

**What Do Social Scientists Know about the Benefits of Marriage?
A Review of Quantitative Methodologies**

**David C. Ribar
Department of Economics
The George Washington University
2201 G Street, NW
Washington, DC 20052
e-mail: dcr7@gwu.edu**

September 2003

What Do Social Scientists Know about the Benefits of Marriage?
A Review of Quantitative Methodologies

Table of Contents

Acknowledgements.....	iii
Executive Summary.....	iv
I. Introduction.....	1
II. Conceptual Models of the Effects of Family Structure.....	3
Children’s well-being.....	4
Adult’s economic and material well-being.....	6
Adults’ physical and mental health.....	8
Marriage as a decision.....	9
Implications for empirical analyses.....	9
III. Statistical Methods for Examining the Effects of Family Structure.....	11
Cross-section methods.....	12
Longitudinal/panel methods.....	17
More complicated descriptions of family structure.....	21
Nonlinear models.....	21
Summary.....	23
IV. Empirical Studies of the Effects of Family Structure.....	23
Children’s well-being.....	24
Adults’ earnings.....	38
Adults’ physical health and mortality.....	48
V. Conclusion.....	58
What the research from each domain can contribute to the others.....	58
Appendix A. Overview of Statistical Techniques for More Complicated Models.....	62
More complicated descriptions of family structure.....	62
Nonlinear models.....	64
References.....	68

Acknowledgements

This review was prepared while the author was a research analyst with the Office of Planning, Research and Evaluation in the Administration for Children and Families of the U.S. Department of Health and Human Services. The author thanks Sung Un Kim for research assistance and for preparing an initial draft of section IV.2. He also thanks Joseph Grubbs for helpful advice and substantive discussions with initial drafts. Other staff at the ACF, including Naomi Goldstein, Brendan Kelly and Howard Rolston, also provided useful suggestions. The author is indebted to several colleagues, including Bob Lerman, Daniel Lichter, David Loughran and Donna Ruane Morrison, who generously supplied references to innovative articles and studies along with helpful comments. Preliminary drafts of this report were presented at the 2003 ACF Annual Research Conference and the 2003 Annual Congress of the European Society of Population Economics. The views expressed are the author's own and are not necessarily shared by the U.S. Department of Health and Human Services.

Executive Summary

Marriage is positively associated with a large number of outcomes including improved cognitive, emotional and physical well-being for children, better mental and physical health for adults, and greater earnings and consumption for family members. While the associations between marriage and various measures of well-being have been convincingly established, they do not, by themselves, make a compelling case that marriage has beneficial effects. As with many other types of social science data, the empirical relationships are likely to be confounded by problems of reverse causality and spurious correlation from omitted variables. Because of this, we cannot be sure whether the observed relationships reflect marriage making people better off, better-off people being more likely to marry, or some combination of the two. These issues have long been recognized by researchers. Some researchers have simply acknowledged the problem and interpreted their results accordingly, while others have tried to address the problem statistically.

This study critically reviews quantitative methods that have been employed and evidence that has been gathered to assess the benefits of marriage and consequences of other family structures. The study begins by describing theoretical models of the determinants of different well-being outcomes and the role of family structure in producing those outcomes. It also discusses models of the determinants of marriage. The study then overviews specific statistical techniques that have been applied in empirical analyses of the effects of marriage, including standard regression, instrumental variables, selection and switching models, matching, non-parametric bounds, fixed effects, and latent factor (correlated random effects) methods. The study then reviews selected studies that have been completed in three domains of well-being outcomes: children's well-being, adults' earnings, and adults' physical health.

Theories. Theoretical models are important because they form the lens through which researchers view the data and make causal interpretations. They can also alert us to potential empirical problems. The study considers economic, rational choice models of the determinants of children's well-being, adults' economic success, and adults' physical health and augments these models to incorporate hypotheses and insights from other social science disciplines.

To examine children's well-being, the study adopts a household production model in which parents combine inputs of goods, services and their own time to produce beneficial outcomes for children. In this model, marriage can affect children's well-being by increasing the financial and time resources available within a household. Marriage may also change the way that inputs are combined so that they are used more effectively. Beyond the household production model, marriage may improve children's well-being by reducing instability and stress or by providing a favorable environment to socialize children.

Five hypotheses are offered to explain why marriage may affect adults' earnings. The first is that marriage allows spouses to concentrate on and become more productive in activities in which each has a relative advantage. The higher productivity for spouses who specialize in market work would lead to higher earnings. A second hypothesis is that a spouse may provide instrumental support that increases the other's productivity or augments his or her career. The third hypothesis is that marriage is a stabilizing or maturing influence, which leads to better and

more consistent individual work habits. The fourth hypothesis is that married people look for different amenities and disamenities in their jobs and receive different compensation as a result. The final hypothesis is that the earnings differences reflect discrimination by employers.

For adults' physical health, the study considers a variant of the household production model called the health production model. In the health production model, adults combine inputs of goods and time to produce health outcomes for themselves and other household members. The implications of a change in family structure are similar to those from the model for children's well-being. Marriage increases the resources to adults and possibly the productivity of those resources. Marriage may also reduce stress, allow spouses to monitor each others' behavior, or change individual health habits.

The study next considers the reasons why marriages are formed or maintained. In an economic model, potential spouses compare their expected valuations of the economic, social, and health outcomes associated with entering into or continuing a marriage with those of remaining single or divorcing. Marriages occur or continue if the perceived value of marriage exceeds that of the alternative. Couples who face good prospects within marriage are likely to marry, while couples who face bad prospects are not. This provides a rationale for researchers' concerns regarding selectivity—namely, that well-being outcomes could drive marriage outcomes rather than the other way around.

Statistical methods. Standard regression and discrete-choice models are commonly used by empirical researchers. These models specify well-being as an outcome and family structure and other measures as explanatory variables. The models rely on an assumption that family structure and the other observed explanatory variables are not related to any unobserved determinants of well-being. This assumption will be violated and the estimated impact of family structure will be biased if family structure is misreported, affected by well-being, or influenced by other factors that also affect well-being—that is, if there is measurement error, reverse causality, or relevant omitted variables.

To address problems associated with omitted variables, researchers often include direct and indirect controls for these variables. The strategy is sensible but is only successful if the researcher knows which variables are missing and can find the corresponding measures. The strategy does not address biases that arise from reverse causality.

Researchers have also considered special circumstances, or *natural experiments*, in which marriages might be thought to occur or break up independently of other well-being outcomes. Analyses of these situations can be useful if the assumptions regarding the circumstances are correct and if the people who experience them are representative of the general population. However, it is very difficult to find situations that meet these requirements.

Instrumental variables estimators, which rely on the researcher's ability to identify situations that alter the chances that marriages will be formed or dissolved independently of well-being outcomes, are closely related to natural experiments comparisons. These estimators use variables that are related to family structure but are otherwise unrelated to well-being to predict living arrangements. The estimators are helpful because they address problems associated with reverse causality and omitted variables. A practical difficulty arises, however, in coming up with

suitable instruments. There are also problems with the technique if the effects of family structure vary across people and if these differences affect people's family formation decisions.

Matching methods use data on observable characteristics to form comparable groups of married and unmarried individuals. These methods do not require strong modeling assumptions about how the observable characteristics contribute to selection and do not require variable exclusion restrictions. Moreover, the approach is easy to explain to non-statisticians and improves transparency because the researcher shows exactly how comparisons are being made. The disadvantages of the approach are that it requires comparable observations and does not account for selection based on unobservable characteristics.

An alternative strategy is to specify an empirical model for the marriage decision and its relationship to well-being, that is, specify a model for the selection process. This approach addresses selectivity associated with unobserved characteristics but requires some assumptions regarding how the unobservable characteristics are distributed.

Interval estimates, or nonparametric bounds, of the effects of family structure can be generated without any assumptions regarding the distribution of the unobserved components. However, the intervals are generally too wide to be of much use to policymakers.

Fixed-effects methods can be used when observations are available over time for an individual or across individuals within a group. These methods mitigate biases that arise from omitted variables that are common across the observations, such as a permanent characteristic in longitudinal data for individuals or a shared trait in data for a group. Researchers using these techniques do not need to specify which variables are omitted; they only need to describe the properties of the variables. The techniques do not address biases that arise from other error structures and can exacerbate biases associated with reverse causality and measurement error. It is also difficult to apply fixed effects controls in non-linear specifications, like probit models.

Correlated random effects can be used in situations where fixed effects controls are appropriate but impractical. The correlated random effects procedures require stronger assumptions than the fixed effects procedures but can be applied in non-linear models, such as probit and survival models.

Review of Empirical Studies. *Children's well-being.* A vast number of studies have examined the relationship between family structure and children's well-being. While this research has generally found that marriage is associated with better outcomes for children under most circumstances, the evidence is based mainly on analyses that failed to account for selectivity. Selectivity appears to be more than a hypothetical concern. Comparisons across regression specifications indicate that the addition of controls for family background and circumstances reduces the association between marriage and children's outcomes. The measured association also falls when longitudinal data are employed to account for the characteristics of children before they experience family disruptions. Some quantitative research on children's outcomes has moved beyond standard regression analyses and longitudinal comparisons. For instance, research has shown that children's well-being is negatively associated with a more permissive divorce environment, a plausible instrument for family structure. The findings from

these studies provide more convincing evidence of a causal link between family structure and children's well-being.

Adults' earnings. Numerous studies, including studies with statistical controls for selectivity, have found that marriage is positively associated with men's earnings. The evidence regarding women's earnings is less clear with studies finding positive and negative associations as well as no associations at all. Many of the studies have also concluded that selectivity is a substantive concern. While controls for selectivity have not generally eliminated the association between marriage and earnings, they have reduced the size of the estimated relationships. The remaining associations are consistent with possible causal effects.

Adults' physical health and mortality. Studies have consistently found that being married and being in a satisfying marriage are positively associated with health and negatively associated with mortality. The introduction of longitudinal data has not overturned these long-standing findings. These results notwithstanding, the existing body of research on the effects of family structure on physical health and mortality has to be judged to be methodologically weak. Although several recent studies have adopted more sophisticated statistical methods like hazard, longitudinal and structural equations models, research based on simple descriptive analyses continues to be published in peer-reviewed journals. Even when advanced techniques are used, the studies in this area generally do not address issues associated with selection and omitted variables bias.

The research on marital interactions, as opposed to marriage status and family structure, contains more solid evidence of a causal link to health outcomes. This research, which has included some experimental designs, has been able to show that interactions affect stress. Although these experimental results have not been convincingly tied to larger health outcomes, they suggest that programs that improve relationship skills and marital interactions might improve health.

Conclusions. While there is evidence throughout the literature that marriage is associated with positive outcomes for adults and children, most of it comes from regression analyses that account for observed covariates but do not account for other sources of selectivity. More and more evidence is coming, however, from studies that employ selectivity controls, such as longitudinal comparisons, instrumental variables methods, switching regression models and matching methods. Researchers have generally found that the use of such controls reduces the association between marriage and well-being, which validates initial concerns regarding selectivity, but does not eliminate the association, which is consistent with causal explanations.

The study draws other conclusions regarding how research in each of the areas examined can borrow from strengths found in the others.

Statistical methods. Statistical methods for addressing selectivity have been applied in each of the three domains that were reviewed. However, the techniques have been used more consistently in the research on adults' earnings and less consistently in the other two areas. Research on children's well-being and adults' health and mortality could be strengthened by a more careful consideration of the selectivity issue. This not only means adopting statistical

techniques from other fields but also testing, where possible, for the assumptions underlying the techniques as well as for differences in results across methods.

Instrumental variables. Instrumental variables methods are attractive to researchers because they can address bias from several sources including omitted variables, reverse causality, and measurement error. Their application has been limited, however, by the practical problem of finding suitable variables that are correlated with marriage but otherwise unrelated to well-being outcomes. Results from several studies have been undercut by questionable choices regarding instruments. The most promising source of instruments for current research on outcomes in the United States appears to be state laws and policies. Experimental evaluations that will soon be undertaken by the U.S. Department of Health and Human Services of programs to strengthen marriages and prepare couples to enter healthy marriages may yield additional sets of instrument conditions.

Relationship quality. Numerous studies have found that children's well-being and adults' health and mortality are strongly related to both the existence and quality of a marital relationship. Considerations of relationship quality, which have been almost entirely absent from the research on earnings, should be studied more carefully in this literature. Marital quality is also a dimension along which some degree of experimental manipulation is possible; thus, it may be possible to sidestep selection issues altogether in this type of research.

Direct and indirect effects. Much of the research on the association between family structure and well-being has included controls for mediating factors such as economic resources, work skills, and parenting practices. This research has generally found that the addition of these variables leads to smaller estimates of the direct association between family structure and well-being. While the approach has been valuable for examining the sensitivity of family structure to alternative specification assumptions, researchers have not been as careful in considering this aspect of their empirical models as others. Many of the mediators that have been examined are behaviorally determined and therefore endogenous; future empirical analyses need to account for this. Researchers have also frequently overlooked the total association between family structure and well-being, which would include the direct and indirect associations. An understanding of the total association is important for policy purposes. It could also lead to a larger estimate of the possible effects of marriage.

What Do Social Scientists Know about the Benefits of Marriage? A Review of Quantitative Methodologies

I. Introduction

Marriage is positively associated with a large number of outcomes including improved cognitive, emotional and physical well-being for children, better mental and physical health for adults, and greater earnings and consumption for family members. These associations have been documented in hundreds of quantitative studies covering different time periods and different countries. The studies date back to at least the mid-nineteenth century, when William Farr reported an inverse relationship between marriage and mortality in France (referenced in Goldman 1993). More recent evidence on the benefits of marriage has been summarized in a number of influential books, reports and scholarly presentations including those by Glenn et al. (2002), McLanahan and Sandefur (1994), Waite (1995) and Waite and Gallagher (2000).

While the associations between marriage and various measures of well-being have been convincingly established, the associations do not, by themselves, make a compelling case that marriage has beneficial effects. As with many other types of social science data, the empirical relationships are likely to be confounded by problems of reverse causality and spurious correlation from omitted variables. Because of these problems, we cannot be sure whether the observed relationships reflect marriage making people better off, better-off people being more likely to marry, or some combination of the two. These issues have long been recognized by researchers. For example, Farr tempered the interpretation of his findings by commenting that society encouraged marriage among healthy people and discouraged it among the less able (see Goldman 1993, p. 190). Nearly every subsequent study of the possible effects of marriage has either acknowledged or attempted to overcome these difficulties. The implications of reverse causality and omitted variables continue to be hotly debated (Cherlin 1999).

This study critically reviews quantitative methods that have been employed and evidence that has been gathered to assess the benefits of marriage and consequences of other family structures. The study begins by describing theoretical models of the determinants of different well-being outcomes and the role of family structure in producing those outcomes. It also discusses models of the determinants of marriage. Theoretical models are important because they form the lens through which researchers view the data and make causal interpretations. Theoretical analyses alert us to potential empirical problems. In this case, the discussion shows how problems associated with omitted variables and selection into marriage arise. The models also help us to evaluate the trade-offs that are inherent in different estimation methods.

The study then overviews specific statistical techniques. Because of the ubiquity of the selection problem, there are already several comprehensive reviews of technical approaches (e.g., Card 1999; Heckman et al. 1999; Heckman and Robb 1995; Manski 1995) and epistemological issues (e.g., Marini and Singer 1988) associated with drawing causal inferences from observational data. Accordingly, the study provides relatively brief descriptions of the techniques and limits its focus to approaches that have actually been applied in empirical analyses of the effects of marriage. These include standard regression, instrumental variables,

selection and switching models, matching, non-parametric bounds, fixed effects, and latent factor (correlated random effects) methods. The study considers the properties of these techniques in the canonical situation of a two-way comparison between married and unmarried people. It then considers their applicability to comparisons involving finer categorizations of family structure and analyses involving detailed marital and relationship histories.

