Distribution Agreement

Signature:

______________________________ _______________

Handie Peng Date

Essays on the Economics of Risky Health Behaviors

By

Handie Peng
Doctor of Philosophy

Economics

Hugo M. Mialon, Ph.D.
Advisor

Andrew M. Francis, Ph.D.
Committee Member

David E. Frisvold, Ph.D.
Committee Member

Matthew J. Higgins, Ph.D.
Committee Member

Accepted:

Lisa A. Tedesco, Ph.D.
Dean of the James T. Laney School of Graduate Studies

Date

Essays on the Economics of Risky Health Behaviors

By

Handie Peng
B.A., Sichuan University, 2004
M.A., Emory University, 2011

Advisor: Hugo M. Mialon, Ph.D.

An abstract of
A dissertation submitted to the Faculty of the
James T. Laney School of Graduate Studies of Emory University
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy
in Economics
2013

Abstract

Essays on the Economics of Risky Health Behaviors
By Handie Peng

The first chapter *The Effects of Maternal Employment on the Initiation of Adolescent Risky Behaviors* studies the effects of maternal employment on the initiation of adolescent risky behaviors. We find that maternal employment after the first few years of a child's life has a positive and significant impact on the initiation of alcohol, cigarette and marijuana consumption as well as sexual activity during adolescence. These effects remain significant when the average maternal employment over a child's life is used as the independent variable. Subgroup analyses indicate that the positive and significant results appear to be driven mostly by households with higher socioeconomic status.

The second chapter *In Sickness and in Health: Same-Sex Marriage Laws and Sexually Transmitted Infections* analyzes the relationship between same-sex marriage laws and sexually transmitted infections in the United States. Our findings may be summarized as follows. Laws banning same-sex marriage are unrelated to gonorrhea rates, which are a proxy for risky heterosexual behavior. They are positively associated with syphilis rates, which are a proxy for risky homosexual behavior. However, these estimates are smaller and less statistically significant when we exclude California. Also, laws permitting same-sex marriage are unrelated to gonorrhea or syphilis, but variation in these laws is insufficient to yield precise estimates.

The third chapter *Weight and Labor Market Outcomes: Understanding the Role of Teenage Weight Status* looks at whether the relationship between weight and labor market outcomes differs for individuals with different teenage weight statuses. The results indicate that there is indeed a heterogeneous effect: specifically, being overweight is associated with a wage *premium* (penalty) for adult white males who were healthy weight teens or HWT (obese or overweight teens or OOT). We also look at how much of the wage and employment differentials between HWT and OOT are explained by observed characteristics. We find that for white males and non-white females who were OOT, their lower wages are explained by their lower education. White females who were OOT tend to have a lower employment probability, and this differential is explained by their higher current BMI, lower education, and less work experience.

Essays on the Economics of Risky Health Behaviors

By

Handie Peng
B.A., Sichuan University, 2004
M.A., Emory University, 2011

Advisor: Hugo M. Mialon, Ph.D.

A dissertation submitted to the Faculty of the
James T. Laney School of Graduate Studies of Emory University
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy
in Economics
2013

Acknowledgements

First, I want to thank my advisor Dr. Hugo Mialon for his kind advice and guidance throughout my Ph.D. experience. Second, I would like to thank my committee members and mentors Dr. Andrew Francis, Dr. David Frisvold and Dr. Matthew Higgins for their support. Finally, I wish to extend sincere gratitude to my parents Zulin Peng and Yongsu Liao and my husband Stephen M.I. Kunen for their love and encouragement.

Table of Contents

List of Tables

List of Figures

Chapter 1

The Effects of Maternal Employment on the Initiation of Adolescent Risky Behaviors

Handie Peng*

This paper studies the effects of maternal employment on the initiation of adolescent risky behaviors[1] using data from the National Longitudinal Survey of Youths 1979 (NLSY79) and its children and young adult supplement. We find that maternal employment after the first few years of a child's life has a positive and significant impact on the initiation of alcohol, cigarette and marijuana consumption as well as sexual activity during adolescence. These effects remain significant when the average maternal employment over a child's life is used as the independent variable. Subgroup analyses indicate that the positive and significant results appear to be driven mostly by households with higher socioeconomic status.

JEL codes: I12, J22.
Keywords: Maternal Employment, Initiation of Adolescent Risky Behaviors.

---

* Handie Peng, Department of Economics, Emory University, Atlanta, GA 30322-2240. Email: hpeng4@emory.edu. The author is extremely grateful to Andrew Francis, David Frisvold, Matthew Higgins, Sara Markowitz, and Hugo Mialon for helpful comments, and to Alison Aughinbaugh, Kristin Butcher, and Adam Looney for data assistance.

[1] Adolescent in this paper is loosely defined. At times, risky behaviors are initiated when someone has not yet reached adolescence (e.g. 10 years of age) as normally defined. That is, adolescent risky behaviors and child risky behaviors are used interchangeably.

I. Introduction

Adolescent risky behaviors, including alcohol, tobacco and drug use, criminal activities and risky sexual behaviors, contribute to the leading causes of morbidity and mortality among youth and adults (CDC, 2009). As a result, researchers and policy makers have been interested in detecting factors that contribute to these behaviors. One such factor that has been continually identified in academic research and the media is maternal employment. For example, According to the Washington Post, "two-thirds of the people surveyed said that although it may be necessary for a mother to work, it would be better for her family if she could stay home and care for the house and children (Grimsley and Melton, 1998)."[2]

An analysis in the tradition of Becker (1976) can be used to motivate why maternal employment might have an impact on adolescent risky behaviors. A working mother might knowingly "sacrifice" certain aspects of a child's development in order to maximize the family utility as a whole, or she might unintentionally do so because market goods are not perfect substitutes for maternal care. Empirical evidence is mixed. On the one hand, Aughinbaugh and Gittleman (2004) do not find that maternal employment – whether early in a child's life or during adolescence – affects the likelihood of participation in risky behaviors. On the other hand, Ruhm (2004a) finds that maternal labor supply has deleterious effects on risky behaviors such as smoking and drinking, although the effects are not statistically significant due to small incidence rates.

[2] According to New York Times, "... rather than relying on them (younger adolescents) to make reasoned choices (regarding risky behaviors) or to learn from the school of hard knocks, a better approach is to supervise them (Brody 2007)", pointing to the importance of parental supervision in reducing adolescent risk-taking. ABC News also reported, "[t]hose (teens) who spent seven hours or less per week with their parents were twice as likely to use alcohol and twice as likely to say they expected to try drugs, compared with teens who spent 21 hours or more per week with their parents (Francis 2011)."

The purpose of this paper is to build on previous research and determine whether maternal employment affects the *initiation* of adolescent risky behaviors. The adolescent risky behaviors studied here include consumption of alcohol, cigarette and marijuana, sexual activity, and criminal activity. Additionally, this paper explores how these effects differ at various stages of a child's life and along socioeconomic dimensions.

There are several reasons why we choose to focus on risky behavior *initiation*. Medical research finds that early initiators of certain risky behaviors are more likely to be involved in other risky behaviors than late initiators. For example, Stanton et al. (2001) conclude that "early initiators of sex were significantly more likely to report involvement in substance use and drug-delivery/sales than were late initiators" among urban, low-income African-American adolescents. Therefore, understanding what contributes to the initiation of certain risky behaviors would not only help prevent or postpone the initiation of these behaviors, but also reduce the likelihood of engaging in other risky behaviors by the adolescents. Moreover, economists find that there is a gateway effect of using certain substances, such as cigarette use leading to marijuana use and marijuana use leading to hard drug use (Beenstock and Rahav, 2002). Given that the gateway effect could lead to more problematic and hence more costly behaviors, preventing the less problematic behaviors from happening in the first place might prove the most cost-effective approach. Lastly, there will be no incidence without initiation. Similar to the previous argument, preventing initiation might prove to be more cost effective than terminating incidence in general.

This paper extends previous literature in a number of important ways. Firstly, although previous research has looked at the relationship between maternal employment and the incidence of adolescent risky behaviors, this is the first study to investigate the association between maternal employment and the *initiation* of adolescent risky behaviors. The reasons that risky behavior initiation is as important as, if not more important than, risky behavior incidence are stated above. Second, this paper exploits variations in maternal health limits that restrict the kind or amount of work they could do as instruments. The instruments employed in prior literature rely on variations in state economic conditions, which usually do not have enough statistical power (e.g., Aughinbaugh and Gittleman, 2003). In comparison, our instruments vary at the individual-level. Third, we use alternative measures of maternal employment. Besides the number of maternal working hours, the percent time employed and employment status (full-time, part-time, unemployed and out of the labor force) of the mother are also utilized. These alternative measures capture the effects at the extensive margin. Fourth, this paper not only examines the effects of maternal employment in a child's early years but also in subsequent years. Aughinbaugh and Gittleman (2004) use maternal employment in a child's early years and adolescent years but omit the years in between. However, our results suggest that the years in between might be crucial. Finally, we investigate how the effects of maternal employment on the initiation of adolescent risky behaviors differ along socioeconomic status (SES) dimensions.

We find that maternal employment after the first few years of a child's life has a positive and statistically significant impact on the probability of an adolescent initiating alcohol, cigarette and marijuana consumption, as well as sexual activity. The results suggest that the years in middle childhood and adolescence are important in shaping decisions on risk-taking. When average maternal employment over a child's life is used, the effects remain robust, suggesting that the years following the first few years of a child's life are driving the results. Subgroup analyses indicate that this positive relationship appears to be driven mostly by households with higher SES.

The rest of the paper is organized as follows. Section II reviews related literature. Section III proposes hypotheses about the relationship between maternal employment and adolescent risky behavior initiation. Section IV describes the data. Section V presents the methodology. Section VI displays the summary statistics and empirical results. Section VII performs robustness checks, and Section VIII concludes.

II. Related Literature

A. Early Childhood Experiences

There are recent and related health and labor literatures focusing on the early experiences of children. These literatures find not only a short- and medium-term, but also a long-lasting, effect of early childhood experiences on their subsequent outcomes.[3] With respect to short- and medium-term outcomes, studies on the effects of maternity leave (Ruhm, 2000; Tanaka, 2005; Rossin, 2011) provide a good example of how maternal care in a child's early years is beneficial to their development. For instance, Rossin (2011) finds that maternity leave for college-educated and married mothers has a positive impact on a range of infant health outcomes. Several other studies also find that early maternal employment has adverse effects on cognitive and behavioral outcomes of children (Waldfogel et al., 2002; Brooks-Gunn et al., 2002; Ruhm, 2004b).

Over the long term, Chetty et al. (2011) find that the quality of kindergarten

[3] These outcomes include health, cognitive development, behavioral problems, educational achievement, criminal activity and labor market performance.

education is predictive of future labor market success. Early intervention programs such as Head Start are found to have long-term effects on health, education, criminal activity and labor market outcomes of low SES children (Garces et al., 2002; Anderson et al., 2010). Some studies even attribute middle-age labor market outcomes to what happened when the individual was in their mother's uterus (e.g., Almond, 2006).

The current trend of focusing on early childhood years is partly spurred by the welfare reform under the Personal Responsibility and Work Opportunity Reconciliation Act (PRWORA) of 1996 and the concurrent expansions of Earned Income Tax Credit (EITC) and Medicaid. These reforms have provided greater incentives for low SES women with young children to enter or return to the work force.[4] Such policy changes have motivated a wide range of research investigating the effects of maternal employment on childhood development, especially cognitive achievement (Waldfogel et al., 2002; Gennetian et al., 2002), health (Gordon et al., 2007; Gennetian et al., 2010), and obesity (Anderson et al., 2003).

Besides the aforementioned childhood development measures, adolescent risky behaviors could also be affected by a child's early experiences. Aughinbaugh and Gittleman (2004) argue that in the first 3 years a child might develop important emotional

---

[4] In the PRWORA, for example, the work requirements state that "(b)y 2002, at least 50 percent of all recipient families and 90 percent of two-parent families were required to be working or in work preparation programs" (Blank 2002, p. 1106). "The combination of increased minimum wages and increased EITC subsidies meant that the real earnings plus wage subsidy (in 2000 dollars) received by a woman with one child working full time at the minimum wage rose from $10,568 in 1989 to $12,653 in 2000, a 19.7 percent increase. For a similar woman with two or more children, real earnings and subsidies rose from $10,568 in 1989 to $14,188 in 2000, a 34.3 percent increase. These changes should have greatly increased the work incentives for low-wage single mothers with children" (Blank 2002, p.1108). With respect to Medicaid expansion, "from the mid-1980s on, access to public health insurance became increasingly delinked from participation in cash public assistance programs. By 1999, all children in families whose income was below 100 percent of the poverty line were eligible for Medicaid", … "Because many eligible children did not appear to be accessing Medicaid, in 1997 Congress funded a $24 billion, five-year program known as the Children's Health Insurance Program (CHIP), providing incentives and funding to states to expand health-care usage and health-insurance access among low-income children" (Blank 2002, p.1108).

relationships, attain a basic sense of self and others, and have a variety of learning experiences, and therefore, these early years could impact adolescent risky behaviors. To buttress their argument, the authors cite movie director and child advocate, Rob Reiner:

> "If we want to have a real significant impact, not only on children's success in school and later on in life, healthy relationships, but also an impact on reduction in crime, teen pregnancy, drug abuse, child abuse, welfare, homelessness and a variety of other social ills, we are going to have to address the first 3 years of life. There is no getting around it. All roads lead to Rome (Bruer 1999, p.8)."

B. Middle Childhood and Adolescent Experiences

Children's experiences after the first few years of life could also be important in shaping their decisions on risk-taking. A number of studies from developmental psychology provide supportive evidence (Erikson, 1968; Steinberg and Silverberg, 1986; Barnes and Farrell, 1992; Eccles, 1999; Moore and Chase-Lansdale, 2001; Chase-Lansdale and Pittman, 2002; Coley et al., 2008; Wang et al., 2011). For example, Steinberg and Silverberg (1986) find that susceptibility to peer influence follows an inverted U-shaped function, increasing between childhood and early adolescence, peaking sometime around age 14, and declining during the early high school years. Without maternal time devotion in these critical years, a child or adolescent might find it hard to rise above such peer influence.

Chase-Lansdale and Pittman (2002) have identified dimensions of parenting that are linked to childhood outcomes. These include warmth and responsiveness; control and discipline; cognitive stimulation; modeling of attitudes, values, and behaviors; gate-keeping; and family routines and traditions. All these dimensions can potentially play an important role in shaping children's or young adolescents' decisions on risk-taking and are likely to be affected by maternal employment. Specifically, if a mother's job is low-paid and/or stressful, maternal employment may serve to reduce her warmth and responsiveness. If work makes it hard to outline specific rules and expectations and enforce them consistently, it may hinder her ability to control and discipline her child. If work reduces her time to be verbally engaging or provide the child with cognitively stimulating toys, activities, and interactions, it may render her less capable of stimulating the child's cognitive development. If a mother appears aggressive and stressed due to pressure from work, she may set a negative model for her child. If work reduces her involvement in influencing her child's social network and outside activities, it may negatively affect her role of gate-keeping. Finally, work may interrupt family routines and traditions.

In contrast, maternal working could also be beneficial, if the child sees her working mother as a role model, and/or obtains independence. For example, Montemayor and Clayton (1983) argue:

> "… adolescents with working mothers may develop a greater degree of autonomy and adult maturity than those with nonworking mothers. Also, employed mothers may be models of feminine competence for their adolescents, who might develop less stereotyped and traditional sex-role concepts. Finally, employed mothers may have less stressful relationship with their adolescents because they would not be as fully invested in child-rearing and therefore could more easily relinquish maternal control than could full-time homemakers."

C. Maternal Employment

Aughinbaugh and Gittleman (2003 & 2004) and Ruhm (2004a) are the papers most closely related to this paper. Aughinbaugh and Gittleman (2003 & 2004) analyze the impact of maternal employment during a child's first 3 years and during adolescence (3 calendar years preceding the interview) on his or her decisions to engage in a range of risky activities: smoking cigarettes, drinking alcohol, using marijuana and other drugs, engaging in sex and committing crimes. The authors use data from the National Longitudinal Survey of Youths 79 (NLSY79) and NLSY79 YA, the young adult supplement, which interviews the children of the NLSY79 female respondents when they are at least 15 years old. The authors conclude that maternal employment – either early in a child's life or during adolescence – does not seem to affect the likelihood of participation in risky behaviors. However, they are quick to note that insufficient statistical power makes it hard to rule out potential effects. Moreover, the authors do not specifically model the years in between the first 3 years and the most recent 3 years of life, which could be an important stage for child development.

Focusing on children 10 to 11 years old and using data from NLSY79 and NLSY79 Children, Ruhm (2004a) studies how maternal employment is related to a child's cognitive development, obesity, behavior problems and risky behaviors such as smoking and drinking. In order to see how the effects differ along SES dimensions, Ruhm (2004a) carries out subgroup analyses for "advantaged" (top 50% SES) and "disadvantaged" (bottom 50% SES) children. Ruhm finds that the effects differ greatly by SES. For low SES youths, maternal employment does not affect their development negatively, and even affects it positively when a mother is working less than half-time. In contrast, maternal labor supply is associated with substantial negative consequences for high SES youths, even when maternal working hours are limited. A possible drawback of Ruhm (2004a) is that the causal effects might be unidentified, although the author uses an extensive list of controls and maternal fixed-effect and propensity score models in an effort to deal with the omitted variable problem. Even if these methods control for the omitted variable bias, endogeneity might still be a concern due to reverse causality: not only could maternal employment decisions affect adolescent risky behaviors, the latter could also affect the former.

Complementing these papers is research that investigates the effects of maternal employment on a child's cognitive development (Blau and Grossberg, 1992; Waldfogel et al., 2002; Bernal, 2008), health (Gordon et al., 2007; Gennetian et al., 2010; Morrill, 2011), and obesity (Anderson et al., 2003; Cawley and Liu, 2007; Chia, 2008; Ruhm, 2008; Fertig et al., 2009; Liu et al., 2009). The abundance of literature makes clear the importance of understanding the effects of maternal employment on children and adolescents.

III. Hypotheses

Becker (1976) provides an economic framework that can help motivate why maternal employment may have an impact on adolescent risky behaviors. Child rearing can be seen as a non-market economic activity that requires optimization under financial and time constraints. When maximizing the family utility as a whole, or when maximizing her own utility, which incorporates her child's utility, a working mother faces the trade-off between working and taking care of her child. Working serves as a way to provide the financial means to afford market goods, which is utility-enhancing. Taking care of her child requires spending time to provide high-quality food, help with school work, and monitor after school activities. All of these activities might serve to increase a child's utility through better physical health, academic achievement, mental health, etc., which in turn increases a mother's utility. Time spent working means time lost with a child, and vice versa. However, it does not mean that a mother who works more would undoubtedly have children with worse outcomes. The mother could purchase non-home-cooked high quality food, services from nannies or tutors, and extracurricular activities monitored by professionals, all of which can substitute for the care she herself could otherwise provide. To the extent that these market goods are not perfect substitutes for maternal care, the effects of maternal employment on adolescent risky behaviors are ambiguous.

In total, there are three relevant effects at play. The first one is the time effect, which refers to the effect caused by reduced time investment in a child that serves to increase his or her probability of initiating risk-taking. The two channels through which this effect takes place are lack of direct monitoring and/or internalizing moral values. While direct monitoring is a straightforward channel, internalizing moral values means that the mother has spent enough time during a certain stage of a child's life, so that the child obtains a sense of morality that would guide him or her in the future even when the mother is absent. Therefore, when direct monitoring is not available, the internalized moral values may still help to reduce the probability of initiating risk-taking by the child/adolescent.

The second effect at play is the income effect, which refers to the effect caused by higher income that serves to increase or reduce the probability of initiating risk-taking by a child/adolescent. The mechanism through which income increases the probability of initiating risk-taking is that a child/adolescent from a wealthier family can better afford risky activities such as drinking and smoking. The two channels through which income decreases the probability of initiating risk-taking are purchases of market goods that are substitutes for maternal care, and relief of financial strain that might affect a child's/adolescent's psychological wellbeing, the lack of which is a risk factor for risky behaviors.

The third effect is the role model effect, which refers to the effect brought about by the fact that a working mother is seen as a role model by her child, which can potentially reduce the probability that he or she initiates risky behaviors. There is a substantial literature in psychology studying the effects of role models or the lack thereof. For example, Harris et al. (2002, p. 1010) argue, "[t]he economic and social roles that parents assume serve as models and influence their children's attitudes and expectations for their own futures." Yancey et al. (2002) conclude that role model selection is associated with protective psychosocial characteristics, and having a role model was associated with decreased substance use. Additionally, Bryant and Zimmerman (2003) find that the presence of female role models was associated with adolescent psychological well-being and lower levels of distress. Another line of psychology literature points to the negative effects of parental unemployment. For example, maternal unemployment in single mother families is found to negatively affect adolescent socioemotional functioning through maternal psychological functioning, and in turn, parenting behavior and mother-child relations (McLoyd et al., 1994).

For two reasons, we hypothesize that the role model effect from maternal working is larger for households with lower SES. First, employment is more difficult to obtain and secure for women with low SES. Second, children from high SES families have a lower probability of initiating most of the risky behaviors studied here.[5] This means that the

[5] According to the summary statistics by socioeconomic status (not shown here): children whose mothers have high school education or less are significantly more likely than children whose mothers have some college education or more to initiate smoking cigarettes, using marijuana, being convicted of crime, having sex and becoming pregnant; children whose mothers are single are significantly more likely than children whose mothers are married to initiate drinking alcohol, drinking at least once a month, smoking cigarettes, using marijuana, being convicted of crime, having sex and becoming pregnant; children whose average family income is in the lower 50th percentile are significantly more likely than children whose average

benchmark a child from a high SES family is compared to is lower. Thus it is easier for the researcher to detect the detrimental effects of maternal working. It may be for such reasons that the aforementioned psychology literatures on role models or parental unemployment tend to focus on low SES households, such as racial minority groups (Bryant and Zimmerman, 2003) and single mother families (McLoyd et al., 1994).

To sum up, among the three effects, the time effect tends to increase the probability of initiating risky behaviors, the role model effect tends to decrease it, and the income effect is ambiguous. Therefore, maternal employment can have positive, negative, or no effects on the probability that a child/adolescent initiates risky behaviors.

IV. Data

A. Data Source

The data used in this study comes from the National Longitudinal Survey of Youths 1979 (NLSY79) and NLSY79 Children and Young Adults, administered by the Bureau of Labor Statistics. The NLSY79 is a nationally representative sample of 12,686 young men and women who were 14-22 years old when they were first surveyed in 1979. These individuals were interviewed annually until 1994 and biennially afterwards. NLSY79 Children has surveyed the children born to the female respondents of NLSY79 starting from 1986. The Children survey of NLSY79 includes assessments of each child as well as additional demographic and development information collected from either the mother or the child. Biennially (since 1994), children ages 15 and older completed a

---

family income is in the upper 50th percentile to initiate smoking cigarettes, using marijuana, being convicted of crime, having sex and becoming pregnant; and children who are non-white are significantly more likely than children who are white to initiate having sex and becoming pregnant.

lengthy child interview modeled on the NLSY79 questionnaire (NLSY79 Young Adults, or NLSY79 YA).

The variables relating to risky behavior initiations were obtained from both NLSY79 Children and NLSY79 YA. Maternal employment measures are collected from NLSY79. All the other independent variables come from NLSY79, NLSY79 Children, or NLSY79 YA. Table 1.1 describes the definition and source of each variable used in this paper. The time period this paper covers is 1984 to 2008, with biennial availability after 1994.[6]

An issue with the data is that there are some observations in NLSY79 YA for respondents who are no longer teenagers. The reason is that NLSY79 YA surveys all the descendents of the female NLSY79 respondents as long as they are at least 15 years of age. These observations are dropped from the analyses since the focus of this paper is *adolescent* risky behaviors.

B. Dependent Variables

Starting from 1988, children born to the female respondents of NLSY79 have been asked about their initiation age of a wide range of risky behaviors (initiation questions about criminal activity and pregnancy have been asked since 1994). From these questions, we construct the outcome variables, which include *Drinking Alcohol, Drinking at Least Once A Month, Smoking Cigarettes, Using Marijuana, Being Convicted of Crime*, *Having Sex* and *Becoming Pregnant*. The variable *Drinking Alcohol* is defined as having a glass of beer or wine or a drink of liquor, such as whiskey, gin, scotch (not

[6] The questions about risky behavior initiations were asked retrospectively, therefore, the outcomes are available all the way back to when these children were born. However, several control variables (e.g. *Mother Chronic Marijuana Use* and *Mother Cocaine*) are not available until 1984, which restricts the time period this paper could cover.

including childhood sips that one might have had from an older person's drink). The variable *Being Convicted of Crime* is defined as being convicted of any charges other than a minor traffic violation. The other variables are aptly named. Note that the variables may be subject to measurement errors since they are constructed from retrospective questions. However, there is a literature in psychology that describes that more salient events suffer less recall bias. For example, Eisenhower et al. (1991) argue:

> "The more salient the event, the stronger the memory trace and subsequently, the less effort to search of memory necessary to retrieve the information. This suggests that salient events are less subject to errors of recall decay than less salient events."

Since initiations of risky behaviors are likely more salient than any regular event, recall bias should be less of an issue, especially for risky behaviors such as *Having Sex*, *Becoming Pregnant* and *Being Convicted of Crime*. Even though inconsistencies across years do occur at times, the pattern of inconsistencies is such that adjacent ages tend to be reported, reflecting the nature of imprecise recall. However, there does not seem to be systematic over-reporting or under-reporting, indicating that the measurement errors are likely classical. In the baseline, we use the youngest initiation age reported. To make sure that measurement errors do not impose a problem, we also use the mean initiation age reported or the maximum initiation age reported.

C. Independent Variables

Maternal employment is our main independendant variable of interest and is defined in three ways. First, we consider annual maternal working hours, which is constructed by dividing the number of hours worked in the last calendar year by the percent time accounted for.[7] For example, if the mother reports that she worked 1000 hours in the last calendar year, but only 50% of the time was accounted for, then annual maternal working hours is 2000 for the mother in question. There are cases where the values for the annual hours are extremely large. As a robustness check, a top-coded hours variable[8] is also employed to check whether it would change the results. Additionally, for ease of interpretation, the hours variable used in the estimation is measured in units of 1000.

Maternal employment in a single year might suffer from reporting errors, hence we average the hours across years in hopes of averaging out any classical measurement errors. Specifically, we construct a variable *Hours over Life*, which is defined as the average annual hours worked by the mother over a child's life in the year prior to the interview. For example, if the outcome is in 2008, the main independent variable is the average annual hours worked by the mother from a child's birth to 2007.

In addition to *Hours over Life*, two hours variables are used together in a separate regression: *Hours over First 3 Years* and *Hours over Subsequent Years*,[9] which are defined as the average annual hours worked by the mother in the first 3 years of a child's

[7] The annual hours variable is created from an array of variables that record the week-by-week working hours of the respondent. However, due to missing data or indeterminate status, some weeks are not accounted for. See http://www.nlsinfo.org/nlsy79/docs/79html/79text/workhist.htm.

[8] The hours variable is top-coded to be 4160 (80 × 52). Less than 1% of the sample is affected.

[9] The age range included in the "subsequent years" depends on the youngest initiation age of the risky behavior, which is determined by the data. For example, for *Drinking Alcohol*, the age range included in the "subsequent years" could be 3 to 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, or 18 years old. The youngest initiation age for *Drinking Alcohol, Drinking at Least Once A Month, Smoking Cigarettes, Using Marijuana, Being Convicted of Crime, Having Sex* and *Becoming Pregnant* is 7, 11, 7, 10, 13, 12, and 14 years old respectively. The reason that the higher end of the age range is one year less than the youngest initiation age is because maternal employment is measured in the year prior to the interview. For example, for *Drinking Alcohol*, the smallest age range in the "subsequent years" is 3 to 6, while the youngest initiation age is 7.

The robustness check section also shows the results with the average annual hours worked by the mother in the first 1, 2, 4, 5 or 6 years of a child's life and in the correspondent subsequent years until one year prior to the interview as the main independent variables in the regressions.

life and in the subsequent years until one year prior to the interview, respectively.

Average maternal working hours is calculated from the years available. For example, if information on maternal working hours is only available for 2 years out of the first 3 years of a child's life, then we calculate the average from the 2 years, instead of setting it to missing because one year is missing.

Second, we consider percent time employed (PTE) of the mother. A mother is defined as employed in a certain year as long as she worked more than 0 hours in that year. Maternal employment measures in this category include *PTE over Life* and *PTE over First 3 Years* and *PTE over Subsequent Years*.[10] And they are defined as percent time employed of the mother over a child's life in the year prior to the interview, percent time employed of the mother in the first 3 years of a child's life and percent time employed of the mother in the subsequent years until one year prior to the interview, respectively. Missing values for percent time employed are treated the same way as above.

Finally, we consider employment status of the mother: *Full-Time, Part-Time, Unemployed* and *Out of the Labor Force* (with the omitted group being *Unemployed*). Mothers who work more than or equal to 1750 hours (35 × 50) annually are considered *Full-Time* workers for that year; those who work more than 0 but less than 1750 hours annually are considered *Part-Time* workers for that year; and those who report that they are out of the labor force for 52 weeks of the year are considered *Out of the Labor Force*

[10] Although all three variables of this set of maternal employment measure are reported in the summary statistics, only *PTE over Life* is used in the regressions. In results not shown, when *PTE over First 3 Years* and *PTE over Subsequent Years* are used in the same regressions, the pattern is very similar to that when *Hours over First 3 Years* and *Hours over Subsequent Years* are used in the same regressions. Therefore, due to limited space, these results are not reported.

for that year.[11] The omitted category consists of those who report 0 hours worked but are still in the labor force for at least part of the year, and hence are defined as *Unemployed*. To capture the average over a child's life, we define a set of variables that include *Always Full-Time, Always Part-Time,* and *Always Out of the Labor Force,* over a child's life in the year prior to the interview (the omitted category consists of *Employment Status Changers*, defined as mothers who change their employment status over a child's life in the year prior to the interview).[12] Missing values are treated the same way as above.

E. Instrumental Variables

Maternal employment might be endogenous due to omitted variables or reverse causality. If there are some unobserved characteristics that affect both maternal employment and adolescent risky behaviors, or if adolescent risky behaviors affect maternal labor supply decisions, the coefficients on maternal employment are estimated with a bias. For the latter concern, even if we use lagged maternal employment measures, reverse causality might still be a problem if mothers can predict the initiations of adolescent risky behaviors before it happens. This paper endeavors to deal with the endogeneity problem by using instruments.

The instruments utilized are *Health Limiting Kind of Work* and *Health Limiting*

---

[11] Since the data are compiled by the Bureau of Labor Statistics, the definition of "out of the labor force" in the data is consistent with the official definition. If the respondent is not out of the labor force for 52 weeks of a certain year, it means that they are either employed or unemployed for part of the year, and hence are not out of the labor force for the entire year.

[12] Strictly speaking, there is another category *Always Unemployed*. However, *Always Unemployed* is a very small group, consisting of only around 1 percent of the sample. This is because *Always Unemployed* means that the mother is never working but is always looking for work over the child's entire life, which is rarely the case given that most mothers would eventually obtain employment or drop out of the labor force altogether. Although we included it as a control, it does not have any predicting power, that is, it falls out of the regression results and the coefficient on it does not show up.

Again, due to limited space, only the life-time averages of this set of maternal employment measure: *Always Full-Time, Always Part-Time,* and *Always Out of the Labor Force,* are used in the summary statistics and regressions.

Amount of Work, which are defined as whether health limits the kind of work the mother could do and whether health limits the amount of work the mother could do, in the last calendar year (or after 1994, in the last survey year). This is to make sure that the instruments have relevance, since *Hours over Life* is defined in the year prior to the interview, while health limits questions were asked for the current year. For example, to instrument average maternal employment measured in 1996, which was with regard to 1995 and before, we use the health limits in 1994, the last available survey year.

For the instruments to be valid, they have to be correlated with maternal employment and uncorrelated with the error term. The former is obviously the case because of how these variables are defined. However, if there are unobservables in the error term that are correlated with the instruments and also affect adolescent risky behaviors, then the instruments are not valid. The unobservables that first come to mind include maternal ability and productivity, as well as maternal risk-seeking personality. Omitting these unobservables would bias the coefficient on maternal employment downward. However, the results actually show a significant positive association between maternal employment and the initiation of certain adolescent risky behaviors, hence even if these are indeed omitted variables, the conclusion that maternal employment increases the initiation hazards of these risky behaviors is not changed qualitatively.

Another unobservable that is potentially more problematic is a career-oriented but not family-oriented personality. Omitting this unobservable may bias the coefficient on maternal employment upward, possibly rendering a spurious positive relationship between maternal employment and adolescent risky behavior initiation. However, some of the characteristics that are commonly associated with employability such as responsibility, efficient time management, multi-tasking ability, etc., should enhance a mother's ability to take care of her child rather than reduce it. Psychology studies on the association between the Big Five [13] personality traits and career success find that conscientiousness positively predicted intrinsic and extrinsic success, and neuroticism negatively predicted extrinsic success (e.g. Judge et al., 1999). Conscientiousness is the state of being thorough, careful, or vigilant; it implies a desire to do a task well (Thompson, 2008). It is manifested in characteristic behaviors such as being efficient, organized, neat, and systematic (Thompson, 2008). Neuroticism, on the other hand, is manifested by characteristics of anxiety, moodiness, worry, envy and jealousy (Thompson, 2008). Individuals who score high on neuroticism are more likely than the average to experience such feelings as anxiety, anger, envy, guilt, and depressed mood (Matthews and Deary, 1998). Conscientiousness is likely to enhance maternal ability to take care of her child, while neuroticism is likely to reduce it. That is to say, personality traits should affect maternal employability and child-caring ability in the same direction.

To sum up, there are potentially two sources of bias for the OLS coefficients. The first source is less problematic because it biases the coefficients downward while we are finding positive effects. The second source is more problematic but it is very unlikely to exist. However, we still have to rule out that the first type of omitted variables are correlated with the instruments in order to make sure that the instruments are valid and further, to obtain unbiased estimates. To do so, we utilize statistical tools and regressions

---

[13] The Big Five Model, a.k.a. the Five Factor Model, is a comprehensive, empirical, data-driven research finding (Digman 1990). The five broad factors were discovered and defined by several independent sets of researchers (Digman 1990). The factors of the Big Five and their constituent traits include: Openness to experience – *inventive/curious* vs. *consistent/cautious*; Conscientiousness – *efficient/organized* vs. *easy-going/careless*; Extraversion – *outgoing/energetic* vs. *solitary/reserved*; Agreeableness – *friendly/compassionate* vs. *cold/unkind*; Neuroticism – *sensitive/nervous* vs. *secure/confident* (Atkinson et al. 2000).

to test the exogeneity of the instruments, which will be discussed in detail in the Methodology section.

F. Other Control Variables

Demographic characteristics of a child can influence his or her risky behavior initiation. Therefore, in the regressions, we include the following variables: a child's race/ethnicity (*White*, *Black* and *Hispanic*, with *White* being the omitted category), *Female*, *Child's Age* (dummies), *First-Born*, and *Residing with Mother*, which is defined as whether the usual residence of the child is with the mother.

Secondly, the background of the mother when she grew up might potentially affect the probability that her child initiates risky behaviors, and is likely exogenous. We include the following controls, which reflect the mother's background: *US Born*; *US 14* (whether she lived in the US at age 14); *Urban 14* (whether she lived in an urban area at age 14); *Magazines 14* (whether her household member received magazines regularly when she was 14); *Newspapers 14* (whether her household member received newspapers regularly when she was 14); *Library Card 14* (whether her household member had a library card when she was 14); *Working Female 14* (whether the adult female in her household worked when she was 14); her living arrangement when she was 14 (*Both Parents*, *Two Adults*, *One Adult*, and *Other Arrangement*, which are defined as whether she lived with both parents, whether she lived with two adults, whether she lived with one adult and whether she had other living arrangement, when she was 14, with the first one being the omitted category); her mother's education level (*Grandmother No High School*, *Grandmother High School*, and *Grandmother Some College Or More*, with the first one being the omitted category); her father's education level (*Grandfather No High*

School, *Grandfather High School*, and *Grandfather Some College Or More*, with the first one being the omitted category); her religious affiliation in 1979 (*No Religious Affiliation*, *Protestant*, *Baptist*, *Episcopalian*, *Lutheran*, *Methodist*, *Presbyterian*, *Roman Catholic*, *Jewish* and *Other Religious Affiliation*, with the first one being the omitted category); her religious attendance in 1979 (*Not At All, Infrequently, Once Per Month, 2-3 Times Per Month*, *Once Per Week* and *More than Once Per Week*, with the first one being the omitted category); and *AFQT 1981* (her AFQT score in 1981).

Moreover, current maternal characteristics are likely associated with the mother's employment status as well as her child's risky behavior patterns, and hence should be controlled for in the regressions (Thornberry, 1987; Anderson et al., 2003; Aughinbaugh and Gittleman, 2004). In this paper, current maternal characteristics controlled for include: maternal education level (*Less than High School*, *High School*, *Some College*, and *College Or More*, with the first one being the omitted category); *Maternal Age at Birth*; marital status (*Single*, *Married* and *Other Marital Status*, with the first one being the omitted category); *Mother Regular Drinking*;[14] *Mother Chronic Marijuana Use*;[15] *Mother Cocaine Use*;[15] *Mother Illegal Activity*;[16] and *Mother Obese*.[17]

---

[14] This is a dummy variable indicating whether a mother drank at least once a month. Questions about maternal alcohol use are asked in 1982, 1983, 1984, 1985, 1988, 1989, 1992, 1994, 2002, 2006 and 2008. We impute maternal alcohol use for those years in the gaps by using the information from the most recent previous year available. For example, maternal alcohol use in 1986 or 1987 is the same as that in 1985. Maternal alcohol use before 1982 is not imputed.

