significantly so), which increases the magnitude of the effect of paternal absence. In fact, in column 1, Panel D, girls who smoked regularly in Wave I and whose biological father had left between ages 0 and 5 had an approximately 5 percentage point lower hazard to quit. This effect, however, was rendered insignificant with the inclusion of the maternal smoking variable (column 2). Furthermore, in disaggregated paternal absence in boys who ever reported regularly smoking during the Wave I interview (Panel C; columns 5 and 6), those whose father was never present had a 5 percentage point lower hazard of quitting. Upon the inclusion of other males (Panel D), however, the effect becomes insignificant. This is likely due to the negative direction of the marginal effects of the entrance of other males on quitting behavior in boys. By contrast, in girls, the inclusion of other males actually heightens the negative effects of paternal absence, since the presence of another

male actually increases the likelihood of quitting (though, having more than one such other male decreases this likelihood).²⁶

In addition to family structure potentially affecting quitting propensities among adolescent smokers, it is conceivable that paternal absence (or other male entrance) may serve to increase the likelihood that adolescent non-smokers will subsequently start smoking. In girls (columns 3 and 4), aggregated paternal absence (Panel A) has a marginally significant effect of 2.4 percentage points in raising the hazard of regularly smoking subsequent to Wave I for adolescent non-smokers. However, this marginal result is rendered insignificant with the inclusion of other males (Panel B) and with the disaggregation of the paternal absence measures (Panel C). However, in Panel C, all the coefficients are positively signed and all are 2 percentage points or greater, suggesting that paternal absence plays a positive, albeit small, role in causing girls to start smoking after Wave I. With the inclusion of other males (Panel D), nonsmoking adolescent girls whose father left between ages 6 and 10 were marginally more likely to start smoking by 5 percentage points. Other male entrance generally plays a limited role in affecting the hazard for starting smoking in both boys and girls; however, girls who face the entrance of a step-father or cohabiting man between ages 11 and 15 (Panel D, columns 3 and 4), are significantly less likely to start smoking after Wave I. The highest propensity for starting smoking occurred during Wave IV, while the lowest was during Wave II (results not shown).

Combined, the results from our discrete-time hazard models for smoking revealed that paternal absence has a more limited role in smoking transitions after Wave I than it does during

²⁶ In the quitting specifications, boys and girls quit less in Wave II compared to Wave I. As with the baseline models (Table 4), maternal smoking behavior plays a large role in the quitting trajectory in girls. Omitting maternal smoking yields a larger effect size for paternal absence on quitting behavior in girls. However, maternal smoking plays no significant role in quitting behavior in boys. Maternal smoking also marginally increases the propensity to starting smoking in boys, but not girls.

Wave I. However, paternal absence for the entirety of a child's life up to age 15 significantly reduces the propensity of boys to quit smoking, while paternal absence later in adolescence significantly reduces the likelihood of girls quitting, by approximately 5 percentage points in both cases. Additionally, in girls, aggregate paternal absence also appears to have a small, positive propensity in the hazard for starting to smoke after Wave I. These results have important implications, especially since family structure has never been studied (to our knowledge) in the context of quitting smoking. While quitting smoking is a difficult task for many previous smokers, it appears that paternal absence may increase the difficulty. While adolescent smoking can have health consequences, long-term smoking may put these individuals at even greater risk for adverse health outcomes later in adulthood, and possibly mortality.

Similar to the case of smoking, the descriptive analysis in Figure 2 suggests that adolescents differ in their long-term physical and mental health depending on the presence or absence of their biological father. Tables 6 (girls) and 7 (boys) show the effect of family structure on physical and mental health, which as with smoking, can be dynamic, and may improve or decline subsequent to Wave I. The measures of family structure are presented similarly to Tables 4 and 5. We begin by considering the role of family structure on the improvement of self-reported physical health status after adolescence, that is, the transition to "very good" or "excellent" among adolescents reporting "good" or lower health status at Wave I (column 1). In girls (Table 6), family structure has a relatively limited effect on this upward transition. While girls whose father was never present had a negative (though insignificant) hazard for this upward transition (Panel C), if the biological father left during the latter part of the girls' life (between ages 11-15), girls were significantly more likely to improve their health after Wave I. However, in boys (Table 7), there is an aggregate, marginally significant negative

effect for paternal absence (Panel A), though this effect disappears upon the inclusion of other males, which appears to dominate the negative effect albeit not reaching significance (Panel B). In Panel C, also among boys, early paternal absence has a strong negative effect on this upward transition: boys whose biological father left between the ages of 0 and 5 were 13 percentage points less likely to report “very good” or “excellent” health after Wave I compared with boys growing up in intact, two-biological-parent families. In boys, the results reported in Panel D suggest that other male entrance diminishes the effect of biological father exit, paralleling the results of other models discussed earlier. Overall, these results suggest that boys who “start” in good or lower health states during adolescence are more likely to stay that way if they experienced the absence of their biological father growing up, especially during the early portion of their lives.

Alternatively, it is conceivable that adolescents reporting very good or excellent health may subsequently report lower health outcomes after Wave I (Tables 6 and 7; column 2). However, there is little evidence to suggest that family structure plays a role in this transition, in either boys or girls²⁷. Though the aggregate effect of paternal absence on the transition to lower health status is in the positive direction for both boys and girls, the effects are insignificant, even in the disaggregated state. However, in boys, upon inclusion of disaggregated measures of other male entrance (Panel D), the entrance of another male between ages 11 and 15 had a positive effect on the hazard of reporting worse health after Wave I. In girls who reported very good or excellent health during Wave I, those who had more than one other male were less likely to “drop” the rating of their own health to good, fair or poor status.

²⁷ These results notwithstanding, static models considering young adult outcome as a function of father absence and control variables revealed that girls whose father was ever absent were consistently less likely to report being in very good or excellent health from Wave I to Wave IV. These results are available from the authors upon request.

The other physical health measure that we consider, overweight or obese status (columns 3 and 4) does not appear to be substantially affected by family structure. This is notwithstanding the high likelihood that adolescents classified as normal weight status will ultimately become classified as either overweight or obese (see e.g., Figure 1). In girls, father absence during ages 6 and 10 played a protective role in lowering the hazard for normal weight adolescents to be classified later as overweight or obese (Table 6, Panels C and D; column 3). In boys (Table 7), family structure played no significant role in changes in weight status (in either direction) subsequent to adolescence. The most important demographic characteristics that contributed to transitions between overweight/obese status and normal weight were race and ethnicity, especially in girls (results not shown). Most notably, Black females tended to be less likely to lose weight if they were obese or overweight during adolescence, and more likely to become overweight or obese if they were classified as normal weight during adolescence.²⁸

Finally, we consider the hazard of depression non-diagnosis and diagnosis among adolescents who were classified as depressed or not depressed, respectively (columns 5 and 6). As for the weight outcomes, paternal absence tended to have little impact on depression diagnosis subsequent to adolescence in either boys or girls. However, the direction of the aggregate effects of father absence (Panel A) for both boys and girls (though insignificant) tended to shift both depressed and non-depressed adolescents towards a depression diagnosis in Waves III or IV. Demographic and socioeconomic factors that influenced depression transitions (not shown) in girls included minority status and a higher income, both associated with a lower

²⁸ The baseline hazard for transitioning into overweight or obese status was higher in Wave IV relative to the earlier periods. Similar trends were found for adverse self-reported health status (i.e., good or lower health).

likelihood of being diagnosed with depression. In boys, having a U.S.-born mother had a strong role in increasing the propensity to receive a depression diagnosis.²⁹

These results support the notion that family structure can have small, but long-standing implications on health past adolescence. Additionally, we find that some of the effects of paternal absence on health diminish substantially with the addition of other males, especially with regard to smoking. This suggests that studies which ignore the issue of step-fathers or cohabiting males, or alternatively, use a very crude measure to capture these fathers, may be overstating the role of paternal absence on some youth outcomes.

3.5. Conclusions

We explored the role of family structure—specifically biological father departure and step-father and cohabiting- male entrance—in the evolution of smoking, and physical and mental health outcomes from adolescence into adulthood. This study addressed several gaps in the literature concerning the interaction between family structure and child outcomes generally, and health outcomes and behaviors in particular. First, we construct a more accurate set of measures reflecting the child’s living arrangements through age 15 by using maternal questionnaires that explicitly dealt with the biological father, rather than using a marital history alone. Second, we incorporate other males into the analyses, whose presence is frequently ignored. Third, we consider the impact of family structure on the long-term trajectory of health outcomes and smoking. Finally, we consider these effects separately by gender.

From the descriptive analysis, the negative association between paternal absence and adolescent health is seen: children growing up without their father tend to have higher rates of

²⁹ The hazard for transitioning into a depression diagnosis did not exhibit significant duration dependence.

depressive symptomatology and a lower subjective rating of their own health. The unadjusted disparities in health outcomes between children growing up in intact families and those without the father continue as the adolescents grow into young adulthood, especially in regard to smoking and self-reported health status.

Baseline estimates of paternal absence on adolescent health paralleled findings in the literature, revealing that adolescents spending time without their father were more likely to report worse physical health status, higher rates of smoking, and higher rates of depression. However, we find that during adolescence, girls tended to be much more sensitive to paternal absence than boys. We find little evidence that spending time without their biological fathers has any adverse effect on adolescent weight outcomes, including overweight and obese status (combined).

To examine the long-term effects of paternal absence (i.e., subsequent to the Add Health Wave I interview), we used discrete-time hazard models and found that adolescents spending time without their fathers were more likely to transition from better health to worse health, though the effect differed by outcome. Furthermore, boys who regularly smoked during adolescence were more likely to continue smoking subsequently if their father had left during the early portion of their lives. However, maternal smoking mediated the effect of paternal absence on smoking persistence. Similarly, boys reporting good, fair, or poor health during adolescence were more likely to stay in these categories if their father was absent during the early part of their lives. In girls who did not report smoking in adolescence, there was a marginal increase in starting to smoke regularly subsequent to Wave I if their father had ever been absent.

Overall, our results suggest that while most of the adverse health consequences of paternal absence are evident by adolescence or earlier, there is evidence to suggest that paternal

absence may serve a small role in increasing the persistence of smoking and worse self-reported health into young adulthood. Policy interventions directed at smoking cessation and the improvement of adolescent health may be especially beneficial among children growing up without a father.

Our study has several limitations. First, because of sample size limitations, we choose to measure family structure only up to the age of 15. While we do not believe this leads to substantial misclassification, it can lead to smaller effects for individuals whose biological father exited, or step-father entered, the household from 16-18. Second, most of the time-varying “pre-disruption” variables, capturing household heterogeneity before the father left, are unobserved, the most important of which is income. The measure of income that we do have is likely a noisy one, particularly for families in which the biological father left. Thus, it is possible that unobserved factors associated with paternal absence and child health may influence the observed results, especially in the static models. Finally, the data we use has large gaps, especially between Waves II and III, and Waves III and IV, which makes studying the dynamics of health and health behaviors difficult using more sophisticated longitudinal statistical techniques.

3.6. References

- Angel, R. and J.L. Worobey. 1988. "Single Motherhood and Children's Health." *Journal of Health and Social Behavior* 29(1):38-52.
- Antecol, H. and K. Bedard. 2007. "Does Single Parenthood Increase the Probability of Teenage Promiscuity, Substance Use, and Crime?" *Journal of Population Economics* 20(1):55-71.
- Becker, G.S. 1965. "A Theory of the Allocation of Time." *The Economic Journal* 75(299):493-517.
- Beller, A.H. and J.W. Graham. 2003. "The Economics of Child Support." in *Marriage and the economy: Theory and evidence from advanced industrial societies*, edited by S. Grossbard-Shechtman. Cambridge, UK: Cambridge University Press.

- Bjarnason, T., A.G. Davidaviciene, P. Miller, A. Nociar, A. Pavlakis, and E. Stergar. 2003. "Family Structure and Adolescent Cigarette Smoking in Eleven European Countries." *Addiction* 98(6):815-824.
- Blau, D.M., D.K. Guilkey, and B.M. Popkin. 1996. "Infant Health and the Labor Supply of Mothers." *The Journal of Human Resources* 31(1):90-139.
- Bramlett, M.D. and S.J. Blumberg. 2007. "Family Structure And Children's Physical And Mental Health." *Health Affairs* 26(2):549-558.
- Brown, S.L. 2006. "Family Structure Transitions and Adolescent Well-being." *Demography* 43(3):447-461.
- Bzostek, S.H. and A.N. Beck. 2011. "Familial instability and young children's physical health." *Social Science & Medicine* 73(2):282-292.
- Case, A., A. Fertig, and C. Paxson. 2005. "The Lasting Impact of Childhood Health and Circumstance." *Journal of Health Economics* 24(2):365-389.
- Case, A. and C. Paxson. 2001. "Mothers and Others: Who Invests in Children's Health?" *Journal of Health Economics* 20(3):301-328.
- Chen, A.Y. and J.J. Escarce. 2006. "Effects of Family Structure on Children's Use of Ambulatory Visits and Prescription Medications." *Health Services Research* 41(5):1895-1914.
- Contoyannis, P. and J. Li. 2011. "The Evolution of Health Outcomes from Childhood to Adolescence." *Journal of Health Economics* 30(1):11-32.
- Conway, K. and M. Li. 2011. "Family Structure and Child Outcomes: A High Definition, Wide Angle "snapshot". " *Review of Economics of the Household*:1-30.
- Crossley, T.F. and S. Kennedy. 2002. "The Reliability of Self-assessed Health Status." *Journal of Health Economics* 21(4):643-658.
- Crossman, A., D.A. Sullivan, and M. Benin. 2006. "The Family Environment and American Adolescents' Risk of Obesity as Young Adults." *Social Science & Medicine* 63(9):2255-2267.
- Dawson, D.A. 1991. "Family Structure and Children's Health and Well-Being: Data from the 1988 National Health Interview Survey on Child Health " *Journal of Marriage and Family* 53(3):573-584.
- Fairbrother, G., G. Kenney, K. Hanson, and L. Dubay. 2005. "How Do Stressful Family Environments Relate to Reported Access and Use of Health Care by Low-Income Children?" *Medical Care Research and Review* 62(2):205-230.
- Fletcher, J. and J. Sindelar. 2012. "The Effects of Family Stressors on Substance Use Initiation in Adolescence." *Review of Economics of the Household* 10(1):99-114.
- Fletcher, J.M. 2009. "Childhood Mistreatment and Adolescent and Young Adult Depression." *Social Science & Medicine* 68(5):799-806.
- Fletcher, J.M., J.C. Green, and M.J. Neidell. 2010. "Long term effects of childhood asthma on adult health." *Journal of Health Economics* 29(3):377-387.
- Francesconi, M., S.P. Jenkins, and T. Siedler. 2010. "The effect of lone motherhood on the smoking behavior of young adults." *Health Economics* 19(11):1377-1384.
- Gilman, S.E., R. Rende, J. Boergers, D.B. Abrams, S.L. Buka, M.A. Clark, S.M. Colby, B. Hitsman, A.N. Kazura, L.P. Lipsitt, E.E. Lloyd-Richardson, M.L. Rogers, C.A. Stanton, L.R. Stroud, and R.S. Niaura. 2009. "Parental Smoking and Adolescent Smoking Initiation: An Intergenerational Perspective on Tobacco Control." *Pediatrics* 123(2):e274-e281.

- Ginther, D. and R. Pollak. 2004. "Family structure and children's educational outcomes: Blended families, stylized facts, and descriptive regressions." *Demography* 41(4):671-696.
- Göhlmann, S., C.M. Schmidt, and H. Tauchmann. 2010. "Smoking initiation in Germany: the role of intergenerational transmission." *Health Economics* 19(2):227-242.
- Goodman, E. and R.C. Whitaker. 2002. "A Prospective Study of the Role of Depression in the Development and Persistence of Adolescent Obesity." *Pediatrics* 110(3):497-504.
- Griesbach, D., A. Amos, and C. Currie. 2003. "Adolescent Smoking and Family Structure in Europe." *Social Science & Medicine* 56(1):41-52.
- Grossman, M. 1972. "On the Concept of Health Capital and the Demand for Health." *Journal of Political Economy* 80(2):223.
- Harknett, K. 2009. "Why are Children with Married Parents Healthier? The Case of Pediatric Asthma." *Population Research and Policy Review* 28(3):347-365.
- Haveman, R. and B. Wolfe. 1994. *Succeeding Generations: On the Effects of Investments in Children*. New York: Russell Sage Foundation.
- Heard, H., B. Gorman, and C. Kapinus. 2008. "Family Structure and Self-Rated Health in Adolescence and Young Adulthood." *Population Research and Policy Review* 27(6):773-797.
- Heckman, J.J. 2007. "The Economics, Technology, and Neuroscience of Human Capability Formation." *Proceedings of the National Academy of Sciences* 104(33):13250-13255.
- Hill, M.S., W.J. Yeung, and G.J. Duncan. 2001. "Childhood Family Structure and Young Adult Behaviors." *Journal of Population Economics* 14(2):271-299.
- Hoffmann, J.P. and R.A. Johnson. 1998. "A National Portrait of Family Structure and Adolescent Drug Use." *Journal of Marriage and Family* 60(3):633-645.
- Jacobson, K.C. and D.C. Rowe. 1999. "Genetic and Environmental Influences on the Relationships between Family Connectedness, School Connectedness, and Adolescent Depressed Mood: Sex Differences." *Developmental Psychology* 35(4):926-939.
- Johnson-Down, L., J. O'Loughlin, K.G. Koski, and K. Gray-Donald. 1997. "High Prevalence of Obesity in Low Income and Multiethnic Schoolchildren: A Diet and Physical Activity Assessment." *Journal of Nutrition* 127(12):2310-2315.
- Kreider, R. 2008. "Living Arrangements of Children: 2004." *Current population reports* 70.
- Krein, S.F. and A.H. Beller. 1988. "Educational Attainment of Children From Single-Parent Families: Differences by Exposure, Gender, and Race." *Demography* 25(2):221-234.
- Kuczmarski, R.J., C.L. Ogden, S.S. Guo, L.M. Grummer-Strawn, K.M. Flegal, Z. Mei, R. Wei, L.R. Curtin, A.F. Roche, and C.L. Johnson. 2002. "2000 CDC Growth Charts for the United States: methods and development." *Vital Health Stat* 11(246):1-190.
- Lammers, C., M. Ireland, M. Resnick, and R. Blum. 2000. "Influences on Adolescents' Decision to Postpone Onset of Sexual Intercourse: A Survival Analysis of Virginity Among Youths Aged 13 to 18 years." *Journal of Adolescent Health* 26(1):42-48.
- Lang, K. and J.L. Zagorsky. 2001. "Does Growing up with a Parent Absent Really Hurt?" *The Journal of Human Resources* 36(2):253-273.
- Langton, C.E. and L.M. Berger. 2011. "Family Structure and Adolescent Physical Health, Behavior, and Emotional Well-Being." *Social Service Review* 85(3):323-357.
- Leininger, L.J. and K.M. Ziol-Guest. 2008. "Reexamining the Effects of Family Structure on Children's Access to Care: The Single-Father Family." *Health Services Research* 43(1p1):117-133.

- Mauldon, J. 1990. "The Effect of Marital Disruption on Children's Health." *Demography* 27(3):431-446.
- Montgomery, L.E., J.L. Kiely, and G. Pappas. 1996. "The Effects of Poverty, Race, and Family Structure on US Children's Health: Data from the NHIS, 1978 through 1980 and 1989 through 1991." *American Journal of Public Health* 86(10):1401-1405.
- Norton, A.J. and P.C. Glick. 1986. "One Parent Families: A Social and Economic Profile." *Family Relations* 35(1):9-17.
- Ogden, C.L. and M.D. Carroll. 2010. "Prevalence of Overweight, Obesity, and Extreme Obesity among Adults: United States, Trends 1960-1962 through 2007-2008.", edited by C.f.D.C.a. Prevention.
- Polsky, D., J.A. Doshi, J. Escarce, W. Manning, S.M. Paddock, L. Cen, and J. Rogowski. 2009. "The Health Effects of Medicare for the Near-Elderly Uninsured." *Health Services Research* 44(3):926-945.
- Radloff, L.S. 1977. "The CES-D Scale: A Self-Report Depression Scale for Research in the General Population." *Applied Psychological Measurement* 1(3):385-401.
- Roberts, R.E., P.M. Lewinsohn, and J.R. Seeley. 1991. "Screening for Adolescent Depression: A Comparison of Depression Scales." *J Am Acad Child Adolesc Psychiatry* 30(1):58-66.
- Ruhm, C.J. 2004. "Parental Employment and Child Cognitive Development." *Journal of Human Resources* XXXIX(1):155-192.
- Spruijt, E. and M. de Goede. 1997. "Transitions in Family Structure and Adolescent Well-being." *Adolescence* 32(128):897-911.
- Stewart, S.D. and C.L. Menning. 2009. "Family Structure, Nonresident Father Involvement, and Adolescent Eating Patterns." *Journal of Adolescent Health* 45(2):193-201.
- U.S. Census Bureau 1997. "Children With Single Parents- How They Fare." in *Census Briefs*: U.S. Census Bureau, Washington, D.C.
- Wolfe, B., R. Haveman, D. Ginther, and C.B. An. 1996. "The "Window Problem" in Studies of Children's Attainments: A Methodological Exploration." *Journal of the American Statistical Association* 91(435):970-982.