With this theoretical and statistical background in hand, the study reviews selected studies that have been conducted in three domains of well-being outcomes: children's well-being, adults' earnings, and adults' physical health and mortality. It must be stated from the outset that the domains and selected studies encompass only a miniscule part of the vast body of research on the outcomes associated with marriage and family structure. The choice of domains and studies is guided by a desire to illustrate the range of methodologies that have been and are being applied by quantitative researchers. Selection and omitted variables problems are present and recognized in each of the chosen domains; nevertheless, the literatures differ widely in terms of their current practices for addressing these issues. In particular, the studies of the effects of marriage on men's and women's earnings have been the most sensitive to selectivity and unobserved heterogeneity and incorporated the most sophisticated techniques. In contrast, studies examining adults' health outcomes have only recently begun to account for these problems, while studies of children's well-being fall somewhere in between.

This study is intended to serve several purposes. First and most importantly, it summarizes and evaluates research evidence regarding the benefits of marriage focusing on the methods used to control for selectivity and omitted variables. Previous reviews, such as the book by Waite and Gallagher (2000), have evaluated research on the basis of the populations that were examined and the observed controls that were included—crucial methodological elements to be sure. However, they have not carefully considered the implications of different selection controls. While no single, fool-proof methodological strategy exists for addressing the selection issue in observational research, many alternative methods are available. Some of the assumptions of these methods can be tested, and the sensitivity of results can be compared. If nothing else, analyzing the differences in methods and results gives us insight into the amount of uncertainty in the research literature.

Second, the review brings together research from several different social science disciplines. An advantage of a cross-cutting research topic like the benefits of marriage is that it draws researchers with different perspectives and different tools. A disadvantage, however, is that it can be difficult to communicate across disciplines; so, advances from one field are not necessarily transmitted to the others. The inclination of researchers to specialize in narrow sub-areas further reduces communication. A methodological review can bridge these gaps by informing researchers of important developments and useful techniques in related areas.

Third, new developments in data, methods, and software mean that the time is ripe to reconsider some approaches. With respect to data, new longitudinal surveys, such as the Fragile Families and Child Wellbeing Study and the National Longitudinal Survey of Youth 1997 cohort, are providing fresh information on family processes over time. Other on-going surveys, such as the Panel Study of Income Dynamics (PSID) and the Survey of Income and Program Participation, have added new detailed modules for child and family well-being. Statistical

methods, such as kernel and propensity-score matching, have been introduced into the social sciences. Some other types of statistical controls, such as correlated random effects and discrete-factor approximations for unobserved heterogeneity, are now incorporated into easy-to-use software packages. These developments provide new opportunities to test theories, unpack previous findings, and tackle problems associated with selectivity.

Finally, there is intense public interest in the potential benefits of marriage. The promotion of healthy marriages is a centerpiece of President Bush's welfare reauthorization plan. The Department of Health and Human Services has committed millions of dollars to a series of novel interventions and demonstration projects to test approaches for stabilizing and strengthening marriages among disadvantaged families. Several state governments are also pursuing their own initiatives, such as establishing "covenant marriages," eliminating disincentives to married parents' participation in cash assistance programs, and supporting marriage counseling and relationship training (Gardiner et al. 2002). Non-profit organizations like Marriage Savers and the Coalition for Marriage, Family and Couples Education have also been active in promoting marriage. Positive research findings regarding marriage motivate much of this advocacy. A clearer understanding of the theory and evidence underlying these results is important to evaluating marriage policies.

II. Conceptual Models of the Effects of Family Structure

A conceptual analysis of the possible effects of family structure begins with theoretical models of how each outcome is produced. Unfortunately, because of the differences in the outcomes, it is not reasonable to use a simple, generalized model that treats all of the outcomes as manifestations of a single, global indicator of household or family well-being. Moreover, because of the number of outcomes, it is not feasible to work with a single model at all. Accordingly, the review begins with separate models broadly grouped into three domains: children's general well-being, adults' earnings, and adults' physical health and mortality.

Each of the outcomes crosses several disciplinary boundaries including psychology, sociology, biology, demography and economics. These disciplines, in turn, have their own distinct conceptual approaches with attendant strengths and weaknesses. Where possible, this review frames its analyses in terms of the economic approach but modifies the economic approach to incorporate other perspectives.

The distinguishing feature of the economic approach is its assumption of *rational behavior*. By rational behavior, we mean that adults make conscious decisions that maximize the fulfillment of their perceived needs and desires subject to constraints on their time, budgets, information, and other resources. Simply stated, rational choice models assume that people do the best they can with the resources they have. People do not necessarily make the "right decisions" but are assumed to do what they feel is right for them.

When applied to simple consumer choices such as deciding how much bread or milk to buy at the supermarket, the rational-choice approach seems reasonable. However, when applied to topics outside those traditionally studied by economists such as marriage behavior or

children's development, the approach may appear less suitable. What, after all, can the economic approach say about a couple's views of commitment in marriage or a parent's ability to provide children with a warm but structured environment? Quite a lot, it turns out. First, there are some purely economic elements that bear on each of these outcomes such as the tax implications of marriage and the expense of childrearing. Second, by categorizing the determinants of personal and family outcomes into factors associated with objectives and constraints, the economic approach provides a convenient organizing framework. It is generally easier to incorporate other disciplinary perspectives into the economic approach than vice versa. As a result, rational choice models have been increasingly adopted by sociologists and family researchers (Hechter and Kanazawa 1997). Finally, even if we are uncomfortable with the underlying assumption of rationality, what alternative would we offer to describe *general* patterns of behavior? To borrow from Winston Churchill's famous remark on democracy, rational choice is the worst assumption except for all those other assumptions that have been tried from time to time.

Children's well-being

The theoretical workhorse for economists examining children's physical, material, emotional, and cognitive well-being is Gary Becker's (1965) time allocation, or household production, model.¹ The essential features of Becker's model, as they apply to children's outcomes, are that

- (a) the decision-makers in a household value a number of outcomes including children's well-being,
- (b) the enjoyment and production of those outcomes requires purchases of goods and services and contributions of the household members' time,
- (c) the household members face constraints on the uses of their time, and
- (d) the members face constraints in their available prices, wages and non-labor income.

The model recognizes that parents can combine and substitute purchases of goods and services with contributions of their own time to produce children's well-being. In this respect, the model borrows from economic theories of business behavior. Just as firms have production functions and choose levels of labor and intermediate goods to produce final goods, households also have production functions and choose levels of time, goods, and services to produce certain outputs. For instance, to develop pre-reading skills in young children, parents could provide their own materials and home instruction, which would require some purchases but also a substantial investment of time. Alternatively, they could place their children in pre-schools or engage tutors, which would involve greater out-of-pocket expense but less of their own time. Other combinations of inputs are also possible that would lead to the same level of skills.

The household production framework neatly accommodates several sociological perspectives. For example, the active socialization of children can be viewed as a production process in which parents contribute inputs of time and possibly goods or services to produce developmental outcomes. Similarly, social control and monitoring of children's behavior can be viewed as processes that involve inputs of time and goods.

¹ Haveman and Wolfe (1995) provide a comprehensive overview of the household production model and other non-economic approaches for analyzing children's socioeconomic attainments.

The household production model can also be extended to incorporate insights from the developmental psychology literature regarding the continuity of development and the timing of critical events (Chase-Lansdale 1998). Continuity refers to the intertemporal dependence of children's developmental outcomes and trajectories. A child's physical, emotional, and cognitive development in one period set the stage for development in the next period. Once a child begins along a particular developmental trajectory, he or she is likely to continue along that trajectory. Applying this to the household production model, well-being outcomes in a given period not only depend on contemporaneous inputs of goods and time but also on the entire history of outcomes and, hence, previous inputs of goods and time. Another important consideration is timing, or the notion that certain inputs matter more at different points of childhood than others. The production functions that characterize outcomes in early childhood are not the same as those for middle childhood, adolescence, or young adulthood.

It is straightforward to show how different family structures can affect children's well-being in such a model. Compare the circumstances of a child who grows up with both parents to those of a child who lives with only one parent and has no contact with or support from the other parent. An obvious difference lies in the resources that are available in each environment. The two-parent household would have more time than the one-parent household to allocate between the production of child well-being and other activities, such as paid employment. The two-parent household would also have greater financial resources. The advantages in time and money are likely to lead to better outcomes for the child. Note that these advantages stem from the presence of an additional adult—any adult—and the resources he or she brings. The explanation does not distinguish between adults who are married, cohabiting, or part of some other relationship such as a three-generation household. Indeed, the reasoning suggests that a single parent with sufficient resources can produce the same developmental outcomes as two parents.

We can also modify the household production function so that family structure itself and changes in family structure matter directly. For instance, the absence of a parent or conflicted relationships between parents might provide examples that children observe, internalize and subsequently emulate. Alternatively, disruptions in the family could be a source of instability or stress with negative consequences for children (Wu and Martinson 1993). Besides directly affecting children's outcomes, family structure could also change the shape of the production function, altering the way in which time and goods inputs are transformed into well-being. Evidence that parenting styles and behaviors differ between married and unmarried parents (Astone and McLanahan 1991) can be read as supporting this contention.

Along with the resource and stability advantages, we can consider other characteristics of marriage that might operate through the household production model. Waite and Gallagher (2000) summarize evidence that marriage affects parents' ability to function by allowing them to specialize in work or household production and making them physically and mentally healthier. These types of changes would increase the marginal effectiveness of a given set of time and goods inputs, leading to better results for children. Marriage may also facilitate cooperation among parents that mitigates public goods problems associated with children's well-being (Weiss and Willis 1985). If parents each value their children's well-being, investments in children by one parent not only benefit that parent but the other as well. In the absence of

coordination, each parent faces incentives to under-invest in children's well-being—that is, to “free-ride” off of the other's contributions.

Marriage might also affect the way in which household decisions are made. Standard time allocation models assume that a household acts to maximize a common or shared set of goals. More recently, analysts have considered cooperative and non-cooperative bargaining models that recognize that individuals in a household have different objectives (see, e.g., Bergstrom's 1996 review). In these models, the incentives that married and unmarried parents face affect how they reconcile their differences. If parents differ in their relative valuation of children's well-being, changes in bargaining strength may affect children's outcomes.

While these theories suggest numerous ways that marriage might benefit children, we must also acknowledge that there might be negative effects under some circumstances. For instance, children who grow up in married stepfamilies do not fare as well on average as children who grow up in two-parent, continuously married families. From a purely economic perspective, this evidence is puzzling because married stepfamilies should bring many of the same advantages as other married families—greater resources, higher household productivity, etc. Negative results might arise because stepfamily arrangements involve more instability (at least one additional change in family structure) than other arrangements. They might also arise if step-parents value stepchildren's well-being less than biological parents. The lower valuation would lead to less investment in children and less cooperation within marriage.

Outcomes for children also appear to be worse in married families with high amounts of parental conflict, both relative to married families with less conflict and unmarried families. Conflict between parents would reduce their cooperation in household and economic tasks. It might also interfere with their ability to perform these tasks or be a direct source of stress to children.

Finally, the findings of Wu and Martinson (1993) and Wu (1996) indicate that transitions themselves may be harmful. Thus, there could be consequences if two parents decided to marry sometime after the birth their child or children. Evidence of children's difficulties in adjusting to the return of their fathers from military service (see the review in Herzog and Sudia 1973) also suggests that transitions into two-parent households may not be uniformly positive.

Adult's economic and material well-being

On average, married couples enjoy larger incomes, greater wealth, and higher living standards than do single adults. Differences in living standards are the easiest to explain. Living together introduces economies of scale in household production—it takes fewer resources to produce outcomes for the second person in a household than for the first. A couple living together can get by with fewer goods than two people living on their own (one residence instead of two, one set of appliances, etc.). It also takes less time to produce the second and subsequent units of some outputs, such as the second serving of a meal or second set of clean clothes, than the first. Economies of scale provide a powerful incentive for people to join together to form a single household. These advantages increase with people's willingness and ability to share

resources. To the extent that marriage facilitates greater sharing than other arrangements, married couples would benefit more from economies of scale than cohabiting couples.

The more intriguing statistic is that average incomes and financial holdings in households headed by married couples are greater than the sum of the averages for unmarried men and unmarried women. Hourly wages for married men are also generally higher than those for unmarried men. Waite and Gallagher (2000) cite numerous studies that indicate that these economic premia reflect more than just pre-existing differences in individuals' economic abilities—that is, higher rates of marriage among high-wage than low-wage people. It is not clear whether these earnings advantages extend to women. Some researchers have found insignificant or small negative associations between marriage and earnings for women (Hewitt, Western and Baxter 2002; Hill 1979; Korenman and Neumark 1992), while others have found small positive associations (Budig and England 2001; Waldfogel 1997).

Five hypotheses have been offered to explain these wage patterns. The first is primarily economic and involves specialization of activities. Marriage allows spouses to concentrate on activities in which each has a relative advantage. One spouse can specialize in non-market, household activities while the other specializes in market work. Specialization, by itself, does not explain gender roles within marriage. However, the theory suggests that small, initial differences in relative abilities or circumstances can lead to complete specialization. Thus, if women are brought up to have a slight advantage in household production or alternatively, if childbearing places women at a temporary disadvantage in market labor, there could be profound gender differences in specialized activities. As people specialize in home or market production, they may increase their skills through experience or training and become even more productive in these activities over time, amplifying the initial differences in earnings.

The second hypothesis is that a spouse provides instrumental support that increases the other spouse's productivity or augments his or her career (Bellas 1992, Daniel 1995). For instance, one spouse might help the other to prepare a resumé, make a work-related decision, or entertain co-workers or clients. If these activities reduce the time available for the helper's own career, there could be a bifurcated outcome in which one spouse—most likely the one with higher initial earnings—receives a premium while the other suffers a penalty.

The third hypothesis also involves productivity but is more sociological and focuses on marriage as a stabilizing or maturing influence, especially for men. Married men are much less likely to abuse drugs and alcohol than single men. They also are more likely to stay home at night, keep out of trouble, and show up for work in the morning. There are several reasons why marriage may prompt these changes. First, married men may value their discretionary home time more highly than single men, if for no other reason than that they have someone at home to share time with. Second, married men may feel a greater sense of responsibility regarding their family's well-being and be more concerned about family members' feelings. Either of these motivations would lead married men to be more other-oriented and, thus, more considerate of the impacts that their actions have on others. Third, marriage, which entails a lifetime commitment, may make men more future-oriented and, thus, more sensitive to the long-term consequences of their actions. Finally, the effects may simply reflect increased scrutiny and monitoring of husbands' behavior by wives. Note that because women engage in fewer risky and damaging behaviors to begin with, there is less scope for positive changes within marriage for them.

The fourth hypothesis is that there are amenities or disamenities associated with particular jobs that result in compensating differentials in pay. Married men may feel a financial responsibility that makes them more willing than unmarried men to work in jobs with undesirable characteristics, like inflexible hours, a crowded office, or stressful conditions. The opposite might be true of married women, who might desire flexible hours or accommodating schedules more strongly than unmarried women. The critical assumption in this hypothesis is that preferences regarding work conditions differ by both marital status and gender.

