Marriage, Depression, and Mortality Across the Life-Span

Christopher Ryan Beam
Los Angeles, California

B.A., Seattle Pacific University, 2002
M.A., New York University, 2008
M.A., University of Virginia, 2011

A Dissertation presented to the Graduate Faculty
of the University of Virginia in Candidacy for the Degree of
Doctor of Philosophy

Department of Psychology

University of Virginia
August, 2015

_____
Robert E. Emery, Ph.D., Chair

_____
Eric N. Turkheimer, Ph.D.

_____
Steven M. Boker, Ph.D.

_____
Allison J. Pugh, Ph.D.

Abstract

This dissertation addresses the causal influence of marital formation and marital dissolution on depression and risk of early mortality, including suicide. Although many studies have shown that marriage is correlated with fewer physical and mental health problems, the causal role of marriage in producing these favorable outcomes is uncertain. Spouses provide support, which may be one reason why marriage has been found to be beneficial. However, healthier people may select into marriage more frequently than unhealthy people, minimizing the causal role of marriage. The general research questions addressed include: (1) Does marrying decrease the severity of adults' depression over the lifespan? (2) Does becoming widowed increase the severity of adults' depression in late adulthood? (3) Does divorce increase the risk of early mortality, especially by committing suicide? The overall hypothesis tested in this dissertation is the following: Marriage is a protective factor against depression and mortality across the marital lifespan.

This dissertation extends prior work on the benefits of marriage by using several longitudinal twin samples to parse genetic and environmental selection from the potential causal influences of marriage. By adjusting for genetic and environmental selection factors that make twins similar to one another, twin designs provide the closest approximation to random assignment to within-family environmental "conditions" in nonexperimental research.

Three longitudinal twin studies were conducted to test the protective role of marriage over the lifespan. The first study addressed whether entering marriage

predicts lower severity in subsequent measures of depression over the lifespan, using data from the National Longitudinal Study of Adolescent Health (Add Health) twin sample and data from the Swedish Adoption/Twin Study on Aging (SATSA) sample. The second study addressed whether becoming widowed predicts increases in subsequent depression during the second half of the lifespan, using five waves of data from the SATSA sample. Additionally, twins who were discordant for becoming widowed were used to test whether depression mediated the association between widowhood and risk of early mortality. The third study addressed whether divorce raises the risk of early mortality, particularly by suicide, using twins' birth records, death records, and marital histories from two cohorts in the Swedish Twin Registry (STR).

Table of Contents

Dedication

There are many people who have supported my academic pursuits, and without them this dissertation would not have been possible. I am indebted to Drs. Eric Turkheimer and Robert Emery for modeling how to challenge ideas and beliefs in a scientific manner. I am grateful to both of them for helping me develop my conceptual, analytical, and organizational skills to present my arguments sensibly to others. Eric and Bob have also provided professional and personal support to me beyond anything I imagined before arriving at the University of Virginia.

I am also indebted to Drs. John Nesselroade and Steve Boker for supporting me with their NIA T32 training grant and in the International Max Planck Research School on the Life Course (LIFE). They have imparted knowledge to me about how to approach simple developmental questions with advanced statistical techniques in a thoughtful (and kind) way. Their support throughout my doctoral training at UVA and in LIFE has provided me with many life-changing opportunities, importantly the opportunity to meet and get to know my wife, Bettina Sonnenberg. Because of them, I also gained an important mentor and adviser at Freie Universität-Berlin, Dr. Michael Eid.

I am grateful to Dr. Patrick Shrout at New York University, who invited me to join his Couples Lab, taught me the initial statistical tools to study the interrelations among psychological phenomena, and modeled a passion for academic work that has guided me since our first meeting. Pat patiently answered my questions and helped mold my interest in relationship functioning and the development of depression. I am also grateful to Trudy Steinfeld at the NYU Wasserman Center for Career Development for her

willingness to take me in as one of her own and encouragement to “find places where you are valued.”

Dr. Kurt Fredrickson planted the first seeds of philosophical and existential inquiry in me during weekly meetings where we would discuss theological missives and commentaries. I am indebted to Dr. Kenneth Einar Himma, my philosophy mentor at Seattle Pacific University, for encouraging me to think critically and sensitively about philosophical, psychological, and interpersonal issues. Ken’s commitment to my development is immeasurable, beginning with his lesson that communicating one good, clear thought is an iterative process. I also acknowledge psychology faculty at SPU. During and since my time at SPU, I have entrusted all thoughts and motivations (the good and the bad) to Jon Wheeler. Thank you.

I have been fortunate to belong to many supportive clinical communities in New York and Virginia. By far, the clinical area at UVA has provided me with a deep well of clinical knowledge to help others and cope with the stress of graduate school. I am grateful to Dr. Lee Llewellyn, the emotional ballast for many students in the clinical area. I am also grateful to my supervisors, Drs. Joe Allen, Lee Ann Bass, Matt Zimmerman, and Andy Thomson for inspiring me to blend research with clinical work. Pat Shrout, Barry Rosenfeld, Ph.D., and Sherif Abdelmessih also were instrumental in my decision to pursue the study and practice of clinical psychology.

I am most grateful to my family, as they have been my greatest resource. My parents, brother, and grandparents have provided me with gold standard examples of high quality, happy marriages. They have supported every step of my academic

pursuits, from furnishing environments that promoted good study skills to the means to keep studying over the years. I am grateful to my father, Ronald Beam, for modeling that academic success requires a solid work ethic and follow-through, in addition to lively discussions. Thank you for asking, “What do you do?” I am most grateful to my mother, Gerilyn Beam, for being a bedrock, taking the time to listen and talk through my concerns, and teaching me the most basic interpersonal skill needed to be a clinician: “It’s not *what* you say, it’s *how* you say it.”

While the Max Planck LIFE Academy helped me to develop the ideas and skills needed to complete this dissertation, the microcosm of LIFE that Bettina Sonnenberg and I have created on a daily basis has provided me with the greatest source of support and intellectual inspiration. Her acumen for conceptual, design, and data analysis issues is something I only dream of attaining. Her intellect is matched only by her sense of humor for scientifically weak psychological studies.

I dedicate this dissertation to my grandparents, Salvatore (“Sam”) Pietro DiBlasi, Jr. and Rose Marie DiBlasi (*née* Marchese; 1930-2011), who taught (and continue to teach) me that adoration between spouses is what matters most for both spouses’ well-being.

Acknowledgments

This research uses data from Add Health, a program project directed by Kathleen Mullan Harris and designed by J. Richard Udry, Peter S. Bearman, and Kathleen Mullan Harris at the University of North Carolina at Chapel Hill, and funded by grant P01-HD31921 from the Eunice Kennedy Shriver National Institute of Child Health and Human Development, with cooperative funding from 23 other federal agencies and foundations. Special acknowledgment is due Ronald R. Rindfuss and Barbara Entwisle for assistance in the original design. Information on how to obtain the Add Health data files is available on the Add Health website (http://www.cpc.unc.edu/addhealth). No direct support was received from grant P01-HD31921 for this analysis.

This research also uses data from the Swedish Adoption/Twin Study on Aging, a project designed by Nancy Pedersen and funded by the National Institute on Aging (AG04563, AG10175), the MacArthur Foundation Research Network on Successful Aging, the Swedish Council for Working Life and Social Research (FAS) (97:0147:1B, 2009-0795), and the Swedish Research Council (825-2007-7460, 825-2009-6141).

This dissertation work was supported by grants from the National Institute on Aging to Christopher Beam, M.A. (1F31AG044047-01A1) and John Nesselroade, Ph.D. (T32AG020500). Additional grant support was provided by the National Institute of Child Health and Human Development to Robert E. Emery (1R01HD056354-01, 2RO1-HD056354) and Eric Turkheimer (5R01HD053550-02). The International Max Planck Research School on the Life Course and Michael Eid's Lab at Freie Universität-Berlin also provided support for the development of this dissertation work.

General Introduction

In 2013, the U.S. Supreme Court overturned the Defense of Marriage Act, the federal prohibition of same-sex marital unions. Within one year of this landmark decision, 10 individual state bans on same-sex were lifted, with pending court dates in more states. Part of the motivation to extend marital rights to all adults consists of the belief that prohibiting relationships from evolving into marriage denies "a status of immense import," which confers rights, privileges, and benefits that bolster physical and mental health to spouses and children (*United States v. Windsor*, 2013). Sociological and psychological research confirms this belief. Stably married people and their offspring are better adjusted than those in discordant marriages and divorced families (Emery, 1999; Waite & Gallagher, 2000), in terms of better physical health (Kiecolt-Glaser & Newton, 2001; K. Williams & Umberson, 2004), psychological health (Gove, Hughes, & Style, 1983; Johnson & Wu, 2002; Leonard & Mudar, 2003), longevity (Sbarra, Law, & Portley, 2011; Sbarra & Nietert, 2009), and economic benefits (Waite & Gallagher, 2000). While marriage might be correlated with better health, delineating whether marriage causally influences good physical and mental health outcomes is a more challenging matter.

Theoretical support for the positive effects of marriage

Despite the correlational nature of marital research, several theories have been proposed to explain the underlying causal mechanisms. The *main effect model* proposes that the mere connection with another person, like eating dinner at the kitchen table with one's spouse rather than over the kitchen sink. The connection promotes

improvements in health by, for example, enhancing self-worth associated with being with another person (Cohen, Gottlieb, & Underwood, 2000). Alternatively, the *buffering model of social support* proposes that marriage provides access to social networks, which dampen people's depressive reactions to stress (Aneshensel & Stone, 1982). Having a close, confiding spouse can soften the depressogenic effects of actual and perceived stress (Booth & Amato, 1991; Kessler & Essex, 1982; Myers, Lindenthal, & Pepper, 1975), through overt or covert support mechanisms (Bolger, Zuckerman, & Kessler, 2000; Shrout, Herman, & Bolger, 2006). Finally, *social pain models* suggest that the presence of a partner or spouse may prevent the neural activation of pain regulatory systems associated with social exclusion and abandonment (Panksepp, 2005). Single, divorced, and widowed adults lack partner support hypothesized to inhibit the activation of these networks.

Correlational nature of marital research

All marital studies rely on nonexperimental (or correlational) study designs. There is an inherent uphill methodological battle in distinguishing causal processes from selective ones to test the *social causation hypothesis* – the belief that marriage itself protects people against physical, psychological and mortality risks (Mastekaasa, 2006). In contrast, the *social selection hypothesis* posits that some people are more likely to select into marriage (and remain married) at a higher rate than others. Happier, healthier, prettier, and wealthier people probably do make more appealing mates and select into marriage at a higher rate than the homely and infirm. Selection can occur for a number of reasons, which are described in the next section. The goal in the present

research, however, is not to explain away the potential causal benefits of marriage with selection. The objective is to test the plausibility of the social causation hypothesis by appropriately controlling for selection factors that may partially account for why married people experience better health outcomes – namely, fewer depressive symptoms (Booth & Amato, 1991; Gove et al., 1983) and longevity (Lillard & Waite, 1995) – than their unmarried counterparts.

Using twins to parse social and genetic selection from causation

Social selection is a broad term that may result from *observed* processes, like measured educational attainment (Liu & Reczek, 2012), childhood problems (Gove et al., 1983), personality measures (Marks, 1996), attitudes toward marriage (Horwitz & White, 1998), and pre-marital levels of depression (Aseltine & Kessler, 1993; Lamb, Lee, & DeMaris, 2003; Menaghan & Lieberman, 1986; Simon & Marcussen, 1999) or *unobserved* processes, like unmeasured personality characteristics (Verbeek, 1990), unshakable socioeconomic status conditions (Verbeek & Nijman, 1992), and unmeasured familial background factors such as genes and family environments (Burt et al., 2010; Horn, Xu, Beam, Turkheimer, & Emery, 2013; Prescott & Kendler, 2001). Twin studies are powerful study designs for the reason that the genetic and environmental factors that make twins raised in the same family more similar to one another can be easily parsed from the within-family environmental factors that make them different from one another (Turkheimer & Harden, 2013). In this way, twin designs provide one of the most rigorous tests of the causal features of marriage.

Other methodological considerations

Besides using twin designs to parse selection processes from causal mechanisms in the associations among marriage, depression, and mortality risk (Rutter, Pickles, Murray, & Eaves, 2001), there are several methodological issues worth taking into consideration to put marital research in a lifespan developmental context (Baltes, 1987). These include: 1) clearly explicating the mediating factors between predictors and outcomes; 2) using quasi-experimental research designs to adjust for mediating effects of selection factors; 3) reliability and measurement considerations; and 4) embedding research in a lifespan developmental context.

Clear explication of the range of possible mediators, particularly selection, is important for all scientific inquiry, but is especially critical in marital research where random assignment of participants is unavailable and unethical. Selection effects often are found in studies on marriage and depression (Blekesaune, 2008; Booth & Amato, 1991; Horn et al., 2013; Mastekaasa, 1992). Even when selection processes have not been found to be statistically significant mediators (Horwitz & White, 1998; Mastekaasa, 1994b; Menaghan & Lieberman, 1986), taking into account nonsignificant selection processes still helps to strengthen causal arguments that marriage promotes better health outcomes. People's propensity to increase or decrease in depression as they age also may mediate the correlation between marriage and depression (Blazer, 2003). Modern statistical modeling techniques for longitudinal data, like latent growth curve models, are available to take into account this mediating factor.

Quasi-experimental research designs can disaggregate nonrandom processes from causal ones (Shadish, Cook, & Campbell, 2002). Twin studies function as a type of natural "experiment", because twins are matched for their familial background (Turkheimer & Harden, 2013). As an example, take two identical twins raised in the same household, where one is married and the other is single. The twins are genetically matched and share many of the same environments (e.g., shared the same prenatal environment, were raised by the same parents, and reached the same developmental milestones around the same time). If the married twin was happier than the single twin, the difference could not be due to genetic or shared environmental differences, as these factors are shared between identical twins. Their differences in happiness can be considered to result from "within-family random assignment" to different marital conditions. The advantage of using twins is that once genetic and environmental confounds are known, stronger conclusions can be made regarding the causal effects of marriage on depression (Beam et al., 2011; Heath, Eaves, & Martin, 1998; Horn et al., 2013; Lichtenstein, Gatz, Pedersen, Berg, & McClearn, 1996; Nes, Røysamb, Harris, Czajkowski, & Tambs, 2010; Osler, McGue, Lund, & Christensen, 2008) and mortality (Lichtenstein, Gatz, & Berg, 1998; Marenberg, Risch, Berkman, Floderus, & De Faire, 1994; Sbarra et al., 2011).

Whereas mortality can be reliably assessed with a single indicator, reliability and measurement considerations must be made when studying depression (Frech & Williams, 2007; Horn et al., 2013; Mastekaasa, 1995; Musick & Bumpass, 2012). Depression is a psychological disorder (or trait) that cannot be directly measured or

quantified. Using reliable assessments, like the Center for Epidemiologic Studies Depression Scale (CES-D;Radloff, 1977; Shafer, 2006) or latent variable models (Blekesaune, 2008; Menaghan & Lieberman, 1986) for measuring depression across multiple time points obviates problems associated with unreliability (Shrout, 2002).

There are many ways to embed marital research in a *lifespan developmental context*. In this dissertation, gender differences, sampling characteristics, plasticity and multidirectionality of developmental processes, and historical differences were considered for how marital status may differentially influence depression and mortality risk.

Gender differences in marital reactions traditionally have been considered since Bernard (1982) observed discrepancies between husbands' and wives' health and proclaimed that marriage was "good" for men's health but "bad" for women's health. When observed, marriage has been found to lower men's level of depression (Horwitz, White, & Howell-White, 1996; Marks, 1996; D. R. Williams, Takeuchi, & Adair, 1992) and increase their longevity (Choi & Marks, 2011; Kaplan & Kronick, 2006; Sbarra & Nietert, 2009; Van Poppel & Joung, 2001), whereas the same typically has not been found for women. Gender differences, however, have not always been observed (Blekesaune, 2008; Cheung, 2000; Clayton, Halikas, & Maurice, 1972; Frech & Williams, 2007; Gove et al., 1983; Manzoli, Villari, Pirone, & Boccia, 2007; Mastekaasa, 1992), which emphasizes a need for continued assessment in each sample.

Sampling characteristics often are overlooked in marital research. Age-heterogeneous samples frequently are pooled in the interest of increasing the overall

sample size for the sake of greater power (Karney & Bradbury, 1995). The trade-off, however, is that pooled samples compromise the ability to generalize findings to specific populations of interest. As a result developmental processes that change over the lifespan are ignored. The longitudinal effects of divorce, for example, have been found to depend on gender and age (Blekesaune, 2008). Collapsing across age (Lorenz, Wickrama, Conger, & Elder, 2006; Wade & Cairney, 2000) and gender (Musick & Bumpass, 2012; Wade & Pevalin, 2004) may miss these important features of development.

Plasticity and multidirectionality of developmental processes refers to people's adaptability and capability of change over the course of development (plasticity) and diversity of developmental trajectories (multidirectionality) both between and within groups (Baltes, 1987). The effects of marital transitions, like marrying and becoming widowed, may be time limited, lasting for short periods (Lichtenstein et al., 1996; Strohschein, McDonough, Monette, & Shao, 2005) or long periods of time (Horwitz et al., 1996; Lucas, Clark, Georgellis, & Diener, 2003).

The historical context in which people live also can influence their psychological development and associated outcomes (Baltes, 1987). Partly on the basis of historical events, like the Great Depression and World Wars I and II, the meaning of marriage changed in two important ways during the 20th century: first, from marriage as an institution to marriage as companionship, and second, from companionate marriages to individualized marriages (Cherlin, 2004). The social influence of marital transitions on depression and mortality risk may depend on the zeitgeist. As a result, cohort effects

were considered as potential moderators of the effects of marriage on depression and early mortality risk.

Longitudinal Associations between Marriage, Depression, and Mortality

The correlation between marriage and depression is well studied (Coombs, 1991), as is the correlation between marital status and longevity (Manzoli et al., 2007). Marrying tends to be associated with fewer symptoms of depression than never-marrying, cohabiting, divorcing, remarrying, separating, and becoming widowed (Jose, Daniel O'Leary, & Moyer, 2010; Kessler & Essex, 1982; Menaghan, 1985; Pearlin & Johnson, 1977; Ross, 1995; Weissman, 1987). Entering marriage has the potential to "right the ship" for people on otherwise depressive trajectories over short (Mastekaasa, 2006) and long intervals of time (Frech & Williams, 2007; Kim & McKenry, 2002; Marks & Lambert, 1998), but these effects may be time-limited (Musick & Bumpass, 2012). Widowhood, conversely, can increase the severity of depressive symptoms (Lichtenstein et al., 1996; Osler et al., 2008), but the negative effects of bereavement also appear to dissipate with the passage of time (Bonanno, 2004; Lucas et al., 2003). Finally, divorce has been found to increase the risk of all-cause early death (Sbarra et al., 2011), with elevations in suicide among younger divorced men (Yip et al., 2012).

Dissertation study aims

The studies in this dissertation do not address new questions about the effects of marriage on depression and longevity, but rather use longitudinal twin designs to control for confounds left unaddressed in prior research. The general question studied was: Does having a close, confiding spouse protect adults from depression and the

potentially mortal consequences of social isolation? *Study 1* addressed whether entering marriage decreases the subsequent severity of depression over short and long intervals of time across the lifespan. *Study 2* extended prior work on the effects of becoming widowed on depression in late adulthood (Lichtenstein et al., 1996) as well as tested whether severity of depression mediated the association between widowhood and risk of early mortality (Lichtenstein et al., 1998). *Study 3* expanded work on the association between divorce and early death by using a twin sample to test the causal influence of divorce on mortality risk. Suicide, moreover, was explored as a possible mechanism for the increased risk in mortality among divorced adults.

General Description of the Methodology

Longitudinal twin studies were used for two purposes in this dissertation: (1) to test the social causation hypothesis while adjusting for genetic and environmental selection and (2) to test the effects of marriage in real time.

Two twin methods were used to adjust for selection effects: multivariate longitudinal twin analyses (referred to as “longitudinal quasi-causal models”) and co-twin-control analyses. Longitudinal quasi-causal models were conducted to test whether marrying and becoming widowed influences short-term and long-term changes in people’s CES-D scores over time, statistically adjusting for the genetic and environmental selection effects that may partially (or fully) mediate the relation between marital status and depression. Significant within-family environmental effects are regarded as “quasi-causal,” as the effects are based on correlational data, but do help

to strengthen the argument that marital changes socially influence depression outcomes.

Co-twin-control analyses were used to test for significant differences in depression scores and mortality rates between twins who became widowed, for example, and their co-twins who did not become widowed (Lichtenstein et al., 1996; Robins, Greenland, & Breslow, 1986; Silman, Newman, & MacGregor, 1996). Paired *t*-tests were then performed to formally test whether twins who experienced marital change had lower depression (if they married), higher depression (if they become widowed), and higher risks of mortality and suicide (if they divorced) than their co-twins who experienced none of these changes.

In addition to twin methodologies, latent growth curve (LGC) models were used to take into account intraindividual growth in depressive symptoms over time (Duncan, Duncan, & Strycker, 2006; Willett & Sayer, 1994) while testing whether marital change – specifically, marrying or becoming widowed – predicted changes in subsequent depressive symptoms over short- and long-term intervals.

The prospective data were used to test whether marital transitions had effects on depression and mortality in real time. Where depressive symptom severity was the outcome (Studies 1 and 2), marital transitions were operationalized as taking place between measurements and used to predict subsequent depression outcomes. In this way, the time-limited effects of marrying (Study 1) and becoming widowed (Study 2) on depression were tested. Where mortality was the outcome (Study 3), twins who were discordant for divorce and married status were identified, and Cox proportional hazard

regression analysis (survival analysis) was used to test whether married twins had a higher probability of longevity than their divorced co-twins.

Isolation, depression, and mortality

The three studies presented in this dissertation fulfill two purposes. The first purpose was to explore the hypothesis that marriage confers advantages to adults by using a variety of longitudinal twin samples to test whether marriage protects people from depression and early mortality risk. Part of the first aim was to place people's depressive and mortality reactions to marrying, widowhood, and divorce within a lifespan context by considering sampling characteristics, mediators, gender differences, and generational differences. The second purpose extends beyond marital benefits to provide knowledge of the health consequences of isolation and loneliness. Social relationships often are taken for granted until people experience the emotional and physical sequelae associated with social isolation (Blazer, 2003). Marriage alone does not prevent depression and early death. Rather, social relationships and integration into communities – of which marriage is one component – more generally inhibit depression and lower mortality risk (Schwarzbach, Luppa, Forstmeier, König, & Riedel-Heller, 2013; Schwarzbach, Luppa, Sikorski, et al., 2013). Marriage, for many, is one node in the constellation of social relations that prevents loneliness and isolation from speeding along depression and death.

Entering First Marriage and Depression

Part of the United States Supreme Court decision to overturn the Defense of Marriage Act (DOMA) was based on the belief that marriage provides benefits that ought to be available to all adults and families (*United States v. Windsor, 2013*). Married people have been found to have better physical health (House, Landis, & Umberson, 1988; Kiecolt-Glaser & Newton, 2001; K. Williams & Umberson, 2004); lower risks for excessive alcohol use (Leonard & Mudar, 2003); higher subjective well-being (Diener & Seligman, 2002); longer lives (Lillard & Waite, 1995); and lower depressive symptomatology (Frech & Williams, 2007). The correlation between depressive symptoms (hereafter referred to as *depression* for ease of presentation) and entering marriage across the lifespan is the focus of this paper. Yet, while the *correlation* between marriage and depression has been widely explored (Horn et al., 2013; Mastekaasa, 1994a; Pearlin & Johnson, 1977; Waite & Gallagher, 2000), marital research is in constant need of empirical work to support the argument that marriage *causes* lower depression.

Entering marriage may increase happiness for otherwise depressed people who lack social support (Frech & Williams, 2007; Thoits, 1984). Marriage is correlated with decreases in depression relative to remaining single (Mastekaasa, 2006) and cohabiting (Lamb et al., 2003; Musick & Bumpass, 2012). The correlation between marriage and depression, however, may also be time-limited. Entering marriage correlates with modest, short bursts of happiness and life satisfaction, but the association diminishes to pre-marital levels with the passage of time (Lucas et al., 2003; Luhmann, Hofmann, Eid,

& Lucas, 2012). The temporal relation between marriage and depression has been studied but contains inconsistent findings. In a sample of young adults, entering and remaining married predicted less severe symptomatology 15 years later when compared to single controls (Horwitz et al., 1996).

Does marriage cobble together a lifestyle that *causes* decreases in depression? Or, are healthier people more likely to select into marriage in the first place? The latter – referred to as the *social selection hypothesis* – is the expectation that third variable confounds (measured or unmeasured) account for the correlation between marriage and depression. Some people, for example, might be more likely to select into marriage because they are happier, healthier, prettier, and wealthier, which makes them more appealing mates than the homely and infirm. Conversely, the *social causation hypothesis* suggests that marriage itself protects people against physical, psychological and mortality risks (Mastekaasa, 2006). The goal in marital research is to appropriately control for selection factors that partially account for why married people experience better health outcomes – like fewer depressive symptoms (Booth & Amato, 1991; Gove et al., 1983) – to test the social causation hypothesis.

Genetically informed studies are one way to take into account genetic and environmental selection effects that might confound the causal influence of entering marriage on depression. Twin studies in particular parse genetic and common environmental selection effects that make twins similar to one another from within environmental effects that make twins – in particular identical twins – different from one another (Turkheimer & Harden, 2013). Twin studies have been used in marital research

to show that entering marriage decreases antisocial behavior (Burt et al., 2010; Horn et al., 2013) and curbs alcohol consumption (Prescott & Kendler, 2001), adjusting for genetic and environmental selection confounds. Horn et al. found a within-family environmental effect of coupled status (married or cohabiting) on subsequent decreases in self-reported depression. In a pair of identical twins, for example, this means that the twin who coupled was also the twin who reported less depression at follow-up compared to the noncoupled twin.

One issue that often goes overlooked in marital studies is that entering marriage can occur at any point in the lifespan. While there are many studies that have focused specifically on marrying in early adulthood (Burt et al., 2010; Horn et al., 2013; Horwitz & White, 1991), there are no studies investigating the positive effects of entering first marriage in older adulthood. Instead, age heterogeneous samples frequently are pooled to increase overall sample size for detecting small effects (Karney & Bradbury, 1995). Pooled samples compromise the ability to generalize findings to specific populations of interest. For example, the effect of entering marriage on depression in a 25-year-old may not generalize to a 55-year-old's experience. The value of entering marriage for a 25-year-old may consist of financial stability and the potential to form a family whereas the value of marriage for a 55-year-old may be partner support and caretaking during the elderly years. Thus, the age people enter marriage may be crucial for whether marriage protects against depression.

Gender also must be considered in marital research, as the debate over "his" and "hers" marriage is on-going (Bernard, 1982; Kiecolt-Glaser & Newton, 2001). Marriage

has been found to lower men's depression (Horwitz et al., 1996; Marks, 1996; D. R. Williams et al., 1992) whereas women do not enjoy the same benefit. These differences, however, are not always observed (Blekesaune, 2008; Frech & Williams, 2007; Gove et al., 1983; Mastekaasa, 1992). Given the inconsistent replication of gender differences in marital studies, differential effects of entering marriage on depression must be considered in men and women.

Finally, cultural differences may also be a factor in who marriage benefits. For example, in nations with social welfare systems, like Scandinavian and Northern European countries (e.g., Sweden, Germany, and The Netherlands), marriage might not be as highly correlated with physical and mental health because many social benefits (e.g., financial and healthcare insurance) are provided by the state. In nonsocial welfare states, like the United States, entering marriage may be more highly correlated with health outcomes because it provides propitious environments associated with improved health. The present study used samples from Sweden and the United States to explore possible cultural differences in the effects entering marriage has on depression.

Present Study

The purpose of the present study was to use two longitudinal twin samples from different cultures to test the hypothesis that entering marriage causally influences lower depression in men and women over the lifespan. The following research questions were addressed:

1. Does the trajectory of depression decrease in early adulthood and increase in late adulthood?

2. Above and beyond the tendency for people to increase or decrease in depression as they age, does entering marriage predict subsequent decreases in depression?

3. Above and beyond genetic and environmental selection factors, does entering marriage predict subsequent decreases in depression?

4. Does entering marriage have different effects on depression for those in late adulthood as opposed to early adulthood?

5. Does entering marriage have different effects on depression for men and women?

Lifespan and cultural comparisons were made to the extent that the samples used could support such comparisons. The U.S. sample consisted of adults less than 35 years of age whereas the Swedish sample primarily consisted of adults over the age of 45.

Method

Sample

As noted above, two twin samples were used in the current study. The first sample consisted of 707 twin pairs from the National Longitudinal Study of Adolescent Health (Add Health; Harris, 2009), a nationally representative school-based study designed to assess health and risk behaviors of middle- and high-school aged youth (grades 7-12). The first wave of interviews was conducted in 1994 and 1995 (Wave 1), with follow-up interviews in 1996 (Wave 2), 2001-2002 (Wave 3), and 2008-2009 (Wave 4). Sample sizes at each wave are given by gender and zygosity in Table 1.

Table 1

Number of complete, incomplete, and missing twin pairs in Waves 1-4 of Add Health

	Complete Pairs			
	Wave 1	Wave 2	Wave 3	Wave 4
MZM	107	96	81	73
MZF	106	96	86	83
DZM	124	105	94	93
DZF	129	123	101	109
	Incomplete Pairs			
	Wave 1	Wave 2	Wave 3	Wave 4
MZM	2	5	10	20
MZF	2	4	16	18
DZM	1	7	21	21
DZF	3	5	14	12
	Missing Pairs			
	Wave 1	Wave 2	Wave 3	Wave 4
MZM	0	8	18	16
MZF	0	8	6	7
DZM	0	13	10	11
DZF	0	4	17	11

Note. MZM = monozygotic male; MZF = monozygotic female; DZM = dizygotic male; DZF = dizygotic female. Sample sizes were computed based on depressive symptom outcome scores.

The mean age of the male and female twins at each wave was as follows (based on complete and incomplete pairs): Wave 1: M_{Men} = 16.15 (SD = 1.62) and M_{Women} = 15.99 (SD = 1.54); Wave 2: M_{Men} = 17.09 (SD = 1.62) and M_{Women} = 16.94 (SD = 1.55);

Wave 3: M_{Men} = 22.50 (SD = 1.66) and M_{Women} = 22.32 (SD = 1.58); and Wave 4: M_{Men} = 29.03 (SD = 1.69) and M_{Women} = 28.80 (SD = 1.56).

The second sample consisted of 1,919 twin pairs from the Swedish Adoption/Twin Study of Aging (SATSA; Finkel & Pedersen, 2004), a gerontological study designed to assess physical and psychological health in a representative sample of middle- and late-aged Swedish adults. The questionnaire-based component was used in the current study, which began in 1984, with follow-up measures sent to all still-living participants every three years until 1993. Two additional follow-up questionnaires were sent to the still-living participants in 2004 and 2007. Marital status information was collected at six waves (1984, 1987, 1990, 1993, 2004, and 2007) and depressive symptom measures were administered at the last five waves. Sample sizes at each wave are given by gender and zygosity in Table 2.

Table 2

Number of complete, incomplete, and missing twin pairs from 1987-2007 in the SATSA

	Complete Pairs				
	1987	1990	1993	2004	2007
MZM	80	71	74	32	24
MZF	93	84	75	50	35
DZM	114	100	107	52	35
DZF	187	172	165	90	77
	Incomplete Pairs				
	1987	1990	1993	2004	2007
MZM	73	82	68	73	69
MZF	92	88	90	56	64
DZM	170	170	160	148	149
DZF	216	207	214	165	157
	Missing Pairs				
	1987	1990	1993	2004	2007
MZM	106	106	117	154	166
MZF	145	158	165	224	231
DZM	204	218	221	288	304
DZF	303	327	327	451	472

Note. MZM = monozygotic male; MZF = monozygotic female; DZM = dizygotic male; DZF = dizygotic female. Sample sizes were computed based on depressive symptom and marital status information.

Given the aging focus of the SATSA, mortality is a major cause of attrition across the 20-year-measurement window. In 1987, 3.10% (n = 119) of the total sample was deceased, with increases at each follow-up wave: 7.22% (n = 277) in 1990, 10.34% (n =

397) in 1993, 26.73% (n = 1026) in 2004, and 30.41% (n = 1167) in 2007. The age range of the sample in 1984 was 26 years to 93 years. The mean age of the male and female participants at each follow-up wave was as follows (based on complete and incomplete pairs): in 1987, M_{Men} = 60.66 (*SD* = 13.00) and M_{Women} = 62.12 (*SD* = 13.91); in 1990, M_{Men} = 62.73 (*SD* = 12.25) and M_{Women} = 64.22 (*SD* = 13.54); in 1993, M_{Men} = 63.76 (*SD* = 12.33) and M_{Women} = 65.83 (*SD* = 13.40); in 2004, M_{Men} = 69.81 (*SD* = 10.47) and M_{Women} = 70.10 (*SD* = 11.75); and in 2007, M_{Men} = 71.47 (*SD* = 9.79) and M_{Women} = 71.97 (*SD* = 10.83). The gradual increase in mean ages over time reflects the higher retention of younger participants and attrition due to death among older participants.

Measures

Depressive Symptoms

The Center for Epidemiological Studies–Depression (CES-D) scale is a 20-item scale widely used as a screening measure of depressive symptomatology (Radloff, 1977). Participants rate how often they experienced each symptom over the past 7 days on a scale of 0-3, with 0 = *rarely or none of the time (less than 1 day),* 1 = *some or a little of the time (1-2 days),* 2 = *occasionally or a moderate amount of time (3-4 days); and* 3 = *most or all of the time (5-7 days)*. The reliability and validity of the CES-D scale has been well-studied, with 4 facets typically found to comprise the total scale: depressed affect, lack of well-being, somatic and retarded activity, and interpersonal difficulties (Hertzog, Alstine, Usala, Hultsch, & Dixon, 1990; Shafer, 2006). All 20-items are presented in Appendix A.

In the Add Health data set, there were 9 items common across all 4 waves. An exploratory factor analysis of the items revealed that the subset of items across all waves was generally consistent with the four-factor solution found in prior studies (Shafer, 2006). As a result, the total sum score was used as an overall depressive symptom score index (CES-D-9). The log of the total sum score was used, as the distribution of the raw total scores was positively skewed. Histograms of the log scores are presented in Appendix B. In keeping with conventional reliability value standards (Shrout, 2002), the test-retest reliability across all five measurements was moderate for both men (McDonald's $\omega = .70$) and women (McDonald's $\omega = .71$) (McDonald, 1999).