[15] This is a dummy variable indicating whether a mother used marijuana for at least 100 times in her lifetime (whether a mother has ever used cocaine). Questions about maternal marijuana (cocaine) use are asked in 1984, 1988, 1992, 1994 and 1998. We impute maternal marijuana (cocaine) use for those years in the gaps and after 1998 by using the information from the most recent previous year available. Maternal marijuana (cocaine) use before 1984 is not imputed.

[16] This is a dummy variable indicating whether a mother was ever charged with illegal activity before 1980. In the 1980 survey, the mother is asked whether she was ever charged with illegal activity except for a minor traffic offense. We impute this variable for all the years after 1980 using the information from 1980.

[17] This is a dummy variable indicating whether a mother is obese. Maternal weight information is collected in 1981, 1982, 1985, 1986, 1988, 1989, 1990, 1992, 1993, 1994, 1996, 1998, 2000, 2002, 2004, 2006 and

Controlling for household characteristics is also important (Blum et al., 2000, Anderson et al., 2003, Aughinbaugh and Gittleman, 2004). To this end, we include *Father Figure*;[18] *Number of Children*; *Number of Adults*; *Age of Youngest Child*; and *Family Income over Life,* defined as the average family income from a child's birth to one year prior to the interview, measured in units of 1000, converted to year 2001 dollars.

Finally, *Urban* (whether mother resides in an urban area), maternal region of residence (*Northeast*, *North Central*, *South* and *West*, with the first one being the omitted category), and a linear time trend are included in the regressions.

V. Methodology

A discrete time piecewise-constant hazard model is used to study how maternal employment is associated with the hazards of initiating risky behaviors by the child/adolescent. The model is a variation of the proportional hazards model, which is commonly used in survival analysis. Survival analysis examines and models the time it takes for events to occur. The method is employed in a variety of disciplines under various rubrics – for example, 'event-history analysis' in sociology (Fox, 2002). In health economics, the method is considered an appropriate statistical technique for examining the structural determinants of the transition from one state to another, and has been used to study the initiation of risky behaviors such as smoking (Cawley et al., 2004).

2008, and maternal height information is collected in 1981, 1982 and 1985, by which time the youngest mothers should have reached adult height. We correct for the reporting errors in weight and height using the method outlined in Cawley (2004, p.455). We use the corrected weight and height to calculate BMI, which is weight in kilograms divided by height in meters squared. Being obese is defined as having a BMI greater than or equal to 30. We impute maternal obesity status for those years in the gaps by using the information from the most recent previous year available.

[18] This is a dummy variable indicating whether a father figure is in the household. For a married mother, father figure refers to her spouse; for a single mother, father figure refers to her partner.

Proportional hazard model was first introduced by Cox (1972):

$$\frac{\lambda(t_j;x_i)}{1-\lambda(t_j;x_i)} = \exp(x_i\beta)\frac{\lambda_0(t_j)}{1-\lambda_0(t_j)} \quad (1),$$

where $\lambda(t_j) = \Pr(T = t_j \mid T > t_j - 1)$ is the hazard function, defined as the conditional probability of failure at time t_j given that one has survived through time t_j - 1. $\lambda(t_j;x_i)$, accordingly, is the hazard at time t_j for an individual with covariates x_i. $\lambda_0(t_j;x_i)$ is the baseline hazard at time t_j, and $\exp(x_i\beta)$ is the relative risk associated with covariates x_i. In the case of time-varying covariates, only the values of the covariates at the discrete times are relevant.

Taking natural logarithm on both sides of equation (1), one can obtain a model on the logit of the hazard at time t_j given survival up to that time:

$$\log it\lambda(t_j;x_i) = \alpha_j + x_i\beta \quad (2),$$

where α_j = logit $\lambda_0(t_j;x_i)$ is the logit of the baseline hazard, and $x_i\beta$ is the effects of the covariates on the logit of the hazard. The model essentially treats time as a discrete factor by introducing one parameter α_j for each possible time of failure t_j.

One can calculate the hazard $\lambda(t_j;x_i)$ once the estimates of α_j and β are obtained:

$$\lambda(t_j;x_i) = \frac{1}{1+\exp(-x_i\beta)\cdot\exp(-\alpha_j)} \quad (3).$$

Since this paper studies the initiation of risky behaviors, age rather than year is the more appropriate discrete factor. A piecewise constant hazard model with age as the discrete factor is used in order to allow the baseline logit hazard to shift up or down at each age, because the probability of initiation tends to grow with age, not necessarily year. That is, a group of age dummies are introduced in the model.[19]

In order to pin down the causal effects of maternal employment on the initiation of risky behaviors by the child/adolescent, we employ the instrumental variable approach. Before doing that, however, we carry out a group of OLS regressions to check whether the marginal effects from the logit regressions (hazard model) are similar to the OLS coefficients, and furthermore, to check whether the OLS and IV coefficients are statistically different using a Durbin-Wu-Hausman (DWH) test.

In terms of strength, the instruments perform well in the related statistical tests: when the sample is not stratified by SES, the first stage F statistic on the excluded instruments is at least 16, above the rule-of-thumb threshold of 10 suggested by Staiger and Stock (1997). However, the instruments do not perform as well when the sample is stratified by SES, especially for the subgroup with higher SES. Specifically, in 2 out of 28 cases, the first stage F statistic on the excluded instruments falls below 5. These two cases occur when the sample is stratified by *Family Income over Life* and for the sub-sample in the upper 50th percentile of the income distribution. When the outcome is *Drinking Alcohol* (*Smoking Cigarettes*), the first stage F statistic on the excluded instruments is 3.06 (4.48). Interestingly, the fact that the instruments work better for the group with lower SES suggests that the instruments are likely exogenous, because it shows that health limits are likely physical constraints that prevent one from carrying out manual work rather than a proxy for ability, productivity or personality. The idea is that women with lower SES are more likely to engage in the type of jobs that require physical strength, while physical well being affects women with higher SES less.

[19] Because of the high correlation between age and year, we cannot control for year fixed effects and age dummies at the same time. Instead, a linear time trend is included in the regressions to reflect the effect of time on adolescent risky behavior initiation.

In terms of exogeneity, we utilize Hansen's overidentification test (Hansen's J statistic). The null-hypothesis of Hansen's overidentification test is that the instruments are valid, i.e., they are uncorrelated with the error term, and that the excluded instruments are correctly excluded from the estimated equation. When the sample is not stratified, only for one outcome *Becoming Pregnant* is the p-value of the Hansen's J statistic below 0.05. A p-value below 0.05 means that one can reject the null hypothesis at the 5% level, i.e., the instruments are endogenous. However, for this outcome, the coefficient on maternal employment is never significant. Again, for the subgroup analyses, the instruments do not perform as well; specifically, the p-value of the Hansen's J statistic is sometimes below 0.05 when the outcome is initiation of *Smoking Cigarettes* or *Becoming Pregnant*.

To further test the exogeneity of the instruments, we also conduct a group of "suggestive" tests. The endogeneity of the instruments comes from the possibility that there are some unobserved maternal or adolescent characteristics that are associated with both maternal health limits and the initiation of risky behaviors by the child/adolescent. While these unobserved characteristics are not directly available, provided that they exist, it can be informative to examine whether health limits are correlated with observables that are believed to be related to these unobserved factors. For the question at hand, three groups of such observables could be relevant. First, general intelligence and education level may be correlated with maternal ability to take care of and educate her children, and may also affect maternal employment. Second, maternal risky behaviors may be correlated with risk-seeking personality that serves to dampen employment prospect and is also passed down to children through genes or a common household environment.

Third, maternal productivity in child-caring activities is possibly related to productivity in paid work, which obviously affects maternal employment. To implement these "suggestive" tests, maternal *AFQT Score* and *Years of Schooling*, a group of maternal risky behaviors,[20] and a group of parental investment measures[21] are regressed on the two instruments and the other regressors, which include all the independent variables in the main regressions except for maternal employment measures. When maternal *AFQT score*, *Years of Schooling* or a certain maternal risky behavior is the dependent variable, we do not control for children-related variables, since these dependent variables were measured early on when most of the female respondents of NLSY79 had not yet borne children.[22] Moreover, since parental investment measures do not distinguish between maternal and paternal investment, we also run regressions with the sample restricted to single mothers. The results are shown in Appendix Table 1.1 and Appendix Table 1.2.

Although this is not a definitive test, the regression results are generally consistent with the identifying assumption: in none of these regressions is *Health Limiting Kind of Work* or *Health Limiting Amount of Work* significant at the 5% level or below; out of the 82 coefficients on either *Health Limiting Kind of Work* or *Health Limiting Amount of*

---

[20] These include: *Mother Chronic Smoking* (whether a mother smoked at least 100 cigarettes in her lifetime), *Mother Regular Drinking*, *Mother Chronic Marijuana Use*, *Mother Cocaine Use*, and *Mother Illegal Activity*. Other than the first variable, the rest are all controls in the main regressions. The reason that we exclude the first one as a control in the main regressions is because this variable is only available after 1992, and we cannot impute the values backwards for the years before then. That is, if this variable is included, it would reduce the sample size of the main regressions by quite a bit.

[21] See the appendix for how these variables are constructed.

[22] For example, *AFQT score* was measured in 1980, during which year around 70% of the sample did not have children yet. Therefore, we eliminate the following controls in the exogeneity test regressions: child's race/ethnicity (*White*, *Black* and *Hispanic*, with *White* being the omitted category), *Female*, *Child's Age* (dummies), *First-Born*, *Residing with Mother*, *Maternal Age at Birth*, *Number of Children*, *Age of Youngest Child*, *Average Income Over Life*, and we replace child's race/ethnicity with mother's race/ethnicity, and we replace *Average Income Over Life* with maternal family income in the past calendar year.

Work, only 8 are significant at the 10% level (2 of them positive).[23]

Lastly, to test the heterogeneous role model effect, we carry out subgroup analyses along the following SES dimensions: maternal education level: high school or less vs. at least some college; maternal marital status: single (never married, separated, divorced or widowed) vs. married; *Family Income over Life*: lower 50th percentile vs. upper 50th percentile; and child's race/ethnicity: non-white (black or Hispanic) vs. white.

VI. Empirical Results

A. Summary Statistics

Table 1.2 shows the summary statistics by maternal employment status in the last calendar year.[24] Specifically, mothers who were not employed in the last calendar year include both those who were unemployed and those who were out of the labor force in the last calendar year. As can be seen from the data, for five out of the seven risky behaviors, *Drinking Alcohol*, *Drinking at Least Once A Month*, *Smoking Cigarettes*, *Using Marijuana* and *Having Sex*, initiation rates are higher for children whose mothers

---

[23] In the Data section, we mentioned that there might be another unobservable that could bias the estimates upward but is very unlikely to exist. We refer to this unobservable as a career-oriented but not family-oriented personality. Although it is impossible to measure this unobservable, given that one's geographic location and religious affiliation can influence their views towards family and career, we use region of residence (*Northeast, North Central*, *South* and *West*,) and religious affiliation (*No Religious Affiliation*, *Protestant, Baptist*, *Episcopalian*, *Lutheran*, *Methodist*, *Presbyterian*, *Roman Catholic*, *Jewish* and *Other Religious Affiliation*) as proxies for this unobservable and carry out another set of "suggestive" tests. The results are again consistent with the hypothesis that the instruments are not correlated with the unobservable. Specifically, out of the 14 regressions, none of the coefficients on *Health Limiting Kind of Work* or *Health Limiting Amount of Work* is significant at the 5% level or less; only when the outcome is *Methodist* is the coefficient on *Health Limiting Amount of Work* significant at the 10% level. The results are not shown here but available upon request.

[24] Table 1.2 uses the sample for which initiating *Drinking Alcohol* is the outcome. A hazard model is employed in this study, which means that observations must be dropped once initiation starts, hence the sample is different for each risky behavior. Summary statistics for samples by other risky behaviors are available upon request.

were employed in the last calendar year, and the difference between the two groups is statistically significant for all but *Using Marijuana*.

Panels two through four present the summary statistics for the average annual hours worked by the mother, the average percent time employed of the mother and the employment status of the mother. Mothers who were not employed in the last calendar year worked more hours and were employed for a larger percent of the time during the first 3 years of a child's life (304 hours and 34%) than during subsequent years (242 hours and 24%). Mothers who were employed in the last calendar year worked fewer hours and were employed for a smaller percent of the time during the first 3 years of a child's life (943 hours and 70%) than during subsequent years (1383 hours and 86%). The cross group comparison shows that, at any point in a child's life, mothers who were employed in the last calendar year worked more hours and were employed for a larger percent of the time than those who were not employed in the last calendar year (e.g., *Hours over Life*: 1241 hours vs. 269 hours, *PTE over Life*: 80% vs. 28%). From the summary statistics for the average employment status of the mother, one can see that regardless of maternal employment status in the last calendar year, the majority of mothers change their employment status over a child's life (81% for mothers who were not employed in the last calendar year and 88% for mothers who were employed in the last calendar year).

The final panel shows the means and standard deviations of the control variables. Mothers who were employed in the last calendar year are more likely to be born and raised in the US, have better-educated parents, have higher AFQT scores, and are better educated, on average, than mothers who were not employed in the last calendar year.

B. Regression Results

The baseline results are reported in Table 1.3. The first panel displays the coefficients on *Hours over Life*, *Family Income over Life*, and *Age* dummies from the discrete time piecewise-constant hazard model, with the marginal effects at the mean of each control variable shown in square brackets. The second and third panels display the coefficients on *Hours over Life* from the linear probability model and the instrumental variable model. The other control variables used have been described in the Data section, but their coefficients are not reported due to space constraint. Robust standard errors clustered at the household level are reported in parentheses.

In terms of magnitudes, the marginal effects from the hazard model and the OLS coefficients are very similar. Moreover, the DWH test cannot reject the null hypothesis that the OLS coefficient is consistent for all but initiating *Drinking at Least Once A Month* and *Having Sex,* at the 10% level. However, for these two outcomes, the IV coefficients are also statistically significant, and much larger than their OLS counterparts. Therefore, the OLS results are used to quantify the causal effects other than in the two exceptions where the IV results will be used. Ceteris paribus, a 1000 hours increase in *Hours over Life* increases the initiation hazard of *Drinking Alcohol*, *Drinking at Least Once A Month*, *Smoking Cigarettes*, *Using Marijuana* and *Having Sex* by 1.7, 16.3, 0.9, 0.7, and 10.6 percentage points respectively.

The magnitudes of the coefficients from the IV estimation are large, but they are not entirely implausible. Although the summary statistics show that the means for initiating *Drinking at Least Once A Month* and *Having Sex* are 13% and 19%, respectively, these means are “diluted” by those who never initiated these risky behaviors. That is to say, because of the utilization of the hazard model, individuals are dropped once they initiated, but for those who never initiate, they are kept until they are 19 years of age. However, from summary statistics not shown here, around 73% of the observations have initiated *Drinking at Least Once A Month* before age 19, and around 82% of the observations have initiated *Having Sex* before age 19. Moreover, if the "compliers" of the instruments are especially susceptible to maternal employment, or if the "non-compliers" of the instruments are immune to maternal employment, it might explain the large IV coefficients. The latter is more likely, since the "non-complying" mothers, who will always work or not work regardless of their health conditions, might be extremely resourceful or extremely disadvantaged, and their working status might indeed not matter as much in affecting their children's risky behavior initiation, rendering the OLS coefficient diluted.

Family Income over Life significantly increases the probability that the adolescent initiates *Drinking Alcohol*, but significantly decreases the probability that the adolescent initiates *Having Sex* and *Becoming Pregnant*. These results point to the mixed effects of income. On the one hand, higher income potentially provides the means for the adolescents to afford some of these risky activities more easily, which serve to increase the initiation hazards. On the other hand, wealthier families can better afford substitutes for maternal care, such as care from nannies or after-school activities, which serve to decrease the initiation hazards of certain risky behaviors.

Age tends to increase the initiation hazards of risky behaviors except for *Using Marijuana* and *Having Sex*, the initiation hazards of which peak at 16 and 18 years of age respectively, and then start declining. Figure 1.1, which shows the average predicted hazards at each age for every risky behavior, provides a visualization of this relationship.

This supports the argument that maternal supervision in middle childhood and adolescence could be critical.

Table 1.4 shows the results with *PTE over Life* as the main independent variable. The results in Table 1.4 closely mirror those in Table 1.3. When a mother changes from not working at all to working one hundred percent of the time over a child's life in the year prior to the interview, the initiation hazard of *Drinking Alcohol*, *Drinking at Least Once A Month*, *Smoking Cigarettes*, *Using Marijuana* and *Having Sex* is increased by 3.0, 32.6, 1.7, 1.5 and 21.8 percentage points, respectively. Again, the magnitudes of the IV coefficients for initiating *Drinking at Least Once A Month* and *Having Sex* are large. However, a change from not working at all to working one hundred percent of the time is substantial, especially if this change is caused by health limits alone (for the "compliers" of the instruments), and therefore it might indeed cause a large change in the initiation probabilities. In terms of *Family Income over Life* and *Age*, they have similar effects on the initiation hazards when *PTE over Life* is the main independent variable compared to when *Hours over Life* is the main independent variable.

Table 1.5 presents the results with maternal employment status as the main independent variables. Since the number of endogenous variables exceeds the number of instruments, we cannot run IV regressions, and thus only the results from the hazard model are shown. Although there are only a few significant coefficients, one can see that *Always Full-Time* tends to increase the initiation hazards of most of the risky behaviors, while *Always Part-Time* or *Always Out of the Labor Force* tends to decrease them. This is consistent with the notion that a high level of maternal employment is harmful, while a modest level of maternal employment is beneficial, for preventing the initiation of certain risky behaviors.

Table 1.6 shows the results from the regressions for which *Hours over First 3 Years* and *Hours over Subsequent Years* are the main independent variables. The first panel displays the results where these two variables are entered together in the regressions, while the second panel shows the results where these two variables are entered separately. Marginal effects at the mean of each control variable are reported in square brackets, the variance inflation factors (VIF) are reported in curly brackets, and the correlation between and the p-value of the joint F-test on the two main independent variables are also reported.

As can be seen from Table 1.6, the VIFs of the two variables are low, well below the rule-of-thumb cutoff of 10, suggesting that multicollinearity is not severe, although the correlation between the two variables is high. However, if only *Hours over First 3 Years* is used, it would pick up the effect of *Hours over Subsequent Years*. Although this is also true if only *Hours over Subsequent Years* is used, the problem of omitted variable bias is less severe. That is, the magnitudes of the coefficients on *Hours over Subsequent Years* remain similar to those in the first panel, and the statistical significance of the coefficients remains as well. One might not be able to conclude for certain whether maternal employment in the first 3 years has an effect due to the double problems of multicollinearity and omitted variable bias, but it seems safe to conclude that the subsequent years matter. Specifically, a 1000 hours increase in *Hours over Subsequent Years* is associated with a 1.3, 0.8, 0.7, 0.5, and 2.2 percentage points increase in the initiation hazard of *Drinking Alcohol*, *Drinking at Least Once A Month*, *Smoking Cigarettes*, *Using Marijuana*, and *Having Sex*, respectively.

Table 1.7 through Table 1.10 display the results from the subgroup analyses. The evidence is generally supportive of the hypothesis that the role model effect is larger for households with lower SES. Appendix Table 1.3 presents a high-level summary of these results. Specifically, the effect of maternal employment on initiating *Drinking Alcohol* is larger for households with better educated mothers and white children; the effect of maternal employment on initiating *Drinking at Least Once A Month* is larger for households with better educated mothers, married parents and white children; the effect of maternal employment on initiating *Smoking Cigarettes* is larger for households with better educated mothers, married parents, higher income and white children; the effect of maternal employment on initiating *Marijuana Use* is larger for households with better educated mothers; and the effect of maternal employment on initiating *Having Sex* is larger for households with married parents, higher income and white children.

Vii. Robustness Checks

Several robustness checks are carried out, which are shown in Appendix Table 1.4 and Appendix Table 1.5. Firstly, the cutoff between the first few years and the subsequent years is varied. The cutoff is chosen to be 1, 2, 3, 4, 5 or 6. The baseline results, i.e., the results with the cutoff being 3, are also included in the table for ease of comparison. It can be seen that the general pattern stays the same when the cutoff varies. Specifically, *Hours over Subsequent Years* is positively and significantly associated with the probability that the adolescent initiates the following risky behaviors: *Drinking Alcohol*, *Drinking at Least Once A Month*, *Smoking Cigarettes, Using Marijuana*, and

Having Sex, with the only exception being when initiating *Using Marijuana* is the outcome and the cutoff is 1. Interestingly, as the cutoff gets larger, the coefficient on *Hours over First Few Years* generally becomes larger, although mostly not statistically significant. This suggests that the effects of maternal employment might indeed increase with child's age.

Additional robustness checks are reported in Appendix Table 1.5. In this table, *Hours over Life* is always the main independent variable. Again, the baseline results are included in the table for ease of comparison, which are shown in Panel a.

Panel b displays the results with maternal work shifts as additional controls. Mothers have been asked about the shift usually worked at their current or most recent job every year since 1979. The variable *Day Shift* is coded as 1 if a mother's usual shift is the regular day shift, and 0 otherwise.[25] However, between 1986 and 1989, instead of being asked about the shift, mothers were asked about the time they usually began and ended work at the current or most recent job. For these years, the variable *Day Shift* is coded as 1 if a mother's working hours are between 6 am and 6 pm, and 0 otherwise. Another variable *Non-Day Shift* is coded as 1 if *Day Shift* equals 0, and 0 if *Day Shift* equals 1. For those mothers who have never worked, both variables are coded as 0. The main independent variable is an average over a child's life, but the shift variables are about the current or most recent job. A better approach may be to take the average of *Day Shift* and *Non-Day Shift* over a child's life. However, using that method still does not solve the problem of inconsistency between the hours variable and the shift variables. This inconsistency is due to the number of working hours being measured with regard to

---

[25] These include the regular evening shift, the regular night shift, a split shift and varying hours.

the last calendar year and the shift worked being measured with regard to the current or most recent job, which might not be in the last calendar year. It is for this reason that these variables are not included in the baseline regressions. To check whether the omission of them would bias the results, *Day Shift* and *Non-Day Shift* (rather than the life-time averages of the two variables) are included in the regressions. The results indicate that the coefficients on maternal working hours remain very similar both in terms of magnitudes and significance levels. The only major difference is that the coefficient on *Hours over Life* loses significance when the outcome is initiating *Using Marijuana*.

Panel c displays the results with *Child Working For Pay* as an additional control. One may argue that it is important to control for the income that the child/adolescent has at their disposal, with the ideal measure being the amount of allowance they have. In NLSY79 Children, the respondents are asked each year: “Who usually makes the decisions about how much allowance you get?” rather than the specific amount. In NLSY79 Young Adult, there are no questions about allowance. Since the sample for this paper consists of observations from both NLSY79 Children (observations for individuals who are younger than 15 years old) and NLSY79 Young Adult (observations for individuals who are at least 15 years old), it is not possible to control for allowance-related variables. Another way to proxy for a child’s/adolescent’s income is to use their earned income. This poses the difficulty that young adults are asked about their income on both regular and odd jobs, while children are only asked about their income on odd jobs. Therefore, in an effort to make this measure consistent across observations, a dummy variable *Child Working For Pay* is constructed, which takes on the value 1 if a child or adolescent works at any paid job, regular or odd, and 0 otherwise.

Due to this variable's potential inconsistency, it is not included in the main regressions. However, the robustness check, which includes this variable as a rough proxy for a child's/adolescent's own financial resource, indicates that omitting this measure does not bias the results.

Panel d displays the results with top-coded maternal working hours as the main independent variable. Specifically, annual maternal working hours is top-coded to be 4160 (80×52)[26] in this robustness check. The resulting coefficients are only slightly smaller, and the significance or insignificance is unchanged for all the outcomes other than initiating *Using Marijuana*, for which the coefficient on *Hours over Life* becomes insignificant.

Panel e displays the results with interpolated independent variables, which are employed to make sure that missing values do not cause biases in the estimates. The full sample should consist of observations for which the information on at least one risky behavior is available. However, because the control variables might have missing values, the sample actually used in the baseline regressions is smaller than the full sample. For each of the independent variables, the missing values are filled in with the sample means, and an indicator variable denoting whether the value is missing is added to the regressions. The results are again consistent with the baseline.

The final three panels display the results from the sample for which maternal health limits vary over time, that is, the values for *Health Limiting Kind of Work* or *Health Limiting Amount of Work* is not always 0s or 1s for the same mother. The idea is that the instruments are more likely to be exogenous if it is not a permanent condition that

[26] Less than 1% of the sample has annual working hours greater than 4160.

prevents a mother from being able to work. Panel f shows the results from the hazard model, Panel g shows the OLS results, and Panel h shows the IV results, using this restricted sample. The results from the hazard model remain very similar to those from the baseline hazard model, with the only exception being that the coefficient on *Hours over Life* loses significance when initiating *Using Marijuana* is the outcome. The marginal effects from the hazard model are very similar to the OLS coefficients. Moreover, the DWH test cannot reject that OLS is efficient except when the outcome is initiating *Drinking at Least Once A Month*, for which the IV coefficient is larger than the OLS coefficient and also significant. Overall, these results tend to confirm the robustness of the baseline.

VIII. Conclusions

This paper studies the effects of maternal employment on the initiation of adolescent risky behaviors. The results indicate that, overall, maternal employment has a positive and significant impact on the initiation of alcohol, cigarette and marijuana consumption, as well as sexual activity. Furthermore, maternal employment after the first few years of a child's life seems to matter more than that in the first few years. Unlike other measures of childhood development, risky behaviors may be affected by middle childhood and adolescence experiences more than early childhood experiences. In addition, the effects of maternal employment on the initiation of adolescent risky behaviors are often times larger for high SES households, suggesting that the role model effect is greater for low SES households. Finally, the results from the baseline model are robust to including additional controls, using alternative independent variables, and using an alternative sample.

References

Almond, D. (2006). Is the 1918 influenza pandemic over? Long-term effects of in utero influenza exposure in the post-1940 U.S. population. *Journal of Political Economy*, 114(4), 672-712.

Anderson, P. M., Butcher, K. F., & Levine, P. B. (2003). Maternal employment and overweight children. *Journal of Health Economics*, 22(3), 477-504.

Anderson, K. H., Foster, J. E., & Frisvold, D. E. (2010). Investing in health: The long-term impact of Head Start on smoking. *Economic Inquiry*, 48(3), 587-602.

Atkinson, R. L., Atkinson, R. C., Smith, E. E., Bem, D. J., & Nolen-Hoeksema, S. (2000). *Hilgard's Introduction to Psychology* (13 ed.). Orlando, Florida: Harcourt College Publishers. pp. 437.

Aughinbaugh, A., & Gittleman, M. (2003). Maternal employment and adolescent risky behavior. BLS working paper 366.

Aughinbaugh, A., & Gittleman, M. (2004). Maternal employment and adolescent risky behavior. *Journal of Health Economics*, 23(4), 815-838.

Barnes, G. M., & Farrell, M. P. (1992). Parental support and control as predictors of adolescent drinking, delinquency, and related problem behaviors. *Journal of Marriage & the Family*, 54(4), 763-776.

Becker, G. S. (1976). *An Economic Approach to Human Behavior*. Chicago IL: University of Chicago Press.

Beenstock, M., & Rahav, G. (2002). Testing gateway theory: do cigarette prices affect illicit drug use?. *Journal of Health Economics*, 21(4), 679-698.

Bernal, R. (2008). The effect of maternal employment and child care on children's cognitive development. *International Economic Review*, 49(4), 1173-1209.

Blank R. (2002) Evaluating welfare reform in the United States. *Journal of Economic Literature*, 40(4), 1105-1166.

Blau, F. D., & Grossberg, A. J. (1992). Maternal labor supply and children's cognitive development. *Review of Economics and Statistics*, 74(3), 474-481.

Blum R. W., Beuhring T., Shew M. L., Bearinger L. H., Sieving R. E., Resnick M. D. (2000). The effects of race/ethnicity, income, and family structure on adolescent risk behaviors. *American Journal of Public Health*, 90(12), 1879–1884.

Bryant, A. L., & Zimmerman, M. A. (2003). Role models and psychosocial outcomes among African American adolescents. *Journal of Adolescent Research*, 18(1), 36-67.

Brody, J. E. (2007, December 18). Teenage risks, and how to avoid them. *The New York Times*. Retrieved from http://www.nytimes.com/2007/12/18/health/18brod.html?pagewanted=all#.

Brooks-Gunn, J., Han, W. & Waldfogel, J. (2002). Maternal employment and child cognitive outcomes in the first three years of life: the NICHD study of early child care. *Child Development*, 73(4), 1052–72.

Bruer, J. T. (1999). *The Myth of the First Three Years: A New Understanding of Early Brain Development and Lifelong Learning*. New York: Free Press.

Cawley, J., (2004). The impact of obesity on wages. *Journal of Human Resources*, 39 (2), 451-474.

Cawley, J., Markowitz, S., & Tauras, J. (2004). Lighting up and slimming down: the effects of body weight and cigarette prices on adolescent smoking initiation. *Journal of Health Economics*, 23(2), 293-311.

Cawley, J. & Feng, L. (2007). Maternal employment and childhood obesity: A search for mechanisms in time use data. NBER Working Paper 13600.

Centers for Disease Control and Prevention (CDC). (2009). Youth risk behavior surveillance—United States, 2009. *MMWR 2010*, 59(SS-5), 1-142.

Chase-Lansdale, P., & Pittman, L. D. (2002). Welfare reform and parenting: Reasonable expectations. *The Future of Children*, 12(1), 167-185.

Chetty, R., Friedman, J. N., Hilger, N., Saez, E., Schanzenbach, D., & Yagan, D. (2011). How does your kindergarten classroom affect your earnings? Evidence from Project STAR. *Quarterly Journal of Economics*, 126(4), 1593-1660.

Chia, Y. (2008). Maternal labour supply and childhood obesity in Canada: Evidence from the NLSCY. *Canadian Journal of Economics*, 41, 217–244.

Coley, R., Medeiros, B. L., & Schindler, H. S. (2008). Using sibling differences to estimate effects of parenting on adolescent sexual risk behaviors. *Journal of Adolescent Health*, 43(2), 133-140.

Cox, D. R. (1972). Regression models and life tables. *Journal of the Royal Statistical Society, Series B*, 34(2), 187-220.

Digman, J. M. (1990). Personality structure: Emergence of the five-factor model. *Annual Review of Psychology*, 41, 417–440.

Eccles, J. S. (1999). The development of children ages 6 to 14. *The Future of Children*, 9(2), 30-44.

Eisenhower D., Mathiowetz N. A., Morganstein D. (1991). Recall error: Sources and bias reduction techniques. In P. P. Biemer, R. M. Groves, L. E. Lyberg, N. A. Mathiowetz, & S. Sudman (Eds.), *Measurement errors in surveys* (pp. 127-144). New York: John Wiley & Sons.

Erikson, E. H. (1968). *Identity, youth and crisis.* New York: W.W. Norton and Company.

Fertig, A., Glomm, G., & Tchernis, R. (2009). The connection between maternal employment and childhood obesity: Inspecting the mechanisms. *Review of Economics of the Household*, 7(3), 227-255.

Fox J. (2002). Cox proportional-hazard regression for survival data. Appendix to an R and S-PLUS companion to applied regression. Retrieved from http://cran.r-project.org/doc/contrib/Fox-Companion/appendix-cox-regression.pdf. Accessed October 2012.

Francis, E. (2011, September 22). Family dinners linked to less risky behavior in teens. *ABC News*. Retrieved from http://abcnews.go.com/Health/Parenting/family-dinners-linked-risky-behavior-teens/story?id=14583590#.UAOCeRz2Css.

Garces, E., Thomas, D., Currie, J., (2002). Longer term effects of Head Start. *American Economic Review*, 92(4), 999–1012.

Gennetian, L. A., Duncan, G. J., Knox, V. W., Vargas, W. G., Clark-Kauffman, E., & London, A. S. (2002). *How Welfare and Work Policies for Parents Affect Adolescents: A Synthesis of Research.* New York: Manpower Demonstration Research Corporation.

Gennetian, L. A., Hill, H. D., London, A. S., & Lopoo, L. M. (2010). Maternal employment and the health of low-income young children. *Journal of Health Economics*, 29(3), 353-363.

Gordon, R. A., Kaestner, R., & Korenman, S. (2007). The effects of maternal employment on child injuries and infectious disease. *Demography*, 44(2), 307-333.

Grimsley, K. D., Melton, R. H. (1998, March 22). Homemakers Get Plenty of Support. *The Washington Post*, pp. A16.

Harris, K., Duncan, G. J., & Boisjoly, J. (2002). Evaluating the role of 'nothing to lose' attitudes on risky behavior in adolescence. *Social Forces*, 80(3), 1005-1039.

Judge, T. A., Higgins, C. A., Thoresen, C. J., and Barrick, M. R. (1999). The Big Five personality traits, general mental ability, and career success across the life span. *Personnel Psychology*, 52(3), 621–652.

Matthews G. and Deary I. J. (1998). *Personality traits*. Cambridge, UK: Cambridge University Press.

McLoyd, V. C., & Jayaratne, T. (1994). Unemployment and work interruption among African American single mothers: Effects on parenting and adolescent socioemotional functioning. *Child Development*, 65(2), 562-589.

Montemayor, R., & Clayton, M. D. (1983). Maternal employment and adolescent development. *Theory Into Practice*, 22(2), 112-118.

Moore, M. R., & Chase-Lansdale, P. (2001). Sexual intercourse and pregnancy among African American girls in high-poverty neighborhoods: The role of family and perceived community environment. *Journal of Marriage & The Family*, 63(4), 1146-1157.

Morrill, M. (2011). The effects of maternal employment on the health of school-age children. *Journal of Health Economics*, 30(2), 240-257.

Rossin, M. (2011). The effects of maternity leave on children's birth and infant health outcomes in the United States. *Journal of Health Economics*, 30(2), 221-239.

Ruhm, C. J. (2000). Parental leave and child health. *Journal of Health Economics*, 19(6), 931-960.

Ruhm, C. J. (2004a). Maternal employment and adolescent development. NBER Working Paper 10691.

Ruhm, C. J. (2004b). Parental employment and child cognitive development. *Journal of Human Resources*, 39(1), 155–92.

Ruhm, C. J. (2008). Maternal employment and adolescent development. *Labour Economics*, 15(5), 958-983.

Staiger, D., & Stock, J. H. (1997). Instrumental variables regression with weak instruments. *Econometrica*, 65(3), 557-586.

Stanton B, Li X, Cottrell L, & Kaljee L. (2001). Early initiation of sex, drug-related risk behaviors, and sensation-seeking among urban, low-income African-American adolescents. *Journal of the National Medical Association*. 93(4):129–138.

Steinberg, L., & Silverberg, S. B. (1986). The vicissitudes of autonomy in early adolescence. *Child Development*, 57(4), 841-851.

Tanaka, S. (2005). Parental leave and child health across OECD countries. *Economic Journal*, 115(501), F7-28.

Thompson, E. R. (2008). Development and validation of an international English Big-Five Mini-Markers. *Personality and Individual Differences*, 45 (6), 542–548.

Thornberry, T. P. (1987). Toward an interactional theory of delinquency. *Criminology*, 25(4), 863-891.

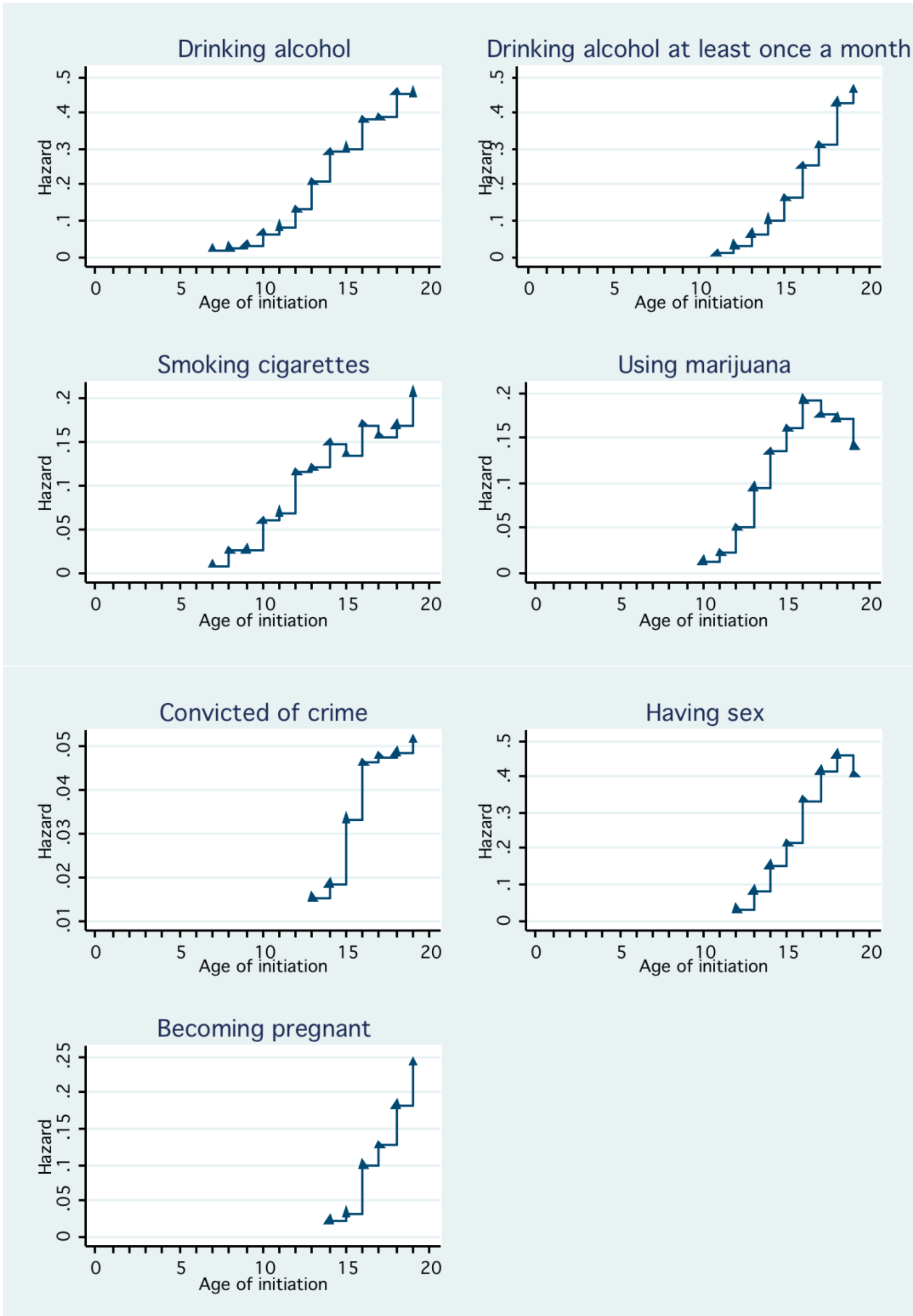
Waldfogel, J., Han, W., & Brooks-Gunn, J. (2002). The effects of early maternal employment on child cognitive development. *Demography*, 39(2), 369-392.