3.7. Tables and Figures

Table 3.1. Family structure through childhood by age at Wave I interview (1995)

	Girls					Boys				
	All girls (15-18)	15	16	17	18	All boys (15-18)	15	16	17	18
Panel A: Employing marital history and biological father questionnaire										
Father never present	0.105	0.105	0.094	0.124	0.095	0.088	0.077	0.099	0.092	0.082
Father left between ages 0 and 5	0.099	0.114	0.097	0.097	0.085	0.101	0.121	0.110	0.084	0.089
Father left between ages 6 and 10	0.061	0.086	0.052	0.047	0.056	0.048	0.058	0.057	0.034	0.042
Father left between ages 11 and 15	0.055	0.052	0.055	0.056	0.058	0.050	0.036	0.059	0.062	0.044
Father left between ages 16 and 18	0.012		0.006	0.018	0.026	0.016		0.009	0.026	0.033
Father ever left up to age 15	0.320	0.357	0.298	0.324	0.294	0.287	0.292	0.325	0.272	0.257
Other male entered between ages 0 and 5	0.068	0.072	0.064	0.065	0.069	0.070	0.086	0.069	0.060	0.063
Other male entered between ages 6 and 10	0.081	0.105	0.062	0.079	0.078	0.075	0.084	0.081	0.064	0.068
Other male entered between ages 11 and 15	0.068	0.062	0.070	0.074	0.066	0.061	0.060	0.077	0.051	0.056
Other male entered between ages 16 and 18	0.016		0.009	0.037	0.024	0.017		0.012	0.019	0.038
More than one other male present up to age 15	0.036	0.036	0.034	0.050	0.050	0.035	0.043	0.034	0.033	0.045
At least one other male entered up to age 15	0.184	0.204	0.174	0.203	0.193	0.173	0.189	0.203	0.164	0.185
Observations	4315	1032	1133	1158	992	4156	910	1128	1138	980
Panel B: Employing marital history only										
Father never present	0.138	0.139	0.116	0.146	0.156	0.126	0.112	0.130	0.117	0.148
Father left between ages 0 and 5	0.111	0.140	0.099	0.109	0.089	0.108	0.130	0.116	0.100	0.084
Father left between ages 6 and 10	0.062	0.066	0.073	0.053	0.054	0.054	0.052	0.062	0.049	0.053
Father left between ages 11 and 15	0.044	0.048	0.036	0.042	0.050	0.046	0.040	0.054	0.048	0.044
Father left between ages 16 and 18	0.010		0.011	0.013	0.019	0.010	0.000	0.005	0.013	0.024
Father ever left up to age 15	0.355	0.393	0.324	0.35	0.349	0.334	0.334	0.362	0.314	0.329
Other male entered between ages 0 and 5	0.095	0.110	0.088	0.100	0.080	0.102	0.125	0.103	0.088	0.089
Other male entered between ages 6 and 10	0.106	0.130	0.095	0.104	0.090	0.092	0.096	0.090	0.083	0.098
Other male entered between ages 11 and 15	0.080	0.075	0.087	0.085	0.073	0.076	0.074	0.092	0.065	0.071
Other male entered between ages 16 and 18	0.020		0.012	0.040	0.031	0.022	0.000	0.016	0.026	0.052
More than one other male present up to age 15	0.053	0.055	0.053	0.051	0.051	0.053	0.056	0.062	0.036	0.057
At least one other male entered up to age 15	0.238	0.265	0.231	0.243	0.206	0.223	0.239	0.239	0.206	0.207
Observations	4058	976	1075	1085	922	3892	857	1060	1065	910

Note: Other male refers to a man with whom the mother shared a marriage or marriage-like relationship after the child's biological left.

Table 3.2. Health and demographic characteristics of Add Health respondents in adolescence (Wave I): Girls

	All girls		Father ever left (32 % of girls)		Father never left (68 % of girls)	
Health measures						
Overweight or obese status	0.234	(0.423)	0.247	(0.431)	0.228	(0.420)
Very good or excellent health ***	0.656	(0.475)	0.582	(0.493)	0.691	(0.462)
Self reported health (1=excellent health) ***	2.181	(0.899)	2.292	(0.952)	2.128	(0.867)
Meets depression criteria ***	0.103	(0.304)	0.151	(0.358)	0.081	(0.272)
Ever regularly smoked cigarettes by 1995 ***	0.235	(0.424)	0.285	(0.451)	0.211	(0.408)
Family structure measures						
Father never present	0.107	(0.309)	0.332	(0.471)		
Father left between ages 0 and 5	0.099	(0.298)	0.306	(0.461)		
Father left between ages 6 and 10	0.061	(0.239)	0.188	(0.391)		
Father left between ages 11 and 15	0.056	(0.230)	0.174	(0.379)		
Other male entered between ages 0 and 5	0.067	(0.250)	0.209	(0.406)		
Other male entered between ages 6 and 10	0.082	(0.275)	0.255	(0.436)		
Other male entered between ages 11 and 15	0.068	(0.251)	0.210	(0.407)		
More than one other male present up to age 15	0.036	(0.187)	0.113	(0.316)		
At least one other male entered by age 15	0.185	(0.388)	0.575	(0.495)		
Demographic measures and maternal characteristics						
Age	16.411	(1.114)	16.353	(1.119)	16.439	(1.111)
Asian ***	0.031	(0.173)	0.013	(0.114)	0.039	(0.194)
Black ***	0.149	(0.356)	0.257	(0.437)	0.098	(0.297)
Hispanic	0.110	(0.313)	0.093	(0.291)	0.118	(0.322)
Other race	0.093	(0.291)	0.093	(0.291)	0.093	(0.290)
Birth order ***	1.843	(1.179)	1.575	(0.923)	1.970	(1.263)
Mother U.S. born ***	0.881	(0.324)	0.917	(0.276)	0.864	(0.343)
Income in 1995 (thousands) ***	47.366	(41.079)	33.010	(29.731)	54.190	(43.878)
Missing income in 1995	0.110	(0.313)	0.100	(0.300)	0.114	(0.318)
Mother's highest grade: high school	0.432	(0.495)	0.437	(0.496)	0.429	(0.495)
Mother's highest grade: some college ***	0.185	(0.388)	0.218	(0.413)	0.169	(0.375)
Mother's highest grade: college or beyond ***	0.220	(0.415)	0.168	(0.374)	0.245	(0.430)
Mother worked outside the home in past year	0.793	(0.406)	0.806	(0.395)	0.786	(0.410)
Mother ever employed full-time in past year ***	0.607	(0.489)	0.663	(0.473)	0.580	(0.494)
Mother age at birth ***	25.224	(5.126)	23.423	(4.934)	26.080	(4.993)
Number of siblings in household in 1995 **	1.422	(1.171)	1.294	(1.195)	1.483	(1.154)
Mother religious attendance: at least weekly ***	0.382	(0.486)	0.270	(0.444)	0.435	(0.496)
Mother religious attendance: between weekly and monthly	0.180	(0.385)	0.198	(0.398)	0.172	(0.378)
Mother religious attendance: less than monthly ***	0.229	(0.420)	0.276	(0.447)	0.206	(0.405)
Mother smokes ***	0.280	(0.449)	0.411	(0.492)	0.218	(0.413)
Mother reports very good or better health ***	0.563	(0.496)	0.496	(0.500)	0.595	(0.491)
Observations	4091		1337		2754	

Standard deviations in parentheses. Girls under the age of 15 at the time of the Wave I interview are excluded

Father presence refers to the child's biological father up to the year the child turned 15.

Proportion tests for categorical variables and t-tests for continuous variables were used to calculate significant differences between the sample of girls whose biological fathers ever left and those whose fathers never left.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 3.3. Health and demographic characteristics of Add Health respondents in adolescence (Wave I): Boys

	All boys		Father ever left (29 % of boys)		Father never left (71 % of boys)	
Health measures						
Overweight or obese status	0.275	(0.447)	0.258	(0.438)	0.283	(0.450)
Very good or excellent health **	0.728	(0.445)	0.689	(0.463)	0.744	(0.437)
Self reported health (1=excellent health) ***	1.987	(0.873)	2.093	(0.896)	1.944	(0.860)
Meets depression criteria ***	0.075	(0.263)	0.102	(0.303)	0.063	(0.244)
Ever regularly smoked cigarettes by 1995 ***	0.240	(0.427)	0.284	(0.451)	0.222	(0.416)
Family structure measures						
Father never present	0.087	(0.282)	0.303	(0.460)		
Father left between ages 0 and 5	0.102	(0.303)	0.356	(0.479)		
Father left between ages 6 and 10	0.048	(0.214)	0.168	(.374)		
Father left between ages 11 and 15	0.050	(0.218)	0.174	(0.379)		
Other male entered between ages 0 and 5	0.072	(0.258)	0.248	(0.432)		
Other male entered between ages 6 and 10	0.076	(0.266)	0.265	(0.441)		
Other male entered between ages 11 and 15	0.063	(0.244)	0.220	(0.414)		
More than one other male present up to age 15	0.036	(0.187)	0.126	(0.332)		
At least one Other male entered up by age 15	0.177	(0.382)	0.615	(0.487)		
Demographic measures and maternal characteristics						
Age	16.452	(1.122)	16.391	(1.103)	16.477	(1.129)
Asian ***	0.033	(0.178)	0.018	(0.134)	0.039	(0.193)
Black ***	0.120	(0.325)	0.242	(0.428)	0.071	(0.256)
Hispanic	0.110	(0.313)	0.113	(0.316)	0.110	(0.312)
Other race	0.093	(0.290)	0.092	(0.289)	0.093	(0.291)
Birth order ***	1.796	(1.072)	1.586	(0.897)	1.881	(1.124)
Mother U.S. born	0.882	(0.323)	0.916	(0.277)	0.868	(0.338)
Income in 1995 (thousands) ***	48.593	(40.420)	35.771	(35.312)	53.783	(41.199)
Missing income in 1995 ***	0.110	(0.313)	0.083	(0.276)	0.121	(0.326)
Mother's highest grade: high school	0.438	(0.496)	0.454	(0.498)	0.432	(0.495)
Mother's highest grade: some college	0.183	(0.387)	0.198	(0.398)	0.178	(0.382)
Mother's highest grade: college or beyond ***	0.237	(0.426)	0.185	(0.388)	0.259	(0.438)
Mother worked outside the home in past year	0.802	(0.399)	0.811	(0.391)	0.798	(0.402)
Mother ever employed full-time in past year ***	0.597	(0.491)	0.671	(0.470)	0.567	(0.496)
Mother age at birth ***	25.063	(5.104)	23.431	(5.237)	25.723	(4.899)
Number of siblings in household in 1995 *	1.425	(1.144)	1.340	(1.280)	1.459	(1.083)
Mother religious attendance: at least weekly ***	0.391	(0.488)	0.295	(0.456)	0.430	(0.495)
Mother religious attendance: between weekly and monthly	0.178	(0.383)	0.178	(0.383)	0.178	(0.383)
Mother religious attendance: less than monthly ***	0.248	(0.432)	0.307	(0.461)	0.224	(0.417)
Mother smokes ***	0.272	(0.445)	0.429	(0.495)	0.208	(0.406)
Mother reports very good or better health ***	0.592	(0.492)	0.506	(0.500)	0.626	(0.484)
Observations	3920		1182		2738	

Standard deviations in parentheses. Boys under the age of 15 at the time of the Wave I interview are excluded

Father presence refers to the child's biological father up to the year the child turned 15.

Proportion tests for categorical variables and t-tests for continuous variables were used to calculate significant differences between the sample of boys whose biological fathers ever left and those whose fathers never left.

* p<0.10, ** p<0.05, *** p<0.01.

Table 3.4. Baseline logit models of adolescent smoking, weight status, self-reported health status, and depression at Wave I interview

	Girls				Boys					
	Smoking (1)	Smoking (2)	Overweight/ Obese status (3)	Very good or excellent health (4)	Depression (5)	Smoking (6)	Smoking (7)	Overweight/ Obese status (8)	Very good or excellent health (9)	Depression (10)
Panel A: Biological father absence										
Father ever absent up to age 15	0.083*** (0.022)	0.079*** (0.021)	-0.012 (0.025)	-0.095*** (0.022)	0.064*** (0.015)	0.063*** (0.023)	0.051*** (0.023)	-0.051** (0.021)	-0.044* (0.023)	0.024* (0.013)
Mother smoked in 1995		0.051** (0.022)					0.086*** (0.024)			
Panel B: Biological father absence and Other male entrance										
Father ever absent up to age 15	0.067* (0.035)	0.066* (0.035)	0.003 (0.032)	-0.091*** (0.033)	0.065*** (0.018)	0.056* (0.030)	0.050* (0.030)	-0.058** (0.029)	-0.035 (0.033)	0.025 (0.017)
Other male ever present up to age 15	0.024 (0.034)	0.020 (0.034)	-0.026 (0.029)	-0.006 (0.039)	-0.002 (0.022)	0.010 (0.036)	0.001 (0.036)	0.014 (0.038)	-0.013 (0.041)	-0.001 (0.014)
Mother smoked in 1995		0.050** (0.022)					0.086*** (0.024)			
Panel C: Biological father departure, disaggregated										
Father never present	0.139*** (0.035)	0.135*** (0.034)	-0.020 (0.030)	-0.124*** (0.032)	0.057** (0.026)	0.043 (0.036)	0.031 (0.035)	-0.020 (0.033)	-0.077** (0.037)	0.048** (0.023)
Father left between ages 0 and 5	0.077** (0.033)	0.072** (0.033)	-0.007 (0.027)	-0.115*** (0.033)	0.089*** (0.028)	0.074** (0.037)	0.059 (0.037)	-0.039 (0.029)	0.024 (0.032)	0.014 (0.019)
Father left between ages 6 and 10	0.022 (0.040)	0.017 (0.039)	0.001 (0.054)	-0.044 (0.051)	0.073* (0.038)	0.061 (0.053)	0.051 (0.052)	-0.067* (0.038)	-0.076 (0.048)	0.009 (0.026)
Father left between ages 11 and 15	0.105* (0.056)	0.104* (0.056)	-0.018 (0.045)	-0.093** (0.044)	0.064* (0.037)	0.084* (0.047)	0.074* (0.045)	-0.080** (0.036)	-0.074 (0.051)	0.045 (0.031)
Mother smoked in 1995		0.052** (0.022)					0.084*** (0.024)			

Table 3.4. (continued) Baseline logit models of adolescent smoking, weight status, self-reported health status, and depression at Wave I interview

	Girls				Boys				
	Smoking		Overweight/ Obese status		Smoking		Overweight/ Obese status		Depression
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	
Panel D: Biological father departure and Other male entrance, disaggregated									
Father never present	0.083* (0.045)	0.081* (0.045)	0.008 (0.040)	-0.121** (0.048)	0.050* (0.030)	-0.007 (0.044)	-0.013 (0.044)	-0.034 (0.038)	-0.021 (0.050)
Father left between ages 0 and 5	0.039 (0.047)	0.036 (0.046)	0.022 (0.034)	-0.117*** (0.042)	0.075** (0.035)	0.033 (0.040)	0.025 (0.040)	-0.052 (0.038)	0.066 (0.040)
Father left between ages 6 and 10	0.018 (0.047)	0.014 (0.047)	0.015 (0.058)	-0.049 (0.053)	0.061* (0.032)	0.053 (0.049)	0.047 (0.049)	-0.059 (0.041)	-0.044 (0.049)
Father left between ages 11 and 15	0.101* (0.056)	0.100* (0.056)	-0.015 (0.047)	-0.093** (0.045)	0.062 (0.038)	0.080* (0.048)	0.072 (0.047)	-0.073* (0.040)	-0.058 (0.053)
Other male entered between ages 0 and 5	0.118* (0.060)	0.119** (0.060)	-0.050 (0.039)	-0.008 (0.055)	0.022 (0.035)	0.093 (0.071)	0.084 (0.071)	0.056 (0.063)	-0.092 (0.065)
Other male entered between ages 6 and 10	-0.007 (0.048)	-0.010 (0.048)	-0.041 (0.040)	0.021 (0.050)	0.023 (0.035)	0.012 (0.044)	0.008 (0.044)	0.030 (0.049)	-0.033 (0.063)
Other male entered between ages 11 and 15	0.023 (0.045)	0.020 (0.045)	-0.014 (0.041)	-0.000 (0.053)	0.010 (0.028)	0.009 (0.044)	0.001 (0.044)	-0.056 (0.045)	-0.074 (0.062)
More than one other male present up to age 15	-0.018 (0.049)	-0.021 (0.050)	0.040 (0.071)	-0.029 (0.077)	-0.045** (0.018)	0.005 (0.068)	0.011 (0.069)	-0.025 (0.067)	0.053 (0.065)
Mother smoked in 1995		0.052** (0.022)					0.083*** (0.024)		
Number of adolescents	4104	4093	4125	4123	4124	3922	3920	3954	3954

Note: Marginal effects from logit models reported. Estimates adjusted for covariates in Tables 2 and 3.

Sample sizes are for diasaggregated specifications (shown in Panels C and D). Sample sizes are slightly higher in Panels A and B.

* p<0.10, ** p<0.05, *** p<0.01. Standard errors in parentheses.

Smoking is defined as ever smoking at least one cigarette per day for 30 days.

Adolescents who are aged 15 to 18 at the time of the Add Health Wave I interview (1995) are included in the sample.