In the SATSA, the log of the sum of all 20-items was used at each wave, as the distribution of the raw total scores was positively skewed. Histograms of the log scores are presented in Appendix B. The test-retest reliability across all five measurements is substantial for both men (McDonald's $\omega = .87$) and women (McDonald's $\omega = .85$).

Entering Marriage

In Add Health, the *entering marriage* variable was a binary coded variable operationalized as a change from nonmarried status (*never married*) at wave 3 to married status (*married*) at wave 4. Few participants in Add Health were of marriageable age at Wave 1 and Wave 2. Participants who transitioned into marriage between any two waves were given a code of 1. Participants who reported being never married between adjacent waves were considered to have remained unmarried and were coded as 0 (no change). Participants who divorced or remarried between adjacent waves were also coded 0. Of the 524 male twins, none of the male twins were already married by

wave 1, 0.57% entered marriage between wave 1 and wave 2, 9.73% entered marriage between wave 2 and wave 3, and 23.09% entered marriage between wave 3 and wave 4. Of the 528 female twins, 0.38% were already married by wave 1, 0.76% entered marriage between wave 1 and wave 2, 15.34% entered marriage between wave 2 and wave 3, and 25.00% entered marriage between wave 3 and wave 4.

In the SATSA, the *entering marriage* variable also was a binary coded variable operationalized as a change from a nonmarried status at time *t* (e.g., 1987) to married status at time *t*+1 (e.g., 1990). Participants who transitioned into marriage between any two waves were coded 1. Few people in the SATSA married at all during the study period, so participants who entered first marriages or remarried were considered to "enter marriage." The decision to include participants who remarried may overestimate the positive effects of entering marriage, as remarriage has been found to increase well-being among widowed (Gentry & Shulman, 1988) and divorced people (Blekesaune, 2008). Participants who reported the same marital status or were missing at adjacent waves of measurement were coded 0. Few participants entered marriage in the SATSA. Of the 1,494 male twins, 5.22% entered marriage between 1984 and 1987, 3.88% entered marriage between 1987 and 1990, 3.55% entered marriage between 1990 and 1993, 1.20% entered marriage between 1993 and 2004, and 0.40% entered marriage between 2004 and 2007. Of the 2,072 female twins, 4.01% (n = 2,072) entered marriage between 1984 and 1987, 2.51% entered marriage between 1987 and 1990, 2.85% entered marriage between 1990 and 1993, 0.53% entered marriage between 1993 and 2004, and 0.63% entered marriage between 2004 and 2007.

Data Analysis

The analyses for this study consisted of two steps: individual-level analyses and twin-level analyses. The individual-level analyses addressed the first and second research questions, and the twin-level analyses addressed the third research question. The fourth question regarding generational differences was addressed by comparing the results from both analyses. The fifth research question regarding gender differences were addressed in both steps of the analyses.

In the individual-level analyses, participants' mean CES-D scores were first plotted by age group and gender. Age effects were evaluated by examining longitudinal trajectory plots of the participants' CES-D scores, first by wave of measurement and second by age of measurement. Latent growth curve (LGC) models were then fit to the male and female samples separately with change in marital status at each wave included as covariates. Slightly different LGC models were fit to the Add Health (Figure 1) and SATSA (Figure 2) samples, because of differences in the measurement of marital status.

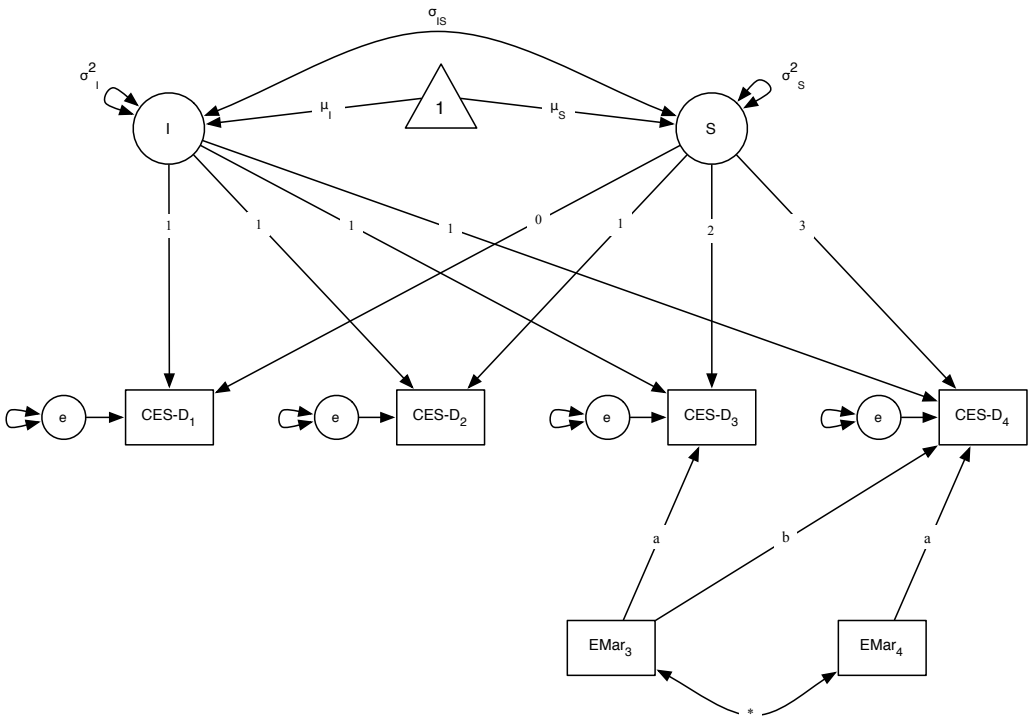

Figure 1. Linear Growth Curve Model. Squares represent observed variables, circles represent unobserved latent constructs, and the triangle represents the mean structure of the latent variables. EMar = entered marriage; CES-D = log total Center for Epidemiological Studies–Depression scale score; I = intercept; S = slope; 1-4 = wave of measurement

Latent growth curve (LGC) models were fit using the M*plus* 7.11 program (Muthén & Muthén, 2014). The TYPE=COMPLEX function was used in the LGC analyses to take into account the clustered family data in the estimation of the standard errors and chi-square tests of model fit. While twins are necessarily age-matched, the ages at which the twins were measured varied within each wave. Age heterogeneity within each wave of measurement violates the assumption that the intervals between measurements are the same for the entire sample (Duncan et al., 2006). To correct for this assumption violation, the individually-varying times of observation across individuals were taken into account in the model by applying multivariate SEM methods and

multilevel modeling methods. The TSCORES option in M*plus* achieves this by identifying the variables in the data on which the CES-D measures individually varied.

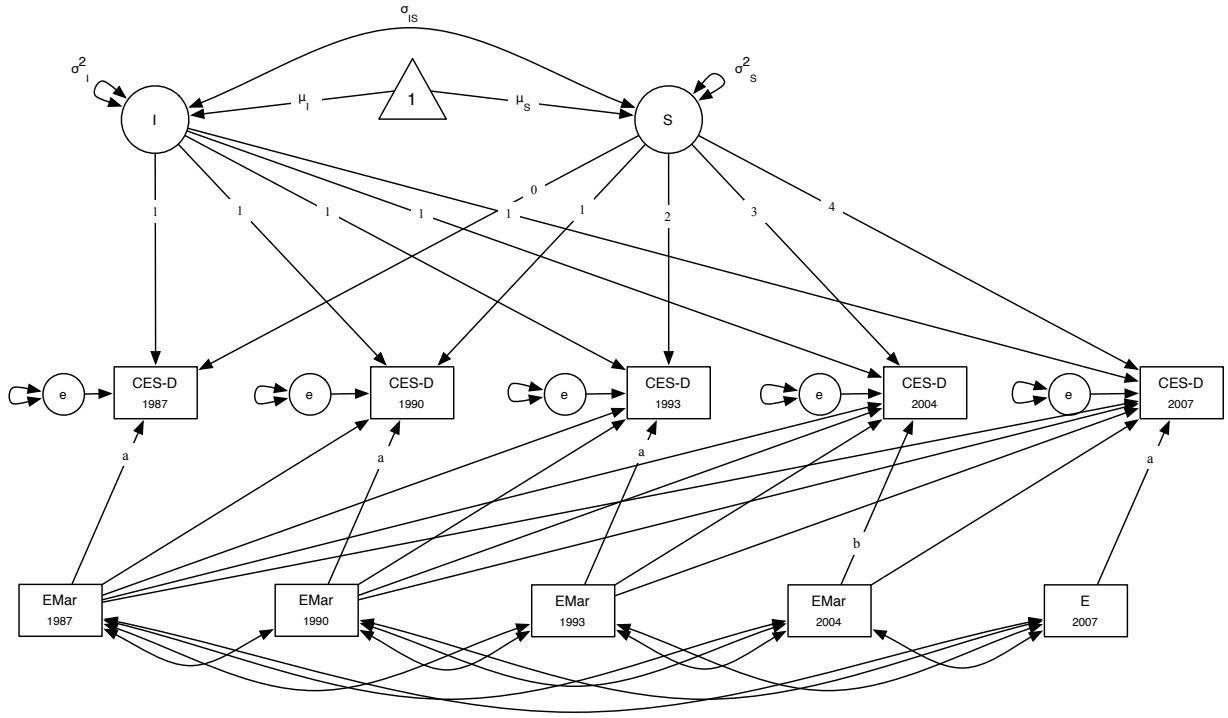

Figure 2. Linear Growth Curve Model. Squares represent observed variables, circles represent unobserved latent constructs, and the triangle represents the mean structure of the latent variables. EMar = entered marriage; CES-D = log total Center for Epidemiological Studies–Depression scale score; I = intercept; S = slope; a = short-term effect of entering marriage on depressive symptom scores; b = long-term effect of entering marriage on depression symptom scores; all non-labeled regression paths and covariances were freely estimated in the model.

The LGC models served two purposes. First, they were used to answer the first research question regarding intraindividual trends in depression over the second half of the lifespan. Second, they were used to test the short-term and long-term phenotypic effects of entering marriage on CES-D scores, above and beyond the individual linear growth in depression over time. Short-term and long-term effects were operationalized differently in Add Health and the SATSA, because the intervals of time between waves

of measurement were different in each study. In Add Health, short-term effects were defined as entering marriage within a 6-year period prior to follow-up. Long-term effects were defined as entering marriage within a 12-year period prior to follow-up. In the SATSA, short-term effects were defined as entering marriage within a 3-year period prior to follow-up whereas long-term effects were defined as entering marriage within an 11-year period prior to fellow-up. The time lags were operationalized this way because participants' dates of marriage were not recorded. If either the short-term or long-term phenotypic effect of marrying on CES-D scores was significant in the LGC analysis, twin models were then fit to the data to test the social causation hypothesis.

In the twin-level analyses, the first step consisted of a co-twin control study. Twins who were discordant for marital status (entered marriage and did not enter marriage) were identified at each wave of measurement. Intrapair differences were then calculated by subtracting the twin-who-entered-marriage's CES-D score from the twin-who-did-not-enter-marriage's CES-D score. Paired *t*-tests were used to test whether the twin who entered marriage had significantly lower CES-D scores than the twin who did not enter marriage. All co-twin-control analyses were conducted in R 3.0.2 (R Core Team, 2013).

Next, multivariate longitudinal twin analyses (referred to as "longitudinal quasi-causal models") were fit to the data to test whether within-family environmental differences in entering marriage mediated the short-term and long-term effects on CES-D scores, holding constant the genetic and environmental selection effects underlying their association. The model is presented in Figure 3. For simplicity of presentation, only

two time points are shown from the Add Health analysis. These models provided a formal test of the social causation hypothesis of marriage.

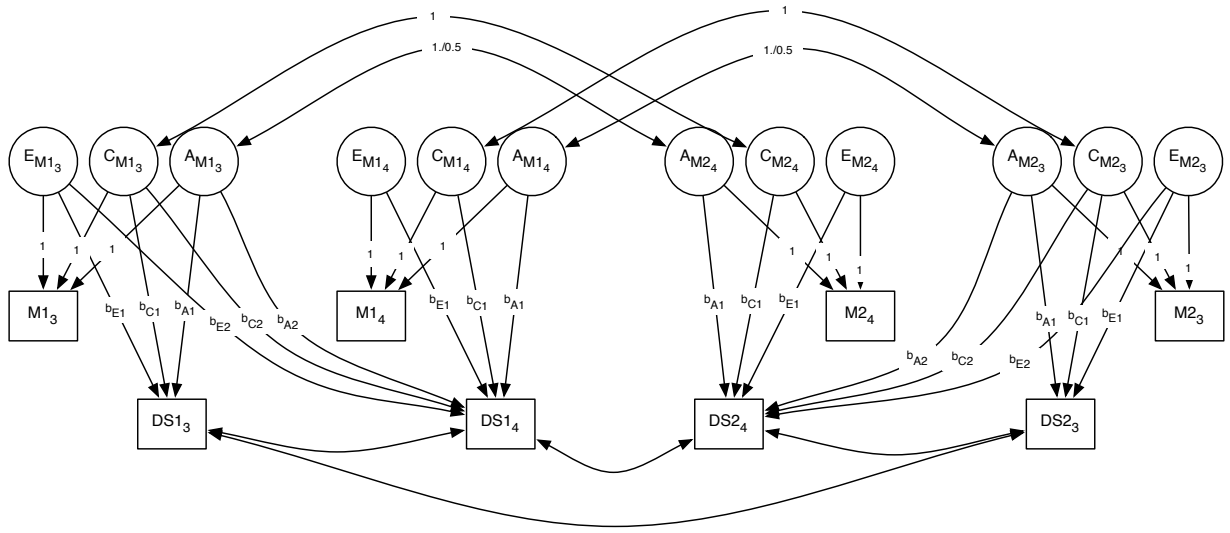

Figure 3. Longitudinal quasi-causal model. Only two waves of measurement (for example, subscripts "3" and "4" represent waves 3 and 4 in Add Health) are shown for simplicity. Actual models were estimated using all waves of measurement in Add Health and the SATSA. Squares represent observed variables, and circles represent unobserved latent constructs. M = entered marriage variable; DS = log total Center for Epidemiological Studies–Depression scale score; A = additive genetic factor; C = common environmental factor; E = nonshared environmental factor; b_1 = short-term regression coefficient; b_2 = long-term regression coefficient. Variable names are appended with "1" or "2" to correspond with twin 1's scores and twin 2's scores.

As can be seen in Figure 3, the binary marriage variables were decomposed into three latent variables (or effects): genetic (A) effects, common environmental (C) effects, and nonshared environmental (E) effects. Genetic effects refer to the common genes twins share that make them resemble one another. Monozygotic (MZ) twins share 100% of their genes whereas dizygotic (DZ) twins share 50% of their genes, on average. In standard twin models, the twins' genetic effects are correlated 1.0 in the MZ

group and 0.5 in the DZ group. Common environmental effects refer to any environmental factor (e.g., parental SES, rearing environment) that makes twins from the same family more similar to one another. Common environmental effects are assumed to affect MZ and DZ twins equally and are correlated 1.0 in both zygosity groups. Finally, nonshared environment effects (E) refer to any environmental factor that makes twins different from one another, including measurement error, and are uncorrelated between twins. In identical twin pairs, the nonshared environment is equal to the unreliability between identical twin pairs – that is, one minus the MZ correlation.

Conventional twin models make three additional assumptions (Neale & Cardon, 1992). First, the ACE variance components are uncorrelated with one another. Second, the ACE components do not interact with one another. Third, mating between the twins' parents is random, which is implied by the assumption that DZ twins share 50% of their genes.

The genetic and common environmental regression effects (labeled b_{A1} and b_{C1} in Figure 3), if significant, represent support for the selection hypothesis. The nonshared environmental regression effects (b_{E1} in Figure 3), if significant, represent support for the causal hypothesis. Only random within-family environmental differences in marrying can explain observable differences in CES-D scores between twins in the same family. The Add Health and SATSA studies are nonexperimental designs, so it is important to keep in mind that strict causal inferences cannot be made. Other unmeasured, third variable factors remain unaccounted for in the models, so significant within-family environmental effects are considered to be "quasi-causal."

Ideally, the longitudinal quasi-causal model and LGC model could have been fit in a single model. The study design features of SATSA, however, preclude incorporating the TSCORES function in M*plus* to take into account the individually-varying times of observation across individuals *and* using dependent categorical variables necessary for decomposing binary variable into latent ACE variables in the same model. Although this model is possible in other SEM software (e.g., OpenMx), the LGC models and longitudinal quasi-causal models were fit separately in the present study. The descriptive results suggest that the substantive conclusions are preserved regardless of using a one-step or two-step modeling procedure.

Missing data were assumed to be missing at random. Full-information maximum likelihood estimation with robust standard errors (MLR) for the LGC models and weighted least squares mean (WLSM) estimation for the longitudinal quasi-causal models were used to handle missingness. With both estimators, the Satorra-Bentler scaled chi-square difference test was used to calculate a chi-square distributed test statistic to compare nested models (Satorra & Bentler, 2001). Additionally, the Root Mean Square Error of Approximation (RMSEA) was used to evaluate model fit (Browne & Cudeck, 1992). Models with RMSEA values below 0.05 are considered to fit the data well, and models with values below 0.08 are considered to fit the data adequately. The Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) also were used to evaluate model fit (Burnham & Anderson, 2004). Both the AIC and BIC indexes are computed to balance model parsimony and model complexity (Kline, 2005). Lower values indicate better model fit to the data. The comparative fit index (CFI) was used as

a final measure of model fit, as it evaluates the incremental improvement in model fit between an alternative model (the model under consideration) and a baseline model (the null model). Values range from 0-1 with higher values indicating better model fit; a criterion cutoff of .90 is the conventional lower limit for model selection (L. Hu & Bentler, 1995).

Results

Entering Marriage: Young Adulthood

Individual-level Analysis

A description of the male and female participants' CES-D-9 means was the initial analysis used to answer the first research question (*Does the trajectory of depression decrease in early adulthood?*). Figure 4 presents longitudinal trajectory plots of the male participants' CES-D-9 scores across all four waves, with the wave means overlaid in red.

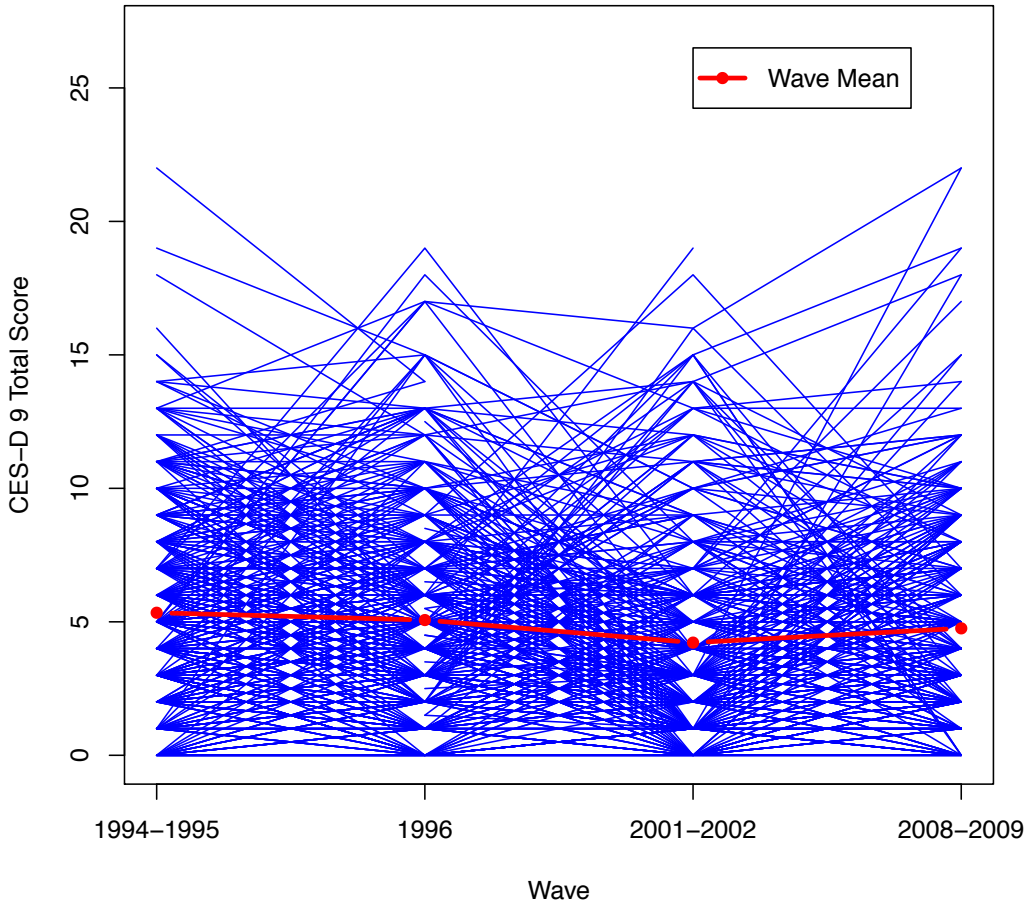

Figure 4. Longitudinal trajectory plots of male participants' total CES-D-9 scores across waves 1-4 in Add Health. The bolded red line represents the mean total CES-D-9 score at each wave.

As can be seen from the figure, the CES-D-9 means are relatively flat over the entire study period. Rather, the participants' scores oscillate which suggests an auto-regressive pattern rather than a growth pattern. Plotting the longitudinal scores across waves, however, ignores the age heterogeneity within each wave of measurement. The male participants' scores were plotted by age, with each twin contributing four data points (at most) to the plot (Figure 5).

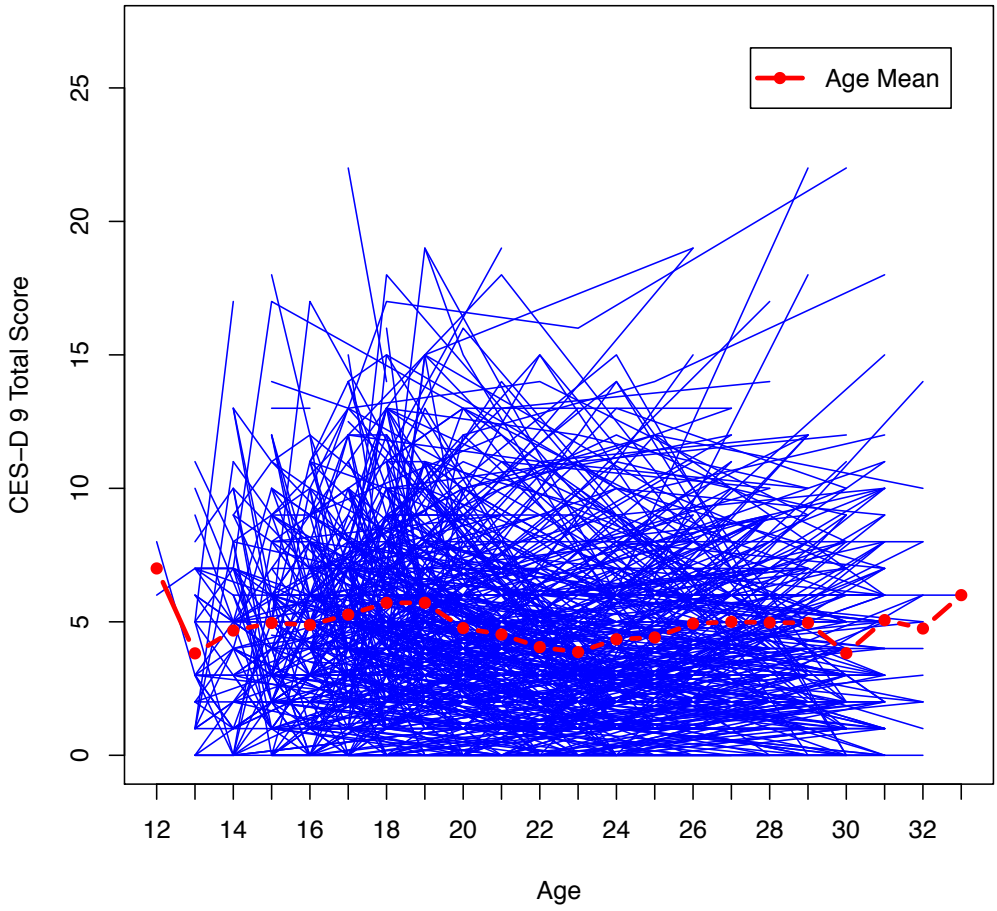

Figure 5. Longitudinal trajectory plots of male participants' total CES-D-9 scores strung out by age in Add Health. Yearly age bins were created from age 12-33. The bolded dotted red line represents the mean total CES-D-9 score for each age bin.

The red line illustrates that the CES-D-9 mean scores did not substantially change from age 12 to age 33. Although there was a slight overall decrease in scores over the entire age window, with an increase from age 13 to age 19 and a decrease from age 20 to age 23, the participants' patterns again suggest an auto-regressive pattern of CES-D-9 scores over time rather than growth.

Figure 6 presents the longitudinal trajectory plots of the female participants' CES-D-9 scores across the four waves with the wave means overlaid in red.

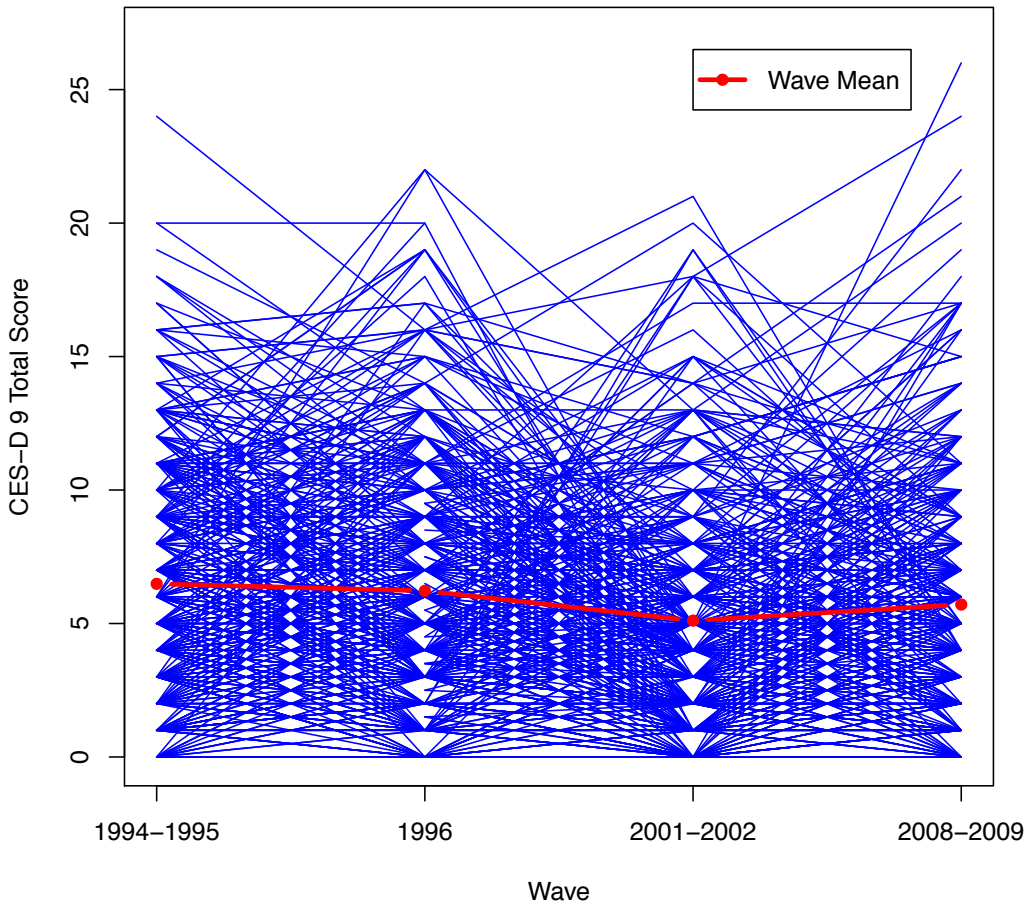

Figure 6. Longitudinal trajectory plots of female participants' total CES-D-9 scores across waves 1-4 in Add Health. The bolded red line represents the mean total CES-D-9 score at each wave.

As in the men, the participants' means followed an auto-regressive pattern, as their scores swung from high to low across the measurement occasions. As a result, the overall mean remained relatively flat across the four waves. Nearly the same impression

of change emerged from early adolescence through young adulthood in the female participants when the participant's scores were strung out by age (Figure 7).

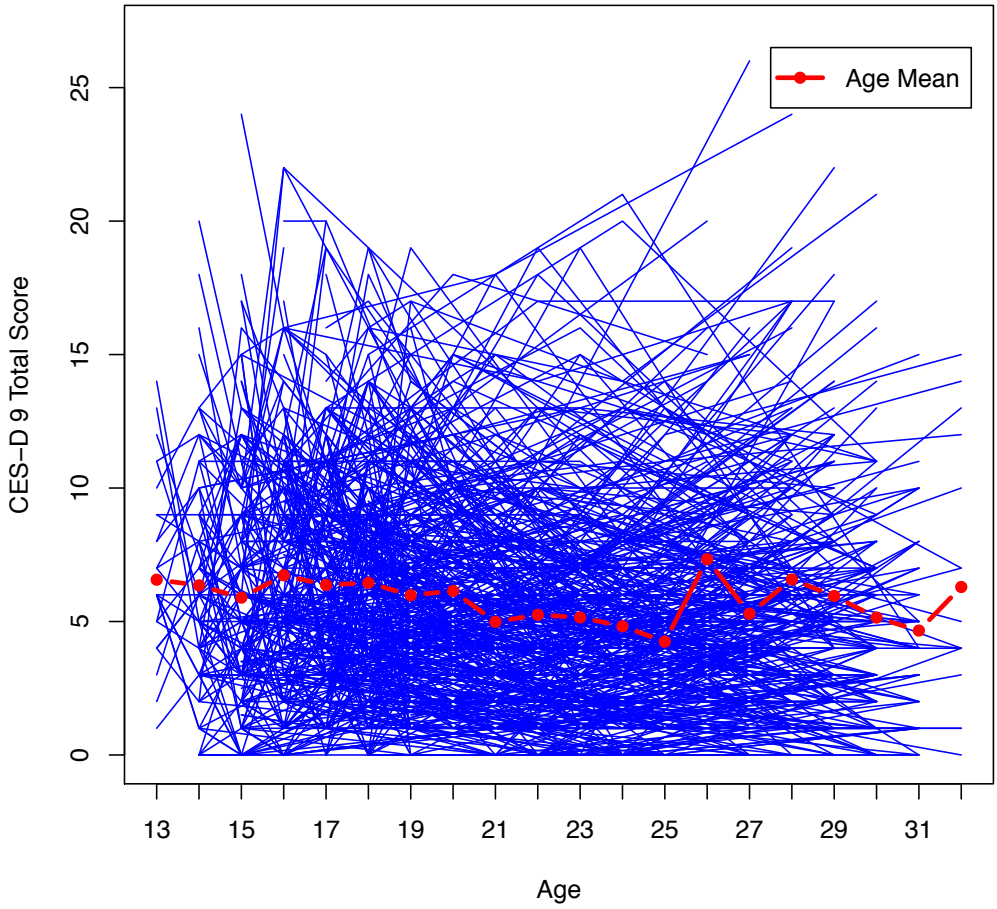

Figure 7. Longitudinal trajectory plots of female participants' total CES-D-9 scores strung out by age in Add Health. Yearly age bins were created from age 13-32. The bolded dotted red line represents the mean total CES-D-9 score for each age bin.

Next, the CES-D 9 means of male and female participants who entered marriage were compared to the CES-D-9 means of their counterparts who did not enter marriage. Assuming that marriage lowers people's depression, the expectation was that the CES-D-9 means of those who entered marriage would be lower than the CES-D-9 means of

those who did not enter marriage. Table 3 presents the means, split into three cohorts to take into account potential age effects. The *young cohort* consists of all twins who were between the ages of 12 and 15 in Wave 1; the *middle cohort* consists of all twins who were between the ages of 15 and 18 in Wave 1; and the *old cohort* consists of all twins who were older than 18 years of age in Wave 1.

Table 3

Descriptive CES-D-9 results for Add Health participants who did and did not enter marriage, by gender and age group

	Men						Women					
	No Marital Status Change			Entered Marriage			No Marital Status Change			Entered Marriage		
	n	*M*	*SD*	n	*M*	*SD*	n	*M*	*SD*	n	*M*	*SD*
Wave 1												
Young (12-14)	129	4.81	3.57	-	-	-	137	6.20	4.09	-	-	-
Middle (15-17)	307	5.28	3.65	-	-	-	288	6.56	4.50	2	5.50	3.54
Old (>18)	61	6.75	3.82	-	-	-	58	6.83	4.43	-	-	-
Wave 2												
Young (12-14)	56	4.95	3.38	-	-	-	69	5.39	3.99	-	-	-
Middle (15-17)	286	5.08	3.83	-	-	-	295	6.50	4.68	3	7.67	5.51
Old (>18)	53	6.35	3.85	2	2.50	3.54	56	5.99	4.71	1	3.00	NA
Wave 3												
Young (12-14)	99	4.12	3.48	3	3.33	1.53	102	5.07	4.16	13	8.46	4.41
Middle (15-17)	186	4.23	3.77	34	2.91	2.70	167	5.16	4.15	48	4.88	5.08
Old (>18)	35	4.94	3.22	14	3.43	2.62	23	4.30	3.87	19	4.32	2.73
Wave 4												
Young (12-14)	59	5.46	3.90	22	3.77	3.52	72	6.31	4.97	29	5.41	4.67
Middle (15-17)	127	4.56	3.84	80	4.45	3.76	126	5.64	4.00	89	5.90	4.72
Old (>18)	28	5.46	4.03	19	4.26	3.41	29	5.14	3.83	14	4.50	4.00

Due to the youthfulness of the sample through Wave 2, few twins married between Waves 1 and 2. At Wave 3, men who entered marriage had lower CES-D-9 scores than

those who did not enter marriage, which held across all cohorts. At Wave 4, men from the young and old cohorts had lower CES-D-9 scores, on average, than those who did not enter marriage.

For the women, a different pattern of results emerged than in the male participants (Table 3). First, women who entered marriage had higher CES-D-9 scores after entry into marriage than women who did not enter marriage (middle cohort in Wave 2 and young cohort in Wave 3). The means of the married group, however, were based on a small number of participants and should be interpreted with caution. Second, at Waves 3 and 4, the mean differences after entry into marriage were not very different from controls who did not enter marriage, particularly in the middle and old cohorts.

Overall, the individual-level results suggest that entering marriage was associated with a slight general decrease in depression from early adolescence through young adulthood in men and women, although the change generally followed an auto-regressive pattern than a linear growth pattern over time. The analysis of CES-D-9 means, however, suggests that entering marriage is associated with lower subsequent CES-D-9 scores in men after entry into marriage but few differences in favor of entering marriage for women, a finding consistent with Bernard's gender difference argument (Bernard, 1982).