Wang, M., Dishion, T. J., Stormshak, E. A., & Willett, J. B. (2011). Trajectories of family management practices and early adolescent behavioral outcomes. *Developmental Psychology*, 47(5), 1324-1341.

Yancey, A. K., Siegel, J. M., & McDaniel, K. L. (2002). Role models, ethnic identity, and health-risk behaviors in urban adolescents. *Archives of Pediatrics & Adolescent Medicine*, 156(1), 55.

Figure 1.1

Mean Initiation Hazards and Age



Notes: The graph shows the averages of the predicted initiation hazards at each age, calculated using the estimated coefficients from the hazard model, for each risky behavior.

Table 1.1 Variable definition and source

Variable Name	Definition	Source
Drinking Alcohol	Drinking alcohol	NLSY79 Children&YA
Drinking at Least Once A Month	Drinking at least once a month	NLSY79 Children&YA
Smoking Cigarettes	Smoking cigarettes	NLSY79 Children&YA
Using Marijuana	Using marijuana	NLSY79 Children&YA
Being Convicted of Crime	Being convicted of crime	NLSY79 Children&YA
Having Sex	Having sex	NLSY79 Children&YA
Becoming Pregnant	Becoming pregnant	NLSY79 Children&YA
Health Limiting Kind of Work	Whether health limits the kind of work the mother could do	NLSY79
Health Limiting Amount of Work	Whether health limits the amount of work the mother could do	NLSY79
Hours over First 3 Years	The average annual hours worked by the mother in the first 3 years of a child's life	NLSY79
Hours over Subsequent Years	The average annual hours worked by the mother following the first 3 years of a child's life in the year prior to the interview	NLSY79
Hours over Life	The average annual hours worked by the mother over a child's life in the year prior to the interview	NLSY79
PTE over First 3 Years	Percent time employed of the mother in the first 3 years of a child's life	NLSY79
PTE over Subsequent Years	Percent time employed of the mother following the first 3 years of a child's life in the year prior to the interview	NLSY79
PTE over Life	Percent time employed of the mother over a child's life in the year prior to the interview	NLSY79
Always Full-Time	Mother always worked full-time (more than or equal to 1750 hours annually) over a child's life in the year prior to the interview	NLSY79
Always Part-Time	Mother always worked part-time (greater than 0 hours and less than 1750 hours annually) over a child's life in the year prior to the interview	NLSY79
Always Unemployed	Mother was always unemployed (0 hours annually but not out of the labor force) over a child's life in the year prior to the interview	NLSY79
Always Out of the Labor Force	Mother was always out of the labor force (for the entire year) over a child's life in the year prior to the interview	NLSY79
White	Child is white	NLSY79 Children&YA
Black	Child is black	NLSY79 Children&YA
Hispanic	Child is Hispanic	NLSY79 Children&YA
Female	Child is female	NLSY79 Children&YA
First Born	Child is first born	NLSY79 Children&YA
Child's Age	Age of child	NLSY79 Children&YA
Residing with Mother	Whether the usual residence of the child is with the mother	NLSY79 Children
US Born	Whether mother was born in the US	NLSY79
US 14	Whether mother lived in the US at age 14	NLSY79
Urban 14	Whether mother lived in an urban area at age 14	NLSY79
Magazines 14	Whether mother's household member received magazines regularly when she was 14	NLSY79
Newspapers 14	Whether mother's household member received newspapers regularly when she was 14	NLSY79

Library Card 14	Whether mother's household member had a library card when she was 14	NLSY79
Working Female 14	Whether the adult female in mother's household worked when she was 14	NLSY79
Both Parents	Mother lived with both parents when she was 14	NLSY79
Two Adults	Mother lived with two adults when she was 14	NLSY79
One Adult	Mother lived with one adult when she was 14	NLSY79
Other Arrangement	Mother had other living arrangement when she was 14	NLSY79
Grandmother Less than High School	Grandmother's years of schooling is less than 12	NLSY79
Grandmother High School	Grandmother's years of schooling is equal to12	NLSY79
Grandmother Some College Or More	Grandmother's years of schooling is greater than 12	NLSY79
Grandfather Less than High School	Grandfather's years of schooling is less than 12	NLSY79
Grandfather High School	Grandfather's years of schooling is equal to12	NLSY79
Grandfather Some College Or More	Grandfather's years of schooling is greater than 12	NLSY79
No Religious Affiliation	Mother had no religious affiliation in 1979	NLSY79
Protestant	Mother was Protestant in 1979	NLSY79
Baptist	Mother was Baptist in 1979	NLSY79
Episcopalian	Mother was Episcopalian in 1979	NLSY79
Lutheran	Mother was Lutheran in 1979	NLSY79
Methodist	Mother was Methodist in 1979	NLSY79
Presbyterian	Mother was Presbyterian in 1979	NLSY79
Roman Catholic	Mother was Roman Catholic in 1979	NLSY79
Jewish	Mother was Jewish in 1979	NLSY79
Other Religious Affiliation	Mother had other religious affiliation in 1979	NLSY79
Attendance: Not At All	Mother's religious attendance in 1979: not at all	NLSY79
Attendance: Infrequently	Mother's religious attendance in 1979: infrequently	NLSY79
Attendance: Once Per Month	Mother's religious attendance in 1979: once per month	NLSY79
Attendance: 2-3 Times Per Month	Mother's religious attendance in 1979: 2-3 times per month	NLSY79
Attendance: Once Per Week	Mother's religious attendance in 1979: once per week	NLSY79
Attendance: More than Once Per Week	Mother's religious attendance in 1979: more than once per week	NLSY79
AFQT Score	Mother's AFQT score	NLSY79
Mother Less than High School	Mother's years of schooling is less than 12	NLSY79
Mother High School	Mother's years of schooling is equal to12	NLSY79
Mother Some College	Mother's years of schooling is greater than 12 and less than 16	NLSY79
Mother College Or More	Mother's years of schooling is greater than 16	NLSY79
Maternal Age at Birth	Mother's age at child's birth	NLSY79 Children&YA
Single	Mother is single (never married)	NLSY79
Married	Mother is married	NLSY79
Other Marital Status	Mother has other marital status	NLSY79
Mother Regular Drinking	Mother drank at least once a month	NLSY79
Mother Chronic Smoking	Mother smoked at least 100 cigarettes in lifetime	
Mother Chronic Marijuana Use	Mother used marijuana for at least 100 times in lifetime	NLSY79
Mother Cocaine Use	Mother ever used cocaine	NLSY79
Mother Illegal Activity	Mother ever charged with illegal activity before	NLSY79

	1980	
Mother Obese	Mother is obese	NLSY79
Father Figure	Whether a father figure (spouse or partner of mother) is in the household	NLSY79 Children
Number of Children	Number of children in the household	NLSY79
Number of Adults	Number of adults in the household	NLSY79
Age of Youngest Child	Age of youngest child in the household	NLSY79
Average Income over Life	Average family income from child's birth to the year prior to the interview (in units of a thousand, $2001)	NLSY79
Urban	Mother resides in an urban area	NLSY79
Northeast	Mother's region of residence is northeast	NLSY79
North Central	Mother's region of residence is north central	NLSY79
South	Mother's region of residence is south	NLSY79
West	Mother's region of residence is west	NLSY79
Day Shift	Mother's usual shift is a regular day shift	NLSY79
Non-Day Shift	Mother does not have a regular day shift	NLSY79
Child Working for Pay	Child or adolescent works at any paid job	NLSY79 Children&YA

Table 1.2 Summary statistics by mother's employment status in the last calendar year

	Mother not employed in the last calendar year		Mother employed in the last calendar year		TOM
	Mean	Std. Dev.	Mean	Std. Dev.	
Risky behavior initiation					
Drinking Alcohol	0.11	0.31	0.14	0.35	*
Drinking at Least Once A Month	0.13	0.33	0.15	0.36	*
Smoking Cigarettes	0.07	0.26	0.09	0.28	*
Using Marijuana	0.09	0.29	0.10	0.30	
Being Convicted of Crime	0.04	0.20	0.03	0.18	*
Having Sex	0.19	0.39	0.22	0.41	*
Becoming Pregnant	0.13	0.33	0.10	0.30	*
Average annual hours worked by the mother					
Hours over First 3 Years	304	538	943	821	*
Hours over Subsequent Years	242	400	1383	722	*
Hours over Life	269	397	1241	684	*
Percent time employed of the mother					
PTE over First 3 Years	0.34	0.39	0.70	0.39	*
PTE over Subsequent Years	0.24	0.28	0.86	0.23	*
PTE over Life	0.28	0.28	0.80	0.25	*
Employment status of the mother					
Always Full-Time	0.00	0.00	0.08	0.28	*
Always Part-Time	0.00	0.00	0.04	0.20	*
Always unemployed	0.00	0.02	0.00	0.00	*
Always Out of the Labor Force	0.19	0.39	0.00	0.00	*
Control variables					*
White	0.36	0.48	0.45	0.50	*
Black	0.36	0.48	0.34	0.47	*
Hispanic	0.28	0.45	0.21	0.41	*
Female	0.49	0.50	0.51	0.50	*
First Born	0.40	0.49	0.44	0.50	*
Child's Age	10.67	3.10	11.19	3.21	*
Residing with Mother	0.93	0.25	0.94	0.23	*
US Born	0.90	0.30	0.93	0.25	*
US 14	0.96	0.20	0.97	0.16	*
Urban 14	0.83	0.38	0.79	0.41	*
Magazines 14	0.42	0.49	0.51	0.50	*
Newspapers 14	0.62	0.49	0.72	0.45	*
Library Card 14	0.63	0.48	0.70	0.46	*
Working Female 14	0.45	0.50	0.54	0.50	*
Both Parents	0.68	0.47	0.74	0.44	*
Two Adults	0.06	0.23	0.04	0.20	*
One Adult	0.26	0.44	0.21	0.41	*
Other Arrangement	0.00	0.06	0.00	0.05	
Grandmother's Years of Schooling	9.61	3.52	10.40	3.20	*
Grandfather's Years of Schooling	9.43	4.30	10.25	3.96	*
No Religious Affiliation	0.09	0.28	0.08	0.28	
Protestant	0.06	0.23	0.04	0.21	*
Baptist	0.27	0.44	0.29	0.45	*
Episcopalian	0.01	0.08	0.01	0.10	

Lutheran	0.02	0.12	0.04	0.20	*
Methodist	0.05	0.23	0.07	0.25	*
Presbyterian	0.01	0.09	0.02	0.14	*
Roman Catholic	0.35	0.48	0.32	0.46	*
Jewish	0.00	0.05	0.01	0.07	*
Other Religious Affiliation	0.14	0.35	0.13	0.33	*
Attendance: Not At All	0.15	0.36	0.13	0.34	*
Attendance: Infrequently	0.24	0.43	0.24	0.43	
Attendance: Once Per Month	0.10	0.30	0.09	0.29	*
Attendance: 2-3 Times Per Month	0.14	0.35	0.15	0.36	*
Attendance: Once Per Week	0.24	0.43	0.26	0.44	*
Attendance: More than Once Per Week	0.13	0.33	0.12	0.33	
AFQT Score	26.98	26.73	37.26	26.70	*
Maternal Years of Schooling	11.70	2.42	12.80	2.19	*
Maternal Age at Birth	23.51	4.24	24.22	4.22	*
Single	0.25	0.43	0.11	0.32	*
Married	0.52	0.50	0.61	0.49	*
Other Marital Status	0.24	0.42	0.28	0.45	*
Mother Regular Drinking	0.40	0.49	0.48	0.50	*
Mother Chronic Marijuana Use	0.24	0.43	0.25	0.43	
Mother Cocaine Use	0.22	0.41	0.21	0.41	
Mother Illegal Activity	0.04	0.21	0.03	0.17	*
Mother Obese	0.34	0.47	0.30	0.46	*
Father Figure	0.61	0.49	0.67	0.47	*
Number of Children	2.95	1.47	2.42	1.13	*
Number of Adults	1.83	1.03	1.83	0.82	
Age of Youngest Child	6.21	4.22	7.93	4.17	*
Average Income over Life	36.69	60.75	45.11	44.64	*
Urban	0.76	0.43	0.76	0.43	
Northeast	0.17	0.38	0.13	0.34	*
North Central	0.24	0.43	0.25	0.43	
South	0.37	0.48	0.43	0.49	*
West	0.22	0.41	0.19	0.39	*

Notes: The last column shows the results from the two-tailed T-tests of means (TOM). * denotes a p-value less than 0.01.

Table 1.3 *Hours over Life* and adolescent risky behavior initiation

	Drinking alcohol	Drinking at least once a month	Smoking cigarettes	Using Marijuana	Convicted of crime	Having sex	Becoming pregnant
Discrete Time Piecewise-Constant Hazard Model							
Hours	0.19339***	0.16950***	0.14771***	0.09694*	-0.05962	0.21457***	0.00659
	(0.03817)	(0.04888)	(0.04450)	(0.05178)	(0.09330)	(0.04772)	(0.11181)
	[0.01366]	[0.01165]	[0.00700]	[0.00509]	[-0.00133]	[0.02404]	[0.00027]
Income	0.00081*	0.00075	-0.00126	-0.00020	-0.00306	-0.00261***	-0.01042***
	(0.00048)	(0.00072)	(0.00088)	(0.00079)	(0.00242)	(0.00093)	(0.00399)
	[0.00006]	[0.00005]	[-0.00006]	[-0.00001]	[-0.00007]	[-0.00029]	[-0.00043]
Age=7	-2.68008***		-2.03925***				
	(0.39161)		(0.46214)				
Age=8	-2.46445***		-0.94093**				
	(0.39671)		(0.41337)				
Age=9	-2.13257***		-0.84939**				
	(0.38299)		(0.40909)				
Age=10	-1.36591***		0.03316	-1.68595***			
	(0.37928)		(0.40241)	(0.47189)			
Age=11	-1.08688***	-4.00392***	0.22688	-1.03717**			
	(0.38122)	(0.46549)	(0.39935)	(0.45290)			
Age=12	-0.52227	-2.72751***	0.83334**	-0.14275		-0.67020	
	(0.37729)	(0.44701)	(0.39921)	(0.44253)		(0.44423)	
Age=13	0.09727	-1.97997***	0.96895**	0.68305	-2.26707***	0.53324	
	(0.37883)	(0.44065)	(0.40056)	(0.44097)	(0.72218)	(0.43204)	
Age=14	0.57758	-1.43465***	1.25940***	1.17144***	-2.10670***	1.38519***	-1.34045
	(0.38011)	(0.43256)	(0.40546)	(0.44517)	(0.73955)	(0.43047)	(0.92903)
Age=15	0.69461*	-0.84803*	1.23708***	1.54331***	-1.44728*	2.03293***	-0.75399
	(0.38538)	(0.43908)	(0.41644)	(0.44733)	(0.74520)	(0.44118)	(0.96638)
Age=16	1.09111***	-0.28425	1.54264***	1.83046***	-1.12019	2.79074***	0.53172
	(0.38641)	(0.43912)	(0.42002)	(0.44515)	(0.74041)	(0.44198)	(0.94217)
Age=17	1.17941***	0.04631	1.44073***	1.76688***	-1.09392	3.26498***	0.88312
	(0.38939)	(0.44272)	(0.42625)	(0.45371)	(0.74423)	(0.44852)	(0.95636)
Age=18	1.51683***	0.60734	1.56754***	1.78767***	-1.12686	3.56322***	1.35555
	(0.39902)	(0.44338)	(0.43124)	(0.45430)	(0.75894)	(0.44967)	(0.95522)
Age=19	1.52430***	0.79306*	1.78288***	1.56614***	-1.15290	3.31692***	1.76310*
	(0.40353)	(0.44476)	(0.43436)	(0.45783)	(0.76966)	(0.45699)	(0.96209)
N	19793	12787	21722	15674	10859	10541	3667
L-Likelihood	-6145.162	-4087.257	-5364.359	-4206.017	-1474.696	-4177.032	-891.879
Linear Probability Model							
Hours	0.01730***	0.01612***	0.00983***	0.00752**	-0.00213	0.02361***	-0.00930
	(0.00346)	(0.00461)	(0.00287)	(0.00371)	(0.00273)	(0.00563)	(0.00707)
Income	0.00007*	0.00006	-0.00005	-0.00002	-0.00005	-0.00021***	-0.00017**
	(0.00004)	(0.00007)	(0.00004)	(0.00005)	(0.00004)	(0.00007)	(0.00008)
N	19793	12787	21722	15674	10912	10541	3694
R-squared	0.275	0.290	0.144	0.173	0.061	0.348	0.204
Instrumental variable model							
Hours	0.06607	0.16332***	0.01353	0.05375*	0.01224	0.10560**	-0.00423
	(0.04516)	(0.04452)	(0.03488)	(0.03251)	(0.02166)	(0.04734)	(0.06111)

Income	0.00005	-0.00003	-0.00005	-0.00004	-0.00005	-0.00022***	-0.00015*
	(0.00005)	(0.00011)	(0.00004)	(0.00006)	(0.00004)	(0.00008)	(0.00008)
N	18941	12204	20834	14993	10402	10079	3518
Partial R^2	0.008	0.014	0.008	0.014	0.018	0.016	0.015
First stage F	21.00	29.40	20.71	31.08	33.74	32.38	16.15
Hansen's J	0.799	0.121	0.094	0.835	0.639	0.499	0.027
DWH	0.280	0.001	0.885	0.160	0.501	0.096	0.867

Notes:
1. Hours stands for *Hours over Life*.
2. Income stands for *Average Income over Life*.
3. See text for the other control variables used.
4. The marginal effects at the mean of each control variable are displayed in square brackets.
5. The reported statistics include: number of observations, partial R-squared on the excluded instruments, first stage F statistics on the excluded instruments, the p-value for the Hansen’s J statistic, and the p-value for the DWH test.
6. L-likelihood stands for log likelihood.
7. Robust standard errors clustered at the household level are displayed in parentheses. * 10%, ** 5%, *** 1%.

Table 1.4 *PTE over Life* and adolescent risky behavior initiation

	Drinking alcohol	Drinking at least once a month	Smoking cigarettes	Using Marijuana	Convicted of crime	Having sex	Becoming pregnant
Discrete Time Piecewise-Constant Hazard Model							
PTE	0.35869***	0.28459**	0.27131***	0.19150	-0.09147	0.34615***	-0.21068
	(0.08967)	(0.11434)	(0.10493)	(0.12126)	(0.19915)	(0.11409)	(0.25788)
	[0.02538]	[0.01959]	[0.01287]	[0.01006]	[-0.00204]	[0.03890]	[-0.00873]
Income	0.00090*	0.00086	-0.00115	-0.00014	-0.00313	-0.00243***	-0.00988**
	(0.00047)	(0.00070)	(0.00087)	(0.00078)	(0.00246)	(0.00091)	(0.00390)
	[0.00006]	[0.00006]	[-0.00005]	[-0.00001]	[-0.00007]	[-0.00027]	[-0.00041]
Age=7	-2.79198***		-2.13752***				
	(0.39541)		(0.46827)				
Age=8	-2.57171***		-1.03568**				
	(0.40178)		(0.41949)				
Age=9	-2.23913***		-0.94298**				
	(0.38782)		(0.41559)				
Age=10	-1.47262***		-0.06129	-1.75331***			
	(0.38439)		(0.40991)	(0.47756)			
Age=11	-1.19086***	-4.07538***	0.13501	-1.10315**			
	(0.38641)	(0.47096)	(0.40636)	(0.45740)			
Age=12	-0.62785	-2.79997***	0.74015*	-0.20877		-0.76256*	
	(0.38256)	(0.45153)	(0.40557)	(0.44740)		(0.44834)	
Age=13	-0.00891	-2.05094***	0.87666**	0.61742	-2.23851***	0.44288	
	(0.38377)	(0.44603)	(0.40702)	(0.44613)	(0.72889)	(0.43662)	
Age=14	0.46985	-1.50467***	1.16618***	1.10583**	-2.07845***	1.29648***	-1.21086
	(0.38545)	(0.43764)	(0.41208)	(0.45025)	(0.74568)	(0.43499)	(0.93401)
Age=15	0.58846	-0.91834**	1.14746***	1.47736***	-1.41859*	1.94425***	-0.62951
	(0.39019)	(0.44438)	(0.42297)	(0.45212)	(0.75165)	(0.44553)	(0.97128)
Age=16	0.97966**	-0.35758	1.44610***	1.76289***	-1.09047	2.69849***	0.66299
	(0.39141)	(0.44501)	(0.42646)	(0.45039)	(0.74848)	(0.44636)	(0.94918)
Age=17	1.07169***	-0.02689	1.35003***	1.70084***	-1.06459	3.17275***	1.00944
	(0.39378)	(0.44854)	(0.43248)	(0.45906)	(0.75080)	(0.45248)	(0.96279)
Age=18	1.40842***	0.53215	1.46968***	1.71914***	-1.09660	3.46621***	1.48595
	(0.40350)	(0.44913)	(0.43748)	(0.45950)	(0.76702)	(0.45257)	(0.96234)
Age=19	1.41018***	0.71804	1.69226***	1.50130***	-1.12285	3.22330***	1.89018*
	(0.40780)	(0.45025)	(0.44048)	(0.46299)	(0.77816)	(0.46032)	(0.96866)
N	19793	12787	21722	15674	10859	10541	3667
L-Likelihood	-6150.234	-4090.599	-5366.761	-4206.630	-1474.814	-4183.250	-891.536
Linear Probability Model							
PTE	0.02992***	0.02595**	0.01749***	0.01520*	-0.00414	0.03835***	-0.02936*
	(0.00757)	(0.01026)	(0.00631)	(0.00838)	(0.00622)	(0.01308)	(0.01755)
Income	0.00008*	0.00007	-0.00005	-0.00001	-0.00005	-0.00020***	-0.00017**
	(0.00004)	(0.00007)	(0.00004)	(0.00005)	(0.00004)	(0.00007)	(0.00008)
N	19793	12787	21722	15674	10912	10541	3694
R-squared	0.274	0.289	0.144	0.173	0.061	0.348	0.204
Instrumental variable model							
PTE	0.13327	0.32640***	0.02472	0.10791*	0.02292	0.21847**	-0.03469
	(0.09015)	(0.08602)	(0.06948)	(0.06364)	(0.04432)	(0.09613)	(0.13063)

Income	0.00008*	0.00008	-0.00004	-0.00001	-0.00004	-0.00016*	-0.00016**
	(0.00005)	(0.00011)	(0.00004)	(0.00006)	(0.00004)	(0.00008)	(0.00007)
N	18941	12204	20834	14993	10402	10079	3518
Partial R^2	0.010	0.019	0.011	0.019	0.024	0.022	0.019
First stage F	20.36	26.34	20.67	27.93	28.02	25.96	13.87
Hansen's J	0.793	0.200	0.091	0.964	0.607	0.581	0.026
DWH	0.253	0.000	0.900	0.155	0.557	0.072	0.915

Notes:
1. PTE stands for *PTE over Life*.
2. See notes 2 to 7 below Table 1.3.

Table 1.5 Employment status over a child's life and adolescent risky behavior initiation

	Drinking alcohol	Drinking at least once a month	Smoking cigarettes	Using Marijuana	Convicted of crime	Having sex	Becoming pregnant
Discrete Time Piecewise-Constant Hazard Model							
Always FT	0.26332**	-0.00044	0.07524	0.14629	-0.08163	0.20083	0.16855
	(0.10446)	(0.14465)	(0.12525)	(0.14436)	(0.32421)	(0.13346)	(0.36468)
	[0.02073]	[-0.00003]	[0.00366]	[0.00811]	[-0.00176]	[0.02414]	[0.00765]
Always PT	-0.17924	-0.24374	-0.30036	-0.30191	-0.27458	-0.46096**	
	(0.16365)	(0.22911)	(0.18310)	(0.26388)	(0.55124)	(0.21872)	
	[-0.01195]	[-0.01534]	[-0.01255]	[-0.01393]	[-0.00542]	[-0.04403]	
Always OLF	-0.04391	-0.02410	-0.47108**	-0.17307	-0.45347	-0.35108*	0.46023
	(0.17325)	(0.21863)	(0.23329)	(0.24590)	(0.51322)	(0.20002)	(0.42409)
	[-0.00309]	[-0.00166]	[-0.01839]	[-0.00843]	[-0.00829]	[-0.03492]	[0.02397]
Income	0.00090*	0.00081	-0.00111	-0.00024	-0.00303	-0.00243***	-0.01022**
	(0.00049)	(0.00070)	(0.00088)	(0.00079)	(0.00251)	(0.00090)	(0.00402)
	[0.00006]	[0.00006]	[-0.00005]	[-0.00001]	[-0.00007]	[-0.00027]	[-0.00043]
Age=7	-2.52072***		-2.07604***				
	(0.38492)		(0.45486)				
Age=8	-2.29541***		-0.94566**				
	(0.38816)		(0.40882)				
Age=9	-1.94612***		-0.87071**				
	(0.37612)		(0.40690)				
Age=10	-1.19699***		0.01912	-1.67995***			
	(0.36988)		(0.39585)	(0.46677)			
Age=11	-0.91481**	-3.94298***	0.19408	-1.01441**			
	(0.37425)	(0.46238)	(0.39343)	(0.45519)			
Age=12	-0.33500	-2.70221***	0.82200**	-0.12486		-0.51732	
	(0.36988)	(0.44587)	(0.39530)	(0.44347)		(0.45331)	
Age=13	0.27129	-1.91841***	0.95359**	0.69756	-2.20855***	0.66976	
	(0.37194)	(0.43719)	(0.39689)	(0.43878)	(0.73165)	(0.43938)	
Age=14	0.75278**	-1.37851***	1.24598***	1.19116***	-2.01851***	1.52721***	-1.39875
	(0.37386)	(0.42856)	(0.40102)	(0.44331)	(0.74340)	(0.43851)	(0.93245)
Age=15	0.87543**	-0.79322*	1.21625***	1.53512***	-1.35457*	2.15991***	-0.73851
	(0.37960)	(0.43468)	(0.41204)	(0.44656)	(0.75096)	(0.44897)	(0.96282)
Age=16	1.26307***	-0.23850	1.53720***	1.83378***	-1.03136	2.94340***	0.51276
	(0.38071)	(0.43537)	(0.41407)	(0.44430)	(0.74338)	(0.45097)	(0.93321)
Age=17	1.36310***	0.09323	1.43503***	1.75318***	-1.00685	3.40376***	0.87688
	(0.38357)	(0.43904)	(0.42219)	(0.45331)	(0.75087)	(0.45761)	(0.94956)
Age=18	1.70268***	0.67087	1.55669***	1.76821***	-1.06270	3.71650***	1.36744
	(0.39222)	(0.43911)	(0.42536)	(0.45239)	(0.76437)	(0.45905)	(0.94507)
Age=19	1.69894***	0.83265*	1.77112***	1.54994***	-1.09476	3.45789***	1.81154*
	(0.39801)	(0.44152)	(0.42885)	(0.45718)	(0.77493)	(0.46621)	(0.95144)
N	19414	12578	21322	15421	10683	10368	3540
L-Likelihood	-6061.992	-4041.326	-5268.636	-4135.019	-1459.256	-4117.504	-871.257

Notes:

1. Always FT (always PT) [always OLF] stands for *Always Full-time* (*Always Part-time*) [*Always Out of the Labor Force*].
2. See notes 2 to 7 below Table 1.3.

Table 1.6 Hours in different stages of a child's life and adolescent risky behavior initiation

	Drinking alcohol	Drinking at least once a month	Smoking cigarettes	Using Marijuana	Convicted of crime	Having sex	Becoming pregnant
***Hours over First 3 Years* and *Hours over Subsequent Years*, entered together**							
First 3 years	0.013	0.044	-0.004	-0.009	0.160*	0.006	-0.088
	(0.039)	(0.049)	(0.044)	(0.049)	(0.095)	(0.048)	(0.109)
	[0.001]	[0.003]	[-0.000]	[-0.000]	[0.004]	[0.001]	[-0.004]
	{1.904}	{1.786}	{1.877}	{1.808}	{1.758}	{1.779}	{1.868}
Subsequent	0.180***	0.120**	0.150***	0.102*	-0.178	0.197***	0.080
years	(0.042)	(0.054)	(0.048)	(0.055)	(0.111)	(0.052)	(0.114)
	[0.013]	[0.008]	[0.007]	[0.005]	[-0.004]	[0.022]	[0.003]
	{1.943}	{1.846}	{1.926}	{1.868}	{1.821}	{1.838}	{1.929}
Correlation	0.651	0.628	0.647	0.632	0.621	0.625	0.622
Joint F-test	0.000	0.004	0.002	0.120	0.183	0.000	0.675
N	19660	12741	21585	15605	10833	10511	3660
L-likelihood	-6060.908	-4051.567	-5287.413	-4172.451	-1472.030	-4167.901	-891.333
***Hours over First 3 Years* and *Hours over Subsequent Years*, entered separately**							
First 3 years	0.109***	0.103**	0.074**	0.042	0.074	0.101**	-0.054
	(0.033)	(0.041)	(0.038)	(0.043)	(0.077)	(0.041)	(0.097)
	[0.008]	[0.007]	[0.003]	[0.002]	[0.002]	[0.011]	[-0.002]
N	19662	12741	21587	15605	10833	10511	3660
L-likelihood	-6070.883	-4054.426	-5292.951	-4174.494	-1473.715	-4176.120	-891.565
Subsequent	0.183***	0.145***	0.145***	0.097**	-0.091	0.203***	0.036
years	(0.035)	(0.045)	(0.041)	(0.048)	(0.089)	(0.044)	(0.102)
	[0.013]	[0.010]	[0.007]	[0.005]	[-0.002]	[0.023]	[0.001]
N	19714	12776	21641	15651	10859	10541	3667
L-likelihood	-6082.468	-4059.232	-5299.068	-4176.072	-1474.308	-4176.502	-891.821

Notes:
1. The marginal effects at the mean of each control variable are displayed in square brackets.
2. The variance inflation factors are displayed in curly brackets.
3. The correlation between and the p-value of the joint F-test on *Hours over First 3 Years* and *Hours over Subsequent Years* are reported.
4. N stands for number of observations.
5. L-likelihood stands for log likelihood.
6. Robust standard errors clustered at the household level are displayed in parentheses. * 10%, ** 5%, *** 1%.

Table 1.7 *Hours over Life* and adolescent risky behavior initiation, by maternal education

	High School or Less			At Least Some College		
	Hazard	LP	IV	Hazard	LP	IV
Drinking Alcohol	0.124**	0.010**	0.065	0.269***	0.024***	0.030
	(0.053)	(0.005)	(0.059)	(0.056)	(0.005)	(0.061)
	[0.009]			[0.018]		
N	11657	11657	11030	8136	8136	7911
Partial R-squared			0.009			0.010
1st Stage F Statistics			13.29			10.14
Hansen's J Stat p-value			0.193			0.096
DWH test p-value			0.347			0.824
Drinking at Least Once A Month	0.123*	0.011*	0.122**	0.240***	0.021***	0.187***
	(0.065)	(0.006)	(0.050)	(0.076)	(0.007)	(0.069)
	[0.009]			[0.015]		
N	7322	7322	6909	5465	5465	5295
Partial R-squared			0.016			0.016
1st Stage F Statistics			17.44			13.03
Hansen's J Stat p-value			0.122			0.808
DWH test p-value			0.035			0.007
Smoking Cigarettes	0.105*	0.007	0.023	0.231***	0.012***	-0.002
	(0.057)	(0.004)	(0.046)	(0.078)	(0.004)	(0.047)
	[0.005]			[0.009]		
N	12661	12661	12015	9061	9061	8819
Partial R-squared			0.009			0.010
1st Stage F Statistics			12.45			10.64
Hansen's J Stat p-value			0.613			0.075
DWH test p-value			0.756			0.910
Using Marijuana	0.049	0.004	0.050	0.143*	0.008	0.043
	(0.069)	(0.005)	(0.041)	(0.085)	(0.005)	(0.049)
	[0.003]			[0.006]		
N	8916	8916	8421	6758	6758	6572
Partial R-squared			0.018			0.012
1st Stage F Statistics			21.33			11.32
Hansen's J Stat p-value			0.118			0.364
DWH test p-value			0.215			0.430
Being Convicted of Crime	-0.041	-0.002	0.013	-0.071	-0.002	0.009
	(0.120)	(0.004)	(0.031)	(0.173)	(0.003)	(0.025)
	[-0.001]			[-0.001]		
N	6066	6066	5711	4752	4846	4691
Partial R-squared			0.021			0.018
1st Stage F Statistics			18.84			14.97
Hansen's J Stat p-value			0.359			0.963
DWH test p-value			0.617			0.686

Having Sex	0.211***	0.026***	0.068	0.238***	0.022***	0.131*
	(0.065)	(0.008)	(0.058)	(0.073)	(0.008)	(0.070)
	[0.028]			[0.020]		
N	5754	5754	5434	4787	4787	4645
Partial R-squared			0.020			0.016
1st Stage F Statistics			19.45			15.06
Hansen's J Stat p-value			0.931			0.343
DWH test p-value			0.483			0.087
Becoming Pregnant	-0.024	-0.007	-0.096	-0.082	-0.015*	0.072
	(0.153)	(0.011)	(0.086)	(0.179)	(0.009)	(0.085)
	[-0.001]			[-0.001]		
N	2003	2003	1870	1649	1691	1648
Partial R-squared			0.022			0.010
1st Stage F Statistics			11.66			6.26
Hansen's J Stat p-value			0.013			0.946
DWH test p-value			0.405			0.250

Notes:
1. The main independent variable is *Hours over Life*.
2. Hazard stands for discrete-time piecewise constant hazard model.
3. LP stands for linear probability model.
4. IV stands for instrumental variable model.
5. The reported statistics include: number of observations, partial R-squared on the excluded instruments, first stage F statistics on the excluded instruments, the p-value for the Hansen's J statistic, and the p-value for the DWH test.
6. Robust standard errors clustered at the household level are displayed in parentheses. * 10%, ** 5%, *** 1%.

Table 1.8 *Hours over Life* and adolescent risky behavior initiation, by maternal marital status

	Single mothers			Married mothers		
	Hazard	LP	IV	Hazard	LP	IV
Drinking Alcohol	0.209***	0.019***	0.066	0.164***	0.014***	0.056
	(0.069)	(0.007)	(0.055)	(0.048)	(0.004)	(0.075)
	[0.016]			[0.011]		
N	6868	6868	6462	12925	12925	12479
Partial R-squared			0.021			0.005
1st Stage F Statistics			17.91			10.34
Hansen's J Stat p-value			0.598			0.744
DWH test p-value			0.357			0.572
Drinking at Least Once A Month	0.101	0.009	0.156***	0.177***	0.017***	0.149**
	(0.083)	(0.008)	(0.058)	(0.062)	(0.006)	(0.067)
	[0.007]			[0.012]		
N	4621	4621	4310	8166	8166	7894
Partial R-squared			0.027			0.009
1st Stage F Statistics			18.80			17.36
Hansen's J Stat p-value			0.579			0.103
DWH test p-value			0.007			0.048
Smoking Cigarettes	0.058	0.005	0.036	0.165***	0.010***	0.001
	(0.074)	(0.006)	(0.044)	(0.058)	(0.003)	(0.054)
	[0.003]			[0.006]		
N	7491	7491	7079	14231	14231	13755
Partial R-squared			0.020			0.005
1st Stage F Statistics			16.32			11.13
Hansen's J Stat p-value			0.049			0.589
DWH test p-value			0.520			0.909
Using Marijuana	0.078	0.009	0.042	0.108	0.007*	0.063
	(0.085)	(0.008)	(0.047)	(0.067)	(0.004)	(0.049)
	[0.006]			[0.004]		
N	5519	5519	5173	10155	10155	9820
Partial R-squared			0.029			0.008
1st Stage F Statistics			21.70			15.92
Hansen's J Stat p-value			0.639			0.205
DWH test p-value			0.466			0.198
Being Convicted of Crime	-0.150	-0.007	0.012	0.023	0.001	0.033
	(0.150)	(0.006)	(0.031)	(0.133)	(0.003)	(0.031)
	[-0.005]			[0.000]		
N	3961	3961	3692	6861	6951	6710
Partial R-squared			0.035			0.011
1st Stage F Statistics			19.57			15.89
Hansen's J Stat p-value			0.194			0.344
DWH test p-value			0.447			0.290

Having Sex	0.060	0.012	0.038	0.278***	0.028***	0.145**
	(0.081)	(0.012)	(0.067)	(0.060)	(0.006)	(0.064)
	[0.010]			[0.024]		
N	3519	3519	3280	7022	7022	6799
Partial R-squared			0.031			0.012
1st Stage F Statistics			19.38			20.16
Hansen's J Stat p-value			0.855			0.376
DWH test p-value			0.748			0.073
Becoming Pregnant	-0.156	-0.014	-0.103	0.062	-0.005	0.113
	(0.164)	(0.016)	(0.093)	(0.165)	(0.007)	(0.071)
	[-0.012]			[0.001]		
N	1316	1316	1217	2332	2378	2301
Partial R-squared			0.032			0.014
1st Stage F Statistics			9.89			10.52
Hansen's J Stat p-value			0.661			0.048
DWH test p-value			0.330			0.116

Notes:
See notes below Table 1.7.