Table 3.5. Discrete-time hazard models of smoking cessation and

	Girls				Boys			
	Wave I Smoker		Wave I Non-smoker		Wave I Smoker		Wave I Non-smoker	
	Quitting after Wave I (1)	Wave I (2)	Starting after Wave I (3)	Wave I (4)	Quitting after Wave I (5)	Wave I (6)	Starting after Wave I (7)	Wave I (8)
Panel A: Biological father absence								
Father ever absent up to age 15	-0.018 (0.020)	-0.009 (0.020)	0.024* (0.012)	0.023* (0.013)	-0.006 (0.022)	-0.005 (0.022)	0.018 (0.016)	0.015 (0.016)
Mother smoked in 1995		-0.066*** (0.018)		0.009 (0.012)		-0.006 (0.020)		0.031* (0.017)
Panel B: Biological father absence and Other male entrance								
Father ever absent up to age 15	-0.033 (0.025)	-0.026 (0.026)	0.021 (0.016)	0.021 (0.016)	0.014 (0.029)	0.014 (0.029)	0.000 (0.024)	-0.001 (0.024)
Other male ever present up to age 15	0.026 (0.028)	0.028 (0.029)	0.004 (0.020)	0.004 (0.020)	-0.030 (0.025)	-0.030 (0.025)	0.028 (0.029)	0.026 (0.029)
Mother smoked in 1995		-0.066*** (0.018)		0.009 (0.012)		-0.005 (0.020)		0.030* (0.017)
Panel C: Biological father departure, disaggregated								
Father never present	-0.011 (0.026)	-0.005 (0.026)	0.020 (0.018)	0.020 (0.018)	-0.054*** (0.020)	-0.053** (0.020)	0.018 (0.027)	0.017 (0.027)
Father left between ages 0 and 5	-0.035 (0.024)	-0.023 (0.025)	0.023 (0.020)	0.022 (0.020)	0.008 (0.027)	0.010 (0.028)	0.023 (0.025)	0.018 (0.025)
Father left between ages 6 and 10	0.058 (0.056)	0.064 (0.056)	0.034 (0.028)	0.033 (0.028)	-0.003 (0.042)	-0.002 (0.042)	0.038 (0.034)	0.035 (0.034)
Father left between ages 11 and 15	-0.047** (0.024)	-0.044* (0.024)	0.022 (0.025)	0.022 (0.025)	0.023 (0.034)	0.024 (0.034)	-0.023 (0.030)	-0.025 (0.031)
Mother smoked in 1995		-0.065*** (0.018)		0.007 (0.013)		-0.012 (0.020)		0.031* (0.017)

Table 3.5.(continued): Discrete-time hazard models of smoking cessation and initiation

	Girls				Boys			
	Wave I Smoker		Wave I Non-smoker		Wave I Smoker		Wave I Non-smoker	
	Quitting after Wave I (1)	(2)	Starting after Wave I (3)	(4)	Quitting after Wave I (5)	(6)	Starting after Wave I (7)	(8)
Panel D: Biological father departure and Other male entrance, disaggregated								
Father never present	-0.038 (0.031)	-0.034 (0.031)	0.018 (0.023)	0.017 (0.023)	-0.026 (0.038)	-0.026 (0.038)	-0.012 (0.032)	-0.011 (0.032)
Father left between ages 0 and 5	-0.049** (0.024)	-0.039 (0.025)	0.025 (0.026)	0.024 (0.026)	0.035 (0.051)	0.036 (0.051)	-0.004 (0.035)	-0.007 (0.034)
Father left between ages 6 and 10	0.051 (0.060)	0.057 (0.058)	0.050* (0.029)	0.049* (0.029)	0.018 (0.053)	0.018 (0.053)	0.028 (0.037)	0.026 (0.037)
Father left between ages 11 and 15	-0.050** (0.023)	-0.047* (0.025)	0.029 (0.025)	0.029 (0.025)	0.027 (0.035)	0.028 (0.035)	-0.025 (0.031)	-0.028 (0.031)
Other male entered between ages 0 and 5	0.074 (0.056)	0.070 (0.056)	0.023 (0.036)	0.024 (0.036)	-0.030 (0.043)	-0.030 (0.044)	0.024 (0.051)	0.022 (0.051)
Other male entered between ages 6 and 10	0.000 (0.043)	-0.001 (0.042)	0.016 (0.027)	0.016 (0.027)	-0.010 (0.054)	-0.009 (0.055)	0.029 (0.042)	0.027 (0.041)
Other male entered between ages 11 and 15	0.021 (0.040)	0.020 (0.041)	-0.059*** (0.018)	-0.059*** (0.018)	-0.029 (0.035)	-0.028 (0.036)	0.012 (0.034)	0.011 (0.034)
More than one other male present up to age 15	-0.043 (0.038)	-0.033 (0.040)	0.028 (0.045)	0.026 (0.045)	-0.041 (0.056)	-0.043 (0.055)	0.042 (0.067)	0.042 (0.067)
Mother smoked in 1995		-0.064*** (0.018)		0.007 (0.012)		-0.012 (0.020)		0.029* (0.017)
Number of observations (person-time)	2268	2260	7077	7059	2361	2359	6119	6118

Note: Marginal effects from logit models reported. Estimates adjusted for covariates in Tables 2 and 3, including wave indicator variables. trend. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses.

Sample sizes are for disaggregated specifications (shown in Panels C and D). Sample sizes are slightly higher in Panels A and B.

Smoking is defined as ever smoking at least one cigarette per day for 30 days. Quitting is defined as smoking no cigarettes for the past 30 days. Adolescents who are age 15 or above at the time of the Add Health Wave I interview (1995) are included in the sample.

Table 3.6. Discrete-time hazard models for physical health transitions in girls

	Adolescents reporting good or lower health: hazard of excellent or very good health status after Wave I	Adolescents reporting excellent or very good health: hazard of good or lower health status after Wave I	Normal-weight adolescents: hazard of overweight or obese status after Wave I	Overweight or obese adolescents: hazard of normal weight status after Wave I	Adolescents not classified as depressed: hazard of depression diagnosis after Wave I	Adolescents classified as depressed: hazard of no depression diagnosis after Wave I
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Biological father absence						
Father ever absent up to age 15	0.014 (0.032)	0.017 (0.017)	-0.024 (0.016)	0.005 (0.021)	0.017 (0.028)	-0.056 (0.089)
Panel B: Biological father absence and Other male entrance						
Father ever absent up to age 15	0.035 (0.038)	-0.004 (0.024)	-0.022 (0.023)	0.011 (0.025)	0.027 (0.037)	-0.092 (0.092)
Other male ever present up to age 15	-0.035 (0.045)	0.037 (0.031)	-0.003 (0.029)	-0.010 (0.016)	-0.016 (0.036)	0.062 (0.102)
Panel C: Biological father absence, disaggregated						
Father never present	-0.048 (0.042)	-0.014 (0.023)	0.002 (0.027)	0.015 (0.026)	-0.004 (0.037)	-0.153 (0.119)
Father left between ages 0 and 5	0.046 (0.042)	0.031 (0.025)	-0.022 (0.020)	-0.010 (0.020)	0.042 (0.046)	0.005 (0.124)
Father left between ages 6 and 10	0.009 (0.063)	0.039 (0.033)	-0.061*** (0.023)	0.052 (0.063)	-0.041 (0.036)	-0.051 (0.158)
Father left between ages 11 and 15	0.124** (0.054)	0.012 (0.039)	-0.030 (0.033)	-0.031 (0.020)	0.063 (0.056)	-0.007 (0.154)

Table 3.6. (continued): Discrete-time hazard models for physical health transitions in girls

	Adolescents reporting good or lower health: hazard of excellent or very good health status after Wave I	Adolescents reporting excellent or very good health: hazard of good or lower health status after Wave I	Normal-weight adolescents: hazard of overweight or obese status after Wave I	Overweight or obese adolescents: hazard of normal weight status after Wave I	Adolescents not classified as depressed: hazard of depression diagnosis after Wave I	Adolescents classified as depressed: hazard of no depression diagnosis after Wave I
	(1)	(2)	(3)	(4)	(5)	(6)
Panel D: Biological father absence and Other male entrance, disaggregated						
Father never present	-0.044 (0.050)	-0.045 (0.028)	-0.005 (0.031)	0.030 (0.037)	-0.007 (0.040)	-0.190 (0.132)
Father left between ages 0 and 5	0.067 (0.055)	-0.008 (0.029)	-0.025 (0.029)	0.001 (0.029)	0.048 (0.062)	-0.015 (0.151)
Father left between ages 6 and 10	0.032 (0.061)	0.018 (0.033)	-0.056** (0.025)	0.060 (0.066)	-0.027 (0.035)	-0.069 (0.168)
Father left between ages 11 and 15	0.128** (0.053)	0.002 (0.038)	-0.028 (0.034)	-0.031 (0.020)	0.071 (0.058)	-0.005 (0.144)
Other male entered between ages 0 and 5	-0.053 (0.067)	0.082 (0.056)	0.051 (0.047)	-0.033* (0.018)	-0.022 (0.042)	0.077 (0.134)
Other male entered between ages 6 and 10	-0.048 (0.057)	0.042 (0.043)	-0.011 (0.036)	-0.021 (0.019)	-0.001 (0.047)	0.032 (0.138)
Other male entered between ages 11 and 15	-0.027 (0.072)	0.061 (0.043)	-0.011 (0.038)	0.003 (0.032)	-0.061* (0.036)	-0.019 (0.189)
More than one other male present up to age 15	0.167 (0.105)	-0.074** (0.037)	-0.021 (0.054)	0.049 (0.059)	0.130 (0.095)	-0.040 (0.279)
Number of observations (person-time)	2413	5527	6389	2012	3735	463

Note: Marginal effects from logit models reported. Estimates adjusted for covariates in Tables 2 and 3 (except maternal smoking), including a wave indicator variables.

* p<0.10, ** p<0.05, *** p<0.01. Standard errors in parentheses.

Table 3.7. Discrete-time hazard models for physical and mental health transitions in boys

	Adolescents reporting good or lower health: hazard of excellent or very good health status after Wave I	Adolescents reporting excellent or very good health: hazard of good or lower health status after Wave I	Normal-weight adolescents: hazard of overweight or obese status after Wave I	Overweight or obese adolescents: hazard of normal weight status after Wave I	Adolescents not classified as depressed: hazard of depression diagnosis after Wave I	Adolescents classified as depressed: hazard of no depression diagnosis after Wave I
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Biological father absence						
Father ever absent up to age 15	-0.062* (0.032)	0.028 (0.017)	0.009 (0.019)	0.013 (0.015)	0.011 (0.016)	-0.093 (0.080)
Panel B: Biological father absence and Other male entrance						
Father ever absent up to age 15	-0.013 (0.048)	0.019 (0.022)	0.015 (0.023)	0.013 (0.015)	0.005 (0.026)	-0.096 (0.103)
Other male ever present up to age 15	-0.079 (0.058)	0.014 (0.025)	-0.010 (0.027)	0.000 (0.019)	0.009 (0.033)	0.005 (0.087)
Panel C: Biological father absence disaggregated						
Father never present	-0.046 (0.053)	0.031 (0.028)	0.012 (0.029)	-0.001 (0.019)	0.038 (0.037)	-0.029 (0.127)
Father left between ages 0 and 5	-0.130*** (0.050)	0.031 (0.023)	0.016 (0.025)	0.032 (0.024)	-0.016 (0.022)	-0.091 (0.143)
Father left between ages 6 and 10	-0.100 (0.067)	0.020 (0.032)	-0.011 (0.042)	0.006 (0.028)	0.047 (0.044)	-0.205 (0.158)
Father left between ages 11 and 15	0.072 (0.064)	0.032 (0.033)	0.005 (0.032)	-0.002 (0.035)	-0.015 (0.024)	-0.166 (0.170)

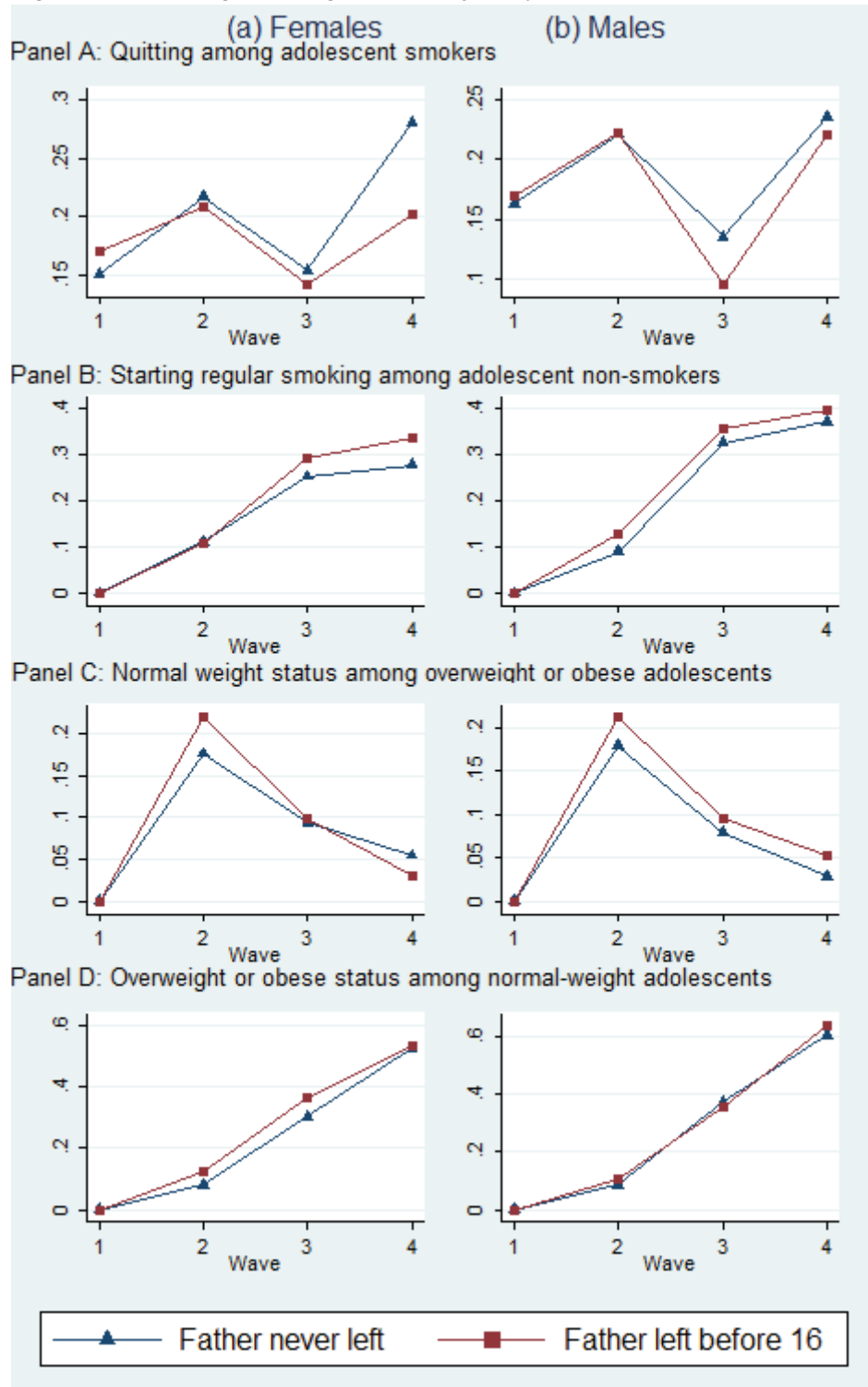
Table 3.7. (continued): Discrete-time hazard models for physical and mental health transitions in boys

	Adolescents reporting good or lower health: hazard of excellent or very good health status after Wave I	Adolescents reporting excellent or very good health: hazard of good or lower health status after Wave I	Normal-weight adolescents: hazard of overweight or obese status after Wave I	Overweight or obese adolescents: hazard of normal weight status after Wave I	Adolescents not classified as depressed: hazard of depression diagnosis after Wave I	Adolescents classified as depressed: hazard of no depression diagnosis after Wave I
	(1)	(2)	(3)	(4)	(5)	(6)
Panel D: Biological father absence and Other male entrance, disaggregated						
Father never present	-0.036 (0.085)	0.022 (0.035)	0.029 (0.034)	-0.006 (0.023)	0.032 (0.061)	0.086 (0.104)
Father left between ages 0 and 5	-0.106 (0.068)	0.016 (0.028)	0.031 (0.036)	0.030 (0.030)	-0.027 (0.025)	-0.012 (0.134)
Father left between ages 6 and 10	-0.066 (0.083)	-0.004 (0.031)	0.003 (0.045)	0.014 (0.035)	0.017 (0.041)	-0.217 (0.181)
Father left between ages 11 and 15	0.074 (0.063)	0.017 (0.033)	0.012 (0.033)	0.001 (0.036)	-0.022 (0.023)	-0.157 (0.165)
Other male entered between ages 0 and 5	0.033 (0.116)	-0.009 (0.033)	0.001 (0.046)	0.012 (0.036)	-0.020 (0.039)	-0.207 (0.274)
Other male entered between ages 6 and 10	-0.110 (0.070)	0.034 (0.034)	-0.014 (0.035)	-0.004 (0.019)	0.040 (0.052)	0.014 (0.161)
Other male entered between ages 11 and 15	-0.012 (0.082)	0.081* (0.045)	-0.037 (0.036)	-0.018 (0.022)	0.055 (0.066)	0.047 (0.095)
More than one other male present up to age 15	0.031 (0.131)	-0.050 (0.039)	-0.010 (0.061)	0.018 (0.058)	-0.028 (0.036)	-0.142 (0.300)
Number of observations (person-time)	1671	5796	5487	2205	3385	275

Note: Marginal effects from logit models reported. Estimates adjusted for covariates in Tables 2 and 3 (except maternal smoking), including a wave indicator variables.

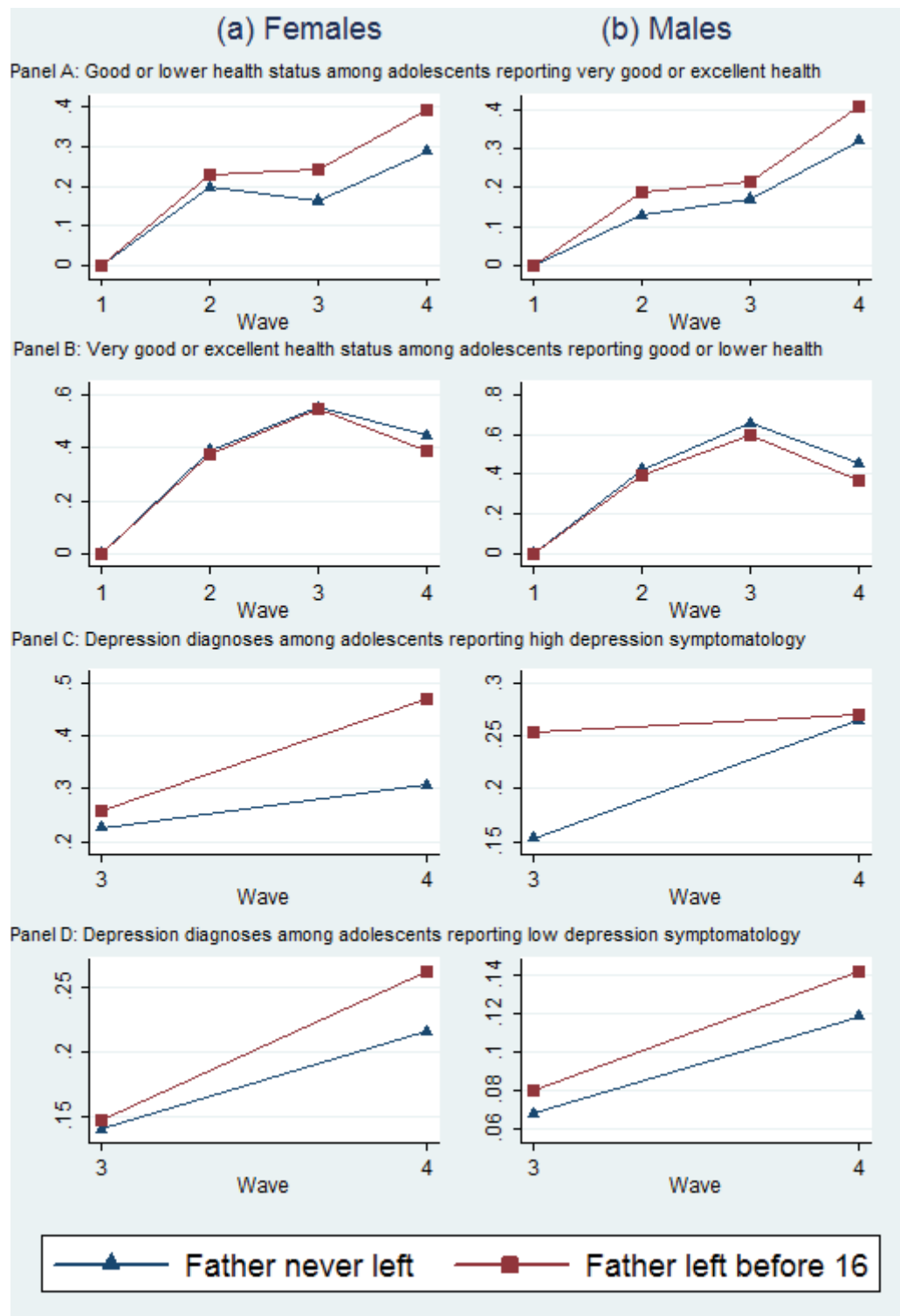
* p<0.10, ** p<0.05, *** p<0.01.

Figure 3.1. Smoking and weight status trajectory after adolescence



Note: Adolescents reported ever smoking and ever quitting (if applicable) during the Wave I interview. As a result, the mean quitting probability among adolescents who ever smoked is slightly above zero.

Figure 3.2. Self-reported health status and depression diagnosis trajectory after adolescence



Note: Depression diagnoses are only measured during Waves III and IV.