Next, LGC models were fit to the data to answer the second research question (*Above and beyond the tendency for people to decrease in depression over the early adult years, does entering marriage predict subsequent decreases in depression scores?*). Model fitting results are presented for men and women in Table 4.

Table 4

Latent growth curve model fitting results in Add Health

Male Twins	-2LL	Parameters	Scaling correction factor	Model Comparison	LR	S-B LR test	AIC	BIC
1. Baseline model (Correlated intercept and slope with unstructured correlations among measured variables)	-2166.66	22	1.16				4377.31	4471.02
2. Short-term correlations constrained to be equal	-2166.66	21	1.16	1	0.00	1.000	4375.31	4464.76
3. Constrain short-term correlation to be zero	-2171.28	20	1.17	2	9.86	0.002	4382.56	4467.75
4. Constrain long-term correlation to be zero	-2167.00	20	1.18	2	0.93	0.336	4374.00	4459.19
Female Twins								
1. Baseline model (Correlated intercept and slope with unstructured correlations among measured variables)	-2348.52	22	1.10				4741.04	4834.84
2. Short-term correlations constrained to be equal	-2348.70	21	1.11	1	0.34	0.561	4739.39	4828.92
3. Constrain short-term correlation to be zero	-2348.72	20	1.10	2	0.05	0.829	4737.44	4822.71
4. Constrain long-term correlation to be zero	-2350.74	19	1.11	3	4.44	0.035	4739.48	4820.49

Note. The scaling correction factor is a correction that better approximates chi-square under conditions of multivariate nonnormality (Satorra & Bentler, 2001). -2LL = -2 log-likelihood; LR = likelihood ratio; S-B = Satorra-Bentler likelihood ratio difference for nested model comparison; AIC = Akaike Information Criterion; BIC = Bayesian Information Criterion.

In the men, the short-term effect of marriage (paths labeled *a* in Figure 1) could be constrained to be the same without a significant reduction in model fit (Model 2) and was significantly different from zero (Model 3). The long-term effect of entering marriage (path labeled *b* in Figure 1) was not significantly different from zero and could be dropped from the model without significant loss of model fit (Model 4). The parameter estimates from Model 4 are presented in Table 5. The short-term effect of entering marriage on subsequent log CES-D-9 scores is -0.03 ($SE = 0.01$, $p < .001$), adjusting for the variance accounted for by the intercept and slope of the CES-D-9 scores. This means that entering marriage predicted a 0.03 unit decrease in log CES-D-9 scores

over the short-term, above and beyond intraindividual trends in depressive symptoms across all waves.

Table 5

Parameter estimates from the best fitting LGC models for male (Model 3) and female (Model 4) participants in Add Health

	MLR estimates			
	Male		Female	
Parameter	Est.	*SE*	Est.	*SE*
μ_I	1.74	0.08	1.97	0.07
μ_S	-0.01	0.003	-0.01	0.003
σ^2_I	1.04	0.16	0.61	0.18
σ^2_S	0.002	0.00	0.001	0.00
σ_{IS}	-0.04	0.01	-0.02	0.01
$\sigma_{M,D}$ - short-term	-0.03	0.01	0.00	0.01
$\sigma_{M,D}$ - long-term	-0.01	0.01	0.03	0.01
σ^2_{eD1}	0.20	0.02	0.25	0.03
σ^2_{eD2}	0.23	0.02	0.26	0.03
σ^2_{eD3}	0.38	0.03	0.39	0.03
σ^2_{eD4}	0.22	0.05	0.24	0.05
σ^2_{eM3}	0.10	0.01	0.15	0.01
σ^2_{eM4}	0.20	0.01	0.21	0.01

Note. μ = mean; σ^2 = variance; σ = covariance; I = intercept; S = slope; M = entered marriage; D = log CES-D score; 1-4 = wave of measurement. Short-term refers to the effects of entering marriage up to 6 years later. Long-term refers to the effects of entering marriage up to 11 years later.

In the women, the short-term effects of marriage (paths labeled *a* in Figure 1) could be constrained to be the same without a significant reduction in model fit (Model 2

in Table 4), but ultimately could be dropped from the model without loss of model fit (Model 3). The long-term effect of entering marriage (path labeled *b* in Figure 1), however, could not be set to zero without significant reduction in model fit (Model 4). The parameter estimates from Model 3 are presented in Table 5. The long-term effect of entering marriage on CES-D-9 score was 0.02 (*SE* = 0.01, *p* = .042), adjusting for the variance accounted for by the intercept and slope of the CES-D-9 scores across all waves. This means that entering marriage predicted a 0.02 unit *increase* in log CES-D-9 scores over the long-term, above and beyond intraindividual trends in depressive symptoms across all waves.

Twin-level Analysis

Co-twin-control analyses were used as the first step toward answering the third research question (*Above and beyond genetic and environmental selection factors, does entering marriage predict subsequent decreases in depression scores?*). Table 6 presents a contingency table of the total number of twin pairs concordant for marital status (either both twins entered marriage or both twins remained single) and discordant for marital status, separated by zygosity and gender groups (monozygotic male, monozygotic female, dizygotic male, and dizygotic female) at each wave. For male and female twins, only Wave 3 (male = 27 pairs, female = 43 pairs) and Wave 4 (male = 54, female = 57 pairs) had large enough samples of discordant twin pairs to carry out the co-twin control analysis.

Table 6

Contingency tables of male and female twin pairs concordant and discordant for marital status in Add Health

		Wave 1							
		Twin 1							
		MZM		MZF		DZM		DZF	
		No Δ MS	Married	No Δ MS	Married	No Δ MS	Married	No Δ MS	Married
Twin 2	No Δ MS	118	-	113	-	141	-	145	-
	Married	0	0	2	0	0	0	0	0
		Wave 2							
		Twin 1							
		MZM		MZF		DZM		DZF	
		No Δ MS	Married	No Δ MS	Married	No Δ MS	Married	No Δ MS	Married
Twin 2	No Δ MS	92	-	89	-	100	-	116	-
	Married	1	0	1	1	1	0	0	0
		Wave 3							
		Twin 1							
		MZM		MZF		DZM		DZF	
		No Δ MS	Married	No Δ MS	Married	No Δ MS	Married	No Δ MS	Married
Twin 2	No Δ MS	84	-	60	-	92	-	84	-
	Married	16	3	22	5	11	6	21	6
		Wave 4							
		Twin 1							
		MZM		MZF		DZM		DZF	
		No Δ MS	Married	No Δ MS	Married	No Δ MS	Married	No Δ MS	Married
Twin 2	No Δ MS	54	-	49	-	57	-	48	-
	Married	30	5	22	12	24	15	35	14

The co-twin-control analysis tested the differences in CES-D-9 mean scores of the twins who entered marriage to the mean scores of their co-twins who did not enter marriage. The results are presented in Table 7. Positive scores indicate that the twin who entered marriage was less depressed, on average, than the co-twin.

Table 7

Co-twin-control analysis of differences on CES-D-9 depression scale between twins who entered marriage and twins who did not enter marriage in Add Health

No Marital Status Change vs. Entered Marriage			
	Pairs	Mean Intrapair Difference	*t*
Wave 1			
Men	-	-	-
Women	2	2.50	0.93
Wave 2			
Men	2	3.00	1.18
Women	1	-1.00	NA
Wave 3			
Men	27	1.90	2.02
Women	43	0.79	0.76
Wave 4			
Men	54	0.53	0.58
Women	57	0.21	0.25

In the male twins, the direction of the effects generally suggest that twins who entered marriage were less depressed than their co-twins who did not enter marriage. Only the paired *t*-test at Wave 3 for the male twins was statistically significant, suggesting that the twin who married was less depressed than the co-twin who did not enter marriage (*p*

= .050). There were no significant differences in the CES-D-9 means between pairs of female twins discordant for entering marriage.

A graphical examination (bar plots) of the differences between discordant MZ and DZ twin pairs can provide a better understanding of the within-family effects of entering marriage on depression as well as the potential genetic and environmental selection effects underlying the association (Figure 8 and Figure 9). When the within-family difference in CES-D-9 scores between discordant MZ pairs is less than the difference between discordant DZ pairs, the interpretation is that genetic effects partially confound the effect of entering marriage on depression. To the extent that the MZ twin who entered marriage has a lower CES-D-9 score than the MZ twin who did not enter marriage, there is evidence of a nonshared environmental effect, consistent with the social causation hypothesis.

Figure 8 presents the bar plots of discordant male twins' CES-D-9 scores. Overall, MZ and DZ twins who entered marriage generally reported lower depression than their co-twins who did not enter marriage. At Wave 3, discordant MZ twins generally were no more different than discordant DZ twins. The difference between unrelated groups discordant for entering marriage, however, was smaller than the twin pair differences, suggesting that common environmental selection effects may have suppressed the within-family environmental benefits of marrying. This finding, unfortunately, is unreliable and cannot be trusted. Upon further analysis, the sample sizes used to calculate the mean differences in discordant MZ ($n = 15$) and DZ ($n = 9$) pairs and the unrelated groups ($n_{\text{Married}} = 332$, $n_{\text{Not Married}} = 1{,}652$) likely affected the

mean estimates. When, for example, the CES-D-9 mean of 15 randomly selected individuals who entered marriage was compared to the CES-D-9 mean of 15 randomly selected individuals who did not enter marriage – an *N* of 15 was chosen to match the sample of discordant MZ pairs of twins – the difference was 2.88 units (compared to a difference of 1.99 in MZ pairs and a difference of 1.76 in DZ pairs), with those who did not enter marriage reporting higher depression scores than those who entered marriage. In contrast to suppression, this pattern of results suggests a common environmental selection effect underlying the correlation between marriage and depression. Thus, the small number of discordant MZ and DZ pairs used to calculate the CES-D-9 means may have inflated the within-family differences between them.

The small difference in CES-D-9 means between discordant MZ twins at Wave 4 does not support the social causation hypothesis (in fact, the twin who entered marriage was slightly more depressed than the twin who did not enter marriage). The smaller absolute difference between MZ twins than DZ twins, moreover, suggests an underlying genetic selection process that mediated the correlation between marrying and CES-D-9 scores. Again, the sample sizes used to calculate the means of discordant MZ and DZ pairs were small, so the estimates should not be considered reliable.

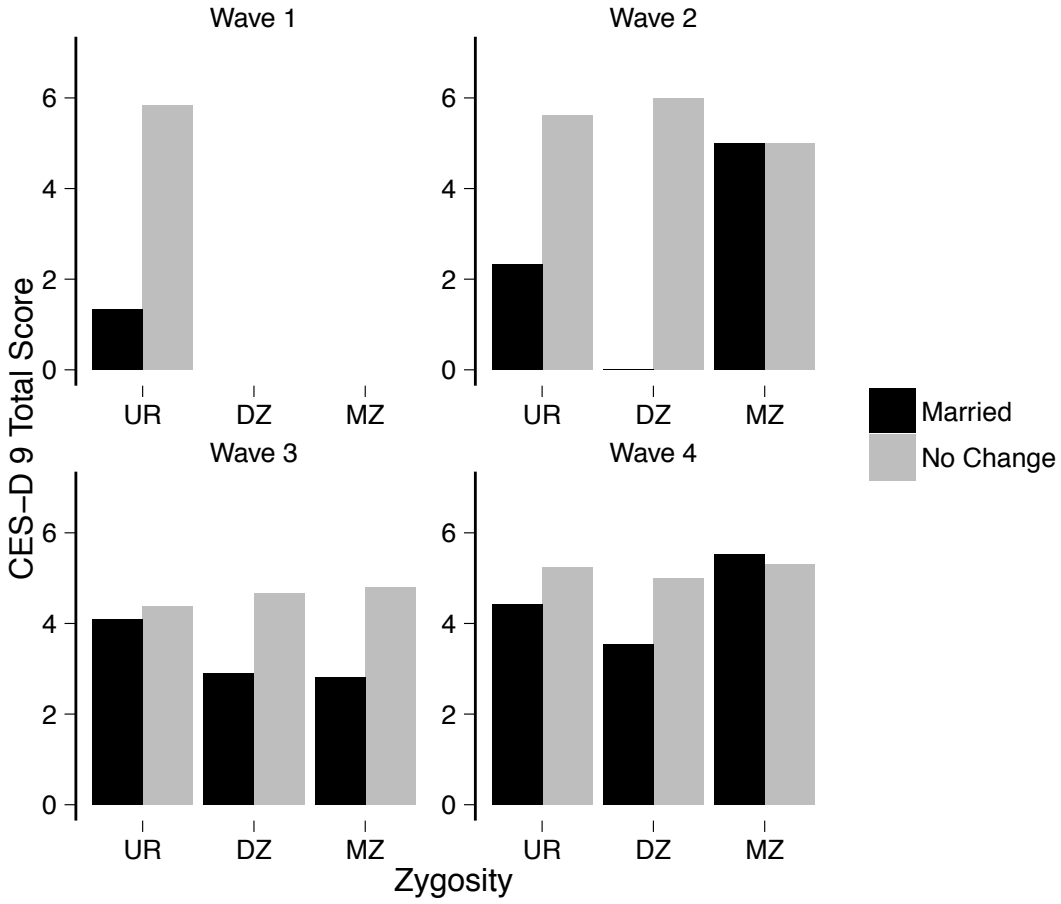

Figure 8. Bar plots of differences in mean total CES-D-9 scores at each wave between pairs of MZ and DZ male twins discordant for entering marriage. Pair differences were compared to the mean differences in unrelated groups discordant for entering marriage.

In the female twins discordant for entering marriage (Figure 9), the differences in Waves 1 and 2 are not interpretable, as the plots only consist of 1 pair of MZ twins in each wave. The differences in Wave 3 suggest that genetic selection processes mediated the correlation between entering marriage and CES-D-9 scores. The slight difference between discordant pairs of MZ twins, however, is inconsistent with the social causation hypothesis. Additionally, at Wave 4, the difference neither supports the selection hypothesis or the social causation hypothesis, as there was no difference in

CES-D-9 scores between twins discordant for entering marriage or between the means of the unrelated groups.

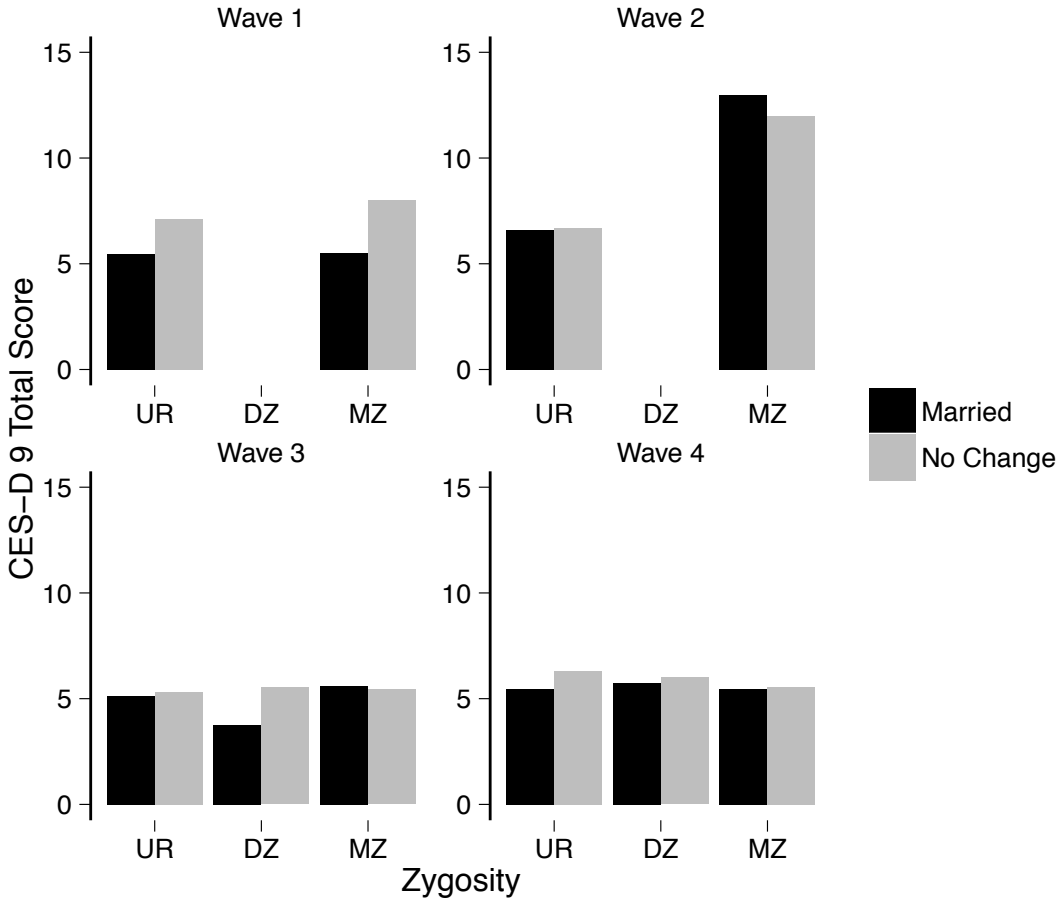

Figure 9. Bar plots of differences in mean total CES-D-9 scores at each wave between pairs of MZ and DZ female twins discordant for entering marriage. Pair differences were compared to the mean differences in unrelated groups discordant for entering marriage.

The final step in the twin-level analysis consisted of fitting longitudinal quasi-causal models to test the short-term effect of entering marriage on CES-D-9 scores in male twins and the long-term effect of entering marriage on CES-D-9 scores in female twins. Table 8 presents the model fitting results. All phenotypic variables in the model

were regressed on twins' age so that the estimates were age-adjusted to compensate for the effects of individual growth in depression associated with age.

Table 8

Longitudinal quasi-causal model fitting results in Add Health

Male Twins	χ^2	df	Scaling correction factor	Model Comparison	LR	S-B LR test	RMSEA	CFI
1. Baseline Bivariate ACE model	125.22	143	0.91				0.00	0.99
2. Short-term genetic (A) regression set to zero	132.73	146	0.91	1	6.77	0.079	0.00	1.00
3. All A parameters set to zero	139.38	152	0.90	2	7.46	0.281	0.00	1.00
4. Short-term quasi-causal (E) regression set to zero	139.35	153	0.90	3	0.08	0.776	0.00	1.00
5. Short-term common cnvironmental (C) regression set to zero	145.11	154	0.90	4	8.90	0.003	0.00	1.00
Female Twins								
1. Baseline Bivariate ACE model	136.14	140	0.92				0.00	1.00
2. Long-term genetic (A) regression and Wave 3 A variance set to zero	136.56	142	0.92	1	0.53	0.767	0.00	1.00
3. Long-term common environmental (C) regression set to zero	136.92	143	0.92	2	0.08	0.779	0.00	1.00
4. Long-term quasi-causal (E) regression set to zero	138.20	144	0.92	3	1.45	0.484	0.00	1.00

Note. Baseline model included all genetic, common environmental, and nonshared environmental regression estimates for the short-term effects constrained to be the same across all waves (men only). Only the genetic, common environmental, and nonshared environmental long-term effects were tested in female twins. The scaling correction factor is a correction that better approximates chi-square under conditions of multivariate nonnormality (Satorra & Bentler, 2001). -2LL = -2 log-likelihood; LR = likelihood ratio; S-B = Satorra-Bentler likelihood ratio difference for nested model comparisons; RMSEA = Root Mean Square Error of Approximation; CFI = Comparative Fit Index.

In the male twins, the genetic regression effects (Model 2) and all other genetic parameters (Model 3) could be set to zero without a significant reduction in model fit. As the genetic variance estimates were negative, various starting values and constraints (e.g., lower bound constraint of zero) were tested that led to the conclusion that all genetic parameters could be safely dropped from the model. Next, the quasi-causal

effect (Model 4) could be set to zero without a significant reduction in model fit. The nonsignificant quasi-causal effect does not support conclusions regarding the causal influence of marriage. The common environmental effect, however, could not be removed from the model without a significant loss of model fit (Model 5). Environmental factors common to both twins appeared to mediate the entire correlation between entering marriage and log CES-D-9 scores (see Model 4 parameter estimates in Table 9). As this analysis was based on a larger sample of male twins (n_{MZ} = 121, n_{DZ} = 141), the findings here are regarded to be more trustworthy than the co-twin-control analysis results presented above.

In the female twins, a different set of models was tested than in the male twins, as it seemed likely that both the selection and quasi-causal parameters of interest would not be correlated with CES-D-9 scores. As suspected, all genetic, common environmental, and quasi-causal long-term effects were not significantly different from zero (i.e., the regression coefficients could be set to zero without a significant reduction in model fit). Thus, based on this longitudinal quasi-causal twin analysis, entering marriage seems to be uncorrelated with CES-D-9 scores in female twins, as indicated by the large standard errors relative to the parameter estimates in the final model (Model 4, Table 9).

Table 9

WLSM parameter estimates from the baseline model and best fitting longitudinal quasi-causal model for male and female twins in Add Health

	WLSM estimates							
	Male				Female			
	Model 1		Model 4		Model 1		Model 4	
Parameter	Est.	*SE*	Est.	*SE*	Est.	*SE*	Est.	*SE*
σ^2_{A3}	-0.16	0.11	-	-	-0.06	0.10	-	-
σ^2_{A4}	-0.13	0.10	-	-	0.16	0.20	0.15	0.19
σ^2_{C3}	0.20	0.10	0.10	0.04	0.11	0.09	0.07	0.04
σ^2_{C4}	0.14	0.09	0.08	0.05	0.01	0.11	0.01	0.10
σ^2_{E3}	0.11	0.06	0.05	0.02	0.13	0.07	0.11	0.06
σ^2_{E4}	0.12	0.08	0.10	0.07	0.14	0.14	0.13	0.12
A_{Reg} - short-term	-0.23	0.40	-	-	-0.03	0.32	0.15	0.24
C_{Reg} - short-term	-0.19	0.25	-0.29	0.13	0.15	0.28	0.20	0.27
E_{Reg} - short-term	-0.30	0.25	-	-	-0.28	0.25	-0.18	0.20
A_{Reg} - long-term	-	-	-	-	-0.12	1.50	-	-
C_{Reg} - long-term	-	-	-	-	0.08	0.64	-	-
E_{Reg} - long-term	-	-	-	-	-0.34	0.42	-	-
σ^2_{eD1}	0.44	0.03	0.43	0.03	0.45	0.04	0.45	0.04
σ^2_{eD2}	0.31	0.02	0.30	0.02	0.33	0.03	0.33	0.05
σ^2_{eD3}	0.44	0.04	0.42	0.03	0.49	0.05	0.50	0.05
σ^2_{eD4}	0.39	0.05	0.35	0.03	0.44	0.06	0.44	0.05

Note. A = additive genetic; C = common environment; E = nonshared environment; D = log total CES-D scale score; e = residual variance; Reg = regression coefficient; 1-4 = wave of measurement. Short-term refers to the effects of entering marriage up to 6 years later. Long-term refers to the effects of entering marriage up to 11 years later.

Entering Marriage: Late Adulthood

Individual-level Analysis

A description of the male and female participants' CES-D means in the SATSA was the initial analysis used to answer the first research question (*Does the trajectory of depression increase in later adulthood?*). Figure 10 presents the longitudinal trajectory plots of the male twins' CES-D scores across all five waves, with the wave means overlaid in red.

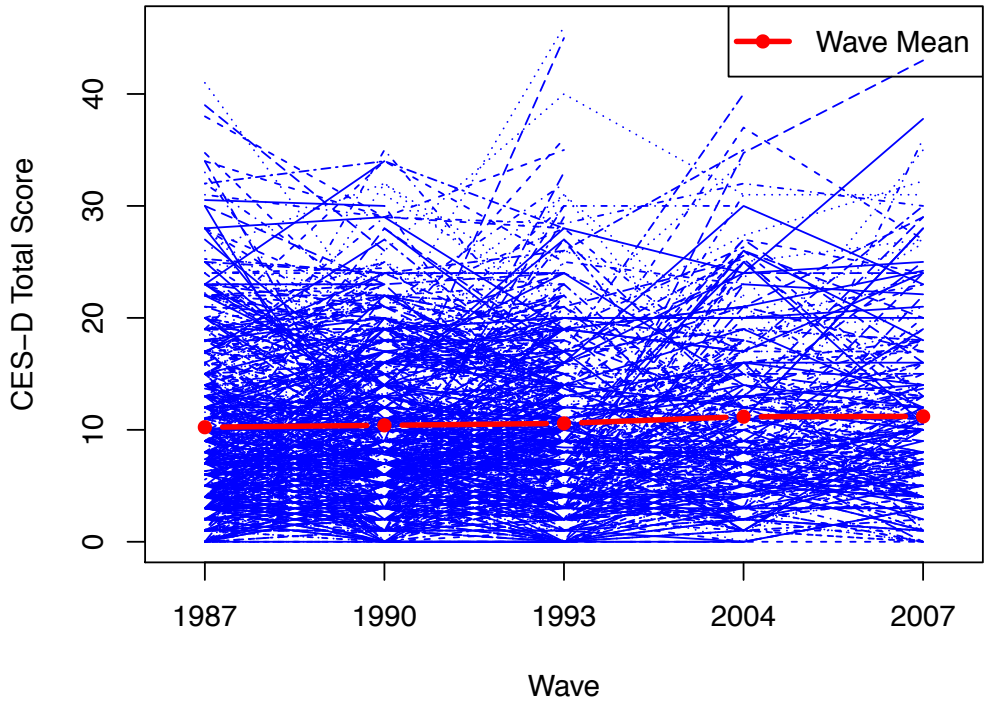

Figure 10. Longitudinal trajectory plots of male participants' total CES-D scores across wave in SATSA. The bolded red line represents the mean total CES-D scores.

As in the young adult men in the Add Health sample shown above, the male participants' scores oscillated over the measurement, which suggests an auto-

regressive pattern of association over time rather than linear growth. The male participants' scores were subsequently plotted by age (Figure 11) to take into account age heterogeneity within each wave of measurement. While there was a slight tendency of the overall mean to increase (the red line) after about age 60, the oscillatory pattern of the participants' scores (the blue lines) is consistent with an auto-regressive pattern of change as in Figure 10.

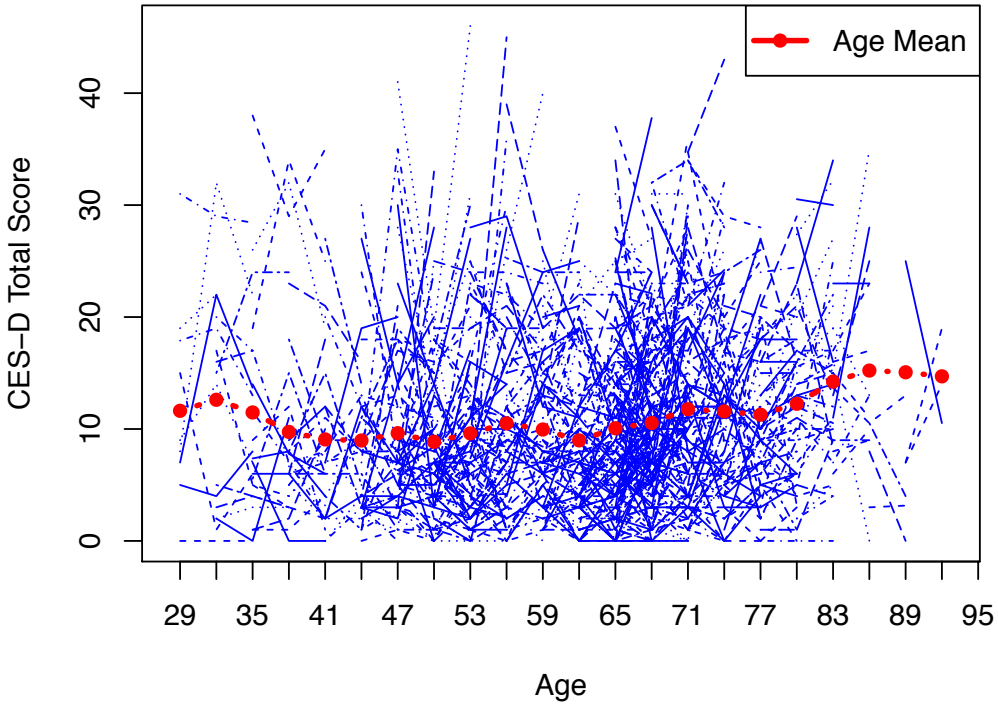

Figure 11. Longitudinal trajectory plots of male participants' total CES-D scores strung out by age in SATSA. Twenty-four age bins were created in 3-year intervals from age 29-101. The bolded dotted red line represents the mean total CES-D score for each age bin.

Figure 12 presents the longitudinal trajectory plots of the female participants' CES-D scores across the five waves (blue lines), with wave means overlaid (red line).

As in the male twins, when viewed by wave, participants' CES-D scores oscillated over time following an auto-regressive pattern.

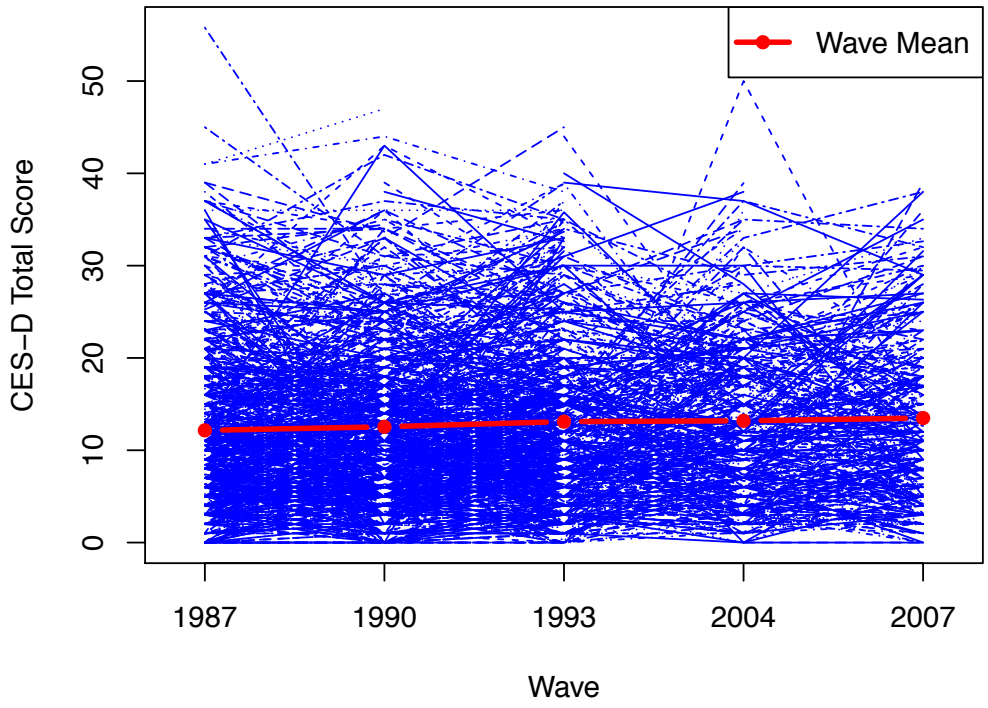

Figure 12. Longitudinal trajectory plots of female participants' total CES-D scores across wave in SATSA. The bolded red line represents the mean total CES-D score.

When the female twins' CES-D scores were plotted by age (Figure 13), mean CES-D scores still appeared to follow an auto-regressive pattern (blue lines) even though the overall mean slightly increased over the second half of the lifespan, starting to rise in the mid-40s.

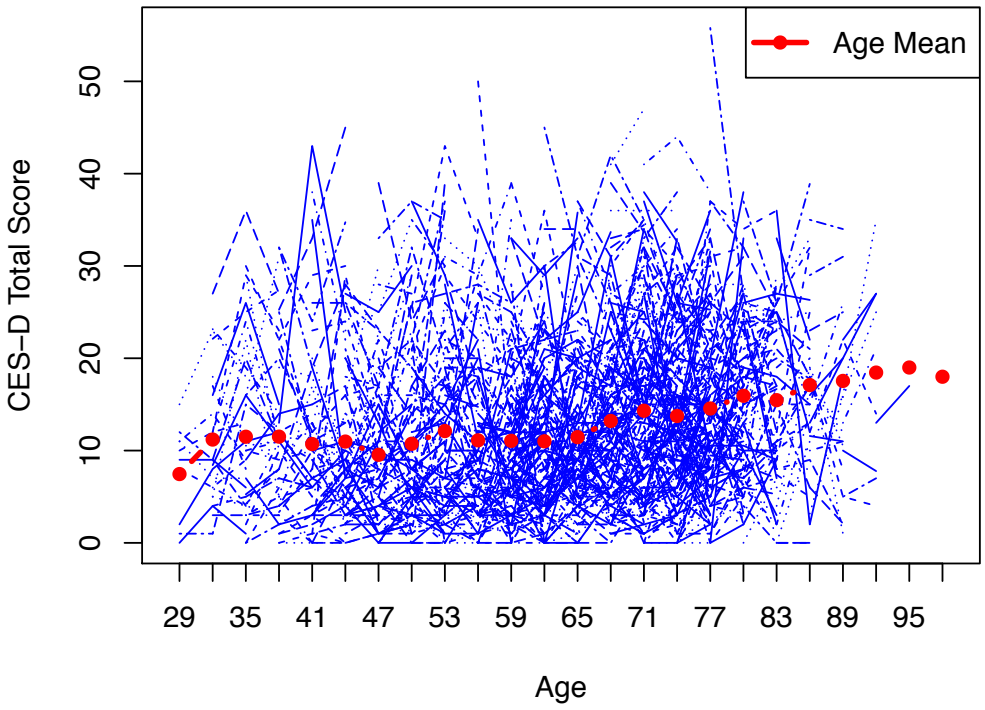

Figure 13. Longitudinal trajectory plots of female participants' total CES-D scores strung out by age in SATSA. Twenty-four age bins were created in 3-year intervals from age 29-101. The bolded dotted red line represents the mean total CES-D score for each age bin.

Next, the CES-D means of those who entered marriage were compared to the CES-D means of those who did not enter marriage (Table 10). The expectation was that the CES-D scores of those who entered marriage would be lower than the CES-D scores of those who did not enter marriage at each wave. The scores were divided into 4 age groups to also examine potential age differences in the effects of entering marriage on CES-D scores.