The fifth hypothesis is that the earnings patterns reflect discrimination. Business owners and managers, who are disproportionately married and male, might favor people who are like themselves or be sympathetic to men who are supporting a family. This might lead to greater economic opportunities for married men. At the same time, employers with “traditional” beliefs might be antagonistic toward married women or mothers and less willing to hire or pay them. A shortcoming of the discrimination hypothesis is that it presupposes that employers have certain preferences and beliefs. Another weakness is that earnings differences unrelated to productivity or hiring costs are tough to sustain in a competitive economy. By paying unnecessary premiums and failing to hire the most productive employees, prejudiced employers place themselves at a competitive disadvantage relative to other employers (see Cain’s 1986 discussion). Nevertheless, earnings differences can appear if discriminating employers have market power or if market frictions limit workers’ abilities to seek out non-discriminators.

Adults’ physical and mental health

To examine health outcomes, economists rely on a model developed by Grossman (1972) that combines elements of Becker’s theories regarding household production and human capital. Like the household production model, Grossman’s model recognizes that health outcomes are produced with inputs of goods, such as medicines, nutritious foods, and physician services, and inputs of time, such as time spent exercising, monitoring symptoms, or visiting a doctor. Like the human capital model (Becker 1993), Grossman’s approach also recognizes that current investments in health have long-term benefits and that health depreciates over time.

The implications of Grossman’s “health capital” model are similar to those of the extended household production model that we used to examine children’s well-being. Because goods and services contribute to the production of health, the higher incomes of married couples put them at an advantage relative to single adults. Married couples would also benefit because there would be a greater chance that one of the adults had access to a job with health insurance. Similarly, the increased time available in a married household would allow for greater investments in health. Beyond these resource explanations, specialization might allow married families to become more productive with the time and goods that they have or allow a spouse to search more intensively for a job with insurance benefits.

Also like the analysis of children’s outcomes, marriage could have other direct, non-economic impacts on health. For example, the stability associated with marriage and the attendant reduction in stress might directly benefit adults’ emotional and physical well-being (Gove 1973, Kiecolt-Glaser 2001). Alternatively, the simple knowledge that there is another adult who cares and is concerned about you could provide a sense of security and psychological

well-being. Our earlier discussion of marriage's effect on unhealthy behaviors including drinking, drug use, and irregular hours would also be relevant (Umberson 1987).

Negative health effects are also possible. Of particular concern is the increased opportunity for domestic violence against women. A wife's view of the commitment associated with marriage, her specialization in household activities, and the possible depreciation of her market skills over time, might make it difficult to leave an abusive relationship. Other health consequences could occur because of the specialization decision itself. A spouse who specializes in home production gives up at least part of a career. Besides the feelings of loss, specialization could increase isolation from other adults or leave the person feeling vulnerable. These feelings in turn could contribute to emotional or psychological distress. Waite and Gallagher (2000) point out that violence and emotional health problems for single or cohabiting adults may be worse; nevertheless, they acknowledge evidence that marriage may produce bifurcated health outcomes in which husbands gain and wives lose.

Marriage as a decision

The foregoing discussion reveals that there are a number of avenues by which marriage might affect children's and adults' well-being. A critical concern for empirical research, however, is that marriage itself is a decision, which may in turn be influenced by expectations or experiences regarding well-being. The simplest economic models of these decisions assume that potential spouses compare their expected valuations of the economic, social, and health outcomes associated with entering into or continuing a marriage with those of remaining single or divorcing. Marriages occur or continue if the perceived value of marriage exceeds that of the alternative. This means that couples who face good prospects within marriage are likely to marry, while couples who face bad prospects are not. If this is the case, well-being outcomes could well drive marriage outcomes rather than the other way around.

Consider the evidence that married men earn more than unmarried men. Is this a consequence of marriage? That is, does marriage make men more productive? The hypotheses regarding specialization and stabilizing influences within marriage imply that marriage affects earnings. However, an alternative explanation for the statistical association between marriage and earnings is that women are simply more willing to accept marriage proposals from successful men. Yet another explanation is that there are characteristics of economically successful men, such as disciplined personal habits and strong interpersonal skills, that make them better candidates for marriage. Under this explanation, marriage and earnings do not directly affect one another but instead are affected by one or more additional variables.

Implications for empirical analyses

As the discussion in this section has shown, statistical associations between marriage and well-being outcomes could occur because

- (a) marriage affects well-being for children and adults,
- (b) well-being for children and adults affects marriage, or
- (c) other variables affect both marriage and well-being.

Most of the empirical research on the consequences of family structure (effect a) has been sensitive to the issues of reverse causality (effect b) and omitted variables bias and multiple causation (effect c). This sensitivity has been manifested in several ways. Some research simply acknowledges that the statistical problems exist and interprets findings in that light. These studies are usually careful to report their findings in terms of associations rather than effects. Other research has employed alternative statistical techniques to control for biases from reverse causality and omitted variables. Unfortunately, these techniques require additional data or assumptions and can be quite sensitive to departures from those requirements. Indeed, it is not unusual for researchers looking at the same question and facing the same statistical problems to employ different empirical techniques and generate conflicting findings.

The conceptual models have other implications for empirical analyses of well-being outcomes. All of the models include marriage as one among many potential determinants of personal or family well-being. In each model, marriage is embedded in a more general theory of how well-being is achieved. These models, in turn, yield important insights regarding how empirical analyses should be specified. For instance, the models point to other variables, besides marriage, that should be included in the empirical analyses. If we consider the example of men's earnings, the theoretical model indicates that productivity is a key determinant and that marriage affects earnings through productivity. This implies that an empirical analysis of earnings should not only control for marriage but also for other indicators of productivity that preceded marriage, such as education and work experience.

A related implication of these models is that marriage may generate indirect effects—that is, work through other variables—as well as direct effects. Because of this, researchers must be careful not to over-control for other variables. Let us continue with the earnings example and again assume that marriage affects earnings through productivity. Consider an extreme example in which current productivity (a) is the only direct determinant of earnings and (b) is perfectly measured and accounted for in the empirical analysis. Under these conditions, we would expect to observe no additional, independent effect of marriage. The absence of a direct effect might be incorrectly interpreted as evidence of no effect at all.

Finally, by indicating whether and how particular variables enter an empirical relationship, theoretical analyses help us to implement and justify more sophisticated and comprehensive statistical investigations. For instance, instrumental variable methods require variable exclusion restrictions (e.g., finding variables that directly affect marriage but only affect well-being through marriage). Fixed-effect techniques require strong assumptions on the unobserved variables, or error terms, in regressions. The assumptions necessary for these approaches may be reasonable under some circumstances but not under others. The theoretical analysis assists in identifying those circumstances. More generally, theory assists in identifying hypothetical conditions under which the effects of marriage should and should not be operative. Relying on a given theory, an empirical analysis can look for the existence of a relationship between marriage and well-being when it is expected to occur and for the absence of a relationship elsewhere. Similarly, if marriage is theorized to work in conjunction with some other condition to produce an outcome, the empirical analysis can test for this pattern. A careful theoretical analysis should generate several sharp (refutable) predictions. When such an analysis is followed by a comprehensive empirical investigation that tests for and supports all of the

predictions, we can have greater confidence that we have uncovered a causal explanation (Freeman 1991; Marini and Singer 1988). Freeman (1991) has described the ideal process as good detective work supplemented with lots of “shoe leather.”

III. Statistical Methods for Examining the Effects of Family Structure

This section reviews empirical methods that have been used to evaluate the effects of marriage and other types of family arrangements on children’s and adults’ well-being. The review is somewhat abstract and stylized. Much of the discussion is framed around an arbitrary well-being outcome, Y , which could represent any of the social, economic, health or development outcomes that were discussed in the previous section. The objective is to introduce various statistical techniques, state the conditions under which they produce unbiased or consistent results, and discuss the implications when these conditions are not met. Because the discussion concerns relatively sophisticated statistical methods and their properties, some technical detail is unavoidable. Nevertheless, the review tries to steer clear of highly technical discussions and maintain a general tone wherever possible.² The next section in this report provides more context by considering specific applications of these techniques.

The review begins by categorizing family structure into two, mutually-exclusive outcomes: marriage and non-marriage. This simplification is adopted partly to help the discussion along, as it is easier to describe statistical techniques that account for fewer outcomes than many. The simplification also serves a methodological purpose. Models that compare outcomes across two types of outcomes fall into a widely-used class of models called *treatment effect* models. Because of their broad applicability, statisticians have devoted enormous effort to developing methods to estimate these models and studying their properties. These particular methods have, in turn, found their way into numerous studies of the effects of marriage. Once it has discussed techniques for estimating models with two comparison groups, the review briefly considers complications that arise in models that include more elaborate descriptions of family structure; a more detailed treatment of these complications is provided in Appendix A.

Some of the empirical techniques require special data. The review initially considers methods that can be applied when micro-level, cross-section data on well-being are available—that is, when the data describe well-being outcomes for many people or households at a single point in time. An example of this type of data is the information that is available from a single survey that asks respondents about their current health or income and their current marital status. The review then discusses techniques that can be implemented using micro-level, longitudinal data on well-being outcomes—that is, data at several points in time for the same individuals or households. Examples of longitudinal data include surveys with repeated follow-ups, such as the National Longitudinal Survey of Youth and Panel Study of Income Dynamics, and certain types of administrative data such as school records and welfare caseload files.³

² Readers can find more detailed descriptions of these techniques in the reviews by Card (1999), Greene (2003), Heckman, LaLonde and Smith (1999), and Heckman and Robb (1985).

³ Heckman, LaLonde and Smith (1999) discuss methods for these types of micro-level data as well as other types such as time series (one person or household followed over different time periods) and repeated cross-sections (data on different people at different time periods). They also review methods that can be applied to aggregated data.

Cross-section methods

Consider a data set that consists of observations on well-being, marital status, and other measures for N individuals (or households). Let Y_i be a continuous variable that represents well-being for the i th person in the sample ($i = 1, N$); let M_i be a binary variable that describes the person's marital status (the variable equals one if the person is married and zero otherwise); and let \mathbf{X}_i be a vector of variables that represent other observed characteristics, such as the person's age and ethnicity. The goal of much of the research on the effects of marriage has been to estimate equations of the form

$$Y_i = \alpha M_i + \mathbf{B}'\mathbf{X}_i + \varepsilon_i \quad (1)$$

where ε_i is a variable that represents unobserved characteristics and α and \mathbf{B} are coefficients to be estimated. When this equation is estimated using Ordinary Least Squares (OLS) regression, the estimate of α captures the difference in the average well-being of married and unmarried people holding the other observed characteristics constant. For instance, if we regressed measures of people's health status on indicators for their marital status, age, and ethnicity, the coefficient on marriage would represent the average difference in health status for married and unmarried people net of the influences of age and ethnicity.

The estimated coefficient on marital status is unbiased if the model is correctly specified and the explanatory variables are uncorrelated with the error term ε_i .⁴ These, of course, are two strong assumptions. With respect to specification, theory guides us in the selection of the estimating model but rarely provides us with an exact functional form. Many studies adopt linear specifications like equation (1) for reasons of convenience and as a rough approximation to the true relationship. Depending on the theoretical framework, however, the specification may be inappropriate. For instance, equation (1), which implies that the effect of marriage is uniform across individuals, would be inconsistent with theories of family formation that emphasize differences in the benefits of marriage.

Although both assumptions are critical, quantitative researchers have generally been more concerned with the second assumption regarding the independence of the explanatory variables and error term. The independence, or exogeneity, assumption is valid in experimental settings in which researchers randomly assign treatments to some subjects and withhold them from others. For obvious ethical reasons, however, marital status cannot be assigned this way. Even with observational data, the independence assumption would still be valid if marriage were a purely random process, like winning the lottery or flipping a coin. While there is undoubtedly some random variation in the incidence of marriage, the theoretical analysis tells us that there is also a purposeful component. The theoretical model predicts that actual or expected well-being will enter into marriage decisions. It also indicates that several factors related to well-being, such as people's family backgrounds, attitudes and abilities, will affect marriage.

⁴ Recall that the estimated coefficient is a function of the other observed variables. If these are random variables—that is, variables with numerical outcomes and probability distributions defined over those outcomes—the estimator will also be a random variable with a probability (sampling) distribution. The estimator is unbiased if its expected value (the mean of the sampling distribution) is equal to the true value of the coefficient.

We can consider how reverse causality and omitted relevant variables each lead to a violation of the independence assumption.⁵ Assume that marital status is an *endogenous* outcome—that is, an outcome that depends on a set of observed and unobserved characteristics, just as well-being does. If the two outcomes are characterized by reverse causality, well-being is an explanatory variable for marriage. This implies that marriage is related to all of the determinants of well-being, including the error term. In the case of omitted relevant variables, problems arise if the empirical analysis leaves out one or more determinants of well-being which are also determinants of marriage. The omitted characteristics effectively become part of the error term in the estimating equation for well-being. Because the characteristics also affect marriage, there is a correlation with the error term. Researchers are generally aware of these explanations and understand that reverse causality and omitted variables bias the results from OLS regressions. Researchers sometimes, however, fail to distinguish between the sources of bias. This can be a crucial mistake because methods that address one source of bias do not necessarily address the other.

Adding variables. If the estimating equation omits a variable that is related to both well-being and marriage, the most straightforward solution is to add the variable to the analysis. While the solution is obvious, it is seldom practical because the appropriate measures either do not exist or were not collected as part of the survey or data set. Lacking direct measures for relevant variables, researchers sometimes include indirect measures, or proxy variables. This approach has the potential to reduce bias. However, to the extent that indirect measures are imperfect and capture only some of the variation in the omitted characteristics, biases will remain. Of course, both approaches—using direct measures or proxy variables—require that the researcher know precisely which variables have been omitted. This may be unrealistic when considering outcomes like people’s physical or economic well-being that are thought to have numerous causes. The approaches also do not address biases associated with reverse causality.

Concerns about omitted variables lead many researchers to adopt a “better safe than sorry” strategy in which they include lots of variables that can serve as potential indirect controls. While there is some logic to this approach, there are also downsides. First, even with a large set of controls, researchers can never be sure that they have accounted for every confounding influence. Because there is no natural stopping point, the strategy can get out of hand and mutate into an approach of “throwing in everything but the kitchen sink.” Second, unnecessary variables reduce the precision of the coefficient estimates resulting in larger standard errors and fewer significant results. Finally, the injudicious use of controls increases the chances of multicollinearity and of over-accounting for effects.

Even when it is done properly, adding extra control variables into a regression model can complicate the interpretation of results. Consider again the hypothetical analysis of adults’ earnings that included measures of marital status and productivity as explanatory variables. The coefficient on marital status in such a model would provide an estimate of the *direct association* between marital status and earnings—the association holding productivity and the other explanatory variables constant. However, it seems reasonable to suppose that there might also be

⁵ Mismeasurement of the marital status variable would also lead to a correlation with the error term and bias. Korenman, Goldman and Fu (1997) provide a powerful example, in the context of older women’s mortality, of the prevalence of misclassification of marital status in surveys and its effects on estimation results.

an *indirect association* that occurs through an effect of marriage on productivity. The *total association*, which consists of the direct and all indirect associations, could be large even if the direct association is small or non-existent. As they interpret estimation results, analysts need to consider exactly what is being held constant and take care not to overlook indirect associations.