Table 1.9 ***Hours over Life*** **and adolescent risky behavior initiation, by** ***Family Income over Life***

	Lower 50th percentile			Upper 50th percentile		
	Hazard	LP	IV	Hazard	LP	IV
Drinking Alcohol	0.247***	0.021***	0.057	0.176***	0.015***	0.150
	(0.068)	(0.006)	(0.055)	(0.051)	(0.004)	(0.118)
	[0.018]			[0.012]		
N	7968	7969	7411	11824	11824	11530
Partial R-squared			0.019			0.002
1st Stage F Statistics			25.16			3.06
Hansen's J Stat p-value			0.277			0.683
DWH test p-value			0.492			0.210
Drinking at Least Once A Month	0.178**	0.016*	0.182***	0.158**	0.015**	0.224**
	(0.089)	(0.008)	(0.056)	(0.064)	(0.006)	(0.114)
	[0.010]			[0.011]		
N	4982	4983	4595	7804	7804	7609
Partial R-squared			0.027			0.004
1st Stage F Statistics			23.82			7.16
Hansen's J Stat p-value			0.482			0.173
DWH test p-value			0.003			0.043
Smoking Cigarettes	0.118	0.008	-0.003	0.196***	0.012***	0.082
	(0.077)	(0.006)	(0.047)	(0.060)	(0.003)	(0.074)
	[0.007]			[0.008]		
N	8445	8446	7888	13276	13276	12946
Partial R-squared			0.018			0.003
1st Stage F Statistics			22.11			4.48
Hansen's J Stat p-value			0.008			0.810
DWH test p-value			0.636			0.313
Using Marijuana	0.123	0.011	0.043	0.093	0.007	0.122*
	(0.089)	(0.007)	(0.048)	(0.071)	(0.005)	(0.065)
	[0.008]			[0.004]		
N	5916	5917	5477	9757	9757	9516
Partial R-squared			0.030			0.005
1st Stage F Statistics			28.01			7.16
Hansen's J Stat p-value			0.586			0.407
DWH test p-value			0.497			0.052
Being Convicted of Crime	-0.193	-0.008	0.005	0.106	0.002	0.044
	(0.152)	(0.006)	(0.029)	(0.131)	(0.003)	(0.041)
	[-0.007]			[0.002]		
N	3975	4033	3705	6827	6879	6697
Partial R-squared			0.040			0.006
1st Stage F Statistics			27.39			8.44
Hansen's J Stat p-value			0.371			0.745
DWH test p-value			0.688			0.292

Having Sex	0.041	0.007	0.147**	0.311***	0.032***	0.090
	(0.084)	(0.012)	(0.061)	(0.063)	(0.007)	(0.095)
	[0.006]			[0.027]		
N	3666	3667	3373	6874	6874	6706
Partial R-squared			0.036			0.007
1st Stage F Statistics			25.51			9.47
Hansen's J Stat p-value			0.887			0.393
DWH test p-value			0.022			0.516
Becoming Pregnant	-0.144	-0.015	0.020	0.133	0.002	-0.026
	(0.162)	(0.017)	(0.087)	(0.176)	(0.007)	(0.115)
	[-0.013]			[0.002]		
N	1365	1365	1252	2266	2329	2266
Partial R-squared			0.031			0.007
1st Stage F Statistics			11.77			5.28
Hansen's J Stat p-value			0.221			0.045
DWH test p-value			0.605			0.817

Notes:
See notes below Table 1.7.

Table 1.10 ***Hours over Life*** **and adolescent risky behavior initiation, by child's race/ethnicity**

	Non-white Hazard	LP	IV	White Hazard	LP	IV
Drinking Alcohol	0.151***	0.012**	0.065	0.195***	0.018***	0.101
	(0.057)	(0.005)	(0.058)	(0.055)	(0.005)	(0.075)
	[0.010]			[0.014]		
N	9599	9599	9098	10194	10194	9843
Partial R-squared			0.011			0.006
1st Stage F Statistics			17.14			8.23
Hansen's J Stat p-value			0.202			0.512
DWH test p-value			0.369			0.253
Drinking at Least Once A Month	0.022	0.002	0.129**	0.210***	0.020***	0.249***
	(0.076)	(0.007)	(0.054)	(0.069)	(0.007)	(0.085)
	[0.001]			[0.015]		
N	6263	6263	5927	6524	6524	6277
Partial R-squared			0.020			0.009
1st Stage F Statistics			21.07			13.04
Hansen's J Stat p-value			0.472			0.215
DWH test p-value			0.014			0.002
Smoking Cigarettes	0.028	0.002	0.002	0.189***	0.012***	0.047
	(0.071)	(0.004)	(0.047)	(0.061)	(0.004)	(0.054)
	[0.001]			[0.010]		
N	10659	10659	10126	11063	11063	10708
Partial R-squared			0.010			0.007
1st Stage F Statistics			15.87			8.97
Hansen's J Stat p-value			0.036			0.578
DWH test p-value			0.934			0.494
Using Marijuana	0.081	0.007	-0.023	0.058	0.004	0.158***
	(0.077)	(0.006)	(0.050)	(0.076)	(0.005)	(0.050)
	[0.004]			[0.003]		
N	7456	7456	7062	8218	8218	7931
Partial R-squared			0.018			0.011
1st Stage F Statistics			18.83			16.38
Hansen's J Stat p-value			0.726			0.751
DWH test p-value			0.503			0.001
Being Convicted of Crime	-0.189	-0.005	0.010	0.019	-0.001	0.015
	(0.137)	(0.004)	(0.024)	(0.140)	(0.004)	(0.038)
	[-0.004]			[0.000]		
N	5189	5189	4907	5596	5723	5495
Partial R-squared			0.025			0.014
1st Stage F Statistics			22.60			17.82
Hansen's J Stat p-value			0.869			0.700
DWH test p-value			0.538			0.652

Having Sex	0.134*	0.020*	0.088	0.273***	0.027***	0.147**
	(0.073)	(0.010)	(0.068)	(0.067)	(0.007)	(0.068)
	[0.019]			[0.022]		
N	4681	4681	4429	5860	5860	5650
Partial R-squared			0.021			0.014
1st Stage F Statistics			17.77			19.50
Hansen's J Stat p-value			0.986			0.433
DWH test p-value			0.336			0.077
Becoming Pregnant	-0.089	-0.011	0.002	-0.035	-0.007	0.005
	(0.151)	(0.013)	(0.088)	(0.187)	(0.008)	(0.084)
	[-0.006]			[-0.001]		
N	1728	1732	1636	1920	1962	1882
Partial R-squared			0.018			0.015
1st Stage F Statistics			7.41			10.30
Hansen's J Stat p-value			0.379			0.055
DWH test p-value			0.859			0.770

Notes:
See notes below Table 1.7.

Appendix

The 17 variables regarding parental investment are constructed using the following 3 groups of questions.
(1). In the current or most recent school year, how often do/did your parents do the following? Would you say often, sometimes, rarely, or never?
Check on whether you have done your homework
Help you with your homework
Give you special privileges because of good grades
Limit privileges because of poor grades
Require you to do work or chores around the home
Limit the amount of time you can spend watching TV or playing video games
Limit the amount of time you go out with friends on school nights
(2). In the current or most recent school year, how often have you discussed the following with either or both of your parents or guardians? Would you say often, sometimes, rarely, or never?
Selecting courses or programs at school
School activities or events of particular interest to you
Your grades or report card
Going to college
Community, national or world events
Things that are troubling you
(3). In the current or most recent school year, how many times do/did either your parents or guardians do any of the following? Was it never, once a term, once a month, or more than once a month?
Attend a school meeting
Phone or speak to your teacher or counselor
Attend a school event in which you participated
Act as a volunteer at your school
If the variable comes from the first two groups, we recode the variable to be 1 if the respondent answers “often” or “sometimes”, and 0 if the respondent answers “rarely” or “never”; if the variable comes from the third group, we recode the variable to be 1 if the respondent answers “more than once a month” or “once a month”, and 0 if the respondent answers “once a term” or “never”.

Appendix Table 1.1 Suggestive test for the exogeneity of the instruments

Mother's attributes							
	AFQT Score	*Years of Schooling*	*Mother Chronic Smoking*	*Mother Regular Drinking*	*Mother Chronic Marijuana Use*	*Mother Cocaine Use*	*Mother Illegal Activity*
Health Limiting	-1.71156	-0.04957	-0.02260	-0.04773*	-0.04048	-0.03359	-0.00284
Kind of Work	(3.21904)	(0.08905)	(0.04624)	(0.02436)	(0.02702)	(0.02983)	(0.03177)
Health Limiting	1.78199	-0.15166	0.06327	-0.02602	0.04590	0.03966	-0.00313
Amount of Work	(4.35790)	(0.09457)	(0.05279)	(0.02764)	(0.03319)	(0.03596)	(0.03054)
N	6269	65156	15308	55000	26980	26966	5837
R^2	0.488	0.472	0.153	0.129	0.100	0.134	0.058
Joint F-test	0.868	0.047	0.454	0.000	0.283	0.482	0.974
Parental Investment Measures							
	Check Whether Homework Is Done	*Help with Homework*	*Reward Good Grades*	*Punish Bad Grades*	*Require to Do Housework*	*Limit TV Or Video Game Time*	*Limit Time Out on School Night*
Health Limiting	-0.00697	-0.02404	-0.03958	-0.01924	0.01659	-0.02731	-0.04377
Kind of Work	(0.04205)	(0.04287)	(0.04213)	(0.04541)	(0.02574)	(0.04017)	(0.04089)
Health limiting	-0.00116	0.02032	0.02891	0.02347	-0.00081	0.06085	0.04057
Amount of Work	(0.04561)	(0.04773)	(0.04783)	(0.04889)	(0.03007)	(0.04424)	(0.04516)
N	5438	5440	5429	5425	5441	5438	5433
R^2	0.027	0.029	0.047	0.025	0.031	0.062	0.018
Joint F-test	0.956	0.854	0.620	0.890	0.473	0.334	0.564
	Discuss Selecting Courses	*Discuss School Activities*	*Discuss Grades*	*Discuss Going to College*	*Discuss Current Events*	*Discuss Troubling Things*	*Attend School Meeting*
Health Limiting	0.02454	-0.07163*	-0.05259*	0.01617	-0.04336	0.00432	-0.03938
Kind of Work	(0.03831)	(0.03796)	(0.02985)	(0.03270)	(0.03998)	(0.03905)	(0.04506)
Health Limiting	-0.04348	0.04915	0.04057	-0.00160	0.07200	0.02826	0.07334
Amount of Work	(0.04185)	(0.03877)	(0.03231)	(0.03610)	(0.04468)	(0.04232)	(0.04992)
N	5435	5436	5437	5436	5434	5440	3432
R^2	0.026	0.038	0.019	0.052	0.044	0.019	0.048
Joint F-test	0.568	0.153	0.197	0.745	0.263	0.435	0.313
	Speak to Teacher	*Attend School Event*	*Volunteer at School*				
Health Limiting	0.02272	-0.06184	-0.06824*				

Kind of Work	(0.04497)	(0.05353)	(0.03568)
Health Limiting	0.00690	0.05309	0.06701*
Amount of Work	(0.05044)	(0.05805)	(0.04053)
N	3437	3441	3443
R^2	0.053	0.091	0.047
Joint F-test	0.649	0.512	0.154

Notes:

1. Each outcome is regressed on the two main explanatory variables: *Health Limiting Kind of Work* and *Health Limiting Amount of Work* in the last year (or after 1994, in the last survey year), and the other regressors (except maternal employment measures).
2. The other regressors are slightly different from those in the baseline regressions. See text for the reason and the list of them.
3. The reported statistics include: number of observations, R-squared, and the p-value of the joint F-test on the two main independent variables.

Appendix Table 1.2 Suggestive test for the exogeneity of the instruments, restricting the sample to single mothers

Parental Investment Measures							
	Check Whether Homework Is Done	*Help with Homework*	*Reward Good Grades*	*Punish Bad Grades*	*Require to Do Housework*	*Limit TV Or Video Game Time*	*Limit Time Out on School Night*
Health Limiting Kind of Work	0.01476	-0.02606	-0.08936	0.00293	0.02989	0.03538	-0.02125
	(0.06698)	(0.06123)	(0.07194)	(0.06897)	(0.02170)	(0.05942)	(0.06803)
Health Limiting Amount of Work	-0.05871	0.07783	0.12553	0.05426	-0.00472	0.01690	0.04175
	(0.07292)	(0.07067)	(0.07778)	(0.07545)	(0.02799)	(0.06429)	(0.08115)
N	1930	1932	1925	1927	1933	1931	1929
R^2	0.053	0.039	0.071	0.049	0.057	0.070	0.037
Joint F-test	0.514	0.468	0.267	0.336	0.257	0.431	0.839
	Discuss Selecting Courses	*Discuss School Activities*	*Discuss Grades*	*Discuss Going to College*	*Discuss Current Events*	*Discuss Troubling Things*	*Attend School Meeting*
Health Limiting Kind of Work	0.07112	-0.03500	0.01934	-0.00070	-0.02518	0.07952	-0.09157*
	(0.05180)	(0.05266)	(0.04586)	(0.05144)	(0.06339)	(0.06067)	(0.05482)
Health Limiting Amount of Work	-0.10588*	0.02157	-0.04143	0.01746	0.04730	-0.06840	0.12298*
	(0.06047)	(0.05660)	(0.05134)	(0.05797)	(0.06662)	(0.06857)	(0.07167)
N	1929	1931	1930	1931	1930	1932	1263
R^2	0.055	0.055	0.035	0.078	0.056	0.042	0.073
Joint F-test	0.211	0.783	0.647	0.906	0.744	0.423	0.187
	Speak to Teacher	*Attend School Event*	*Volunteer at School*				
Health Limiting Kind of Work	0.00837	0.08183	-0.01943				
	(0.06970)	(0.08097)	(0.05297)				
Health Limiting Amount of Work	0.04421	-0.09224	0.03968				
	(0.08610)	(0.09200)	(0.06021)				
N	1262	1264	1266				
R^2	0.075	0.089	0.090				
Joint F-test	0.646	0.576	0.791				

Notes: This table shows the results using the subsample of single mothers. See notes on Appendix Table 1.1.

Appendix Table 1.3 Role model effect high-level summary

	Maternal Education	Maternal Marital Status	Family Income Over Life	Child's Race/Ethnicity
Drinking Alcohol	Yes			Yes
Drinking At Least Once A Month	Yes	Yes		Yes
Smoking Cigarettes	Yes	Yes	Yes	Yes
Marijuana Use	Yes			
Having Sex		Yes	Yes	Yes

Notes: "Yes" means that the coefficient on maternal employment is larger for households with better educated mothers, married parents, higher income or white children.

Appendix Table 1.4 Varying the cutoff between the first few years and the subsequent years

	Drinking Alcohol	*Drinking at Least Once A Month*	*Smoking Cigarettes*	*Using Marijuana*	*Being Convicted of Crime*	*Having Sex*	*Becoming Pregnant*
***Hours over First 1 Year* and *Hours over Subsequent Years*, entered together**							
First 1 year	-0.006	-0.009	-0.059	0.008	0.136	-0.008	-0.039
	(0.035)	(0.046)	(0.041)	(0.047)	(0.086)	(0.043)	(0.097)
	[-0.000]	[-0.001]	[-0.003]	[0.000]	[0.003]	[-0.001]	[-0.002]
Subsequent	0.194***	0.152***	0.188***	0.087	-0.141	0.213***	0.023
years	(0.044)	(0.056)	(0.051)	(0.058)	(0.114)	(0.055)	(0.120)
	[0.014]	[0.010]	[0.009]	[0.005]	[-0.003]	[0.024]	[0.001]
Correlation	0.634	0.603	0.632	0.617	0.607	0.607	0.583
Joint F-test	0.000	0.009	0.001	0.194	0.265	0.000	0.924
N	19138	12429	21025	15229	10582	10253	3573
L-likelihood	-5913.901	-3973.577	-5186.512	-4096.561	-1450.948	-4095.797	-881.899
Hours over First 2 Years* and *Hours over Subsequent Years, entered together							
First 2 years	0.022	0.043	-0.030	-0.008	0.119	-0.000	-0.063
	(0.037)	(0.046)	(0.043)	(0.048)	(0.091)	(0.046)	(0.099)
	[0.002]	[0.003]	[-0.001]	[-0.000]	[0.003]	[-0.000]	[-0.003]
Subsequent	0.174***	0.116**	0.177***	0.101*	-0.137	0.209***	0.054
years	(0.043)	(0.055)	(0.049)	(0.056)	(0.110)	(0.054)	(0.114)
	[0.012]	[0.008]	[0.008]	[0.005]	[-0.003]	[0.024]	[0.002]
Correlation	0.644	0.617	0.641	0.626	0.617	0.619	0.606
Joint F-test	0.000	0.006	0.001	0.142	0.354	0.000	0.795
N	19572	12674	21478	15526	10780	10451	3626
L-likelihood	-6034.580	-4025.309	-5264.123	-4158.403	-1465.666	-4153.648	-887.379
Hours over First 3 Years* and *Hours over Subsequent Years, entered together							
First 3 years	0.013	0.044	-0.004	-0.009	0.160*	0.006	-0.088
	(0.039)	(0.049)	(0.044)	(0.049)	(0.095)	(0.048)	(0.109)
	[0.001]	[0.003]	[-0.000]	[-0.000]	[0.004]	[0.001]	[-0.004]
Subsequent	0.180***	0.120**	0.150***	0.102*	-0.178	0.197***	0.080
years	(0.042)	(0.054)	(0.048)	(0.055)	(0.111)	(0.052)	(0.114)
	[0.013]	[0.008]	[0.007]	[0.005]	[-0.004]	[0.022]	[0.003]
Correlation	0.651	0.628	0.647	0.632	0.621	0.625	0.622
Joint F-test	0.000	0.004	0.002	0.120	0.183	0.000	0.675
N	19660	12741	21585	15605	10833	10511	3660
L-likelihood	-6060.908	-4051.567	-5287.413	-4172.451	-1472.030	-4167.901	-891.333
Hours over First 4 Years* and *Hours over Subsequent Years, entered together							
First 4 years	0.039	0.044	0.018	-0.004	0.141	0.030	-0.049
	(0.040)	(0.050)	(0.047)	(0.051)	(0.100)	(0.049)	(0.116)
	[0.003]	[0.003]	[0.001]	[-0.000]	[0.003]	[0.003]	[-0.002]
Subsequent	0.155***	0.119**	0.135***	0.104*	-0.160	0.179***	0.051
years	(0.041)	(0.052)	(0.048)	(0.053)	(0.111)	(0.050)	(0.116)
	[0.011]	[0.008]	[0.006]	[0.005]	[-0.004]	[0.020]	[0.002]
Correlation	0.655	0.635	0.650	0.636	0.627	0.630	0.627
Joint F-test	0.000	0.003	0.001	0.090	0.288	0.000	0.885
N	19674	12756	21603	15627	10845	10527	3662
L-likelihood	-6070.655	-4054.432	-5289.425	-4174.074	-1472.917	-4173.416	-891.634

***Hours over First 5 Years* and *Hours over Subsequent Years*, entered together**							
First 5 years	0.053	0.047	0.029	-0.009	0.105	0.045	-0.049
	(0.041)	(0.051)	(0.048)	(0.052)	(0.104)	(0.050)	(0.118)
	[0.004]	[0.003]	[0.001]	[-0.000]	[0.002]	[0.005]	[-0.002]
Subsequent	0.136***	0.118**	0.124***	0.110**	-0.141	0.166***	0.058
years	(0.039)	(0.050)	(0.046)	(0.052)	(0.110)	(0.048)	(0.114)
	[0.010]	[0.008]	[0.006]	[0.006]	[-0.003]	[0.019]	[0.002]
Correlation	0.656	0.635	0.652	0.638	0.623	0.632	0.630
Joint F-test	0.000	0.003	0.001	0.066	0.426	0.000	0.864
N	19637	12767	21571	15639	10856	10536	3667
L-likelihood	-6071.552	-4057.935	-5286.918	-4170.327	-1473.643	-4176.161	-891.728
***Hours over First 6 Years* and *Hours over Subsequent Years*, entered together**							
First 6 years	0.086**	0.080	0.045	-0.005	0.097	0.064	-0.065
	(0.042)	(0.052)	(0.048)	(0.053)	(0.104)	(0.050)	(0.118)
	[0.006]	[0.005]	[0.002]	[-0.000]	[0.002]	[0.007]	[-0.003]
Subsequent	0.104***	0.089*	0.111**	0.112**	-0.129	0.151***	0.072
years	(0.038)	(0.049)	(0.045)	(0.050)	(0.109)	(0.047)	(0.109)
	[0.008]	[0.006]	[0.006]	[0.006]	[-0.003]	[0.017]	[0.003]
Correlation	0.648	0.630	0.643	0.632	0.620	0.625	0.622
Joint F-test	0.000	0.003	0.001	0.045	0.484	0.000	0.779
N	18664	12768	20588	15638	10857	10537	3667
L-likelihood	-5963.585	-4058.049	-5247.834	-4169.879	-1473.798	-4176.050	-891.620

Notes:
1. *Hours over First 1 (2, 3, 4, 5, 6) Year(s)* and *Hours over Subsequent Years* (correspondingly) are controlled for in the same regression.
2. The marginal effects at the mean of each control variable are displayed in square brackets.
3. The correlation between and the p-value of the joint F test on *Hours over First 1 (2, 3, 4, 5, 6) Year(s)* and *Hours over Subsequent Years* (correspondingly) are reported.
4. N stands for number of observations.
5. L-likelihood stands for log likelihood.
6. Robust standard errors clustered at the household level are displayed in parentheses. * 10%, ** 5%, *** 1%.

Appendix Table 1.5 Robustness checks

	Drinking Alcohol	*Drinking at Least Once A Month*	*Smoking Cigarettes*	*Using Marijuana*	*Convicted Of Crime*	*Having Sex*	*Becoming Pregnant*
a. Baseline hazard model							
Hours over Life	0.193***	0.170***	0.148***	0.097*	-0.060	0.215***	0.007
	(0.038)	(0.049)	(0.045)	(0.052)	(0.093)	(0.048)	(0.112)
	[0.014]	[0.012]	[0.007]	[0.005]	[-0.001]	[0.024]	[0.000]
N	19793	12787	21722	15674	10859	10541	3667
L-likelihood	-6145.162	-4087.257	-5364.359	-4206.017	-1474.696	-4177.032	-891.879
b. Hazard model with maternal work shifts (*Day Shift* and *Non-Day Shift*) as additional controls							
Hours over Life	0.195***	0.151***	0.105**	0.042	-0.082	0.152***	0.004
	(0.043)	(0.056)	(0.052)	(0.059)	(0.107)	(0.053)	(0.127)
	[0.015]	[0.011]	[0.005]	[0.002]	[-0.002]	[0.018]	[0.000]
N	15837	10682	17478	13038	9268	8915	3164
L-likelihood	-5106.428	-3524.381	-4506.869	-3596.247	-1219.115	-3594.058	-756.074
c. Hazard model with *Child Working for Pay* as an additional control							
Hours over Life	0.200***	0.191***	0.153***	0.128**	-0.044	0.185***	0.022
	(0.045)	(0.050)	(0.052)	(0.054)	(0.098)	(0.050)	(0.117)
	[0.031]	[0.016]	[0.014]	[0.008]	[-0.001]	[0.021]	[0.001]
N	8450	10194	10003	11506	9559	8834	3357
L-likelihood	-4021.119	-3618.838	-3498.985	-3427.136	-1337.515	-3605.186	-822.191
d. Hazard model with top-coded maternal working hours							
Hours over Life	0.151***	0.105***	0.097***	0.104***	-0.056	0.154***	0.091
	(0.029)	(0.036)	(0.033)	(0.037)	(0.067)	(0.036)	(0.079)
	[0.010]	[0.008]	[0.005]	[0.006]	[-0.001]	[0.019]	[0.005]
N	35987	23335	39366	28601	20324	18925	6771
L-Likelihood	-11162.106	-7674.152	-9934.453	-7887.794	-2850.602	-7838.827	-1860.857
e. Hazard model with interpolated independent variables							
Hours over Life	0.234***	0.170**	0.141**	0.100	-0.057	0.263***	0.059
	(0.065)	(0.083)	(0.065)	(0.083)	(0.153)	(0.076)	(0.176)
	[0.017]	[0.010]	[0.007]	[0.005]	[-0.001]	[0.031]	[0.003]
N	7999	5312	8670	6308	4444	4195	1469
L-Likelihood	-2468.419	-1622.757	-2251.224	-1696.240	-672.067	-1709.936	-401.738
f. Using the sample for which health limits have variations (Hazard Model)							
Hours over Life	0.234***	0.170**	0.141**	0.100	-0.057	0.263***	0.059
	(0.065)	(0.083)	(0.065)	(0.083)	(0.153)	(0.076)	(0.176)
	[0.017]	[0.010]	[0.007]	[0.005]	[-0.001]	[0.031]	[0.003]
N	7999	5312	8670	6308	4444	4195	1469
L-Likelihood	-2468.419	-1622.757	-2251.224	-1696.240	-672.067	-1709.936	-401.738
g. Using the sample for which health limits have variations (LPM)							
Hours over Life	0.021***	0.017**	0.010**	0.008	-0.004	0.027***	-0.005
	(0.006)	(0.008)	(0.005)	(0.006)	(0.005)	(0.009)	(0.012)
N	7999	5312	8670	6308	4463	4195	1481
R-squared	0.258	0.288	0.156	0.188	0.075	0.359	0.237
h. Using the sample for which health limits have variations (IV)							
Hours over Life	0.043	0.188***	0.049	0.075	0.037	0.126**	0.042
	(0.062)	(0.066)	(0.050)	(0.046)	(0.030)	(0.063)	(0.077)
N	7370	4872	8025	5810	4083	3852	1347

Partial R-squared	0.011	0.017	0.011	0.020	0.026	0.024	0.028
1st Stage F stat	11.30	14.48	11.26	17.15	19.72	18.87	11.87
Hansen's J p-value	0.617	0.433	0.164	0.276	0.331	0.834	0.153
DWH test p-value	0.737	0.006	0.446	0.139	0.137	0.120	0.517

Notes:
1. The main independent variable is *Hours over Life*
2. N stands for number of observations.
3. L-likelihood stands for log likelihood.
4. The reported statistics for the results in panel h include: number of observations, partial R-squared on the excluded instruments, first stage F statistics on the excluded instruments, the p-value for the Hansen's J statistic, and the p-value for the DWH test.
5. Robust standard errors clustered at the household level are displayed in parentheses. * 10%, ** 5%, *** 1%.

Chapter 2

In Sickness and in Health: Same-Sex Marriage Laws and Sexually Transmitted Infections

Andrew M. Francis, Hugo M. Mialon, and Handie Peng*

This paper analyzes the relationship between same-sex marriage laws and sexually transmitted infections in the United States using state-level data from 1981 to 2008. We hypothesize that same-sex marriage laws may directly affect risky homosexual behavior; may affect or mirror social attitudes toward gays, which in turn may affect homosexual behavior; and may affect or mirror attitudes toward non-marital sex, which may affect risky heterosexual behavior. Our findings may be summarized as follows. Laws banning same-sex marriage are unrelated to gonorrhea rates, which are a proxy for risky heterosexual behavior. They are more closely associated with syphilis rates, which are a proxy for risky homosexual behavior. However, these estimates are smaller and less statistically significant when we exclude California, the state with the largest gay population. Also, laws permitting same-sex marriage are unrelated to gonorrhea or syphilis, but variation in these laws is insufficient to yield precise estimates. In sum, the findings point to a modest positive association—if any at all—between same-sex marriage bans and syphilis.

* Andrew Francis, Hugo Mialon, and Handie Peng, Department of Economics, Emory University, Atlanta, GA 30322-2240 (andrew.francis@emory.edu, hmialon@emory.edu, hpeng4@emory.edu). We are extremely grateful to the Editor, Joanna Coast, two anonymous referees, Dave Frisvold, Mike Grossman, Dan Hamermesh, Ted Joyce, Bob Kaestner, Darius Lakdawalla, Sara Markowitz, Josh Robinson, and seminar participants at the 2011 NBER Health Economics Spring Meeting for wonderfully helpful comments.

I. Introduction

In this paper, we investigate the relationship between same-sex marriage laws and sexually transmitted infections. Today, most U.S. states prohibit marriage between two people of the same sex. For example, Michigan's same-sex marriage ban reads:

> "Marriage is inherently a unique relationship between a man and a woman. As a matter of public policy, this state has a special interest in encouraging, supporting, and protecting that unique relationship in order to promote, among other goals, the stability and welfare of society and its children. A marriage contracted between individuals of the same sex is invalid in this state." (Michigan Compiled Laws Annotated § 551.1)

We hypothesize that same-sex marriage laws may affect homosexual behavior directly, affect it indirectly by changing social attitudes toward gays, or simply reflect changing social attitudes towards gays, which in turn may affect homosexual behavior. We also hypothesize that same-sex marriage laws may affect heterosexual behavior indirectly by changing attitudes toward non-marital sex or reflect changing attitudes toward non-marital sex, which in turn may affect heterosexual behavior. Using a state-level panel dataset from 1981 to 2008, we estimate the connections between same-sex marriage laws and syphilis and gonorrhea. In the paper, syphilis is a proxy for risky homosexual behavior, since syphilis is particularly concentrated among men who have sex with men. 64% of syphilis cases in the US and 79.3% of syphilis cases in California are attributable to men who have sex with men (CDC, 2010a,b; California Department of Public Health, 2010). Gonorrhea is a proxy for risky heterosexual behavior, since more than 90% of gonorrhea cases are attributable to sex between men and women (CDC, 1997).

We summarize our findings as follows. First, there is no association between same-sex marriage bans and gonorrhea. Bans are more closely associated with syphilis.

Second, of the different legal measures, bans on both same-sex marriage and civil union are most consistently associated with syphilis. These are the most restrictive type of ban because they deny same-sex couples access not only to marriage but also to any legal status analogous to marriage. Third, the estimates for same-sex marriage bans are smaller and less significant when we exclude California, the state with the largest gay population. Fourth, there is no evidence of any association between laws permitting same-sex marriage and syphilis or gonorrhea, although there is insufficient variation in these laws to yield precise estimates. Fifth, exploring the causal pathways by which same-sex marriage laws may influence STIs, most evidence suggests that same-sex marriage laws reflect social attitudes toward gays, and some evidence suggests that they may also affect social attitudes. Lastly, there is some evidence that laws affect self-reported sexual behaviors, but these findings are consistent with several interpretations. In sum, the results point to a modest positive association—if any at all—between syphilis and same-sex marriage bans.

This paper contributes to research on risky sexual behavior and STIs (e.g., Ahituv, Hotz, and Philipson, 1996; Cornwell and Cunningham, 2012a, 2012b; Francis and Mialon, 2010; Johnson and Raphael, 2009; Kremer, 1996; Landsburg, 2007; Oster, 2005, 2009; Philipson and Posner, 1994; Portelli, 2004). It also contributes to a burgeoning literature on laws concerning sexual behavior (e.g., Burris and Cameron, 2008; Delavande et al., 2010; Francis and Mialon, 2008; Lazzarini et al., 2002). In particular, a number of studies examine the history and legal attributes of same-sex marriage laws (Brandenburg, 2005; Gonen, 2001; Koppelman, 2005; Kramer, 1997; Metzger, 2007; Ruskay-Kidd, 1997; Schacter, 2009; Schroeder, 2005), while other studies examine the politics and correlates of such laws (Burnett and Salka, 2009; Fleischmann and Moyer, 2009; McVeigh and Diaz, 2009; Soule, 2004).

Dee (2008) is the first study to rigorously evaluate the effects of same-sex marriage laws on STIs. Using panel data on European countries, Dee estimates the effect of the legalization of same-sex partnerships on the incidence of STIs and finds that same-sex marriage laws decreased the incidence of syphilis but not HIV or gonorrhea. He concludes that the evidence suggests that same-sex marriage laws may promote sexual fidelity. Focusing on the US, Langbein and Yost (2009) find that laws permitting same-sex marriage raised the marriage rate and lowered the abortion rate and percentage of children in female-headed households, while laws prohibiting same-sex marriage lowered the divorce rate, abortion rate, and percentage of children in female-headed households. While this study represents an advance, it uses data from only three years (1990, 2000, and 2004); does not estimate the effects of the laws on STIs; and does not take advantage of available information about the laws, e.g., precise year of passage, whether they were prohibitions by statute or constitutional amendment, or whether they prohibited only same-sex marriage or both marriage and civil union.

Francis and Mialon (2010) examine the relationship between tolerance for gays and the spread of HIV. Using a panel of US states from the mid-1970s to the mid-1990s, they find that tolerance is negatively associated with the HIV rate. The HIV rate is estimated using data on the AIDS rate and the median number of years between HIV infection and the onset of AIDS, prior to the development of highly active antiretroviral therapy (HAART) in 1996. Tolerance is quantified using the measure of attitudes toward homosexuals in the GSS. To complement the GSS measure, state bans on gay marriage are used as a proxy for intolerance. However, the study is far from a rigorous evaluation of the effects of same-sex marriage bans. Many state bans on same-sex marriage were introduced in the post-HAART era, which the study does not analyze, and the study does not take advantage of a wealth of information about the laws.

The remainder of the paper is organized as follows. Section II discusses the theoretical framework. Section III describes the data and empirical strategy. Section IV presents the empirical results. Section V concludes.

II. Theory

In light of previous research, there are several reasons to believe that same-sex marriage laws may induce changes in behavior that impact the spread of STIs, even if same-sex couples did not have the positive right to marry or enter into civil union prior to the passage of the laws.

Same-sex marriage laws may directly affect homosexual behavior. Dee (2008) argues that allowing same-sex marriage can alter the behavioral incentives of homosexuals who aspire to form long-term partnerships. He provides evidence that extending marriage to same-sex couples in Europe resulted in a significant reduction in syphilis, which bolsters the notion that same-sex marriage may raise the gains to forming a committed partnership and reduce the gains to engaging in sexual promiscuity. Conversely, same-sex marriage bans might undermine the incentives to behave monogamously by lowering the expectation that gays will be able to enjoy the economic and emotional benefits of marriage in the near future. By discouraging monogamy, same-sex marriage bans may accelerate the spread of STIs. Alternatively, same-sex marriage bans may affect gays' sense of self-worth or value of life by sending them the message that they are not equal to others or that they are not deserving of the rights enjoyed by others (Kawata, 2010). In the context of the HIV epidemic in Africa, Oster (2005, 2009) finds that the lower is the economic value of life, the greater is the willingness to participate in risky behavior. If this principle may be extended to self-worth, then marriage bans may raise the prevalence of risky sex among gays.

Same-sex marriage laws may also affect homosexual behavior indirectly by changing social attitudes toward gays or, alternatively, may simply reflect changing attitudes towards gays, which in turn, may affect behavior. The law can move as well as mirror social attitudes. Recent research documents evidence that laws can influence attitudes (Alesina and Fuchs-Schündeln, 2007; Fong et al. 2006; Gallus et al., 2006; Kotsadam and Jakobsson, 2011; Soss and Schram, 2007; Svallfors, 2009; Tang et al., 2003). For example, Tang et al. (2003) find that a California smoke-free bar law increased support for smoke-free bars among patrons, while Kotsadam and Jakobsson (2011) find that a Norwegian law criminalizing prostitution made people's attitudes toward prostitution more negative.

Theoretical research proposes several potential causal pathways by which laws might impact attitudes. Laws can affect social costs and benefits underlying the creation of social norms; codify as well as signal social values, which people may internalize to gain cooperation opportunities; signal the prevalence of certain attitudes, which may affect the behavior of those who are concerned with approval; and change social norms by providing a focal point (Carbonara, Parisi, and Wangenheim, 2008; Cooter, 1998; McAdams, 2000; Khan and Stinchcombe, 2010; McAdams and Rasmusen, 2007; Posner, 1998, 2000). Hence, same-sex marriage bans may either reflect attitudes towards homosexuality or influence them by signaling socially-unacceptable and socially-acceptable behaviors, magnifying the stigma associated with same-sex partnerships, and/or conveying information about the prevalence of intolerance toward gays in society.

Whether same-sex marriage laws make or mirror social attitudes toward gays, it remains to establish the link between these attitudes and homosexual behavior. Intolerant attitudes toward gays and social stigma associated with homosexuality may increase the spread of STIs among gay men. Increasing social costs of same-sex partnerships may induce some men who have had male partners to have only female partners or no partners at all. If such men at the extensive margin of homosexual behavior are of “low-activity” type, as they exit the pool of same-sex partners, it is possible that the overall rate of STI transmission among gays might rise (Francis and Mialon, 2010; Kremer, 1996; Landsburg, 2007). Moreover, intolerance may drive homosexual behavior underground causing gay men to substitute from relatively safe, open, and socially-mediated interactions toward relatively risky, secret, and socially-disconnected interactions (Francis and Mialon, 2010). Also, intolerance may raise the incentives for gays to cluster in urban areas, which reduces search costs for partners and potentially increases the spread of STIs (Müller, 2002).