Chapter 4

The Relationship between Nonresident Father Involvement and Maternal Depression in Fragile Families

The beneficial effects of nonresident father involvement, including child support payment and visitation, on the welfare of children have been well-characterized. The indirect effects of this involvement on mothers, however, have remained relatively unexplored. This paper utilizes panel data from the Fragile Families and Child Wellbeing Survey to consider the role of nonresident father involvement on maternal depression, which can have adverse consequences on the developmental trajectory of their children. Linear probability models are estimated to clarify the effect of different forms of nonresident father involvement, as measured by any formal child support payment, any informal support, any in-kind support, and visitation, on maternal depression. The role of material hardship and paternal incarceration as potential pathways in this relationship is explored. I find that in-kind support is especially beneficial in reducing the risk for maternal depression. Visitation, on the other hand, can be harmful to some mothers, especially when done so infrequently and if unaccompanied by informal forms of material support. The results of this study underscore the need for the disaggregation of measures of nonresident father involvement, and the need for a better understanding of the spillover effects of father involvement on non-child outcomes.

4.1. Introduction

The prevalence of the “traditional” two-parent family in the United States has eroded significantly over the past several decades. In 2009, over 26% of all children under 21 lived with only one parent (Kreider and Ellis 2011). This change in family structure was largely

precipitated by an increase in the divorce rate over time up to the early 1980's, together with a precipitous rise in births out-of-wedlock (Hamilton, Martin, and Ventura 2011). Much evidence suggests that children who spend time in single-parent families have worse outcomes in educational, cognitive, and possibly health domains compared their counterparts living in intact, two-biological-parent families (e.g. Krein and Beller 1988; Stewart and Menning 2009; Strauss and Knight 1999; Bramlett and Blumberg 2007; Harknett 2009; Langton and Berger 2011). In addition to the well-studied harmful effects that single-parent status can have on children, negative consequences can extend to mothers as well. An important domain of maternal welfare that is sensitive to family structure is her mental health. It has been found that single mothers often have less favorable mental health profiles than married mothers (Burgos et al. 1995; Crosier, Butterworth, and Rodgers 2007). Single mothers also face a greater constellation of stressors that can contribute to poorer mental health outcomes (Avison, Ali, and Walters 2007). As single mothers generally have less financial resources than two-parent households, financial stressors can be more prevalent in these households, which can lead to increased rates of mental illness, including depression (Muntaner et al. 2004; Heflin and Iceland 2009)

In order to alleviate some of the financial consequences to mothers living without resident fathers, the child support system in the United States was designed to ensure that children with nonresident parents receive financial support from both the father and mother. This generally entails a cash payment by a non-custodial parent to the custodial parent, who is usually the mother. Child support payments allow custodial parents to better support their children through shelter, food, and clothing, among others. There is a large literature documenting that child support can enhance child outcomes independently of total household income (which includes child support) (Graham, Beller, and Hernandez 1994; Hernandez, Beller, and Graham

1995). For instance, the mother's receipt of child support has been shown to enhance the child's academic and cognitive performance, and reduce behavioral problems (reviewed by Amato and Gilbreth 1999). Though it is possible some of these findings are driven by selection or unobserved heterogeneity, studies that have addressed this bias (such as the use of variations in state-level child support enforcement as instrumental variables) have found similar effects, especially in regard to cognitive skills (e.g. Argys et al. 1998; Knox 1996).

The majority of research surrounding nonresident father involvement and child support focuses on the welfare of the children living with an absent father, who are the primary beneficiaries of the behavior. However, the role of nonresident father involvement on improving the welfare of the household as a whole, or the mothers themselves, has remained understudied. The health and economic welfare of the mother may be important channels by which nonresident fathers may improve the outcomes of their children. Two studies have focused on the relationship between nonresident father involvement and the availability and management of material resources, such as food and shelter. Garasky and Stewart (2007) use cross-sectional data from the National Survey of America's Families and find that families with nonresident fathers were less likely to suffer from food insecurity if they had received a higher amount of child support or more frequent visitation. Father-child contact (i.e., visitation) had a much greater influence on reducing food insecurity than monetary child support receipts. Nepomnyaschy and Garfinkel (2011) extend the work of Garasky and Stewart by using data from the Fragile Families and Child Wellbeing Study to study the effects of father involvement on household well-being by considering additional measures of material hardship and nonresident father involvement, and by incorporating panel data. Like Garasky and Stewart, they find that

visitation, compared to material forms of involvement, played the strongest role in decreasing measures of material hardship (including measures of food insecurity).

In this paper, I add to the existing literature by investigating how similar measures of nonresident father involvement, including formal child support, informal support, in-kind support, and visitation, affect maternal depression. Panel data from the Fragile Families and Child Wellbeing Study provide a useful sample to ascertain how variations in these involvement measures affect the mental health of mothers with young children. This analysis contributes to the wider literature on child support and human capital outcomes in several ways. First, it expands upon a potential mechanism by which nonresident father involvement can affect children and families: depression. Secondly, the most recent wave of the Fragile Families data is included, which spans up to when the child is nine years old. This offers the advantage of increased variability in father involvement over time to estimate its effects on maternal mental health. It also extends the analysis to older children, compared to most studies using the Fragile Families data that focus on children age five or under. Finally, I explore pathways that are related to both nonresident father involvement and maternal depression, including paternal incarceration and maternal material hardship.

4.1.1. Family-Related Consequences and Causes of Maternal Depression

The World Health Organization has characterized depression as one of the most severe health burdens worldwide. Indeed, it has been suggested that depression was the single leading cause of disability worldwide in 2004 as measured by years lost to disability (YLD) (World Health Organization 2008). Women tend to face a larger share of the depression burden than men. Depression among mothers in the United States is very common, especially among those

with very young children, where the rate of depression can reach 15% (McLennan, Kotelchuck, and Cho 2001). This number is higher among single mothers, where prevalence rates of chronic depression approaching 30% have been reported (Wang et al. 2011). Though the sheer prevalence of maternal depression emphasizes the need to understand its correlates, maternal depression can adversely affect the well-being of her children as well (Chang, Halpern, and Kaufman 2007; Kramer et al. 2009; Surkan, Kawachi, and Peterson 2008; Talge et al. 2007; Turney 2011) .

Though the intrinsic demographic and physical health characteristics of the mother are important in her risk for depression (DeKlyen et al. 2006; Breslau et al. 2005; Kessler and Zhao 1999), a growing literature implicates household- and relationship- characteristics as important correlates as well. Studies comparing depression prevalence of single and partnered mothers find that single mothers have increased depression levels compared to their partnered peers (e.g., Burgos et al. 1995; Crosier, Butterworth, and Rodgers 2007; DeKlyen et al. 2006; Kessler and Zhao 1999) . Increasingly, the dynamics of family structure (including marital entry and exit) are being implicated as an important correlate to maternal depression. Studies using longitudinal data from the National Survey of Families and Households (summarized by Wood, Goesling, and Avellar 2007) find that transitions into marriage leads to lower rates of depression, and marital dissolution contributes to higher rates of depression. Further, studies using these data found that stably married individuals were less likely to experience subsequent depression compared to unmarried individuals.

Evidence from the FFCWS supports the characterization that relationship transitions away from cohabitation or marriage can raise the mother's risk for experiencing subsequent psychological distress. Studies using the panel nature of FFCWS to examine relationship

transitions have shown that departure from a cohabiting or marriage relationship can increase rates of depression and parenting stress compared to consistently married individuals, the effects of which can vary according to the mother's socioeconomic status (Cooper et al. 2009; Meadows, McLanahan, and Brooks-Gunn 2008; Osborne, Berger, and Magnuson 2012). Additional recent evidence suggests that mothers in the FFCWS who have children with more than one father (i.e., multi-partner fertility) may have an increased risk of depression compared to mothers who have children with one man (Turney and Carlson 2011).

4.1.2. Nonresident Father Involvement and Maternal Depression

While direct measures of material nonresident father involvement have not been directly linked to maternal mental health, studies using small samples of young mothers have implicated that increased levels of father social involvement (or the perceptions thereof) is associated with lower levels of maternal psychological distress (Elsenbruch et al. 2007; Jackson 1999; Kalil, Ziol-Guest, and Coley 2005; Kiernan and Pickett 2006; Malik et al. 2007; Smith and Howard 2008). Thus, to the extent that any form of nonresident father involvement implies a social commitment to the mother and child, the results of these studies suggest that nonresident father involvement should yield improvements in maternal depression profiles. However, these studies have several limitations, in addition to their relatively small sample sizes. First, most of these studies (with the exception of Jackson, 1999) focus on mothers who reside both with and without the child's biological father. Since child support and visitation are only germane for families with nonresident fathers, the results of studies considering all family structures may not be sufficiently generalizable to families without a resident father. Furthermore, many of these studies aggregate all forms of father involvement into a single scale incorporating emotional,

material (or instrumental), and social support. Thus if a positive association is observed between father involvement and lower depressive symptoms, it can be difficult to ascertain which specific aspects of the involvement drive this effect.

Studies using larger samples and disaggregated measures of nonresident father involvement to study non-child outcomes have focused on economic-related outcomes of the mother and household. As previously discussed, Garasky and Stewart (2007) and Nepomnyaschy and Garfinkel (2011) find that nonresident father involvement, and especially father-child contact (i.e., visitation), can reduce measures of material hardship, including food insecurity. Food insecurity, for example, can directly increase levels of parental depression (Whitaker, Phillips, and Orzol 2006). More generally, economic strain and material hardship are highly associated with depressive symptomatology (e.g., Sullivan, Turner, and Danziger 2008; Heflin and Iceland 2009; Muntaner et al. 2004; Okechukwu et al. 2012). Thus, material hardship of the mother, which can reflect her economic circumstances, may be an important channel between nonresident father involvement and maternal depression.

Another important, but frequently overlooked element that influences maternal mental health, material hardship, and nonresident father involvement is paternal incarceration, especially in fragile families. Geller, Garfinkel and Western (2011) report that about half of all fathers in the Fragile Families and Child Wellbeing Study (FFCWS) spent some time in prison by the time of the interview corresponding to when the child was approximately five years old. This number was even higher for nonresident fathers (about 55% of such fathers). My analyses using the most recent follow-up version of the Fragile Families data (when the child was approximately nine years old) suggests that about two thirds of nonresident fathers were ever incarcerated by the time the child was approximately nine years old. Paternal incarceration is an important variable

to consider in this sample and study because of its strong ties to family stability, maternal hardship (including depression), and financial contributions. Geller, Garfinkel and Western tied paternal incarceration to substantially fewer contributions to their families (including formal and informal support). In addition to reducing the dollar amount fathers pay to their families (in effect, lowering the mother's income level), paternal incarceration also directly contributes to more strained household resources¹, as measured by the number of material hardship events that the mother experienced (e.g., eviction, inability to pay mortgage or utility bills) (Schwartz-Soicher, Geller, and Garfinkel 2011). Schwartz-Soicher, Geller, and Garfinkel found that both paternal contributions, along with maternal mental health, explained part of the effect that paternal incarceration had on material hardship of the mother. Additional recent evidence using these data suggests that paternal incarceration may directly increase maternal depression (Wildeman, Schnittker, and Turney 2012).

Finally, it should be noted that the contributions that nonresident fathers can make to their children are often diverse, complex, and highly correlated. In addition to formal child support payments, a nonresident father has several material and non-material mechanisms with which to provide support to his child. Most notably, this includes nonmonetary support, such as aiding the mother in child-rearing through visitation, informal support (monetary support not mandated by a formal agreement), or in-kind (non-cash) support, such as purchasing toys, diapers, or medicine for the child. The majority of work linking nonresident father involvement to child outcomes has focused on monetary child support. However, there is evidence to suggest that child support might be highly correlated with other forms of involvement, such as visitation or in-kind (non-cash) support. For instance, Nepomnyaschy (2007) finds that being paid any

¹ Though reverse causality is possible in this context (i.e., non-payment of child support may lead to paternal incarceration), it has been reported that this probability is very low in practice (Geller, Garfinkel, and Western 2011).

child support payment or any informal support significantly enhances the likelihood of subsequent visitation by the father. Likewise, Garasky et al. (2010) find that in-kind support, formal support, and visitation are highly correlated with one another. In addition to formal child support payments, both informal support and visitation are also associated with improved household outcomes. Nepomnyaschy and Garfinkel (2011) report that in the Fragile Families and Child Wellbeing Study, the correlation between days of contact and informal support was more than 0.50. Informal support and in-kind support were also highly correlated (over 0.30). Craigie (2011) highlights the common use of informal support and in-kind support among households headed by Black mothers, and the correlation of informal support with other forms of nonresident father involvement.

In summary, the inter-relationship between paternal involvement (or the lack thereof), maternal economic and mental well-being, and child outcomes underscore the need to ascertain different avenues by which fathers affect their nonresident children. As maternal depression is highly related to her economic circumstances, which can be positively affected by nonresident father involvement, this study tests the hypothesis that this involvement can have positive effects on her mental health, as well.

4.2. Data and Empirical Strategy

The data come from the three most recent waves of the Fragile Families and Child Wellbeing Study (FFCWS). The FFCWS is a nationally representative study of unmarried mothers (when weighted) that follows the mothers and fathers of a group of approximately 4,900 focal children born between 1998 and 2000, the majority of whom (3,710) were born out of wedlock. Mothers and fathers were subsequently re-interviewed by person or telephone when the

child was one-, three-, five-, and nine- years old (Waves 2, 3, 4, and 5 respectively). There were three levels of sampling: major cities, hospitals within cities, and births within hospitals, where the mothers were originally interviewed. Ultimately, 75 hospitals were included within 20 cities with a population greater than 200,000². Of the mothers interviewed at the child's birth (Wave 1), there was an 89, 86, 84, and 72 percent follow-up at Wave 2, 3, 4, and 5, respectively. The response rate for the subset of mothers who gave birth out-of-wedlock was very similar to the response rate for the entire sample.

The FFCWS is well suited to examine nonresident father involvement since it asks mothers and fathers several questions about father involvement, including financial contributions and contact, over a long period of time. Further, the FFCWS oversamples mothers who gave birth out-of-wedlock, which provides a relatively large sample to study activities that affect mothers living without the child's father. Studies employing cross-sectional data using measures of father involvement and child outcomes, especially in regard to nonresident fathers, are often unable to capture some of the dynamics of nonresident father involvement over time, in that fathers tend to shift their modes of involvement over time, often towards formal support and away from visitation and more informal measures of support (as discussed in Section 3).

I examine the effects of nonresident father involvement on maternal mental health by considering pooled, linear, models incorporating a set of fixed and time-varying predictors of maternal depression, including measures of nonresident father involvement (all of which vary over time). I also adopt a more conservative model by including mother-level fixed effects which relies on the changes of time-varying factors. To perform these analyses, I use the three most

² The twenty cities are: Oakland, CA; San Jose, CA; Jacksonville, FL; Chicago, IL; Indianapolis, IN; Boston, MA; Baltimore, MD; Detroit, MI; Newark, NJ; New York City, NY; Toledo, OH; Philadelphia, PA; Pittsburgh, PA; Nashville, TN; Austin, TX ; Corpus Christi, TX; San Antonio, TX; Norfolk, VA; Richmond, VA; and Milwaukee, WI.

recent waves, which were conducted in 2002-2003, 2003-2006, and 2007-2010. The interviews corresponded to when the child was approximately 3, 5, and 9 years old. The analysis is restricted to the most recent three waves because measures of maternal depression and nonresident father involvement were not consistently collected prior to the third year interview. In the first year interview, measures of depression were not collected identically across all twenty cities. Further, measures of informal child support were not collected for the mothers who received formal child support at the time of the first year interview, which would limit the analysis of informal child support only to those who did not have child support awards during this period.

Given that nonresident father involvement is only relevant for mothers and children who are not living with the child's biological father, the sample used consists of mothers who gave birth out-of-wedlock and who were not cohabiting with or married to the child's father at the time of the survey interview, and who were not missing key explanatory variables (i.e., measures of nonresident father involvement)³ or the dependent variable (maternal depression). Mothers are excluded if the focal child (or the biological father of that child) died⁴. The final sample consisted of 2,338 unique mothers and 4,658 observations. This included 1,502 mothers from the third-year interview, 1,704 mothers from the fifth-year interview, and 1,452 mothers from the ninth-year interview. The increase in the sample size from the third- to fifth- year interview

³ Table 10 presents sample characteristics for mothers who were not missing or missing any measures of nonresident father involvement. Overall, the two groups are fairly balanced in important domains such as income and depression levels. Because mothers who are missing data have statistically similar depression levels as those who are not missing data, performing the analysis by excluding these observations (as has been done in this analysis) is unlikely to have a large impact on the results. However, mothers who did not report levels of nonresident father involvement were less likely to report baseline support levels during pregnancy.

⁴ The sample is restricted to mothers who gave birth out-of-wedlock, since baseline (i.e., at birth) father support variables were only asked of unmarried mothers. Furthermore, I use the unmarried sample to be consistent with the existing literature that focuses on the effects of nonresident father involvement with children of out-of-wedlock births in Fragile Families (see, for example, Craigie (2011), Nepomnyaschy (2007), Nepomnyaschy and Garfinkel (2010), and Nepomnyaschy and Garfinkel (2011)).

reflects the relatively low rate of attrition combined with a substantial portion of fathers moving out of cohabitation (or marriage) with the mother. While additional mothers fell out of cohabitation (or marriage) with the biological father from the fifth-year to the ninth-year interview, the higher rate of attrition during this time period drove the overall sample downward. I use only the mother's report for all of the analyses in this paper, since mother's perceptions and reports of father involvement may be more germane to her own health. Additionally, many fathers (especially nonresident fathers) are lost to follow-up, which would severely limit the sample size. Further, including only fathers who complete the survey would introduce a large selection problem, since both observed and unobserved factors may induce the father to complete the survey.

4.2.1. Maternal Depression

The primary outcome measure used in this study is maternal depression status. Mental health was assessed in the context of symptoms for major depressive disorder (MD). In Waves 2 and beyond, the FFCWS administered the short form of the Composite International Diagnostic Interview, Short Form (CIDI-SF), which asks questions about some of the primary symptoms of depression that comprise the Diagnostic and Statistical Manual IV (DSM-IV) criteria. Specifically, mothers are asked about weight gain or loss, losing interest in hobbies or work, trouble concentrating, trouble sleeping, feeling blue, feeling tired, and having ideations about death for a period of two weeks during the past year. The CIDI-SF scale for MD is the sum of these relevant symptoms. A CIDI score of 3 is generally used as a threshold for diagnosing mental illness (Kessler et al. 1998). Therefore, the mother is considered as having evidence of MD if her CIDI score is 3 or above.

In addition to the binary variable indicating if a mother's CIDI-SF score exceeds the cut-point of 3, Kessler et al. (1998) provide a probability of caseness, which maps the count score on the CIDI-SF scale to a probability that the respondent meet the diagnostic criteria were she given the full CIDI battery. The relationship between the CIDI-SF score and the probability of caseness are shown in Figure 2, which demonstrates a precipitous rise in the probability of being a “true” depression case (as measured by the full CIDI scale), as the CIDI-SF score increases, especially after 3. In robustness checks, I explore using the probability of caseness as a dependent variable, to potentially gain more information than a simple binary indicator of depression.

4.2.2. Measures of Independent Variables, Including Nonresident Father Involvement

In this study, involvement among nonresident fathers is characterized by four primary modalities: formal child support, informal child support, in-kind (non-cash) support⁵, and father-child contact (i.e., visitation). In this study, I choose to employ binary measures to denote if any formal, informal, or in-kind support was reported. The use of a binary scale, rather than a continuous dollar amount, has several advantages in the context of this study. First, these measures are all reported by mothers, which will invariably suffer from measurement error which is likely downwardly biased (see Table 9 for a comparison of mother and father reports of father involvement measures). Because of the wording of the questions, it may be difficult for some mothers to be able to fully recall or give the exact amount of support that was received. Additionally, mothers often do not report the amount of the child support arrangement with the time the agreement was reached, which makes it very difficult to create a normalized measure of monthly child support receipt. A binary indicator therefore affords a consistent measure of

⁵ In-kind support is not assigned a numerical dollar value in the FFCWS. I define the father to have provided any in-kind support if he often or sometimes bought any of a list of items listed by the Study for the child, including toys, medicine, food, school or camp tuition, school supplies, entertainment items, personal items, or “anything else.”

nonresident father involvement across all eligible mothers, which is less sensitive to maternal reporting bias.