Circumstances without selectivity. Researchers have considered special circumstances in which marriages might occur or break up independently of other well-being outcomes. These types of situations are sometimes referred to as “natural experiments.” Analyses of these situations can be useful if the assumptions regarding the circumstances are correct and if the people who experience them are representative of the general population. As with analyses involving genuine experimental data, analyses of natural experiments do not require special estimation procedures. They are also easy to describe to policymakers. The only drawback is that the situations are hard to find.

Some researchers including Ginther and Zavodny (2001) and Lerman (2002d) have argued that marriages that immediately follow a premarital pregnancy constitute a natural experiment. These marriages are sometimes described as “shot-gun” marriages—presumably because of the methods that a pregnant bride’s family would use to persuade a reluctant groom to come to the altar. While these types of marriages (and persuasive techniques) have become less common over time, a fair number of “rushed” marriages do still occur. Analysts contend that these marriages meet the requirements of a natural experiment because they are less anticipated and therefore less selective than other types of marriages. The argument, however, is not entirely convincing. The pregnancies may not be unanticipated, and even if they are, the subsequent marriages might still be selective. Some pregnancies occur among people who had already planned to marry. Also, premarital pregnancies occur disproportionately among young and less-educated couples. Finally, the marriages themselves involve an element of choice.

Other changes in family structure, such as parental deaths (Corak 2001; Lang and Zagorsky 2001) and temporary absences of soldiers and sailors (see the review by Herzog and Sudia 1973), have also been considered as natural experiments. Like shot-gun marriages, the appropriateness of using these events in comparisons is debatable.

Instrumental variables. Instrumental variables estimators, which rely on the researcher’s ability to identify situations that alter the *chances* that marriages will be formed or dissolved independently of well-being outcomes, are closely related to natural experiments comparisons. As an illustration, consider state or local requirements, like mandatory waiting periods or blood tests, that increase the cost or inconvenience of getting married but are not likely to be otherwise related to well-being. If the requirements indeed have these properties, the only way that they could affect well-being is through their effect on marriage. Thus, if we were to find that marriage and well-being were lower in areas that adopted such requirements than in areas that did not, we might conclude that the differences represent an effect of marriage.

More formally, an instrumental variable, z_i , is a measure, like one of the local requirements discussed above, that predicts marriage but is not related (conditional on the other observed variables) to ε_i . The method essentially involves replacing actual data on marital status in the estimating model for well-being with predicted values, which are based on the instrument z_i and other \mathbf{X}_i variables. The predicted value is obviously related to marriage. However,

because it is constructed from variables that are conditionally unrelated to ε_i , the prediction itself is conditionally unrelated, and the coefficient estimate on this variable is consistent.⁶

Instrumental variable methods are attractive because they address problems associated with reverse causality and omitted variables. The methods are also familiar to quantitative researchers and included in many statistical packages. A practical difficulty arises, however, in coming up with suitable instruments. As mentioned, instruments for estimating the impact of marriage on well-being must first be strong predictors of marriage and second be unrelated to well-being except through marriage. Few variables meet both of these criteria. With respect to local marriage restrictions, we might not detect any effects on marriage if few people are close enough to the decision margin where the inconvenience of a waiting period or blood test matters. Even if these policies do have measurable impacts on marriage, they might only be enacted in communities with particular economic and cultural characteristics or, worse, as a result of concerns about local marriage and well-being trends.⁷

Card (1999) and Heckman, LaLonde and Smith (1999) point out that instruments can also fail when there are differences across people in the effects of an event, like marriage, which subsequently affect people's decision-making. To appreciate their argument, consider a case in which there are exogenous differences across areas in marriage restrictions. In areas with burdensome restrictions, only people who foresee large gains in well-being will marry, while in areas with fewer restrictions, people who foresee smaller gains will marry. In this way, the size of the effect on well-being varies systematically with the otherwise exogenous costs of marriage.

Matching techniques to form comparison groups. One way to view the problem that researchers face in trying to estimate an effect of marriage is that married and unmarried people are not entirely comparable. That is, their well-being outcomes are likely to differ even in the absence of marriage. Ideally, researchers would like to observe and compare each person's opportunities inside and outside of marriage at a given point in time. However, because a person can never simultaneously occupy both of these states, analysts are left to conjecture what the counterfactual condition would be.

Matching methods use data on observable characteristics to match married and unmarried individuals as closely as possible and, thus, enhance comparability. Matching is advantageous for several reasons. First, the technique is easy for non-statisticians to grasp; the transparency of

⁶ Because the predicted value has less variability than the actual value, the standard errors of the coefficient estimates have to be adjusted (see Greene 2003).

⁷ To address the practical problems in selecting instruments, Gennetian, Bos and Morris (2002) have suggested using randomized treatments from social experiments, such as the recently completed welfare reform experiments. We can consider their strategy in the specific context of marriage. Recall that randomization breaks the link between the treatment and other observed and unobserved characteristics. If the treatment affects the incentives for marriage but does not affect other outcomes, it would meet the requirements for an instrument. While this approach seems promising, it would only be appropriate under a limited set of conditions. First, the experiment would not only have to include marriage as a component, but probably include it as the *only component*. An exception would be an experiment that included multi-stage randomization. Second, the experiment could not include direct economic subsidies for marriage because these would affect income and, hence, well-being. Similarly, it could not operate through any other mechanism that might independently affect well-being. The general point is that randomization is not sufficient by itself to identify and unpack the behavioral processes that lead to different outcomes in a social experiment (Heckman 1996).

the method can be useful in policy discussions and presentations. Second, while all methods for assessing the effects of marriage effectively involve comparing outcomes across groups, matching forces the researcher to be explicit about the comparisons. Third, the method does not require strong modeling assumptions about how the observable characteristics contribute to selection. Fourth, the method does not require variable exclusion restrictions; so, its data requirements are similar to those of the standard regression approach. Fifth, it can accommodate variation in the effects of marriage.

At the same time, there are also some drawbacks. First, while the method is conceptually straightforward, it can be difficult to establish and choose among close matches. Exact matches of the observable characteristics are preferred; however, if there are many characteristics or numerous possible outcomes for each characteristic, there will be few, if any, exact matches in a reasonably-sized sample. Mahalanobis metric (Rubin 1979), propensity score (Rosenbaum and Rubin 1983) and kernel methods (Heckman et al. 1999) have been developed to overcome the dimensionality problem when multiple characteristics are considered. Even if the dimensionality problem is addressed, the method may still break down if there are no matches for certain ranges of the observed characteristics. This can occur if there is only one observation in a given range or if everyone in the range is either married or unmarried.

A more fundamental shortcoming of the matching approach is that it controls for selection on the observed characteristics but not the unobserved characteristics of individuals. This latter type of selection would occur if people considering marriage act upon information on well-being that is not observable to the researcher. Similar problems can occur if the researcher either omits or does not have access to all of the relevant observed characteristics.

Modeling the selection process. Another strategy is to specify an empirical model for the marriage decision and its relationship to well-being. In particular, let marriage be a binary, endogenous outcome with observed and unobserved determinants. A dummy endogenous variable model (Heckman 1978) would combine this specification with equation (1) and allow the unobserved determinants of marriage and well-being to be correlated with each other. In the more general version of this model (Heckman 1979), equation (1) would be replaced with the switching equation

$$Y_{Ui} = \alpha_U + \mathbf{B}_U' \mathbf{X}_{Ui} + \varepsilon_{Ui} \quad \text{if } M_i = 0 \quad (1a)$$

$$Y_{Mi} = \alpha_M + \mathbf{B}_M' \mathbf{X}_{Mi} + \varepsilon_{Mi} \quad \text{if } M_i = 1. \quad (1b)$$

This version of the model would allow the unobserved determinants of marriage to be correlated with both ε_{Ui} and ε_{Mi} . The “standard” versions of these methods assume that the unobserved determinants in the marriage and outcome equations are normally distributed. Maximum likelihood and two-stage estimators for these models are available in econometric software packages such as LimDep and Stata.

The standard selection models have been criticized for imposing strong distributional assumptions on the unobserved variables. If the assumptions are incorrect, the estimators will be biased. The standard two-stage approaches are members, however, of a broader class of

methods, called control function methods, which address selection based on unobserved characteristics (Heckman et al. 1999; Heckman and Robb 1985). Control functions have been developed that incorporate alternative distributional assumptions (see Vella’s 1998 review). In addition, there are semi-parametric approaches for estimating the models that do not require distributional assumptions. The semi-parametric models do, however, require variable exclusions and are limited to cases in which there is a single source of selectivity (single index models).⁸ Maximum likelihood methods have also been developed that incorporate flexible distributional assumptions (see, e.g., Mroz 1999).

Nonparametric bounds. Other researchers, most notably Manski (1995), have advocated examining nonparametric bounds on treatment effects that require minimal assumptions. Consider the difference in expected well-being between married and unmarried people. The difference, conditional on other observed characteristics \mathbf{X}_i , can be decomposed

$$E(Y_{Mi}|\mathbf{X}_i) - E(Y_{Ui}|\mathbf{X}_i) = E(Y_{Mi}|\mathbf{X}_i, M_i=1) \Pr(M_i=1|\mathbf{X}_i) + E(Y_{Mi}|\mathbf{X}_i, M_i=0) \Pr(M_i=0|\mathbf{X}_i) - E(Y_{Ui}|\mathbf{X}_i, M_i=1) \Pr(M_i=1|\mathbf{X}_i) - E(Y_{Ui}|\mathbf{X}_i, M_i=0) \Pr(M_i=0|\mathbf{X}_i). \quad (2)$$

Most of the terms in equation (2) can be estimated directly from the data; however, the two counterfactual terms— $E(Y_{Mi}|\mathbf{X}_i, M_i=0)$ and $E(Y_{Ui}|\mathbf{X}_i, M_i=1)$ —cannot.⁹ Manski’s insight was that if these two terms had upper and lower bounds, such as best and worst case outcomes for Y , then expression (2) would also be bounded.

The most attractive feature of Manski’s technique is that it seems to provide something for nothing, that is, bounds on the effects of marriage without making any strong assumptions. Unfortunately, analysts who apply this technique are often disappointed with the “something” they receive, as the bounds turn out to be too wide to provide directions of impacts or much policy guidance. In some cases, the bounds can still serve a specification purpose and be used to rule out other estimates. The bounds can also be narrowed if the analyst is willing to impose other assumptions. When used in this way, the technique contributes to transparency because it forces the researcher to state and therefore justify his or her assumptions.

Longitudinal/panel methods

We now turn to data that contain observations on well-being, marital status, and other measures for individuals at different points of time, t ($t = 1, T$). We keep the same basic notation as before but add an argument in parentheses to distinguish observations from different time periods. Thus, $Y_i(t)$ denotes well-being for the i th person at time t ; $M_i(t)$ denotes marital status at time t , etc. With this notation, a longitudinal version of equation (1) can be written as

⁸ Examples with two potential sources of selectivity are wages, which differ by marital status and are observed only if a person works, and children’s outcomes, which differ by marital status and are observed only if a child is born.

⁹ For instance, $\Pr(M_i=1|\mathbf{X}_i)$ is the expected probability of being married conditional on the observed variables, \mathbf{X}_i . This probability—and its converse $\Pr(M_i=0|\mathbf{X}_i)$ —could be predicted from a probit regression of marriage on the elements of \mathbf{X}_i . An estimate of the conditional expected level of married people’s well-being among those who marry, $E(Y_{Mi}|\mathbf{X}_i, M_i=1)$, could be predicted from a regression of well-being outcomes on the elements of \mathbf{X}_i run on a sample of married people. An estimate of $E(Y_{Ui}|\mathbf{X}_i, M_i=0)$ could be predicted from a similar regression run on a sample of unmarried people.

$$Y_i(t) = \alpha M_i(t) + B'X_i(t) + \varepsilon_i(t). \quad (3)$$

Once again, the OLS estimate of α captures the difference in the average well-being of married and unmarried people holding the other observed characteristics constant.

All of the cross-section estimators that we have discussed can be implemented using longitudinal data. The added advantage, however, of longitudinal data is that the repeated observations for individuals permit before and after comparisons of outcomes. The information for a person at one point in time can be used to control for unobserved characteristics of the same person at another point in time.

Fixed effects. Fixed-effects methods address biases that arise from particular types of omitted variables. Assume that the unobserved determinants of well-being for each individual at each point in time can be decomposed into separate permanent and transitory components. That is, assume that the error term in equation (3) can be expressed as $\varepsilon_i(t) = \mu_i + v_i(t)$ where μ_i is a time-invariant component and $v_i(t)$ is a transitory component. The time-invariant component could correspond to permanent beliefs, attitudes, abilities or background characteristics of the person that affect well-being. If these are correlated with marital status or the other observed explanatory variables, OLS estimates of the coefficients in equation (3) will be biased.

We could account for these permanent characteristics in a regression model by including separate dummy variables for each person in the sample. While this approach addresses the potential bias, there are two drawbacks: first, the method may be impractical if the data include large numbers of people, and second, estimates of the coefficients on the dummy variables will not be consistent unless there are many observations for each person.

Instead of including the dummy variables, analysts typically use a differencing approach to obtain estimates of the coefficients α and B . Consider a data set in which two observations are available for each person: one at time t and another at $t + 1$. Assume that the data include at least some people for whom well-being, marital status and the other observed measures differ over time. Taking the difference of equation (3) over time, we obtain

$$\begin{aligned} [Y_i(t+1) - Y_i(t)] &= \alpha[M_i(t+1) - M_i(t)] + B'[X_i(t+1) - X_i(t)] + [\mu_i - \mu_i] + [v_i(t+1) - v_i(t)] \\ \Delta Y_i(t+1) &= \alpha \Delta M_i(t+1) + B' \Delta X_i(t+1) + \Delta v_i(t+1). \end{aligned} \quad (4)$$

Differencing sweeps out the unobserved permanent component μ_i , eliminating the source of bias. The requirement that some of the well-being and marriage outcomes vary over time is crucial; if these measures do not vary, they will be swept out along with unobserved permanent effect. Because this technique involves differences in outcomes over time for people who change their marital status as well as differences across people who were ever married or unmarried, it is sometimes referred to as a *difference-in-difference* estimator. In this example, one period differences have been used; however, differences over longer periods or differences from means of the variables for an individual over time can also be used.

Longitudinal fixed-effects estimators are straightforward to implement in most circumstances and are available in statistical software packages such as LimDep and Stata. Their main advantage is that they address omitted variables bias without requiring the researcher to identify exactly which variables are missing; the researcher only needs to specify the structure of the omitted variables. Aside from the data requirements, the principal disadvantage is the assumption that the source of bias is an omitted, time-invariant factor. The estimator does not address biases that arise from other error structures, although it can be extended to address some of these structures. The estimator also does not address biases associated with reverse causality and measurement error; in fact, application of the technique can make these biases worse.

Random effects and latent unobserved variables. Random effects models also decompose the unobserved characteristics in equation (3) into permanent and transitory components. The primary difference between the “random” and “fixed” effects approaches is that the random effects approach treats μ_i as being independent of some or all of the explanatory variables while the fixed effects approach allows for correlations with the explanatory variables. Clearly, the standard random effects version of equation (3), which treats μ_i as being independent of all of the explanatory variables including marital status, does not address omitted variables bias.

Suppose, however, that we also specify a random effects model for marriage. Let the net benefits (the difference between the advantages and disadvantages) of marriage for person i at time t be

$$M_i(t)^* = \mathbf{\Gamma}'\mathbf{Z}_i(t) + \theta_i + \zeta_i(t) \quad (5)$$

where $\mathbf{Z}_i(t)$ is a vector of observed characteristics, θ_i is a permanent unobserved characteristic, and $\zeta_i(t)$ is a transitory unobserved characteristic. The person is married ($M_i(t) = 1$) if the net benefits are positive and unmarried ($M_i(t) = 0$) otherwise.