Furthermore, by codifying traditional family norms and signaling the prevalence of traditional family values, same-sex marriage bans may affect general attitudes toward non-marital sex, or they may simply reflect such attitudes, which may affect the spread of STIs among heterosexuals. Lastly, it has been suggested that same-sex marriage bans may uphold the concept of marriage as an institution committed to sexual fidelity (Girgis, George, and Anderson, 2010; George and Elshtain, 2006; Family Research Council, 2010a, 2010b). If men who have sex with men tend to have sexual partnerships of shorter duration than men who have sex with women, and if they tend to have greater propensity for infidelity, then permitting same-sex couples to marry might weaken expectations of marital fidelity generally, thereby increasing the spread of STIs among heterosexuals. Generally against this hypothesis, Graham and Barr (2008) demonstrate empirical evidence that an increase in same-sex households did not lead to a decrease in heterosexual marriage or an increase in heterosexual cohabitation.

III. Data and Empirical Strategy

A. State Laws on Same-Sex Marriage and Civil Union

Today, 39 states prohibit marriage between two people of the same sex by statute, constitutional amendment, or both; 16 permit same-sex marriage or civil union; and 2 do not have any laws explicitly allowing or disallowing same-sex marriage or civil union. Taking advantage of legal resources (Lexis-Nexis Legal and Hein Online Session Laws Library), we reviewed state statutory law, constitutional law, and court decisions in order to compile a comprehensive database of state laws on same-sex marriage and civil union. Table 2.1 summarizes the history of these laws. The table lists the year of enactment for statutes prohibiting/allowing same-sex marriage/civil union, constitutional amendments prohibiting same-sex marriage/civil union, and supreme court rulings allowing same-sex marriage. For legal references and notes, please see the Legal Appendix, which the reader can find on the corresponding author's website. Table 2.1 illustrates that most states (36) currently have statutory bans on same-sex marriage, all of which were enacted since 1973. More than half of states (30) have constitutional bans on same-sex marriage, all of which were enacted since 1998. About half (27) have both statutory and constitutional bans. 19 states prohibit both same-sex marriage and civil union, while 6 states allow civil union. Seven states allow same-sex marriage by statute (District of Columbia, Maine, Maryland, New Hampshire, New York, Vermont, Washington) and three by court ruling (Connecticut, Iowa, Massachusetts). Only one of these laws was enacted prior to 2008.

In the empirical analysis, we employ six different sets of legal variables. The first (Regression A) is a binary indicator for whether a particular state in a particular year had prohibited same-sex marriage by statute or constitutional amendment. The second (Regression B) consists of binary indicators for whether a state had prohibited same-sex marriage only by statute or by both statute and constitutional amendment. Of the two measures, the latter is the stronger type of ban because not only was it approved by the state legislature (statute) but also by a majority of voters (constitutional amendment). The third (Regression C) consists of binary indicators for whether a state had prohibited only same-sex marriage or both marriage and civil union. Of the two measures, the latter is the stronger type of ban because not only does it deny same-sex couples access to marriage but also to any legal status analogous to marriage. The fourth (Regression D) is analogous to Regression B but uses the percentage by which the constitutional amendment passed instead of a binary indicator. Information about ballot measures was gathered from the National Conference of State Legislatures (NCSL, 2010). Note that the extent by which the measures passed varied from 52% to 86%, and that every ballot measure put to a public vote was approved. The fifth (Regression E) consists of a set of binary variables indicating the number of years since the passage of a same-sex marriage ban. This is used to investigate the dynamic effects of the bans—whether the effects are temporary or permanent, immediate or delayed. The sixth (Regression F) distinguishes a fuller range of laws from the least permissive (prohibition of both same-sex marriage and civil union) to the most permissive (legalization of same-sex marriage).

B. State Panel Analysis: Dependent Variables

Using a state-level panel dataset from 1981 to 2008, our objective is to estimate the association between same-sex marriage bans and STIs. The top panel of Table 2.2 displays summary statistics. Except in this table, both dependent variables are logged in order to normalize their distributions, a practice that follows the emerging precedent in the literature on STIs (e.g., Carpenter, 2005; Chesson et al., 2000; Cornwell and Cunningham, 2012a, 2012b; Dee, 2008).

Dependent variables are the number of syphilis cases per 100,000 population aged 15-44 and the number of gonorrhea cases per 100,000 population aged 15-44 (CDC, 2009).[27] The CDC receives reports of notifiable diseases, including syphilis and gonorrhea, from all state health departments, which obtain data from health care providers and clinical laboratories. Completeness and accuracy of the data depend on factors that determine if and when diagnosis occurs (CDC, 2009). A crucial assumption in this paper is that syphilis is a proxy for risky homosexual behavior. Estimates from 2006 suggest that 64% of all syphilis cases are attributable to men who have sex with men, and the rate of syphilis among men who have sex with men is more than 46 times that of other men and more than 71 times that of women (CDC, 2010a,b). In California, syphilis is even more concentrated among men who have sex with men. Estimates from 2009 indicate that they account for 82.8% of male syphilis cases and 79.3% of cases

[27] Unfortunately, we do not have sufficient data on HIV. Most states did not start reporting HIV until the late 1990s, and many of the large states, including California, New York, and Illinois, did not begin to report HIV until 2001 or later (CDC, 1982-2008).

overall (California Department of Public Health, 2010). In contrast, gonorrhea is a proxy for risky heterosexual behavior since more than 90% of gonorrhea cases are attributable to sex between men and women (CDC, 1997). Figure 2.1 illustrates trends in syphilis and gonorrhea. With standard deviations of 25 and 477, respectively, their extreme variability across the period is testament to the fact that these are social diseases spread through networks.

C. State Panel Analysis: Controls

State panel regressions include state fixed effects, year effects, and are weighted by state population share. Many include state-specific linear time trends or state-specific linear and quadratic time trends. Robust standard errors are adjusted for clustering on states (Bertrand, Duflo, and Mullainathan, 2004).

We include a number of state controls in the regressions. The percentage of people aged 25-49 who completed high school, the percentage who completed some college, and the percentage who completed college or more are controls for education, constructed using IPUMS-CPS (King et al., 2010). The percentage of working-age people in the labor force who were unemployed and average real personal income are also constructed using IPUMS-CPS. The percentage urban, based on the Statistical Abstract of the United States, is interpolated between census years (US Census Bureau, 1981-2005). The percentage of the population that was black, the percentage between ages 15 and 29, and the percentage between ages 30 and 44 are derived from data provided by the US Census Bureau (2010). Since religious attitudes may influence the passage of same-sex marriage laws and sexual behavior, we calculate from the GSS the percentage of people who believed the Bible was the literal word of God, the percentage of people who attended religious services nearly every week or more, and the percentage of people who were Protestant, Catholic, Jewish, and other religion (Davis et al., 2010).

We also add several controls to address specific alternative hypotheses. First, other state laws may be related both to the passage of same-sex marriage laws and to sexual behavior. For this reason, we include in the regressions an indicator for whether states had a law that required sex education programs in school to stress abstinence. We also include indicators for whether states either had a law that required parental notification to legally perform an abortion upon a minor or had a law that required parental consent.

Second, illegal drug use represents a potential confounding influence, since it is positively correlated with STIs. In order to account for this, we control for the drug arrest rate as well as the property crime rate, which is known to be closely associated with patterns of illegal drug use (ONDCP, 2000). Both variables are derived from the Uniform Crime Reports (FBI, 2011; NCOVR, 2011).

Third, same-sex marriage laws may have been passed in response to changes in perceived risky homosexual behavior or changes in visibility of the gay community. Undoubtedly, the observable features of the AIDS epidemic most influence perceived risky homosexual behavior. To some extent, the AIDS epidemic may reflect or affect risky behaviors of gay men. Therefore, we include in regressions the number of AIDS cases per 100,000 population, the principal AIDS statistic publicly reported at the state level (CDC, 1982-2008). Also, the relative size of the gay population may influence perceived homosexual risky behavior and visibility of the gay community. To measure this, we gathered data from historical editions of Damron Men's Travel Guide, the longest and most complete gay men's travel guide (Damron, 1981-1992, 1993-1998, 1999-2008). We include in regressions the state share of total entries (e.g., gay bars, bookstores, restaurants, and churches) listed in the guide.

Fourth, cross-state migration of gay men in response to changes in laws or attitudes might subsequently impact STI rates. Gay men may move from states that pass laws prohibiting same-sex marriage and to states that pass laws permitting same-sex marriage. The inclusion of the state share of total entries listed in the Damron Men's Travel Guide, as described above, can address this issue.

Fifth, it is crucial to consider the emergence of Highly Active Antiretroviral Therapy (HAART) in the mid-1990s. These medical advances significantly extended the life expectancy of those infected with HIV and led to an upsurge of risky sexual behavior among men who have sex with men (Katz et al., 2002; Lakdawalla et al., 2006). It may be important to control for access to HAART even beyond the set of year indicators included in the regressions. AIDS Drug Assistance Programs (ADAPs) provide HIV-related prescription drugs to low income people living with HIV (US Department of Health and Human Services, 2003). The scale and scope of ADAPs grew substantially during the HAART era. Indicative of this trend, total ADAP funding increased from $204 million in 1996 to $1.5 billion in 2008. We make use of the substantial heterogeneity in ADAP financial eligibility across states and time. In the regressions, we include a measure of access to HAART, which before 1996 is zero and after 1996 is ADAP financial eligibility as a percentage of the Federal Poverty Level (NASTAD, 1998-2009).

D. Individual-Level Analysis

To complement the state panel analysis, we explore the potential causal mechanisms underlying the relationship between same-sex marriage laws and STIs using the GSS, a nationally representative repeated cross-sectional survey of adults (Davis et al., 2010). The bottom panel of Table 2.2 displays summary statistics for the individual-level dataset. First, we test whether laws reflect or affect social attitudes toward sex. Measures of tolerance for same-sex sex and tolerance for teen sex are constructed from 1982 to 2008. The GSS provides the longest and most consistent measure of society-wide attitudes towards gays. Tolerance for same-sex sex equals one if a respondent believes sexual relations between two adults of the same sex is "not wrong at all" or "wrong only sometimes" and equals zero if a respondent believes it is "almost always wrong" or "always wrong" (Francis and Mialon, 2010). Tolerance for teen sex is defined analogously. Second, we test whether laws reflect or affect sexual behavior. Measures of sexual behavior are limited in the GSS. Three measures are constructed from 1988 to 2008: whether the respondent had same-sex sex in the past year, had multiple sexual partners in the past year, and had an extra-marital affair ever. Leads of same-sex marriage laws are utilized to examine the possibility that laws reflect social attitudes/sexual behavior, while lags of laws are utilized to examine the possibility that they affect social attitudes/sexual behavior. Since the GSS is conducted roughly every two years, it makes sense to use leads and lags of t+2 and t-2, respectively. Additionally, all individual-level regressions include state fixed effects, year effects, age, gender, race, education, and religious attendance.

E. Empirical Models

In the analysis, we employ OLS and dynamic panel models. Following the literature (e.g., Dee, 2008; Chesson et al., 2000), we implement dynamic panel specifications. In the case of sexually transmitted infections, it may be important to introduce a lagged dependent variable as regressor because the contemporaneous incidence of an infectious disease depends heavily on its prior incidence. We estimate the following equation:

$$s_{i,t} = \gamma s_{i,t-1} + \boldsymbol{\beta}' \boldsymbol{X}_{i,t} + \alpha_i + \mu_t + \varepsilon_{i,t}.$$

The dependent variable, $s_{i,t}$, is the natural log of the syphilis or gonorrhea rate in state *i* and year *t*, α_i and μ_t are state and year fixed effects, respectively, $\boldsymbol{X}_{i,t}$ includes legal variables and state-year controls, and $\varepsilon_{i,t}$ is the error term. Although the inclusion of a lagged dependent variable can create a bias, the magnitude of the bias decreases with the length of the panel, and the least squares fixed effects model with a lagged dependent variable "performs just as well or better than many alternatives when T=30" (Judson and Owen, 1999). Additionally, reverse causality may not be a significant concern when the outcome variables are STIs. Investigating the effects of same-sex marriage laws on STIs in Europe, Dee (2008) argues that such laws were largely exogenous given that the public debates about them centered on issues of "fairness, equality, and morality" but not on issues of public health. Moreover, the broader population is unaware that syphilis is highly concentrated among men who have sex with men, so reverse causality is unlikely. Syphilis has been around much longer than HIV, and many historical figures suspected or known to have had syphilis, including King Henry the 8th, Lord Byron, Al Capone, and Adolf Hitler, are not typically associated with homosexual behavior. Similarly, measurement error in STI rates may not be a concern, since passage of same-sex marriage laws is unlikely correlated with public awareness of STIs and state STI control policies.

IV. Results

Figure 2.2 and Figure 2.3 illustrate the estimated change in syphilis and gonorrhea rates ten years before and ten years after the passage of same-sex marriage bans. To construct the figures, each of the dependent variables was regressed on a set of binary indicators for the number of years before and after enactment, state fixed effects, year effects, and controls. Dotted lines denote confidence intervals. As Figure 2.2 shows, there was an upward trend in syphilis 5 to 10 years prior to the passage of the laws but the trend was relatively flat 1 to 5 years prior. Following passage, syphilis increased and remained at an elevated level for about 5 years. As Figure 2.3 shows, there was an upward trend in gonorrhea throughout most of the period. Thus, it may be important to control for time trends.

Table 2.3 displays regressions of syphilis on bans, state fixed effects, year effects, state-level controls, and state-specific time trends using various measures of same-sex marriage bans. The top panel of the table includes the state of California, and the bottom panel excludes it. In the top panel, several regressions indicate that same-sex marriage bans are positively associated with syphilis. Notably, the coefficients on the various legal measures are significant in dynamic panel models with linear time trends; and the coefficient on "both marriage & civil union ban" is significant in dynamic panel models with and without linear and quadratic time trends. In the bottom panel, the coefficients are positive but smaller and mostly insignificant; the coefficient on "both marriage &

civil union ban" is significant in the dynamic panel model with linear time trends.[28] These findings are consistent with at least three explanations. Same-sex marriage laws may affect behavior directly, may affect behavior indirectly by changing attitudes toward gays, or may simply reflect changing attitudes towards gays, which in turn, affect behavior. Table 2.4 displays analogous regressions for gonorrhea. The table reveals that none of the coefficients on the legal measures are significant. This finding is inconsistent with the hypothesis that same-sex marriage bans reduce the spread of heterosexual STIs.

It is useful to examine the magnitudes of these estimates and place them in context. Our preferred specification is dynamic panel with time trends including California. Estimates suggest that having any same-sex marriage ban is associated with a 16.3% increase in the syphilis rate, which corresponds to a rise in about 2.2 cases of syphilis per 100,000 population aged 15-44. Having a ban on both marriage and civil union is associated with a 20.8% increase in syphilis, which corresponds to a rise in about 2.9 cases per 100,000 population aged 15-44. Estimates excluding California indicate that having a ban on both marriage and civil union is associated with a 12.9% increase in syphilis, which corresponds to a rise in about 1.8 cases per 100,000 population aged 15-44. These are reasonably-sized estimates given that syphilis had exhibited considerable variation during the study period. Between 2000 and 2008, syphilis increased by 215%, and between 1981 and 2008, the standard deviation of syphilis was 25 cases per 100,000 population.

[28] Note that we also employed Arellano and Bond's GMM technique in case the inclusion of a lagged dependent variable was problematic, and we found that most coefficients were not significant, but those that were significant corresponded to the most restrictive type of ban, "both marriage & civil union ban."

Table 2.5 makes use of information about the margin by which the constitutional bans were passed. Note that voter approval ranged from 52% to 86% with mean approval of 68%. The coefficients on "only statutory ban" are consistent with those we found in Table 2.3 and Table 2.4. Additionally, the sign and significance of the coefficients on "vote in favor of constitutional ban" closely mirror those on "both constitutional and statutory ban" in Table 2.3 and Table 2.4. As before, the regression results vary considerably with California. When California is excluded, most of the coefficients are insignificant. While the findings in this table might suggest that social attitudes relate to STIs, they do not inform us about how they do so. Using a Regression Discontinuity design would be much more informative, because it would enable us to distinguish the effect of the passage of the law from the effect of underlying voter attitudes. However, we cannot use this technique since every ballot measure on same-sex marriage put to a public vote was approved. Following our state panel analysis, we use GSS individual-level data to distinguish whether laws affect attitudes or attitudes affect laws.

To investigate the dynamic effects of the laws—whether the effects are temporary or permanent, immediate or delayed—we regress each of the dependent variables on a set of binary variables indicating the number of years since the passage of a same-sex marriage ban. Regressions also utilize state fixed effects, year effects, state-specific trends, and state-year controls. Table 2.6 displays the results. The evidence suggests that when California is included, bans have a statistically significant effect on syphilis in the short and medium term, while bans do not have any effect on gonorrhea at all.

Table 2.7 examines a range of same-sex marriage laws from the least permissive to the most permissive. Four types of laws were identified: laws that prohibit both same-sex marriage and civil union; those that prohibit same-sex marriage but neither prohibit nor allow same-sex civil union; those that prohibit same-sex marriage but allow same-sex civil union; and those that allow same-sex marriage or allow same-sex civil union but neither prohibit nor allow same-sex marriage. As of 2012, states that prohibit same-sex marriage but allow same-sex civil union include California, Hawaii, Illinois, Nevada, and Oregon. States that allow same-sex marriage, or allow same-sex civil union but neither prohibit nor allow same-sex marriage, include Connecticut, District of Columbia, Iowa, Maryland, Massachusetts, New Hampshire, New Jersey, New York, Vermont, and Washington. As the table shows, the coefficient on the most permissive type of law, "allow marriage or allow civil union & not ban marriage," is never significant. In contrast, when syphilis is the dependent variable, the coefficient on the least permissive type of law, "ban marriage & civil union," is positive and significant across dynamic panel models that include California. Nevertheless, without California, most of the estimates in the table are smaller and less significant.

Using GSS individual-level data, we investigate the causal pathways by which same-sex marriage laws may influence STI rates. Table 2.8 displays regressions of attitudes toward same-sex sex and teen sex on state laws, state fixed effects, year effects, and individual-level controls including age, gender, race, education, and religious attendance. To test whether laws reflect social attitudes, the specifications in columns (1) and (2) involve leads of the laws (t+2); to test whether laws affect social attitudes, the specifications in columns (3) and (4) involve lags of the laws (t-2). First, let us consider tolerance for same-sex sex. Although the coefficients on leads of laws permitting same-sex marriage or civil union are not statistically significant, most of the coefficients on leads of laws prohibiting same-sex marriage or civil union are negative and significant. The more restrictive the ban, the larger is the coefficient. Nevertheless, most of the coefficients on lags of the laws remain insignificant. Only the coefficient on "both marriage & civil union ban" is negative and significant, and only the one on "ban marriage & allow civil union" is positive and significant. In contrast to tolerance for same-sex sex, tolerance for teen sex does not appear to be statistically associated with any leads or lags of same-sex marriage laws. All in all, the evidence is largely consistent with the notion that same-sex marriage laws reflect social attitudes toward gays, but some evidence indicates that they may affect social attitudes as well.

Table 2.9 displays regressions of measures of sexual behavior on state laws, state fixed effects, year effects, and individual-level controls. To test whether laws reflect changes in sexual behavior, regressions in the first three columns use leads of the laws (t+2); to test whether laws affect sexual behavior, regressions in the last three columns use lags of the laws (t-2). The top panel includes male respondents and the bottom panel female respondents. As the table shows, almost none of the coefficients on leads of the laws are statistically significant. For both men and women, many of the coefficients on lags of bans are negative and significant when the dependent variable is same-sex sex in the past year. The more restrictive the ban, the larger is the coefficient. When the dependent variable is multiple partners, only one coefficient is significant at the 5% level, the one on "allow marriage or allow civil union & not ban marriage" for male respondents, which is positive and significant. Additionally, when the dependent variable is extra-marital affairs, a number of coefficients on same-sex marriage bans are negative and significant for men and women. It is challenging to interpret these findings because they are consistent with several explanations. It could be that laws prohibiting same-sex marriage caused a decrease in same-sex sex and extra-marital affairs, whereas laws permitting same-sex marriage caused an increase in multiple partners for male respondents. However, it also could be that laws merely influenced cross-state migration and/or reporting of sexual behaviors, since unlike the state panel data, the individual-level data is especially vulnerable to selection and reporting bias.

V. Conclusion

In this paper, we examined the relationship between same-sex marriage laws and sexually transmitted infections. We summarize the evidence as follows. First, there is no evidence of any association between laws prohibiting same-sex marriage/civil union and gonorrhea, a proxy for risky heterosexual behavior. Bans are positively associated with syphilis, a proxy for risky homosexual behavior. Second, of the different legal measures, bans on both same-sex marriage and civil union are most consistently associated with syphilis. These are the most restrictive type of ban because they deny same-sex couples access not only to marriage but also to any legal status analogous to marriage. Third, all the estimates are smaller and less significant when we exclude California, the state with the largest gay population. Fourth, there is no evidence of any association between laws permitting same-sex marriage/civil union and syphilis or gonorrhea, although there is insufficient variation in these laws to yield precise estimates. Fifth, exploring the causal pathways by which same-sex marriage laws may influence STIs, most findings suggest that same-sex marriage laws reflect social attitudes toward gays, and some findings suggest that they may also affect social attitudes. Lastly, there is some evidence that laws affect self-reported sexual behaviors, but these results are consistent with several interpretations.

Thus, the results point to a modest positive association—if any at all—between syphilis and same-sex marriage bans. While we are confident that the effects on syphilis are not large, we remain unsure as to whether or not they are positive. The fragility of the results to the inclusion of California may indicate that there is actually no association between the laws and syphilis or may indicate that there is too little power to determine whether there is an association. In future research, it may be worthwhile to investigate the association between same-sex marriage laws and STIs in California more specifically. Given that most laws permitting same-sex marriage were enacted toward the end of our study period, it may be fruitful to revisit this question in a few years when additional data become available.

References

Ahituv, A., Hotz, V.J., & Philipson, T. (1996). The responsiveness of the demand for condoms to the local prevalence of AIDS. *Journal of Human Resources,* 31(4), 869-897.

Alesina, A., & Fuchs-Schündeln, N. (2007). Good-bye Lenin (or not?): The effect of communism on people’s preferences. *American Economic Review,* 97(4), 1507-1528.

Bertrand, M., Duflo, E., & Mullainathan, S. (2004). How much should we trust difference-in-difference estimates? *Quarterly Journal of Economics,* 119(1), 249-75.

Brandenburg, N.M. (2005). Preachers, politicians, and same-sex couples: Challenging same-sex civil unions and implications on interstate recognition. *Iowa Law Review,* 91(1), 319-348.

Burnett, R.C., & Salka, W.M. (2009). Determinants of electoral support for anti-gay marriage constitutional amendments: An examination of 2006 votes on ballot measures in the states. *Journal of Homosexuality,* 56(8), 1071-1082.

Burris, S., & Cameron, E. (2008). The case against criminalization of HIV transmission. *JAMA-Journal of the American Medical Association*, 300(5), 578-581.

California Department of Public Health. (2010). California syphilis elimination: Surveillance data, 2009." Online, http://www.cdph.ca.gov/data/statistics/Documents/STD-Data-Syphilis-Elimination-Surveillance-Data.pdf

Carbonara, E., Parisi, F., & Wangenheim, G.V. (2008). Lawmakers as norm entrepreneurs. *Review of Law and Economics,* 4(3), 779-799.

Carpenter, C. (2005). Youth alcohol use and risky sexual behavior: Evidence from underage drunk driving laws. *Journal of Health Economics,* 24(3), 613-628.

Centers for Disease Control and Prevention (CDC). (1982-2008). HIV/AIDS surveillance report. Atlanta, GA: CDC.

Centers for Disease Control and Prevention (CDC). (1997) Gonorrhea among men who have sex with men–selected sexually transmitted diseases clinics, 1993-1996. *MMWR Weekly,* 46, 889-892.

Centers for Disease Control and Prevention (CDC). (2009). National Electronic Telecommunications System for Surveillance (NETSS). Atlanta, GA: CDC.

Centers for Disease Control and Prevention (CDC). (2010a). Syphilis and MSM - CDC fact sheet. Atlanta, GA: CDC.

Centers for Disease Control and Prevention (CDC). (2010b). CDC analysis provides new look at disproportionate impact of HIV and syphilis among U.S. gay and bisexual men. CDC Press Release. Atlanta, GA: CDC.

Chesson, H., Harrison, P., & Kassler, W.J. (2000). Sex under the influence: The effect of alcohol policy on sexually transmitted disease rates in the United States. *Journal of Law and Economics*, 43(1), 215-238.

Cooter, R.D. (1998) Expressive law and economics. *Journal of Legal Studies,* 27(2), 585-608.

Cornwell, C.M., & Cunningham, S. (2012a). Sex ratios and risky sexual behavior. Working Paper, Department of Economics, Baylor University.

Cornwell, C.M., & Cunningham, S. (2012b). Abortion legalization and sexually transmitted infections. Working Paper, Department of Economics, Baylor University.

Damron, B. (1981-1992). *Bob Damron's Address Book*. San Francisco, CA: Bob Damron Enterprises.

Damron, B. (1993-1998). *Damron Address Book*. San Francisco, CA: Damron Company.

Damron, B. (1999-2008). *Damron Men's Travel Guide*. San Francisco, CA: Damron Company.

Davis, J.A., Smith, T.W., & Marsden, P.V. (2010). General Social Surveys, 1972-2008 [Cumulative File] [Computer file]. ICPSR25962-v2. Storrs, CT: Roper Center for Public Opinion Resarch, University of Connecticut/Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributors], 2010-02-08. doi:10.3886/ICPSR25962

Dee, T.S. (2008). Forsaking all others? The effects of same-sex marriage partnership laws on risky sex. *Economic Journal*, 118(530), 1055-1078.

Delavande A., Goldman, D., & Sood, N. (2010). Criminal prosecution and human immunodeficiency virus-related risky behavior. *Journal of Law & Economics,* 53(4), 741-782.

Family Research Council. (2010a). *Ten arguments from social science against same-sex 'marriage'*. Available at http://www.frc.org/get.cfm?i=if04g01.

Family Research Council. (2010b). *What's wrong with letting same-sex couples 'marry'?* Available at http://www.frc.org/get.cfm?i=if03h01.

Federal Bureau of Investigation (FBI). (2011). Uniform crime reports (UCR) data online, http://www.ucrdatatool.gov

Fleischmann, A., & Moyer, L. (2009). Competing social movements and local political culture: Voting on ballot propositions to ban same-sex marriage in the US states. *Social Science Quarterly,* 90(1), 134-149.

Fong, G.T., Hyland, A., Borland, R., Hammond, D., Hastings, G., McNeill, A., Anderson, S., Cummings, K.M., Allwright, S., Mulcahy, M., Howell, F., Clancy, L., Thompson, M.E., Connolly, G., & Driesen, P. (2006). Reductions in tobacco smoke pollution and increases in support for smoke-free public places following the implementation of comprehensive smoke-free workplace legislation in the Republic of Ireland: Findings from the ITC Ireland/UK survey. *Tobacco Control*, 15(Supplement 3), 51-58.

Francis, A.M., & Mialon, H.M. (2008). The optimal penalty for sexually transmitting HIV. *American Law and Economics Review,* 10(2), 388-423.

Francis, A.M., & Mialon, H.M. (2010). Tolerance and HIV. *Journal of Health Economics,* 29(2), 250–267.

Gallus, S., Zuccaro, P., Colombo, P., Apolone, G., Pacifici, R., Garattini, S., & La Vecchia, C., 2006. "Effects of new smoking regulations in Italy. *Annals of Oncology,* 17(2), 346-347.

George, R.P., & Elshtain, J.B. (2006). *The Meaning of marriage: Family, state, market, and morals*. Spence Publishing.

Girgis, S., George, R.P., & Anderson, R.T. (2010). What is marriage? *Harvard Journal of Law and Public Policy,* 34(1), 245-287.

Gonen, J.S. (2001). Same-sex unions and domestic partnerships. *Georgetown Journal of Gender and the Law,* 2(1), 329-345.

Graham, J.W., & Barr, J. (2008). Assessing the geographic distribution of same sex and opposite sex couples across the United States: implications for claims of causality between traditional marriage and same sex unions. *Review of Economics of the Household,* 6(4), 347-367.

Johnson, R.C., & Raphael, S. (2009). The effects of male incarceration dynamics on AIDS infection rates among african-american women and men. *Journal of Law and Economics*, 52(2), 251-293.

Judson, R.A., & Owen, A.L. (1999). Estimating dynamic panel data models: A guide for macroeconomists. *Economics Letters*, 65(1), 9-15.

Katz, M.H., Schwarcz, S.K., Kellogg, T.A., Klausner, J.D., Dilley, J.W., Gibson, S., & McFarland, W. (2002). Impact of highly active antiretroviral treatment on HIV seroincidence among men who have sex with men: San Francisco. *American Journal of Public Health*, 92(3), 388-394.

Kawata, P.A. (2010). Why prop 8 decision should be important to the HIV/AIDS community. *The Body: The Complete HIV/AIDS Resource*. August 10, 2010, Available at http://www.thebody.com/content/news/art58034.html.

Khan, U., & Stinchcombe, M.B. (2010). Informative waiting. Working Paper, University of Texas.

King, M., Ruggles, S., Alexander, J.T., Flood, S., Genadek, K., Schroeder, M.B., Trampe, B., & Vick, R. (2010). Integrated Public Use Microdata Series, Current Population Survey: Version 3.0. [Machine-readable database]. Minneapolis: University of Minnesota.

Koppelman, A. (2005). Recognition and enforcement of same-sex marriage. *University of Pennsylvania Law Review*, 153(6), 2143-2194.

Kotsadam, A. & Jakobsson, N. (2011). Do laws affect attitudes? An assessment of the Norwegian prostitution law using longitudinal data. *International Review of Law and Economics,* 31(2), 103-115.

Kramer, L. (1997). Same-sex marriage, conflict of laws, and the unconstitutional public policy exception. *Yale Law Review,* 106(7), 1965.

Kremer, M. (1996). Integrating behavioral choice into epidemiological models of AIDS. *Quarterly Journal of Economics,* 111(2), 549-573.

Lakdawalla, D., Sood, N., & Goldman, D. (2006). HIV breakthroughs and risky sexual behavior. *Quarterly Journal of Economics*, 121(3), 1063-1102.

Landsburg, S.E. (2007). *More sex is safer sex: The unconventional wisdom of economics*. New York: Free Press.

Langbein, L., & Yost, M.A. (2009). Same-sex marriage and negative externalities. *Social Science Quarterly*, 90(2), 292-308.

Lazzarini, Z., Bray, S., & Burris, S. (2002). Evaluating the impact of criminal laws on HIV risk behavior. *Journal of Law, Medicine, and Ethics*, 30(2), 239-253.

McAdams, R.H., & Rasmusen, E.B. (2007). Norms in the law. In M. Polinsky, & S. Shavell (Eds.), *Handbook of Law and Economics*, Elsevier.

McAdams, R.H. (2000). An attitudinal theory of expressive law. *University of Oregon Law Review*, 79(2), 339-390.

McVeigh, R., & Diaz, M.D. (2009). Voting to ban same-sex marriage: Interests, values, and communities. *American Sociological Review,* 74(6), 891-915.

Metzger, G.E. (2007). Congress, article IV, and interstate relations. *Harvard Law Review*, 120(6), 1468-1542.

Müller, C. (2002). An economic analysis of same-sex marriage. German Working Papers in Law and Economics, No. 2002-1-1045.

National Alliance of State and Territorial AIDS Directors (NASTAD). (1998-2009). National ADAP monitoring project annual report. Online, www.kff.org and www.nastad.org

National Conference of State Legislatures (NCSL). (2010). Online database of ballot initiatives: http://www.ncsl.org/LegislaturesElections/ElectionsCampaigns/BallotMeasuresDatabase/tabid/16580/Default.aspx.

National Consortium on Violence Research (NCOVR). (2011). Uniform crime reports (UCR) arrests data. Web-based Data Center of NCOVR, http://www.ncovr.org

Office of National Drug Control Policy (ONDCP). (2000). Drug-Related Crime. ONDCP Drug Policy Information Clearinghouse Fact Sheet.

Oster, E. (2005). Sexually transmitted infections, sexual behavior, and the HIV/AIDS epidemic. *Quarterly Journal of Economics*, 120(2), 467-515.

Oster, E. (2009). HIV and sexual behavior change: Why not Africa? National Bureau of Economic Research, NBER Working Paper 13049.

Philipson, T., & Posner, R.A. (1994). Public spending on AIDS education: An economic analysis. *Journal of Law and Economics,* 37(1), 17-38

Portelli, C.J. (2004). Economic analysis of same-sex marriage. *Journal of Homosexuality*, 47(1), 95-109.

Posner, E.A. (1998). Symbols, signals, and social norms in politics and the law. *Journal of Legal Studies,* 27(Issue 2, Part 2), 765-797.

Posner, E.A. (2000). Law and social norms: The case of tax compliance. *Virginia Law Review* 86(8), 1781-1819.

Ruskay-Kidd, S. (1997). The defense of marriage act and the overextension of congressional authority. *Columbia Law Review,* 97(5), 1435-1482.

Schacter, J.S. (2009). Courts and the politics of backlash: Marriage equality litigation, then and now. *Southern California Law Review*, 82(6), 1153-1223.

Schroeder, T.C. (2005). Does sex matter? Washington's defense of marriage act under the equal rights amendment of the Washington state constitution. *Washington Law Review,* 80(4), 535-563.

Soss, J., & Schram, S.F. (2007). A public transformed? Welfare reform as policy feedback. *American Political Science Review*, 101(1), 111-127.

Soule, S.A. (2004). Going to the chapel? Same-sex marriage bans in the United States, 1973-2000. *Social Problems,* 51(4), 453-477.

Svallfors, S. (2010). Policy feedback, generational replacement and attitudes to state intervention: Eastern and Western Germany, 1990-2006. *European Political Science Review,* 2(1), 119-135.

Tang H., Cowling, D.W., Lloyd, J.C., Rogers, T., Koumjian, K.L., Stevens, C.M., & Bal, D.G. (2003). Changes of attitudes and patronage behaviors in response to a smoke-free bar law. *American Journal of Public Health,* 93(4), 611-617.

US Census Bureau, Population Division. (2010). Population Estimates Program website, Washington, DC.

US Census Bureau. (1981-2005). Statistical abstract of the United States. Washington, DC: US G.P.O.