Table 10

Descriptive CES-D results for SATSA participants who entered marriage and participants who did not enter marriage, by gender and age group

	Men						Women					
	No Marital Status Change			Entered Marriage			No Marital Status Change			Entered Marriage		
	n	*M*	*SD*	n	*M*	*SD*	n	*M*	*SD*	n	*M*	*SD*
1987												
< 60 years	158	8.43	8.14	49	9.71	6.73	218	8.84	7.31	60	9.32	6.76
60 to 69 years	143	9.19	6.89	14	8.29	4.71	155	12.30	9.70	11	11.99	6.28
70 to 79 years	101	10.15	6.96	7	16.38	4.79	171	12.83	9.10	9	13.01	11.39
≥ 80 years	24	15.45	8.89	2.00	6.00	7.07	47	15.35	8.96	1	6.00	NA
1990												
< 60 years	146	8.50	6.66	29	8.84	8.08	192	9.57	8.11	33	10.46	8.17
60 to 69 years	117	8.53	5.72	15	9.41	7.84	153	10.94	7.16	11	14.00	6.72
70 to 79 years	116	11.01	7.10	6	8.33	5.89	167	14.39	9.27	8	12.25	6.98
≥ 80 years	25	12.94	8.02	2	22.50	7.78	51	14.07	8.16	0	-	-
1993												
< 60 years	140	9.25	8.64	28	9.01	5.77	178	9.67	8.10	28	11.42	9.48
60 to 69 years	104	7.88	6.46	9	12.56	5.75	124	11.90	8.31	14	13.63	10.63
70 to 79 years	126	10.76	7.59	8	10.50	5.26	159	14.34	8.61	11	17.18	7.17
≥ 80 years	30	11.73	6.77	8	16.22	12.68	72	15.05	9.78	2	10.5	9.19
2004												
< 60 years	34	8.97	6.18	3	8.33	1.53	66	11.35	9.14	1	26.00	NA
60 to 69 years	77	9.78	7.81	7	11.29	8.54	88	9.21	7.54	5	11.20	10.76
70 to 79 years	56	8.37	6.73	4	11.64	4.50	81	13.43	7.73	2	4.00	2.83
≥ 80 years	40	11.21	7.18	5	12.47	9.88	50	18.35	7.94	4	16.00	12.83
2007												
< 60 years	24	10.38	8.23	0	-	-	36	13.37	11.99	2	8.63	4.76
60 to 69 years	49	9.28	7.43	4	9.84	7.55	72	8.98	7.68	3	9.00	12.17
70 to 79 years	63	10.80	8.01	1	13	NA	77	12.47	7.17	5	4.81	4.2
≥ 80 years	46	12.24	8.75	1	13.68	NA	69	16.72	8.07	3	21.67	14.98

For the men, there was not a strong pattern of differences between means to support the conclusion that entering marriage correlated with lower CES-D scores. Although CES-D scores tended to rise for all age groups across the 20-year measurement window, men who entered marriage did not consistently report lower CES-D scores than those who did not enter marriage. The sample sizes of those who entered marriage at each wave were small in all age groups, so the mean comparisons may not be trustworthy.

For the women, a similar pattern emerged as the men (Table 10). Younger women entered marriage more often than older women (particularly in 1987, 1990, and 1993) but reported higher CES-D scores than those who did not enter marriage. The result contrasts with expectations based on prior research, possibly because they were largely compared to women already in stable marriages. These differences also contrast with previous findings that women who remarried reported subsequent increased happiness (Gentry & Shulman, 1988). Given the small sample sizes and large standard deviations of all groups, the group mean estimates are considered unreliable. Overall, the findings from the individual-level analyses do not suggest that entering marriage negatively correlates with total CES-D score.

Next, LGC models were fit to the data to answer the second research question (*Above and beyond the tendency for people to increase in depression as they age, does entering marriage predict subsequent decreases in depression scores?*). Model fitting results are presented for men and women in Table 11.

Table 11

Latent growth curve model fitting results (SATSA)

Male Twins	-2LL	Parameters	Scaling correction factor	Model Comparison	LR	S-B LR test	AIC	BIC
1. Baseline model (Correlated intercept and slope with unstructured correlations among measured variables)	-2461.64	55	1.82	-	-	-	5033.28	5278.09
2. Short-term correlations constrained to be equal	-2464.16	52	1.88	1	6.57	0.087	5032.33	5272.29
3. Constrain short-term correlation to be zero	-2464.44	51	1.90	2	0.66	0.417	5030.88	5266.27
4. Constrain long-term correlation to be zero	-2464.42	50	1.92	2	-0.04	NA	5028.83	5259.57
Female Twins								
1. Baseline model (Correlated intercept and slope with unstructured correlations among measured variables)	-2946.99	55	2.05	-	-	-	6003.98	6277.06
2. Short-term correlations constrained to be equal	-2948.77	52	2.11	1	3.54	0.316	6001.54	6259.72
3. Constrain short-term correlation to be zero	-2948.79	51	2.13	2	0.04	0.846	5999.59	6252.80
4. Constrain long-term correlation to be zero	-2949.44	50	2.14	2	0.80	0.372	5998.87	6247.13

Note. The scaling correction factor is a correction that better approximates chi-square under conditions of multivariate nonnormality (Satorra & Bentler, 2001). -2LL = -2 log-likelihood; LR = likelihood ratio; S-B = Satorra-Bentler likelihood ratio difference for nested model comparison; AIC = Akaike Information Criterion; BIC = Bayesian Information Criterion.

For the male twins, Model 1 suggested that the intercept of the log CES-D scores (b_I = 1.33, SE = 0.13, $p < .001$) and slope (b_S = 0.01, SE = 0.002, $p < .001$) estimates were significantly different from zero (Table 12), with significant, albeit minute, variation in the slope estimates (σ^2_I = 1.59, SE = 0.43, $p < .001$; σ^2_S = 0.0001, SE = 0.000, p = .012). As the interval of time between the previous wave and the 1987, 1990, 1993, and 2007 waves was the same (3 years), the short-term effects of entering marriage on log CES-D scores were constrained to be equal to one another (Model 2). There was not a significant reduction in model fit between models 1 and 2; the short-term and long-term effects could also be set to zero without a significant reduction in model fit (Model 3 and Model 4). Identical model fitting results were observed in the female twins. Entering

marriage, it seems, was found to be uncorrelated with CES-D scores in men and women in the SATSA, adjusting for individual-level change in depression over time.

Table 12

Parameter estimates from the baseline LGC models for male and female participants in the SATSA

	MLR estimates			
	Male		Female	
Parameter	Est.	*SE*	Est.	*SE*
μ_I	1.34	0.13	1.44	0.11
μ_S	0.01	0.002	0.01	0.001
σ^2_I	1.56	0.43	1.72	0.42
σ^2_S	0.000	0.00	0.000	0.00
σ_{IS}	-0.02	0.01	-0.02	0.01
$\sigma_{M,D}$ - short-term	0.003	0.003	-0.001	0.004
$\sigma_{M,D}$ - long-term	0.000	0.01	-0.01	0.01
σ^2_{eD1987}	0.32	0.02	0.33	0.03
σ^2_{eD1990}	0.29	0.04	0.28	0.03
σ^2_{eD1993}	0.35	0.04	0.28	0.03
σ^2_{eD2004}	0.24	0.04	0.28	0.04
σ^2_{eD2007}	0.30	0.05	0.26	0.04
σ^2_{eM1987}	0.11	0.01	0.08	0.01
σ^2_{eM1990}	0.09	0.01	0.06	0.01
σ^2_{eM1993}	0.08	0.01	0.07	0.01
σ^2_{eM2004}	0.04	0.01	0.02	0.01
σ^2_{eM2007}	0.02	0.01	0.03	0.01

Note. μ = mean; σ^2 = variance; σ = covariance; I = intercept; S = slope; M = entered marriage; D = log CES-D score. Short-term refers to the effects of entering marriage up to 3 years later. Long-term refers to the effects of entering marriage up to 11 years later.

Twin-level Analysis

Co-twin-control analyses were used as the first step toward answering the third research question (*Above and beyond genetic and environmental selection factors, does entering marriage predict subsequent decreases in depression scores?*). Table 13 presents a contingency table of the total number of twin pairs concordant for marital status (either both entered marriage or both did not enter marriage) and discordant for entering marriage, separated by zygosity and gender groups (monozygotic male, monozygotic female, dizygotic male, and dizygotic female) at each wave. For all male twins (MZM and DZM), there were 31 twins discordant for marital status in 1987, 30 discordant pairs in 1990, 28 discordant pairs in 1993, 11 discordant pairs in 2004, and only 3 discordant pairs (DZM only) in 2007. For all female twins (MZF and DZF), there were 49 discordant pairs in 1987, 23 discordant pairs in 1990, 24 discordant pairs in 1993, 9 discordant pairs in 2004, and 9 discordant pairs in 2007. Overall, the sample sizes of discordant twin pairs at each wave were modest. Pair differences must be interpreted with caution.

Table 13

Contingency tables of male and female twin pairs concordant and discordant for entering marriage status in the SATSA

		1987							
		Twin 1							
		MZM		MZF		DZM		DZF	
		No Δ MS	Married	No Δ MS	Married	No Δ MS	Married	No Δ MS	Married
Twin 2	No Δ MS	72	-	91	-	94	-	181	-
	Married	9	1	15	2	22	5	34	3
		1990							
		Twin 1							
		MZM		MZF		DZM		DZF	
		No Δ MS	Married	No Δ MS	Married	No Δ MS	Married	No Δ MS	Married
Twin 2	No Δ MS	63	-	92	-	93	-	186	-
	Married	13	0	9	2	17	1	14	5
		1993							
		Twin 1							
		MZM		MZF		DZM		DZF	
		No Δ MS	Married	No Δ MS	Married	No Δ MS	Married	No Δ MS	Married
Twin 2	No Δ MS	68	-	85	-	99	-	172	-
	Married	10	0	4	1	18	1	20	6
		2004							
		Twin 1							
		MZM		MZF		DZM		DZF	
		No Δ MS	Married	No Δ MS	Married	No Δ MS	Married	No Δ MS	Married
Twin 2	No Δ MS	36	-	59	-	67	-	113	-
	Married	3	0	2	0	8	0	7	0
		2007							
		Twin 1							
		MZM		MZF		DZM		DZF	
		No Δ MS	Married	No Δ MS	Married	No Δ MS	Married	No Δ MS	Married
Twin 2	No Δ MS	33	-	43	-	64	-	101	-
	Married	0	1	3	0	3	0	6	0

The co-twin-control analysis compared the CES-D means of the twins who entered marriage with the means of their co-twins who did not enter marriage (Table 14). Positive scores indicate that twins who entered marriage had a *lower* CES-D score than their co-twins who did not enter marriage.

Table 14

Co-twin-control analysis of CES-D score differences between male and female twins discordant for entering marriage in the SATSA

No Marital Status Change vs. Entered Marriage			
	Pairs	Mean Intrapair Difference	*t*
1987			
Male	31	1.35	0.71
Female	49	1.12	0.64
1990			
Male	30	0.68	0.32
Female	23	-0.22	-0.08
1993			
Male	28	2.22	1.22
Female	24	1.81	0.70
2004			
Male	11	-4.29	-1.43
Female	9	1.18	0.22
2007			
Male	3	-5.33	-1.47
Female	9	6.06	2.09

The co-twin-control analysis results are consistent with the individual-level analysis results. Although male and female twins who entered marriage typically were found to be less depressed than their co-twins who did not enter marriage, male twins who entered marriage had *higher* CES-D scores than their co-twins in 2004 and 2007. None of the pair differences in the male twins discordant for entering marriage were statistically significant ($p < .05$). Female twins who entered marriage were found to be less depressed than their co-twin controls who did not enter marriage (except in 1990), but the only marginal effect was between discordant twins in 2007 ($p = .060$). Female twins who entered marriage reported 6.06 fewer CES-D units than their co-twins who did not enter marriage.

As in the Add Health data, a graphical examination of the differences between discordant MZ and DZ twin pairs can provide a better understanding of the within-family effects of entering marriage on depression, as well as potential genetic and environmental selection effects underlying their association. Figure 14 presents the CES-D means of MZ and DZ male twins discordant for marital status. The means of unrelated groups discordant for entering marriage are provided as a reference. Overall, the results do not reveal clear underlying selection processes or causal mechanisms, as the direction of the pair differences illustrate that twins and individuals who entered marriage were not consistently less depressed than their counterparts who did not enter marriage. It is important to keep in mind that sample sizes in the MZ and DZ groups were small. The differences between discordant MZ twins were smaller than the differences between discordant DZ twins (in 1990, 1993, and 2004), which suggest

genetic selection as one possible mediator between entering marriage and CES-D scores at these time points. The lack of differences between unrelated pairs (compared to the MZ and DZ discordant pairs) suggests that selection processes may have suppressed the effect of entering marriage on CES-D scores. The small sample sizes used to calculate the means, however, make these estimates unreliable. For the 2007 measurement, the results could not be interpreted, as there were no discordant pairs of MZ male twins.

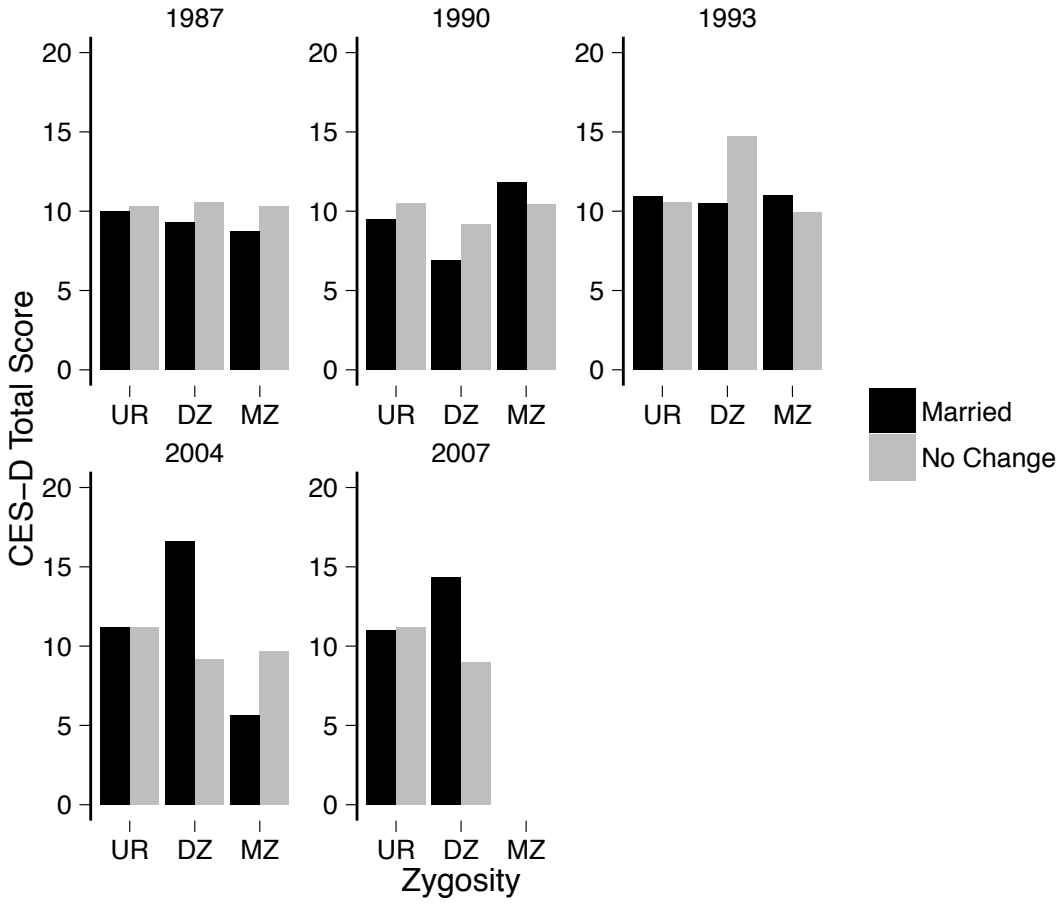

Figure 14. Bar plots of differences in mean total CES-D scores at each wave between pairs of MZ and DZ male twins discordant for entering marriage in the SATSA. Pair differences were compared to the mean differences in unrelated groups discordant for entering marriage.

Figure 15 presents the mean CES-D differences between MZ and DZ female twins discordant for entering marriage. For the short-term effects (1987, 1990, 1993, and 2007), the results are generally consistent with partial confounding. In waves 1987, 1993, and 2007, the differences in CES-D scores in the MZ twins discordant for entering marriage is consistent with the causal interpretation that entering marriage predicts lower CES-D scores.. The greater similarity in CES-D scores between discordant DZ twins than discordant MZ pairs in 1987, 1990, and 1993 suggests that genetic selection

effects *protect* co-twins who did not enter marriage from depression rather than put them at risk. As with the discordant male twins, the inconsistency in the direction of the effects between discordant twin pairs made it difficult to interpret whether entering marriage predicts decreases in CES-D scores at follow-up. Given the lack of consistency in the twin analysis and the null findings in the LGC analyses, longitudinal quasi-causal models were not fit to the data.

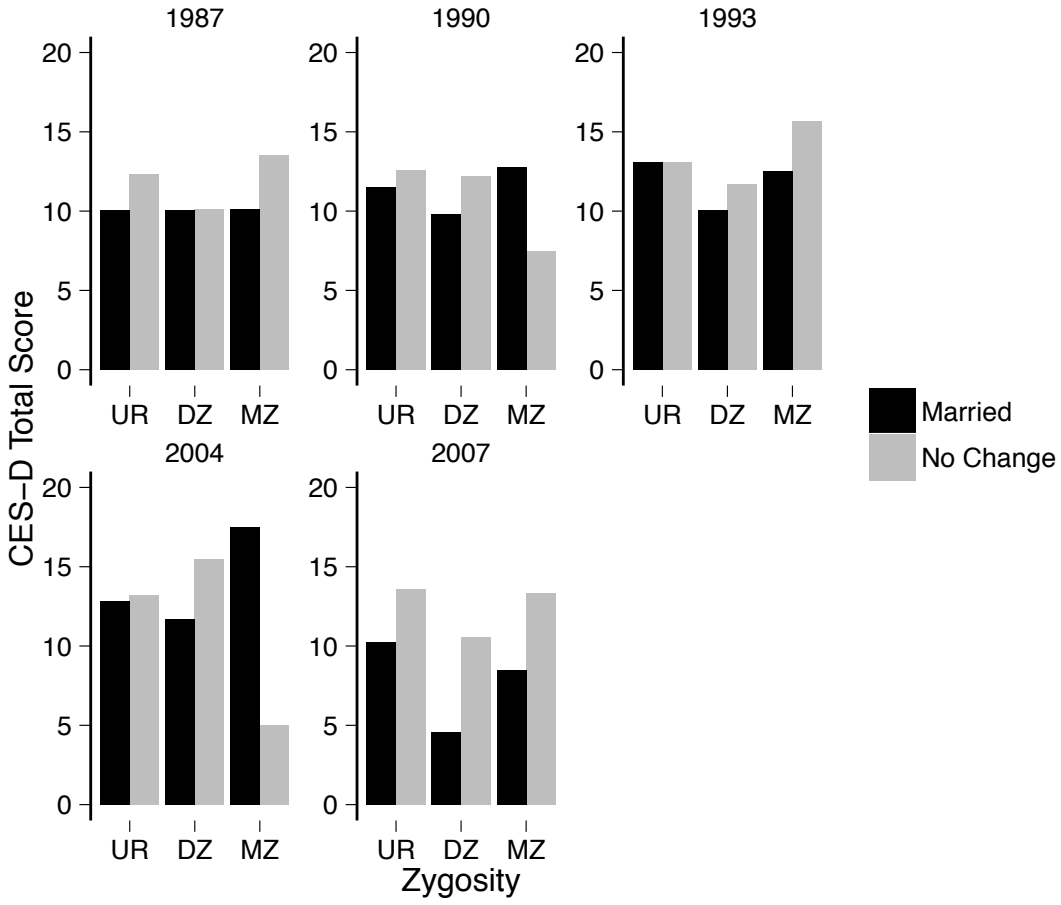

Figure 15. Bar plots of differences in mean total CES-D scores at each wave between pairs of MZ and DZ female twins discordant for entering marriage in the SATSA. Pair differences were compared to the mean differences in unrelated groups discordant for entering marriage.

Discussion

The purpose of the present study was to use two longitudinal twin samples to advance knowledge on the benefits of entering first marriage for lowering depression across the lifespan. Virtually all research on the marriage benefit is limited by nonexperimental designs, leaving countless third variables (e.g., genetic and environmental selection factors) as possible alternative hypotheses to the social causation hypothesis. The quasi-experimental properties of longitudinal twin studies on marriage can help elucidate causality where experimental control is not possible (Turkheimer & Harden, 2013). The use of twins permits the statistical control of genetic and common environmental selection effects while testing whether the effects of entering marriage on depression remains statistically significant within identical twin pairs. Longitudinal data also has the advantage of testing lagged effects (i.e., entering marriage at time *t* negatively predicts depression at time *t*+1). Finally, the effect of marriage on depression was explored in a sample of young adults and in a sample of older adults to clarify whether the benefits of marriage differ across the lifespan.

The findings from the present study add to the existing base of marital research using longitudinal twin samples. Whereas others have found quasi-causal effects of marriage on antisocial behavior (Burt et al., 2010; Horn et al., 2013) and alcohol consumption (Prescott & Kendler, 2001), the results in the present study largely do not support the social causation hypothesis. There was a significant negative short-term phenotypic effect of entering marriage on depression in the young male participants. The twin analyses, unfortunately, did not help to clarify whether potential selection

processes or underlying causal mechanisms accounted for the short-term effect. One finding from the co-twin-control analysis suggested that the twin who married might be less depressed than the twin who did not marry. This effect, although, was only found at Wave 3; the difference in depression scores between discordant twins at Wave 4 was smaller and not statistically different from zero. In contrast, the findings from the longitudinal quasi-causal models supported the social selection hypothesis. There was a significant common environmental selection effect, as the pathway accounted for the entire association between marrying and depression scores over the short-term. Regardless of the underlying mechanism, the significant short-term effect suggests that there may be advantages to marrying earlier in adulthood compared to later in life. Younger people have the benefit of financial and emotional support from parents and other family members, which may be a key mediator for why young men who are happier select into marriage at higher rates than unhappier young men.

The significant short-term reduction in depressive symptoms found in young men stands in contrast with the nonsignificant short-term effect – and significant long-term increase in depressive symptoms – found in young women. The nonsignificant short-term effect in women may indicate that marriage has even shorter time-limited benefits than was captured in the Add Health sample. In samples with shorter intervals between measurements, the positive effects of marrying on depression were observed in women (Frech & Williams, 2007; Mastekaasa, 1992). In the current study, the “short-term” interval could have ranged between 0 and 6 years, so simply *being* married might not have had lasting effects on depression in women. It also has been documented that

marital status matters more for men but *marital quality* matters more for women (Kiecolt-Glaser & Newton, 2001). Marital status, unfortunately, is an inadequate index of the contents of marriage and so might not fully capture the positive benefits of marriage during the early years (Beam, 2008). The protracted effects of marrying may not be beneficial for women, as indicated by the positive long-term phenotypic effect of entering marriage on CES-D scores. The transition to parenthood, as well as drops in marital quality during the early parenting years, however, were not adjusted for in the models and may be one reason for the unexpected long-term result.

There may be generational differences in the effects of marriage on depression too, as only the correlation between entering marriage and depression was found in young men. Marriage typically has been regarded as a stabilizing force for men that put an end to behaviors like carousing, promiscuity, social isolation, and smoking, as a few examples (Waite & Gallagher, 2000). The potential for marriage to have a positive effect on negative behaviors may be restricted to critical periods over the lifespan, with early adulthood being one of these periods. Men from older generations who enter marriage may already possess indelible behaviors or have already abandoned risky behaviors from their youth, lessening the impact of marriage on health.

Cultural differences might have contributed to marital benefits for depressive symptoms in the Americans but not in the Swedish. Although generational differences could not be controlled for in the models, the data were collected during relatively similar measurement windows – the 1990s through the 2000s – which make the historical climates reasonably comparable between the two samples. Marriage might have been

beneficial to American men but not Swedish men for the reason that marriage confers more rights, privileges, and benefits in the United States than in Sweden. For example, married Americans receive tax benefits that promote economic assurance, insurance benefits that support physical and mental health, and financial support for children. In Sweden, the tax structure and social security benefits extend to all citizens regardless of marital status, so marriage does not confer advantages over other statuses. It is not clear why these benefits did not extend to young American women, however, as marriage has been considered to provide economic benefits to both men and women (Waite & Gallagher, 2000).

The current findings should be interpreted in view of three limitations. First, the intervals between measurements in Add Health and the SATSA may have been too broad to capture the effects of marriage over time. Others have found significant effects of marrying on depression, but over shorter time intervals (Kim & McKenry, 2002; Mastekaasa, 1992). Second, there were several confounds in the comparison between young American and older Swedish twins, including generational differences, sample size differences (fewer adults marry in late adulthood). Finally, marital status may be too broad an indicator of social support, as it does not capture what goes on in marriage, making marital quality a more important index of benefits associated with marriage (Kiecolt-Glaser & Newton, 2001). A complete understanding of the marriage benefit would require following a sample of twins discordant for being single and engaged-to-be-married before the co-twin marries and after the co-twin marries. Monthly measures of depressive symptoms and relationship functioning in the 6 months prior to marriage

and in the 18 months following marriage would help clarify whether differences in marrying influence depression within families while indexing whether within-person changes in the married co-twin's relationship functioning correlates with changes in depression over time.

The primary strength of the present study is that it offers a lifespan perspective on the effects of entering marriage on depression. Whereas many studies have focused on marriage in early adulthood (Horwitz & White, 1998; Mastekaasa, 2006) or aggregated samples across a large age range (Frech & Williams, 2007; Lamb et al., 2003), the current study used two age-focused samples to investigate whether marital effects on depression differ across the lifespan. The use of longitudinal twin samples also was a strength insofar that genetic and environmental selection factors could be statistically controlled while testing whether the effects of entering marriage on depression were consistent with the social causation hypothesis in real time.

Widowhood, Depression, and Mortality Risk

Although many adults are resilient and recover from spousal loss after brief periods of bereavement (Bonanno et al., 2002), many widowed adults suffer prolonged depressive symptoms (hereafter referred to as “depression” for simplicity). The aim in widowhood research is to determine whether conjugal loss causally influences mental health outcomes (the *social causation hypothesis*), like depression (Clayton et al., 1972) or whether selection processes (the *social selection hypothesis*) underlie observed correlations. In other words, to what extent do unmeasured or unobserved third variables confound the causal interpretation of the emotionally harmful effects of spousal loss? There is a tendency, as an example, for older adults to experience chronic or recurring depressive symptoms as they age (Blazer, 2003), which makes it difficult to determine whether spousal loss influences people’s subsequent symptom levels or whether they were bound to become depressed despite their loss.

Social selection might also occur for genetic and environmental reasons. Some adults may possess qualities (e.g., emotional sensitivity) that make them more responsive to interpersonal loss in general and more prone to depression. Lichtenstein, Gatz, Pedersen, Berg, and McClearn (1996) controlled for genetic and environmental selection by comparing depression scores in twins who were discordant for widowhood and marriage. In a longitudinal sample of older Swedish adults, they showed that becoming widowed might causally influence depression. Co-twin control studies inherently control for unmeasured genetic and environmental selection effects, for the reason that twins share their genes (monozygotic (MZ) twins share all of their genes

and dizygotic (DZ) twins share one-half of their genes, on average) and their common environments (e.g., parental SES, household and family environments). Any differences in depression scores between discordant twin pairs, therefore, cannot be attributed to genetic or common environmental causes. In the case of MZ twins, observed differences between them must be environmental in nature because they are genetically matched.

Co-twin-control studies are a powerful method for exploring the causal influences of widowhood on depression because people cannot be randomly assigned to widowhood. Yet there are few published twin studies to date. Besides Lichtenstein et al.'s study (1996), Osler, McGue, Lund, and Christensen (2008) used a sample of elderly Danish adult twins discordant for widowhood and found within-family differences in the association between widowhood and depression, consistent with the social causation hypothesis.

Aside from co-twin-control studies, longitudinal studies have advanced knowledge on the causal influence widowhood on depression. The recently widowed have been found to experience more severe depression than their still married counterparts, with grief responses returning to baseline levels after 1-2 years (Bonanno et al., 2002; Lund, Caserta, & Dimond, 1986). Less is known about the protracted effects of widowhood on depression. Some studies have highlighted resiliency among the majority of widows and widowers following the death of a spouse (Bonanno, Wortman, & Nesse, 2004). Others, including Lichtenstein and colleagues (1996), have shown that the effects of widowhood on depression persist beyond a 1-2 year window.

Depression scores among the recently widowed were found to be above married controls for up to 3 years after spousal loss in Swedish twins, with nonsignificant differences 6 years post-spousal loss. In a nationally representative sample of Germans, decreases in life satisfaction (a measure highly correlated with depression) following conjugal loss could take up to 8 years to return to pre-loss levels. Finally, in a sample of older widowed Americans, widowers (but not widows) were found to have elevated depression after conjugal loss, but the effect also diminished with the passage of time since spousal loss (Lee & DeMaris, 2007).

The Swedish Adoption/Twin Study of Aging (SATSA) is a unique data source for studying the longitudinal effects of widowhood on depression. While the prospective nature of the SATSA permits approximating when twins became widowed, the family-level data permit statistical control of genetic and environmental selection effects. Lichtenstein et al. (1996) used the initial 4 waves of data (1984, 1987, 1990, and 1993), but additional follow-up measurements have been conducted since publication of their results. The purpose of the present study is to update and extend Lichtenstein et al.'s study in four ways. First, the two additional measurements since 1996 (one in 2004 and a second in 2007) were added to the study. The SATSA contains two cohorts: the old-old and the young-old. The added benefit in the present study is that depressive responses to widowhood can be better studied among the young-old sample of twins, as many of them have had the unfortunate experience of losing a spouse between 1996 and 2007.

Second, Lichtenstein and colleagues (1996) compared widowed twins to their still married co-twins. In this way, depressive responses to widowhood were interpreted relative to the most logical marital group – those *still* married. While a still married comparison group is the most ideal, it eliminates from the analysis widowed twins who had co-twins in other marital statuses (e.g., divorced, already widowed, never married). As a result, information about these twins' widowhood experiences is lost and sample size is reduced in a study paradigm that already suffers from small sample sizes. Inferences about the widowhood experience, however, can be expanded to intergroup differences within other married subgroups as a way to increase sample size. Spousal loss, for example, is a stressful life event that may be broadly compared to other nonmarital stressful events (Holmes & Rahe, 1967). While broadening the comparison group to those "who did not become widowed" introduces potential confounds related to marital status, the cost is offset with the increase in the sample size available for analysis. Thus, a proband approach was used in the present study so that inferences regarding the effects of becoming widowhood could be drawn relative to all controls who did not experience widowhood.

Third, the present study advances Lichtenstein et al.'s (1996) study by testing different short-term (up to 3 years following spousal death) and long-term effects (up to 11 years following spousal death) of becoming widowed. Although Lichtenstein et al. did not find elevated symptoms among the long-term widowed relative to the still married 6 years after spousal loss, widowhood has been found to predict lower life satisfaction up to 8 years after spousal loss (Lucas et al., 2003).

Finally, while Lichtenstein and colleagues (1996) tested differences between co-twins discordant for widowhood and marriage, they did not explore which selection effects – genetic, environmental, or both – mediated the observed correlation between widowhood and depression. Structural equation modeling methods for twins were used in the current study to identify specific selection processes that explain their association.

The present study also extends research on the mechanisms underlying the effects of widowhood on mortality risk (Lichtenstein et al., 1998) by testing whether increases in severity of depression mediate the association. Elsewhere, it has been shown that widowhood increases risk of early mortality (Manzoli et al., 2007), particularly in the period immediately following spousal loss (Lichtenstein et al., 1998). One hypothesis is that depressive reactions observed among bereaved older adults mediate the risk of mortality associated with becoming widowed. The death records for all SATSA participants were available, so it was possible to test this hypothesis.

Gender Differences

Widows and widowers may have different depressive reactions and mortality risks associated with widowhood (Lee & DeMaris, 2007; Lee, Willetts, & Seccombe, 1998; Lichtenstein et al., 1996). When gender differences have been observed, widowers typically have been found to experience greater subsequent depression than widows. Gender differences have not always been observed (Cheung, 2000; Clayton et al., 1972; Lund et al., 1986; Manzoli et al., 2007), with gender comparisons sometimes not even made (Bonanno et al., 2002; Osler et al., 2008; Zisook et al., 1994), possibly for the reason that wives undergo spousal loss more often than husbands. Although the

SATSA consists of a greater number of widows than widowers, gender comparisons were possible in the present study as in Lichtenstein et al.'s (1996) original study.

Present Study

As noted above, the purpose of the present study was to update Lichtenstein et al.'s (1996) SATSA study that included depressive response to becoming widowed. The current study extends their work in 4 ways: (1) using additional data collected since 1996; (2) using a proband approach to compare the effects of widowhood against all nonwidowed relationship experiences; (3) updating the short-term (3-year) effect and testing a new long-term (11-year) effect of widowhood; and (4) specifying the genetic and environmental selection processes underlying the association between widowhood and depression. The same co-twin-control design that was used in Lichtenstein et al. (1996) and modern structural equation modeling approaches were used to address the following research questions:

1. Does the trajectory of depression generally increase across late adulthood?
2. Above and beyond the general propensity for people to become more depressed as they age, does becoming widowed predict short-term and long-term subsequent increases in depression?
3. Above and beyond genetic and environmental selection effects, does becoming widowed predict short-term and long-term increases in subsequent levels of depression?
4. Is widowhood status associated with reduced longevity? If so, does depression mediate the association?

5. Does becoming widowed have different effects on depression and mortality risk between men and women?

Method

The Swedish Adoption/Twin Study of Aging (Finkel & Pedersen, 2004) is a longitudinal gerontological study designed to assess physical and psychological health in a representative sample of middle- and late-aged Swedish adults. The questionnaire-based component of the SATSA began in 1984, with follow-up measures sent to all still living participants every three years until 1993. Two additional follow-up questionnaires were sent to still living participants in 2004 and 2007. Marital status information was collected at 6 waves (1984, 1987, 1990, 1993, 2004, and 2007) and depressive symptom measures were administered at the last 5 waves. The overall number of twin pairs from which the current study sample was drawn is 1,919 twin pairs. For the present study, the number of twin pairs available at each wave for each zygosity group separated by gender is given in Table 1.

Table 1

Number of completes, incomplete, and missing twin pairs from 1987-2007 in the SATSA

	Complete Pairs				
	1987	1990	1993	2004	2007
MZM	54	47	55	22	18
MZF	65	58	49	31	17
DZM	85	77	76	34	24
DZF	139	122	117	50	50
	Incomplete Pairs				
	1987	1990	1993	2004	2007
MZM	99	103	80	51	39
MZF	118	105	108	55	59
DZM	194	176	176	112	101
DZF	259	248	253	167	133
	Missing Pairs				
	1987	1990	1993	2004	2007
MZM	106	109	124	186	202
MZF	147	167	173	244	254
DZM	209	235	236	342	363
DZF	308	336	336	489	523

Note. MZM = monozygotic male; MZF = monozygotic female; DZM = dizygotic male; DZF = dizygotic female. Sample sizes were computed based on depressive symptom and marital status information.