Suppose also that we estimate this model with the longitudinal well-being equation

$$Y_i(t) = \alpha M_i(t) + \mathbf{B}'\mathbf{X}_i(t) + \mu_i + v_i(t) \quad (3')$$

and that we allow the two time-invariant unobserved components, μ_i and θ_i , to be correlated. If μ_i and θ_i are assumed to be normally distributed, longitudinal equations (5) and (3') can be jointly estimated using an extension of the random effects procedure described by Butler and Moffitt (1982). If the random effects are instead assumed to follow discrete distributions, the equations can be jointly estimated using the procedures described by Heckman and Singer (1984).¹⁰ Through the correlated random effects, the joint model accounts for the association

¹⁰ Each of these procedures uses maximum likelihood (ML) methods. To derive the ML estimator, we begin by specifying probability distribution functions for the unobserved variables in equations (3') and (5)—specifically, μ_i , θ_i , $v_i(t)$, and $\zeta_i(t)$. Given these distributional assumptions, there is a probability that we would observe the set of marriage outcomes, well-being outcomes, and other observed characteristics for all of the people in a particular sample. ML estimators for α , \mathbf{B} , $\mathbf{\Gamma}$ and the other parameters in the model are formed by finding the values of the parameters that maximize this joint probability.

If there were no random effects (i.e., if μ_i and θ_i were not in the model) and the transitory error terms $v_i(t)$ and $\zeta_i(t)$ were independent across individuals, independent over time and independent of one another, the probability of

between marriage and the permanent, unobserved determinants of well-being. Thus, like the fixed-effects estimator, it has the potential to eliminate bias.

There are several advantages with this approach. First, the random-effects specification does not sweep out all of the time-invariant characteristics. So, unlike the fixed-effects model, it is possible to include permanent characteristics like ethnicity and family background. Second, the random-effects approach can be applied to nonlinear models; hence, it can be used in more situations and with more complicated specifications than the fixed-effects approach.

Because it still assumes that the source of bias is a set of omitted, time-invariant factors, the random-effects model shares the main disadvantage of the fixed-effects model. As mentioned, the assumptions required for the random-effects model are actually more stringent—the random effects must be independent of the other observed variables.¹¹ Other disadvantages of the random-effects model are its computational requirements and complexity. Models of this type can take 50-100 times as long to run as comparable models without random effects. Also, it can be much harder to get stable estimates of some of the parameters of the model.

Applications to siblings, peers and neighbors. Estimation methods that were initially developed for panel data have been applied to other types of data in which the observations are repeated across members of a group, such as brothers or sisters in a family, classmates in a school, or neighbors in a community. Instead of controlling for unobserved, permanent characteristics of an individual, these applications account for unobserved characteristics of a group. For instance, a panel model applied to data on siblings would account for their shared family environment and upbringing. A model applied to classmates would account for common characteristics of their classroom, school and neighborhood.

The strengths and weaknesses of these estimation methods in the group context are the same as in the standard longitudinal context. On the plus side, the techniques address problems associated with omitted variables without requiring the researcher to identify a particular variable—the researcher only needs to know that the variables are common across members of the group. On the negative side, the ability of the techniques to eliminate bias stems entirely from this assumption. The techniques do not address biases associated with unobserved characteristics that vary across members of the group and can exacerbate biases associated with reverse causality and measurement error (Griliches 1979).

observing each marriage and well-being outcome would also be independent across individuals and time. Thus, the probability of observing the entire sample would be characterized by the product of $N \times T \times 2$ univariate distributions. If we introduce random effects that are correlated with each other but independent across individuals, the probability is instead characterized by the product of N multivariate ($T \times 2$ -variate) distributions. Computational problems arise because it is substantially more difficult to evaluate multivariate distributions than univariate distributions. Note, however, that if the transitory terms $v_i(t)$ and $\zeta_i(t)$ are independent of the random effects, the probability of observing the outcomes in the sample can be re-expressed as the product of $N \times T \times 2$ conditional (on μ_i and θ_i) univariate distributions and an additional bivariate distribution (the joint marginal distribution of μ_i and θ_i). This greatly simplifies computation of the probabilities and ML estimates. The procedures by Butler and Moffitt (1982) and Heckman and Singer (1984) decompose the problem in this way and then make alternative assumptions about the marginal distribution characterizing the random effects. A relatively new software package, aML (EconWare, Inc.), is available to implement these procedures.

¹¹ Chamberlain (1980) proposes a technique for relaxing this restriction.

The value of applying fixed or correlated random effects models to sibling, peer or neighborhood comparisons seems minimal for research on the impact of family structure. Consider models based on family comparisons. For analyses of the relationship between marriage and well-being for adults (e.g., Loh 1996), family background is essentially a fixed characteristic, and standard longitudinal controls could be employed to account for this and other fixed personal characteristics. There are also difficulties in analyses of children's well-being (see, e.g., Ermisch and Francesconi 2001; Ginther and Pollak 2003; Sandefur and Wells 1999). Sibling comparisons use differences in the length and timing of brothers' and sisters' exposure to alternative arrangements to explain differences in their outcomes. Families with stable intact or non-intact arrangements (a majority of families) do not contribute any usable variation to such an analysis. Biases associated with misreports of the timing of family events are also magnified.

More complicated descriptions of family structure

All of the models that have been considered so far treat family structure as a binary variable; people are modeled simply as being married or not. However, for some types of research questions, it is necessary to distinguish between different non-marital arrangements such as being never-married, cohabiting, separated, divorced or widowed. Similarly, other types of analyses require us to distinguish between first and later marriages or identify blended households. In addition, some research questions involve people's marital or relationship histories, including information on the duration of an existing marriage or whether a marriage was preceded by a period of cohabitation. Some research also focuses on other characteristics of relationships such as whether they are satisfying, healthy, or successful (Glenn 1990).

In principle, all of the estimation methods that were examined in the context of a two-way comparison can be extended to accommodate multiple comparisons or multiple descriptors for family structure. In practice, however, some extensions are more difficult to implement than others. At one end of the spectrum, the OLS (including the method of adding variables to account for omitted variables) and fixed-effects regression methods readily accommodate more extensive descriptors for family structure. The only complications, which are common to all of the estimation approaches, are collecting the additional family structure variables and verifying that they contain independent variation. For the fixed-effects model, there is also a requirement that the indicators change over time for some people in the sample. At the other end of the spectrum, the nonparametric bounds approach has not been extended to consider large numbers of comparison groups. The other methods fall between these extremes. In each case, the estimators have been extended to account for continuous endogenous variables or multiple sets of endogenous variables; however, these extensions introduce significant complications. The complications are discussed in more detail in Appendix A.

Nonlinear models

Researchers examining the benefits and consequences of marriage have also used other specifications besides linear regression models with continuous dependent variables. The review briefly describes some of these models and their properties. It also overviews issues that arise in accounting for omitted variables and reverse causality in these models.

Binary outcome and ordered categorical models. Dependent variables that are characterized by binary, “yes/no” outcomes, such as indicators for whether someone is employed or whether a child is enrolled in school, *can* be modeled using linear regression methods. The same is true of dependent variables that take on ordered categorical values, such as responses to a Likert scale. The chief difficulties with using standard regression methods to model binary and ordered categorical outcomes are that the error terms in the estimating equations are heteroskedastic (i.e., the variances of the errors differ across observations) and the models might predict values outside the original range of the variables. An alternative approach, which addresses these problems, is to estimate a maximum likelihood logit or probit model (or ordered logit or ordered probit model).

Logit and probit models are non-linear and, therefore, more difficult to estimate than linear regression models. This difficulty is not substantial when standard versions of the models are estimated. However, it may be an issue when statistical controls for selectivity are added. For instance, the differencing approach to estimating fixed effects models does not work when applied to nonlinear models. Because of the computational problems, researchers sometimes adopt regression specifications, instead of probit or logit specifications, when they also have to address problems with omitted variables or selectivity. Methods for addressing selectivity in probit and logit models are discussed in more detail in Appendix A.

Hazard and survival models. Hazard and survival models are used to examine how long someone stays in one state before transitioning to another. For example, the models are used in analyses of mortality, spells of unemployment, and periods of program enrollment. Estimating the determinants of spell lengths is straightforward if all spells are measured to their conclusion (e.g., if complete lifetime data were available in the mortality example). In many cases, however, information on spell lengths is incomplete because people are only surveyed for a limited period of time or drop out of a sample before making a transition.

Hazard models address these censoring problems by modeling the probability that someone will make a transition at a particular point in time *conditional* on the person not having made the transition before that point in time. Some hazard models impose distributional assumptions on the spell lengths and are estimated using maximum likelihood. Like other maximum likelihood estimators, these models are very sensitive to specification errors. For instance, random variation arising from any source including omitted variables and imprecise measurement of spell lengths can lead to biased estimates. There are relatively few methods for addressing selectivity and non-independent omitted variables in hazard models. Some of the available methods are discussed in Appendix A.

Structural equations, path, and covariance structure models. All of the models discussed up until now have considered the relationships between individual observed variables, such as the relationship between a single measure of well-being and an indicator for marital status. Structural equations and path models are used in situations where there are multiple indicators for an outcome variable, the explanatory variables, or all of these. An illustration of this type of model is given in Appendix A.

The consequences of omitted variables and reverse causality are essentially the same in these models as in linear regression models. In theory, structural equations models can be adjusted to overcome these problems. The techniques can accommodate an enormous range of specifications including correlated random effects and bi-directional relationships. While models with these features are commonly estimated, flexibility in one area is often accompanied by arbitrary restrictions on other areas. For instance, many applications with simultaneously determined outcomes assume that the unobserved determinants of the outcomes are uncorrelated (Godwin 1988). Variable exclusions and other restrictions are also often imposed without regard to the underlying theory. These faulty restrictions can lead to biased estimates.

Summary

This section reviews statistical methods that are commonly employed to examine the relationship between marriage and well-being. These general approaches include linear regression models when well-being is specified as a single continuous variable, discrete choice models when well-being is specified as a binary or ordered categorical variable, hazard models when well-being is specified as a spell, and structural equations models when either well-being or marriage are indirectly represented by multiple indicators. The section also reviews methods that address biases that arise from omitted variables and endogenous explanatory variables. It examines these methods first in the simple case of a linear regression model for well-being outcomes in which marital status is included as a binary explanatory variable. It then briefly considers the complications that arise in non-linear models and models that include several measures for marriage or family structure.

There is no “magic bullet” statistical technique for determining the effects of marriage. All of the techniques that are considered in this section involve trade offs—not only in terms of their computational complexity but also in terms of their underlying assumptions and data requirements. Faster computers and advanced software packages like Stata and aML are rapidly extending the boundary of what is computationally possible. Even with these advances, however, underlying assumptions and data quality remain critical. As computable models become more complex, the number of important specification decisions will increase. Careful theoretical analyses and data collection can help to guide these decisions.

IV. Empirical Studies of the Effects of Family Structure

This section reviews quantitative studies that have been conducted in the domains of children’s well-being, adults’ earnings, and adults’ physical health and mortality. Although the review characterizes the general findings from each domain, it is far from exhaustive. Instead, the review discusses the broad trajectories in the types of statistical methodologies that have been employed and illustrates how and in what context particular techniques have been applied. The challenges associated with selection and omitted variables have been recognized and acknowledged in all three domains for some time. Accordingly, there have been many parallel developments, such as the wider use of longitudinal data in the last two decades. However, there have also been differences, which are reflected in the current research in each domain. The review discusses these differences and assesses the prevailing practices.

Children's well-being

Overview. The literature examining the relationship between family structure and children's well-being is astonishingly vast. A search through the *Sociological Abstracts* database for English-language journal articles published since 1990 with the keywords "Children" and either "Divorce" or "Family Structure" generated more than 1,100 citations. A similar search through *PsychINFO* produced more than 1,500 citations. Not only has the research in this area resulted in a tremendous number of publications, but the publications themselves have been subjected to multiple reviews and meta-analyses. Among the reviews are the summary of evidence supporting marriage by Waite and Gallagher (2000), specialized reviews on the topics of divorce by Amato (1993, 2000), Hines (1997), and Seltzer (1994), remarriage by Coleman, Ganong and Fine (2000), non-marital families by Seltzer (2000) and economic outcomes by Lerman (2002c), as well as more methodologically-focused reviews by Cherlin (1999), Haveman and Wolfe (1995), Sigle-Ruston and McLanahan (2002b) and Thornton (2001).

There are several reasons why children's well-being is the subject of such intense scrutiny. The most important is that childhood represents a critical and unique period of development. Deficits that appear during childhood may be difficult or impossible to overcome later in life and thus may have long-lasting consequences. If these disadvantages also affect subsequent parenting, the consequences could play out across multiple generations.

A second reason has to do with children's vulnerability. Children are dependent on others for their care and well-being, young children especially so. Moreover, children have little or no voice in many of the decisions that affect them. Economists refer to situations in which choices made by one group affect outcomes for others as *externalities*. Because a decision-maker in such a situation does not gain all of the benefits or bear all of the costs of his or her choice, there is a tendency take fewer positive actions and more negative actions than are socially desirable. The love and concern that parents have for their children overcome some of these problems; however, to the extent that parents do not entirely appreciate or weigh the consequences for children, their decisions may be sub-optimal.

A third reason, which is more scientific and methodological, involves the insights that this type of research can offer into basic theories of how children develop. Empirical research on children's development requires data on outcomes and circumstances; the data can, in turn, be drawn from either experimental or observational settings. While observational data have many drawbacks, they are easier to obtain and often have higher external validity than experimental data. Data on family transitions are especially useful because they represent large, salient and potentially stressful changes in children's circumstances. Even if researchers have trouble unpacking the specific processes that are involved, they can still examine whether children are resilient to these types of changes.

The studies of the relationship between family structure and children's outcomes paint a fairly consistent picture. On average, children who grow up with two married parents enjoy higher living standards, better health and better developmental outcomes than children in other types of families. These children also tend to advance farther in school, behave better, abuse fewer substances and achieve more success as adults, spouses and parents. Naturally, there are

exceptions among the studies. Many studies report insignificant associations for some or all of their indicators.¹² Studies also occasionally report that particular types of non-marital arrangements, such as living with a single mother in a multi-generational household (Deleire and Kalil 2002b), are more beneficial than living with two parents. Some studies have also found that marriage has negative associations for select groups of children, such as those whose parents fight and argue often (Amato, Loomis and Booth 1995, Jekielek 1998, Morrison and Coiro 1999). On balance, however, the evidence indicates that marriage is associated with better outcomes for children under most circumstances.

The principal research question, therefore, is whether the observed association reflects an actual benefit from marriage or is instead an artifact of reverse causality or spurious correlation from other unmeasured factors. Reverse causality may seem like an esoteric concern to many social scientists, especially those in disciplines that assign all of the power and decision-making ability in theoretical models to parents and caregivers. There is also a temporal argument that some type of household structure exists prior to the arrival of a child.

A case for reverse causality, however, can be made on the basis of transactional theories of child development that emphasize interactions between children and those around them (see the discussion in Chase-Lansdale 1998). Poor outcomes for children are surely a source of stress in families, as Harris (1998, p. 296) illustrates with the following joke:

Psychologist: You should be kind to Johnny. He comes from a broken home.

Teacher: I'm not surprised. Johnny could break any home.