Unlike forms of material support, visitation is collected on a continuous scale that represents how many days in the past 30 days (at the time of the survey interview) the father visited the child. This variable was coded as 0 for fathers who did not see their child at all during the one to two year period⁶ prior to the interview wave. In addition to the continuous variable, a binary variable is added to models to represent if fathers had *ever* visited the child in the year(s) prior to the interview. The addition of this dichotomous variable is important because fathers who visit the child zero times in the past 30 days could have either never visited the child at all in the past year(s), or alternatively, could have visited the child in the past year(s), but not in the past 30 days. This distinction is important, and as such, the effect of various functional forms of paternal visitation on maternal depression is explored.

As a result of the substantial overlap between the four modes of nonresident father involvement that are considered here, I present an alternate specification involving a set of dichotomous variables representing mutually exclusive categories of involvement. Though there are 16 possible categories, which are shown in Table 6, some of these categories contain very few mothers. This is especially true among mothers whose child's father does not visit, but does provide specific informal means of support (e.g., in-kind or informal support). As a result, six categories of involvement involving these mothers whose fathers do not visit, but provide either informal or in-kind support are collapsed into one variable, giving a total of 11 mutually exclusive categories of nonresident father involvement. Since a continuous measure of visitation

⁶ In the ninth-year interview, paternal visitation is only collected up to one year prior to the interview. The first-, third-, and fifth- year interviews collected information on visitation during the two years prior to the interview.

is difficult to implement here, the alternate specification only considers a binary measure indicating if the child's father ever visited in the last one- to two- years.

Two important variables that are related to both nonresident father involvement and maternal psychological distress include material hardship and paternal incarceration. To measure material hardship, I define the measure following Nepomyashchy and Garfinkel (2011). The measure represents the sum of eight questions representing if they had experienced any of the following events in the past 12 months prior to the interview: received free food or meals, incomplete payment of rent or mortgage, incomplete payment of utility bill(s), had utility service suspended, had phone service suspended, been evicted for nonpayment of rent or mortgage, stayed in a shelter, abandoned building, or automobile. The scale therefore ranges from zero to eight, with higher values indicating a higher number of reported hardships. The inclusion of these constituents as measures of material hardship has been used in other studies, as well (e.g., Wildeman, Schnittker, and Turney 2012). Questions about paternal incarceration are asked both to the mother and father at each interview wave. If either the mother or father responds affirmatively, the father is considered to have *ever* been incarcerated up to that survey period⁷.

Several time-invariant and time-varying covariates are employed to control for observed heterogeneity. Time-invariant variables included demographic factors such the race and ethnicity of the mother, the mother's educational level at the time of the child's birth, the age of the mother at the child's birth. Other important time-invariant measures that are used include a measure of the mother's cognitive ability (as measured by the Wechsler Adult Intelligence Scale-Revised; WAIS-R), impulsivity, measures of depression among either of the mother's parents,

⁷ In addition to missing values, incarceration is a noisy measure in these data. Misreporting is possible if individuals do not understand what incarceration means or if there is a stigma attached to the questions. Incarceration may also not be internally consistent in that there is heterogeneity in the length and nature of the incarceration. If the measurement error associated with this variable is assumed to be random, the coefficient on this variable would be biased toward the null.

and paternal support measures during pregnancy and at the time of the child's birth⁸. These measures included if the child took the father's last name, if the child's biological father was on the birth certificate, and if he contributed material items during the mother's pregnancy. Time-varying variables include measures of social support of the mother⁹, number of times the mother moved since the previous survey wave, self-reported health status by the mother, multi-partner fertility of the mother¹⁰ and father, maternal income, along with household composition including number of adults, children, and current partner other than the child's father.

Incomplete maternal reports on several variables necessitated imputation of several variables. Specifically, paternal multi-partner fertility, baseline measures of paternal support, paternal incarceration, maternal cognitive ability and impulsivity, and if either of the mother's parents suffered from depression were imputed at the mean values for these measures at each wave. Dichotomous imputation flags were included in all models where imputed variables were used.

4.2.3. Empirical Strategy

The central question of this study is to ascertain the relationship between nonresident father involvement and maternal depression. As a baseline specification, a pooled linear probability model can be estimated whereby maternal depression is regressed on a vector of father involvement and demographic variables in Equation 1.

$$(1) \quad H_{it} = \alpha + \beta F_{it} + \theta X_{it} + \phi Z_{it} + \varepsilon_{it}$$

⁸ Baseline father support measures were available at the baseline interview only for unmarried couples. Maternal impulsivity, cognitive ability, and parental depression were measured during the 3-year follow-up interview. These measures are defined similarly to Nepomnyaschy and Garfinkel (2011).

⁹ Social support is defined as the sum of the mother's answers to three questions including if in the next year, she could count on someone to help her with emergency child care, provide her with a place to live, or loan her \$200.

¹⁰ Maternal multi-partner fertility was directly reported by the mother up to the fifth-year interview. At the ninth-year interview, maternal multi-partner fertility was inferred through a question listing the mother's biological children.

In this specification, H represents if the mother is classified as depressed by meeting or exceeding the cutoff value of the CIDI scale. The variable F represents several measures of father involvement, including visitation, any in-kind support, any informal support and any formal child support. All measures of nonresident father involvement, with the exception of days of visitation, are coded as dichotomous variables¹¹. An additional binary variable indicating if the father had ever visited in the past two years (or one year in the case of the ninth-year survey) is also included. The vectors denoted by Z and X represents time-invariant and time-varying factors, respectively, that may influence both maternal psychological distress and father involvement. Included as a time-varying regressor is household income, which includes both formal and informal receipts. An advantage to this pooled specification is its ability to account for several baseline, or static, factors that are related to father involvement and maternal depression. Such factors include father contributions at birth, commitment to provide support, and maternal impulsivity, among others. In Equation 1, a significant β coefficient would represent an effect of father involvement that is above and beyond the effect of income, which would be consistent with the broader literature as to the effects of child support. Since this regression is pooled over three periods, standard errors are clustered at the mother level to adjust the standard errors for the correlation between multiple observations from the same mother. As previously discussed, it is likely that many of the father involvement measures are highly correlated with each other, especially those surrounding informal forms of involvement, such as informal support, in-kind support, and visitation. I perform two analyses to ascertain the sensitivity of my results to this issue. First, I use a set of eleven mutually exclusive categories of nonresident father involvement combinations (as discussed above). Second, robustness checks

¹¹ Alternate specifications use eleven dichotomous, mutually exclusive, categories representing possible combinations of nonresident father involvement.

are performed in which the models are estimated using only two father involvement variables with relatively low correlations (i.e., visitation and formal support). The informal support measures are then iteratively added to assess the extent to which the estimates become attenuated due to multi-collinearity.

A potential problem with the pooled specification is that the effects of father involvement on maternal health outcomes might be driven in part by time-invariant unobserved heterogeneity of the mother. For instance, mothers who are motivated to collect child support awards (or informally promised funds) might be more likely to ultimately receive such awards. Likewise, mothers who have lower levels of mental distress and who desire a strong father-child relationship might permit the father increased access to the child. Such unobserved time-invariant factors might be positively correlated with parenting stress or depression levels, and may therefore introduce bias into the measured effect of father involvement on maternal distress. Similarly, unobserved factors of the child or father might also influence this relationship as well. However, as the FFCWS only considers one focal child per mother, it is impossible to isolate child- or father- level unobserved heterogeneity using these data.

In dealing with this empirically, I exploit the panel nature of this data set by estimating a three-period fixed-effects model, in Equation 2.

$$(2) \quad H_{it} = \delta + \psi F_{it} + \varphi X_{it} + \mu_i + \lambda_t + u_{it}$$

The inclusion of (mother-level) fixed effects is represented by μ_i , which represent the unobserved, time-invariant factors of the mother that are captured by the model. Time period fixed effects are included as λ_t . The subscript t represents the survey wave (third, fifth, or ninth-

year follow-up interview). Only time-varying covariates are included in X_{it} . The fixed effects estimates are identified by changes in the nonresident father involvement and maternal depression. As with the pooled specifications, sensitivity analyses are performed (with particular attention to the multi-collinearity of the nonresident father involvement measures) with the fixed effects models.

While a fixed-effects estimator is a powerful tool to address unobserved heterogeneity, it can only account for this bias to the extent that omitted variables are time-invariant. In addition to time-invariant unobserved heterogeneity, time-varying heterogeneity and reverse causality may also introduce bias in this relationship. It is possible, for example, for maternal psychological distress to discourage the material and non-material contributions of nonresident fathers. While this study focuses on the effects of nonresident father involvement on maternal depression, it is possible that maternal depression may discourage subsequent father involvement, especially in regards to forms of involvement that are not formally enforced, such as informal support, in-kind support, and visitation.

In this case, another approach such as instrumental variables would be desirable to allow consistent identification of the effects of father involvement on maternal mental health. Though previous literature in cognitive domains have relied upon state-level variation in child support expenditures and enforcement as instruments for child support receipts, there are no similarly widely-used instrumental variables for either child visitation, in-kind support, or informal support (though such instruments may theoretically exist). Furthermore, given the use of multiple measures of nonresident father involvement, there would need to be at least as many instruments as father involvement variables (in this case, five separate instruments), which is very difficult in practice. Further, Nepomnyaschy (2007) argues that including child support expenditure or

enforcement variables for formal child support may bias estimates since higher enforcement of child support may drive fathers to prefer informal support. Further, panel data techniques, such as fixed effects, may provide for identification in the absence of suitable instrumental variables, since many of the unobserved effects that may be related to father involvement and maternal mental distress are likely to be static across time. The extent of the bias of the fixed effects estimator is based on the nature of the unobserved factors driving maternal depression that are related to nonresident father involvement. Since the unobserved factors driving depression are likely to be related to lower (or even under-reported) levels of nonresident father involvement, the true difference is likely larger. In this sense, the estimates of these models provide a lower bound on the effects of nonresident father involvement on maternal depression.

4.3. Results and Discussion

4.3.1. Descriptive Analysis

Table 1 presents several characteristics of mothers according to their depression status at the time of the survey interview. The overall characteristics of the sample, reported in Column 1, suggest that these mothers are quite disadvantaged in many respects. About 20% of the mothers are classified as being depressed, which is higher than national prevalence rates of about 15% in mothers of young children. In considering the time-invariant characteristics of mothers in the entire sample, women were relatively young at the baseline interview (approximately 23 years old), which was done in the hospital shortly after the child was born. Approximately 64% are Black, and 20% were of Hispanic origin (note that both Hispanic Blacks and non-Hispanic Blacks are included in this sample). Further, these women are relatively poorly educated; only about 3% of the sample finished college. However, many women completed some amount of

college. It should be noted that several women completed some training or educational program (including college graduation) throughout the later periods of the FFCWS, so education in this sample isn't truly a "time-invariant" variable, but the changes are modest. A unique characteristic of this sample includes the very high prevalence of multi-partner fertility and paternal incarceration. Overall, nearly 6 out of 10 mothers had a child with an additional man other than the child's father. Similarly, almost six out of every ten children in this sample had a biological father who had children with a woman other than their biological mother. A similar percentage of children had their fathers ever spend time in prison. These high rates of incarceration and multi-partner fertility are due in large part to these events occurring between the fifth- year and ninth- year interviews. Additionally, relatively high rates of material hardship are seen; the mothers in the sample averaged about one hardship measure out of a total of 8 possible hardship scenarios.

In terms of nonresident father involvement, visitation and in-kind support are relatively common among these families, whereas the receipt of any formal or informal support is relatively less common. However, there are fairly marked changes of the patterns of nonresident father involvement in this sample as children and mothers evolve through the Fragile Families Study. Figure 1 depicts some of these changes over time. In Figure 1, the probability of receipt of any formal support rises by more than ten percentage points as children age from 3 to 9 years old. Both informal support and in-kind support decrease precipitously over this time period as well, with informal support dropping to very low levels at the 9-year interview. Father-child contact also declines over time, both in terms of the number of days of visitation as well as any visitation in a 1-2 year period. These similar trends among informal support, in-kind support, and father-child contact gives rise to the possibility of high collinearity between these measures in

this sample. In the analytic sample, the highest levels of collinearity include in-kind support and informal support, as well as in-kind support and visitation, with correlations exceeding 0.50 in both cases. This is not surprising in light of the previous literature reporting the inter-relationships between these measures, as well as literature documenting similarly high correlations. Intuitively, this implies that the vast majority of fathers who provide in-kind or informal support also visit the child more frequently.

The second and third columns of Table 1 report these characteristics of mothers who are classified as being depressed compared to those who are not classified as depressed. In considering the unadjusted differences in nonresident father involvement across maternal depression, mothers who experience depression are much less likely to receive in-kind support from the child's biological father. Mothers who do not suffer from depression tend to report more father visits, but are less likely to report that the father ever visited. Formal support is not statistically different between depressed and non-depressed mothers, while depressed mothers are more likely to report receiving informal support. Other variables that are positively associated with maternal depression include lower self-reported health status, impulsivity, parental depression, number of material hardships and paternal incarceration. Higher maternal education levels, the age of the child, income, and religious attendance are negatively associated with depression.

4.3.2. Baseline and fixed effects models

I begin by employing pooled Ordinary Least Squares (OLS) models to clarify the relationship between father involvement and maternal depression. As described above, the analysis is conducted on the sample of mothers who gave birth out-of-wedlock and were not

cohabiting with (or married to) the child's biological father at the time of the survey interview, representing the primary sample that will benefit from nonresident father involvement. Table 2 presents the results of OLS models predicting maternal depression, as a function of involvement levels of nonresident fathers. Additionally, the results of three models are reported in this Table, each with an expanding control set that includes a baseline specification (Column 1), measures of father involvement at birth and time-invariant predictors of maternal depression (Column 2), as well as the full model, including material hardship and paternal incarceration (Column 3). All OLS specifications include clustered standard errors at the mother level to account for correlations in the mother's responses across time, as well as panel indicator variables.

In the baseline specification, which only controls for the mother's health and demographic characteristics (Column 1), the most important and somewhat surprising result in the context of this model is the strong positive coefficient of about 5.6 percentage points on nonresident fathers visiting the child at all in the past one-to-two years, relative to the fathers who did not visit the child at all in this period. Additionally, there was a strong protective effect of in-kind support on maternal depression. There were no significant effects for the father's payment of either formal or informal support, though the number of days in which the father saw the child is marginally protective against depression. Once a richer set of variables were added, representing the father's baseline involvement and contributions, along with maternal characteristics potentially related to her mental health (Column 2), the magnitude and significance of the father involvement variables dropped slightly, though not substantially so. The number of visits ceased to be significant in this specification. Maternal impulsivity and parental depression were highly significant in these specifications, which is consistent with the descriptive statistics, along with prior work using this data (e.g., Wildeman, Schnittker and

Turney, 2012). When material hardship and paternal incarceration were added (Column 3), these coefficients decreased even further. However, the coefficients for any paternal visitation and in-kind support remained significant at or above the 1% level in all three specifications. The measure denoting if the father was ever incarcerated, surprisingly, had little effect on maternal depression in this sample. However, the drop in effect size for visitation and in-kind support does suggest that these variables, and material hardship in particular, are likely important in the relationship between nonresident father involvement and maternal depression.

While the baseline models of depression did control for a rich set of both time-varying and time-invariant factors that may affect both father involvement and depression, there may be important baseline variables that were omitted in these models, which could potentially bias the estimates obtained. Fixed effects models have the potential to eliminate unobserved heterogeneity, and give the most plausible estimates for the effects of nonresident father involvement on maternal depression. Table 3 reports fixed effects models for the model denoted in Equation 2. Column 1 reports the baseline model, whereas Column 2 includes measures of material hardship and paternal incarceration. Overall, the fixed effects analyses support the findings from the baseline models (reported in Table 2). In-kind support tends to be protective against maternal depression, while any paternal visitation in the past one-two years tends to increase maternal depression propensity. In general, the magnitude of the coefficient for any parental visitation was similar in both the baseline models at around a 5 percentage point harmful effect. However, the effect of in-kind support, while strongly significant in baseline models, is only marginally significant at the 10% level, with a lower magnitude, in the fixed effects models. Further, like the pooled OLS models, the inclusion of material hardship and incarceration made a relatively small impact in the magnitude of the fixed effects coefficient. The fixed effects model,

like the pooled model, showed that the number of material hardships played a large role in determining maternal depression.

Column 2 reports the findings when including two important time-varying variables that may affect the fixed effects estimates of paternal involvement: material hardship, and paternal incarceration. While paternal incarceration is positively related to maternal depression, it is not significantly significant, similarly to the results obtained from the baseline model. Like the pooled models, material hardship was strongly related to maternal depression. However, the coefficients on the paternal involvement variables did not attenuate substantially with the addition of these two variables, suggesting that nonresident father involvement likely has an independent effect in affecting maternal depression levels. Like the baseline models, the fixed effects models suggest small amounts of visitation may be especially harmful to mothers. Additionally, many fathers who visit the child frequently likely also provide in-kind support to the child, which may further reduce maternal depression.

Table 4 presents the results of alternate specifications that avoid multi-collinearity in the father involvement measures. As discussed above, combinations of the father involvement modalities are formed into 11 mutually exclusive categories. The category of no father involvement (i.e., no visitation, formal support, informal support, or in-kind support), serves as the reference category in these models. The results from column 1 support the general characterization of an adverse effect of visitation found in the models reported in Table 2. Specifically, relative to fathers who are completely uninvolved, fathers who visit have a positively signed effect on maternal depression (with the exception of those fathers who visit and provide in-kind support, but no informal support). Mothers whose child's father visits and provides informal support are about 9.2 percentage points more likely to suffer from depression

relative to mothers whose child's father remains uninvolved. Fathers who visited without providing any material support increased maternal depression by 3.9 percentage points, while fathers who visited and provided formal support increased the probability of maternal depression by 5.6 percentage points. The results of the fixed effects analyses, which are reported in Table 2, generally mirror the static models, and provide a larger effect size for each of the categories that were statistically significant in Column 1. While the models presented in Table 3 support the conclusion that visitation may be harmful, these models do not support the notion that in-kind support may be independently protective. Rather, on average, mothers whose child's father visits and provides in-kind support are less likely to suffer from depression compared to mothers whose child's father visits, but does not provide any material support.¹²

Table 5 explores which mothers are most sensitive to the involvement of the child's nonresident father. Overall, paternal visitation remained at least marginally positively significant in all specifications with the exception of the mothers who had relatively high rates of education (Column 2). Interestingly, mothers who did not live with another partner tended to be more strongly affected by paternal visitation, mothers where the child's father ever visited the child were 6 percentage points more likely to suffer from depression compared to fathers who never visited. Mothers who reported no hardships reported were slightly less adversely affected by paternal visitation, while mothers whose child's father never was incarcerated tended to have a slightly higher sensitivity to paternal visitation.¹³

While it is plausible that increased involvement through in-kind support may reduce levels of depression, it is somewhat surprising that father visitation, especially infrequently,

¹² This difference is significant at the 5% level.

¹³ Additional robustness checks revealed that mothers whose child's biological father visited in the past year, but not in the past 30 days were driving a large portion of the effects. These results are available from the author upon request.

seems to adversely affect mothers. The fathers whose visitation is harmful to the mother may represent a subset of fathers with whom the mother has a very poor relationship with, and that any agitation in the relationship may induce the mother to become more depressed under these circumstances. Alternatively, mothers who had the child with a desirable partner may “miss” his presence which might contribute to her depression. Further, the substance of father’s visitation may be harmful to the welfare of the child, in which the father may not participate in any investments to the welfare of the household or child. It is important to note that these effects persist even after accounting for material hardship and paternal incarceration, which could plausibly affect both father involvement and maternal depression. Other variables, such as ethnicity, race, maternal education, and lower self-reported health status tend to increase the risk of depression, as the literature has suggested (e.g., Turney and Carlson 2011).

In comparing my results to those of Nepomnyaschy and Garfinkel (2011), who use the same data set and a very similar sample, they find a net beneficial effect of father-child visitation on maternal material hardship. However, they treat visitation as a purely linear measure (i.e., days of contact). Given that Garasky and Stewart (2007) found that paternal visitation reduced food insecurity only when done so frequently (i.e., more than once per week), it is likely that the (relatively few) fathers who visit the most frequently were driving much of the beneficial effects of visitation documented by Nepomnyaschy and Garfinkel¹⁴. In this study, similar to the aforementioned two studies, visitation does not appear to affect maternal depression in a linear fashion, though the net effect of high amounts visitation may reduce depression levels in some mothers, especially those receiving either informal support or in-kind support. Additionally, while I find that in-kind support is arguably the most protective against maternal depression,

¹⁴ Replication of the results of Nepomnyaschy and Garfinkel (2011) using my sample revealed that the protective effect of nonresident father visitation on material hardship was driven by fathers who visited very frequently (i.e., close to 30 days in the past month). These fathers did not play a similar protective role on maternal depression.