Given the aging focus of the SATSA, mortality was a major cause of attrition across the 20-year measurement window. In 1987, 3.10% (n = 119) of the total sample was deceased, with increases at each follow-up wave: 7.22% (n = 277) in 1990, 10.34% (n = 397) in 1993, 26.73% (n = 1026) in 2004, and 30.41% (n = 1167) in 2007. The age range of the sample in 1984 was 26 years to 93 years. The mean age of the male and female twins at each follow-up wave was as follows (based on complete and incomplete pairs): in 1987, M_{Men} = 60.66 (SD = 13.00) and M_{Women} = 62.12 (SD = 13.91); in 1990, M_{Men} = 62.73 (SD = 12.25) and M_{Women} = 64.22 (SD = 13.54); in 1993, M_{Men} = 63.76 (SD = 12.33) and M_{Women} = 65.83 (SD = 13.40); in 2004, M_{Men} = 69.81 (SD = 10.47) and M_{Women} = 70.10 (SD = 11.75); and in 2007, M_{Men} = 71.47 (SD = 9.79) and M_{Women} = 71.97 (SD = 10.83). The gradual increase in mean ages over time reflects higher retention of younger participants and the attrition of older twins due to death.

Measures

The Center for Epidemiological Studies–Depression (CES-D) scale is a 20-item scale widely used as a screening measure of depressive symptomatology (Radloff, 1977). Participants rate how often they experienced each symptom over the past 7 days on a scale of 0-3, with 0 = *rarely or none of the time (less than 1 day),* 1 = *some or a little of the time (1-2 days),* 2 = *occasionally or a moderate amount of time (3-4 days); and* 3 = *most or all of the time (5-7 days)*. The reliability and validity of the CES-D scale has been well-studied, with 4 facets typically found to comprise the total scale: depressed affect, lack of well-being, somatic and retarded activity, and interpersonal difficulties (Hertzog et al., 1990; Shafer, 2006). All 20-items are presented in Appendix

A. The log of the sum of all 20-items was used, as the distribution of raw total scores was positively skewed. The histograms of the log scores for the male and female twins are presented in Appendix B. The test-retest reliability across all five measurements is substantial for both men (McDonald's ω = .87) and women (McDonald's ω = .85) (McDonald, 1999).

The *becoming widowed* variable is a binary coded variable that was operationalized as a change from being married at time *t* (e.g., 1987) to widowhood status at time *t*+1 (e.g., 1990). Participants who became widowed between any two waves were coded 1. All participants who reported the same marital status or were missing at an adjacent wave of measurement were coded 0. Participants were only given the "becoming widowed" code once during the study period, regardless of whether they remarried and became widowed a second time during the study. (For example, one scenario might be a participant who became widowed between 1987 and 1990, was remarried by 1993, and became widowed a second time between 2004 and 2007.) The purpose of this coding scheme was to capture the negative effects of people's first known widowhood experience on depression.

Of the 1,494 male twins, 1.41% became widowed between 1984 and 1987, 0.60% between 1987 and 1990, 1.07% between 1990 and 1993, 1.47% between 1993 and 2004, and 0.80% between 2004 and 2007. Of the 2,072 female twins, 3.28% became widowed between 1984 and 1987, 3.09% between 1987 and 1990, 2.90% between 1990 and 1993, 2.75% between 1993 and 2004, and 1.54% between 2004 and 2007.

Data Analysis

The analytic approach used for the study in this chapter was nearly identical to the approach used in the previous chapter. For both the individual-level analyses and the twin-level analyses, the reader is referred to the *Data Analysis* section in the previous chapter (pages 23-30). As a review, individual-level analyses and twin-level analyses were conducted. The individual-level analyses consisted of an examination of the longitudinal trajectory plots of the participants' mean CES-D scores, a comparison of the CES-D means of men and women who became widowed to men and women who did not become widowed, and latent growth curve (LGC) analyses (see Figure 1). These analyses addressed the first research question (*Does the trajectory of depression generally increase across late adulthood?*) and the second research question (*Above and beyond the general propensity for people to become more depressed as they age, does becoming widowed predict short-term and long-term increases in subsequent levels of depression?*)

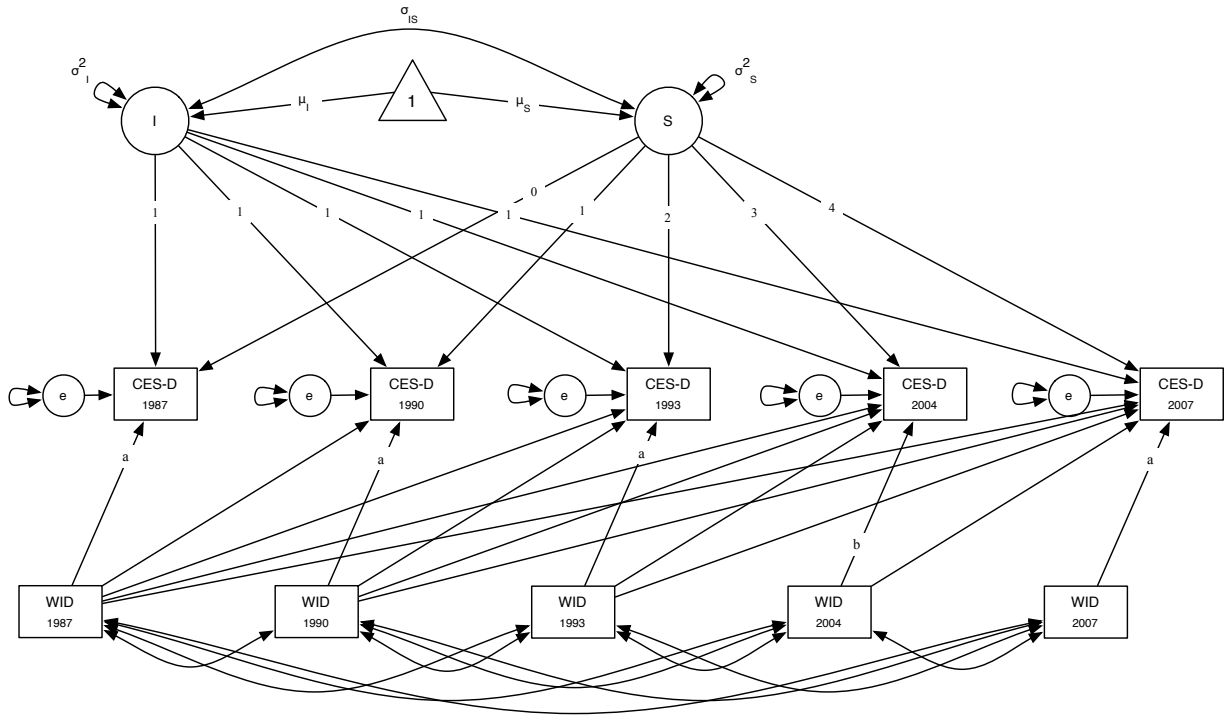

Figure 1. Linear Growth Curve Model. Squares represent observed variables, circles represent unobserved latent constructs, and the triangle represents the mean structure of the latent variables. WID = became widowed; CES-D = log total Center for Epidemiological Studies–Depression scale score; I = intercept; S = slope.

The twin-level analyses consisted of co-twin control analyses and multivariate longitudinal twin analyses (here again referred to as "longitudinal quasi-causal models"; see Figure 2). These analyses addressed the third research question (*Above and beyond genetic and environmental selection effects, does becoming widowed predict short-term and long-term increases in subsequent levels of depression?*) and the fourth research question (*Is widowhood status associated with reduced longevity? If so, does depression mediate the association?*). The fifth research question (*Does becoming widowed have different effects on depression and mortality risk for men and women?*) regarding gender differences was addressed in both analyses.

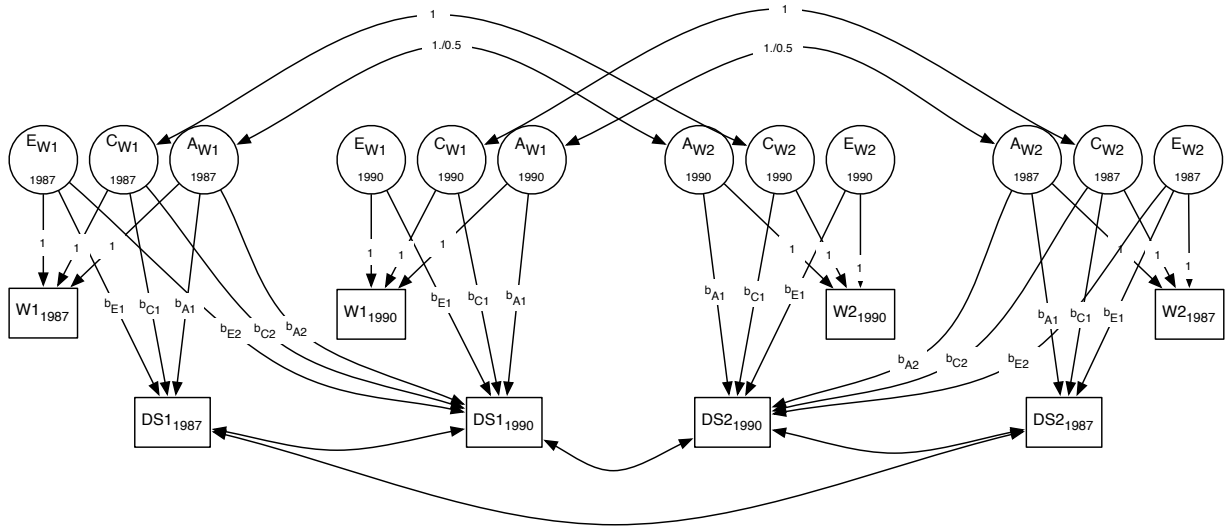

Figure 2. Longitudinal quasi-causal model. Only two waves of measurement (1987 and 1990) are shown for simplicity. Actual models were estimated using all 5 measurements. Squares represent observed variables, and circles represent unobserved latent constructs. W = became widowed variable; DS = log total Center for Epidemiological Studies–Depression scale score; A = additive genetic factor; C = common environmental factor; E = nonshared environmental factor; b = regression coefficient (both short-term and long-term effects). Variable names are appended with "1" or "2" to correspond with twin 1's scores and twin 2's scores.

Results

Individual-level analyses

A description of the male and female participants' CES-D means in the SATSA was the initial analysis used to answer the first research question (*Does the trajectory of depression generally increase across late adulthood?*) and consisted of a series of longitudinal trajectory plots (by wave and age). The findings are identical to those presented in the previous chapter and can be found in the *Results* section of the first chapter (pp. 49-52). As a brief review, the longitudinal trajectory plots of both the male

and female participants tended to follow oscillatory patterns regardless of whether their scores were plotted by wave or age. The trajectories were consistent with an auto-regressive pattern of CES-D scores over time rather than linear growth.

Next, the CES-D means of those who became widowed were compared to the CES-D means of those who did not become widowed (Table 2). The expectation was that the CES-D scores of the widowed would be higher than the CES-D scores of the nonwidowed at each wave. The scores were divided into 4 age groups to examine whether there were age effects, too.

Table 2

Descriptive CES-D results for male and female SATSA participants who did and did not become widowed

	Men						Women					
	No Marital Status Change			Became Widowed			No Marital Status Change			Became Widowed		
	n	*M*	*SD*	n	*M*	*SD*	n	*M*	*SD*	n	*M*	*SD*
1987												
< 60 years	158	8.43	8.14	1	18.00	NA	218	8.84	7.31	9	15.44	10.70
60 to 69 years	143	9.19	6.89	7	12.57	9.38	155	12.30	9.70	12	17.85	10.48
70 to 79 years	101	10.15	6.96	8	14.25	7.19	171	12.83	9.10	26	17.04	8.96
≥ 80 years	24	15.45	8.89	2	17.00	4.24	47	15.35	8.96	14	14.45	11.24
1990												
< 60 years	146	8.50	6.66	1	9.00	NA	192	9.57	8.11	3	15.67	17.01
60 to 69 years	117	8.53	5.72	3	13.33	9.02	153	10.94	7.16	8	16.50	10.54
70 to 79 years	116	11.01	7.10	3	18.33	12.22	167	14.39	9.27	31	18.25	9.91
≥ 80 years	25	12.94	8.02	2	15.28	0.39	51	14.07	8.16	12	14.92	10.27
1993												
< 60 years	140	9.25	8.64	0	-	-	178	9.67	8.10	1	5	NA
60 to 69 years	104	7.88	6.46	4	12.00	8.08	124	11.90	8.31	11	14.64	7.78
70 to 79 years	126	10.76	7.59	6	14.00	8.32	159	14.34	8.61	23	17.84	11.74
≥ 80 years	30	11.73	6.77	3	15.47	6.80	72	15.05	9.78	18	15.39	7.41
2004												
< 60 years	34	8.97	6.18	0	-	-	66	11.35	9.14	0	-	-
60 to 69 years	77	9.78	7.81	5	11.40	10.64	88	9.21	7.54	8	15.25	6.80
70 to 79 years	56	8.37	6.73	8	15.12	6.92	81	13.43	7.73	21	15.10	8.99
≥ 80 years	40	11.21	7.18	7	17.08	5.88	50	18.35	7.94	27	17.61	7.81
2007												
< 60 years	24	10.38	8.23	0	-	-	36	13.37	11.99	1	21.00	NA
60 to 69 years	49	9.28	7.43	3	16.00	9.85	72	8.98	7.68	3	9.67	4.04
70 to 79 years	63	10.80	8.01	4	11.08	7.76	77	12.47	7.17	14	11.21	6.45
≥ 80 years	46	12.24	8.75	4	13.88	2.77	69	16.72	8.07	14	21.36	9.52

For both the men and women, the differences in mean CES-D scores between those who became widowed and those who did not become widowed suggests that the widowed reported higher CES-D scores at the next measurement occasion, regardless of the interval between waves. Although the sample of widowed men was small ($n < 10$ for all age groups at all waves), they had higher CES-D scores across all age groups and waves. The youngest age group (< 60 years old) cannot be interpreted, as only one participant became widowed in 1987, one in 1990, and none in 1993, 2004, and 2007.

The sample of widowed women was larger, with clear differences emerging between young and old women. The differences between widowed and nonwidowed women were larger in the youngest two age groups (< 60 years old and 60-69 years old), with widowed women reporting higher CES-D scores. As with the youngest widowed male group, only differences in 1987 and 1990 can be interpreted, as there were two few widows at later measurements to make meaningful comparisons. Women who became widowed at older ages (70-79 years old and > 80 years old) did not consistently report higher CES-D scores than their nonwidowed controls. Although older widows typically reported higher scores than older nonwidows, the differences typically were small. Overall, these results suggest that becoming widowed correlates with higher CES-D scores at subsequent measurements.

Next, LGC models were fit to the data to answer the second research question (*Above and beyond the general propensity for people to become more depressed as they age, does becoming widowed predict short-term and long-term increases in subsequent levels of depression?*). Model fitting results are presented in Table 3.

Table 3

Latent growth curve model fitting results in the SATSA

Men	-2LL	Parameters	Scaling correction factor	Model Comparison	LR	S-B LR test	AIC	BIC
1. Baseline model (Correlated intercept and slope with unstructured correlations among measured variables)	-1895.00	55	1.87	-	-	-	3901.91	4155.33
2. Short-term correlations constrained to be equal	-1896.19	53	1.91	1	2.98	0.226	3898.37	4142.67
3. Constrain short-term correlation to be zero	-1900.28	51	1.94	2	7.07	0.029	3904.56	4144.25
4. Constrain long-term correlation to be zero	-1897.13	51	1.97	2	3.40	0.182	3896.25	4131.38
Women								
1. Baseline model (Correlated intercept and slope with unstructured correlations among measured variables)	-3791.87	55	1.20	-	-	-	7693.74	7966.14
2. Short-term correlations constrained to be equal	-3792.15	52	1.21	1	0.55	0.908	7688.29	7945.84
3. Constrain short-term correlation to be zero	-3798.34	51	1.23	2	34.20	0.000	7698.69	7951.27
4. Constrain long-term correlation to be zero	-3793.11	51	1.21	2	1.65	0.200	7688.22	7940.81

Note. The scaling correction factor is a correction that better approximates chi-square under conditions of multivariate nonnormality (Satorra & Bentler, 2001). -2LL = -2 log-likelihood; LR = likelihood ratio; S-B = Satorra-Bentler likelihood ratio difference for nested model comparison; AIC = Akaike Information Criterion; BIC = Bayesian Information Criterion.

In men, the interval of time between the previous wave and the 1987, 1990, and 1993 waves was the same (3 years), so the short-term effects of becoming widowed on log CES-D scores at 1987, 1990, and 1993 were constrained to be equal to one another (Model 2). There was not a significant reduction in model fit by constraining these parameters to be equal. Although there was a 3-year interval between the 2004 and 2007 waves, it was not constrained along with the 1987, 1990, and 1993 effects, as too few men became widowed during this period. When included in the model, the Satorra-Bentler likelihood ratio tests produced nonsensical values. As a result, the wave 2007 parameter was freely estimated in the model but not interpreted. The short-term effect of becoming widowed could not be set to zero without a significant reduction in model fit

(Model 3). The long-term effect of becoming widowed, however, could be set to zero without a significant decrement in model fit (Model 4). Overall, the results suggest a small tendency for widowed men to experience short-term increases in CES-D symptoms.

In women, the short-term effects of becoming widowed on log CES-D scores at 1987, 1990, 1993, and 2007 could be set equal to one another without a significant decrement in model fit (Model 2). The effect could not be set to zero without a significant reduction in model fit (Model 3) whereas the long-term effect of becoming widowed – wave 2004 – could be constrained to zero without significantly reducing the fit of Model 2 (Model 4). As in the men, the results suggest a small tendency for widowed women to experience short-term increases in CES-D symptoms.

Table 4

Parameter estimates from the best fitting LGC model in male (Model 3) and female (Model 3) participants in the SATSA

	MLR estimates			
	Male		Female	
Parameter	Est.	*SE*	Est.	*SE*
μ_I	1.37	0.13	1.51	0.10
μ_S	0.01	0.002	0.01	0.001
σ^2_I	1.40	0.42	1.64	0.40
σ^2_S	0.000	0.00	0.000	0.00
σ_{IS}	-0.02	0.01	-0.02	0.01
$\sigma_{W,D}$ - short-term	0.01	0.003	0.02	0.01
$\sigma_{W,D}$ - long-term	0.03	0.02	0.02	0.02
σ^2_{eD1987}	0.32	0.03	0.33	0.03
σ^2_{eD1990}	0.28	0.03	0.28	0.03
σ^2_{eD1993}	0.34	0.04	0.28	0.03
σ^2_{eD2004}	0.29	0.07	0.30	0.04
σ^2_{eD2007}	0.33	0.06	0.28	0.04
σ^2_{eW1987}	0.04	0.01	0.08	0.01
σ^2_{eW1990}	0.02	0.01	0.08	0.01
σ^2_{eW1993}	0.03	0.01	0.08	0.01
σ^2_{eW2004}	0.08	0.02	0.14	0.02
σ^2_{eW2007}	0.06	0.01	0.10	0.01

Note. μ = mean; σ^2 = variance; σ = covariance; I = intercept; S = slope; W = became widowed; D = log CES-D score. *Short-term* refers to the effects of becoming widowed up to 3 years post-spousal loss. *Long-term* refers to the effects of becoming widowed up to 11 years post-spousal loss.

Twin-level analyses

A co-twin-control analysis was used as the first step toward answering the third research question (*Above and beyond genetic and environmental selection confounds, does becoming widowed predict short-term and long-term increases in subsequent levels of depression?*). Table 5 presents a contingency table of the total number of twin pairs concordant for widowhood status (i.e., both twins widowed or both twins not widowed) and discordant for widowhood status, separated by zygosity and gender groups at each wave. For all male twins (MZM and DZM), there were 7 twin pairs discordant for widowhood status in 1987, 4 discordant pairs in 1990, 8 discordant pairs in 1993, 11 discordant pairs in 2004, and 3 discordant pairs (DZM only) in 2007. For all female twins (MZF and DZF), there were 34 twin pairs discordant for widowhood status in 1987, 30 discordant pairs in 1990, 30 discordant pairs in 1993, 15 discordant pairs in 2004, and 13 discordant pairs in 2007. Overall, the sample sizes of discordant twin pairs at each wave were modest, particularly for the male twins. Co-twin differences in CES-D scores at each wave must be interpreted with caution.

Table 5

Contingency tables of male and female twin pairs in the SATSA who were concordant and discordant for widowhood status

		1987							
		Twin 1							
		MZM		MZF		DZM		DZF	
		Not Widowed	Widowed	Not Widowed	Widowed	Not Widowed	Widowed	Not Widowed	Widowed
Twin 2	Not Widowed	50	-	61	-	87	-	145	-
	Widowed	3	2	14	0	4	0	20	1
		1990							
		Twin 1							
		MZM		MZF		DZM		DZF	
		Not Widowed	Widowed	Not Widowed	Widowed	Not Widowed	Widowed	Not Widowed	Widowed
Twin 2	Not Widowed	47	-	55	-	76	-	122	-
	Widowed	1	0	13	0	3	0	17	2
		1993							
		Twin 1							
		MZM		MZF		DZM		DZF	
		Not Widowed	Widowed	Not Widowed	Widowed	Not Widowed	Widowed	Not Widowed	Widowed
Twin 2	Not Widowed	50	-	46	-	75	-	113	-
	Widowed	5	0	8	1	3	0	22	3
		2004							
		Twin 1							
		MZM		MZF		DZM		DZF	
		Not Widowed	Widowed	Not Widowed	Widowed	Not Widowed	Widowed	Not Widowed	Widowed
Twin 2	Not Widowed	19	-	24	-	26	-	43	-
	Widowed	3	0	4	3	8	0	11	1
		2007							
		Twin 1							
		MZM		MZF		DZM		DZF	
		Not Widowed	Widowed	Not Widowed	Widowed	Not Widowed	Widowed	Not Widowed	Widowed
Twin 2	Not Widowed	17	-	17	-	23	-	42	-
	Widowed	1	0	4	0	2	0	9	1

The co-twin-control analysis compared the CES-D total means of the twins who became widowed with their nonwidowed co-twins (Table 6). Negative scores indicate that widowed twins reported higher CES-D scores than their nonwidowed co-twins.

Table 6

Co-twin-control analysis of CES-D differences between widowed and nonwidowed twins in the SATSA

	Widowed vs. Nonwidowed		
	Pairs	Mean Intrapair Difference	*t*
1987			
Male	7	-7.83	-1.96
Female	34	-3.16	-1.45
1990			
Male	4	-2.25	-0.40
Female	30	-6.45	-2.74
1993			
Male	8	-6.34	-1.66
Female	30	0.35	0.13
2004			
Male	11	-5.24	-1.88
Female	15	-1.16	-0.35
2007			
Male	3	-6.78	-0.87
Female	13	-2.28	-0.79

The co-twin-control analysis results are consistent with the descriptive results and the LGC model results. Widowed male twins reported higher CES-D scores than

their nonwidowed co-twins across all waves, but only the difference in 1987 approached statistical significance ($p = 0.082$). Similarly, widowed female twins had higher CES-D scores than their nonwidowed co-twins in all waves except for wave 1993. Only widowed female twins at wave 1990 had significantly higher CES-D scores than their nonwidowed co-twins ($p < 0.009$). Overall, while it appears that the direction of the within-family differences between CES-D scores of widowed and nonwidowed discordant pairs is consistent with a causal interpretation, the lack of sample size at each wave was insufficient to detect the small differences between male and female twins discordant for widowhood status at each wave.

A graphical examination of the differences between discordant MZ and DZ twin pairs can provide a deeper understanding of how within-family differences in becoming widowed influence depression scores (Figure 3 and Figure 4). When the within-family difference in CES-D scores between discordant MZ pairs is less than the difference between discordant DZ pairs, there is evidence of partial confounding by genetic selection effects. To the extent that the widowed MZ twin has a higher CES-D score than the nonwidowed MZ twin, there is evidence of a nonshared (or quasi-causal) effect, consistent with the social causation hypothesis.

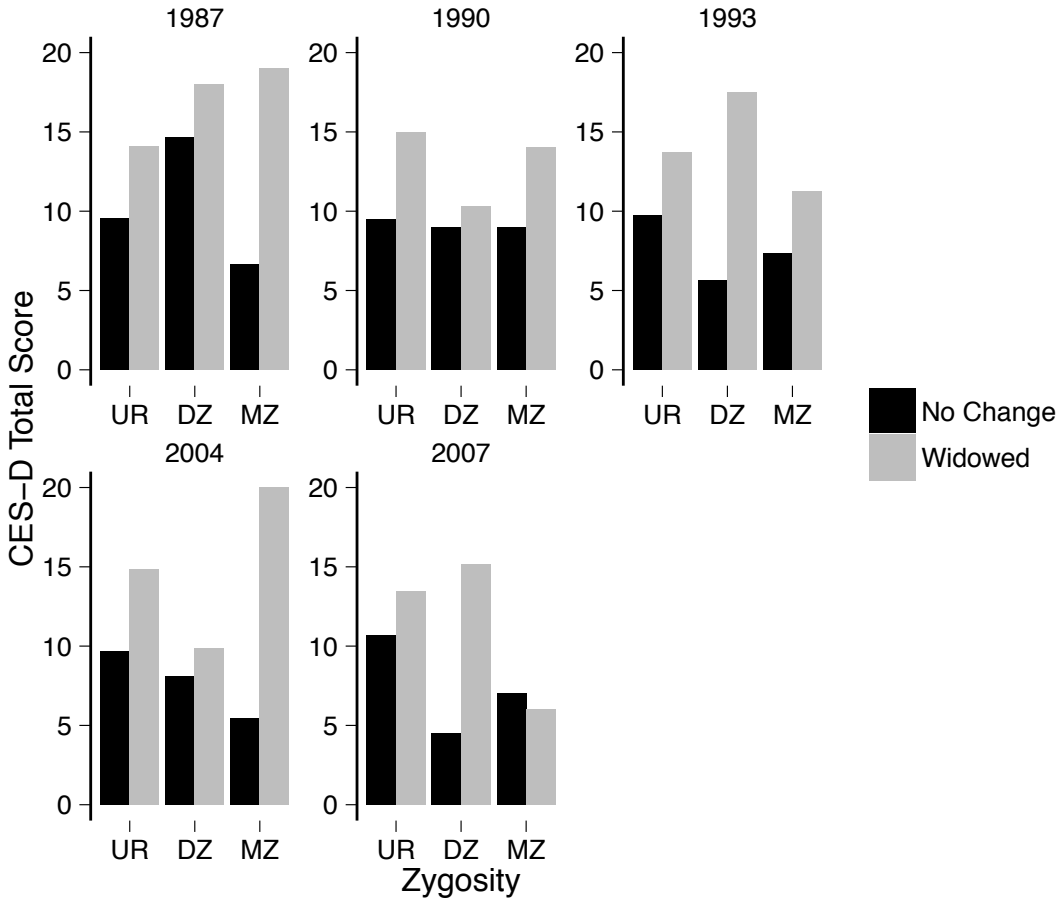

Figure 3. Bar plots of differences in mean total CES-D scores at each wave between pairs of MZ and DZ male twins discordant for becoming widowed in the SATSA. Pair differences were compared to the mean differences in unrelated groups discordant for becoming widowed.

For the male twins, Figure 3 illustrates that widowers reported higher CES-D symptom scores than their nonwidowed co-twins in all waves except 2007. The pattern of results, however, did not provide a clear impression of the underlying selection processes, as the differences in CES-D means of discordant DZ twins often were less than in the discordant MZ twins. This was the case in 1987, 1990, and 2004. This pattern suggests that genetic selection effects may have protected widowed twins from developing depression rather than raising their risk. Again, the co-twin-control analysis

was conducted with only a few discordant pairs of MZ and DZ twins. The estimates should be regarded as unreliable and interpreted cautiously.

For the female twins, the pattern of differences is similar as in the discordant male twins (Figure 4). Widowed twins generally reported more severe CES-D symptoms than their nonwidowed co-twins. The differences in means between discordant DZ twins were less than the differences in discordant MZ twins at waves 1987, 1993, 2004, and 2007, suggesting possible genetic protection against depression.

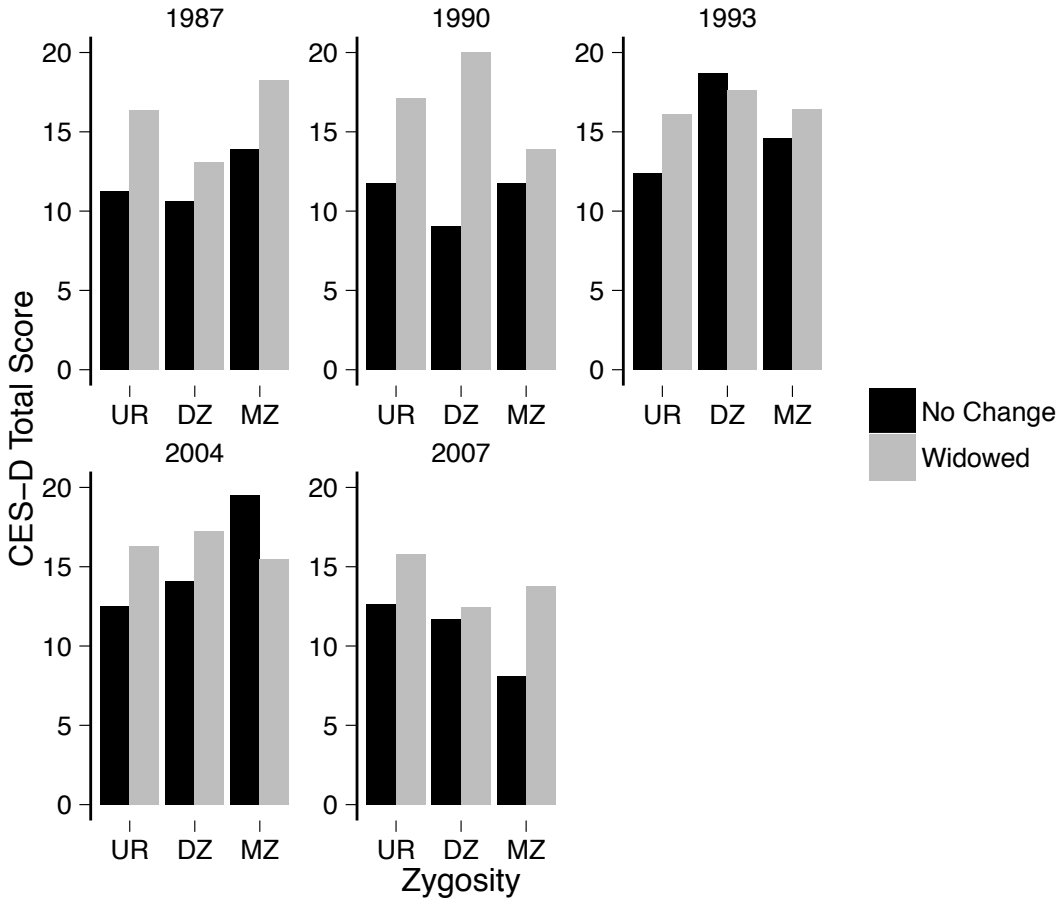

Figure 4. Bar plots of differences in mean total CES-D scores at each wave between pairs of MZ and DZ female twins discordant for becoming widowed in the SATSA. Pair differences were compared to the mean differences in unrelated groups discordant for becoming widowed.

The final step in the twin-level analysis consisted of fitting longitudinal quasi-causal models to test the short-term effect of widowhood on CES-D symptoms in male and female twins. Table 7 presents the model fitting results. All phenotypic variables in the model were regressed on twins' age at 1984 so that the estimates were age-adjusted.

Table 7

Longitudinal quasi-causal model fitting results in the SATSA

Male Twins	χ^2	df	Scaling correction factor	Model Comparison	LR	S-B LR test	RMSEA	CFI
1. Baseline (short-term only)	205.97	152	0.89	-	-	-	0.04	0.93
2. Genetic regression and variances set to zero	193.22	156	0.91	1	-5.94	-	0.03	0.95
3. Common environmental (C) regression set to zero	205.68	157	0.91	2	2.39	0.793	0.04	0.93
Female Twins								
1. Baseline ACE model	554.60	361	0.94	-	-	-	0.04	0.92
2. Genetic (A) regression set to zero	549.64	362	0.94	1	-1.71	-	0.04	0.93

Note. Baseline model included all genetic, common environmental, and nonshared environmental regression estimates for the short-term effects constrained to be the same across all waves. The scaling correction factor is a correction that better approximates chi-square under conditions of multivariate nonnormality (Satorra & Bentler, 2001). -2LL = -2 log-likelihood; LR = likelihood ratio; S-B = Satorra-Bentler likelihood ratio difference for nested model comparison; RMSEA = Root Mean Square Error of Approximation; CFI = Comparative Fit Index.

In the male twins, the baseline model (Male Twins Model 1, Table 7) consisted of only three waves, 1987, 1990, and 1993, as the short-term effect of becoming widowed on log CES-D scores was only considered in these waves. The baseline model (Model 1) fit the data well. The genetic variances, however, were negative in the baseline model. Several models were estimated with various starting values, lower bound constraints (e.g., variances estimates could not be estimated at values less than .01),

and fixed values constraints (e.g., setting the variances to extremely small values like 0.001) to determine whether local minima had been reached in the maximum likelihood estimation or whether the genetic variances were statistically no different from zero. Overall, these models (not presented) led to the conclusion that the genetic variances (and consequently, the genetic regression effects) could be set to zero (Model 2). The likelihood ratio test, moreover, was negative, indicating that the chi-square estimate of Model 2 was smaller than Model 1 and therefore more preferable. Model 3 included genetic variances, genetic regression, and common environmental regression coefficients constrained to zero, which fit as well as Model 1 ($\Delta\chi^2 = 2.39$, $df = 5$, $p = .793$). Model 3 was regarded as the best fitting model. The short-term quasi-causal effect of widowhood, however, was not significantly different from zero (see E_{Reg}-short-term estimate in Table 8), suggesting no statistically significant within-family environmental effect of becoming widowed on depression.

In the female twins, the baseline model (Female Twins Model 1, Table 7) consisted of all waves. Only the short-term effect of becoming widowed on log CES-D was tested, as it was the significant effect in the phenotypic LGC models. The baseline model (Model 1) fit the data well. The genetic regression could be set to zero without a significant reduction in model fit (Model 2). Moreover, the genetic variances were negative. Again, after estimating several models with different starting values, lower bound constraints, and fixed values to discern whether the negative genetic variance estimates could be attributed to local minima found in the maximum likelihood estimation, it was concluded that the genetic effects underlying becoming widowed were

not statistically different from zero. The common environmental component could not be set to zero without a significant reduction in model fit; Model 2 was regarded as the best fitting (final) model. The C component significantly mediated the phenotypic correlation between becoming widowed and the log CES-D scores (Table 8, C_{Reg} = 0.49, *SE* = 0.17, $p < .01$). The short-term quasi-causal effect, however, was not significant (p = .200). The model may have been underpowered to detect a significant short-term quasi-causal effect because there were only a small number of widowed female twins.