US Department of Health and Human Services. (2003). The ADAP manual: AIDS drug assistance program of the Ryan White CARE Act." Online, ftp://ftp.hrsa.gov/hab/ADAP2003.pdf

Figure 2.1

Trends in Syphilis and Gonorrhea, 1981-2008

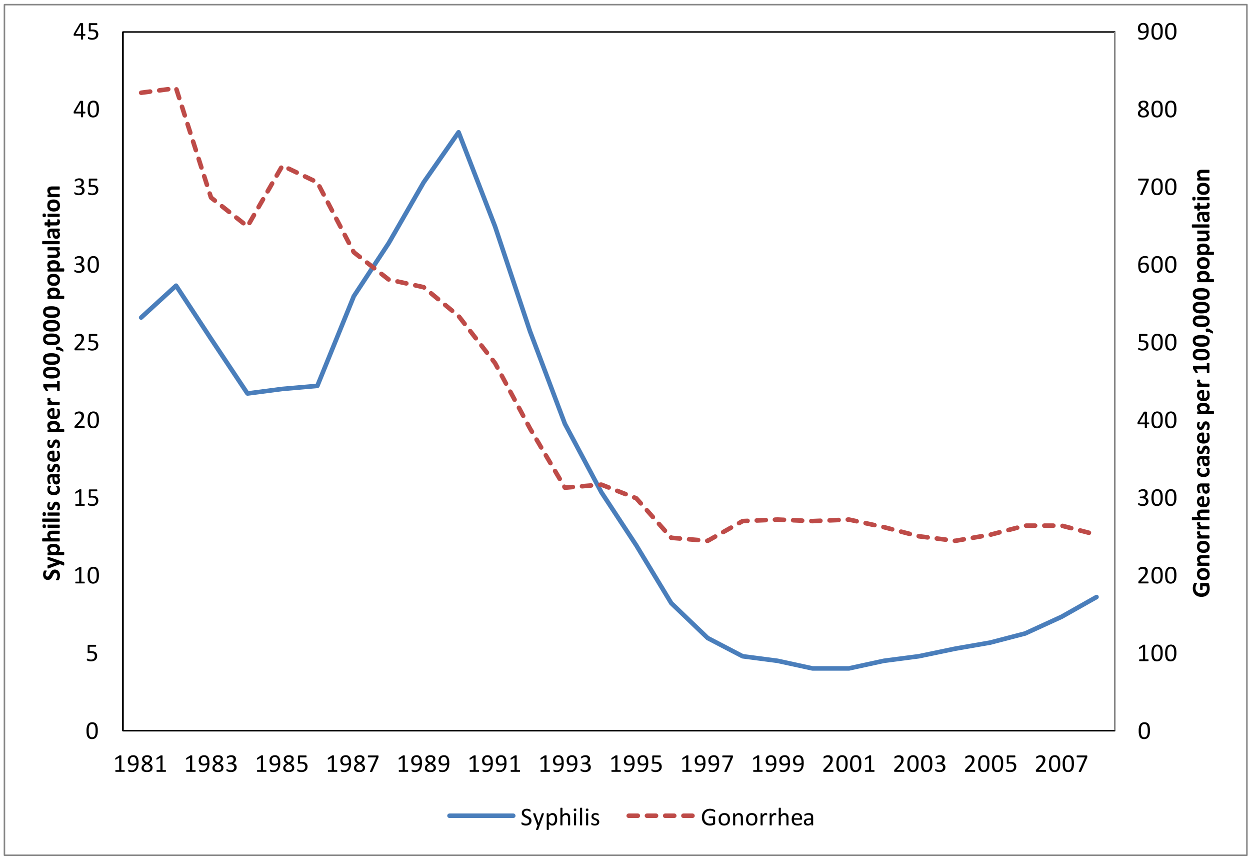


Figure 2.2

Estimated Change in Syphilis Rates Relative to Timing of Same-Sex Marriage Bans

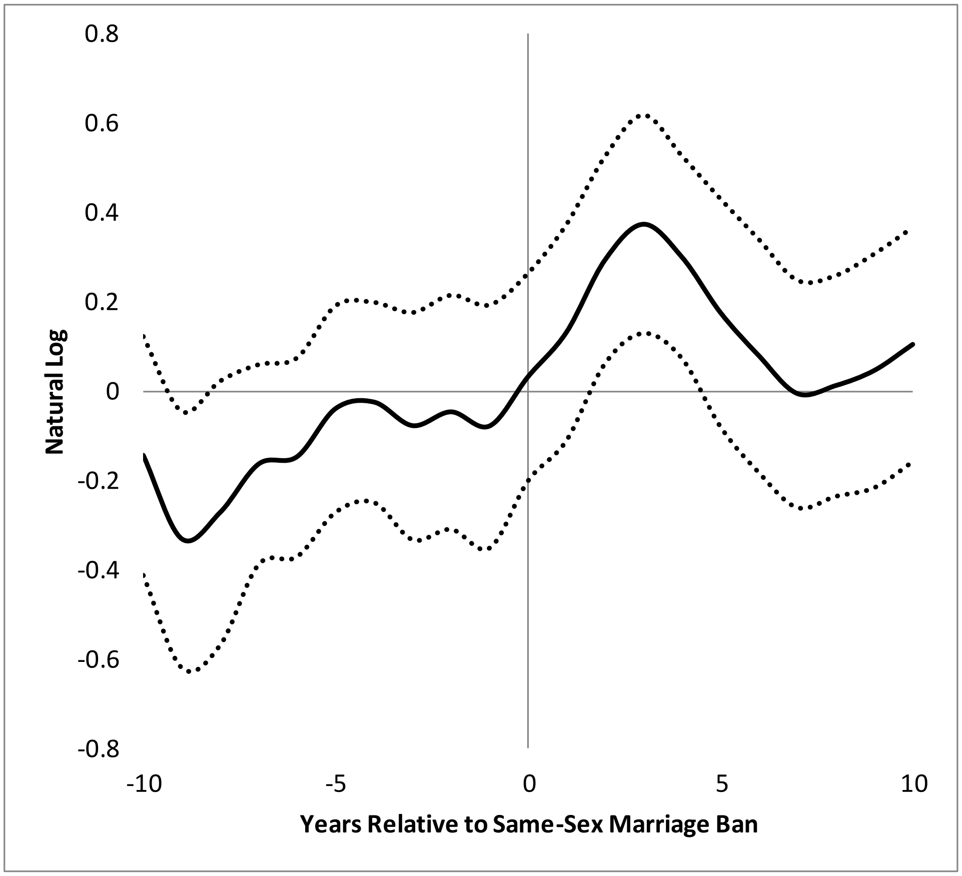


Figure 2.3

Estimated Change in Gonorrhea Rates Relative to Timing of Same-Sex Marriage Bans

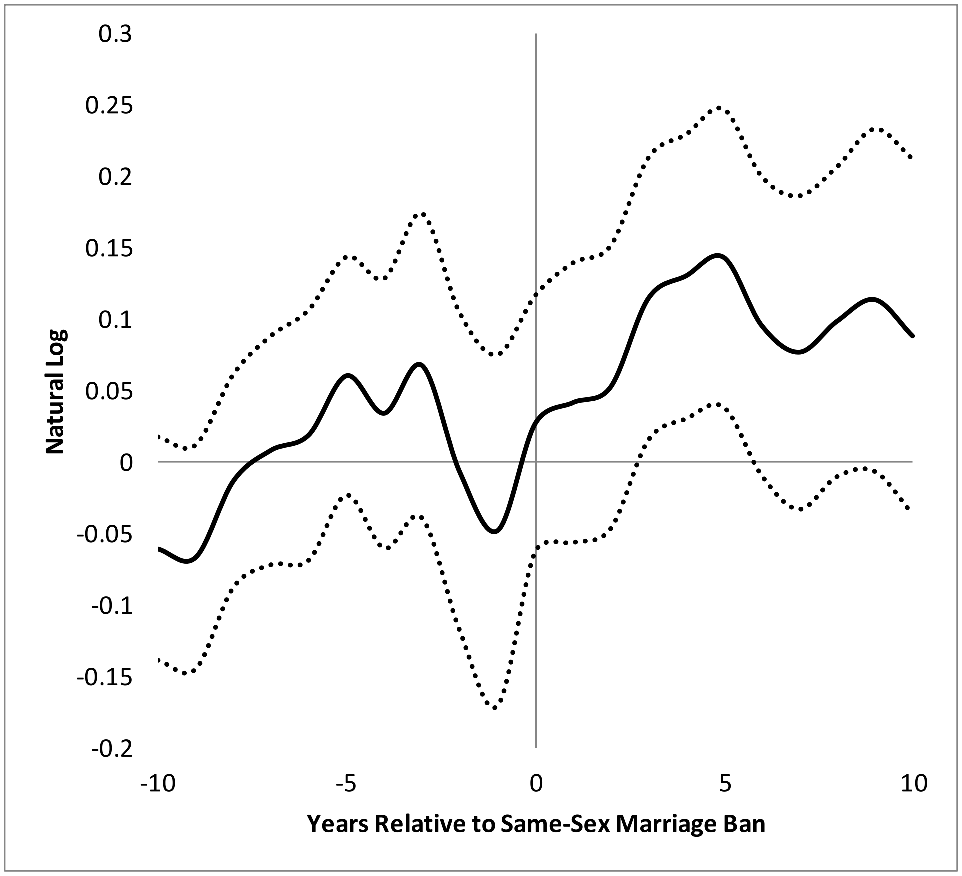


Table 2.1 State Laws on Same-Sex Marriage and Civil Union

	STATE STATUTES				STATE CONSTITUTIONAL AMENDMENTS		STATE SUPREME COURT RULINGS
State	Prohibit Marriage	Prohibit Civil Union	Allow Marriage	Allow Civil Union	Prohibit Marriage	Prohibit Civil Union	Allow Marriage
Alabama	1998				2006	2006	
Alaska	1996	1996			1998		
Arizona	1996				2008		
Arkansas	1997				2004	2004	
California	2000-2008			2003	2008		
Colorado	2000				2006		
Connecticut	2000-2007			2005-2008			2008
Delaware	1996						
DC			2010	2002-2010			
Florida	1977				2008	2008	
Georgia	1996	1996			2004	2004	
Hawaii	1994			2011			
Idaho	1996				2006		
Illinois	1996			2011			
Indiana	1997						
Iowa	1998-2008						2009
Kansas	1996				2005	2005	
Kentucky	1998				2004	2004	
Louisiana	1988	1988			2004	2004	
Maine	1997-2011		2012				
Maryland	1973-2011		2012				
Massachusetts							2003
Michigan	1996				2004	2004	
Minnesota	1997						
Mississippi	1997				2004		
Missouri	1996-98, 2001				2004		
Montana	1997	1997			2004		
Nebraska					2000	2000	
Nevada				2009	2002		

Table 2.1 State Laws on Same-Sex Marriage and Civil Union (continued)

	STATE STATUTES				STATE CONSTITUTIONAL AMENDMENTS		STATE SUPREME COURT RULINGS
State	Prohibit Marriage	Prohibit Civil Union	Allow Marriage	Allow Civil Union	Prohibit Marriage	Prohibit Civil Union	Allow Marriage
New Hampshire	1987-2008		2009	2007-2009			
New Jersey				2006			
New Mexico							
New York			2011				
North Carolina	1995				2012		
North Dakota	1997				2004	2004	
Ohio	2004	2004			2004	2004	
Oklahoma	1975				2004		
Oregon	1975			2007	2004		
Pennsylvania	1996						
Rhode Island							
South Carolina	1996				2006	2006	
South Dakota	1996				2006	2006	
Tennessee	1996				2006		
Texas	1997	2003			2005	2005	
Utah	1977	2004			2004	2004	
Vermont	2000-2008		2009	1999-2009			
Virginia	1997	2004			2006	2006	
Washington	1998-2011		2012	2009-2011			
West Virginia	2000						
Wisconsin	1979				2006	2006	
Wyoming	1977						

NOTE. Please see the Legal Appendix, which the reader can find on the corresponding author’s website, for legal references and notes.

Table 2.2 Summary Statistics, 1981-2008

Variable	Sample Size	Mean	Standard Deviation
STATE PANEL DATASET			
Syphilis cases per 100,000 population aged 15-44	1372	13.72	25.14
Gonorrhea cases per 100,000 population aged 15-44	1372	405.02	477.20
AIDS cases per 100,000 population	1372	11.94	22.25
High school %	1372	37.04	6.26
Some college %	1372	25.22	6.07
College %	1372	26.41	6.18
Bible literal word of God %	1372	36.91	25.20
Church attendance nearly every week or more %	1372	38.84	22.26
Protestant %	1372	60.62	25.04
Catholic %	1372	26.20	23.65
Jewish %	1372	1.28	3.76
Other religion %	1372	6.92	14.87
State share of entries in Damron %	1372	2.01	2.64
Abstinence-stressed sex education law	1372	0.27	0.45
Parental consent abortion law	1372	0.39	0.49
Parental notification abortion law	1372	0.25	0.43
ADAP eligibility (% of Federal Poverty Level)	1372	307.53	95.51
Drug arrests per 100,000 population	1372	443.96	293.87
Property crimes per 100,000 population	1372	4142.82	1220.19
Unemployment %	1372	6.40	2.32
Average real personal income (thousands)	1372	18.14	3.09
Black %	1372	11.29	12.08
Urban %	1372	70.34	15.35
Population aged 15-29 %	1372	22.71	2.63
Population aged 30-44 %	1372	22.46	2.12
INDIVIDUAL-LEVEL DATASET			
Tolerance for same-sex sex	23191	0.29	0.45
Tolerance for teen sex	19000	0.13	0.34
Same-sex sex in past year (men)	9090	0.04	0.19
Multiple partners in past year (men)	10778	0.18	0.39
Extra-marital affair ever (men)	6711	0.23	0.42
Same-sex sex in past year (women)	10283	0.03	0.16
Multiple partners in past year (women)	13885	0.09	0.29
Extra-marital affair ever (women)	9411	0.14	0.34

NOTE. Mean ADAP eligibility is calculated for the HAART era, 1996 and later. All individual-level variables are binary. Same-sex sex, multiple partners, and extra-marital affair are only available from 1988 to 2008.

Table 2.3 Same-Sex Marriage Bans and Syphilis

	Dependent Variable: Syphilis rate					
	OLS			DYNAMIC PANEL		
	(1)	(2)	(3)	(4)	(5)	(6)
Regression A						
Any ban	0.245	0.305	0.237	0.081	0.151	0.123
	(0.128) *	(0.150) **	(0.129) *	(0.055)	(0.082) *	(0.084)
N	1317	1317	1317	1236	1236	1236
Regression B						
Only statutory ban	0.256	0.307	0.234	0.080	0.150	0.121
	(0.128) *	(0.150) **	(0.127) *	(0.055)	(0.082) *	(0.083)
Both constitutional & statutory ban	0.116	0.254	0.329	0.083	0.179	0.159
	(0.145)	(0.212)	(0.201)	(0.066)	(0.094) *	(0.115)
N	1317	1317	1317	1236	1236	1236
Regression C						
Only marriage ban	0.263	0.313	0.229	0.077	0.145	0.110
	(0.138) *	(0.155) **	(0.129) *	(0.059)	(0.085) *	(0.083)
Both marriage & civil union ban	0.149	0.250	0.306	0.101	0.189	0.228
	(0.120)	(0.190)	(0.200)	(0.050) **	(0.084) **	(0.121) *
N	1317	1317	1317	1236	1236	1236
State-Specific Trend	No	Yes	Yes	No	Yes	Yes
State-Specific Trend ^2	No	No	Yes	No	No	Yes
Includes California	Yes	Yes	Yes	Yes	Yes	Yes
	OLS			DYNAMIC PANEL		
	(1)	(2)	(3)	(4)	(5)	(6)
Regression A						
Any ban	0.171	0.152	0.174	0.048	0.061	0.079
	(0.124)	(0.139)	(0.140)	(0.050)	(0.071)	(0.083)
N	1289	1289	1289	1209	1209	1209
Regression B						
Only statutory ban	0.182	0.151	0.169	0.048	0.059	0.077
	(0.125)	(0.138)	(0.135)	(0.051)	(0.070)	(0.081)
Both constitutional & statutory ban	0.068	0.160	0.345	0.047	0.110	0.134
	(0.144)	(0.226)	(0.234)	(0.062)	(0.087)	(0.121)
N	1289	1289	1289	1209	1209	1209
Regression C						
Only marriage ban	0.186	0.150	0.162	0.043	0.048	0.067
	(0.136)	(0.143)	(0.139)	(0.055)	(0.074)	(0.082)
Both marriage & civil union ban	0.105	0.159	0.262	0.069	0.121	0.166
	(0.115)	(0.190)	(0.202)	(0.044)	(0.069) *	(0.109)
N	1289	1289	1289	1209	1209	1209
State-Specific Trend	No	Yes	Yes	No	Yes	Yes
State-Specific Trend ^2	No	No	Yes	No	No	Yes
Includes California	No	No	No	No	No	No

NOTE. Dependent variables are logged. All regressions include state fixed effects, year effects, and state-year controls and are weighted by state population share. Numbers in parentheses are robust standard errors adjusted for clustering on states. A double asterisk indicates significance at the 5% level, and a single asterisk indicates significance at the 10% level.

Table 2.4 Same-Sex Marriage Bans and Gonorrhea

	Dependent Variable: Gonorrhea rate					
	OLS			DYNAMIC PANEL		
	(1)	(2)	(3)	(4)	(5)	(6)
Regression A						
Any ban	0.073	0.031	-0.002	0.019	0.032	0.010
	(0.058)	(0.063)	(0.046)	(0.028)	(0.038)	(0.037)
N	1372	1372	1372	1323	1323	1323
Regression B						
Only statutory ban	0.074	0.031	-0.002	0.018	0.031	0.010
	(0.057)	(0.064)	(0.047)	(0.028)	(0.039)	(0.038)
Both constitutional &	0.061	0.031	0.003	0.032	0.059	0.016
statutory ban	(0.094)	(0.069)	(0.054)	(0.041)	(0.041)	(0.045)
N	1372	1372	1372	1323	1323	1323
Regression C						
Only marriage ban	0.077	0.038	0.002	0.018	0.030	0.009
	(0.057)	(0.066)	(0.050)	(0.028)	(0.041)	(0.041)
Both marriage & civil	0.054	-0.011	-0.036	0.025	0.042	0.014
union ban	(0.090)	(0.065)	(0.051)	(0.041)	(0.040)	(0.039)
N	1372	1372	1372	1323	1323	1323
State-Specific Trend	No	Yes	Yes	No	Yes	Yes
State-Specific Trend ^2	No	No	Yes	No	No	Yes
Includes California	Yes	Yes	Yes	Yes	Yes	Yes
	OLS			DYNAMIC PANEL		
	(1)	(2)	(3)	(4)	(5)	(6)
Regression A						
Any ban	0.037	-0.065	-0.034	-0.003	-0.024	-0.019
	(0.058)	(0.051)	(0.047)	(0.023)	(0.033)	(0.037)
N	1344	1344	1344	1296	1296	1296
Regression B						
Only statutory ban	0.035	-0.067	-0.036	-0.005	-0.026	-0.020
	(0.058)	(0.052)	(0.047)	(0.022)	(0.035)	(0.037)
Both constitutional &	0.054	-0.004	0.043	0.024	0.038	0.034
statutory ban	(0.094)	(0.079)	(0.054)	(0.040)	(0.044)	(0.046)
N	1344	1344	1344	1296	1296	1296
Regression C						
Only marriage ban	0.038	-0.066	-0.033	-0.005	-0.031	-0.021
	(0.059)	(0.054)	(0.052)	(0.023)	(0.037)	(0.041)
Both marriage & civil	0.032	-0.060	-0.043	0.009	0.010	-0.000
union ban	(0.085)	(0.067)	(0.052)	(0.036)	(0.037)	(0.035)
N	1344	1344	1344	1296	1296	1296
State-Specific Trend	No	Yes	Yes	No	Yes	Yes
State-Specific Trend ^2	No	No	Yes	No	No	Yes
Includes California	No	No	No	No	No	No

NOTE. Dependent variables are logged. All regressions include state fixed effects, year effects, and state-year controls and are weighted by state population share. Numbers in parentheses are robust standard errors adjusted for clustering on states. A double asterisk indicates significance at the 5% level, and a single asterisk indicates significance at the 10% level.

Table 2.5 Vote in Favor of Same-Sex Marriage Bans

	Dependent Variable: Syphilis rate					
	OLS			DYNAMIC PANEL		
	(1)	(2)	(3)	(4)	(5)	(6)
Regression D						
Only statutory ban	0.219	0.307	0.261	0.080	0.160	0.141
	(0.127) *	(0.150) **	(0.127) **	(0.055)	(0.081) *	(0.082) *
Vote in favor of	0.000	0.004	0.006	0.001	0.003	0.003
constitutional ban	(0.002)	(0.003)	(0.003) *	(0.001)	(0.001) **	(0.002) *
N	1317	1317	1317	1236	1236	1236
State-Specific Trend	No	Yes	Yes	No	Yes	Yes
State-Specific Trend ^2	No	No	Yes	No	No	Yes
Includes California	Yes	Yes	Yes	Yes	Yes	Yes
Regression D						
Only statutory ban	0.147	0.151	0.190	0.047	0.069	0.094
	(0.126)	(0.142)	(0.137)	(0.051)	(0.071)	(0.082)
Vote in favor of	-0.000	0.002	0.006	0.001	0.002	0.003
constitutional ban	(0.002)	(0.003)	(0.003) *	(0.001)	(0.001)	(0.002)
N	1289	1289	1289	1209	1209	1209
State-Specific Trend	No	Yes	Yes	No	Yes	Yes
State-Specific Trend ^2	No	No	Yes	No	No	Yes
Includes California	No	No	No	No	No	No
	Gonorrhea rate					
	OLS			DYNAMIC PANEL		
	(1)	(2)	(3)	(4)	(5)	(6)
Regression D						
Only statutory ban	0.058	0.027	0.005	0.015	0.033	0.016
	(0.058)	(0.066)	(0.049)	(0.028)	(0.041)	(0.039)
Vote in favor of	0.000	0.000	0.000	0.000	0.001	0.000
constitutional ban	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
N	1372	1372	1372	1323	1323	1323
State-Specific Trend	No	Yes	Yes	No	Yes	Yes
State-Specific Trend ^2	No	No	Yes	No	No	Yes
Includes California	Yes	Yes	Yes	Yes	Yes	Yes
Regression D						
Only statutory ban	0.020	-0.074	-0.035	-0.009	-0.027	-0.018
	(0.058)	(0.055)	(0.047)	(0.022)	(0.036)	(0.038)
Vote in favor of	0.000	-0.000	0.001	0.000	0.001	0.001
constitutional ban	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
N	1344	1344	1344	1296	1296	1296
State-Specific Trend	No	Yes	Yes	No	Yes	Yes
State-Specific Trend ^2	No	No	Yes	No	No	Yes
Includes California	No	No	No	No	No	No

NOTE. Dependent variables are logged. All regressions include state fixed effects, year effects, and state-year controls and are weighted by state population share. Numbers in parentheses are robust standard errors adjusted for clustering on states. A double asterisk indicates significance at the 5% level, and a single asterisk indicates significance at the 10% level.

Table 2.6 Dynamic Effects of Same-Sex Marriage Bans

	Dependent Variable					
	Syphilis rate					
Regression E	(1)	(2)	(3)	(4)	(5)	(6)
First 2 years after ban	0.219	0.229	0.178	0.176	0.106	0.121
	(0.104) **	(0.135) *	(0.128)	(0.111)	(0.126)	(0.133)
Years 3-6	0.300	0.306	0.152	0.181	0.084	0.053
	(0.171) *	(0.230)	(0.222)	(0.154)	(0.203)	(0.234)
Year 7 onwards	0.147	0.179	0.075	0.080	-0.009	-0.007
	(0.188)	(0.269)	(0.248)	(0.184)	(0.257)	(0.258)
N	1317	1317	1317	1289	1289	1289
State-Specific Trend	No	Yes	Yes	No	Yes	Yes
State-Specific Trend ^2	No	No	Yes	No	No	Yes
Includes California	Yes	Yes	Yes	No	No	No
	Gonorrhea rate					
Regression E	(1)	(2)	(3)	(4)	(5)	(6)
First 2 years after ban	0.036	0.023	-0.004	0.017	-0.045	-0.023
	(0.052)	(0.059)	(0.044)	(0.055)	(0.062)	(0.049)
Years 3-6	0.117	0.068	0.005	0.044	-0.069	-0.050
	(0.083)	(0.102)	(0.070)	(0.074)	(0.091)	(0.059)
Year 7 onwards	0.090	0.022	0.007	0.054	-0.077	-0.005
	(0.082)	(0.117)	(0.078)	(0.084)	(0.130)	(0.077)
N	1372	1372	1372	1344	1344	1344
State-Specific Trend	No	Yes	Yes	No	Yes	Yes
State-Specific Trend ^2	No	No	Yes	No	No	Yes
Includes California	Yes	Yes	Yes	No	No	No

NOTE. Dependent variables are logged. All regressions include state fixed effects, year effects, and state-year controls and are weighted by state population share. Numbers in parentheses are robust standard errors adjusted for clustering on states. A double asterisk indicates significance at the 5% level, and a single asterisk indicates significance at the 10% level.

Table 2.7 Same-Sex Marriage and Civil Union Laws

	Dependent Variable: Syphilis rate						Gonorrhea rate					
	OLS			DYNAMIC PANEL			OLS			DYNAMIC PANEL		
Regression F	(1)	(2)	(3)	(4)	(5)	(6)	(1)	(2)	(3)	(4)	(5)	(6)
Ban marriage & civil union	0.149	0.247	0.312	0.095	0.185	0.227	0.042	-0.017	-0.033	0.021	0.040	0.015
	(0.115)	(0.199)	(0.206)	(0.049) *	(0.087) **	(0.120) *	(0.090)	(0.070)	(0.052)	(0.041)	(0.041)	(0.040)
Ban marriage & not ban	0.210	0.237	0.215	0.065	0.129	0.113	0.046	-0.011	-0.005	0.007	0.011	0.005
or allow civil union	(0.130)	(0.138) *	(0.128) *	(0.059)	(0.084)	(0.081)	(0.055)	(0.057)	(0.049)	(0.024)	(0.036)	(0.039)
Ban marriage & allow	0.806	1.015	0.630	0.153	0.307	0.032	0.295	0.474	0.214	0.109	0.217	0.146
civil union	(0.253) **	(0.305) **	(0.307) **	(0.094)	(0.115) **	(0.142)	(0.113) **	(0.102) **	(0.134)	(0.069)	(0.083) **	(0.120)
Allow marriage or allow civil	0.218	0.115	-0.077	-0.025	-0.085	-0.086	-0.033	0.006	-0.042	0.000	0.015	0.023
union & not ban marriage	(0.215)	(0.252)	(0.281)	(0.096)	(0.126)	(0.174)	(0.079)	(0.109)	(0.096)	(0.038)	(0.040)	(0.055)
N	1317	1317	1317	1236	1236	1236	1372	1372	1372	1323	1323	1323
State-Specific Trend	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
State-Specific Trend ^2	No	No	Yes	No	No	Yes	No	No	Yes	No	No	Yes
Includes California	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
	OLS			DYNAMIC PANEL			OLS			DYNAMIC PANEL		
Regression F	(1)	(2)	(3)	(4)	(5)	(6)	(1)	(2)	(3)	(4)	(5)	(6)
Ban marriage & civil union	0.121	0.162	0.263	0.063	0.119	0.170	0.022	-0.062	-0.041	0.005	0.011	0.003
	(0.113)	(0.190)	(0.204)	(0.045)	(0.069)*	(0.110)	(0.086)	(0.069)	(0.051)	(0.037)	(0.037)	(0.035)
Ban marriage & not ban	0.197	0.149	0.167	0.039	0.051	0.075	0.031	-0.072	-0.028	-0.007	-0.030	-0.017
or allow civil union	(0.135)	(0.145)	(0.140)	(0.056)	(0.074)	(0.081)	(0.061)	(0.055)	(0.051)	(0.024)	(0.037)	(0.041)
Ban marriage & allow	0.289	0.190	0.058	0.005	-0.027	-0.124	0.065	0.181	-0.158	-0.058	-0.058	-0.174
civil union	(0.320)	(0.217)	(0.512)	(0.134)	(0.098)	(0.221)	(0.147)	(0.223)	(0.121)	(0.059)	(0.097)	(0.087) *
Allow marriage or allow civil	0.150	0.088	-0.149	-0.051	-0.104	-0.187	-0.095	0.008	-0.070	-0.034	0.004	-0.015
union & not ban marriage	(0.202)	(0.259)	(0.283)	(0.098)	(0.128)	(0.146)	(0.088)	(0.104)	(0.085)	(0.036)	(0.037)	(0.041)
N	1289	1289	1289	1209	1209	1209	1344	1344	1344	1296	1296	1296
State-Specific Trend	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
State-Specific Trend ^2	No	No	Yes	No	No	Yes	No	No	Yes	No	No	Yes
Includes California	No	No	No	No	No	No	No	No	No	No	No	No

NOTE. Dependent variables are logged. All regressions include state fixed effects, year effects, and state-year controls and are weighted by state population share. Numbers in parentheses are robust standard errors adjusted for clustering on states. A double asterisk indicates significance at the 5% level, and a single asterisk indicates significance at the 10% level.

Table 2.8 Tolerance for Same-Sex Sex and Teen Sex Using GSS Microdata

	Dependent Variable							
	Tolerance for same-sex sex		Tolerance for teen sex		Tolerance for same-sex sex		Tolerance for teen sex	
	(1)		(2)		(3)		(4)	
	Leads of Laws (t+2)				Lags of Laws (t-2)			
Regression A								
Any ban	-0.033		-0.006		-0.008		0.003	
	(0.012)	**	(0.010)		(0.013)		(0.011)	
Regression B								
Only statutory ban	-0.033		0.003		-0.007		0.003	
	(0.012)	**	(0.011)		(0.012)		(0.011)	
Both constitutional &	-0.045		-0.009		-0.030		0.006	
statutory ban	(0.022)	**	(0.016)		(0.026)		(0.020)	
Regression C								
Only marriage ban	-0.032		-0.005		-0.000		0.005	
	(0.012)	**	(0.010)		(0.013)		(0.012)	
Both marriage & civil	-0.043		-0.014		-0.041		-0.007	
union ban	(0.020)	**	(0.015)		(0.024)	*	(0.017)	
Regression F								
Ban marriage & civil union	-0.044		-0.010		-0.038		-0.006	
	(0.020)	**	(0.017)		(0.024)		(0.016)	
Ban marriage & not ban	-0.038		-0.005		-0.005		0.013	
or allow civil union	(0.012)	**	(0.010)		(0.013)		(0.012)	
Ban marriage & allow	-0.002		0.012		0.038		-0.043	
civil union	(0.023)		(0.029)		(0.021)	*	(0.028)	
Allow marriage or allow civil	-0.022		0.012		-0.003		0.053	
union & not ban marriage	(0.034)		(0.017)		(0.039)		(0.038)	
N	22845		18674		22845		18674	

NOTE. All specifications include state fixed effects, year effects, and individual-level controls, including age, gender, race, education, and religious attendance. Numbers in parentheses are robust standard errors adjusted for clustering on state-year. A double asterisk indicates significance at the 5% level, and a single asterisk indicates significance at the 10% level.

Table 2.9 Sexual Behaviors Using GSS Microdata

	Dependent Variable					
	Same-sex sex in past year	Multiple partners in past year	Extra-marital affair ever	Same-sex sex in past year	Multiple partners in past year	Extra-marital affair ever
	(1)	(2)	(3)	(4)	(5)	(6)
	Leads of Laws (t+2)			Lags of Laws (t-2)		
	Male respondents					
Regression A						
Any ban	-0.010	-0.009	0.012	-0.018	-0.010	-0.048
	(0.007)	(0.014)	(0.022)	(0.008) **	(0.015)	(0.019) **
Regression B						
Only statutory ban	-0.008	-0.021	-0.007	-0.017	-0.011	-0.049
	(0.007)	(0.014)	(0.023)	(0.008) **	(0.015)	(0.019) **
Both constitutional & statutory ban	-0.013	-0.025	0.007	-0.035	-0.003	-0.037
	(0.011)	(0.022)	(0.030)	(0.011) **	(0.026)	(0.031)
Regression C						
Only marriage ban	-0.009	-0.007	0.015	-0.016	-0.011	-0.037
	(0.007)	(0.014)	(0.022)	(0.008) **	(0.015)	(0.019) *
Both marriage & civil union ban	-0.017	-0.018	-0.003	-0.030	-0.006	-0.111
	(0.012)	(0.020)	(0.025)	(0.011) **	(0.023)	(0.024) **
Regression F						
Ban marriage & civil union	-0.014	-0.014	-0.005	-0.031	0.003	-0.100
	(0.013)	(0.020)	(0.027)	(0.011) **	(0.024)	(0.024) **
Ban marriage & not ban or allow civil union	-0.010	-0.008	0.005	-0.016	-0.014	-0.044
	(0.007)	(0.014)	(0.023)	(0.008) **	(0.014)	(0.019) **
Ban marriage & allow civil union	0.004	0.007	0.054	-0.019	0.044	0.060
	(0.011)	(0.030)	(0.037)	(0.011) *	(0.051)	(0.043)
Allow marriage or allow civil union & not ban marriage	0.013	0.017	-0.016	-0.013	0.085	0.085
	(0.015)	(0.030)	(0.032)	(0.015)	(0.030) **	(0.054)
N	8956	10620	6618	8956	10620	6618
	Female respondents					
Regression A						
Any ban	-0.001	-0.007	0.011	-0.023	-0.013	-0.033
	(0.006)	(0.009)	(0.012)	(0.006) **	(0.008)	(0.012) **
Regression B						
Only statutory ban	-0.005	-0.009	0.017	-0.022	-0.013	-0.031
	(0.007)	(0.009)	(0.012)	(0.006) **	(0.008)	(0.012) **
Both constitutional & statutory ban	-0.013	-0.023	0.004	-0.037	-0.015	-0.063
	(0.010)	(0.014) *	(0.021)	(0.012) **	(0.017)	(0.020) **
Regression C						
Only marriage ban	-0.002	-0.008	0.014	-0.022	-0.017	-0.030
	(0.006)	(0.009)	(0.012)	(0.007) **	(0.009) *	(0.012) **
Both marriage & civil union ban	0.008	0.003	-0.004	-0.030	0.008	-0.047
	(0.008)	(0.013)	(0.020)	(0.010) **	(0.015)	(0.023) **
Regression F						
Ban marriage & civil union	0.013	0.007	-0.005	-0.031	0.011	-0.051
	(0.009)	(0.013)	(0.021)	(0.010) **	(0.015)	(0.023) **

Ban marriage & not ban	0.000	-0.006	0.016	-0.024	-0.015	-0.026
or allow civil union	(0.006)	(0.009)	(0.012)	(0.006) **	(0.008) *	(0.012) **
Ban marriage & allow	-0.002	-0.009	0.005	-0.005	-0.015	-0.089
civil union	(0.016)	(0.019)	(0.023)	(0.028)	(0.039)	(0.017) **
Allow marriage or allow civil	0.020	0.021	-0.000	-0.025	0.057	-0.031
union & not ban marriage	(0.016)	(0.016)	(0.029)	(0.012) **	(0.037)	(0.026)
N	10134	13654	9247	10134	13654	9247

NOTE. All specifications include state fixed effects, year effects, and individual-level controls, including age, gender, race, education, and religious attendance. Numbers in parentheses are robust standard errors adjusted for clustering on state-year. A double asterisk indicates significance at the 5% level, and a single asterisk indicates significance at the 10% level.

Appendix

Legal References and Notes (current as of end of year 2012)

ALABAMA
Alabama Code § 30-1-19 (1998)
Alabama Const. art. I, §36.03 (2006)

ALASKA
Alaska Stat. § 25.05.013 (1996)
Alaska Const. art. 1, § 25 (1998)

ARIZONA
Arizona Rev. Stat. Ann. § 25-101 (1996)
Arizona Rev. Stat. Ann. § 25-112 (1996)
Arizona Rev. Stat. Const. art. 30, § 1 (2008)

ARKANSAS
Arkansas Code Ann. § 9-11-208 (1997)
Arkansas Code Ann. § 9-11-107 (1997)
Arkansas Code Ann. § 9-11-109 (1997)
Arkansas Const. 83 § 1-3 (2004)

CALIFORNIA
California Fam. Code § 308.5 (2000)
California Fam. Code § 299.2 (2003)
California Fam. Code § 297.5 (2003)
In re Marriage Cases (Supreme Court, May 15, 2008)
California Const. art. 1, § 7.5 (2008)
Strauss v. Horton (CA Supreme Court, May 26, 2009)

NOTE. In 2008, state court struck down 2000 law. Some marriages occurred. Later that year, constitutional ban was passed. Those marriages were still honored. In 2003, a bill allowing civil unions was passed but became effective in 2005.

COLORADO
Colorado Rev. Stat. § 14-2-104 (2000)
Colorado Const., art. II, §31 (2006)

NOTE. In 2009, CO passed a law giving limited benefits to same-gender couples (Colorado Rev. Stat. § 15-22-101 et seq).

CONNECTICUT
Connecticut Gen. Stat. Ann. § 45a-727a (2000)
Connecticut Gen. Stat. Ann. § 46b-38aa-oo (2005)
Kerrigan v. Comm'r of Pub. Health et al., 289 Conn. 135, 957 A.2d 407 (Conn. Oct. 10, 2008)

DELAWARE
Delaware Code Ann. tit. 13, § 101 (1996)

DISTRICT OF COLUMBIA
District of Columbia Code § 32-702 (1992)
District of Columbia Code § 32-704, 32-705, and 32-706
56 District of Columbia Reg. 3797 (2009)

NOTE. D.C. Code § 32-702 was effective in 2002, but passed in 1992. The Religious Freedom & Civil Marriage Equality Amendment Act of 2009 was signed into law on Dec. 15, 2009 by Mayor Adrian Fenty. Licenses became available in 2010.

FLORIDA
Florida Stat. Ann. § 741.04 (1977)
Florida Stat. Ann. § 741.212 (1997)
Florida Const. art. I, § 27 (2008)

GEORGIA
Georgia Code Ann. § 19-3-30 (1996)
Georgia Code Ann. § 19-3-3.1 (1996)
Georgia Const. art. 1, § 4, para. 1 (2004)

HAWAII
Hawaii Rev. Stat. Ann. § 572-1 (1994)
Hawaii Rev. Stat. Ann. § 572-1.6 (1994)
Hawaii Const. art. I, § 23 (1998)
Hawaii Senate Bill 232 (2011)

NOTE. HA had a "reciprocal beneficiaries" (1997) law, and legalized civil unions in 2011.

IDAHO
Idaho Code § 32-209 (1996)
Idaho Const. art. III, §28 (2006)

ILLINOIS
Illinois Comp. Stat. Ann. 750 § 5/201 (1977)
Illinois Comp. Stat. Ann. 750 § 5/213.1 (1996)
Illinois Comp. Stat. Ann. 750 § 5/212 (1996)
Illinois Senate Bill 1716 (2011)

NOTE. IL had a 1977 law validating marriage between a man and a woman, which was probably not meant to rule out same-sex marriage, so we code IL as introducing a law restricting marriage to one man and one woman in 1996.

INDIANA
Indiana Code Ann. § 31-11-1-1 (1997)

IOWA
Iowa Code Ann. § 595.2 (1998)
Varnum v. Brien (Iowa Supreme Court 2009, 763 N.W.2d 862, 2009 Iowa Sup. LEXIS 31)

KANSAS
Kansas Stat. Ann. § 23-101 (1980)
Kansas Stat. Ann. § 23-115 (1996)
Kansas Const. art. 15, sec. 16 (2005)

NOTE. KS had a 1980 law validating marriage between a man and a woman, which was probably not meant to rule out same-sex marriage, so we code KS as introducing a law restricting marriage to one man and one woman in 1996.

KENTUCKY
Kentucky Rev. Stat. Ann. § 402.005 (1998)
Kentucky Rev. Stat. Ann. § 402.020 (1998)
Kentucky Rev. Stat. Ann. § 402.040 (1998)
Kentucky Rev. Stat. Ann. § 402.045 (1998)
Kentucky Const. § 233A (2004)

LOUISIANA
Louisiana Civ. Code Ann. art. 89 (1988)
Louisiana Civ. Code Ann. art. 96 (1988)
H. Con. Res. 124, 1996 Reg. Sess.
Louisiana Civ. Code Ann. art. 3520(B) (1999)
Louisiana Const. art. XII, § 15 (2004)

NOTE. 1999 statute 3520 also bans marriage between two persons of the same gender.

MAINE
Maine Rev. Stat. Ann. tit. 19-A, § 650, 651, 701 (2012)

NOTE. In 2004, domestic partnerships registry was established. Some but not all rights are awarded (ME. REV. STAT. ANN. tit. 22, §2710). In 2009, a same-sex marriage bill was signed into law, but it was approved by voters only in 2012.

MARYLAND
Maryland Code Ann., Fam. Law § 2-201 (2012)

NOTE. In 2008, limited rights were given to same-gender couples. See Public Law No. 590 and 599 (2008). In 2012, a same-sex marriage bill was signed into law and was upheld by voters.