Nepomnyaschy and Garfinkel found that in-kind support was not uniformly beneficial to reducing hardship. In fact, the receipt of in-kind support was actually found to marginally increase the probability of experiencing certain measures of material hardship, such as eviction. Thus, even though there is a strong relationship between material hardship and depression, father involvement appears to affect these measures in different ways. While formal child support was found to be beneficial to the welfare of children, my results, combined with the results of related literature, suggest that the indirect effects on mothers may be somewhat limited. Possible explanations include the relatively disadvantaged sample, the inability of fathers to make adequate monetary payments, as well as reverse causality. It is also possible that mothers seeking child support may become involved in the welfare system, which may carry a stigma that can affect mental health.

These models provide support that for some mothers, visitation may actually serve to do more harm than good, especially among fathers that are relatively uninvolved and do not visit frequently. Conversely, receiving any in-kind support appears to be the most beneficial form of material involvement, rather than any monetary support. The results of these models also suggests the importance of considering disaggregated forms of involvement, since not all forms of nonresident father involvement have similar effects, in either magnitude or direction, on maternal depression.

4.3.3. Robustness Checks

In addition to the alternate specifications of pooled and fixed effects linear probability models discussed previously, I test the robustness of my results to several other assumptions I made in the paper, including the use of a linear model, my treatment of missing values (i.e.,

imputation at the mean), as well as the use of a binary scale for the classification of depression. First, the use of an unbalanced panel, as opposed to the balanced panel of mothers responding to the third-, fifth-, and ninth- year interviews, can introduce bias through sample loss, and specifically through nonrandom mother nonresponse patterns. The unbalanced panel also includes mothers who recently stopped living with the child's father, whereas the balanced panel focuses on mothers who have been living apart from the child's father for long periods of time. The results of these models restricted to the fully balanced sample are largely similar to the results for the full analytic sample (see Panel A of Table 7). Further, to ascertain the potential attenuation of the effects of the involvement variables due to multi-collinearity, I report several specifications varying the inclusion of the nonresident father involvement measures, starting with the two most commonly studied measures (formal support and visitation) in Column 1 of Table 7, and iteratively adding informal and in-kind support. Overall, the major effects of any father-child contact and any in-kind support remain robust across these specifications. However, the number of days of visitation, which was previously marginally protective against maternal depression, ceased to be significant with the inclusion of in-kind support.

Additionally, I compare the results obtained from the binary dependent variable to those obtained using a dependent variable representing the numerical probability that a respondent would meet the full depression criterion. These probabilities are imputed using the CIDI-SF score following Kessler et al. (1998)¹⁵. Column 1 of Table 8 shows OLS results when a continuous probability of depression is used to consider the effects of nonresident father involvement on the continuous depression variable. Overall, there is very little difference in the results using the continuous variable, with the exception of the fixed effects case (Panel B),

¹⁵ Kessler et al. (1998) uses the National Comorbidity Study (NCS) to generate the probability of baseness based on a respondent's CIDI-SF score. The NCS is a nationally representative survey, and as such the imputed probability of caseness for each CIDI-SF score may be somewhat different for the sample used in this study.

where in-kind support plays a slightly lower role in alleviating depression, and any visitation has a slightly larger effect in aggravating depression, compared to the binary case.

To ascertain the sensitivity of my results to the treatment of missing data, Column 2 of Table 8 shows the re-estimation of the model using only the observations for which the explanatory variables were non-missing, rather than imputation at the mean, which consider only observations with non-missing values on for all variables specified in the model. Overall, there is relatively little change in the significance of in-kind support and any visitation. Obtaining significant results using two relatively conservative samples (using observations with only non-missing values as well as the balanced panel) suggest that either attrition or imputation unduly affect the main results obtained in this study.

Finally, it is possible that nonresident father involvement may act on maternal mental health in a non-linear manner. Furthermore, a non-linear estimator (such as a logit estimator) may produce different estimates than a linear model, which could potentially change results. Results from the estimation of the model using a logit (reported in Column 3 of Table 8), rather than linear, model do not substantively change the results in the pooled specifications. In fixed effects case, in-kind support ceases be significant, but its direction is still negative. Furthermore, the use of a binary indicator for the various forms of father involvement preclude more sophisticated functional forms (e.g., quadratic or cubic terms). In addition to the alternate specifications using mutually exclusive categories presented in Table 4, I explore interaction terms of formal support with the other types of material support and visitation that are considered in other models. The addition of the formal support interaction terms (Column 4 of Table 8) does not change the overall effects of father involvement on maternal depression as reported in the baseline or linear fixed effects models.

Notwithstanding the sensitivity analyses, this study has several limitations. Firstly, some of the results presented may not be sufficiently generalizable for mothers with higher incomes or who are living in more rural or suburban locations, since data used in my analyses were only collected in cities for a sample of relatively disadvantaged mothers. Secondly, though fixed effects analyses address bias resulting from time-invariant unobserved heterogeneity, other forms of bias may be present, the most of severe of which being reverse causality- maternal depression discouraging father involvement. Finally, because of the high correlation between in-kind support, formal support, informal support, and visitation, it is difficult to identify specific effects of each of these activities on maternal mental distress. However, this study provides evidence that father involvement may have indirect effects (though not always positive) on the mother's well-being.

4.4. Conclusions

A large body of literature indicates that single mothers overwhelmingly face an increased risk of adverse psychological consequences compared to their married counterparts. Further, adverse maternal health outcomes have been shown to be related to a myriad of child behavioral, psychological and cognitive deficiencies. On the other hand, father involvement and child support has been implicated in positive welfare gains for children, especially in cognitive and behavioral domains, and reductions in material hardship, such as food insecurity, for the mother. This study used three waves of panel data from the Fragile Families and Child Wellbeing Study to study how various forms of nonresident father involvement affect depression.

From the descriptive analysis, I find that unadjusted measures of father involvement are associated with maternal depression. While the receipt of any formal support is not substantially

related to maternal depression, increased amounts of other forms of involvement, such as the number of days of father-child contact, along with in-kind or informal support, are associated with lower levels of maternal depression.

Baseline linear probability models showed that in-kind support had the strongest protective role against maternal depression. Further, father-child contact, while having been shown to be beneficial in reducing measures of food insecurity in households, if infrequent, may in fact be harmful in promoting maternal depression in fragile families, especially among households in which the father is otherwise uninvolved. However, in households in which the father was involved in informal or in-kind support, visitation had a small beneficial effect on maternal depression. The receipt of in-kind support had the largest protective effect against depression among mothers. These results were robust to the inclusion of mother fixed effects and alternative estimation strategies.

The results of this study suggest that visitation should not be universally “forced” among children living with nonresident fathers. Small amounts of visitation, for certain families, could serve to increase depression in mothers, which can not only harm the mother, but negatively affect the child as well. Further, the encouragement of fathers to contribute to their families informally, especially those of lower incomes, may have benefits for the mother’s mental health as opposed to more formal support mechanisms. Finally, future studies can benefit from the study of these and other disaggregated forms of nonresident father involvement to better understand the indirect effects of these behaviors on mothers and families.

4.5. References

Amato, Paul R., and Joan G. Gilbreth. 1999. "Nonresident Fathers and Children's Well-Being: A Meta-Analysis." *Journal of Marriage and Family* no. 61 (3):557-573.

- Argys, L. M., H. E. Peters, J. Brooks-Gunn, and J. R. Smith. 1998. "The impact of child support on cognitive outcomes of young children." *Demography* no. 35 (2):159-73.
- Astone, Nan Marie, and Sara S. McLanahan. 1991. "Family Structure, Parental Practices and High School Completion." *American Sociological Review* no. 56 (3):309-320.
- Avison, William R., Jennifer Ali, and David Walters. 2007. "Family Structure, Stress, and Psychological Distress: A Demonstration of the Impact of Differential Exposure." *Journal of Health and Social Behavior* no. 48 (3):301-317.
- Bramlett, Matthew D., and Stephen J. Blumberg. 2007. "Family Structure And Children's Physical And Mental Health." *Health Affairs* no. 26 (2):549-558. doi: 10.1377/hlthaff.26.2.549.
- Breslau, J., K. S. Kendler, M. Su, S. Gaxiola-Aguilar, and R. C. Kessler. 2005. "Lifetime risk and persistence of psychiatric disorders across ethnic groups in the United States." *Psychol Med* no. 35 (3):317-27.
- Burgos, N. M., M. C. Lennon, M. Bravo, and J. Guzman. 1995. "Depressive symptomatology in single women heads of households in Puerto Rico: a comparative analysis." *Women Health* no. 23 (3):1-18.
- Chang, Jen Jen, Carolyn T. Halpern, and Jay S. Kaufman. 2007. "Maternal Depressive Symptoms, Father's Involvement, and the Trajectories of Child Problem Behaviors in a US National Sample." *Arch Pediatr Adolesc Med* no. 161 (7):697-703. doi: 10.1001/archpedi.161.7.697.
- Cooper, Carey E., Sara S. McLanahan, Sarah O. Meadows, and Jeanne Brooks-Gunn. 2009. "Family Structure Transitions and Maternal Parenting Stress." *Journal of Marriage and Family* no. 71 (3):558-574. doi: 10.1111/j.1741-3737.2009.00619.x.
- Craigie, Terry-Ann. 2011. "Informal Child Support Contributions in Black Female-Headed Families." *The Review of Black Political Economy*:1-7. doi: 10.1007/s12114-011-9115-1.
- Crosier, Timothy, Peter Butterworth, and Bryan Rodgers. 2007. "Mental health problems among single and partnered mothers." *Social Psychiatry and Psychiatric Epidemiology* no. 42 (1):6-13. doi: 10.1007/s00127-006-0125-4.
- DeKlyen, M., J. Brooks-Gunn, S. McLanahan, and J. Knab. 2006. "The mental health of married, cohabiting, and non-coresident parents with infants." *Am J Public Health* no. 96 (10):1836-41. doi: 10.2105/AJPH.2004.049296.
- Elsenbruch, S., S. Benson, M. Rucke, M. Rose, J. Dudenhausen, M. K. Pincus-Knackstedt, B. F. Klapp, and P. C. Arck. 2007. "Social support during pregnancy: effects on maternal depressive symptoms, smoking and pregnancy outcome." *Hum Reprod* no. 22 (3):869-77. doi: 10.1093/humrep/del432.
- Garasky, Steven, and Susan Stewart. 2007. "Evidence of the Effectiveness of Child Support and Visitation: Examining Food Insecurity among Children with Nonresident Fathers." *Journal of Family and Economic Issues* no. 28 (1):105-121. doi: 10.1007/s10834-006-9049-0.
- Garasky, Steven, Susan Stewart, Craig Gundersen, and Brenda Lohman. 2010. "Toward a Fuller Understanding of Nonresident Father Involvement: An Examination of Child Support, In-Kind Support, and Visitation." *Population Research and Policy Review* no. 29 (3):363-393. doi: 10.1007/s11113-009-9148-3.
- Geller, A., I. Garfinkel, and B. Western. 2011. "Paternal incarceration and support for children in fragile families." *Demography* no. 48 (1):25-47. doi: 10.1007/s13524-010-0009-9.

- Graham, John W., Andrea H. Beller, and Pedro M. Hernandez. 1994. "The Effects of Child Support on Educational Attainment." In *Child Support and Child Well-Being*, edited by I. Garfinkel, S.S. McLanahan and P.K. Robins. Washington, D.C.: The Urban Institute Press.
- Hamilton, Brady E., Joyce A. Martin, and Stephanie J. Ventura. 2011. Births: Preliminary Data for 2010. In *National Vital Statistics Reports 60*, no. 2, edited by National Center for Health Statistics. Hyattsville, MD.
- Harknett, Kristen. 2009. "Why are Children with Married Parents Healthier? The Case of Pediatric Asthma." *Population Research and Policy Review* no. 28 (3):347-365. doi: 10.1007/s11113-008-9102-9.
- Heflin, Colleen M., and John Iceland. 2009. "Poverty, Material Hardship, and Depression*." *Social Science Quarterly* no. 90 (5):1051-1071. doi: 10.1111/j.1540-6237.2009.00645.x.
- Hernandez, Pedro M., Andrea H. Beller, and John W. Graham. 1995. "Changes in the Relationship between Child Support Payments and Educational Attainment of Offspring, 1979-1988." *Demography* no. 32 (2):249-260.
- Jackson, A. P. 1999. "The effects of nonresident father involvement on single black mothers and their young children." *Soc Work* no. 44 (2):156-66.
- Kalil, Ariel, Kathleen M. Ziol-Guest, and Rebekah Levine Coley. 2005. "Perceptions of father involvement patterns in teenage-mother families: Predictors and links to mothers' psychological adjustment*." *Family Relations* no. 54 (2):197-211.
- Kessler, R. C., and S. Zhao. 1999. "Overview of descriptive epidemiology of mental disorders." In *Handbook of the sociology of mental health*, edited by C.S. Aneshensel and J.C. Phelan. New York: Kluwer Academic/Plenum Publishers.
- Kessler, Ronald C., Gavin Andrews, Daniel Mroczek, Bedirhan Ustun, and Hans-Ulrich Wittchen. 1998. "The World Health Organization Composite International Diagnostic Interview short-form (CIDI-SF)." *International Journal of Methods in Psychiatric Research* no. 7 (4):171-185. doi: 10.1002/mp.47.
- Kiernan, K., and K. E. Pickett. 2006. "Marital status disparities in maternal smoking during pregnancy, breastfeeding and maternal depression." *Soc Sci Med* no. 63 (2):335-46. doi: 10.1016/j.socscimed.2006.01.006.
- Knox, Virginia W. 1996. "The Effects of Child Support Payments on Developmental Outcomes for Elementary School-Age Children." *The Journal of Human Resources* no. 31 (4):816-840.
- Kramer, Michael S., John Lydon, Louise Séguin, Lise Goulet, Susan R. Kahn, Helen McNamara, Jacques Genest, Clément Dassa, Moy Fong Chen, Shakti Sharma, Michael J. Meaney, Steven Thomson, Stan Van Uum, Gideon Koren, Mourad Dahhou, Julie Lamoureux, and Robert W. Platt. 2009. "Stress Pathways to Spontaneous Preterm Birth: The Role of Stressors, Psychological Distress, and Stress Hormones." *American Journal of Epidemiology* no. 169 (11):1319-1326. doi: 10.1093/aje/kwp061.
- Kreider, R.M., and R. Ellis. 2011. "Living arrangements of children: 2009." *Current population reports*, no. P70-126.
- Krein, Sheila Fitzgerald, and Andrea H. Beller. 1988. "Educational Attainment of Children From Single-Parent Families: Differences by Exposure, Gender, and Race." *Demography* no. 25 (2):221-234.
- Langton, Callie E., and Lawrence M. Berger. 2011. "Family Structure and Adolescent Physical Health, Behavior, and Emotional Well-Being." *Social Service Review* no. 85 (3):323-357.

- Malik, Neena M., Neil W. Boris, Sherryl S. Heller, Brenda Jones Harden, Jane Squires, Rachel Chazan-Cohen, Linda S. Beeber, and Karen J. Kaczynski. 2007. "Risk for maternal depression and child aggression in Early Head Start families: A test of ecological models." *Infant Mental Health Journal* no. 28 (2):171-191. doi: 10.1002/imhj.20128.
- McLennan, J. D., M. Kotelchuck, and H. Cho. 2001. "Prevalence, persistence, and correlates of depressive symptoms in a national sample of mothers of toddlers." *J Am Acad Child Adolesc Psychiatry* no. 40 (11):1316-23. doi: 10.1097/00004583-200111000-00012.
- Meadows, Sarah O., Sara S. McLanahan, and Jeanne Brooks-Gunn. 2008. "Stability and Change in Family Structure and Maternal Health Trajectories." *American Sociological Review* no. 73 (2):314-334. doi: 10.1177/000312240807300207.
- Muntaner, Carles, William W. Eaton, Richard Miech, and Patricia O'Campo. 2004. "Socioeconomic Position and Major Mental Disorders." *Epidemiologic Reviews* no. 26 (1):53-62. doi: 10.1093/epirev/mxh001.
- Nepomnyaschy, L. 2007. "Child support and father-child contact: testing reciprocal pathways." *Demography* no. 44 (1):93-112.
- Nepomnyaschy, L., and I. Garfinkel. 2011. "Fathers' Involvement with Their Nonresident Children and Material Hardship." *Soc Serv Rev* no. 85 (1):3-38. doi: 10.1086/658394.
- Okechukwu, Cassandra A., Alison M. El Ayadi, Sara L. Tamers, Erika L. Sabbath, and Lisa Berkman. 2012. "Household Food Insufficiency, Financial Strain, Work-Family Spillover, and Depressive Symptoms in the Working Class: The Work, Family, and Health Network Study." *Am J Public Health* no. 102 (1):126-133. doi: 10.2105/ajph.2011.300323.
- Osborne, Cynthia, Lawrence Berger, and Katherine Magnuson. 2012. "Family Structure Transitions and Changes in Maternal Resources and Well-being." *Demography* no. 49 (1):23-47. doi: 10.1007/s13524-011-0080-x.
- Schwartz-Soicher, Ofira, Amanda Geller, and Irwin Garfinkel. 2011. "The Effect of Paternal Incarceration on Material Hardship." *Social Service Review* no. 85 (3):447-473.
- Smith, L. E., and K. S. Howard. 2008. "Continuity of paternal social support and depressive symptoms among new mothers." *J Fam Psychol* no. 22 (5):763-73. doi: 10.1037/a0013581.
- Stewart, Susan D., and Chadwick L. Menning. 2009. "Family Structure, Nonresident Father Involvement, and Adolescent Eating Patterns." *Journal of Adolescent Health* no. 45 (2):193-201.
- Strauss, R. S., and J. Knight. 1999. "Influence of the home environment on the development of obesity in children." *Pediatrics* no. 103 (6):e85.
- Sullivan, James X., Lesley Turner, and Sheldon Danziger. 2008. "The relationship between income and material hardship." *Journal of Policy Analysis and Management* no. 27 (1):63-81. doi: 10.1002/pam.20307.
- Surkan, P. J., I. Kawachi, and K. E. Peterson. 2008. "Childhood overweight and maternal depressive symptoms." *J Epidemiol Community Health* no. 62 (5):e11.
- Talge, Nicole M., Charles Neal, Vivette Glover, Translational Research the Early Stress, Fetal Prevention Science Network, Child Neonatal Experience on, and Health Adolescent Mental. 2007. "Antenatal maternal stress and long-term effects on child neurodevelopment: how and why?" *Journal of Child Psychology and Psychiatry* no. 48 (3-4):245-261.

- Turney, Kristin. 2011. "Maternal Depression and Childhood Health Inequalities." *Journal of Health and Social Behavior* no. 52 (3):314-332. doi: 10.1177/0022146511408096.
- Turney, Kristin, and Marcia J. Carlson. 2011. "Multipartnered Fertility and Depression Among Fragile Families." *Journal of Marriage and Family* no. 73 (3):570-587. doi: 10.1111/j.1741-3737.2011.00828.x.
- Wang, L, T. Wu, J.L. Anderson, and J.E. Florence. 2011. "Prevalence and Risk Factors of Maternal Depression During the First Three Years of Child Rearing." *Journal of Women's Health* no. 20 (5):711-8.
- Whitaker, Robert C., Shannon M. Phillips, and Sean M. Orzol. 2006. "Food Insecurity and the Risks of Depression and Anxiety in Mothers and Behavior Problems in their Preschool-Aged Children." *Pediatrics* no. 118 (3):e859-868. doi: 10.1542/peds.2006-0239.
- Wildeman, Christopher, Jason Schnittker, and Kristin Turney. 2012. "Despair by Association? The Mental Health of Mothers with Children by Recently Incarcerated Fathers." *American Sociological Review* no. 77 (2):216-243. doi: 10.1177/0003122411436234.
- Wood, R.G., B. Goesling, and S. Avellar. 2007. The effects of marriage on health: A synthesis of recent research evidence. *Mathematica Policy Research*.
- World Health Organization. 2008. The global burden of disease: 2004 update. edited by World Health Organization. Geneva.