Table 8

WLSM parameter estimates from the baseline and best fitting longitudinal quasi-causal models for male and female twins in the SATSA

	WLSM estimates							
	Male				Female			
	Model 1		Model 3		Model 1		Model 2	
Parameter	Est.	*SE*	Est.	*SE*	Est.	*SE*	Est.	*SE*
σ^2_{A1987}	0.05	10.70	-	-	0.02	0.20	0.00	0.09
σ^2_{A1990}	0.00	13.88	-	-	0.01	0.22	0.01	0.09
σ^2_{A1993}	0.02	2.71	-	-	0.02	0.15	0.01	0.08
σ^2_{A2004}	-	-	-	-	0.11	0.14	0.11	0.09
σ^2_{A2007}	-	-	-	-	0.03	0.50	0.13	0.19
σ^2_{C1987}	0.24	7.84	0.29	2.69	0.01	0.13	0.00	0.06
σ^2_{C1990}	0.00	16.76	-0.17	0.66	0.01	0.14	0.00	0.06
σ^2_{C1993}	0.02	0.61	0.00	0.52	0.02	0.10	0.02	0.05
σ^2_{C2004}	-	-	-	-	0.02	0.12	0.03	0.08
σ^2_{C2007}	-	-	-	-	0.04	0.35	0.04	0.13
σ^2_{E1987}	0.02	1.11	0.03	1.17	0.15	0.09	0.17	0.08
σ^2_{E1990}	0.45	9.66	0.41	4.18	0.15	0.10	0.15	0.07
σ^2_{E1993}	0.19	3.58	0.26	2.67	0.12	0.07	0.13	0.05
σ^2_{E2004}	-	-	-	-	0.07	0.04	0.06	0.04
σ^2_{E2007}	-	-	-	-	0.40	0.26	0.32	0.22
A_{Reg} - short-term	0.50	58.02	-	-	0.31	1.68	-	-
C_{Reg} - short-term	0.501	8.34	0.45	11.16	0.421	0.15	0.49	0.17
E_{Reg} - short-term	0.55	0.39	0.37	0.91	0.20	0.19	0.23	0.18
A_{Reg} - long-term	-	-	-	-	0.82	1.80	1.366	2.028
C_{Reg} - long-term	-	-	-	-	-0.44	1.07	-1.44	3.91
E_{Reg} - long-term	-	-	-	-	-0.77	1.07	-0.87	1.48
σ^2_{eD1987}	0.58	3.49	0.59	2.03	0.66	0.04	0.66	0.04
σ^2_{eD1990}	0.26	1.64	0.37	1.04	0.45	0.04	0.46	0.06
σ^2_{eD1993}	0.33	0.21	0.36	0.28	0.36	0.04	0.42	0.11
σ^2_{eD2004}	-	-	-	-	0.28	0.32	0.27	0.75
σ^2_{eD2007}	-	-	-	-	0.40	0.13	0.54	0.26

Note. A = additive genetic; C = common environment; E = nonshared environment; D = log total CES-D scale score; e = residual variance; Reg = regression coefficient. *Short-term* refers to the effects of becoming widowed up to 3 years post-spousal loss. *Long-term* refers to the effects of becoming widowed up to 11 years post-spousal loss.

The final analysis in this chapter addressed the fourth research question (*Is widowhood status associated with reduced longevity? And if so, does depression mediate the association?*). First, the mean age of death of ever-widowed participants were compared to the mean age of death of never-widowed participants (Table 9). Ever-widowed participants were found to outlive the never-widowed participants, regardless of the age cohort.

Table 9

Descriptive results of age of death in male and female SATSA participants who never and ever experienced becoming widowed, by age group

	Men						Women					
	Never-widowed			Ever-widowed			Never-widowed			Ever-widowed		
	n	*M*	*SD*	n	*M*	*SD*	n	*M*	*SD*	n	*M*	*SD*
< 60 years	654	74.71	11.84	48	82.96	7.40	91	67.41	11.47	9	72.89	7.80
60 to 69 years	275	74.79	10.52	21	81.57	7.28	215	78.08	10.50	68	84.69	5.60
70 to 79 years	198	78.67	9.83	16	88.19	5.64	308	81.44	9.74	67	88.94	4.87
≥ 80 years	45	84.71	8.89	1	88.00	NA	117	85.50	8.35	13	93.15	5.40

A co-twin-control study was used to test whether ever-widowed participants outlived never-widowed participants. Pairs of twins who were discordant for ever-becoming widowed status were identified. Only ever-widowed twins with a never-widowed twin were used in this analysis. For pairs where at least one member had a known age of death, the twin who died first was identified (e.g., the ever-widowed twin died before the never-widowed twin). The discordant pairs were then divided into three groups based on the length of time the ever-widowed twin was widowed before death: ≤ 2 years, 2-4 years, and > 4 years. Length of time in widowed status is a *very* rough

estimate based on an estimate of twins' age of becoming widowed and age of death. Age of becoming widowed was calculated by subtracting twins' year of birth from the first measurement year twins' reported being widowed. It is not a precise indicator of age of becoming widowed, as spousal loss could have occurred within a 3-year period between most measurements and within an 11-year period between the farthest two measurements (1993 and 2004). Length of time in widowed status was then calculated by subtracting twins' estimated age of becoming widowed from their age of death.

The associations between widowhood and mortality are presented in Table 10. For male and female discordant pairs across the three groups, a greater proportion of ever-widowed twins outlived their never-widowed co-twins. Odds ratios less than 1 suggest that never-widowed twins had a greater risk of early mortality compared to their ever-widowed co-twins. The small sample sizes of discordant twin pairs and wide 95% confidence intervals, however, suggest that the estimates are imprecise and should be interpreted cautiously. Based on results in the opposite direction of what was expected, no tests were performed to determine whether increases in depression mediated the correlation between widowhood and mortality risk.

Table 10

Associations between ever-widowed and early death in SATSA

Twin pair types	Number of twin pairs	Number of ever-widowed–never-widowed pairs		Odds Ratio	95% CI
		EWID died first-NWID died second	EWID died second-NWID died first		
All	277	60	117	0.51	0.38-0.70
MZ	91	21	39	0.54	0.32-0.92
DZ	186	39	78	0.50	0.34-0.73
Male - widowed ≤ 2 years	11	4	7	0.57	0.17-1.95
MZ	6	2	4	0.50	0.09-2.73
DZ	5	2	3	0.67	0.11-3.99
Female - widowed ≤ 2 years	22	8	10	0.80	0.32-2.03
MZ	6	3	3	1.00	0.20-4.95
DZ	16	5	7	0.71	0.23-2.25
Male - widowed 2-4 years	13	2	7	0.29	0.06-1.38
MZ	3	1	0	-	-
DZ	10	1	7	0.14	0.02-1.16
Female - widowed 2-4 years	34	5	11	0.45	0.16-1.31
MZ	11	2	5	0.40	0.08-2.06
DZ	23	3	6	0.50	0.13-2.00
Male - widowed > 4 years	50	10	23	0.43	0.21-0.91
MZ	20	3	10	0.30	0.08-1.09
DZ	30	7	13	0.54	0.21-1.35
Female - widowed > 4 years	147	31	59	0.53	0.34-0.81
MZ	45	10	17	0.59	0.27-1.28
DZ	102	21	42	0.50	0.30-0.84

Note. EWID = ever-widowed twin; NWID = never-widowed twin.

Discussion

The strength of the SATSA is that it possesses two crucial components for testing whether widowhood causally influences depression: twins and repeated measurements of marital status and depressive symptoms. In view of recent prospective studies on bereavement and resiliency (Bonanno, 2004), the present study was designed to provide an update to Lichtenstein et al.'s (1996) study with a broader comparison group, additional data, newly tested time-lagged effects of becoming widowed on depression, and exploration of specific selection processes (genetic, environmental, or both). Additionally, depression was considered as a possible mediator in the association between widowhood status and mortality risk (Lichtenstein et al., 1998).

The main distinguishing feature between the present study and Lichtenstein et al.'s (1996) study is the use of a different comparison group. Rather than comparing widowed to still-married controls, widowed twins were compared to all nonwidowed controls so that the experience of widowhood as a stressful life event (Holmes & Rahe, 1967) could be interpreted against all nonwidowed people's relationship experiences (not just against a married control group). Whereas Lichtenstein and colleagues found that short-term widowed male and female twins experienced significantly higher depression compared to their still married co-twins, the same was not found when comparing them to all nonwidowed co-twins. Widowers were only found to be modestly more depressed than their nonbereaved co-twins in 1987; widows were only found to be significantly more depressed than their nonwidowed co-twins in 1990. One reason for

the difference is that the data from each wave were not pooled, as was done in Lichtenstein et al.'s study, but examined at each wave of measurement. When the short-term (3-year) effect of widowhood was tested in the multivariate SEMs, the effect of becoming widowed on depression was replicated in both men and women. The twin models, unfortunately, were underpowered to detect small within-family environmental effects of becoming widowed on depression.

The addition of two more waves of data did not obviously appear to extend knowledge of the short-term effects of becoming widowed on depression. The null long-term finding, however, may advance an understanding of resiliency after conjugal loss. The long-term effect was not significantly different from zero, suggesting that widowed twins were not subsequently more depressed than nonwidowed twins. Widows and widowers may be resilient to spousal loss over the long-term (Bonanno et al., 2002; Lucas et al., 2003), possibly because they create meaning from their loss (Bonanno et al., 2004) or successfully construct adaptive post-widowhood identities (Martin-Matthews, 2011). Thus, depressive risks associated with becoming widowed likely occur in the few months and years after spousal loss. Future research should investigate how the genetic and environmental risks for depression change in the months and years following spousal loss.

Finally, the present study identified potential selection processes that mediated the association between becoming widowhood and depression, a component Lichtenstein et al. (1996) did not address. In both male and female twins, there was a significant common environmental selection process that mediated the short-term

association between becoming widowed and depression. This result is unexpected, as common environmental influences on human behavior often diminish after adolescence (Plomin & Spinath, 2004). One possibility is that twins from lower socioeconomic backgrounds may be exposed to environments that increase the likelihood of becoming widowed and depression (e.g., poverty and illness). Alternatively, twins from families who encourage marriage necessarily satisfy the precondition for widowhood (i.e., being married), increasing the likelihood of experiencing widowhood. Despite selection processes, becoming widowed likely predicted subsequent increases in depression in women.

Two additional findings are worthy of note. First, the results from the LGC models illustrate that depressive responses associated with widowhood occur over the short-term, above and beyond the tendency for adults to gradually increase in depression over the lifespan (Alexopoulos, 2005; Beekman et al., 2002; Blazer, 2003). Age trends in depression may confound cross-sectional and longitudinal findings, so these results help to disentangle the effects of within-person growth in depression from the effects of becoming widowed. Future research should test the short-term effects of widowhood on depression, adjusting for genetic and environmental selection effects as well as age trends in depression.

Second, the mediating role of depression in the association between widowhood and risk of early mortality should be explored in other samples, as even the association between widowhood and early death could not be replicated in the current study using a similar sample (Lichtenstein et al., 1998). Instead, the present results suggest that ever-

widowed twins lived longer than their never-widowed co-twins. Three reasons might explain these counterintuitive findings. First, pairs of twins in which widowed twins had co-twins in other marital statuses other than the still married category (e.g., divorced and never married) were included in the current study. While the benefit of including these discordant twin pairs was an increase in sample size, one cost is the introduction of confounds related to marital status. One of the benefits of marriage, for example, is longevity (Manzoli et al., 2007), so ever-widowed twins might have lived longer than their never-widowed co-twins because the never-widowed control group contained never married and divorced co-twins, two groups found to have increased risk for early death (Y. Hu & Goldman, 1990). Ideally, only widowed twins and their still married co-twins would have been included in the analysis to eliminate this possible confound.

Second, the average age at which widowed twins' spouses died may have been a potential confound that explains why widowed twins were found to live longer than their never-widowed co-twins. The older the average age of the widowed twins' spouses' death, the older these twins must have been when they died. By adjusting for the average age of spouses' death, confounds associated with the fact that these twins necessarily have longer lives are taken into account. Future twin studies on early mortality risk associated with widowhood, then, should take into account the average age of spouses' death.

Finally, recent widowhood – not lifetime widowhood – has been found to correlate with mortality. After a brief critical period of increased mortality risk, widows were actually found to outlive their still married co-twins (Lichtenstein et al., 1998). Thus,

there may be a resilient majority who ably weather spousal loss and continue to live productive lives.

The same limitations cited in Lichtenstein et al. (1996) apply to the present study, including insufficient sample sizes to divide the sample by rearing status (there are subsamples of twins reared together and reared apart in the SATSA) and problems with generalizing the results to the general population. There are three additional limitations in the present study. First, sample attrition due to death may have influenced the results. Although missing data were assumed to be missing at random (MAR) in the present study, mortality is not a randomly distributed variable, so death may have occurred systematically (e.g., the most depressed may have died the earliest). Ideally, future studies should only include twins with completed data over all measurements as a means of ruling out systematic bias associated with attrition due to mortality. Second, widowhood was calculated in a way that only considered each twin as "widowed" once in the entire study. A twin considered widowed at one measurement was considered non-widowed at the next measurement. A nonresilient long-term widowed twin's CES-D score potentially biases the nonwidowed CES-D mean score so that the absolute difference between the widowed and nonwidowed group means was reduced. As a result, the short-term effect of widowhood on depression may be a conservative estimate. Third, the full range of time-lagged effects of widowhood on depressive symptoms was not tested. Although long-term effects of widowhood were not found in the present study or in Lichtenstein et al.'s study, no formal tests of other time lags were conducted.

The strengths of the current study, however, outweigh the limitations. In an era of research that has begun to emphasize the resiliency of widows and widowers, it is important to recognize that for some widowhood comprises a social loss that has severe depressive consequences. There is still much to be learned about how spousal loss affects at-risk subpopulations so that interventions can be developed to help widows and widowers navigate bereavement healthfully and productively. Larger longitudinal twin studies with more intensive measurements are needed to identify the periods of greatest risk after spousal death as well as the mechanisms (e.g., social isolation, stress associated with caregiving, or poor health) that mediate the causal relation between becoming widowed and depression.

Divorce and Early Death

Divorce is a stressful life event that purportedly withdraws the physical, mental, and social benefits associated with staying married. Although many who negotiate divorce are resilient and navigate the structural and emotional transitions without trouble (Sbarra, 2012), research suggests gloomy physical and mental health outcomes for others (Waite & Gallagher, 2000; Waite, Luo, & Lewin, 2009). Increased risk of early mortality is one such outcome (Choi & Marks, 2011; Manzoli et al., 2007; Tucker, Friedman, Wingard, & Schwartz, 1996), with divorced people found to have an average increased risk of all-cause death by one-quarter relative to married people (Sbarra et al., 2011). In a recent meta-analysis on the relation between divorce and early death, Sbarra et al. outlined a research agenda for determining the mechanisms underlying divorce and risk of early mortality. The purpose of this paper is to address the first step of their proposed agenda by clarifying the causal relation between divorce and risk of early mortality.

Divorce may be correlated with increased risk of all-cause death for two reasons: social selection and social causation. The *social selection hypothesis* posits that divorce is not randomly distributed, meaning that some people are more likely to divorce than others, even for genetic reasons (Mcgue & Lykken, 1992). The same set of selection effects that put people at risk for divorce (e.g., aggressive temperament, impulsive personality traits) may also put them at risk for early death (Herskind et al, 1996; McGue, Vaupel, Holm, & Harvald, 1993). The *social causation hypothesis* suggests that divorce increases mortality risk, most likely by triggering a chain reaction of structural

(e.g. financial), health (e.g., smoking), and emotional (e.g., chronic stress) changes that are causally related to mortality (Sbarra et al., 2011). Social selection and social causation are not competing hypotheses, but separable mechanisms that operate in tandem to explain increased mortality risk among divorced people.

Co-twin control studies are one method for disentangling selection processes from causal ones (Lichtenstein et al., 1998, 1996). Twins share common genetic ancestry and exposure to similar, sometimes the same, environments (if raised in the same family), so any differences in death rates between twins discordant for divorce status must be attributed to within-family causes. Consider a scenario in which identical twins differ only in marital status – one is married and the other is divorced. The twins not only are genetically matched (100% of all genes are shared between MZ twins and 50% of all genes are shared between DZ twins), but they also are matched for ethnicity, age and other shared environmental factors (e.g., parents' SES and parents' rearing practices). The married twin has the advantage of emotional and practical support from a spouse that confer health advantages (Lillard & Waite, 1995) whereas the divorced twin lacks this support mechanism. The expectation is that the married twin, by virtue of spousal support, has a higher likelihood of outliving the divorced twin, as all the factors on which they are matched are held constant between them.

Identical (or monozygotic (MZ)) twins offer the clearest explanation for how twin studies take into account genetic and environmental selection confounds, as they share all of their genes and the same rearing environments. Any differences that arise between them must be environmental in nature. Fraternal (or dizygotic (DZ)) twins, on

the other hand, may differ for genetic reasons, as they only share half of their genes, on average. By comparing the death rates of MZ and DZ twins who are discordant for divorce and marriage, it is possible to disentangle within-family environmental effects of divorce on early death from genetically mediated selection effects. Whereas co-twin-control studies have been used to study the causal influence of bereavement on longevity (Lichtenstein et al., 1998), the present study is the first to use twins to test whether divorce causally influences risk of early mortality.

Other covariates may also confound a full understanding of the effects of divorce on early death. First, marriage and divorce have been found to affect men and women differently (Choi & Marks, 2011; Y. Hu & Goldman, 1990; Sbarra et al., 2011; K. Williams & Umberson, 2004). Notably, some researchers have not found gender differences in the association between divorce and early death (Manzoli et al., 2007; Tucker et al., 1996). Where gender differences have been observed, men typically have been found to have greater mortality risks than women (Sbarra et al.). Gender differences in the association between divorce and early death, therefore, were tested in the current study, with currently divorced men predicted to have the greatest risk (Sbarra & Nietert, 2009).

Second, sampling characteristics, like age and cohort effects, also may confound knowledge regarding the causal influence of divorce on early death. For example, divorce that occurs in late adulthood has been found to be less correlated with early death than in midlife and early adulthood (Manzoli et al., 2007; Sbarra et al., 2011). Additionally, the relation between divorce and health outcomes is embedded in the

social and cultural context. In Sweden, for example, no-fault divorce was legalized in 1915 (the Act on Celebration and Dissolution of Marriage; Simonsson & Sandstrom, 2011) that made divorces easier to obtain as long as both spouses agreed that the marriage was “a permanent failure.” After 1973, divorces became easier to secure, with only one spouse required to view the marriage as a permanent failure. The changes in the social surface of marriage and divorce in Sweden likely normalized the divorce experience, possibly lessening health consequences. Thus, there may be cohort differences in the association between divorce and early death.

Finally, divorced people may have a greater risk of early mortality because of a higher risk of completed suicide (Kposowa, 2000). Suicide is difficult to study because of its low base rates (Miller, Azrael, & Barber, 2012), but the sample used in the present study is large enough for studying the association between divorce and suicide. The Swedish Twin Registry (STR) is a national database of all Swedish citizens from 1886 to the present that includes the birth and death (including cause of death) records of all citizens (Lichtenstein et al., 2002). The magnitude of the sample size in the STR made it possible to explore whether divorced adults have an increased risk of suicide compared to married and unmarried people, particularly in divorced men (Schrijvers, Bollen, & Sabbe, 2012).

Present Study

Sbarra et al. (2011) argued that “selection effects [probably] do operate to predict the likelihood of future divorce and health outcomes...” (p. 464). A key first step toward understanding whether divorce influences risk in all-cause mortality is to conduct a co-

twin control study. Several co-twin control studies have been used to elucidate the causal effects of bereavement on physical and psychological health (Lichtenstein et al., 1996; Osler et al., 2008) and mortality (Lichtenstein et al., 1998). No study to date, however, has been published in the area of divorce. The purpose of the present study was to fill this gap in the literature. The following research questions were addressed:

1. Is the association between divorce and early death replicated in a sample of Swedish twins?
2. In twin pairs discordant for marriage and divorce, does the divorced twin have a higher risk of early mortality than the married co-twin?
3. In twin pairs discordant for remarriage and current divorce status, do currently divorced twins have a higher risk of early mortality than their remarried co-twins?
4. Is there an association between divorced status and suicide risk relative to married and never married statuses?

Method

Sample

The sample used for the current study comes from the Swedish Twin Registry (Lichtenstein et al., 2006). Participants in the present study belong to one of two cohorts. The first cohort (old cohort) was born between 1886-1925 and the second cohort (young cohort) was born between 1926-1958. The old cohort consists of 24,601 participants (female = 55.46%) who primarily were interviewed in 1961, 1963, 1967, and 1970. Surviving participants were again interviewed in the Screening Across the Lifespan Telephone (SALT) study conducted between 1998-2002 (Lichtenstein et al.,

2002). The mean age at the first measurement (1961) was 48.58 (*SD* = 9.76) for men and 49.29 (*SD* = 10.05) for women. The young cohort consists of 56,707 participants (female = 51.10%) who were interviewed in 1973 and in the SALT study. The mean age at the first measurement (1973) was 28.43 (*SD* = 9.79) for men and 28.99 (*SD* = 9.88) for women.

In the total sample, there were 69,167 (85.07%) participants who reported their marital status in at least one of the measurements. From this number, 41,737 participants were married (60.34%), 3,571 were cohabiting (5.16%), and 6,739 reported being divorced at some point in their lives (9.74%). Among the divorced participants, 4,923 were currently divorced (73.05%), 696 were remarried (10.33%), and 1,120 were in nonmarital partnerships (16.62%).

For the co-twin control analysis, only pairs where both twins had reported marital status information (marriage, cohabitation, or one of the divorce status categories) were used. Opposite-sexed pairs were excluded from the analysis in order to control for gender effects. Table 1 presents the number of male and female twin pairs in each cohort by their pairwise death status (both deceased, one twin deceased, neither deceased).

Table 1

Number of STR twin pairs concordant and discordant for death status

	Both deceased		One twin deceased		Neither deceased	
	Old Cohort	Young Cohort	Old Cohort	Young Cohort	Old Cohort	Young Cohort
MZM	869	192	72	412	17	1309
MZF	602	114	90	402	35	1715
DZM	1450	245	140	777	13	1855
DZF	1007	147	211	703	44	2298

Note. MZM = monozygotic male; MZF = monozygotic female; DZM = dizygotic male; DZF = dizygotic female.

As can be seen in Table 1, the majority of twin pairs in the old cohort were deceased by the date the data were compiled with the most recent death records (2013-04-09) whereas the majority of twins in the young cohort were still alive. The number of pairs in each zygosity category includes pairs who were concordant (e.g., both married or both divorced) and discordant (e.g., one married, one divorced) for marital status. The co-twin control analysis (described below) included only twins who were discordant for marriage and divorce status.

Measures

Marital status and dates of death were obtained through annually collected data from the Swedish population registry, similar to Lichtenstein et al.'s (1998) bereavement study. Cause of death was obtained through the cause of death registry using each participant's unique Swedish personal number.

Marital status information was collected in 1961, 1967, 1970, 1973, and in the SALT study. For the old cohort, marital status information was collected in 1961, 1967, 1970, and in the SALT survey. For the young cohort, marital status information was collected in 1973 and in the SALT survey. At each measurement, participants reported whether they were *unmarried*, *married*, *separated*, *divorced*, *widowed*, *remarried*, or *cohabiting*. Where participants only reported marital status at one measurement, this marital status was taken as their final status. Where participants reported marital status at multiple measurements, algorithms were created to properly code participants as *married* (never divorced) and *ever divorced* (i.e., self-reported divorce in 1961 and self-reported married in 1970). Among the ever divorced, participants were coded as *currently divorced* if they reported divorce in their most recent measurement, *remarried* if they last reported being married and had ever reported divorce, or being in a nonmarital repartnership if they last reported cohabitation and had ever reported divorce.

Age of death was calculated by subtracting participants' date of death from their date of birth. Participants who were alive at the time the data were compiled were assigned a missing value ("NA") for age of death. Completed suicide is a binary variable (1 = *suicide*, 0 = *other cause of death*) and was coded according to criteria in the ICD-10 (World Health Organization, 1992). The suicide rate in the sample is 0.19% (N = 81,308), which is higher than the estimated suicide rate in Sweden (Björkenstam, Edberg, Ayoubi, & Rosén, 2005) and the United States (Miller et al., 2012).

Data Analysis

The descriptive results consisted of the number of deaths and relative risk (RR) estimates of early death associated with being *ever* divorced and *currently* divorced groups relative to married and partnered groups. Like the odds ratio, the relative risk statistic is the proportion of deceased divorced people from each divorced group (total number of deceased divorced divided by the total number of divorced people) divided by the proportion of deceased married people. The RRs associated with divorce were calculated relative to a married group and an all partnered (married and cohabiting) group. The analyses were performed separately for men and women in the old and young cohorts to test for gender and cohort effects. The same analytic approach was used for suicide risk, except cohort was not considered to increase the power of rejecting the null hypothesis in the male and female subgroups.

The co-twin control analysis used all pairs of MZ and DZ twins who were discordant for being married and divorced status. Divorced twins (probands) and their corresponding twin (regardless of marital status) were first identified. Only pairs discordant for divorce and being partnered were used to test whether divorce status corresponds with a higher risk of a death than partnered status. Discordant twins were tabulated as follows: (1) both twins were still alive, (2) the divorced twin died *before* the married co-twin, or (3) the divorced twin died *after* the married co-twin (or was still alive). Mantel-Haenszel odds ratios for matched pairs were calculated (Robins et al., 1986), as has been done in other co-twin control designs (Silman et al., 1996).

Finally, Cox proportional hazard regression models were used to analyze within-family differences in age of death between twins discordant for being divorced and partnered (Cox, 1972). Whereas the above analysis compared twins who were ever and currently divorced, this analysis only included discordant pairs where the divorced twin was currently divorced. The unit of analysis was the twin pair. Within-family differences in age of death were calculated by subtracting the age of death for the twin who died second (twin 2) from the age of death for the twin who died first (twin 1). Twins were sorted so that higher scores indicated the length of time twin 2 lived beyond twin 1. Differences in age of death were right censored in pairs where twin 2 was still alive and twin 1 was deceased. A set of covariates was also considered, including smoking status, excessive alcohol use, education status, body mass index, cardiovascular disease status, respiratory disease status, and other chronic disease status, similar to the approach used by Lichtenstein et al. (1998). Within-family differences in the covariates were calculated in the same way as the twins' difference in age of death. The following model was fit to the twins' within-family difference in age of death scores, where the difference in death, *w*, was modeled as a continuous-time hazard process (failure time of twin 2's death after twin 1's death):

$$\ln(Y(w)_j) = b_0 + b_1 MS(w)_j + b_2 X(w)_j + e_j$$

The regression estimates (b_0, b_1, and b_2) are the log odds of twin 2's death after twin 1's death. The log odds residual, e_{ij}, is assumed to be homoscedastic. The twin death order covariate, *MS*, is a binary coded variable where -1 indicates that the partnered twin outlived the divorced twin and +1 indicates that the divorced twin outlived the partnered

twin. Positive values of b_1 indicate a higher probability of death of divorced twins after their partnered twin's death. The covariates are represented by the generic variable, *X*. If the main effect of twin death order (Model 1) was significant, then a second model was tested to determine whether the probability of death after twin 1 was statistically significant adjusting for the effects of the covariates, b_2 (Model 2). The corresponding odds ratios (e.g., $e^{b}{}_{1}$) are interpreted as the predicted probability of twin 2's death given the time elapsed since twin 1's death. All analyses were performed in R 3.0.2 (R Core Team, 2013) using the survival() package (Therneau, 2013). Maximum likelihood estimation was used for all cox proportional hazard regression models.

Results

Table 2 presents the relative risk (RR) statistics and corresponding 95% confidence intervals for death among divorced people relative to married and partnered (i.e., married or cohabiting) people. For each marital status category, the total number of deaths, total number of participants, and the death rate are given. As can be seen for the old cohort (cohort 1), neither divorced men nor women had an increased risk of early death relative to their married or partnered counterparts. Conversely, in the young cohort, men who had ever divorced were 21% more likely to die earlier than both married and partnered men. The mortality risk associated with divorce was driven by young men who reported current divorce (RR = 1.30). Repartnered men did not have a greater risk of early death than either their married or partnered counterparts.

A similar pattern of mortality risk associated with divorce was observed for women in the young cohort. Currently divorced women were 27% more likely to die

earlier than married and partnered women whereas increased risk was not found for repartnered women. The RRs among the young cohort are consistent with Sbarra and Nietert's (2009) finding that separated/divorced people who later remarried did not have a greater risk of death than people who were stably married.

Table 2

Number of deaths, death rate, and relative risks of death with 95% confidence intervals by gender and cohort in the STR

	Born before 1926					At least born in 1926				
	Deaths	Total	Death Rate	Relative Risk	95% CI	Deaths	Total	Death Rate	Relative Risk	95% CI
Men										
Married	6208	6590	0.94	1.00		2576	14107	0.18	1.00	
Total Partnered	6247	6645	0.94	1.00		2714	15997	0.17	1.00	
Ever Divorced	702	753	0.93	0.99	0.97-1.01	419	2039	0.21	1.21	1.10-1.33
				0.99	0.97-1.01				1.21	1.10-1.33
Divorced (Currently)	435	460	0.95	1.00	0.98-1.03	322	1458	0.22	1.30	1.17-1.44
				1.01	0.98-1.03				1.30	1.18-1.44
Divorced (Repartnered)	267	293	0.91	0.97	0.93-1.01	97	581	0.17	0.98	0.82-1.18
				0.97	0.93-1.01				0.98	0.82-1.18
Women										
Married	5359	6042	0.89	1.00		1978	14998	0.13	1.00	
Total Partnered	5373	6061	0.89	1.00		2075	16605	0.12	1.00	
Ever Divorced	785	916	0.86	0.97	0.94-0.99	442	3031	0.15	1.17	1.06-1.28
				0.97	0.94-0.99				1.17	1.06-1.28
Divorced (Currently)	597	690	0.87	0.98	0.95-1.01	366	2315	0.16	1.27	1.14-1.40
				0.98	0.95-1.01				1.27	1.14-1.40
Divorced (Repartnered)	188	226	0.83	0.94	0.88-1.00	76	716	0.11	0.85	0.68-1.05
				0.94	0.88-1.00				0.85	0.68-1.05

Whereas the population-based RR estimates provide a first impression of the observed correlation between divorce and early death, the next analysis advanced knowledge of the causal influence of divorce on mortality risk by using a co-twin-control

design. Table 3 presents the odds ratios and 95% confidence intervals for male and female pairs of twins who were discordant for ever-divorced status and marriage. The odds ratios represent the risk of death associated with being the divorced twin. Odds ratios greater than one indicate that divorced twins had a higher risk of dying before their married co-twins compared to the risk of married twins dying before their divorced co-twins. The confidence intervals provide the upper and lower bounds in which there is 95% certainty that the true value ranges. Intervals that include 1 suggest that the married co-twins are no more likely to outlive divorced twins.

Table 3

Association between ever-divorced and early death in the STR

Twin pair types	Number of twin pairs	Number of ever divorced-partnered pairs		Odds Ratio	95% CI
		DIV died first-PAR died second	DIV died second-PAR died first		
All	3222	766	739	1.04	0.94-1.15
MZ	1264	286	282	1.01	0.86-1.20
DZ	1958	480	457	1.05	0.92-1.19
Male - Born < 1926	468	228	233	0.98	0.82-1.17
MZ	177	88	85	1.04	0.77-1.39
DZ	291	143	145	0.99	0.78-1.24
Female - Born < 1926	406	206	175	1.18	0.96-1.44
MZ	156	79	68	1.16	0.84-1.61
DZ	250	127	107	1.19	0.92-1.54
Male - Born ≥ 1926	1000	171	176	0.97	0.79-1.20
MZ	394	60	78	0.77	0.55-1.08
DZ	606	111	98	1.13	0.86-1.49
Female - Born ≥ 1926	1348	161	155	1.04	0.83-1.30
MZ	537	62	48	1.29	0.89-1.88
DZ	811	99	107	0.93	0.70-1.22

Note. DIV = ever-divorced twin; PAR = partnered twin.

There was not a significant risk associated with early death and ever-divorced status. Although many of the odds ratios were greater than 1, the 95% confidence intervals contain 1 as a plausible true value, meaning that the risk of death was found to be equal within families of twins discordant for marriage and ever-divorced status.

Table 4 presents the findings for pairs of twins who were discordant for being currently divorced and married. In both male and female twins, when all discordant twin

pairs were analyzed together, there was a significant association between current divorce and death (OR: 1.19). When analyzed by zygosity, however, the effect only appears in the DZ group (OR: 1.25). The difference between the nonsignificant effect in the MZ group and the significant effect in the DZ group suggests that genetic selection effects accounted for the association between current divorce status and early death. Selection processes, it seems, had a role in the correlation; the finding is inconsistent with a causal interpretation. The pattern of effects was only observed in the young cohort.

Table 4

Association between current divorced status and early death in the STR

Twin pair types	Number of twin pairs	Number of currently divorced-partnered pairs		Odds Ratio	95% CI
		DIV died first-PAR died second	DIV died second-PAR died first		
All pairs	2088	538	454	1.19	1.05-1.34
MZ	834	196	180	1.09	0.89-1.33
DZ	1254	342	274	1.25	1.06-1.46
Male - Born < 1926	267	136	128	1.06	0.83-1.35
MZ	99	47	50	0.94	0.63-1.40
DZ	168	89	78	1.14	0.84-1.55
Female - Born < 1926	285	148	120	1.23	0.97-1.57
MZ	109	58	45	1.29	0.87-1.90
DZ	176	90	75	1.20	0.88-1.63
Male - Born ≥ 1926	608	128	99	1.29	0.99-1.68
MZ	238	42	45	0.93	0.61-1.42
DZ	370	86	54	1.59	1.13-2.24
Female - Born ≥ 1926	928	126	107	1.18	0.91-1.52
MZ	388	49	40	1.23	0.81-1.86
DZ	540	77	67	1.15	0.83-1.59

Note. DIV = current divorced twin; PAR = partnered twin.