Studies have examined whether children's illnesses (Corman and Kaestner 1992) and physical impairments (Mauldon 1992) destabilize marriages. Additional evidence of possible effects of children comes from research that connects the number (Lillard and Waite 1993), ages (Waite and Lillard 1991), and even gender (Katzew, Warner and Acock 1994) of children to subsequent marital instability. While many of these studies can also be criticized for ignoring reciprocal causality, the point remains that researchers can interpret the data in different ways.¹³

Spurious correlation from omitted variables is a more widely acknowledged concern in this literature. Studies typically compare results from raw correlations or parsimonious regressions with those from more fully-specified models. They almost always find that the strength of the association between family structure and children's outcomes falls as controls are added. In one of the best-known studies, McLanahan and Sandefur (1994) found that differences in financial resources, parenting behavior and community resources accounted for most of the family structure differentials in youths' schooling, economic inactivity and early family formation. It is not difficult to suppose that additional controls could erase the remaining

¹² Herzog and Sudia (1973) sounded an early caution regarding the evidence on the consequences of father absence. They pointed out numerous insignificant, equivocal and methodologically flawed results in the then-existing studies. More recent reviewers, such as Amato (1993), continue to catalog a large number of insignificant findings.

¹³ Indeed, an earlier (1990) analysis by Mauldon examined the effect of marital dissolution on children's health. Thus, in one analysis health was modeled as affecting marriage outcomes, while in another marriage was modeled as affecting health. While the structure of these analyses implied that there was reciprocal causality, neither analysis adopted statistical controls to account for this possibility.

differences. For example, Harris (1998) has argued that genetic similarity between parents and children and peer influences could account for the effects generally attributed to parent behavior. Biblarz and Raftery (1999) have shown that the selective use of controls eliminates the negative association between being raised by a single mother and subsequent socioeconomic status.

Studies without selectivity controls. By a wide margin, the prevailing statistical methodology in this domain has been and continues to be standard regression or discrete-choice modeling. A few recent examples are the studies by Astone and McLanahan (1991, 1994) of schooling, Biblarz and Raftery (1999) of socioeconomic success, Dawson (1991) of health and behavioral outcomes, Deleire and Kalil (2002b) of multigenerational families, Ehrle, Kortenkamp and Stagner (2003) of outcomes in welfare families, Haveman and Wolfe (1994) and McLanahan and Sandefur (1994) of youths' attainments, and Lerman (2002a, b) of hardship experiences. These analyses were all carefully done and innovative. However, the innovations mostly involved utilizing new data sources to examine different outcomes or special populations or incorporating new explanatory variables to account for confounding or mediating influences.

Dawson's (1991) article illustrates the first type of innovation. Using new data from the Child Health supplement of the National Health Interview Survey, Dawson examined 17 separate measures (many built up from multi-item inventories) of problems with children's physical health, emotional well-being and behavior. Her regressions indicated that family structure was associated with nearly every one of these indicators. Ehrle et al. (2003) used data from a small sample of long-term welfare families in California to show that few of the positive associations between marriage and children's well-being extended to low-income families. Updating and extending research by Bauman (1999) that examined how material hardship varied with living arrangements, Lerman looked at hardship among families with children using new data from the SIPP (2002a) and National Survey of America's Families (2002b).

The articles by Astone and McLanahan, which introduced detailed measures of parenting practices and aspirations (1991) and residential mobility (1994) into models of youths' schooling outcomes, illustrate the second type of innovation. In each case, the new measures were strongly related to both family structure and schooling, and their addition into the regressions reduced but failed to eliminate the association between family structure and schooling.¹⁴ The studies by Biblarz and Raftery (1999), Haveman and Wolfe (1994), and McLanahan and Sandefur (1994), which formally compared numerous specifications involving different sets of explanatory variables, were more comprehensive but essentially fit the same pattern. Deleire and Kalil (2002b) took a slightly different tack. Instead of incorporating new mediating variables, they examined narrower categories of family structure, including multigenerational arrangements.

An extension of this approach is to adopt Freedman's (1991) "shoe leather" strategy of using theory to form a series of sharp hypotheses and then using new samples or measures to test all of them. Nearly every empirical study offers some type of conceptual framework; however, these usually only serve to predict a few gross relationships in the data, such as a negative effect from divorce or a mediating effect from a chosen variable. A handful of studies including Amato (1993), Wu and Martinson (1993) and Wu (1996) have followed Freedman's admonition

¹⁴ The 1994 article also illustrates a common methodological oversight. Although the study relied on the same survey as the 1991 article, it did not incorporate the parenting variables that were featured in the earlier analysis.

and considered more detailed predictions involving several measures. These studies did not employ statistical controls to account for selectivity. Instead, they considered whether the evidence was consistent with a broad set of predictions and focused on points of departure between theories.

Amato's (1993) analysis of children's adjustments to divorce is especially interesting because it involved a re-examination of published studies rather than new estimation. Amato considered theories of parental loss, parental adjustments, interparental conflict, economic hardship, and stress and drew 3-5 testable implications from each theory. While none of the studies that he reviewed provided a comprehensive test of all of the implications, the studies did test some of them. Amato counted the number of partial test results that supported the theories and the number that did not. Each of the theories could claim some empirical support, but theories of interparental conflict had the most consistent support.

Wu and Martinson (1993) examined the relationship between family structure and young women's premarital childbearing. They constructed family history measures that corresponded to theories of socialization, social control, and instability and stress. They examined the measures separately and in combination and found that the measures associated with instability and stress were consistently significant. Other measures were significant when they appeared by themselves, but these results vanished when the instability measures were added. A weakness of the study was that it lacked measures of economic circumstances in the youths' families and, thus, could not test economic theories. Subsequent research by Wu (1996) used data from the National Longitudinal Survey of Youth on family histories and incomes to address this problem. Wu found that a history of family instability continued to be a significant determinant of premarital childbearing but that the level of economic resources was also important.

Studies with longitudinal controls. Among the studies that have used statistical controls for selectivity, most have employed longitudinal methods. The potential utility of these methods was demonstrated by Cherlin et al. (1991) who compared test results and behavior problems among British and American children at one point in time with measures recorded later after some of their parents had divorced. They found that pre-existing differences in achievement and behavioral problems accounted for up to half of the difference in post-divorce outcomes for boys and a smaller portion of the difference for girls. Initial differences in family problems also accounted for some of the difference in post-divorce outcomes.

Several other longitudinal studies have led to results that are consistent with both selectivity and possible causal effects. For instance, analyses by Cherlin, Kiernan and Chase-Lansdale (1995) and Fronstin, Greenberg and Robins (2001) of early adult outcomes among respondents in the British National Child Development Study indicated that divorce was associated with higher rates of home-leaving, cohabitation and premarital childbearing and lower levels of education, even after controlling for children's initial socioeconomic circumstances, emotional problems and cognitive skills. Amato and Booth (1996) examined longitudinal data on parent-child relations and found that relations in the initial observation period were worse in families that eventually divorced. After accounting for these conditions, they found that divorce was associated with poorer subsequent relationships with fathers but not mothers. Painter and Levine (2002) investigated how high school completion, college attendance and out-of-wedlock

childbearing were associated with family structure. Controls for children's initial characteristics reduced and in some cases eliminated the associations with family structure.

The findings of Morrison and Cherlin (1995) represent an exception to this pattern of results. Morrison and Cherlin examined changes over two years in children's behavioral problems and reading comprehension using data from the Child Supplement of the 1979 NLSY. They found that marital disruptions were associated with worse behavior for boys and that this result was not attenuated by controls for pre-existing conditions. Marital disruptions were not significantly associated with behavior for girls or with reading for either gender. The short observational window may have contributed to some of the weak findings in this study.

Longitudinal comparisons have also been used in research on the role of interparental conflict in divorce outcomes. As mentioned, Cherlin et al. (1991) found that the higher incidence of family problems in divorcing households explained part of the difference in children's subsequent outcomes. More recent studies have not only looked at the initial levels of conflict but also at the interaction between conflict and divorce. Amato, Loomis and Booth (1995), Jekielek (1998) and Morrison and Coiro (1999) estimated models that indicated that divorce among low-conflict couples was associated with negative outcomes for children while divorce among high-conflict couples was associated with positive outcomes. Hanson (1999) also found some evidence of interactions, though his results were weaker.¹⁵

Although longitudinal data have been featured in several studies of children's well-being, few of these studies have actually estimated fixed effects specifications. Fixed effects models, which require repeated measures of and longitudinal variation in the outcome and explanatory variables, are difficult to apply in children's research. Many developmental measures are age-dependent and, thus, not comparable from one period in a child's life to another. Other measures, such as high school completion, family formation, and employment among teenagers and young adults, are only observed after disruptions have occurred. Similarly, some types of family circumstances, like whether a child was born out-of-wedlock, are time-invariant and consequently absorbed by fixed-effects controls. One example of a fixed-effects study was Morrison's and Ritaulo's (2000) analysis of children's economic outcomes following divorce, which indicated that remarriage and cohabitation might confer similar financial benefits.

Deleire and Kalil (2002a) also recently used fixed effects methods to compare the types of expenditures made by single, married and cohabiting parents. On average, cohabiting parents spent more on alcohol and tobacco and less on educational services and health care than either single or married parents; however, these differences were substantially reduced when fixed effects controls were introduced. Deleire and Kalil also estimated propensity score matching models. For tobacco/alcohol and education, the results from the matching models fell in between

¹⁵ Hanson examined 20 specifications involving 14 distinct school, health and behavioral outcomes (he distinguished between parents' and children's reports for some outcomes and estimated gender-specific models for some others). He reported significant interaction coefficients in 8 specifications.

those from standard regression and fixed effects models.¹⁶ The pattern of results was consistent with selection on both observable and unobservable characteristics.

Studies with other controls for selectivity. Sandefur and Wells (1999) analyzed the relationship between family structure and schooling using sibling comparisons to account for unobserved family-specific factors. Specifically, they estimated a structural equations model in which educational attainment for each sibling depended on the person's gender and family structure at age 14 as well as a shared latent variable for family background. The shared background variable, in turn, reflected the family structures experienced by the siblings and other characteristics of the parents and family. Although the authors did not cite Chamberlain (1980), their model effectively implemented his fixed effects procedure. Using this procedure to account for background characteristics reduced the association between family structure and schooling but did not eliminate it. Ermisch and Francesconi (2001) similarly found that sibling controls weakened but did not eliminate the associations between family structure and several youth outcome measures including schooling. In contrast, Ginther and Pollak (2003) reported that family structure was not significantly associated with youths' educational outcomes in models that included sibling controls and other family background measures. Ginther and Pollak did, however, find associations between family arrangements and younger children's reading and behavior.

Manski, Sandefur, McLanahan and Powers (1992) estimated a variety of specifications including standard probit models, endogenous switching regressions, and sets of non-parametric bounds to examine the impact of family structure on high school completion. Results from the probit and switching regression models were nearly identical and indicated that growing up in a non-intact family reduced teenagers' chances of graduating from high school. The results from the models also fell within the non-parametric bounds. There was no evidence from these findings or from formal specification tests that family structure was selective. The only notable weakness in this study was the questionable set of instruments—indicators for region of residence and differences in parents' education—for family structure in the endogenous switching regression model.

McLanahan and Sandefur (1994) also estimated endogenous switching regression models in their analysis of the consequences of single parenthood. As in their earlier study with Manski and Powers, they found few differences between the results from these models and standard probit models. However, also as in the earlier study, they used regional indicators and the difference in parents' education as instruments in the switching specifications.

Two recent studies have used the death of a parent as a natural experiment for single parenthood. Corak (2001) used Canadian tax data to compare earnings, employment and marital outcomes for people who lost a parent through death, lost a parent through divorce, or never lost a parent. Corak found that outcomes were best for people raised in intact households, slightly worse for people who had lost a parent through death, and much worse for people whose parents had divorced. Lang and Zagorsky (2001) generated similar results using data from the NLSY.

¹⁶ Deleire and Kalil estimated bivariate matching models of the differences between cohabiting and married parents and cohabiting and single parents, respectively. They did not employ Imbens' (2000) or Lechner's (2001) procedures for multiple contrasts.

Each study interpreted its findings as indicating a large selective effect but little or no causal effect from single parenthood. There are several grounds, however, for questioning the appropriateness of these comparisons. First, there are differences in the financial and social support that divorced and widowed single-parent families receive. Second, there are also differences in the stresses associated with each type of disruption. Third, the behaviors modeled by the parents in each situation differ. Finally, the underlying assumption that death is exogenous is open to question given the research (which we subsequently review) that links marriage and marital quality to mortality.

Mindful of these limitations, Corak also examined how outcomes varied across people who grew up before and after divorce laws were relaxed in Canada. He found virtually no evidence that outcomes changed after the laws were reformed. Given that the divorce rate appeared to respond to changes in the law, the results suggested that divorce had little impact on children's subsequent attainments. Unfortunately, the results were far from definitive because changes in other aggregate conditions could not be ruled out.

Gruber (2000) also investigated the relationship between divorce laws and children's later socioeconomic attainments but used U.S. rather than Canadian data. Differences in the dates at which states enacted no-fault, unilateral divorce statutes provided sources of both longitudinal and cross-section variation. Thus, Gruber's empirical analysis could account for aggregate time series effects as well as time-invariant, state-specific effects.¹⁷ Gruber found that children who grew up in states that relaxed their divorce laws attained less schooling and lower incomes but were also more likely to marry than children raised in other states.

Assessment. Numerous studies have examined the relationship between family structure and children's well-being. While this research has generally found that marriage is associated with better outcomes for children under most circumstances, the evidence is based mainly on analyses that failed to account for selectivity. Selectivity appears to be more than a hypothetical concern in this literature. Comparisons across regression specifications indicate that adding controls for other aspects of family circumstances greatly reduces the association between marriage and children's outcomes. The measured association also falls when longitudinal data are employed to account for the characteristics of children before they experience family disruptions. The associations that remain are consistent with small direct causal effects, though they could also reflect additional selectivity. It is also possible that the total causal effects are larger than those reported because many of the studies include variables for other outcomes that may be impacted by marriage, like income and neighborhood quality, as controls.

There has been some quantitative research that has moved beyond standard regression analyses and longitudinal comparisons. From a methodological standpoint, Gruber's (2000) analysis is perhaps the most exciting because it indicates that divorce laws might be a worthwhile instrument for family structure. More substantively, the findings by several of these recent studies that marriage continues to be associated with children's well-being provide stronger evidence of a causal link.

¹⁷ Some of Gruber's analyses also accounted for state-specific linear trend effects.