4.6. Tables and Figures

Table 4.1. Descriptive statistics of mothers in fragile families by depression status

	All mothers	Classified as depressed (20% of mothers)	Not classified as depressed (80% of mothers)
Probability of caseness using full CIDI scale***	0.177 (0.351)	0.864 (0.091)	0.001 (0.015)
Any formal support	0.299 (0.458)	0.296 (0.457)	0.300 (0.458)
Any informal support *	0.336 (0.472)	0.308 (0.462)	0.343 (0.475)
Any in-kind support ***	0.440 (0.496)	0.378 (0.485)	0.456 (0.498)
Any father-child contact in past year(s) +	0.747 (0.435)	0.771 (0.420)	0.741 (0.438)
Days of father-child contact in past 30 days **	6.145 (9.444)	5.315 (8.855)	6.358 (9.579)
Mother is white *	0.133 (0.340)	0.155 (0.362)	0.127 (0.333)
Mother is black	0.640 (0.480)	0.631 (0.483)	0.643 (0.479)
Mother is Hispanic	0.203 (0.402)	0.186 (0.389)	0.207 (0.405)
Mother is other race	0.024 (0.154)	0.028 (0.166)	0.023 (0.150)
Mother's highest grade: less than high school *	0.379 (0.485)	0.401 (0.490)	0.374 (0.484)
Mother's highest grade: high school or equivalent	0.356 (0.479)	0.359 (0.480)	0.355 (0.479)
Mother's highest grade: some college	0.236 (0.425)	0.228 (0.420)	0.238 (0.426)
Mother's highest grade: college or better **	0.029 (0.167)	0.013 (0.112)	0.033 (0.178)
Mother is U.S. Born *	0.932 (0.252)	0.950 (0.219)	0.928 (0.259)
Mother currently with another partner	0.479 (0.500)	0.461 (0.499)	0.484 (0.500)
Number of adults in household	1.769 (0.923)	1.745 (0.937)	1.776 (0.920)
Number of children in household	2.519 (1.403)	2.525 (1.439)	2.518 (1.395)
Mother's self-rated health: excellent ***	0.229 (0.420)	0.116 (0.321)	0.258 (0.438)
Mother's self-rated health: very good ***	0.302 (0.459)	0.193 (0.395)	0.331 (0.470)
Mother's self-rated health: good ***	0.302 (0.459)	0.356 (0.479)	0.288 (0.453)
Mother's self-rated health: fair ***	0.143 (0.350)	0.270 (0.444)	0.111 (0.314)
Mother's self-rated health: poor ***	0.023 (0.151)	0.065 (0.247)	0.013 (0.112)
Religious attendance: weekly or more frequently **	0.344 (0.475)	0.312 (0.463)	0.352 (0.478)
Religious attendance: a few times per month *	0.220 (0.414)	0.194 (0.396)	0.226 (0.418)
Religious attendance: a few times per year +	0.217 (0.412)	0.234 (0.424)	0.212 (0.409)
Religious attendance: never or less than yearly ***	0.220 (0.415)	0.260 (0.439)	0.210 (0.407)
Mother's household income ***	2.438 (2.275)	2.078 (1.980)	2.531 (2.336)
Child is male	0.521 (0.500)	0.523 (0.500)	0.521 (0.500)
Mother's age at baseline interview	23.460 (5.392)	23.255 (5.050)	23.513 (5.476)
Child's age (in months) **	69.046 (31.028)	65.756 (31.365)	69.891 (30.889)

Table 4.1.(continued): Descriptive statistics of mothers in fragile families by depression status

	All mothers	Classified as depressed (20% of mothers)	Not classified as depressed (80% of mothers)
Father- multi-partner fertility ^a	0.620	(0.470) 0.632	(0.469) 0.617
Impulsivity Scale ^a ***	2.080	(0.605) 2.258	(0.683) 2.034
WAIS Cognitive Index ^a	6.504	(2.429) 6.627	(2.514) 6.473
Either of mother's parents suffered from depression ^a ***	0.314	(0.438) 0.482	(0.473) 0.270
Child has father's last name ^a	0.744	(0.427) 0.734	(0.432) 0.747
Father is on child's birth certificate ^a +	0.851	(0.347) 0.835	(0.363) 0.856
Father contributed cash during pregnancy ^a +	0.776	(0.414) 0.755	(0.427) 0.781
Father contributed other items during pregnancy ^a	0.938	(0.237) 0.938	(0.237) 0.938
Social support ^a ***	2.497	(0.893) 2.220	(1.055) 2.569
Number of moves since last survey wave***	0.970	(1.175) 1.228	(1.562) 0.904
Mother used one or more drugs***	0.072	(0.259) 0.158	(0.365) 0.050
Number of material hardships ***	1.093	(1.388) 1.790	(1.679) 0.914
Child's father ever incarcerated ^a ***	0.623	(0.478) 0.680	(0.462) 0.609
Number of observations	4665	953	3712

Notes: Means reported with standard deviations in parentheses.

Sample represents mothers of children born out-of-wedlock who are not living with the child's biological father.

Depression classification based on Composite International Diagnostic Interview - Short Form (CIDI-SF) score of three or above.

^a imputed at the mean

significance of t-test of means presented for continuous variables; results from proportion tests presented for dichotomous variables

*** p<0.001

** p<0.01

* p<0.05

+ p<0.10

Table 4.2. Pooled linear probability models of maternal depression in fragile families

	(1)	(2)	(3)
Any formal support	0.010 (0.013)	0.012 (0.013)	0.014 (0.013)
Any informal support	0.009 (0.016)	0.011 (0.016)	0.012 (0.015)
Any in-kind support	-0.051** (0.016)	-0.043** (0.016)	-0.040** (0.016)
Any father-child contact in past year(s)	0.056*** (0.016)	0.046** (0.016)	0.040** (0.015)
Days of father-child contact in past 30 days	-0.001+ (0.001)	-0.001 (0.001)	-0.001 (0.001)
Mother is Black	-0.028 (0.022)	0.003 (0.020)	0.009 (0.020)
Mother is Hispanic	-0.046+ (0.025)	-0.005 (0.024)	0.005 (0.023)
Mother is other race	-0.000 (0.044)	0.015 (0.042)	0.014 (0.041)
Mother's highest grade: high school or equivalent	0.002 (0.016)	0.022 (0.015)	0.022 (0.015)
Mother's highest grade: some college	0.011 (0.019)	0.024 (0.018)	0.015 (0.018)
Mother's highest grade: college or better	-0.038 (0.034)	-0.021 (0.032)	-0.031 (0.032)
Mother is U.S. Born	0.033 (0.026)	0.023 (0.025)	0.018 (0.025)
Mother currently with another partner	-0.017 (0.013)	-0.019 (0.012)	-0.016 (0.012)
Number of adults in household	-0.003 (0.007)	-0.000 (0.007)	0.002 (0.007)
Number of children in household	-0.006 (0.005)	-0.009* (0.005)	-0.010* (0.005)
Mother's self-rated health: excellent	-0.474*** (0.052)	-0.380*** (0.052)	-0.335*** (0.050)
Mother's self-rated health: very good	-0.442*** (0.052)	-0.367*** (0.051)	-0.329*** (0.050)
Mother's self-rated health: good	-0.332*** (0.052)	-0.274*** (0.052)	-0.242*** (0.050)
Mother's self-rated health: fair	-0.191*** (0.054)	-0.165** (0.053)	-0.142** (0.051)
Religious attendance: weekly or more frequently	-0.038* (0.017)	-0.015 (0.017)	-0.018 (0.017)
Religious attendance: a few times per month	-0.045* (0.018)	-0.030+ (0.018)	-0.033+ (0.018)
Religious attendance: a few times per year	-0.021 (0.019)	-0.008 (0.018)	-0.013 (0.018)
Mother's household income	-0.007** (0.003)	-0.003 (0.003)	0.000 (0.003)
Child is male	-0.001 (0.013)	-0.003 (0.013)	-0.002 (0.012)
Mother's age at baseline interview	-0.003*** (0.001)	-0.003* (0.001)	-0.003* (0.001)
Child's age (in months)	0.000 (0.002)	-0.001 (0.002)	-0.002 (0.002)
Mother- multi-partner fertility	0.017 (0.015)	0.013 (0.014)	0.009 (0.014)
Father- multi-partner fertility ^a	0.003 (0.014)	0.003 (0.013)	-0.001 (0.013)
Impulsivity Scale ^a		0.065*** (0.011)	0.064*** (0.011)
WAIS Cognitive Index ^a		0.005* (0.003)	0.005+ (0.003)

Table 4.2. (continued): Pooled linear probability models of maternal depression in fragile families

	(1)	(2)	(3)
Either of mother's parents suffered from depression ^a		0.115*** (0.017)	0.100*** (0.016)
Child has father's last name ^a		0.013 (0.018)	0.014 (0.018)
Father is on child's birth certificate ^a		-0.029 (0.025)	-0.032 (0.024)
Father contributed cash during pregnancy ^a		0.000 (0.019)	-0.002 (0.019)
Father contributed other items during pregnancy		0.028 (0.032)	0.029 (0.031)
Social support ^a		-0.043*** (0.008)	-0.033*** (0.008)
Number of moves since last survey wave		0.019*** (0.006)	0.011* (0.006)
Mother used one or more drugs		0.169*** (0.028)	0.146*** (0.028)
Number of material hardships			0.042*** (0.005)
Child's father ever incarcerated ^a			-0.001 (0.013)
Number of mothers	4653	4653	4653
R-sq	0.098	0.156	0.174

Note: Robust standard errors, clustered at the mother level, in parentheses.

Sample represents mothers of children born out-of-wedlock who are not living with the child's biological father.

Depression classification based on Composite International Diagnostic Interview - Short Form (CIDI-SF) score of three or above.

Survey year fixed effects included in all models.

^a imputed at the mean. Imputation flags included in all models.

*** p<0.001

** p<0.01

* p<0.05

+ p<0.10

Table 4.3. Fixed effects linear probability models of maternal depression in fragile families

	(1)		(2)	
Any formal support	0.012	(0.020)	0.013	(0.019)
Any informal support	0.033+	(0.020)	0.036+	(0.020)
Any in-kind support	-0.035+	(0.021)	-0.036+	(0.021)
Any father-child contact in past year(s)	0.060**	(0.021)	0.056**	(0.021)
Days of father-child contact in past 30 days	-0.002	(0.001)	-0.001	(0.001)
Mother currently with another partner	-0.021	(0.017)	-0.020	(0.016)
Number of adults in household	-0.011	(0.009)	-0.009	(0.009)
Number of children in household	-0.014	(0.008)	-0.013	(0.008)
Mother's self-rated health: excellent	-0.286***	(0.069)	-0.261***	(0.069)
Mother's self-rated health: very good	-0.241***	(0.068)	-0.220**	(0.068)
Mother's self-rated health: good	-0.167*	(0.068)	-0.146*	(0.068)
Mother's self-rated health: fair	-0.106	(0.068)	-0.092	(0.067)
Religious attendance: weekly or more frequently	-0.035	(0.024)	-0.029	(0.024)
Religious attendance: a few times per month	-0.016	(0.023)	-0.014	(0.023)
Religious attendance: a few times per year	0.012	(0.022)	0.013	(0.022)
Mother's household income	0.000	(0.004)	0.003	(0.004)
Child's age (in months)	-0.001	(0.002)	-0.002	(0.002)
Mother- multi-partner fertility	-0.012	(0.029)	-0.013	(0.029)
Father- multi-partner fertility ^a	0.074*	(0.031)	0.070*	(0.031)
Social support ^a	-0.019	(0.012)	-0.015	(0.012)
Number of moves since last survey wave	0.014+	(0.007)	0.011	(0.007)
Mother used one or more drugs	0.085*	(0.040)	0.074+	(0.039)
Number of material hardships			0.037***	(0.007)
Child's father ever incarcerated ^a			0.010	(0.037)
Constant	0.497+	(0.254)	0.509*	(0.259)
Number of mothers	4653		4653	
R-sq	0.061		0.075	

Note: Robust standard errors, clustered at the mother level, in parentheses.

Sample represents mothers of children born out-of-wedlock who are not living with the child's biological father.

Depression classification based on Composite International Diagnostic Interview - Short Form (CIDI-SF) score of three or above.

Survey year fixed effects included in all models.

^a imputed at the mean. Imputation flags included in all models.

*** p<0.001

** p<0.01

* p<0.05

+ p<0.10

Table 4.4. Alternate specification of the role of nonresident father involvement on maternal depression

	(1)	(2)
	Pooled LPM	Fixed effects LPM
Formal support, but no visitation, informal support, or in-kind support	0.019 (0.038)	0.023 (0.027)
In-kind or informal support, but no visitation	0.060 (0.055)	0.036 (0.044)
Visitation, but no informal, formal or in-kind support	0.066** (0.025)	0.039* (0.019)
Formal support and visitation, but no informal or in-kind support	0.074* (0.033)	0.056* (0.024)
Informal support and visitation, but no formal or in-kind support	0.101* (0.041)	0.092** (0.032)
In-kind support and visitation, but no formal support or informal support	0.013 (0.033)	-0.006 (0.023)
Visitation, formal, and informal support, but no in-kind support	0.097 (0.065)	0.092 (0.056)
Visitation, formal and in-kind support, but no informal support	0.030 (0.036)	0.041 (0.025)
Visitation, informal and in-kind support, but no formal support	0.052+ (0.029)	0.011 (0.019)
Visitation, formal, informal, and in-kind support	0.051 (0.036)	-0.010 (0.023)
Mother currently with another partner	-0.019 (0.017)	-0.016 (0.012)
Number of adults in household	-0.009 (0.009)	0.002 (0.007)
Number of children in household	-0.012 (0.008)	-0.010* (0.005)
Mother's self-rated health: excellent	-0.261*** (0.069)	-0.334*** (0.050)
Mother's self-rated health: very good	-0.219** (0.068)	-0.329*** (0.050)
Mother's self-rated health: good	-0.146* (0.068)	-0.242*** (0.050)
Mother's self-rated health: fair	-0.092 (0.067)	-0.142** (0.051)
Religious attendance: weekly or more frequently	-0.026 (0.023)	-0.016 (0.017)
Religious attendance: a few times per month	-0.012 (0.023)	-0.031+ (0.018)
Religious attendance: a few times per year	0.014 (0.022)	-0.011 (0.018)
Mother's household income	0.003 (0.004)	0.000 (0.003)
Child's age (in months)	-0.002 (0.002)	-0.002 (0.002)
Mother- multi-partner fertility	-0.014 (0.029)	0.010 (0.014)
Father- multi-partner fertility ^a	0.072* (0.031)	-0.000 (0.013)
Social support ^a	-0.015 (0.012)	-0.033*** (0.008)
Number of moves since last survey wave	0.012 (0.007)	0.012* (0.006)
Mother used one or more drugs	0.074+ (0.039)	0.146*** (0.028)
Number of material hardships	0.037*** (0.007)	0.042*** (0.005)

Table 4.4. (continued): Alternate specification of the role of nonresident father involvement on maternal depression

	(1)	(2)
	Pooled LPM	Fixed effects LPM
Child's father ever incarcerated ^a		-0.000 (0.013)
Mother is Black	0.013 (0.037)	0.010 (0.020)
Mother is Hispanic		0.006 (0.024)
Mother is other race		0.017 (0.041)
Mother's highest grade: high school or equivalent		0.021 (0.015)
Mother's highest grade: some college		0.014 (0.018)
Mother's highest grade: college or better		-0.034 (0.032)
Mother is U.S. Born		0.018 (0.025)
Child is male		-0.002 (0.012)
Mother's age at baseline interview		-0.003* (0.001)
Impulsivity Scale ^a		0.065*** (0.011)
WAIS Cognitive Index ^a		0.005+ (0.003)
Either of mother's parents suffered from depression ^a		0.101*** (0.016)
Child has father's last name ^a		0.013 (0.018)
Father is on child's birth certificate ^a		-0.031 (0.024)
Father contributed cash during pregnancy ^a		-0.003 (0.018)
Father contributed other items during pregnancy		0.027 (0.031)
Constant	0.489+ (0.260)	0.491* (0.208)
Number of mothers (person-time)	4653	4653
R-sq	0.075	0.175

Note: Robust standard errors, clustered at the mother level, in parentheses.

Reference group for all categories of nonresident father involvement is the group of fathers who do not visit or provide any material (e.g., formal, informal or in-kind) support.

Sample represents mothers of children born out-of-wedlock who are not living with the child's biological father.

Depression classification based on Composite International Diagnostic Interview - Short Form

(CIDI-SF) score of three or above.

Survey year fixed effects included in all models.

^a imputed at the mean. Imputation flags included in all models.

*** p<0.001

** p<0.01

* p<0.05

+ p<0.10

Table 4.5. Heterogeneity of the effects of nonresident involvement on maternal depression by maternal characteristics

	Non-black	Some college or higher	No current partner	No hardships	Father never incarcerated
Any formal support	0.025 (0.022)	0.012 (0.023)	0.014 (0.018)	0.025 (0.016)	-0.007 (0.019)
Any informal support	0.011 (0.027)	0.020 (0.027)	-0.013 (0.021)	0.023 (0.020)	0.003 (0.024)
Any in-kind support	-0.015 (0.026)	-0.034 (0.028)	-0.024 (0.023)	-0.044* (0.020)	-0.018 (0.026)
Dad ever saw child in last year(s)	0.046+ (0.025)	-0.007 (0.030)	0.060** (0.023)	0.038* (0.019)	0.057* (0.027)
Number of visits in last 30 days	-0.001 (0.001)	-0.001 (0.001)	-0.001 (0.001)	0.001 (0.001)	-0.001 (0.001)
Number of observations (nT)	1674	1234	2424	2235	1708

Note: Robust standard errors, clustered at the mother level, in parentheses.

Depression classification based on Composite International Diagnostic Interview - Short Form (CIDI-SF) score of three or above.

All models include survey year fixed effects and the full set of covariates used in Table 2, Column included in all models.

*** p<0.001

** p<0.01

* p<0.05

+ p<0.10

Table 4.6. Nonresident Father Involvement Categories in Fragile Families

	Mean (all mothers)	Frequency (all mothers)
FATHER DID NOT VISIT IN PAST YEAR(S)	0.253	1181
No visitation, informal, formal, or in-kind support	0.196	915
Formal support, but no visitation, informal support, or in-kind support	0.039	184
Informal support, but no visitation, formal support, or in-kind support *	0.005	22
In-kind support, but no visitation, formal support, or in-kind support *	0.004	19
Formal and informal support, but no in-kind support or visitation *	0.000	2
Formal and in-kind support, but no informal support or visitation *	0.003	15
Informal and in-kind support, but no visitation or formal support *	0.004	19
Formal, informal, and in-kind support, but no visitation *	0.001	5
FATHER VISITED IN PAST YEAR(S)	0.747	3484
Visitation, but no informal, formal or in-kind support	0.174	810
Formal support and visitation, but no informal or in-kind support	0.092	429
Informal support and visitation, but no formal or in-kind support	0.041	191
In-kind support and visitation, but no formal support or informal support	0.084	390
Visitation, formal, and informal support, but no in-kind support	0.013	60
Visitation, formal and in-kind support, but no informal support	0.072	337
Visitation, informal and in-kind support, but no formal support	0.194	905
Visitation, formal, informal, and in-kind support	0.078	362
Total (all mothers)	1.000	4665

* Because of the relatively low frequency of these categories, these six groups are condensed into a single group including 82 fathers who provided some form of informal or in-kind support, but did not visit.