MASSACHUSETTS
Goodridge v. Dep't of Pub. Health (2003) 440 Mass 309, 798 NE2d 941, 2003 Mass LEXIS 814 (Massachusetts Superior Court, Nov. 18, 2003)

MICHIGAN
Michigan Comp. Laws Ann. § 551.1 (1996)
Michigan Comp. Laws Ann. § 551.271 (1996)
Michigan Comp. Laws Ann. § 551.272 (1996)
Michigan Comp. Laws Ann. § 551.3 (1996)
Michigan Comp. Laws Ann. § 551.4 (1996)
Michigan Const. art. I, § 25 (2004)

MINNESOTA
Minnesota Stat. Ann. § 517.01 (1997)
Minnesota Stat. Ann. § 517.03 (1997)

NOTE. In the chapter of Minnesota Stat. Ann. that addresses human rights issues, it is written that: "Nothing in this chapter shall be construed to: (1) mean the state of Minnesota condones homosexuality or bisexuality or any equivalent lifestyle" (MINN. STAT. ANN. § 363A.27, 1993). In November 2012, for the first time nationally, voters disapproved of a proposed constitutional amendment banning same-sex marriage.

MISSISSIPPI
Mississippi Code Ann. § 93-1-1 (1997)
Mississippi Const. art. 14, § 263A (2004)

MISSOURI
Missouri Ann. Stat. § 451.022 (1996, 2001)
Missouri Const. art. I, § 33 (2004)

NOTE. The 1996 law was ruled unconstitutional in St. Louis Health Care Network v. State, 968 S.W.2d 145 (1998) because the title of the act was misleading. The 2001 law was passed with the same language as the 1996 law.

MONTANA
Montana Code Ann. § 40-1-103 (1997)
Montana Code Ann. § 40-1-401 (1997)
Montana Const. art. XIII, § 7 (2004)

NEBRASKA
Nebraska Const. art. I, § 29 (2000)

NOTE. In May 2005, a federal court declared the amendment unconstitutional. A federal appeals court then reinstated it in July 2006.

NEVADA
Nevada Const. art. I, § 21 (2002)
Nevada Senate Bill 283 (2009)

NOTE. In 2009, the Assembly overrode the veto of Gov. Jim Gibbons on a bill giving legal rights to domestic partners. The constitutional amendment was also ratified in 2000.

NEW HAMPSHIRE
New Hampshire Rev. Stat. Ann. § 457:1 (1987)
New Hampshire Rev. Stat. Ann. § 457:2 (1987)
New Hampshire Rev. Stat. Ann. 457-A (2007)
New Hampshire HB 436 (2009)

New Hampshire HB 73 (2009)

NOTE. In 2009, the 2007 civil unions law was repealed but civil unions were issued until 2010.

NEW JERSEY
New Jersey Stat. Ann. §37:1; §37:2; §26:8 (2006)

NOTE. This law was prompted by the case of Lewis v. Harris (2006).

NEW MEXICO
NOTE. NM does not have any explicit same-sex marriage restrictions. In 2004, more than 60 marriage licenses were issued in one county but they were declared invalid by the state attorney general.

NEW YORK
A8354-2011 (Marriage Equality Act, 2011)

NOTE. In 2008, NY began to recognize out of state same-sex marriage licenses. Please see Martinez v. County of Monroe, 850 NYS2d 740 (4th Dept. 2008). In 2011, New York passed a law allowing same-sex marriage.

NORTH CAROLINA
North Carolina Gen. Stat. § 51-1.2 (1995)
North Carolina Const. art. 14, § 6 (2012)

NORTH DAKOTA
North Dakota Cent. Code § 14-03-01 (1997)
North Dakota Cent. Code § 14-03-08 (1997)
North Dakota Const. art. XI, § 28 (2004)

OHIO
Ohio Rev. Code Ann. § 3101.01 (2004)
Ohio Const. art. XV, § 11 (2004)

OKLAHOMA
Oklahoma Stat. Ann. tit. 43, § 3 (1975)
Oklahoma Stat. Ann. tit. 43, § 3.1 (1996)
Oklahoma Const. art. II, § 35 (2004)

NOTE. Amendment also says that no part of law should be construed to confer the benefits of marriage on unmarried couples.

OREGON
Oregon Rev. Stat. § 106.010 (1975)
Oregon Const. art. XV, § 5a (2004)
Oregon Family Fairness Act, Public Law Number 99, 2007

PENNSYLVANIA
Pennsylvania Cons. Stat. Ann. 23 § 1704 (1996)

RHODE ISLAND
NOTE. RI does not have any explicit same-sex marriage restrictions, but it has limited benefits for same-gender couples. See R.I. GEN. LAWS §28-48-1, 36-12-4, 44-30-12, 45-49-4.3, 5-33.2-24.

SOUTH CAROLINA
South Carolina Code Ann. § 20-1-10 (1996)
South Carolina Code Ann. § 20-1-15 (1996)
South Carolina Const. Ann. art. XVII, § 15 (2006)

NOTE. The constitutional amendment was voted on in 2006 but enacted in 2007.

SOUTH DAKOTA
South Dakota Cod. Laws § 25-1-1 (1996)

South Dakota Cod. Laws § 25-1-38 (2000)
South Dakota Const. art XXI, § 9 (2006)

TENNESSEE
Tennessee Code Ann. § 36-3-113 (1996)
Tennessee Const. art. XI, § 18 (2006)

TEXAS
Texas Fam. Code Ann. § 2.001 (1997)
Texas Fam. Code Ann. § 6.204 (2003)
Texas Const. art. I, § 32 (2005)

NOTE. Texas originally banned marriage between two persons of the same sex in 1973 and renewed the law in 1997.

UTAH
Utah Code Ann. § 30-1-2 (1977)
Utah Code Ann. § 30-1-4 (1996)
Utah Code Ann. § 30-1-4.1 (2004)
Utah Const. art. I, § 29 (2004)

VERMONT
Vermont Stat. Ann. tit. 15, § 8 (2000, 2009)
Vermont Stat. Ann. tit. 15 § 1201-1207 (1999)

NOTE. In 2009, the Vermont Legislature overrode the Governor's veto legalizing same-sex marriage.

VIRGINIA
Virginia Code Ann. § 20-45.2 (1975, 1997)
Virginia Code Ann. S 20-45.3 (2004)
Virginia Const. art. I, § 15-A (2006)

NOTE. In 1997, VA amended and reenacted their 1975 ban on marriage between two persons of the same sex.

WASHINGTON
Washington Rev. Code § 26.04.010 (1998)
Washington Rev. Code § 26.04.020 (1998)
Washington Rev. Code § 1.12.080 (2009)
Washington Rev. Code § 26.04.010 (2012)
Washington Rev. Code § 26.04.020 (2012)

NOTE. Between 2007 and 2009, WA passed laws giving same-gender couples the same rights as married opposite-gender couples. HB 2516 was passed by the legislature, signed into law by the governor in February 2012, and was upheld by voters in November 2012.

WEST VIRGINIA
West Virginia Code Ann. § 48-2-104 (2000)
West Virginia Code Ann. § 48-2-603 (2000)

NOTE. The 2000 WV law requires that every marriage license include the "the full names of both the female and male parties" and the following statement: "Marriage is designed to be a loving and lifelong union between a woman and a man."

WISCONSIN
Wisconsin Stat. § 765.01 (1979)
Wisconsin Const. art. XIII, § 13 (2006)

NOTE. WI offers some benefits to same-gender couples but not full benefits.

WYOMING
Wyoming Stat. Ann. § 20-1-101 (1977)

Chapter 3

Weight and Labor Market Outcomes: Understanding the Role of Teenage Weight Status

Handie Peng*

Teenage obesity has increasingly become a public health concern. Therefore, understanding the role that teenage weight status plays is important. This paper looks at whether the relationship between weight and labor market outcomes differs for individuals with different teenage weight statuses. We run separate regressions on stratified samples to test whether the returns on various characteristics differ for individuals with different teenage weight statuses. The results indicate that there is indeed a heterogeneous effect: specifically, being overweight is associated with a wage *premium* for adult white males who were healthy weight teens, while it is associated with a wage penalty for adult white males who were obese or overweight teens. We also look at how much of the wage and employment differentials between individuals with different teenage weight statuses are explained by observed characteristics. To this end, we carry out a Blinder-Oaxaca decomposition. We find that for white males and non-white females who were obese or overweight teens, their lower wages are explained by their lower educational attainment. White females who were obese or overweight teens tend to have a lower employment probability, and this differential is explained by their higher current body mass index (BMI), lower education, and fewer years of work experience.

JEL codes: I12, J31
Keywords: Wages, Employment, BMI, Overweight, Obesity, Teenage Weight Status.

---

* Handie Peng, Department of Economics, Emory University, Atlanta, GA 30322-2240. Email: hpeng4@emory.edu. The author is extremely grateful to Andrew Francis, David Frisvold, Matthew Higgins, Sara Markowitz, and Hugo Mialon for helpful comments.

I. Introduction

Pediatric obesity[29] has increasingly become a public health concern, given its many negative health consequences. Pediatric obesity is associated with psychological or psychiatric problems, low self-esteem and behavioral problems, cardiovascular risk factors, asthma, type 1 diabetes, and low grade systemic inflammation (Reilly et al., 2003). In addition, the effects of pediatric obesity persist through adulthood (Reilly et al., 2003). Given rapidly increasing health care costs (Centers for Medicare and Medicaid Services, 2012), reduced earning potential caused by pediatric obesity means that some, if not most, of the financial burden would have to be shared by society. Understanding how pediatric obesity influences individuals' subsequent labor market outcomes is important not only for these individuals but also for the general public.

The wage penalty associated with excessive weight is well documented in the economics literature (Gortmaker et al., 1993; Averett and Korenman, 1996; Pagan and Davila, 1997; Cawley, 2004; Baum and Ford, 2004; Conley and Glauber, 2005; Mocan and Tekin, 2009; Han et al., 2011). This paper endeavors to improve on this literature by investigating whether the labor market penalty associated with excessive weight differs by teenage weight status. In the public health literature, it is found that obesity in childhood or adolescence is associated with lower self-esteem, and weight loss is associated with increased self-esteem for obese children or adolescents (French et al., 1995). Therefore, since obese or overweight teens might have very different

[29] Pediatric obesity includes childhood obesity and teenage obesity.

psychological experiences growing up, they might also have very different underlying returns on their characteristics, such as BMI, educational attainment and human capital.[30]

Previous literature generally finds that being obese or overweight has a negative impact on wages for white females, with mixed evidence for other gender/race groups (Gortmaker et al., 1993; Averett and Korenman, 1996; Cawley, 2004; Mocan and Tekin, 2009). By and large, researchers have not been able to find consistent effects of excessive weight on wages for men. For example, Mocan and Tekin (2009) conclude in their study that “[m]en’s wages are not impacted by their body weight.” Another interesting finding from previous research is that being obese or overweight in adolescence or young adulthood is negatively associated with a person’s subsequent labor market outcomes (Gortmaker et al., 1993; Averett and Korenman, 1996). Specifically, Averett and Korenman (1996) find that “the largest penalty is associated with obesity at the younger ages” instead of with current obesity. However, less attention has been paid to the mechanisms through which teenage weight status affects adult wages (Han et al., 2011) and whether employment probability is also affected by weight status (Cawley, 2000; Morris, 2007).

This paper looks at whether returns on employment-determining and wage-determining characteristics, for example, education, occupation, and contemporaneous weight, vary across groups with different teenage weight statuses: obese or overweight teens (OOT) vs. healthy weight teens (HWT).[31] Following the existent literature, we split the sample by gender and race/ethnicity. First, ordinary least squares regressions are

[30] Human capital in this paper refers to years of work experience and job tenure.

[31] Observations for individuals who are underweight are dropped from the analyses since this group is very different from the others, and is not the focus of this paper. Only less than 2% of the sample is eliminated in the process.

carried out separately for individuals with different teenage weight statuses, for each gender and race/ethnicity group. Then, individual fixed effects models are used in an effort to deal with endogeneity. Lastly, a Blinder-Oaxaca decomposition is utilized to see how much of the adult employment probability (wage) differential between OOT and HWT is attributable to the differences in their characteristics ("explained"), and how much is attributable to the differences in the returns on their characteristics ("unexplained").

Our empirical analyses make the following three contributions. First, we recognize the possibility that individuals with different teenage weight statuses might have disparate employment probability equations or wage equations with unequal returns on not only current weight, but also other characteristics such as education, human capital and occupation. Second, we focus on employment probability besides wages, which provides some insight into whether there is differential selection into the labor force by teenage weight status. Finally, we endeavor to find the mechanisms through which teenage weight status affects adult wage, above and beyond the effect of current weight status. The focus on teenage weight status might help us provide further evidence to support public policies that strive to reduce early-onset obesity/overweightness.

We find that for males, higher BMI is associated with lower wages if they were OOT, yet it is associated with *higher* wages if they were HWT. This is consistent with the explanation that there is statistical discrimination at individuals' younger ages; yet it is not consistent with the explanation that excessive weight reduces productivity, at least for white males. The idea is that when individuals first entered the labor market, their weight status was likely used as a selection criterion, hence HWT had higher wages than OOT;

over time, although some HWT become heavier, their wages also become higher than the wages for those who retain a healthy weight, which means the increased weight has not reduced their productivity. However, for females, higher BMI is associated with lower wages regardless of their teenage weight status. Due to insufficient variations in within-individual weight changes, most of the results, for either males or females, become insignificant when individual fixed effects models are employed.

The results from the Blinder-Oaxaca decomposition suggest that white males and non-white females who were OOT tend to invest less in their education, and black females who were OOT also tend to sort themselves into a blue-collar occupation, which is lower-paid. This is consistent with Han et al. (2011), who find that there is an indirect effect of weight on wages, which stems from reduced educational investment and occupational sorting by the obese teens. The improvement of our method, however, is that we are holding current weight constant and looking for the effect above and beyond that induced by current weight, while Han et al. (2011) use teen weight as the main explanatory variable, which does not take into account the role that contemporaneous weight can play. This distinction matters because we also find "unexplained" differential returns on contemporaneous weight for males. Finally, there is some suggestive evidence that white adult females who were OOT are more likely to drop out of the labor force than those who were HWT.

The rest of the paper is organized as follows: Section II briefly reviews the previous literature; Section III describes and summarizes the data; Section IV introduces the methods and econometric models; Section V displays the results; Section VI discusses the findings; and Section VI concludes.

II. Literature Review

The first study that investigates the association between weight and wages is Gortmaker et al. (1993). They focus on the relationship between being overweight in adolescence and young adulthood and subsequent educational attainment, marital status, household income, and self-esteem. They find that women who had been overweight had completed fewer years of school, were less likely to have married, had lower household incomes, and had higher rates of household poverty; while overweight men were less likely to have married. They find no evidence for an effect of being overweight on self-esteem. Similarly, our results suggest that obese or overweight teens grow up to have lower wages; and white males and non-white females who were OOT tend to have completed fewer years of education than their counterparts who were HWT.

Using the same data as Gortmaker et al. (1993), Averett and Korenman (1996) use both lagged BMI and same-sex sibling difference to deal with endogeneity. When the authors include both lagged BMI in 1981 and current BMI in 1988, they find that women who were obese or overweight at ages 16 to 24 had, at ages 23 to 31, lower family income and lower hourly wages. When using only current weight they find similar but smaller effects. This suggests that the effect of current weight on wages is not biased downward excessively by reverse causality. Employing an interaction term of obesity in 1981 and obesity in 1988, they confirm that the largest penalty is associated with obesity at younger ages. Among others, one important contribution of this study is that it points out the racial as well as gender differences in the wage penalty for obese persons. Specifically, white females suffer from the strongest wage penalty. Like Averett and Korenman (1996), we are also interested in separating the effects of lagged weight and contemporaneous weight on wages. Our improvement, however, is that we also hypothesize and actually find that the returns on other characteristics, such as education and human capital, are also different for individuals with different teenage weight statuses.

Following prior work, Cawley (2000) uses the NLSY79 (1998) data along with the child supplement. In this work, the author uses the child supplement to NLSY79 as the source of instruments to deal with endogeneity. Specifically, the BMI of children aged 6 to 9, interacted with their gender and age, serve as instruments for maternal weight. To justify the validity of the instruments, the author offers several arguments: (1) previous literature finds no consistent pattern between childhood obesity and socioeconomic status; (2) previous literature finds no measurable effect of the common household environment on body weight; and (3) by regressing maternal education and general intelligence on the set of instruments and other regressors, the author finds no significant correlation between the instruments and the proxies for the potential confounders (maternal education and general intelligence), suggesting that a child's weight is uncorrelated with the maternal wage residual. The finding of this study confirms that weight lowers wages for white women, although the IV estimates are not significant due to large standard errors. Cawley (2000) also finds weak evidence that a higher BMI reduces employment probability for females. While in theory it is possible for us to utilize Cawley's instruments in our study, in practice, because our sample is very stratified, the instruments do not work well. In addition, since the instruments are only applicable to women who have borne children, as the author himself points out, the empirical results obtained are likely not generalizable.

Baum and Ford (2004) attempt to understand the source of the wage differential for obese persons. The authors employ multiple empirical techniques in order to determine the effect of obesity on wages. Their results indicate that obese workers suffer a wage penalty in the range of 0.7 to 6.3 percent, although the sibling-difference model does not provide significant evidence of an obesity wage penalty. They test four hypotheses concerning the source of the wage differential for obese persons: (1) obese workers are less productive; (2) obese workers are more economically myopic; (3) employers fear it will be more costly to provide such workers with health care; and (4) obese workers are discriminated against by customers. Their results lend some support to the second hypothesis but not the other three. Although we do not specifically test these hypotheses, our results also lend no support to the hypothesis that workers with excessive weight are less productive.

Building upon his prior work, Cawley (2004) further investigates the wage penalty associated with excessive weight. He also uses multiple empirical techniques and makes similar arguments to Cawley (2000) to justify the validity of his instruments, which are the BMI, age and gender of a sibling. The author finds that, regardless of the econometric method used, weight appears to lower wages for white females; while for other gender and race/ethnicity groups, the findings are mixed. Besides the innovative instruments, another important contribution of this study is that it corrects for the reporting errors in weight and height, a convention that has since been followed by many studies in this area, including ours.

The last paper reviewed here is Han et al. (2011). Using data from NLSY79 (1998), the authors look for the direct and indirect effects of teen weight on wages. They assume that excessive weight in late-teens affects educational investment and occupational choices (which are also affected by education), and in turn wages later on. To obtain the total effect, they carry out regressions with education, types of occupations, and whether the occupation requires social interaction as the outcome variables, and BMI (or whether obese) in late teens as the main explanatory variable (for the latter two outcomes, education is also a main explanatory variable), and then they add these indirect effects to the direct effects. The authors find that late-teen obesity is indirectly associated with 3.5 percent lower hourly wages for both women and men, suggesting that the total effect of obesity is much larger than what has been found previously. Using a different method from that in Han et al. (2011), we also find an indirect effect of obesity or overweightness on wages. Specifically, white males and non-white females who were OOT tend to invest less in their education, and black females who were OOT tend to sort themselves into a blue-collar occupation.

III. Data and Summary Statistics

Similar to prior work, this paper also utilizes data from the NLSY79, which surveys a nationally representative sample of 12,686 young men and women who were 14-22 years old when they were first interviewed in 1979. These individuals were followed up annually through 1994 and are currently interviewed on a biennial basis. Respondents were asked to report their weight in 1981, 1982, 1985, 1986 and all the years after 1988 (except 1991). Height information was only collected in 1981, 1982 and 1985, however, by this time the youngest respondents should have reached their adult height.

This paper focuses on the time period between 1981 and 2008.[32] Following Cawley (2004), we correct for reporting errors in weight and height. We use the BMI classification by the U.S. National Institutes of Health to construct the following categories of weight: healthy weight ($18.5 \leq BMI < 25$), overweight ($25 \leq BMI < 30$) and obese ($BMI \geq 30$). When the weight and height information first became available in 1981, the respondents were 16 to 24 years old; in 1982, they were 17 to 25 years old; in 1985, they were 20 to 28 years old. That is, as of 1985, the youngest respondents have reached adulthood. Our definition of being OOT is being obese or overweight in either 1981 or 1982 and under age 19; and our definition of being HWT is being healthy weight in *both* 1981 and 1982 and under age 19. As a result, respondents who were older than 19 in 1982 are dropped from the analyses.

In this paper, wage is defined as the hourly rate of pay at the current or most recent job, in 2008 dollars. Hourly rate of pay is top-coded as 500 if it is greater than 500, and bottom-coded as 1 if it is smaller than 1.[33] In the summary statistics, the measure is used directly to calculate the means. However, in the regression analyses, the natural log of wage is used as the outcome variable.

The main explanatory variable is BMI or the clinical classification of weight, which includes being obese, being overweight and being healthy weight (with being healthy weight as the omitted category). The reason we use both measures is that BMI might not have a linear effect on wages, at least for some gender and race/ethnicity groups (Cawley, 2004).

---

[32] Note that after 1994, the data have been collected biennially, therefore there are 17 years in total for which we can calculate the BMI, and these years are 1981, 1982, 1985, 1986, 1988, 1989, 1990, 1992, 1993, 1994, 1996, 1998, 2000, 2002, 2004, 2006 and 2008.

[33] As a result, 49 observations are top-coded, and 294 observations are bottom-coded.

Education and human capital have important impacts on wages (Mincer, 1974), and hence should be included in the analyses. Specifically, the respondent's general intelligence, current enrollment status, educational attainment, work experience and job tenure are controlled for in the regressions. The respondent's general intelligence is measured by the Armed Forces Qualification Test (AFQT) Percentile Score in 1981. Educational attainment is measured by four dummy variables: less than high school, high school, some college, and college or more, with the first one being the omitted category. Work experience is defined as years of work experience, and job tenure is defined as weeks of job tenure at the current or most recent job. Two other work-related variables are also included: part-time work (less than 20 hours per week) and white-collar occupation.[34]

A recent line of literature points to the strong association between being obese or overweight and risky behaviors (Averett et al., 2010; Flegal et al., 2013). Given that risky behaviors also affect labor market outcomes (Kaestner, 1991; Mullahy and Sindelar,

[34] White-collar occupation, before 2000 (including 2000), is equal to 1 if the respondent's occupation is professional, technical and kindred; managers, officials and proprietors; sales workers; or clerical and kindred; and 0 otherwise. In 2002, white-collar occupation is equal to 1 if the respondent's occupation is management; business and financial operations; computer and mathematical; architecture and engineering; life, physical and social services; community and social services; legal; education, training and library; arts, design, entertainment, sports and media; healthcare practitioners and technical; healthcare support; sales and related; or office and administrative support; and 0 otherwise. After 2002, white-collar occupation is equal to 1 if the respondent's occupation is executive, administrative and managerial occupations; management related occupations; mathematical and computer scientists; engineers, architects, surveyors, engineering and related technicians; physical scientists; social scientists and related workers; life, physical and social science technicians; counselors, social and religious workers; lawyers, judges and legal support workers; teacher; education, training and library workers; entertainers and performers, sports and related workers; media and communications workers; health diagnosing and treating practitioners; health care technical and support occupations; entertainment attendants and related workers; sales and related workers; or office and administrative support workers; and 0 otherwise.

1991), we control for a group of risky behaviors available from the data, including regular drinking,[35] chronic marijuana use,[36] cocaine use[36] and illegal activity.[37]

Demographic variables are also controlled for in the regressions to take into account the background differences of the respondents, which include: race/ethnicity,[38] female, age, citizenship status in 1984, education attainments of both parents, marital status,[39] number of children in household, youngest child younger than 6, residing in an urban area, residing in a Standard Metropolitan Statistical Area (SMSA), and region of residence.[40]

Other regressors include health limiting kind of work,[41] health limiting amount of work,[42] and a linear time trend.

Several corrections are made to the data. Firstly, since pregnancy could increase a woman's weight significantly, we drop the observations for respondents who are

---

[35] This is a dummy variable indicating whether the respondent drank at least once a month. Questions about alcohol use are asked in 1982, 1983, 1984, 1985, 1988, 1989, 1992, 1994, 2002, 2006 and 2008. We impute alcohol use for those years in the gaps by using the information from the most recent previous year available. For example, alcohol use in 1986 or 1987 is the same as that in 1985. Since we use data from 1985 and onward in the regressions, we do not have to impute alcohol use before 1982.

[36] This is a dummy variable indicating whether the respondent used marijuana for at least 100 times in his or her lifetime (whether the respondent has ever used cocaine). Questions about marijuana (cocaine) use are asked in 1984, 1988, 1992, 1994 and 1998. We impute marijuana (cocaine) use for those years in the gaps and after 1998 by using the information from the most recent previous year available. Since we use data from 1985 and onward in the regressions, we do not have to impute marijuana (cocaine) use before 1984.

[37] This is a dummy variable indicating whether the respondent was ever charged with illegal activity before 1980. In the 1980 survey, the respondent is asked whether he or she was ever charged with illegal activity except for a minor traffic offense. We impute this variable for all the years after 1985 (including 1985) using the information from 1980.

[38] Race/ethnicity includes white, black or Hispanic, with the first one being the omitted category.

[39] Marital status includes single, married and other marital status, with the first one being the omitted category.

[40] Region of residence includes Northeast, North Central, South and West, with the first one being the omitted category.

[41] This variable is defined as whether health limits the kind of work the respondent could do.

[42] This variable is defined as whether health limits the amount of work the respondent could do.

pregnant during a particular survey year. Secondly, military personnel are also dropped from the analyses due to their special pay scale. Finally, we fill in the missing values with the sample means of the correspondent variables. In the regressions, we add indicator variables for whether the value is missing for all independent variables. The only exception is BMI (or the clinical classification of weight), the missing values of which are not imputed.

Summary statistics are provided in Table 3.1. The sample is split and presented for OOT and HWT. The table shows that, compared to HWT, once OOT reach adulthood they earn lower hourly wages; have higher BMI; are less likely to have engaged in risky behaviors; are less likely to be married; are more likely to have health limits; and are less likely to work in a white-collar occupation. At the same time, they also had lower AFQT scores, completed fewer years of schooling, and had less educated parents than their lighter counterparts.

IV. Methods and Econometric Models

Two econometric models are used in this paper: OLS and fixed effects.[43] In addition, we make use of the Blinder-Oaxaca decomposition to understand where the employment probability (wage) differential for individuals with different teenage weight statuses comes from, a method of which very few papers in this literature have taken advantage so far.[44]

---

[43] We do not use lagged weight measures due to the design of our methodology, which attempts to look for the heterogeneity of the obesity/overweightness wage penalty and hence has to stratify the sample by lagged weight measures. We also do not use IV because the sample is very stratified and certain subgroups have too little variation for the instruments commonly used in this literature to work.

[44] To the author's knowledge, the only paper that also uses Blinder-Oaxaca decomposition is Daouli et al. (2010). However, Daouli et al. (2010) use current weight status to separate individuals into obese and non-

The baseline OLS model is as follows:

$$P(emp)_{it} = \beta_0 + \beta_1 BMI_{it} + X_{it}\beta_2 + Z_i\beta_3 + \alpha_{it} + \varepsilon_{it}$$

$$\ln w_{it} = \beta_0 + \beta_1 BMI_{it} + X_{it}\beta_2 + Z_i\beta_3 + \alpha_{it} + \varepsilon_{it},$$

where $P(emp)_{it}$ is the probability of being employed in the week before the survey week for individual i at time t; $\ln w_{it}$ is the natural log of hourly wage at the current or most recent job for individual i at time t; BMI_{it} is the main independent variable; X_{it} is a vector of time variant control variables for individual i at time t; Z_i is a vector of time invariant control variables for individual i; α_{it} is the unobservable associated with both BMI and wages, such as being economically myopic; and ε_{it} is an idiosyncratic error term with mean zero. Specifically, BMI_{it} denotes the BMI for individual i at time t. The time variant control variables X_{it} include: current enrollment status, educational attainment, work experience and its square term, job tenure and its square term, part-time work, white-collar occupation, regular drinking, chronic marijuana use, cocaine use, age and age squared, marital status, number of children in household, youngest child younger than 6, residing in an urban area, residing in an SMSA, region of residence, health limiting kind of work, health limiting amount of work, and a linear time trend.[45] The time invariant control variables Z_i include AFQT score in 1981, illegal activity in 1980, race/ethnicity, female, citizenship status in 1984, maternal education,[46] and paternal education.[47] The OLS model is estimated for each subgroup, which is obtained through stratifying the full

obese groups. Although they also use lagged weight status as a robustness check, they are not holding current weight status constant.

[45] Job tenure at the current or most recent job and its square term, part-time work, and white-collar occupation are not included as controls when employment probability is the outcome variable; otherwise, it restricts the sample size quite a bit.

[46] This variable is defined as the highest grade completed by the respondent's mother.

[47] This variable is defined as the highest grade completed by the respondent's father.

sample by gender, teenage weight status, and race/ethnicity. β_1 is the coefficient of interest, which measures the return on BMI.

The estimates obtained through OLS are biased if α_{it} is not equal to zero. If we assume that α_{it} is constant over time, i.e., α_{it} is essentially α_i, an individual fixed effects model can be used to eliminate the unobservable and yield an unbiased estimate of β_1:

$$P(emp)_{it} - \overline{P(emp)_i} = \beta_1(BMI_{it} - \overline{BMI_i}) + \beta_2(X_{it} - \overline{X_i}) + (\varepsilon_{it} - \overline{\varepsilon_i})$$

$$\ln w_{it} - \overline{\ln w_i} = \beta_1(BMI_{it} - \overline{BMI_i}) + \beta_2(X_{it} - \overline{X_i}) + (\varepsilon_{it} - \overline{\varepsilon_i}),$$

where $\overline{P(emp)_i} = \frac{1}{T}\sum_{t=1}^{T} P(emp)_{it}$, $\overline{\ln w_i} = \frac{1}{T}\sum_{t=1}^{T} \ln w_{it}$, $\overline{BMI_i} = \frac{1}{T}\sum_{t=1}^{T} BMI_{it}$, $\overline{X_i} = \frac{1}{T}\sum_{t=1}^{T} X_{it}$,

and $\overline{\varepsilon_i} = \frac{1}{T}\sum_{t=1}^{T} \varepsilon_{it}$, which denote the cross-time averages of employment probability, natural log of wage, BMI, the other control variables and the error term, for each individual. α_i is differenced out in this process.

To test whether weight status has a non-linear effect on employment probability or wage, BMI is replaced with the clinical classification of weight (being obese, being overweight and being healthy weight, with being healthy weight as the omitted category) in a separate set of regressions, for both OLS and FE models.

To better understand the heterogeneity by teenage weight status, we carry out the Blinder-Oaxaca decomposition (Blinder, 1973; Oaxaca, 1973). To start with, we specify separate employment probability equations and wage equations for adults who were HWT (H) and those who were OOT (O).

$$P(emp)_{it}^{H} = X_{it}^{H}\beta_1^{H} + \varepsilon_{it}^{H} \;\&\; P(emp)_{it}^{O} = X_{it}^{O}\beta_1^{O} + \varepsilon_{it}^{O};$$

$$\ln w_{it}^{H} = X_{it}^{H}\beta_1^{H} + \varepsilon_{it}^{H} \ \& \ \ln w_{it}^{O} = X_{it}^{O}\beta_1^{O} + \varepsilon_{it}^{O}.$$ [48]

Taking the difference of either the employment probability equations or the wage equations, we can decompose the differential at the mean of each covariate into two parts:

$$\overline{P(emp)}_{it}^{H} - \overline{P(emp)}_{it}^{O} = (\overline{X}_{it}^{H} - \overline{X}_{it}^{O})\beta_1^{H} + \overline{X}_{it}^{O}(\beta_1^{H} - \beta_1^{O});$$

$$\overline{\ln w}_{it}^{H} - \overline{\ln w}_{it}^{O} = (\overline{X}_{it}^{H} - \overline{X}_{it}^{O})\beta_1^{H} + \overline{X}_{it}^{O}(\beta_1^{H} - \beta_1^{O}).$$

The first part of the differential comes from differences in the explanatory variables, which is referred to as the "endowment effect" or the "explained part." The second part of the differential comes from differences in the returns on the explanatory variables, which is referred to as the "unexplained part." In the first term, the differences of the average "endowments" are multiplied by the returns for the adults who were HWT. In the second term, the differences of the returns are multiplied by the average "endowments" of the adults who were OOT. The first term is a measure of additional employment probability or additional wages of the adults who were OOT, assuming that they had the "endowments" of the adults who were HWT. The second term is a measure of additional employment probability or additional wages of the adults who were OOT, assuming that they had the returns of the adults who were HWT.

V. Results

The hypotheses that the two groups (OOT and HWT) have equal coefficients on BMI, education, human capital [49] and occupation [50] are tested for each gender and

---

[48] For simplicity, the intercept, BMI and the time invariant covariates are subsumed in X_{it} along with the time variant covariates, and the unobservable is omitted.

race/ethnicity group.[51] Appendix Table 3.1 provides a summary of the test results for the explanatory variables of interest. Overall, the null hypothesis of equal coefficients is more likely to be rejected for males than for females. However, in the rest of the analyses, we still separate females by teenage weight status for the sake of completeness.

For each methodology employed, we start with presenting the results for which wage is the outcome and then move on to those for which employment probability is the outcome. The former are more interesting, and in addition, wage as an outcome has been more of a focus in the previous literature than employment probability.

The baseline OLS results are presented in Table 3.2. The results from the unstratistified sample, reported in column 2 and column 5, are qualitatively similar to those from the previous literature. Specifically, males are not affected and females are negatively affected by excessive weight. On the margin, a one unit increase in BMI is associated with a 0.2 and 0.5 percent decrease in hourly wage for white and Hispanic males, yet it is associated with a 0.1 percent *increase* in hourly wage for black males, which are very similar to the 0.1 and 0.7 percent decrease for white and Hispanic males and the 0.4 percent *increase* for black males reported in Cawley (2004). The major difference is that our results are not significant for males of any race/ethnicity, but

[49] Human capital includes work experience and job tenure. The latter is not included as a control when employment probability is the outcome variable, and hence the test of equal coefficients is not carried out for job tenure when employment probability is the outcome variable.

[50] White-collar occupation is not included as a control when employment probability is the outcome variable, and hence the test of equal coefficients is not carried out for white-collar occupation when employment probability is the outcome variable.

[51] To test whether the coefficients for the two subgroups are equal, we run regressions with the sample that includes observations from both subgroups, and include in the regressions the dummy variable for being OOT and the interaction terms between this dummy and all the other control variables (including BMI), besides all the original independent variables. If the coefficient on the interaction term between the dummy and BMI is significant, it means that for the two groups (OOT and HWT), BMI has significantly different coefficients; otherwise, it does not. The results are not shown here but available upon request.

Cawley (2004) actually finds significant results for black and Hispanic males. In addition, a one unit increase in BMI is significantly associated with a 0.4, 0.3 and 0.4 percent decrease in hourly wages for white, black and Hispanic females, which are also very similar to the 0.8, 0.4 and 0.6 percent reported in Cawley (2004). The discrepancies between our results and those from Cawley (2004) come from two sources. First, we eliminated observations who were not teenagers before 1982. Second, we eliminated the underweight teens. The eliminations render us not only a different but also a smaller sample, which explains why we find slightly different and fewer significant results.

Interestingly, the results from the stratified samples show that while BMI is negatively correlated with wages for males who were OOT, it is actually *positively* correlated with wages for males who were HWT (although the coefficient is statistically significant only for white and black males). On the margin, a one unit increase in BMI is associated with a 1.0 percent to 2.3 percent decrease in hourly wages for males who were OOT. A one unit increase in BMI is associated with a 0.6 percent to 0.8 percent *increase* in hourly wages for males who were HWT. However, BMI is negatively correlated with wages for females regardless of their teenage weight status. On the margin, a one unit increase in BMI is associated with a 0.4 percent to 1.0 percent decrease in hourly wages for females. The null hypothesis of equal coefficients on BMI is rejected for males of every race/ethnicity, but not for females of any race/ethnicity.

When employment probability is the outcome variable, fewer significant results appear for either men or women, and the pattern is less clear. If anything, however, BMI tends to be positively associated with the employment probability for men, while negatively associated with the employment probability for women. On the margin, the significant associations are the following: a one-unit increase in BMI is associated with a 0.4 percentage points *higher* employment probability for black males who were HWT; a one-unit increase in BMI is associated with a 0.6 percentage points *higher* employment probability for Hispanic males who were OOT; a one-unit increase in BMI is associated with a 0.3 percentage points lower employment probability for white females who were HWT; and a one-unit increase in BMI is associated with a 0.5 percentage points lower employment probability for Hispanic females who were HWT. The tests of equal coefficients again indicate that males of every race/ethnicity have statistically different coefficients on BMI by teenage weight status, but it is not the case for females of any race/ethnicity.

The estimates yielded by OLS are biased if there is unobserved heterogeneity or reverse causality. If we assume that the unobserved heterogeneity is time invariant and there is no reverse causality, then an FE model can be used to eliminate the bias. Table 3.3 shows the results from the individual FE model, with BMI as the main explanatory variable. No group has a significant coefficient on BMI when the outcome is natural log of wage except for Hispanic females, for whom a one unit increase in BMI decreases the hourly wages by 0.7 percent. Unsurprisingly, the null hypothesis for the test of equal coefficients on BMI is not rejected for males or females of any race/ethnicity. The negative coefficients on BMI are for black and Hispanic males who were OOT, as well as for white and Hispanic females regardless of their teenage weight status. Oddly, black females have positive coefficients on BMI regardless of their teenage weight status, although the coefficients are not significant.

When employment probability is the outcome, the coefficients on BMI are not significant for males of any race/ethnicity. Females (with the exception of blacks), however, seem to be punished more by current excessive weight if they were *HWT*, although the test of equal coefficients on BMI fails to reject the null hypothesis for females of any race/ethnicity. On the margin, a one-unit increase in BMI decreases the employment probability by 0.8 percentage points for white females, 1 percentage point for white females who were HWT, and 1.1 percentage points for Hispanic females who were HWT. These counterintuitive results for females might come from the different baselines that the two groups with different teenage weight statuses are compared with. Females who were HWT have a higher average employment rate than those who were OOT. Actually, this is part of the reason why we carry out the Blinder-Oaxaca decomposition, which shows us the overall employment probability (wage) differential between individuals with different teenage weight statuses.