Next, the risk of death of current-divorced twins was compared to the risk of death of their remarried co-twins (Table 5). The odds ratios are inconsistent, with some estimates below 1 (current-divorced twins were predicted to outlive their repartnered co-twins) and some estimates above 1 (repartnered twins were predicted to outlive their current-divorced co-twins). None of the estimates, however, were statistically significant, as the 95% confidence intervals are broad and contain 1 as a possible true value. This

may be attributed to the small sample of discordant twin pairs (n = 69). These results should be interpreted cautiously, due to unreliability of the estimates.

Table 5

Association between current divorce (relative to repartnered co-twins) and early death in the STR

Twin pair types	Number of twin pairs	Number of currently divorced-repartnered pairs: DIV died first-REPART died second	Number of currently divorced-repartnered pairs: DIV died second-REPART died first	Odds Ratio	95% CI
All	163	32	37	0.86	0.57-1.47
MZ	83	14	16	0.88	0.43-1.79
DZ	80	18	21	0.86	0.46-1.61
Male - Born < 1926	23	11	12	0.92	0.40-2.08
MZ	11	5	6	0.83	0.25-2.73
DZ	12	6	6	1.00	0.32-3.10
Female - Born < 1926	20	7	13	0.54	0.21-1.35
MZ	10	3	7	0.43	0.11-1.66
DZ	10	4	6	0.67	0.19-2.36
Male - Born ≥ 1926	38	5	6	0.83	0.25-2.73
MZ	17	2	2	1.00	0.14-7.10
DZ	21	3	4	0.75	0.17-3.35
Female - Born ≥ 1926	82	9	6	1.50	0.53-4.21
MZ	17	4	1	4.00	0.45-35.79
DZ	37	5	5	1.00	0.29-3.45

Note. DIV = current divorced twin; REPART = divorced but repartnered twin.

Within-family Cox proportional hazard regression analysis

Finally, I evaluated whether marriage has benefits for longevity by using the twins' age of death information to estimate the probability of how much longer partnered

twins outlive their divorced co-twins. Kaplan-Meier curves were used first to model the probability of twin 2's death as a function of time after twin 1's death. Cox proportional hazard regression models were then used to test whether living divorced co-twins (twin 2) had a higher probability of death after their partnered co-twins (twin 1) died than living partnered co-twins (twin 2) probability of death after their divorced co-twins (twin 1) died.. In other words, I answered the following research question: Was there an increased factor by which divorced twins were expected to die after their partnered co-twins assuming null differences in longevity between partnered and divorced twin pairs?

The Kaplan-Meier curves were estimated for men and women in the old and young cohorts. Based on pairs where both twins were deceased in the old cohort, partnered male twins lived 10.18 years (*SD* = 8.05) after their divorced co-twins' death whereas divorced co-twins lived 8.31 years (*SD* = 7.12) after their partnered co-twins' death. Partnered female twins lived 10.07 years (*SD* = 7.67) after their divorced co-twins' death and divorced female twins lived 8.99 years (*SD* = 7.43) after their partnered co-twins' death. The plots (Figure 1a and 1b) illustrate the probability of death for partnered and divorced twins (twin 2) each year past their co-twins' (twin 1) death, taking into account censored pair differences. Although partnered male and female twins (in this case, twin 2) had higher probabilities of outliving their divorced co-twins (twin 1) over time, the confidence intervals (the shaded colored lines) around each curve are broad and overlapping, suggesting no significant differences between them. These nonsignificant differences were confirmed in the Cox proportional hazard regression models, as only the male twin differences were close to being different (log

odds = 0.25, *SE* = 0.13, *p* = 0.056) – with an odds ratio of 1.28 (95% CI: 0.99-1.65) – suggesting that divorced twins had a higher risk of death when they were twin 2 than when partnered twins were twin 2.

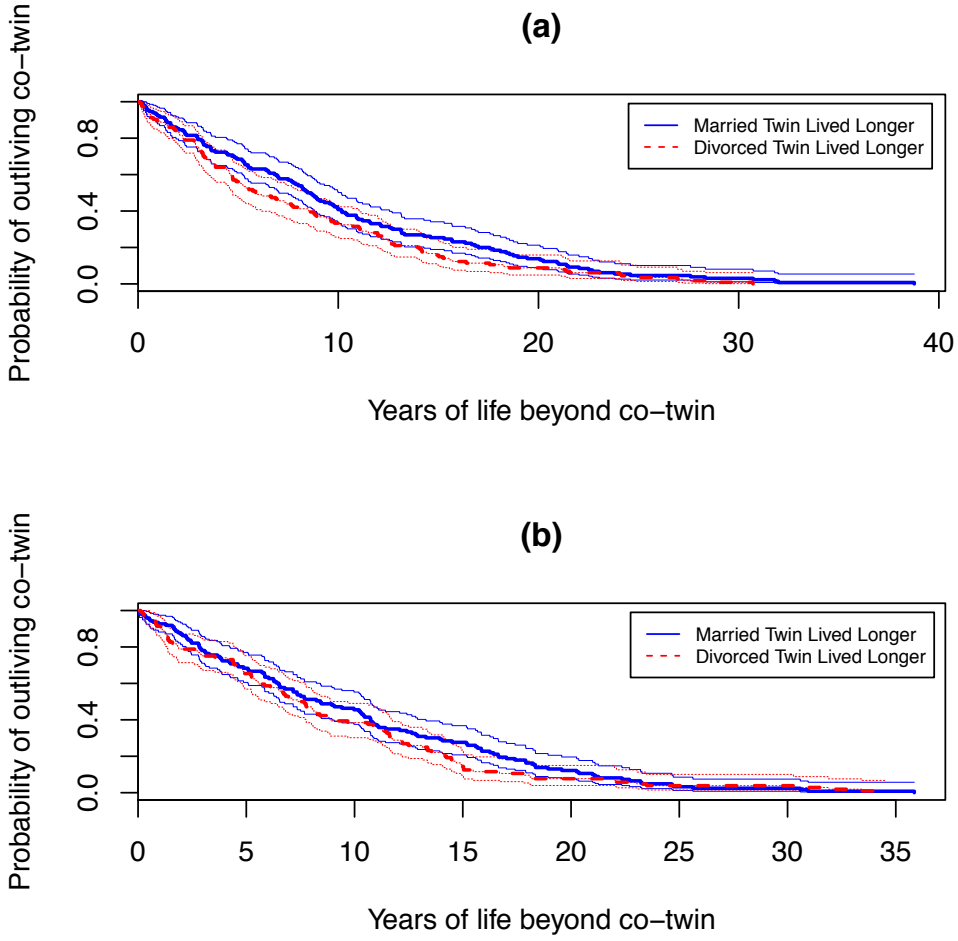

Figure 1. Kaplan-Meier curves for death after co-twin with 95% confidence intervals for discordant male (a) and female (b) twins in the old cohort.

In the young cohort, partnered male twins lived 9.35 years (*SD* = 8.18) after their divorced co-twins' death divorced whereas divorced male twins lived 7.76 years (*SD* = 8.56) after their partnered co-twins' death. Partnered female twins lived 9.13 years (*SD*

= 7.09) after their divorced co-twins' death whereas divorced female twins lived 11.00 years (*SD* = 7.88) after their partnered co-twins' death. These differences, however, are based on small samples of pairs (at most, 32 pairs of discordant twins). The 95% confidence intervals around the mortality curves (Figure 2a and Figure 2b) illustrate that the difference in length of life of the partnered twins beyond divorced twins (and vice versa) were not significantly different. The Cox proportional hazard regression models confirmed that the differences in length of life were nonsignificant, suggesting that partnered twins were no more likely to outlive their deceased divorced co-twins by more years than divorced twins were predicted to outlive their deceased partnered co-twins.

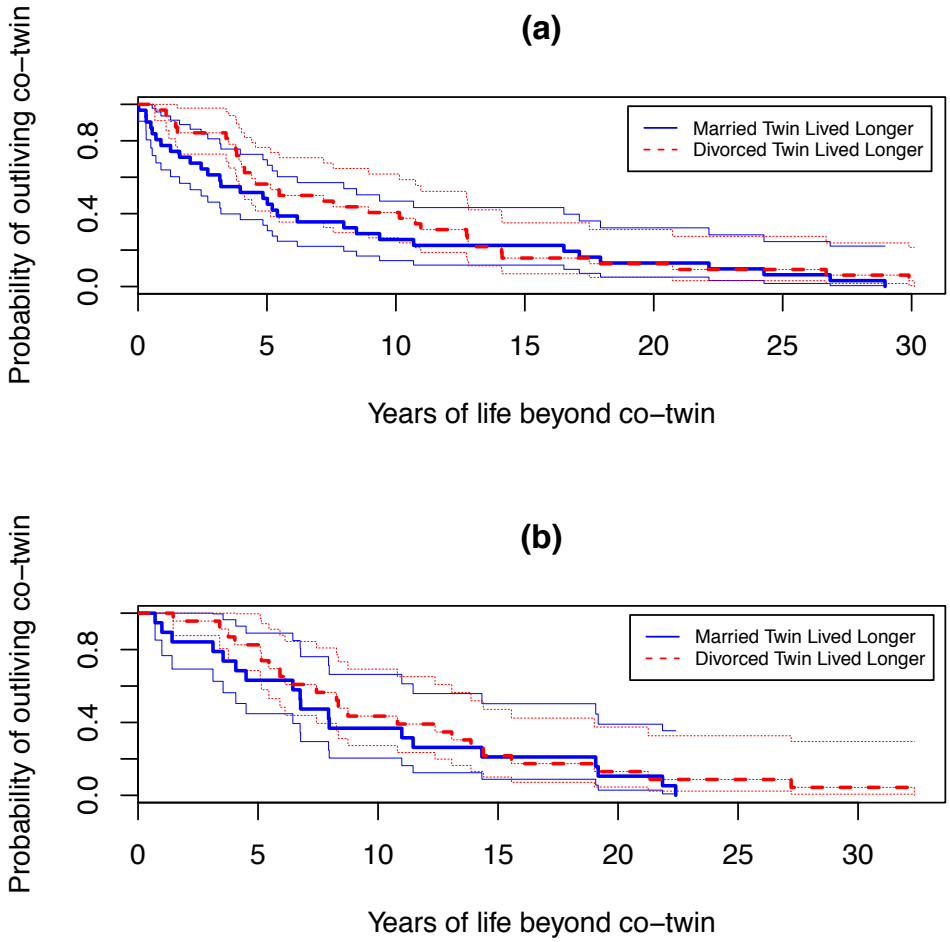

Figure 2. Kaplan-Meier curves for death after co-twin with 95% confidence intervals for discordant male (a) and female (b) twins in the young cohort.

Divorce and Suicide Risk

The final analysis explored the association between divorce status and suicide. Only the RR estimates for the population are presented (Table 6). There were no significant associations between ever-divorced status and suicide relative to being married or partnered. For the current divorced group, however, the RR estimates are greater than one in both men and women, suggesting a possible increase in suicide risk when divorced people do not remarry or repartner. The relative risk of suicide associated with being unmarried (i.e., never married) is also presented to illustrate

whether having ever been married (regardless of divorce) confers protection against suicide. The results suggest that unmarried men had a greater risk of suicide than married, partnered, and those who ever divorced. Unmarried women only had a greater risk of suicide relative to married and partnered people. The 95% confidence intervals are broad, so the estimates are considered imprecise and should be interpreted with caution.

Table 6

Number of suicides, suicide rate, and relative risks of suicide with 95% confidence intervals by gender in the STR

		Men					Women				
	Reference	Suicides	Total	Death Rate	Relative Risk	95% CI	Deaths	Total	Death Rate	Relative Risk	95% CI
Married		44	20697	0.0021	1.00		9	21040	0.0004	1.00	
Total Partnered		50	22642	0.0022	1.00		10	22666	0.0004	1.00	
Ever Divorced	*Married*	5	2792	0.0018	0.84	0.33-2.12	4	3947	0.0010	2.30	0.71-7.46
	Partnered				0.81	0.32-2.03				2.30	0.72-7.32
Divorced (Currently)	*Married*	5	1918	0.0026	1.23	0.49-3.09	4	3005	0.0013	3.02	0.93-9.79
	Partnered				1.18	0.47-2.96				3.02	0.95-9.61
Divorced (Repartnered)	*Married*	0	874	0.0000	0.00		0	942	0.0000	0.00	
	Partnered				0.00					0.00	
Unmarried	*Married*	33	5436	0.0061	2.86	1.82-4.48	10	5053	0.0020	4.63	1.88-11.38
	Partnered				2.75	1.77-4.26				4.49	1.87-10.77
	Ever Divorced				3.39	1.32-8.67				1.95	0.61-6.22
	Currently Divorced				2.33	0.91-5.96				1.49	0.47-4.74

Discussion

The present study was conducted to clarify whether divorce causally increases the risk of early death by conducting a co-twin-control analysis (Sbarra et al., 2011). Twin studies can help rule out genetic and environmental selection effects as explanations for their association – effects that cannot be adjusted for in population-

based samples (Turkheimer & Harden, 2013). Thus, if divorced twins more often die before their married co-twins, it can be concluded with greater confidence that divorce causally influences mortality risk. Gender differences and cohort differences were also tested to embed the association between divorce and early death in a lifespan developmental context (Baltes, 1987).

Overall, the results are inconsistent with a causal interpretation of divorce and early death, at least in a sample of Swedish male and female twin pairs. Although the population-based risk of early mortality (i.e., estimates unadjusted for genetic and environmental selection) was higher for current-divorced men and women in the young cohort relative to their married and partnered counterparts, the co-twin-control analysis results revealed that the increased risk was nonsignificant in all but the DZ male twins. The nonsignificant difference between the odds of early death among divorced twins compared to their married co-twins in the MZ twins, but a significant difference between discordant DZ twins is consistent with a genetic selection interpretation. Additive genetic effects that put younger divorced men at risk for early mortality may also put them at risk for divorce.

The current study does help to distinguish which divorced populations are at increased risk for early death. Sbarra and Nietert (2009) found that only divorced/separated adults who remained divorced had a greater risk of early death whereas mortality risk vanished in divorced adults who later remarried. The current study supports their finding. Only currently divorced Swedish men and women in the young cohort had a greater risk of early death whereas those who remarried did not.

The current results, however, provide a broad view of the difference between divorced populations and do not provide a clear picture of the causal mechanisms underlying the difference. Divorce might render some people “unmarriageable” for the reason that some remain under chronic emotional and financial strain in their post-divorce lives and do not attract new mates (Amato, 2010; Waite & Gallagher, 2000).

Divorced men typically have been found to have a greater risk of early death than divorced women (Choi & Marks, 2011; Hu & Goldman, 1990; Sbarra et al., 2011). At the population level, current divorced men were 30% more likely to die before their married and partnered counterparts, with current divorced women 27% more likely to die before their counterparts. The risk appeared to be approximately equal for current divorced men and women, as has been found in one metaanalysis (Manzoli et al., 2007). In the young cohort, however, there was evidence of genetic selection only in the male twins. One hypothesis is that men are more prone to taking risks (possibly because of higher aggressiveness or impulsivity) than women (Charness & Gneezy, 2012; Stanton, Liening, & Schultheiss, 2011), which might raise their risk of both divorce and early mortality.

There may be generational differences underlying the association between divorce and early death. Only divorced men and women in the younger cohort had a relative increase in mortality risk; divorced people in the older cohort had a relative risk of death no greater than married and partnered adults. There may be historical reasons for the difference (Simonsson & Sandstrom, 2011). From 1973 to the present day, no-fault divorce in Sweden has been easier to receive. It is imaginable that in the older

cohort born between 1886 and 1925 the degree of marital conflict may have needed to be so high that divorce offered release from marital stress and the promise of a more peaceful future for both spouses. In the young cohort, many of whom likely married around or after 1973, people could divorce one partner and remarry as their individual needs changed over time (Cherlin, 2004). As a result, some spouses may be forced into divorce, experiencing marital dissolution as a chronic emotional stressor, especially if one member in the marriage is still in love with the ex-spouse. This interpretation, however, is speculative. Prospective twin studies that include the marital histories of remarried and persistently divorced people across adulthood along with measures of physical, psychological, and structural variables are better suited to answer questions related to such cohort differences.

Suicide was considered as one possible mechanism for early mortality risk. Regarded as exploratory in the present study, there are several implications related to the possible effects of divorce status on risk of suicide completion. First, only currently divorced individuals committed suicide; there were no remarried people whose cause of death was suicide. Although the sample sizes were small, the pattern is consistent with findings here and elsewhere (Sbarra & Nietert, 2009) showing that divorced people who remarry may have lower mortality risks. Second, currently divorced women were three times more likely to commit suicide relative to married and partnered women whereas currently divorced men were only about 25% more likely to commit suicide relative to their married and partnered counterparts. Finally, there may be a protective benefit to having ever been married, particularly for men. Unmarried men had a large relative

increase in risk of suicide compared to married and divorced men. Unmarried women, on the other hand, had an increased risk of suicide only relative to married and partnered women. Based on these results, marriageable men may be less inclined to commit suicide, even if they experience increased risk of all-cause mortality.

There are several limitations to the current study. First, aging effects were not taken into account in the STR. Although cohort effects were explored, there was much age heterogeneity within each cohort – a range of 32 years in the young cohort and 39 years in the old cohort. Thus, timing of divorce might be an important factor for individual differences in how divorce influences mortality risk. Second, the longitudinal nature of the STR data was not used in the present study. Mortality risks associated with divorce might decrease with the passage of time, like bereavement (Lichtenstein et al., 1998; Parkes, Benjamin, & Fitzgerald, 1969), suggesting that mortality risk may be time limited. For example, in the current study there may not have been an effect in the within-family mortality curves in the young cohort because divorced twins lived beyond the period where mortality risk was highest. Finally, the findings are based on older Swedish cohorts and may not generalize to younger cohorts (born after the 1960s) or divorced adults in other cultures. Sweden was one of the first European nations to legalize no-fault divorce and transition to a social welfare state in the 1930s. The social acceptance of divorce may have lowered the negative effects of marital dissolution relative to other nations where divorce is more difficult to obtain. Additionally, the relative wealth of Swedish citizens also may mean that marital status in general has less impact on people's health.

The present study makes an important contribution toward fulfilling the research agenda mapped out by Sbarra et al. (2011). First, I conducted the first co-twin-control study of divorce and early death. Replication in other twin samples is, of course, needed to advance knowledge of the putative selection and causal mechanisms underlying divorce and health. Even if replication were to show that divorce and death correlated only because of selection processes, health research could orient toward relevant factors that might lower genetic and social environmental risks that mediate their association.

General Discussion

The findings from the three studies in this dissertation were inconsistent with a causal interpretation of marriage. There was, in fact, some evidence in support of selection processes in each study. Common environmental effects were found to underlie the relation between entering marriage and depressive symptoms in young U.S. American men, as were they found in the relation between becoming widowed and depressive symptoms in older Swedish women. Genetic effects were observed to mediate the association between divorce status and mortality risk in young Swedish men.

As noted in the General Introduction (and again highlighted in Study 1), nearly every study on the marriage benefit is limited by observational (i.e., nonexperimental) study designs. Countless selection factors (i.e., third variables) lurk in the background as possible alternative hypotheses to the social causation hypothesis. The quasi-experimental properties of twin studies help to pinpoint possible causal relations between marital transitions and physical and psychological health where experimental control is not possible (Turkheimer & Harden, 2013). The use of twin samples and twin designs in this dissertation permitted the statistical control of genetic and common environmental selection effects, so that causal hypotheses regarding the effects of marital transitions on depression and mortality could be more precisely tested.

The lack of support for a causal interpretation in this dissertation may not mean marriage has a negligible effect on people's mood and longevity. Marital transitions probably *do* causally influence change in depression and mortality risk, above and

beyond intraindividual selection factors and family-level selection factors. Even in these results, there was some support for a causal interpretation – marrying negatively predicted later depressive symptoms in young adult men, the transition to widowhood positively predicted depressive symptoms in older adult women, and the divorced male twin had a higher risk of early death than his married co-twin – but the sample sizes were too small to detect the small effects with an alpha set at .05. Even more, the good marriage offers for physical and mental health may be dispersed across many life domains in ways not easily captured by a single status indicator.

Marital transitions, however, are not reliable predictors of mental health disparities, as sometimes transitions correlate with health disparities (Burt et al., 2010; Lichtenstein et al., 1996; Osler et al., 2008; Strohschein et al., 2005) and sometimes transitions do not (Hope, Rodgers, & Power, 1999; Horwitz & White, 1998; Wu & Hart, 2002). There are at least a few reasons for why marriage may be an unreliable predictor. Alone, marital status seems to be a poor indicator of social support. The concepts of plasticity and multidirectionality of development (Baltes, 1987) might be one reason for why marital status is a poor index of social support. The circumstances under which marital transitions take place matters quite a bit. Divorce in the context of a happy or unhappy marriage, for example, has been found to affect the course of depression (Hawkins & Booth, 2005; Waite et al., 2009). Whereas divorce may be particularly distressing to those happily married, ridding oneself of a high conflict marriage may alleviate depression (Menaghan, 1985). Similarly, losing a spouse because of cancer or

Alzheimer's disease may have a different impact from losing a spouse suddenly due to a car accident or heart attack.

Additionally, timing of marital change has important implications for whether marriage is found to reliably influence health outcomes. The benefits of marriage, for example, appear to dissipate after the first few years of marriage (Lucas et al., 2003). Likewise, the negative effects divorce and widowhood transitions purportedly has on well-being (Bonanno et al., 2002; Booth & Amato, 1991) and mortality risk (Lichtenstein et al., 1998) have been found to abate with the passage of time. One hypothesis is that people have an emotional set point (Diener, Suh, Lucas, & Smith, 1999; Headey & Wearing, 1989) in the same way that people have a metabolic set point (Speakman et al., 2011). The shocks and jars – whether positive or negative – can moderate risk for depression and mortality, but people generally return to equilibrium over time. An area for future research, then, is to calculate the amount of time between marital transitions (entering marriage in Study 1, becoming widowed in Study 2, and divorce in Study 3) and measurement of depression and mortality to determine the time-limited protection marriage provides people.

Finally, historical and cultural factors may supply important reasons for why marriage is an unreliable predictor of health disparities. As pointed out in the general introduction and Study 1, the institution of marriage underwent considerable change in the 20th century (Cherlin, 2004), with cohabitation becoming more normative (Jose et al., 2010) and possessing fewer differences in health outcomes compared to marriage (Manning & Cohen, 2012). Also, changes in marriage and divorce laws, like the

legalization of no-fault divorce in Sweden and the United States noted in Study 3, may have diminished the health disparities between married and nonmarried groups as nontraditional marital status became more common. Cultural differences, too, may explain why marriage is sometimes correlated with better health but not always. As noted in Study 1, married American citizens receive more rights, privileges, and benefits than married Swedish citizens. The social welfare system in Sweden automatically provides benefits to all citizens, regardless of marital status, whereas the United States reserves benefits to the married majority. Future research would benefit from comparing married male and female twins from old and young cohorts across multiple cultures. While such a study would be a major undertaking, it would be possible with the various twin registries around the world.

Small correlations between marital transition status indicators and health variables, like depression, mortality, and suicide, ought to be regarded as the norm in marital research. In studies where married people were not found to have social well-being advantages over the nonmarried (Shapiro & Keyes, 2008) and cohabiting (Schimmele & Wu, 2011), the paradox between increased social opportunities *and* social isolation among married people, particularly husbands, may explain the lack of advantage associated with marriage. Marrying could be perceived as broadening social support for some while narrowing social support for others. Similarly, widowhood may be a bittersweet experience for caretaking spouses; some may return to “normal” life, spending more time with extended family and friends whereas others may have severe depressive responses to losing one’s spouse. The resiliency literature tends to support

these subgroup differences among bereaved spouses (Bonanno, 2004). From a statistical perspective, the dispersion of experience between marital transitions and health outcomes, like depression and mortality, varies so much in the general population that expected correlations between marriage and health outcomes should be small.

Group heterogeneity may play a large role for why small correlations between marital status, depression, and mortality are often observed. Some people and cohorts may be more affected by marital transitions than others. Mixing subgroups may threaten the validity of health disparities between different marital status subgroups (Nesselroade & Thompson, 1995). Marriage, for example, has been found to have a larger positive effect on depression for those who were most depressed prior to marrying (Frech & Williams, 2007). Similarly, Bonanno et al. (2002) identified several bereavement patterns, with a little more than half the sample regarded as resilient. Finally, only divorced people who perceived little wrong with their marriages prior to divorce were found to be the most distressed while those who viewed their marriages as problematic found divorce to be a relief (Wheaton, 1990). Future research on the relation between marital transitions and depression (and mortality) should identify cohorts and subgroups for which marriage is most protective, even at the cost in power to detect small effects. Nonsignificant effects, for example, may provide good leads for how marital transitions truly affect health.

Marriage benefits cannot be well understood with measures that merely capture marital status. Marriage, widowhood, and divorce are situated within broader social

support contexts, like extended family, friendships, and community networks. The qualities of different relationships inevitably influence how people experience marital transitions and their subsequent health. A potentially more useful research agenda for the inquiry into marital advantages might steer clear of status indicators and use measures of relationship quality, including marital quality, across major life transitions (i.e., marriage, parental transitions, retirement, and bereavement). That one is married probably counts little toward reaping the rewards of marriage. That one has formed supportive, long-lasting relationships in which marriage can be embedded probably counts a lot toward reaping the rewards of marriage.

References

Alexopoulos, G. S. (2005). Depression in the elderly. *Lancet*, *365*(9475), 1961–1970. doi:10.1016/S0140-6736(05)66665-2

Amato, P. R. (2010). Research on divorce: Continuing trends and new developments. *Journal of Marriage and Family*, *72*(3), 650–666. doi:10.1111/j.1741-3737.2010.00723.x

Aneshensel, C. S., & Stone, J. D. (1982). Stress and depression: A test of the buffering model of social support. *Archives of General Psychiatry*, *39*, 1392–1396.

Aseltine, R. H., & Kessler, R. C. (1993). Marital disruption and depression in a community sample. *Journal of Health and Social Behavior*, *34*(3), 237–251.

Baltes, P. B. (1987). Theoretical propositions of life-span developmental psychology: On the dynamics between growth and decline. *Developmental Psychology*, *23*(5), 611–626. doi:10.1037//0012-1649.23.5.611

Beam, C. R. (2008). *The longitudinal relation between marital happiness and positive affect: A trait-state-error approach* (Master's thesis). New York University, New York, NY.

Beam, C. R., Horn, E. E., Hunt, S. K., Emery, R. E., Turkheimer, E., & Martin, N. (2011). Revisiting the effect of marital support on depressive symptoms in mothers and fathers: A genetically informed study. *Journal of Family Psychology*, *25*(3), 336–344. doi:10.1037/a0023758

Beekman, A. T. F., Geerlings, S. W., Deeg, D. J., Smit, J. H., Schoevers, R. S., de Beurs, E., … van Tilburg, W. (2002). The natural history of late-life depression. *Archives of General Psychiatry*, *59*(2002), 605–611.

Bernard, J. (1982). *The future of marriage* (2nd ed.). New Haven: Yale University Press.

Björkenstam, C., Edberg, A., Ayoubi, S., & Rosén, M. (2005). Are cancer patients at higher suicide risk than the general population? *Scandinavian Journal of Public Health*, *33*(3), 208–214. doi:10.1080/14034940410019226

Blazer, D. G. (2003). Depression in late life: Review and commentary. *The Journals of Gerontology. Series A, Biological Sciences and Medical Sciences*, *58*(3), 249–265. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/12634292

Blekesaune, M. (2008). Partnership transitions and mental distress: Investigating temporal order. *Journal of Marriage and Family*, *70*, 879–890.

Bolger, N., Zuckerman, A., & Kessler, R. C. (2000). Invisible support and adjustment to stress. *Journal of Personality and Social Psychology*, *79*(6), 953–961. doi:10.1037/0022-3514.79.6.953

Bonanno, G. A. (2004). Loss, trauma, and human resilience: Have we underestimated the human capacity to thrive after extremely aversive events? *The American Psychologist*, *59*(1), 20–28. doi:10.1037/0003-066X.59.1.20

Bonanno, G. A., Wortman, C. B., Lehman, D. R., Tweed, R. G., Haring, M., Sonnega, J., … Nesse, R. M. (2002). Resilience to loss and chronic grief: A prospective study from preloss to 18-months postloss. *Journal of Personality and Social Psychology*, *83*(5), 1150–1164. doi:10.1037//0022-3514.83.5.1150

Bonanno, G. A., Wortman, C. B., & Nesse, R. M. (2004). Prospective patterns of resilience and maladjustment during widowhood. *Psychology and Aging*, *19*(2), 260–271. doi:10.1037/0882-7974.19.2.260

Booth, A., & Amato, P. R. (1991). Divorce and psychological stress. *Journal of Health and Social Behavior*, *32*(4), 396–407.

Browne, M. W., & Cudeck, R. (1992). Alternative ways of assessing model fit. *Sociological Methods & Research*, *21*(2), 230–258. doi:10.1177/0049124192021002005

Burnham, K. P., & Anderson, D. R. (2004). Multimodel inference: Understanding AIC and BIC in model selection. *Sociological Methods & Research*, *33*(2), 261–304. doi:10.1177/0049124104268644

Burt, S. A., Donnellan, M. B., Humbad, M. N., Hicks, B. M., Mcgue, M., & Iacono, W. G. (2010). Does marriage inhibit antisocial behavior? *Archives of General Psychiatry*, *67*(12), 1309–1315.

Charness, G., & Gneezy, U. (2012). Strong evidence for gender differences in risk taking. *Journal of Economic Behavior & Organization*, *83*(1), 50–58. doi:10.1016/j.jebo.2011.06.007

Cherlin, A. J. (2004). The deinstitutionalization of American marriage. *Journal of Marriage and Family*, *66*, 848–861.

Cheung, Y. B. (2000). Marital status and mortality in British women: A longitudinal study. *International Journal of Epidemiology*, *29*, 93–99.

Choi, H., & Marks, N. F. (2011). Socioeconomic status, marital status continuity and change, marital conflict, and mortality. *Journal of Aging and Health*, *23*(4), 714–742. doi:10.1177/0898264310393339

Clayton, P. J., Halikas, J. A., & Maurice, W. L. (1972). The Depression of Widowhood. *The British Journal of Psychiatry*, *120*(554), 71–77. doi:10.1192/bjp.120.554.71

Cohen, S., Gottlieb, B. H., & Underwood, L. G. (2000). Social relationships and health. In S. Cohen, B. H. Gottlieb, & L. G. Underwood (Eds.), *Social support measurement and interventions: A guide for health and social scientists* (pp. 3–25). New York: Oxford.

Coombs, R. H. (1991). Marital status and personal well-being: A literature review. *Family Relations*, *40*(1), 97–102.

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

Diener, E., & Seligman, M. E. . (2002). Very happy people. *Psychological Science*, *13*(1), 81–84. doi:10.1111/1467-9280.00415

Diener, E., Suh, E. M., Lucas, R. E., & Smith, H. L. (1999). Subjective well-being: Three decades of progress. *Psychological Bulletin*, *125*(2), 276–302.

Duncan, T. E., Duncan, S. C., & Strycker, L. A. (2006). *An introduction to latent variable growth curve modeling: Concepts, issues, and applications*. *Quantitative methodology series* (2nd ed.). Mahwah, NJ: Lawrence Erlbaum Associates.

Emery, R. E. (1999). *Marriage, divorce, and children's adjustment* (2nd ed.). Thousand Oaks, CA: Sage.

Finkel, D., & Pedersen, N. L. (2004). Processing speed and longitudinal trajectories of change for cognitive abilities: The Swedish Adoption/Twin Study of Aging. *Aging, Neuropsychology, and Cognition*, *11*(2-3), 325–345. doi:10.1080/13825580490511152

Frech, A., & Williams, K. (2007). Depression and the psychological benefits of entering marriage. *Journal of Health and Social Behavior*, *48*(2), 149–163. doi:10.1177/002214650704800204

Gentry, M., & Shulman, A. D. (1988). Remarriage as a coping response for widowhood. *Psychology and Aging*, *3*(2), 191–196. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/3268259

Gove, W. R., Hughes, M., & Style, C. B. (1983). Does marriage have positive effects on the psychological well-being of the individual? *Journal of Health and Social Behavior*, *24*(2), 122–131.

Harris, K. M. (2009). The National Longitudinal Study of Adolescent Health (Add Health), Waves I & II, 1994–1996; Wave III, 2001–2002; Wave IV, 2007–2009 [machine-readable data file and documentation]. Chapel Hill, NC: Carolina Population Center, University of North Carolina at Chapel Hill.

Hawkins, D. N., & Booth, A. (2005). Unhappily ever after: Effects of long-term, low-quality marriages on well-being. *Social Forces*, *84*(1), 451–471. doi:10.1353/sof.2005.0103

Headey, B., & Wearing, A. (1989). Personality, life events, and subjective well-being: Toward a dynamic equilibrium model. *Journal of Personality and Social Psychology*, *57*(4), 731–739.

Heath, A. C., Eaves, L. J., & Martin, N. G. (1998). Interaction of marital status and genetic risk for symptoms of depression. *Twin Research*, *1*, 119–122.

Herskind, A. M., McGue, M., Holm, N. V, Sørensen, T. I., Harvald, B., & Vaupel, J. W. (1996). The heritability of human longevity: A population-based study of 2,872 Danish twin pairs born 1870-1900. *Human Genetics*, *97*(3), 319–323. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/8786073

Hertzog, C., Alstine, J. Van, Usala, P. D., Hultsch, D. F., & Dixon, R. (1990). Measurement properties of the Center for Epidemiological Studies Depression Scale (CES-D) in older populations. *Psychological Assessment*, *2*(1), 64–72.

Holmes, T. H., & Rahe, R. H. (1967). The social readjustment rating scale. *Journal of Psychosomatic Research*, *11*, 213–218.

Hope, S., Rodgers, B., & Power, C. (1999). Marital status transitions and psychological distress: Longitudinal evidence from a national population sample. *Psychological Medicine*, *29*(2), 381–389.

Horn, E. E., Xu, Y., Beam, C. R., Turkheimer, E., & Emery, R. E. (2013). Accounting for the physical and mental health benefits of entry into marriage: A genetically informed study of selection and causation. *Journal of Family Psychology*, *27*(1), 30–41. doi:10.1037/a0029803

Horwitz, A. V, & White, H. R. (1991). Becoming married, depression, and alcohol problems among young adults. *Journal of Health and Social Behavior*, *32*(3), 221–237.

Horwitz, A. V, & White, H. R. (1998). The relationship of cohabitation and mental health: A study of a young adult cohort. *Journal of Marriage and Family*, *60*(2), 505–514.

Horwitz, A. V, White, H. R., & Howell-White, S. (1996). Becoming married and mental health: A longitudinal study of a cohort of young adults. *Journal of Marriage and Family*, *58*(4), 895–907.

House, J. S., Landis, K. R., & Umberson, D. (1988). Social relationships and health. *Science*, *241*(4865), 540–545.