Table 4.7. Alternate specifications of maternal depression

	Full analytic sample			Balanced panel		
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Pooled linear probability models predicting maternal depression						
Any formal support	0.011 (0.013)	0.011 (0.013)	0.014 (0.013)	0.013 (0.019)	0.012 (0.018)	0.017 (0.019)
Any informal support		-0.001 (0.014)	0.012 (0.015)		-0.014 (0.022)	0.008 (0.023)
Any in-kind support			-0.040** (0.016)			-0.069** (0.022)
Any father-child contact in past year(s)	0.031* (0.015)	0.031* (0.015)	0.040** (0.015)	0.044* (0.021)	0.046* (0.021)	0.059** (0.022)
Days of father-child contact in past 30 days	-0.001* (0.001)	-0.001+ (0.001)	-0.001 (0.001)	-0.001 (0.001)	-0.001 (0.001)	0.000 (0.001)
Panel B: Fixed effects linear probability model predicting maternal depression						
Any formal support	0.008 (0.019)	0.010 (0.019)	0.013 (0.019)	-0.001 (0.024)	0.000 (0.024)	0.005 (0.025)
Any informal support		0.026 (0.018)	0.036+ (0.020)		0.018 (0.023)	0.029 (0.025)
Any in-kind support			-0.036+ (0.021)		-0.046+ (0.026)	
Any father-child contact in past year(s)	0.055** (0.021)	0.051* (0.021)	0.056** (0.021)	0.043+ (0.025)	0.040 (0.026)	0.047+ (0.026)
Days of father-child contact in past 30 days	-0.002+ (0.001)	-0.002* (0.001)	-0.001 (0.001)	-0.001 (0.001)	-0.001 (0.001)	-0.001 (0.001)
Number of observations (nT)	4653	4653	4653	2212	2212	2212

Note: Robust standard errors, clustered at the mother level, in parentheses.

Sample represents mothers of children born out-of-wedlock who are not living with the child's biological father.

Depression classification based on Composite International Diagnostic Interview - Short Form (CIDI-SF) score of three or above.

Table 48. Robustness checks

	Probability of depression caseness (1)	List-wise deletion of missing values (2)	Logit model (3)	Logit model: add interaction terms (4)
Panel A: Pooled models predicting maternal depression				
Any formal support	0.013 (0.011)	0.016 (0.015)	0.093 (0.095)	0.152 (0.220)
Any informal support	0.010 (0.013)	0.019 (0.017)	0.099 (0.115)	0.213 (0.135)
Any in-kind support	-0.035** (0.014)	-0.053** (0.018)	-0.280* (0.112)	-0.369** (0.142)
Any father-child contact in past year(s)	0.037** (0.013)	0.048** (0.018)	0.304** (0.114)	0.298* (0.127)
Days of father-child contact in past 30 days	-0.001 (0.001)	-0.001 (0.001)	-0.007 (0.006)	-0.005 (0.007)
Any informal support * Any formal support				-0.348 (0.249)
Any in-kind support * Any formal support				0.232 (0.230)
Any father-child contact in past year(s) * Any formal support				-0.015 (0.265)
Days of father-child contact in past 30 days * Any formal support				-0.007 (0.013)
Panel B: Fixed effects models predicting maternal depression				
Any formal support	0.010 (0.024)	0.010 (0.027)	0.155 (0.213)	0.381 (0.427)
Any informal support	0.028 (0.024)	0.026 (0.026)	0.348 (0.215)	0.395 (0.270)

Table 4.8 (continued): Robustness checks

	Probability of depression caseness (1)	List-wise deletion of missing values (2)	Logit model (3)	Logit model: add interaction terms (4)
Any in-kind support	-0.031 (0.025)		-0.355 (0.218)	-0.363 (0.275)
Any father-child contact in past year(s)	0.051+ (0.026)	0.051+ (0.030)	0.518* (0.225)	0.622* (0.264)
Days of father-child contact in past 30 days	-0.001 (0.001)	-0.002 (0.001)	-0.028* (0.013)	-0.036* (0.015)
Any informal support * Any formal support				
Any in-kind support * Any formal support				-0.142 (0.425)
Any father-child contact in past year(s) * Any formal support				0.050 (0.406)
Days of father-child contact in past 30 days * Any formal support				-0.401 (0.475)
				0.024 (0.024)
Number of mothers	4653	3446/4207	4653/1073	4653/1073

Note: Robust standard errors, clustered at the mother level, in parentheses.

Sample represents mothers of children born out-of-wedlock who are not living with the child's biological father.

Dependent variable is a binary indicator of maternal depression in all specifications except those reported in Column 1

In some specifications, the number of mothers is different between the two panels. In these cases, the numbers on the right represent the number of observations in the specifications shown in Panel B.

Depression classification based on Composite International Diagnostic Interview - Short Form (CIDI-SF) score of three or above.

*** p<0.001

** p<0.01

* p<0.05

+ p<0.10

Table 4.9. Comparison of maternal and paternal reports of father involvement

	Mother report	Father report	Difference
Any formal support	0.506	0.666	-0.16
Any informal support	0.345	0.546	-0.201
Any in-kind support	0.499	0.866	-0.367
Any father-child contact in past year(s)	0.827	0.898	-0.071
Days of father-child contact in past 30 days	6.015	8.775	-2.76

Note: Included in this table are 815 mothers and fathers who both completed the 9-year follow-up survey and had non-missing values for all measures of nonresident father involvement.

Table 4.10. Characteristics of mothers by their reporting of nonresident father involvement

	Not missing measures of nonresident father involvement	Missing any measures of nonresident father involvement
Mother classified as depressed	0.204	0.201
Mother is white *	0.133	0.094
Mother is black	0.640	0.613
Mother is Hispanic **	0.203	0.262
Mother is other race	0.024	0.031
Mother's highest grade: less than high school	0.379	0.416
Mother's highest grade: high school or equivalent	0.356	0.340
Mother's highest grade: some college	0.236	0.213
Mother's highest grade: college or better	0.029	0.031
Mother is U.S. Born ***	0.932	0.879
Mother currently with another partner	0.479	0.461
Number of adults in household	1.769	1.766
Number of children in household	2.519	2.535
Mother's self-rated health: excellent	0.229	0.262
Mother's self-rated health: very good	0.302	0.300
Mother's self-rated health: good	0.302	0.280
Mother's self-rated health: fair	0.143	0.132
Mother's self-rated health: poor	0.023	0.027
Religious attendance: weekly or more frequently	0.344	0.345
Religious attendance: a few times per month *	0.220	0.262
Religious attendance: a few times per year ***	0.217	0.183
Religious attendance: never or less than yearly	0.220	0.210
Mother's household income	2.438	2.415
Child is male	0.521	0.523
Mother's age at baseline interview *	23.460	23.931
Child's age (in months) +	69.046	71.418
Mother- multi-partner fertility	0.596	0.615
Father- multi-partner fertility ^a +	0.620	0.576
Impulsivity Scale ^a	2.080	2.063

Table 4.10. (continued) Characteristics of mothers by their reporting of nonresident father involvement

	Not missing measures of nonresident father involvement	Missing any measures of nonresident father involvement
WAIS Cognitive Index ^a	6.504 (2.429)	6.333 (2.504)
Either of mother's parents suffered from depression ^a +	0.314 (0.438)	0.340 (0.441)
Child has father's last name ^a +	0.744 (0.427)	0.702 (0.445)
Father is on child's birth certificate ^a ***	0.851 (0.347)	0.759 (0.420)
Father contributed cash during pregnancy ^a ***	0.776 (0.414)	0.700 (0.453)
Father contributed other items during pregnancy ^a ***	0.938 (0.237)	0.884 (0.314)
Social support ^a	2.497 (0.893)	2.476 (0.988)
Number of moves since last survey wave	0.970 (1.175)	1.083 (1.260)
Mother used one or more drugs	0.072 (0.259)	0.098 (0.298)
Number of material hardships	1.093 (1.388)	1.056 (1.462)
Child's father ever incarcerated ^a	0.623 (0.478)	0.629 (0.458)
Number of mothers	4665	447

^a imputed at the mean

Significance of t-test of means presented for continuous variables; results from proportion tests presented for dichotomous variables

*** p<0.001

** p<0.01

* p<0.05

+ p<0.10

Figure 4.1. Descriptive means of depression and nonresident father involvement in fragile families

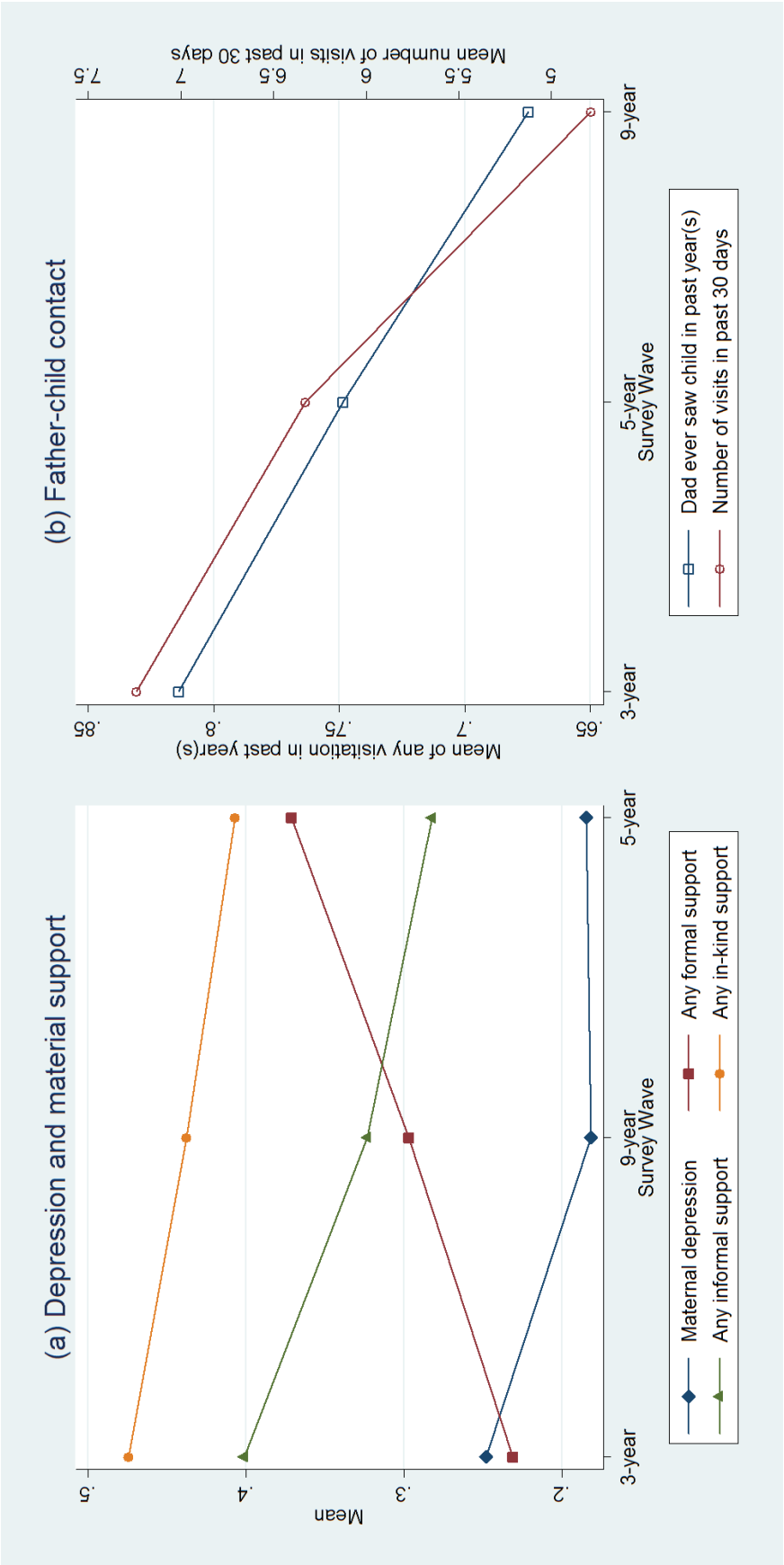
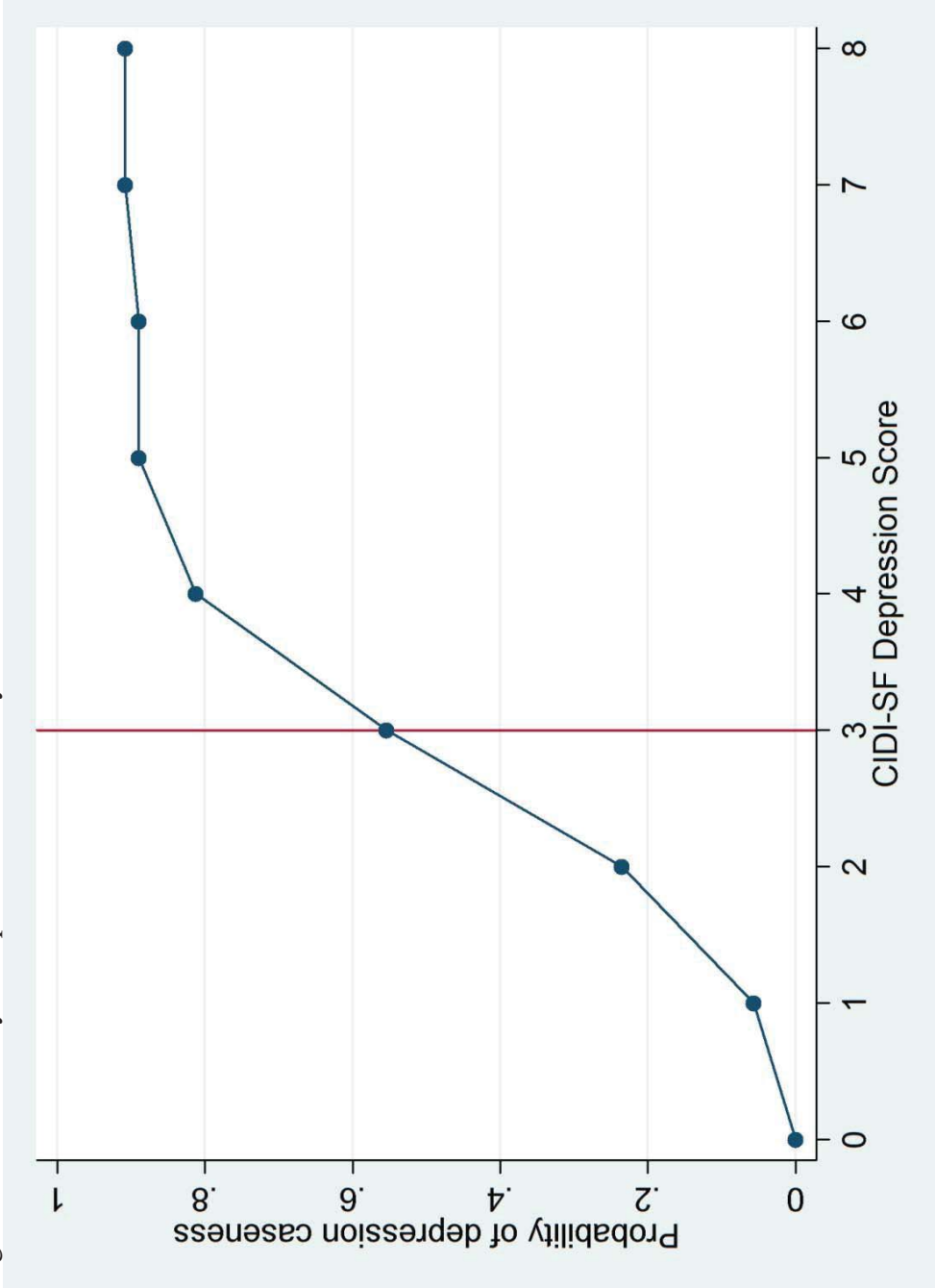


Figure 4.2. Probability of depression caseness by CIDI-SF score



Note: The probability of depression caseness is the probability that given a particular score on the CIDI-SF, that an individual would be classified as being depressed on the full CIDI interview. The red line denotes the CIDI-SF cut-point of 3 or above used to classify depression cases in this study.

Chapter 5

Conclusions

This dissertation explored several important correlates in the health production function of individuals, at different stages of the life cycle, using three different longitudinal samples. Two of the dissertation Chapters (Chapters 3 and 4) focused on factors related the family as an important factor in determining health outcomes and behaviors—family structure (Chapter 3), and nonresident father involvement (Chapter 4). Chapter 2 focuses the role of diabetes: a disease with potentially devastating health consequences, on the production of health outcomes and behaviors.

Chapter 2 used data from the Health and Retirement Study (HRS) to investigate the short- and long- term effects of a diabetes diagnosis on behaviors that have the potential to reduce blood sugar levels and/or the risk for developing diabetic complications over time. These activities include smoking, drinking, frequent exercise, and weight loss. Analysis using dynamic, population average models revealed that individuals make sizable changes in all four activities initially in response to a diabetes diagnosis. Individuals tend to lower cigarette and alcohol consumption subsequent to diagnosis, but exercise levels tend to fall, and weight tends to rise two or more years after diagnosis. The response trajectory varies by age and race, and is somewhat different for those diagnosed with and without medication (especially with regard to alcohol consumption). The policy implications for these results are substantial. In type II diabetes, the disease often follows an insidious course which (unfortunately) can result in the gradual and continual deterioration of organs, including the eyes (retinopathy), kidneys (nephropathy), and the autonomic and peripheral nervous system (neuropathy). Importantly,

elevated blood sugars and organ damage can precede the diagnosis for years, even decades. In fact, it is sometimes the organ damage itself—tingling feet or blurry vision that can lead an individual to seek medical counsel that eventually leads to the diabetes diagnosis. Given that I find that individuals respond to the diagnosis more initially than subsequently, a rapid diagnosis can yield lower rates of complications, in part through health behavior change (which can be precipitated by the diagnostic information). It also suggests the need for intensive management of weight and exercise. As the risk for complications increases as the duration with clinical diabetes endures (i.e., prolonged exposure to high blood sugar), the results of the study suggest that exercise and weight management can elude many patients when they need it the most.

Chapters 3 and 4 consider the family as a producer of health. Family dynamics (especially those concerning the child's biological father) can be important correlates of health and health behaviors. Chapter 3 focuses on the role of biological father absence and other (i.e., step-father or cohabiting-) male entrance on the evolution of smoking, physical and mental health outcomes from adolescence to young adulthood using data from the National Longitudinal Study of Adolescent Health (Add Health). Physical health outcomes included self-reported health status and overweight or obese status (corresponding to a body mass index of 25 or above). The analyses exploited a rich parental questionnaire to develop the most accurate possible picture of family structure from birth through age 15. The baseline analyses revealed a strong association between paternal absence and adolescent health outcomes, which was somewhat stronger in girls. Longitudinal analyses using discrete-time hazard models revealed that boys who reported ever smoking by Wave I (when they were adolescents) tended to be less likely to subsequently quit smoking as young adults. Likewise, boys who reported "good" or lower health at Wave I tended to be less likely to subsequently report "very good" or "excellent"

health in young adulthood. The entrance of step- or cohabiting- fathers diminished the impact of biological-father absence in some cases, and substantially so in the case of smoking.

In addition to the policy implications related to the short- and long- term effects of family structure on child health , Chapter 3 has important practical implications for the wider research community that consider how of family structure influences child outcomes in general. First, we discuss some of the limitations with using only the mother's marital history to glean childhood family structure. Specifically, this method requires the assumption that the spouse of the mother at the time of the child's birth is the child's biological father. This classifies any child born out-of-wedlock as having their biological father never present in the household throughout childhood and adolescence. Furthermore, we highlight the high rate of of step- and cohabiting- father entrance among children who faced the absence of their biological father (over half of these children). Including the entrance of these other males substantially diminishes the estimates for paternal absence in some cases. As many studies completely ignore the presence of step- or cohabiting- fathers in the household, these studies may be overestimating the effects of paternal absence on some child outcomes. Finally, we explore the dynamics of these outcomes over time, and find that paternal absence may have significant (albeit small) long-term effects on young adults.

Chapter 4 explored the role of specific non-resident paternal contributions to the household – child support, informal support, in-kind (non-cash), and visitation on maternal depression. While child support and visitation have been well-studied in the context of the (especially cognitive-related) outcomes of affected children, the indirect effects on mothers has remained understudied. This Chapter used data from the Fragile Families and Child Wellbeing Study (FFCWS), and estimated pooled and fixed-effect linear probability models to clarify the

relationship between several distinct (though correlated) forms of nonresident father involvement on maternal depression. The results suggest that while in-kind support has a protective effect on maternal depression, visitation, especially when done so infrequently, can actually raise depression levels. The policy implications for this paper suggest that for these disadvantaged families living in urban areas, more involvement is not necessarily better for the household. In this sense, visitation should not be “forced” in these households. Further, the encouragement of informal means of material support can be beneficial for some families as well, including those that do not involve cash transfers.