Table 1. Selected Empirical Analyses of the Effects of Marriage on Children's Outcomes

Study	Data	Dependent Variables	Key Explan. Vars.	Technique	Findings
<i>Cross-section studies</i>					
Dawson (1991)	1988 Child Health Supplement of NHIS (families with a child under 18)	health, school behavior, other behavioral and emotional problems	two biological parents, formerly married mother only, never married mother only, mother and stepfather	OLS	living in a non-intact family is consistently associated with negative health and behavioral outcomes
Ehrle, Kortenkamp & Stagner (2003)	1993-2000 California long-term welfare families	behavioral and emotional problems, school success, health	married, married stepfamily, cohabiting, cohabiting-step, single parent formerly married, single parent never married, family transitions	OLS	living with cohabiting parents and experiencing a transition are negatively associated with behavioral and emotional problems; living with a formerly married single parent and experiencing a transition are negatively associated with school success
Ginther & Pollak (2003)	1986-94 NLSY Child Supplement (children 5-15)	reading scores and behavioral problems	proportion of childhood spent living with single mother, both parents but no step-siblings, other blended families; some analyses limited to stable blended families	siblings fixed effects	living in a blended family was associated with worse reading and behavior outcomes
Lerman (2002a)	1996 SIPP (families with children)	poverty and material hardship	married, cohabiting, single parents with and without other adults	OLS & probit	outcomes are best for married couples and worst for single parents with no other adults; results appear across numerous specifications and sub-samples
Lerman (2002b)	1997-99 NSAF (families with children)	material hardship	married, cohabiting and single parent families with both parents, one parent or foster parents (9 categories)	OLS & probit	outcomes are best for two-parent married couples and worst for single-parent families; results appear even after controls for income-to-needs are included

Longitudinal studies

Cherlin et al. (1991)	I: 1958-69 National Child Development Study (U.K. children born in 1958; initially in married families)	behavior problems, reading and math achievement	divorce/separation	OLS	controlling for initial behavior problems, achievement and family problems reduces association between divorce and outcomes for boys; results less pronounced for girls; associations remained for boys and girls
	II: 1976-81 National Survey of Children (U.S. children 7-11 in married families in 1976)	behavior problems	divorce	same	similar results for boys; no association between divorce and behavior for girls with or without controls
Deleire & Kalil (2002a)	1982-98 Consumer Expenditure Survey (families with children)	detailed expenditures	married, cohabiting, single parents	fixed effect & propensity score matching	cohabiting and single-parent couples spent more on alcohol and tobacco and less on health and educational services than married couples; differences were smaller when fixed effects and matching controls were used
Hanson (1999)	1987-94 NSFH (married families with children in 1987-88)	school performance and behavior, delinquency, health outcomes, psychological well-being	divorce, conflict	OLS & logit	divorce and conflict are negatively associated with outcomes for children; interactions of divorce and conflict are weak and inconsistent
Jekielek (1998)	1988-92 NLSY Child Supplement (children 2-10 in intact families in 1992)	behavioral problems index and sub-items	divorce/separation, conflict	OLS	divorce and conflict associated with negative outcomes for children; significant interactions (better outcomes when high-conflict parents divorce); some evidence that impact of divorce fades with time
Mauldon (1990)	1981 Child Health Supplement of NHIS (families with a child under 18; selects non-disabled children born into intact families)	health problems	divorce/separation, remarriage	logit	divorce/separation is positively associated with health problems

Morrison & Cherlin (1995)	1986-88 NLSY Child Supplement (children 3+ in intact families in 1986)	behavioral problems index, reading	divorce/separation	OLS	disruption is negatively associated with behavior for boys but not girls; disruption is not associated with reading
Morrison & Coiro (1999)	1988-94 NLSY Child Supplement (children 4-9 in intact families in 1988)	behavioral problems index	divorce/separation, conflict	OLS	disruption and conflict are positively associated with behavioral problems; behavior improves if high-conflict parents divorce
Morrison & Ritualo (2000)	1986-94 NLSY Child Supplement (children born in intact families but experiencing disruptions)	post-divorce family income	single, cohabiting, remarried, number of transitions, time since disruption	fixed effects	depending on how resources of cohabiting partner are treated, remarriage and cohabitation may confer similar economic benefits

Table 2. Selected Empirical Analyses of the Effects of Marriage on Subsequent Youth and Young Adult Attainments

Study	Data	Dependent Variables	Key Explan. Vars.	Technique	Findings
<i>Cross-section studies</i>					
Astone & McLanahan (1991)	1980-86 High School and Beyond	schooling aspirations, grades, attendance, attitudes, high school completion	two parents, one parent only, step-family, neither parent	OLS & probit	parenting practices partially mediate relationship between family structure and educational outcomes
Astone & McLanahan (1994)	1980-86 High School and Beyond	high school drop-out	two parents, one parent only, step-family, neither parent	OLS & probit	residential mobility partially mediates relationship between family structure and drop-out
Biblarz & Raftery (1999)	I: 1962 Occupational Changes in a Generation Survey	occupational status and education	two parents, mother-only, mother & step-father, father	OLS	controlling for family SES reduces association between family structure and subsequent SES; father-headed structure is negatively associated with occupation; mother-only and father-headed structures are negatively associated with education
	II: 1973 OCG	same	same	same	mother/step-father structure is negatively associated with occupation; all non-intact structures are negatively associated with education
	III: 1986-88 Panels of SIPP	same	same	same	father-headed and mother/step-father structures are negatively associated with occupation; all non-intact structures are negatively associated with education
	IV: 1987-94 NSFH	same	same	same	mother-only and father-headed structures are negatively associated with occupation; all non-marital structures are negatively associated with education
Deleire & Kalil (2002b)	1988-94 NELS	high school graduation, college attendance, smoking, drinking, sexual activity	10 types including never-married and divorced mothers in multigeneration HHs	probit & discrete hazard (probit)	outcomes are better for children with never-married mothers living in multigenerational households; however, sample size (30) for this group is small

Ermisch & Francesconi (2001)	1991-95 British Household Panel Survey (youths living with mother)	education, inactivity, early childbearing, emotional distress, smoking	single parents, other parents (ever and at ages 0-5, 6-10 & 11-16)	sibling fixed effects	exposure to a single-parent household was negatively associated with schooling and positively associated with smoking; early exposure appeared to be important
Ginther & Pollak (2003)	I: 1979-94 NLSY (siblings)	education	proportion of childhood spent living with single mother, both parents but no step-siblings, other blended families, other structures; some analyses limited to stable blended families	siblings fixed effects	living with both parents and no step-siblings was positively associated with schooling in parsimonious sibling FE models; association was not significant when additional family controls were added
Gruber (2000)	II: 1968 1968-93 PSID (siblings born between 1960-70) 1960-90 PUMS	same	same	same	same
Lang & Zagorsky (2001)	1979-93 NLSY	marriage, education, income, labor supply	no-fault, unilateral divorce laws in state of birth and current state of residence	natural experiment	growing up in a state with a unilateral divorce law leads to higher marriage but less education and lower earnings; also associated with lower employment for women
Lang & Zagorsky (2001)	1979-93 NLSY	AFQT, education, marital status, earnings, income, wealth	parents' age at child's birth, years living with each parent; parental death	natural experiment	uses parent's death as a natural experiment; little differential effect associated with parental death—little evidence parent absence affects outcomes
Manski et al. (1992)	NLSY	high school completion	non-intact family	probit, switching model, non-parametric bounds	high school completion lower in non-intact families in probit and switching models; estimated impacts fall within non-parametric bounds; no evidence of selectivity
McLanahan & Sandefur (1994)	HS&B, NLSY, NSFH, PSID	educational attainment and performance, idleness, teenage childbearing	intact, widowed, divorced, initially single households	OLS, logit, switching model	outcomes are best for married families and worst for initially single families; controls for economic conditions, parenting, and community characteristics reduce but do not eliminate associations; results are similar in switching models

Sandefur & Wells (1999)	1979-92 NLSY (siblings)	years of schooling	family structure at age 14, predominant family structure, years in intact family at different ages, transitions	structural equations (sibling contrasts)	non-marital family structures are negatively associated with schooling; sibling controls reduce associations
-------------------------	-------------------------	--------------------	---	--	--

Longitudinal studies

Amato & Booth (1996)	1980-92 longitudinal survey of married couples; initially parents of children under 18	parent-child affection in 1988	divorce, marital happiness	structural equations	divorce reduces affection between fathers and children but not mothers and children; marital happiness is positively associated with parent-child affection
Amato, Loomis & Booth (1995)	1980-92 longitudinal survey of married couples; children of 1992 respondents	psychological distress, relationship happiness, relationships with relatives	divorce, conflict, marital happiness, marital interactions, divorce proneness	OLS	divorce reduces well-being in low-conflict families but increases well-being in high-conflict families
Cherlin, Kiernan, & Chase-Lansdale (1995)	1958-81 National Child Development Study (U.K. children born in 1958; initially in married families)	home-leaving, early family formation	divorce, parent death	logit & multinomial logit	divorce positively associated with home-leaving, premarital cohabitation, premarital parenthood for men and women
Corak (2001)	1982-95 Canadian income tax records (married or separated families with a child 16-19 in 1982-86)	income, earnings, years of work, transfers, marriage	intact, bereaved, divorced families	natural experiment OLS	uses bereaved families and changes in divorce laws as natural experiments; finds small negative associations between bereavement and subsequent outcomes but no associations with divorce laws
Fronstin, Greenberg & Robins (2001)	1958-91 National Child Development Study (U.K. children born in 1958)	education, employment & log hourly wage	exposure to divorce or death of parent	ordered probit, multinomial logit & OLS	association between disruption and youths' outcomes was attenuated by controls for early child characteristics; divorce was negatively associated with men's education and employment and women's education and wages; parental death had weaker associations with youths' outcomes

Painter & Levine (2000)	1988-94 NELS (white, non-hispanic youths)	drop-out, college attendance, non-marital fertility	married, divorce while in high school, persistent female headship, remarriage	probit	divorce and female headship are positively associated with drop-out; headship is positively associated with non-marital fertility; evidence of selectivity
Wu (1996)	1979-89 NLSY (women)	premarital birth	married, mother only, step-family, other all measured at different ages, also transitions	piece-wise linear Gompertz hazard	transitions consistently associated with premarital births; results are not sensitive to the inclusion of economic variables
Wu & Martinson (1993)	1968-90 PSID	premarital birth	married, mother only, step-family, other all measured at different ages, also transitions	piece-wise linear Gompertz hazard	transitions consistently associated with premarital births; other family structures are not significant when transitions are included

*Adults' earnings*¹⁸

Research shows that marriage is associated with improved economic outcomes for adults and families. These outcomes include higher family incomes, greater wealth, lower rates of poverty, and reduced reliance on welfare. One explanation for these outcomes is mechanical—more adults in a household mean more potential resources. Moreover, because of economies of scale in consumption, an additional adult contributes more on average to a household's resources than to its needs. Consider the outcomes for a couple with a 9th-11th grade education and one child in 2001. The median annual income for a woman with this level of education was \$10,330, while the median annual income for a similarly educated man was \$19,434.¹⁹ If the mother and child lived apart from the father, their income would have been below the two-person poverty threshold of \$12,207; however, if the family lived together, their combined income would have exceeded the three-person threshold of \$14,255. The mother and child would have also met the gross income requirement for food stamps if they lived apart from the father but would have been ineligible if they lived with him. Even if the mother had no income and the family just depended on the father's resources, they would have been above the poverty line and ineligible for food stamps if they all lived together. Studies by Lerman (1996), Sigle-Rushton and McLanahan (2002a), and Thomas and Sawhill (2002) used mechanical methods similar to these to match unmarried women with potential mates and calculate the impact on family incomes and poverty. Their analyses all indicated that marriage would raise incomes and reduce poverty.

Another explanation, which is the focus of this section of the review, is that marriage changes the behavior or opportunities of a couple to make them more economically productive. Numerous studies have found that married men earn more than unmarried men. Research has considered whether this premium reflects a productivity effect, a selection effect or something else. The possible productivity effect is hypothesized to come through specialization or more responsible behavior. The outcomes for women are more ambiguous. Women may benefit from the resources their husbands bring in; however, their own earnings and careers are thought to suffer as a result of specialization. Thus, studies of earnings differences for women have usually been framed in the context of investigating whether there are penalties associated with marriage or motherhood. The general patterns of substantial earnings advantages for married men and modest earnings disadvantages for married women have been documented across different developed countries (Schoeni 1995) and over time within the United States (Goldin 1990).

Because earnings are an important general research area in economics, economists have been more active in this domain of the marriage literature than in others. Their involvement has resulted in a greater sensitivity to problems associated with selectivity as well as a stronger tendency to adopt techniques to address these problems.

Marriage as an indicator for productivity. Economists have long used measures of family structure, including marital status and the presence of children, as indirect controls for productivity in analyses of differences in earnings across people, especially women (see the discussion in Hill 1979). While productivity—the amount or value of output associated with a

¹⁸ The initial draft of this section was written by Sung Un Kim.

¹⁹ Figures obtained from <http://www.census.gov/hhes/income/income01/inctab7.html> and <http://www.census.gov/hhes/poverty/threshld/thresh01.html>.

given input—has a straightforward definition, it is difficult to measure at an individual level and rarely available in micro data sets.²⁰ Researchers have instead relied on indirect indicators such as education, training, or work experience to control for productivity differences. In cases where some or all of these indirect variables are unavailable, such as in analyses of the Current Population Survey, researchers have turned to indicators for family structure as proxy measures.

Hill (1979) used a special supplement of the PSID with information on worker experience, training and absenteeism to examine whether marriage and parenthood were suitable proxies for productivity. She found that marriage was positively associated with several productivity measures for men and weakly negatively associated with measures for women. Hill then regressed wages on marital status, parenthood and indicators for productivity. She found sizeable earnings differences between married and unmarried men, even after controlling for the other observed measures. For women, the results differed by race. Black women's wages were negatively associated with marriage, while white women's wages were not. White women's wages were negatively associated with motherhood in some initial specifications, but these associations disappeared when productivity controls were added. For black women, the association between motherhood and wages was actually positive in the specifications with productivity controls. Hill concluded that the positive residual effects of marriage for men and motherhood for black women reflected their responses to greater financial responsibilities.

Studies without selectivity controls. Hill's findings prompted other researchers to examine marriage directly, rather than just as an indirect control for productivity. Much of this subsequent research, including studies by Bellas (1992), Blackburn and Korenman (1994), Cohen (2002), Hewitt, Western and Baxter (2002), Kenny (1983), Loh (1996) and Schoeni (1995) relied on standard regression or discrete-choice models. The models in several of these studies included variables for alternative causal explanations, like productivity or specialization, but the analyses generally did not employ statistical controls for selectivity.²¹

For instance, Kenny (1983) revisited the productivity explanation, focusing on whether men acquired more skills after they married. Kenny conducted an indirect test of this human capital hypothesis by comparing the wage growth for men before and after they were married. He found that married men's wages grew faster than unmarried men's. Blackburn and Korenman (1994) examined whether the decline in the earnings premium for men in the United States over time was associated with aggregate indicators for specialization (the percentage of wives who worked) and selection (the percentage of males who were unmarried) but failed to find strong evidence for either hypothesis. Cohen (2002) also investigated the fall in the male marriage earnings premium but focused on the role of cohabitation. Cohen found that cohabiting men earned more than never-married men but less than married men and concluded that rising rates of cohabitation accounted for part of the decline in the marriage premium.

Hewitt et al. (2002) used cross-section data from Australia and quantile regression techniques to examine whether earnings differences associated with marriage varied across the wage distribution. Theories of specialization suggest that high-wage men should gain more from

²⁰ Government labor productivity figures come from aggregate calculations of the total output for a firm or industry divided by the total hours of work rather than from individual calculations.

²¹ An exception is the study by Loh (1996), which included results from sibling comparison (family fixed effects) models. The results from the sibling comparison models were similar to those from standard regressions.

marriage than low-wage men (Becker 1981); however, Hewitt et al. found that the earnings premium for high-wage men was actually lower than for other men. Bellas (1992) hypothesized that the direct support that married men receive from their wives gives them an advantage in the workplace. While she framed her analysis in terms of support rather than specialization, her tests relied on variables—marital status and wife’s employment—that have been used other specialization studies. She found that men’s earnings were positively associated with being married and having a non-working wife.