Hu, L., & Bentler, P. M. (1995). Evaluating model fit. In R. H. Hoyle (Ed.), *Structural equation modeling: Concepts, issues, and applications*. Thousand Oaks, CA: Sage.

Hu, Y., & Goldman, N. (1990). Mortality differentials by marital status: An international comparison. *Demography*, *27*(2), 233–250.

Johnson, D. R., & Wu, J. (2002). An empirical test of crisis, social selection, and role explanations of the relationship between marital disruption and psychological distress: A pooled time-series analysis of four-wave panel data. *Journal of Marriage and Family*, *64*(1), 211–224.

Jose, A., Daniel O'Leary, K., & Moyer, A. (2010). Does premarital cohabitation predict subsequent marital stability and marital quality? A meta-analysis. *Journal of Marriage and Family*, *72*(1), 105–116. doi:10.1111/j.1741-3737.2009.00686.x

Kaplan, R. M., & Kronick, R. G. (2006). Marital status and longevity in the United States population. *Journal of Epidemiology and Community Health*, *60*(9), 760–765. doi:10. 11 36/jech. 2005.037606

Karney, B. R., & Bradbury, T. N. (1995). The longitudinal course of marital quality and stability: A review of theory, method, and research. *Psychological Bulletin*, *118*(1), 3–34.

Kessler, R. C., & Essex, M. (1982). Marital status and depression: The importance of coping resources. *Social Forces*, *61*(2), 484–507.

Kiecolt-Glaser, J. K., & Newton, T. L. (2001). Marriage and health: His and hers. *Psychological Bulletin*, *127*(4), 472–503. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11439708

Kim, H. K., & McKenry, P. C. (2002). The relationship between marriage and psychological well-being: A longitudinal analysis. *Journal of Family Issues*, *23*(8), 885–911. doi:10.1177/019251302237296

Kline, R. B. (2005). *Principles and practice of structural equation modeling* (2nd Ed.). New York, NY: Guilford Press.

Kposowa, A. J. (2000). Marital status and suicide in the National Longitudinal Mortality Study. *Journal of Epidemiology and Community Health*, *54*(4), 254–261. Retrieved fromhttp://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1731658& tool=pmcentrez&rendertype=abstract

Lamb, K. A., Lee, G. R., & DeMaris, A. (2003). Union formation and depression: Selection and relationship effects. *Journal of Marriage and Family*, *65*, 953–962.

Lee, G. R., & DeMaris, A. (2007). Widowhood, gender, and depression: A longitudinal analysis. *Research on Aging*, *29*(1), 56–72. doi:10.1177/0164027506294098

Lee, G. R., Willetts, M. C., & Seccombe, K. (1998). Widowhood and depression: Gender differences. *Research on Aging*, *20*(5), 611–630. doi:10.1177/0164027598205004

Leonard, K. E., & Mudar, P. (2003). Peer and partner drinking and the transition to marriage: A longitudinal examination of selection and influence processes. *Psychology of Addictive Behaviors*, *17*(2), 115–125. doi:10.1037/0893-164X.17.2.115

Lichtenstein, P., De Faire, U., Floderus, B., Svartengren, M., Svedberg, P., & Pedersen, N. L. (2002). The Swedish Twin Registry: A unique resource for clinical, epidemiological and genetic studies. *Journal of Internal Medicine*, *252*(3), 184–205. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/12270000

Lichtenstein, P., Gatz, M., & Berg, S. (1998). A twin study of mortality after spousal bereavement. *Psychological Medicine*, *28*(3), 635–643. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9626719

Lichtenstein, P., Gatz, M., Pedersen, N. L., Berg, S., & McClearn, G. E. (1996). A co-twin-control study of response to widowhood. *The Journals of Gerontology: Psychological Sciences*, *51B*(5), P279–P289. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/8809004

Lichtenstein, P., Sullivan, P. F., Cnattingius, S., Gatz, M., Johansson, S., Carlström, E., … Pedersen, N. L. (2006). The Swedish Twin Registry in the third millennium: An

update. *Twin Research and Human Genetics*, *9*(6), 875–882. doi:10.1375/183242706779462444

Lillard, L. A., & Waite, L. J. (1995). 'Til death do us part: Marital disruption and mortality. *American Journal of Sociology*, *100*(5), 1131–1156.

Liu, H., & Reczek, C. (2012). Cohabitation and U.S. adult mortality: An examination by gender and race. *Journal of Marriage and Family*, *74*, 794–811. doi:10.1111/j.1741-3737.2012.00983.x

Lorenz, F. O., Wickrama, K. A. S., Conger, R. D., & Elder, G. H. (2006). The short-term and decade-long effects of divorce on women's midlife health. *Journal of Health and Social Behavior*, *47*(2), 111–125. doi:10.1177/002214650604700202

Lucas, R. E., Clark, A. E., Georgellis, Y., & Diener, E. (2003). Reexamining adaptation and the set point model of happiness: Reactions to changes in marital status. *Journal of Personality and Social Psychology*, *84*(3), 527–539. doi:10.1037/0022-3514.84.3.527

Luhmann, M., Hofmann, W., Eid, M., & Lucas, R. E. (2012). Subjective well-being and adaptation to life events: A meta-analysis. *Journal of Personality and Social Psychology*, *102*(3), 592–615. doi:10.1037/a0025948

Lund, D. A., Caserta, M. S., & Dimond, M. F. (1986). Gender differences through two years of bereavement among the elderly. *The Gerontologist*, *26*(3), 314–320.

Manning, W. D., & Cohen, J. A. (2012). Premarital cohabitation and marital dissolution: An examination of recent marriages. *Journal of Marriage and the Family*, *74*(2), 377–387. doi:10.1111/j.1741-3737.2012.00960.x

Manzoli, L., Villari, P., Pirone, G. M., & Boccia, A. (2007). Marital status and mortality in the elderly: A systematic review and meta-analysis. *Social Science & Medicine*, *64*(1), 77–94. doi:10.1016/j.socscimed.2006.08.031

Marenberg, M. E., Risch, N., Berkman, L. F., Floderus, B., & De Faire, U. (1994). Genetic susceptibility to death from coronary heart disease in a study of twins. *The New England Journal of Medicine*, *330*(15), 1041–1046.

Marks, N. F. (1996). Flying solo at midlife: Gender, marital status, and psychological well-being. *Journal of Marriage and Family*, *58*(4), 917–932.

Marks, N. F., & Lambert, J. D. (1998). Marital status continuity and change among young and midlife adults: Longitudinal effects on psychological well-being. *Journal of Family Issues*, *19*(6), 652–686. doi:10.1177/019251398019006001

Martin-Matthews, A. (2011). Revisiting widowhood in later life: Changes in patterns and profiles, advances in research and understanding. *Canadian Journal on Aging*, *30*(3), 339–354. doi:10.1017/S0714980811000201

Mastekaasa, A. (1992). Marriage and psychological well-being: Some evidence on selection into marriage. *Journal of Marriage and Family*, *54*(4), 901–911.

Mastekaasa, A. (1994a). Marital status, distress, and well-being: An international comparison. *Journal of Comparative Family Studies*, *25*(2), 183–205.

Mastekaasa, A. (1994b). Psychological well-being and marital dissolution: Selection effects? *Journal of Family Issues*, *15*(2), 208–228.

Mastekaasa, A. (1995). Marital dissolution and subjective distress: Panel evidence. *European Sociological Review*, *11*(2), 173–185.

Mastekaasa, A. (2006). Is marriage/cohabitation beneficial for young people? Some evidence on psychological distress among Norwegian college students. *Journal of Community & Applied Social Psychology*, *16*(2), 149–165. doi:10.1002/casp.854

McDonald, R. P. (1999). *Test theory: A unified treatment*. Mahwah, NJ: Lawrence Erlbaum Associates.

Mcgue, M., & Lykken, D. T. (1992). Genetic influence on risk of divorce. *Psychological Science*, *3*(6), 368–373.

McGue, M., Vaupel, J. W., Holm, N. V, & Harvald, B. (1993). Longevity is moderately heritable in a sample of Danish twins born 1870-1880. *Journal of Gerontology*, *48*(6), B237–B244. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/8227991

Menaghan, E. G. (1985). Depressive affect and subsequent divorce. *Journal of Family Issues*, *3*(3), 295–306.

Menaghan, E. G., & Lieberman, M. A. (1986). Changes in depression following divorce: A panel study. *Journal of Marriage and Family*, *48*(2), 319–328.

Miller, M., Azrael, D., & Barber, C. (2012). Suicide mortality in the United States: The importance of attending to method in understanding population-level disparities in the burden of suicide. *Annual Review of Public Health*, *33*, 393–408. doi:10.1146/annurev-publhealth-031811-124636

Musick, K., & Bumpass, L. (2012). Re-examining the case for marriage: Union formation and changes in well-being. *Journal of Marriage and the Family*, *74*(1), 1–18. doi:10.1111/j.1741-3737.2011.00873.x

Muthén, L. K., & Muthén, B. O. (1998-2014). Mplus user's guide. Los Angeles, CA: Muthén & Muthén.

Myers, J. K., Lindenthal, J. J., & Pepper, M. P. (1975). Life events, social integration and psychiatric symptomatology. *Journal of Health and Social Behavior*, *16*(4), 421–427.

Neale, M. C., & Cardon, L. R. (1992). *Methodology for genetic studies of twins and families*. Dordrecht: Springer Netherlands. doi:10.1007/978-94-015-8018-2

Nes, R. B., Røysamb, E., Harris, J. R., Czajkowski, N., & Tambs, K. (2010). Mates and marriage matter: Genetic and environmental influences on subjective wellbeing across marital status. *Twin Research and Human Genetics*, *13*(4), 312–321. doi:10.1375/twin.13.4.312

Nesselroade, J. R., & Thompson, W. W. (1995). Selection and related threats to group comparisons: An example comparing factorial structures of higher and lower ability groups of adult twins. *Psychological Bulletin*, *117*(2), 271–284.

Osler, M., McGue, M., Lund, R., & Christensen, K. (2008). Marital status and twins' health and behavior: An analysis of middle-aged Danish twins. *Psychosomatic Medicine*, *70*(4), 482–487. doi:10.1097/PSY.0b013e31816f857b

Panksepp, J. (2005). Why does separation distress hurt? Comment on MacDonald and Leary (2005). *Psychological Bulletin*, *131*(2), 224–230. doi:10.1037/0033-2909.131.2.224

Parkes, C. M., Benjamin, B., & Fitzgerald, R. G. (1969). Broken heart: A statistical study of increased mortality among widowers. *The British Medical Journal*, *1*(5646), 740–743.

Pearlin, L. I., & Johnson, J. S. (1977). Marital status, life-strains and depression. *American Sociological Review*, *42*(5), 704–715.

Plomin, R., & Spinath, F. M. (2004). Intelligence: Genetics, genes, and genomics. *Journal of Personality and Social Psychology*, *86*(1), 112–29. doi:10.1037/0022-3514.86.1.112

Prescott, C. A., & Kendler, K. S. (2001). Associations between marital status and alcohol consumption in a longitudinal study of female twins. *Journal of Studies on Alcohol and Drugs*, *62*, 589–604.

R Core Team. (2013). R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from http://www.r-project.org/

Radloff, L. S. (1977). The CES-D scale: A self-report depression scale for research in the general population. *Applied Psychological Measurement*, *1*(3), 385–401. doi:10.1177/014662167700100306

Robins, J., Greenland, S., & Breslow, N. E. (1986). A general estimator for the variance of the Mantel-Haenszel odds ratio. *American Journal of Epidemiology*, *124*(5), 719–723. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/3766505

Ross, C. E. (1995). Reconceptualizing marital status as a continuum of social attachment. *Journal of Marriage and Family*, *57*(1), 129–140.

Rutter, M., Pickles, A., Murray, R., & Eaves, L. (2001). Testing hypotheses on specific environmental causal effects on behavior. *Psychological Bulletin*, *127*(3), 291–324. doi:10.1037//0033-2909.127.3.291

Satorra, A., & Bentler, P. M. (2001). A scaled difference chi-square test statistic for moment structure analysis. *Psychometrika*, *66*(4), 507–519.

Sbarra, D. A. (2012). Marital dissolution and physical health outcomes: A review of mechanisms. In L. Campbell, J. La Guardia, J. Olson, & M. Zanna (Eds.), *The science of the couple: The Ontario Symposium* (Vol. 12., pp. 205–227). Florence, KY: Psychology Press.

Sbarra, D. A., Law, R. W., & Portley, R. M. (2011). Divorce and death: A meta-analysis and research agenda for clinical, social, and health psychology. *Perspectives on Psychological Science*, *6*(5), 454–474. doi:10.1177/1745691611414724

Sbarra, D. A., & Nietert, P. J. (2009). Divorce and death: Forty years of the Charleston Heart Study. *Psychological Science*, *20*(1), 107–113. doi: 10.1111/j.1467-9280.2008.02252.x

Schimmele, C. M., & Wu, Z. (2011). Cohabitation and social engagement. *Canadian Studies in Population*, *38*, 23–36.

Schrijvers, D. L., Bollen, J., & Sabbe, B. G. C. (2012). The gender paradox in suicidal behavior and its impact on the suicidal process. *Journal of Affective Disorders*, *138*(1-2), 19–26. doi:10.1016/j.jad.2011.03.050

Schwarzbach, M., Luppa, M., Forstmeier, S., König, H.-H., & Riedel-Heller, S. G. (2013). Social relations and depression in late life – A systematic review. *International Journal of Geriatric Psychiatry*, *29*(1), 1–21. doi: 10.1002/gps.3971

Schwarzbach, M., Luppa, M., Sikorski, C., Fuchs, A., Maier, W., Bussche, H. Van Den, … Riedel-heller, S. G. (2013). The relationship between social integration and depression in non-demented primary care patients aged 75 years and older. *Journal of Affective Disorders*, *145*(2), 172–178. doi:10.1016/j.jad.2012.07.025

Shadish, W., Cook, T., & Campbell, D. (2002). Experiments and generalized causal inference. In W. R. Shadish, T. D. Cook, & D. T. Campbell (Eds.), *Experimental and quasi-experimental designs for generalized causal influence* (pp. 1–32). Boston, MA: Houghton Mifflin.

Shafer, A. B. (2006). Meta-analysis of the factor structures of four depression questionnaires: Beck, CES-D, Hamilton, and Zung. *Journal of Clinical Psychology*, *62*(1), 123–146. doi:10.1002/jclp

Shapiro, A., & Keyes, C. L. M. (2008). Marital status and social well-being: Are the married always better off? *Social Indicators Research*, *88*(2), 329–346. doi:10.1007/s11205-007-9194-3

Shrout, P. E. (2002). Reliability. In M. T. Tsuang & M. Tohen (Eds.), *Textbook in psychiatric epidemiology* (2nd ed.), pp. 131–147). New York, NY: John Wiley & Sons, Inc.

Shrout, P. E., Herman, C. M., & Bolger, N. (2006). The costs and benefits of practical and emotional support on adjustment: A daily diary study of couples experiencing

acute stress. *Personal Relationships*, *13*(1), 115–134. doi:10.1111/j.1475-6811.2006.00108.x

Silman, A. J., Newman, J., & MacGregor, A. J. (1996). Cigarette smoking increases the risk of rheumatoid arthritis: Results from a nationwide study of disease-discordant twins. *Arthritis and Rheumatism*, *39*(5), 732–735. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/8639169

Simon, R. W., & Marcussen, K. (1999). Marital transitions, marital beliefs, and mental health. *Journal of Health and Social Behavior*, *40*(2), 111–125. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/10467758

Simonsson, P., & Sandstrom, G. (2011). Ready, willing, and able to divorce: An economic and cultural history of divorce in twentieth-century Sweden. *Journal of Family History*, *36*(2), 210–229. doi:10.1177/0363199010395853

Speakman, J. R., Levitsky, D. A, Allison, D. B., Bray, M. S., de Castro, J. M., Clegg, D. J., … Westerterp-Plantenga, M. S. (2011). Set points, settling points and some alternative models: Theoretical options to understand how genes and environments combine to regulate body adiposity. *Disease Models & Mechanisms*, *4*(6), 733–745. doi:10.1242/dmm.008698

Stanton, S. J., Liening, S. H., & Schultheiss, O. C. (2011). Testosterone is positively associated with risk taking in the Iowa Gambling Task. *Hormones and Behavior*, *59*(2), 252–256. doi:10.1016/j.yhbeh.2010.12.003

Strohschein, L., McDonough, P., Monette, G., & Shao, Q. (2005). Marital transitions and mental health: Are there gender differences in the short-term effects of marital

status change? *Social Science & Medicine*, *61*(11), 2293–2303. doi:10.1016/j.socscimed.2005.07.020

Therneau, T. M. (2013). A package for survival analysis in S. Retrieved from URL: http://CRAN.R-project.org/package=survival

Thoits, P. A. (1984). Explaining distributions of psychological vulnerability: Lack of social support in the face of life stress. *Social Forces*, *63*(2), 453–481.

Tucker, J. S., Friedman, H. S., Wingard, D. L., & Schwartz, J. E. (1996). Marital history at midlife as a predictor of longevity: Alternative explanations to the protective effect of marriage. *Health Psychology*, *15*(2), 94–101. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/8681925

Turkheimer, E., & Harden, K. P. (2013). Behavior genetic research methods: Testing quasi-causal hypotheses using multivariate twin data. In H.T. Reis & C.M. Judd (Eds.), *Handbook of research methods in personality and social psychology* (2nd ed.) (pp. 159–187). New York, NY: Cambridge University Press

United States v. Windsor, 570 U.S. 12-307. (2013).

Van Poppel, F., & Joung, I. (2001). Long-term trends in marital status mortality differences in The Netherlands 1850-1970. *Journal of Biosocial Science*, *33*(2), 279–303. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11284632

Verbeek, M. (1990). On the estimation of a fixed effects model with selectivity bias. *Economics Letters*, *34*(3), 267–270. doi:10.1016/0165-1765(90)90129-O

Verbeek, M., & Nijman, T. (1992). Testing for selectivity bias in panel data models. *International Economic Review*, *33*(3), 681–703.

Wade, T. J., & Cairney, J. (2000). Major depressive disorder and marital transition among mothers: Results from a national panel study. *The Journal of Nervous and Mental Disease*, *188*(11), 741–750. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11093376

Wade, T. J., & Pevalin, D. J. (2004). Marital transitions and mental health. *Journal of Health and Social Behavior*, *45*(2), 155–170.

Waite, L. J., & Gallagher, M. (2000). *The case for marriage: Why married people are happier, healthier, and better off financially*. New York: Doubleday.

Waite, L. J., Luo, Y., & Lewin, A. C. (2009). Marital happiness and marital stability: Consequences for psychological well-being. *Social Science Research*, *38*(1), 201–212. doi:10.1016/j.ssresearch.2008.07.001

Weissman, M. M. (1987). Advances in psychiatric epidemiology: Rates and risks for major depression. *American Journal of Public Health*, *77*(4), 445–451. doi: 10.2105/AJPH.77.4.445

Wheaton, B. (1990). Life transitions, role histories, and mental health. *American Sociological Review*, *55*(2), 209–223.

Willett, J. B., & Sayer, A. G. (1994). Using covariance structure analysis to detect correlates and predictors of individual change over time. *Psychological Bulletin*, *116*(2), 363–381.

Williams, D. R., Takeuchi, D. T., & Adair, R. K. (1992). Marital status and psychiatric disorders among blacks and whites. *Journal of Health and Social Behavior*, *33*(2), 140–157.

Williams, K., & Umberson, D. (2004). Marital status, marital transitions, and health: A gendered life course perspective. *Journal of Health and Social Behavior*, *45*(1), 81–98. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/15179909

Wu, Z., & Hart, R. (2002). The effects of marital and nonmarital union transition on health. *Journal of Marriage and Family*, *64*(2), 420–432.

Yip, P. S. F., Chen, Y.-Y., Yousuf, S., Lee, C. K. M., Kawano, K., Routley, V., … Wu, K. C.-C. (2012). Towards a reassessment of the role of divorce in suicide outcomes: Evidence from five pacific rim populations. *Social Science & Medicine*, *75*(2), 358–366. doi:10.1016/j.socscimed.2012.03.009

Zisook, S., Shuchter, S. R., Irwin, M., Darko, D. F., Sledge, P., & Resovsky, K. (1994). Bereavement, depression, and immune function. *Psychiatry Research*, *52*(1), 1–10. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/8047615

Appendix A

Center for Epidemiological Studies Depression (CES-D) Scale

Below is a list of some of the ways you may have felt or behaved. Please indicate how often you have felt this way during the past week: (circle one number on each line)

During the past week...	Rarely or none of the time (less than 1 day)	Some or a little of the time (1-2 days)	Occasionally or a moderate amount of time (3-4 days)	All of the time (5-7 days)
1. I was bothered by things that usually don't bother me	0	1	2	3
2. I did not feel like eating; my appetite was poor	0	1	2	3
3. I felt that I could not shake off the blues even with help from my family	0	1	2	3
4. I felt that I was just as good as other people	0	1	2	3
5. I had trouble keeping my mind on what I was doing	0	1	2	3
6. I felt depressed	0	1	2	3
7. I felt that everything I did was an effort	0	1	2	3
8. I felt hopeful about the future	0	1	2	3
9. I thought my life had been a failure	0	1	2	3
10. I felt fearful	0	1	2	3
11. My sleep was restless	0	1	2	3
12. I was happy	0	1	2	3
13. I talked less than usual	0	1	2	3
14. I felt lonely	0	1	2	3
15. People were unfriendly	0	1	2	3
16. I enjoyed life	0	1	2	3
17. I had crying spells	0	1	2	3
18. I felt sad	0	1	2	3
19. I felt that people disliked me	0	1	2	3
20. I could not "get going"	0	1	2	3

Notes: All items used at each wave in SATSA. ***Bolded*** *items used at each wave in Add Health.*

Appendix B

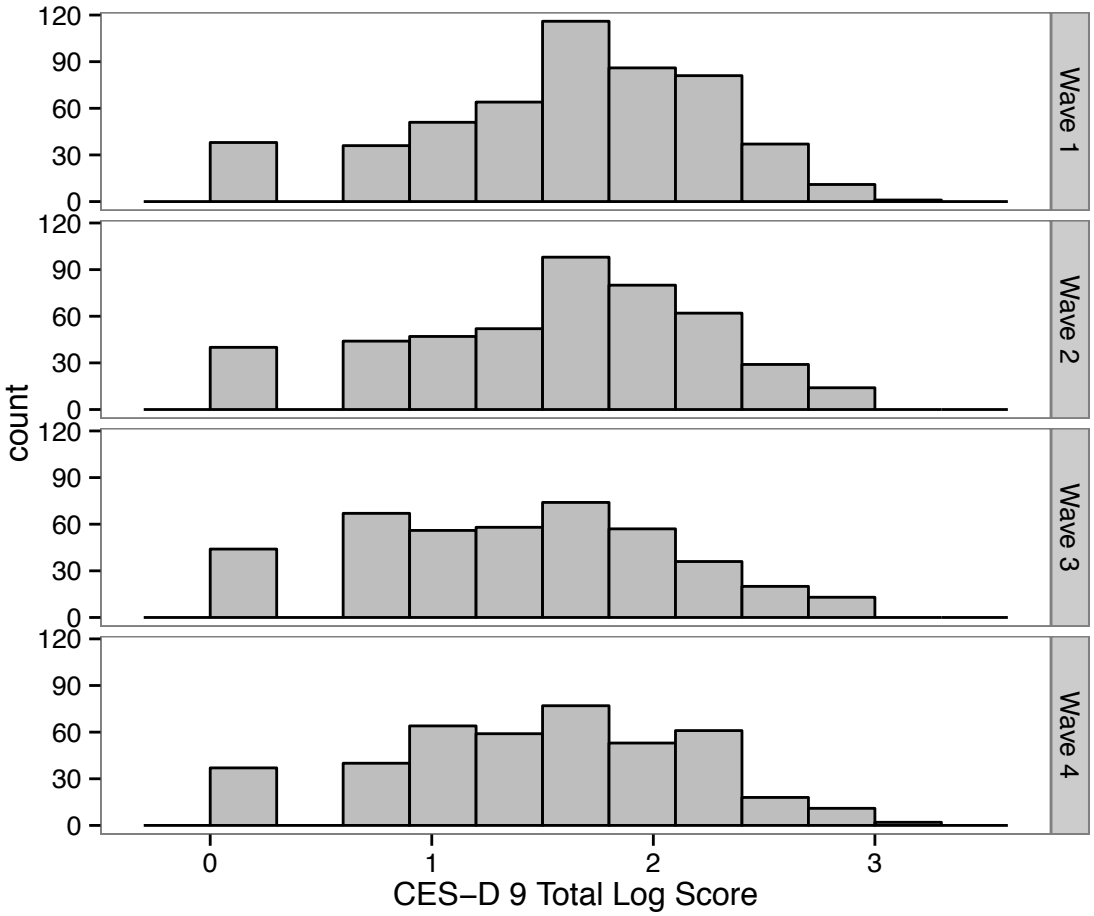

Figure 1. Histogram of Add Health male log total CES-D-9 scores

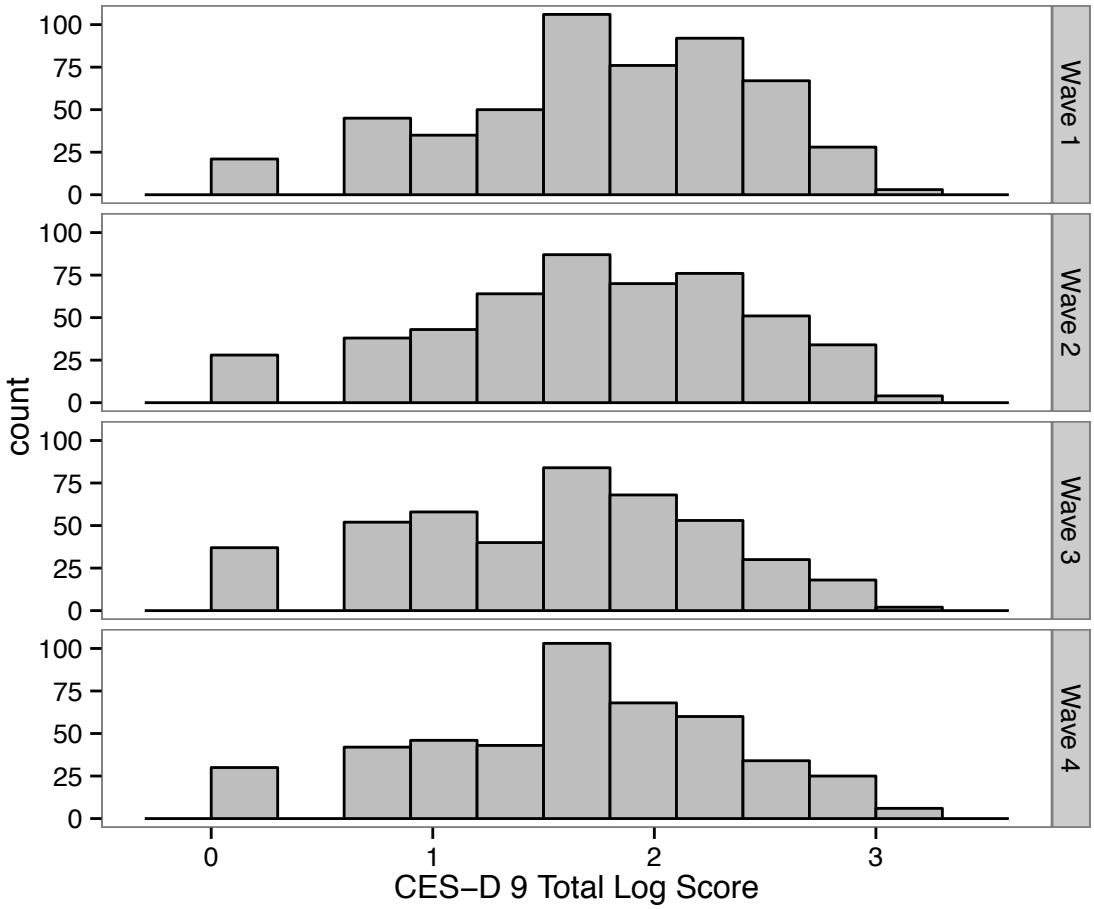

Figure 2. Histogram of Add Health female log total CES-D-9 scores.

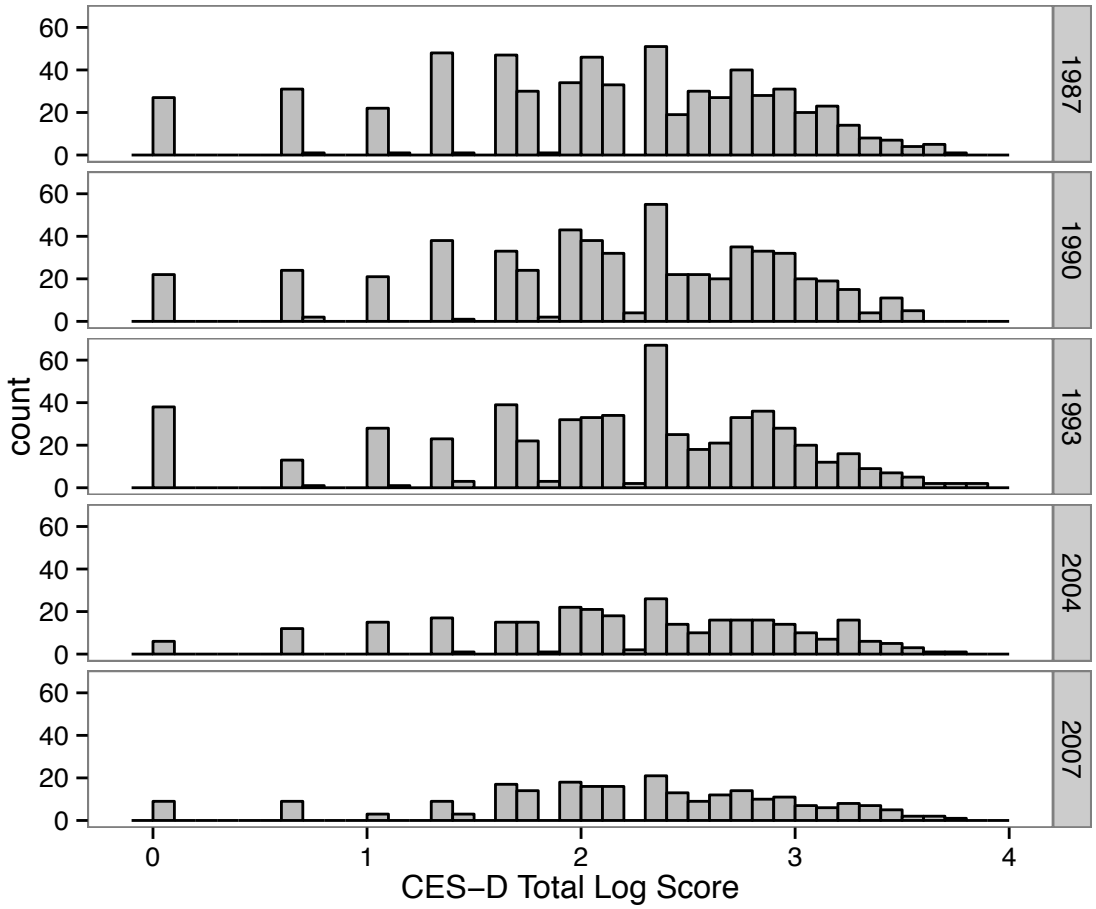

Figure 3. Histogram of SATSA male log total CES-D scores.

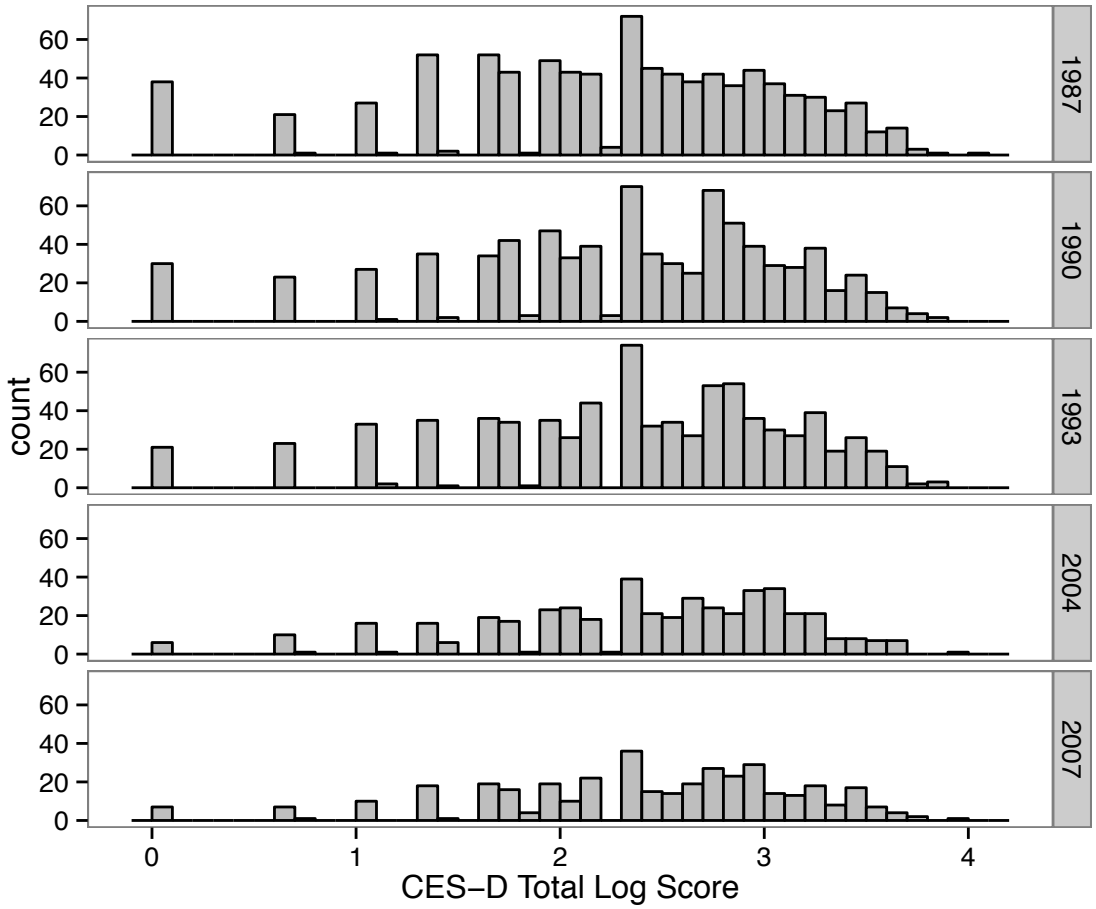

Figure 4. Histogram of SATSA female log total CES-D scores.