We then replace BMI with the clinical classification of weight for both OLS and FE models. The results are presented in Table 3.4 and Table 3.5 respectively. The OLS results from the unstratified sample indicate that for white males, being overweight is significantly associated with a 5.4 percent increase in hourly wages, which is similar to the 3.9 percent reported in Cawley (2004). In contrast, for white, black and Hispanic females, being obese is associated with a 5.9, 6.6 and 5.3 percent decrease in hourly wages, which are also similar to the 11.9, 6.1 and 8.2 percent reported in Cawley (2004). The major difference is that our result for Hispanic females is not significant, but the counterpart result from Cawley (2004) is. Again, our restricted sample might explain why we find slightly different and fewer significant results.

The OLS results from the stratified samples by teenage weight status indicate that, for white males who were OOT, being obese (overweight) as adults is associated with a 20.4 percent (14.1 percent) decrease in hourly wages; for white males who were HWT, being overweight as adults is associated with an 8.6 percent *increase* in hourly wages.[52] For black and Hispanic males, none of the coefficients on weight measures is significant at conventional levels. We again utilize the tests of equal coefficients. The coefficients on being obese are significantly different across teenage weight statuses for white and black males, and the coefficients on being overweight are significantly different across teenage weight statuses for only white males. That is, the strongest evidence of heterogeneity by teenage weight status comes from white males. For females, there is a wage penalty for being obese or being overweight regardless of their teenage weight status, and the size of the penalty ranges from 0.9 percent to 14.0 percent. The null hypothesis for the test of equal coefficients on being obese or being overweight is not rejected for females of any race/ethnicity.

When employment probability is the outcome, again, the pattern is less clear. Similar to the results with BMI as the main explanatory variable, excessive weight tends to be positively associated with the employment probability for men, while negatively associated with the employment probability for women. Specifically, for black males (black males who were HWT), being obese is significantly associated with a 3.5 (5.2) percentage points higher employment probability. For females, the results are more mixed. For white females who were HWT, Hispanic females and Hispanic females who

---

[52] Admittedly, these coefficients are large. However, when we look at the cross-tabs of wages of white males by their teenage and adult weight status, we notice that for the very few respondents who lost weight from being obese or overweight to being healthy weight, their wages are much higher than those who still remained obese or overweight, and they are on par with the group with the highest wages (overweight white adult males who were HWT).

were HWT, being obese is associated with a 4.7, 4.3 and 7.4 percentage points lower employment probability; for white females who were OOT, being overweight is associated with a 6.7 percentage points lower employment probability; and for black females who were HWT, being overweight is associated with a 3.6 percentage points *higher* employment probability.

When we utilize the individual FE model with the clinical classification of weight as the main explanatory variables, again, many significant results disappear. For white males, the wage premium associated with being overweight still remains significant, and the magnitude is only slightly smaller than the counterpart from the OLS model (4.5 percent vs. 5.4 percent). The situation is similar for white males who were HWT (5.7 percent vs. 8.6 percent). For black and Hispanic males, the coefficients on being obese or being overweight are not significant in any case. For males of any race/ethnicity, the coefficients on being obese or being overweight are not statistically different by teenage weight status, with the exception of being overweight for white males. For females, the coefficients on being obese or being overweight are mostly negative, with the following significant results: for white females, white females who were OOT and Hispanic females who were HWT, being obese decreases hourly wages by 6.4, 14.6 and 9.7 percent. Again, for females of any race/ethnicity, the coefficients of interest are not statistically different by teenage weight status.

When employment probability is the outcome, none of the coefficients for males are significant. For white females, white females who were HWT, black females who were OOT and Hispanic females who were HWT, being obese decreases the employment probability by 5.7, 8.8, 13.6 and 10.2 percentage points.

Table 3.6 displays the results from the Blinder-Oaxaca decomposition for the wage equations. The most interesting results are that the effects of BMI are largely "unexplained" for males, while they are "explained" for females. That is, the returns on BMI are different by teenage weight status for males: if adult males who were OOT were to have the returns on BMI as those who were HWT, they would have significantly higher wages. In contrast, if adult females who were OOT were to have the BMI of those who were HWT, they would have significantly higher wages.

Moreover, there is some suggestive evidence that white males and non-white females who were OOT tend to invest less in their education. If they were to have the education level of their respective counterparts who were HWT, they would have significantly higher wages. Black females who were OOT also tend to sort themselves into a blue-collar occupation, which is lower paid. These two findings are consistent with Han et al. (2011). However, the improvement of the current method is that we are also able to find some "unexplained" parts that stem from the different *returns* on the same characteristics. For example, black males and black females who were OOT have a smaller return on being in a white-collar occupation; white females who were OOT have a smaller return on education.

Table 3.7 displays the results from the Blinder-Oaxaca decomposition for the employment probability equations. The most interesting result from this table is that white females who were OOT have significantly lower employment probability than their counterparts who were HWT, and this differential can be explained mostly by differences in their characteristics. Specifically, white females who were OOT have a higher BMI, lower educational attainment and less work experience than white females who were

HWT, although the *return* on education is also lower for them. Previous literature generally finds that white females suffer from the strongest wage penalty. However, this penalty might still be underestimated due to the omission of their lower labor force attachment. That is, being OOT might discourage some white females from investing in education and/or entering (staying in) the labor force. The total earnings loss for this group, therefore, might be underestimated in previous work.

VI. Discussion

The most interesting finding of this paper is that overweight white males who were HWT actually enjoy a wage premium. There are several hypotheses that are consistent with this finding. First, financially successful individuals, when maximizing their overall utility, might intentionally sacrifice some aspects of their wellbeing, which in this case is weight. Because their opportunity cost of exercising and keeping fit is high, it could be a rational decision on their part to spend more time working and less time exercising than other less time-constrained individuals. A study in support of this hypothesis is Meltzer and Jena (2010), who "identify an association between income and exercise intensity that is consistent with the hypothesis that people respond to increased time costs of exercise by increasing intensity."

Second, the wage premium we find for overweight white males who were HWT might be caused by a high percent of fat free mass instead of a high percent of body fat for this group, the former of which actually indicates good health. In contrast, overweight white males who were OOT or obese white males regardless of their teenage weight status might indeed have a high percent body fat, which tends to negatively impact health. Given that good health increases earnings (Smith, 1999), overweight white males who were HWT enjoy the highest earnings exactly because of their best health status among all white males.

Third, the fact that overweight white males who were HWT but not their counterparts who were OOT enjoy a wage premium might be explained by the lower self-esteem of the latter. Heckman et al. (2006) find that "[n]oncognitive ability (including self-esteem) affects the acquisition of skills, productivity in the market and a variety of behaviors" and "noncognitive skills (compared to cognitive skills) are about equally strong in many outcomes and are stronger for some outcomes." Therefore, even if overweight white males who were HWT and their counterparts who were OOT are no different in terms of cognitive skills, the former could earn higher wages because of their higher self-esteem.

As to why there is only a wage premium for overweight white males who were HWT but not for the other overweight gender and race/ethnicity groups who were HWT, an explanation could be differential societal and cultural views towards overweightness in different groups.

Finally, it is worth noting that a single hypothesis alone might not be enough to explain this finding; rather, a combination of some of the aforementioned hypotheses might come closer to the truth.

VII. Conclusions

We find evidence suggesting the existence of heterogeneity by teenage weight status in the effects of excessive weight on wages for males. Specifically, a higher BMI is associated with a wage penalty only if they were OOT, but it is actually associated with a wage premium if they were HWT. When we use the clinical classification of weight as the main independent variables, we find that for white males, being overweight is associated with a wage premium, but not being obese. This suggests that for white males at least, modest weight excess is beneficial, but not when it surpasses a certain threshold. For females, however, a higher BMI is associated with a wage penalty no matter they were OOT or HWT, and the magnitudes of the penalty are not statistically different by teenage weight status, suggesting that the previous convention to group them together does not mask any heterogeneity. The findings for males, however, are still interesting and deserve some attention. Previous literature has not documented a consistent association between weight and wages for men. The results from this paper point to a possible explanation: the opposite effects of the two groups (OOT and HWT) work to cancel each other out.

We also look at how employment probability and weight status are related. While literature in the US generally does not focus on employment probability due to lack of significant findings (Cawley 2000), that in Europe does and has actually identified statistically significant effects of excessive weight on employment probability (e.g. Morris 2007). Our findings suggest that, because employment probability is not taken into account, the magnitude of the wage penalty for white females might be underestimated.

Lastly, we employ the Blinder-Oaxaca decomposition to separate out the "explained" and "unexplained" parts of the employment probability (wage) differential. We find that except for black males, individuals who were OOT tend to have lower wages. For white males and non-white females, this differential is explained by their lower educational attainment. For females of any race/ethnicity who were OOT, the differential is also explained by their higher current BMI. Black females who were OOT tend to sort themselves into a blue-collar occupation, which is lower paid. In addition, white females who were OOT tend to have a lower employment probability, and this differential is explained by their higher current BMI, lower education, and fewer years of work experience.

Several future research directions might prove fruitful. First, a matching exercise might help determine whether the heterogeneity is robust. Specifically, we can match individuals with different teenage weight statuses on observed/unobserved dimensions, and test whether the wage differential still exists. Second, we can take into account the lower employment probability of obese/overweight white females in order to obtain a more accurate estimate of the wage penalty for this group. For example, we can make use of selection models, or impute wages for those who are unemployed in this group following related labor literature. Third, alternative weight measures could be used to check if the results still hold. Economists have come up with imputing methods for obtaining (possibly) more accurate weight measures, such as percent body fat and fat free mass. As mentioned above, the wage premium we find for overweight white males who were HWT might be caused by a high percent of fat free mass instead of a high percent of body fat for this group. Fourth, we can use the results from this paper to facilitate finding the source of the obesity or overweightness wage penalty. For example, the source of the wage penalty might be a form of statistical discrimination. That is, on average, OOT invest less in education and/or are less attached to the labor force, and therefore employers use excessive weight as an elimination standard so that they will not make bad decisions on average. Moreover, the source of the overweightness wage penalty is likely not lower productivity caused by excessive weight, at least for white males. The reason is that overweight white males who were HWT actually have the highest wage among all white males; if wage accurately reflects productivity, the extra weight this group has gained over the years did not make them less productive.

References

Averett, S., & Sanders K. (1996). The economic reality of the beauty myth. *Journal of Human Resources*, 31(2), 304-330.

Averett, S., Corman, H., & Reichman, N. (2010). Effects of overweight on risky sexual behavior of adolescent girls. NBER Working Paper 16172.

Baum, C. L., & William F. F. (2004). The wage effects of obesity: a longitudinal study. *Health Economics*, 13(9), 885-899.

Blinder, A. S. (1973). Wage discrimination: reduced form and structural variables. *Journal of Human Resources*, 8(4), 436-455.

Cawley, J. (2000). Body weight and women's labor market outcomes. NBER Working Paper 7841.

Cawley, J. (2004). The impact of obesity on wages. Journal *of Human Resources*, 39 (2), 451-474.

Centers for Medicare and Medicaid Services, Office of the Actuary, National Health Statistics Group (2012). *National Health Care Expenditures Data*. Retrieved from https://www.cms.gov/Research-Statistics-Data-and-Systems/Statistics-Trends-and-Reports/NationalHealthExpendData/downloads/tables.pdf

Conley, D. & Glauber, R. (2005). Gender, body mass and economic status. NBER Working Paper 11343.

Daouli, J., Davillas, A., Demoussis, M. & Giannakopoulos, N. (2010). Exploring gender specific wage differentials between obese and non-obese adults: Evidence from the NLSY. Unpublished manuscript.

Flegal, K.M., Kit, B.K., Orpana, H., Graubard, B.I. (2013). Association of all-cause mortality with overweight and obesity using standard body mass index categories: a systematic review and meta-analysis. *Journal of the American Medical Association*, 309(1), 71-82.

French S. A., Story M. & Perry C. L. (1995). Self-esteem and obesity in children and adolescents: a literature review. *Obesity Research*, 3(5), 479-490.

Gortmaker, S. L., Must, A., Perrin, J. M., Sobol, A. M., & Dietz, W. H., (1993). Social and economic consequences of overweight in adolescence and young adulthood. *The New England Journal of Medicine*, 329(14), 1008–1012.

Han, E., Norton, E. C., & Powell, L. M. (2011). Direct and indirect effects of body weight on adult wages. *Economics and Human Biology*, 9(4), 381-392.

Heckman, J., Stixrud, N. and Urzua, S. (2006). The effects of cognitive and noncognitive abilities on labor market outcomes and social behavior. *Journal of Labor Economics*, 24 (3), 411-482.

Kaestner, R. (1991). The effect of illicit drug use on the wages of young adults. *Journal of Labor Economics*, 8(4), 318-412.

Meltzer, D. O., & Jena, A. B. (2010). The economics of intense exercise. *Journal Of Health Economics*, *29*(3), 347-352.

Mincer, J. (1974). *Schooling, Experience and Earnings*. New York: Columbia University Press.

Mocan, N. H., & Tekin, E. (2009). Obesity, Self-esteem and Wages. NBER Working Paper 15101.

Morris, S. (2007). The impact of obesity on employment. *Labour Economics*, 14(3), 413-433.

Mullahy, J. & Sindelar, J. L. (1991). Gender differences in labor market effects of alcoholism. *American Economic Review*, 81(2), 161-165.

Oaxaca, R. (1973). Male-female wage differentials in urban labor markets. *International Economic Review*, 14(3), 693-709.

Pagan, J. A., & Davila, A. (1997). Obesity, occupational attainment, and earnings. *Social Science Quarterly*, 78(3), 756–770.

Reilly, J. J., Methven E., McDowell Z. C., Hacking B., Alexander D., Stewart L. & Kelnar C. J. H. (2003). Health consequences of obesity. *Archives of Disease in Childhood*, 88(9), 748–752.Smith, J. P. (1999). Healthy bodies and thick wallets: the dual relation between health and economic status. *Journal of Economic Perspectives*, *13*(2), 145-166.

Table 3.1 Summary Statistics

	OOT		HWT		
	Mean	N	Mean	N	P-value
Hourly wage for the current or most recent job ($2008)	16.57	8391	17.72	25255	0.00
Employed	0.86	6375	0.87	19225	0.03
BMI	31.88	8391	25.51	25255	0.00
AFQT score 1981	41.09	8244	44.77	24949	0.00
Currently enrolled in school	0.06	8391	0.07	25255	0.00
Highest grade completed by respondent	12.75	8386	13.14	25227	0.00
Work experience in years	11.51	8391	11.64	25255	0.14
Job tenure at current or most recent job	4.18	8274	3.98	25012	0.00
Part-time work	0.03	8016	0.04	24139	0.00
White-collar occupation	0.49	8313	0.53	25002	0.00
Regular drinking	0.61	8246	0.67	24808	0.00
Chronic marijuana use	0.27	8259	0.32	24917	0.00
Cocaine use	0.29	8287	0.34	24933	0.00
Illegal activity 1980	0.07	8349	0.08	24937	0.00
White	0.50	8391	0.55	25255	0.00
Black	0.28	8391	0.27	25255	0.05
Hispanic	0.21	8391	0.17	25255	0.00
Female	0.49	8391	0.44	25255	0.00
Age	31.30	8391	31.24	25255	0.52
Citizen 1984	0.96	8269	0.96	24962	0.64
Highest grade completed by mother	10.47	7893	11.07	23758	0.00
Highest grade completed by father	10.49	7315	11.22	21416	0.00
Single (never married)	0.37	8391	0.37	25254	0.64
Married	0.45	8391	0.47	25254	0.00
Other marital status	0.18	8391	0.16	25254	0.00
Number of children in household	1.05	8391	0.93	25255	0.00
Youngest kid's age < 6?	0.28	8369	0.27	25186	0.07
Urban	0.75	8151	0.78	24402	0.00
SMSA	0.58	8104	0.59	24291	0.42
Northeast	0.14	8346	0.16	25018	0.00
North Central	0.22	8346	0.25	25018	0.00
South	0.43	8346	0.39	25018	0.00
West	0.20	8346	0.20	25018	0.76
Health limiting kind of work	0.05	8264	0.03	24948	0.00
Health limiting amount of work	0.03	8262	0.02	24934	0.00

Notes:
P-values are from the two-tailed T-tests of equality for the two groups.

Table 3.2 OLS results with BMI as the main explanatory variable

	OOT & HWT	OOT	HWT	OOT & HWT	OOT	HWT
Outcome: log wage						
	Men			Women		
BMI	-0.002	-0.014***	0.007**	-0.004***	-0.006***	-0.006**
	(0.002)	(0.004)	(0.003)	(0.001)	(0.002)	(0.002)
N	19670	4304	14070	16271	4087	11185
R-squared	0.330	0.302	0.345	0.287	0.336	0.276
	White Men			White Women		
BMI	-0.002	-0.010**	0.006*	-0.004**	-0.006**	-0.008**
	(0.003)	(0.004)	(0.004)	(0.002)	(0.003)	(0.003)
N	10876	2451	7785	8409	1757	6175
R-squared	0.355	0.352	0.362	0.289	0.409	0.270
	Black Men			Black Women		
BMI	0.001	-0.011	0.008*	-0.003*	-0.004	-0.004
	(0.003)	(0.007)	(0.004)	(0.002)	(0.003)	(0.004)
N	5321	917	4025	4641	1469	2878
R-squared	0.291	0.341	0.337	0.311	0.335	0.312
	Hispanic Men			Hispanic Women		
BMI	-0.005	-0.023***	0.007	-0.004*	-0.005	-0.010**
	(0.004)	(0.006)	(0.008)	(0.003)	(0.004)	(0.005)
N	3473	936	2260	3221	861	2132
R-squared	0.278	0.353	0.299	0.298	0.321	0.324
Outcome: employment probability						
	Men			Women		
BMI	0.000	0.001	0.001	-0.001*	-0.001	-0.002
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
N	15094	3313	10797	12237	3062	8428
R-squared	0.170	0.179	0.178	0.141	0.147	0.146
	White Men			White Women		
BMI	-0.001	-0.000	0.000	-0.001	-0.002	-0.003*
	(0.001)	(0.002)	(0.001)	(0.001)	(0.002)	(0.001)
N	8417	1895	6020	6469	1360	4741
R-squared	0.173	0.200	0.172	0.140	0.176	0.138
	Black Men			Black Women		
BMI	0.002	0.002	0.004**	-0.000	0.000	0.001
	(0.001)	(0.003)	(0.002)	(0.001)	(0.002)	(0.003)
N	4020	700	3041	3380	1060	2108
R-squared	0.185	0.238	0.203	0.166	0.177	0.186
	Hispanic Men			Hispanic Women		
BMI	0.000	0.006*	-0.001	-0.002	-0.000	-0.005*
	(0.001)	(0.003)	(0.003)	(0.002)	(0.002)	(0.003)
N	2657	718	1736	2388	642	1579
R-squared	0.167	0.187	0.198	0.141	0.173	0.153

Notes:
Estimation method: OLS model.
Outcome variable: natural log of wage after 1985 or employment probability after 1985.
Main explanatory variable: BMI.
See text for the other explanatory variables used.
Robust standard errors clustered at individual level are displayed in parentheses. * 10%, ** 5%, *** 1%.

Table 3.3 FE results with BMI as the main explanatory variables

	OOT & HWT	OOT	HWT	OOT & HWT	OOT	HWT
Outcome: log wage						
	Men			Women		
BMI	0.000	-0.003	0.003	-0.002	-0.003	-0.000
	(0.002)	(0.004)	(0.003)	(0.002)	(0.002)	(0.003)
N	19670	4304	14070	16271	4087	11185
R-squared	0.588	0.559	0.597	0.530	0.578	0.513
	White Men			White Women		
BMI	0.005	0.003	0.005	-0.004	-0.007	-0.001
	(0.003)	(0.004)	(0.004)	(0.003)	(0.004)	(0.004)
N	10876	2451	7785	8409	1757	6175
R-squared	0.606	0.603	0.601	0.537	0.638	0.512
	Black Men			Black Women		
BMI	0.000	-0.002	0.003	0.002	0.001	0.004
	(0.004)	(0.008)	(0.005)	(0.003)	(0.003)	(0.006)
N	5321	917	4025	4641	1469	2878
R-squared	0.531	0.528	0.571	0.537	0.559	0.528
	Hispanic Men			Hispanic Women		
BMI	0.000	-0.009	0.004	-0.007*	-0.001	-0.009
	(0.005)	(0.006)	(0.007)	(0.004)	(0.005)	(0.005)
N	3473	936	2260	3221	861	2132
R-squared	0.573	0.553	0.589	0.500	0.521	0.513
Outcome: employment probability						
	Men			Women		
BMI	-0.001	0.002	-0.002	-0.003**	-0.000	-0.006***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
N	15094	3313	10797	12237	3062	8428
R-squared	0.334	0.348	0.335	0.318	0.331	0.313
	White Men			White Women		
BMI	-0.001	0.001	-0.001	-0.008***	-0.005	-0.010***
	(0.002)	(0.003)	(0.003)	(0.002)	(0.004)	(0.003)
N	8417	1895	6020	6469	1360	4741
R-squared	0.337	0.365	0.325	0.318	0.360	0.304
	Black Men			Black Women		
BMI	0.001	0.001	0.002	0.003	0.003	0.001
	(0.003)	(0.005)	(0.004)	(0.003)	(0.004)	(0.004)
N	4020	700	3041	3380	1060	2108
R-squared	0.344	0.413	0.349	0.341	0.362	0.353
	Hispanic Men			Hispanic Women		
BMI	-0.001	0.005	-0.009	-0.006	0.002	-0.011*
	(0.004)	(0.006)	(0.006)	(0.004)	(0.006)	(0.006)
N	2657	718	1736	2388	642	1579
R-squared	0.328	0.308	0.363	0.305	0.303	0.311

Notes:
Estimation method: individual fixed effects model.
Outcome variable: natural log of wage after 1985 or employment probability after 1985.
Main explanatory variable: BMI.
See text for the other explanatory variables used.
Robust standard errors clustered at individual level are displayed in parentheses. * 10%, ** 5%, *** 1%.

Table 3.4 OLS results with clinical classification of weight as the main explanatory variables

	OOT & HWT	OOT	HWT	OOT & HWT	OOT	HWT
Outcome: log wage						
	Men			Women		
Obese	-0.012	-0.197***	0.050*	-0.057***	-0.082**	-0.074**
	(0.021)	(0.046)	(0.029)	(0.020)	(0.037)	(0.029)
Overweight	0.047***	-0.118***	0.069***	-0.019	-0.011	-0.028
	(0.014)	(0.039)	(0.016)	(0.016)	(0.031)	(0.020)
N	19670	4304	14070	16271	4087	11185
R-squared	0.331	0.298	0.346	0.287	0.336	0.276
	White Men			White Women		
Obese	-0.026	-0.204***	0.021	-0.059*	-0.085*	-0.073*
	(0.028)	(0.053)	(0.036)	(0.030)	(0.046)	(0.044)
Overweight	0.054***	-0.141***	0.086***	-0.018	0.015	-0.036
	(0.018)	(0.044)	(0.021)	(0.024)	(0.044)	(0.028)
N	10876	2451	7785	8409	1757	6175
R-squared	0.358	0.354	0.365	0.289	0.410	0.270
	Black Men			Black Women		
Obese	0.018	-0.152	0.067	-0.066**	-0.140*	-0.086*
	(0.036)	(0.097)	(0.046)	(0.032)	(0.083)	(0.049)
Overweight	0.022	-0.088	0.032	-0.009	-0.115*	0.007
	(0.023)	(0.079)	(0.026)	(0.027)	(0.067)	(0.034)
N	5321	917	4025	4641	1469	2878
R-squared	0.291	0.339	0.336	0.312	0.337	0.314
	Hispanic Men			Hispanic Women		
Obese	-0.010	-0.032	0.083	-0.053	-0.058	-0.096*
	(0.054)	(0.132)	(0.080)	(0.039)	(0.063)	(0.056)
Overweight	0.050	0.094	0.041	-0.040	0.024	-0.079*
	(0.037)	(0.115)	(0.045)	(0.034)	(0.045)	(0.041)
N	3473	936	2260	3221	861	2132
R-squared	0.279	0.336	0.300	0.298	0.321	0.324
Outcome: employment probability						
	Men			Women		
Obese	0.010	0.002	0.031***	-0.026***	-0.044**	-0.046***
	(0.008)	(0.021)	(0.011)	(0.010)	(0.020)	(0.016)
Overweight	-0.006	-0.010	-0.002	0.004	-0.049**	0.018**
	(0.006)	(0.022)	(0.007)	(0.008)	(0.022)	(0.009)
N	15094	3313	10797	12237	3062	8428
R-squared	0.171	0.180	0.179	0.142	0.148	0.148
	White Men			White Women		
Obese	-0.002	-0.011	0.020	-0.020	-0.040	-0.047**
	(0.010)	(0.022)	(0.013)	(0.013)	(0.027)	(0.020)
Overweight	-0.006	-0.016	-0.002	-0.000	-0.067**	0.011
	(0.007)	(0.024)	(0.008)	(0.010)	(0.030)	(0.012)
N	8417	1895	6020	6469	1360	4741
R-squared	0.173	0.200	0.172	0.140	0.178	0.139
	Black Men			Black Women		
Obese	0.035**	0.070	0.052***	-0.019	-0.039	-0.028

	(0.017)	(0.053)	(0.020)	(0.018)	(0.043)	(0.031)
Overweight	-0.005	0.046	-0.004	0.009	-0.051	0.036**
	(0.014)	(0.050)	(0.015)	(0.016)	(0.049)	(0.018)
N	4020	700	3041	3380	1060	2108
R-squared	0.186	0.241	0.204	0.167	0.178	0.189
	Hispanic Men			Hispanic Women		
Obese	0.014	0.091	0.012	-0.043*	-0.053	-0.074**
	(0.019)	(0.061)	(0.034)	(0.023)	(0.046)	(0.037)
Overweight	-0.012	0.045	-0.003	-0.003	-0.054	0.014
	(0.017)	(0.061)	(0.019)	(0.018)	(0.049)	(0.019)
N	2657	718	1736	2388	642	1579
R-squared	0.168	0.187	0.198	0.141	0.174	0.155

Notes:
Estimation method: OLS model.
Outcome variable: natural log of wage after 1985 or employment probability after 1985.
Main explanatory variables: being obese and being overweight.
See text for the other explanatory variables used.
Robust standard errors clustered at individual level are displayed in parentheses. * 10%, ** 5%, *** 1%.

Table 3.5 FE results with clinical classification of weight as the main explanatory variables

	OOT & HWT	OOT	HWT	OOT & HWT	OOT	HWT
Outcome: log wage						
	Men			Women		
Obese	0.033	-0.016	0.039	-0.039*	-0.067	-0.033
	(0.021)	(0.045)	(0.027)	(0.023)	(0.043)	(0.029)
Overweight	0.021	-0.029	0.024	-0.009	-0.025	-0.008
	(0.014)	(0.038)	(0.015)	(0.017)	(0.037)	(0.020)
N	19670	4304	14070	16271	4087	11185
R-squared	0.588	0.559	0.597	0.530	0.578	0.514
	White Men			White Women		
Obese	0.044	-0.031	0.051	-0.064*	-0.146**	-0.018
	(0.031)	(0.059)	(0.040)	(0.038)	(0.063)	(0.048)
Overweight	0.045**	-0.037	0.057***	-0.014	-0.003	-0.014
	(0.019)	(0.049)	(0.022)	(0.027)	(0.052)	(0.030)
N	10876	2451	7785	8409	1757	6175
R-squared	0.606	0.603	0.602	0.537	0.642	0.512
	Black Men			Black Women		
Obese	0.039	-0.053	0.055	-0.000	-0.018	-0.036
	(0.038)	(0.090)	(0.044)	(0.036)	(0.080)	(0.047)
Overweight	0.002	-0.057	-0.004	0.015	-0.071	0.025
	(0.024)	(0.066)	(0.026)	(0.026)	(0.073)	(0.031)
N	5321	917	4025	4641	1469	2878
R-squared	0.531	0.529	0.571	0.537	0.560	0.529
	Hispanic Men			Hispanic Women		
Obese	0.033	0.049	0.020	-0.063	-0.063	-0.097*
	(0.048)	(0.126)	(0.061)	(0.044)	(0.087)	(0.053)
Overweight	-0.009	-0.010	-0.023	-0.024	-0.047	-0.048
	(0.031)	(0.122)	(0.036)	(0.034)	(0.074)	(0.040)
N	3473	936	2260	3221	861	2132
R-squared	0.574	0.553	0.589	0.500	0.521	0.513
Outcome: employment probability						
	Men			Women		
Obese	-0.005	0.016	-0.002	-0.049***	-0.047	-0.075***
	(0.014)	(0.032)	(0.018)	(0.018)	(0.039)	(0.023)
Overweight	-0.013	0.013	-0.016	-0.004	-0.049	0.011
	(0.009)	(0.029)	(0.011)	(0.012)	(0.035)	(0.013)
N	15094	3313	10797	12237	3062	8428
R-squared	0.335	0.348	0.335	0.319	0.331	0.315
	White Men			White Women		
Obese	-0.002	0.021	0.015	-0.057**	-0.020	-0.088***
	(0.018)	(0.041)	(0.024)	(0.025)	(0.051)	(0.031)
Overweight	-0.002	0.024	-0.006	-0.016	-0.041	-0.004
	(0.012)	(0.038)	(0.013)	(0.017)	(0.046)	(0.018)
N	8417	1895	6020	6469	1360	4741
R-squared	0.337	0.366	0.325	0.317	0.360	0.303
	Black Men			Black Women		
Obese	0.006	0.039	-0.005	-0.045	-0.136*	-0.058

	(0.030)	(0.069)	(0.035)	(0.033)	(0.072)	(0.043)
Overweight	-0.027	0.000	-0.029	0.007	-0.115	0.035
	(0.020)	(0.056)	(0.022)	(0.023)	(0.071)	(0.025)
N	4020	700	3041	3380	1060	2108
R-squared	0.345	0.414	0.349	0.342	0.365	0.357
	Hispanic Men			Hispanic Women		
Obese	-0.019	0.015	-0.045	-0.048	-0.003	-0.102*
	(0.034)	(0.087)	(0.042)	(0.043)	(0.088)	(0.052)
Overweight	-0.025	0.005	-0.023	-0.001	-0.014	0.006
	(0.024)	(0.083)	(0.028)	(0.025)	(0.067)	(0.026)
N	2657	718	1736	2388	642	1579
R-squared	0.329	0.307	0.362	0.305	0.303	0.313

Notes:
Estimation method: individual fixed effects model.
Outcome variable: natural log of wage after 1985 or employment probability after 1985.
Main explanatory variables: being obese and being overweight.
See text for the other explanatory variables used.
Robust standard errors clustered at individual level are displayed in parentheses. * 10%, ** 5%, *** 1%.

Table 3.6 Blinder-Oaxaca decomposition for the wage equations

	Men			Women		
	White men	Black men	Hispanic men	White women	Black women	Hispanic women
Overall						
HWT	2.830***	2.547***	2.724***	2.550***	2.419***	2.581***
	(0.017)	(0.023)	(0.034)	(0.018)	(0.024)	(0.028)
OOT	2.777***	2.578***	2.674***	2.505***	2.357***	2.491***
	(0.028)	(0.041)	(0.043)	(0.035)	(0.032)	(0.043)
Difference	0.053*	-0.031	0.050	0.045	0.062	0.089*
	(0.032)	(0.046)	(0.055)	(0.040)	(0.040)	(0.051)
Explained	0.039	-0.077*	0.081*	0.097***	0.091***	0.142***
	(0.026)	(0.042)	(0.048)	(0.030)	(0.035)	(0.045)
Unexplained	0.014	0.046	-0.031	-0.052	-0.029	-0.053
	(0.028)	(0.043)	(0.054)	(0.033)	(0.037)	(0.045)
Explained						
BMI	0.008	-0.014	0.043	0.044***	0.034*	0.046**
	(0.015)	(0.021)	(0.031)	(0.016)	(0.020)	(0.022)
Education	0.029**	-0.006	-0.004	0.009	0.019*	0.032**
	(0.012)	(0.011)	(0.013)	(0.008)	(0.010)	(0.013)
Human	-0.012	-0.022	0.018	0.007	0.005	0.029
Capital	(0.009)	(0.017)	(0.021)	(0.014)	(0.013)	(0.021)
Occupation	0.003	-0.003	-0.000	0.009	0.009**	0.008
	(0.002)	(0.005)	(0.001)	(0.006)	(0.005)	(0.006)
Unexplained						
BMI	0.458***	0.580**	0.868***	-0.036	-0.007	-0.126
	(0.159)	(0.242)	(0.289)	(0.124)	(0.157)	(0.186)
Education	-0.043	-0.115**	0.036	0.145***	-0.012	0.019
	(0.068)	(0.049)	(0.047)	(0.052)	(0.045)	(0.057)
Human	-0.043	-0.054	0.047	0.281*	0.015	-0.351**
Capital	(0.133)	(0.187)	(0.198)	(0.148)	(0.123)	(0.142)
Occupation	0.024	0.147***	-0.044	0.028	0.095*	0.078
	(0.039)	(0.052)	(0.069)	(0.040)	(0.054)	(0.054)
N	10236	4942	3196	7932	4347	2993

Notes:
The table displays the wage differential between individuals who were HWT and those who were OOT.
The explained part shows how much more log wage individuals who were OOT would earn if they were to have the same characteristics as those who were HWT.
The unexplained part shows how much more log wage individuals who were OOT would earn if they were to have the same returns on the characteristics as those who were HWT.
Education includes three dummies: high school, some college and college or more. Human capital includes work experience and job tenure. Occupation means being in a white-collar occupation.

Table 3.7 Blinder-Oaxaca decomposition for the employment probability equations

	Men			Women		
	White men	Black men	Hispanic men	White women	Black women	Hispanic women
Overall						
HWT	0.901***	0.846***	0.883***	0.878***	0.831***	0.857***
	(0.005)	(0.009)	(0.012)	(0.006)	(0.011)	(0.011)
OOT	0.900***	0.866***	0.884***	0.856***	0.813***	0.820***
	(0.009)	(0.018)	(0.016)	(0.012)	(0.016)	(0.018)
Difference	0.001	-0.020	-0.001	0.022	0.018	0.037*
	(0.010)	(0.020)	(0.020)	(0.013)	(0.019)	(0.021)
Explained	-0.010	-0.019	0.008	0.024**	0.003	0.030
	(0.007)	(0.014)	(0.018)	(0.010)	(0.016)	(0.018)
Unexplained	0.011	-0.001	-0.009	-0.002	0.015	0.007
	(0.009)	(0.019)	(0.022)	(0.014)	(0.021)	(0.021)
Explained						
BMI	-0.000	-0.014*	-0.004	0.013*	-0.006	0.015
	(0.005)	(0.008)	(0.012)	(0.007)	(0.011)	(0.013)
Education	0.003**	-0.000	0.001	0.003*	0.004	-0.001
	(0.001)	(0.003)	(0.005)	(0.002)	(0.003)	(0.003)
Human	-0.008*	-0.007	0.012	0.011*	0.012	0.021**
Capital	(0.005)	(0.009)	(0.009)	(0.006)	(0.009)	(0.010)
Unexplained						
BMI	0.007	0.063	-0.221*	-0.018	0.035	-0.136
	(0.054)	(0.103)	(0.132)	(0.058)	(0.088)	(0.102)
Education	0.039	0.074*	-0.029**	0.042**	0.096***	-0.001
	(0.028)	(0.043)	(0.014)	(0.020)	(0.020)	(0.020)
Human	-0.029	0.027	-0.033	0.008	-0.060	-0.034
Capital	(0.048)	(0.085)	(0.076)	(0.058)	(0.053)	(0.069)
N	7915	3741	2454	6101	3168	2221

Notes:

The table displays the employment probability differential between individuals who were HWT and those who were OOT.

The explained part shows how much more likely individuals who were OOT would get employment if they were to have the same characteristics as those who were HWT.

The unexplained part shows how much more likely individuals who were OOT would get employment if they were to have the same returns on the characteristics as those who were HWT.

Education includes three dummies: high school, some college and college or more. Human capital includes work experience.

Appendix

Appendix Table 3.1 Summary of the results for the tests of equal coefficients by teenage weight status

	White men	Black men	Hispanic men	White women	Black women	Hispanic women
OLS with BMI (wage/employment)						
BMI	Yes/Y	Yes/Y	Yes/Y			
Education		Yes/Y			Yes/Y	Yes
Human Captial		Yes		Yes		
White-collar			Yes			
FE with BMI (wage/employment)						
BMI						
Education	Yes		Yes(+)/Y(+)			
Human Captial		Yes			Yes	
White-collar	Yes(+)					
OLS with clinical classification of weight (wage/employment)						
Obese	Yes/Y	Yes/Y				
Overweight	Yes/Y	Y				Yes(+)/Y(+)
Education		Yes/Y			Yes/Y	Yes
Human Captial		Yes		Yes		
White-collar			Yes			
FE with clinical classification of weight (wage/employment)						
Obese						
Overweight	Yes/Y					
Education	Yes		Yes(+)/Y(+)		Yes	
Human Captial		Yes			Yes	
White-collar	Yes(+)					

Notes:

Yes means that the null hypothesis of equal coefficients is rejected when natural log of wage is the outcome variable.

Y means that the null hypothesis of equal coefficients is rejected when employment probability is the outcome variable.

(+) means the coefficient is more positive or less negative for OOT.

FE stands for (individual) fixed effects model.

Education includes three dummies: high school, some college and college or more. If one can reject the null hypothesis for more than one (including one) of them, it is marked **Yes (Y).**

Human capital includes work experience and job tenure, the latter is not a control when employment probability is the outcome variable. If one can reject the null hypothesis for either work experience or job tenure, it is marked **Yes(Y).**