In contrast to these findings, Loh (1996) reported that the association between marriage and men’s wages was sensitive to different types of observed controls. Loh found that the association became small and insignificant when wives’ education and work experience were included in a regression model. He also found that the wages for self-employed workers were lower for married men than unmarried men. In addition, the wages for cohabiting men were nearly as large as those for married men. Loh’s results were based on a young and selective sample.²² Also, the inclusion of numerous characteristics of the wives, which were conditioned on marital status, complicated interpretation of the marital status variable (some of his marital status coefficients effectively represented the earnings benefit for men whose wives had never completed high school and never worked during their marriage). Despite these shortcomings, his analysis indicated that the evidence regarding the marriage premium for men might be weak.

Studies with cross-section controls for selectivity. An influential article by Nakosteen and Zimmer (1987) indicated that estimates of the earnings premium might be sensitive to controls for selectivity. Nakosteen and Zimmer estimated a switching regression specification in which the determinants of young men’s earnings differed depending on an endogenous marriage decision. Marriage was significantly associated with earnings in a standard regression model but not significantly associated with earnings in the endogenous switching model. Nakosteen and Zimmer interpreted these results as evidence of selectivity. A close examination of the results reveals, however, that the loss of significance was due entirely to an increase in the standard errors (loss of statistical precision) in the switching model and not to a change in the coefficients. Thus, the evidence regarding selection was inconclusive.

Subsequent research by Chun and Lee (2001) also adopted an endogenous switching framework but tried to improve on the identification of this model by using local marriage market conditions and mother’s country of origin to predict marital status. Chun and Lee found that marriage was positively and significantly associated with men’s wages in both standard and switching specifications. Consistent with the specialization hypothesis, they also found that wives’ work hours were negatively associated with husband’s earnings. There was little direct evidence of selectivity in their switching model.

While the results from these cross-section studies were either inconclusive or negative toward the selectivity hypothesis, other research has indicated that selectivity should be considered. For instance, Becker, Landes and Michael (1977) found evidence that men’s earnings were negatively associated with divorce. More recently, Nakosteen and Zimmer (1997)

²² His analysis was based on 1990 data from the NLSY; respondents were 25-33 years old at the time of interview. Loh had to drop nearly 40 percent of the original observations for reasons of attrition or item non-response.

found that that an increase in the male marriage earnings premium increased the chances that single men would marry and decreased the chances that married men would divorce.

Studies with longitudinal controls. In addition to the studies that have used cross-section controls for selectivity, numerous studies of the relationship between family structure and wages have employed longitudinal, fixed effects methods. Longitudinal analyses of men's wages and earnings include Cornwell and Rupert (1997), Ginther and Zavodny (2001), Gray (1997), Hersch and Stratton (2000), Korenman and Neumark (1991), Richardson (2000), and Stratton (2002), while longitudinal analyses of women's wages include Budig and England (2001), Korenman and Neumark (1992), and Waldfogel (1997). Most of the research for men has continued to find a positive association with marriage, while the research for women has been more equivocal.

Korenman and Neumark (1991) conducted a particularly careful analysis of the male earnings premium. They first estimated fixed effects regressions using data on white males from the National Longitudinal Survey (NLS) of Young Men. The controls for fixed effects reduced but did not eliminate the association between marriage and wages. Korenman and Neumark also looked at the duration of marriages and found that wage growth was especially strong in the first few years of marriage. The researchers then examined a second data source consisting of company employment records of managers and professionals. These records had performance evaluations and other information that could be used to control for productivity. Korenman and Neumark found that performance ratings were higher for married men and that controlling for performance sharply reduced the association between marriage and earnings.

Gray (1997) and Hersch and Stratton (2000) examined whether specialization could account for the marriage premium and generated conflicting findings. Gray used comparable samples from late 1970s (NLS of Young Men) and early 1990s (NLSY) and estimated fixed effects models that included measures of marital status and wives' employment. To account for the endogeneity of the latter variable, he used instrumental variables. Gray found that the use of fixed effects reduced the association between marriage and earnings. Marriage was significantly associated with earnings in the late 1970s but not in the 1990s. He attributed this change to a fall in the returns to specialization and the rise in married women's employment. Hersch and Stratton (2000) used measures of men's time in housework as a more direct indicator for specialization. They found that housework was not strongly associated with wages and that controlling for this variable did not diminish estimates of the marriage premium.

Ginther and Zavodny (2001) conducted a longitudinal analysis but distinguished between men whose marriages immediately followed a premarital conception—men in “shot-gun” marriages—and other married men. Ginther and Zavodny hypothesized that shot-gun weddings were unanticipated and therefore less selective than other marriages. They found that men in these marriages had higher earnings than never married men but lower earnings than other married men. They interpreted these findings as evidence of both selective and causal effects.

The longitudinal studies of women's earnings have produced different sets of results. Korenman and Neumark (1992) used multiple methods including standard regression, instrumental variables, and fixed effect models to examine the relationships between marriage, motherhood, and wages among white women from the NLS of Young Women. Similar to the findings of Hill (1979), the estimates from their various specifications indicated that marriage

was not significantly associated with wages. Korenman and Neumark compared estimates from the regression and instrumental variables models and found that they could not reject the assumption that family structure was exogenous. However, comparisons of the regression and fixed effects results led them to reject the regression model (mostly because the results for motherhood changed). While the use of multiple techniques to address selectivity was an important contribution, the study had some weaknesses. It relied on a short, two-year observational window for the main fixed effects analyses and used family background, expectations, and attitudes as instruments for marriage and motherhood.

Waldfogel (1997) also used information from the NLS of Young Women but considered black and Hispanic women in addition to white women and 15 years of data (covering the period 1968-1988) instead of two. Waldfogel's fixed effects estimates indicated that married women enjoyed a small (3-4 percent) but statistically significant wage advantage over never-married women but a slight (1-2 percent) disadvantage relative to divorced women. Waldfogel found that the coefficient on marriage was insignificant in specifications like those used by Korenman and Neumark (1992) that used short, 1-2 year differences in the data to control for fixed effects but significant when longer differences or mean differences were used. An analysis by Budig and England (2001) of data from the NLSY also found a small but statistically significant positive association between marriage and wages for women.

One possible explanation for the different findings across these studies and the studies for men is that the benefits of marriage may vary with age. People who marry later in life may gain more initial job experience and work skills than people who marry earlier and enjoy a higher earnings premium as a result. A preliminary analysis by Loughran and Zissimopoulos (2003) used fixed effects methods to examine this issue. Loughran and Zissimopoulos found that early marriages were associated with lower wages for men and women and that delays in marriage led to higher wages. Marriage was eventually associated with earnings advantages but not until after age 20 for men and age 25 for women.²³

Assessment. Numerous studies, including several with statistical controls for selectivity, have found that marriage is positively associated with men's earnings. The evidence regarding women's earnings is less clear with studies finding positive and negative associations as well as no associations at all. Many of the studies that have examined the issue have also concluded that selectivity is a substantive concern. While controls for selectivity have not generally eliminated the association between marriage and earnings, they have often impacted the estimates. The remaining associations are consistent with possible causal effects.

As a methodological issue, the evidence of selection from the existing studies warrants more attention. Studies have accounted for longitudinal sources of selection but have not looked as closely at time-varying sources. The research by Ginther and Zavodny (2001), which supplemented longitudinal controls with comparisons between "shot-gun" and other marriages, suggests that additional controls may be needed. Qualitative research by Edin (2000), which indicated that low-income mothers put off marriage until their partners hold stable and higher-

²³ Bergstrom and Schoeni (1996) used standard regression methods and also found that annual earnings for men and women were negatively associated with early marriages. They interpreted their results in terms of a "waiting game" in which people delay marriage until the productivities of potential spouses have been revealed.

paying jobs, also indicates that reverse causality needs to be considered. Comprehensive reviews of the research on marriage and adults' earnings by Chiodo and Owyang (2002, 2003) and Lerman (2002c) have similarly concluded that the research in this area can be strengthened.

More substantively, the recent studies of women's earnings by Budig and England (2001) and Waldfogel (1997) are interesting because they challenge the notion that marriage is a mixed blessing for women. These studies, which included longitudinal controls for selectivity, found that marriage had a modest beneficial association with women's earnings. Thus, women might not only gain from having access to their husbands' earnings but also from higher earnings of their own. If these findings hold up to more rigorous statistical tests they would provide compelling and unambiguous evidence of the economic benefits of marriage.

Table 3. Selected Empirical Analyses of the Effects of Marriage on Men’s Earnings

Study	Data	Dependent Variable	Key Explan. Vars.	Technique	Findings
Bellas (1992)	1984 faculty survey (3,694 men)	annual (academic year) salary	never married, formerly married, currently married with and without a working wife	OLS	marriage was positively associated with earnings; earnings were highest for men with nonemployed wives
Blackburn & Korenman (1994)	1968-89 March CPS (non-widowed men 25-54 employed full-time, full-year)	log annual wage	married, never married, divorced/separated	OLS	marriage premium fell over time; trends were not strongly related to aggregate marriage rates
Chun & Lee (2001)	1999 March CPS (unenrolled working males, 18-40)	log hourly wage	married, never married	switching regression	marriage was positively associated with wages; there was no evidence of selection; wife’s work hours were negatively associated with husband’s wages
Cohen (2002)	1976-99 CPS (unenrolled, employed men 25-24)	log hourly wage	never married, cohabiting, married	OLS	wages for cohabiting men were between those of married and never-married men; cohabitation accounted for part of the decline in the marriage premium over time
Cornwell & Rupert (1997)	1971-80 NLS Young Men (666 employed white men, 19-29 in 1971)	log hourly wage	married, never married, divorce, years married/divorced	fixed effects	rejected random effects specification; wages for married and divorced men were similar and above those of never married men
Ginther & Zavodny (2001)	1970-76 NLS Young Men (530 employed white, non-divorced men, 20-30 in 1970)	log hourly wage	married, never married, “shot-gun” marriage	fixed effects & natural experiment	wages for married men, including men in “shot-gun” marriages were higher than those of never-married men
Gray (1997)	I: 1976-80 NLS Young Men (1,248 white men, 24-31 in 1976)	log hourly wage	married, never married, formerly, length of marriage	fixed effects	marriage was positively associated with wages; fixed effects reduced association

	II: 1989-93 NLSY (1,611 white men, 24-31 in 1989)	same	same	same	marriage was not significantly associated with wages; attributed change to reduced specialization
Hersch and Stratton (2000)	1987-94 NSFH (1,373 unenrolled employed white men, 18-59)	log hourly wage	married, never married, formerly married, years married	fixed effects	used housework as a direct control for specialization; accounting for housework did not reduce the marriage premium
Hewitt, Western & Baxter (2002)	1996-97 National Australian Survey (583 men 18-54 employed full-time)	log annual income	married, never married, formerly married	quantile regression	marriage was associated with higher incomes at all but highest income level
Hill (1979)	1976 PSID (5,212 heads and spouses, 18-64, working 500+ hours per year)	log hourly wage	married, never married, divorced, separated, widowed, number of children	OLS	marriage was associated with higher wages for men; results were robust to the inclusion of controls for productivity
Kenny (1983)	1969 Coleman-Rossi Retrospective Life Histories Study (1,233 employed men, 30-40)	average growth rate of monthly earnings	married, unmarried	OLS	wages grew faster after men were married than before
Korenman & Neumark (1991)	I: 1976-80 NLS Young Men (1,228 employed, white men, 24-34 in 1976)	log hourly wage	married, never married, formerly married, years married	fixed effects	marriage was positively associated with wages; associations increased with the length of the marriage
	II: 1976 company personnel records (8,235 white managers and professionals)	log annual salary	married, never married, divorced, widowed	OLS	data included rich performance measures; accounting for performance sharply reduced the association between earnings and marriage
Loh (1996)	1990 NLSY (2,626 employed white men, 25-33)	log hourly wage	married, never married, formerly married	OLS & siblings fixed effects	marriage was not significantly associated with wages when wives' work history or siblings controls were used

Loughran & Zissimopoulos (2003)	1979-2000 NLSY (1,598 ever-married men 35-43 in 2000; excludes military & poor white oversamples & people with nonmarital first births)	log hourly wages	ever-married, never-married, age at first marriage; ever-divorced	fixed effects	later age at marriage was associated with higher wages for men; threshold for wage premium occurred around age 20; divorce was associated with lower wages
Nakosteen & Zimmer (1987)	1977 PSID (576 employed men, 18-24)	log annual earnings	married, unmarried	switching regression	marriage was not significantly associated with earnings in switching models; however, results reflected larger standard errors rather than a change in the coefficients
Richardson (2000)	1968-91 Swedish Level of Living Survey (employed men, 18-65)	log hourly wage	Married, cohabiting, divorced or widowed	fixed effects	Wage premiums positively associated with marriage; little evidence of selection
Schoeni (1995)	1979-86 Luxembourg Income Data Sets (male household heads 24-55 from 14 countries)	log annual earnings	married, unmarried for all countries & married, never married, separated, divorced, widowed for some countries	OLS	Marriage was associated with higher earnings in all 14 countries and significantly associated in 12
Stratton (2002)	1987-94 NSFH (1,358 employed white men, earnings \$2.50-100/hour, 18-64)	log hourly wage	married, never married, formerly married, cohabiting, years in different arrangements	fixed effects	wages were higher for married and formerly married men but not cohabitators

Table 4. Selected Empirical Analyses of the Effects of Marriage on Women's Earnings

Study	Data	Dependent Variable	Key Explan. Vars.	Technique	Findings
Budig & England (2001)	1982-93 NLSY (5,287 women with 2+ years of employment, earnings \$1-200/hour)	log hourly wage	Never married, married, formerly married	fixed effects	article focused on wage penalty for mothers; married women's wages were higher than never married women's and about the same as formerly married women's
Hewitt, Western & Baxter (2002)	1996-97 National Australian Survey (470 full-time and 325 part-time women, 18-54)	log annual income	married, never married, formerly married	quantile regression	marriage was not significantly associated with women's incomes
Hill (1979)	1976 PSID (5,212 heads and spouses, 18-64, working 500+ hours per year)	log hourly wage	married, never married, divorced, separated, widowed, number of children	OLS	marriage was not significantly associated with wages once work history was taken into account; motherhood was associated with higher wages for black women
Loughran & Zissimopoulos (2003)	1979-2000 NLSY (1,390 ever-married women 35-43 in 2000; excludes military & poor white oversamples & people with nonmarital first births)	log hourly wages	ever-married, never-married, age at first marriage; ever-divorced	fixed effects	later age at marriage was associated with higher wages for women; threshold for wage premium occurred around age 25; divorce was associated with higher wages
Korenman & Neumark (1992)	1980-82 NLS Young Women (911 employed white women, 28-38 in 1982)	log hourly wage	married, never married, divorced/separated	OLS, fixed effects & IV	marriage was not significantly associated with wages; tested but could not reject exogeneity of family structure
Waldfogel (1997)	1968-88 NLS Young Women (2,133 employed women, 14-24 in 1968)	log hourly wage	married, never married separated, divorced, widowed	fixed effects	article focused on wage penalty for mothers; married women's wages were higher than never married women's but lower than divorced women's

Adults' physical health and mortality

This section reviews studies from the large literature on the relationship between marriage and adults' physical health and mortality. Marriage has consistently been found to be strongly associated with better health and lower mortality. Summarizing the evidence on marriage and other social relationships, House, Landis and Umberson (1988 p. 541) wrote

The theory and evidence on social relationships and health increasingly approximate that available at the time of the U.S. Surgeon General's 1964 report on smoking and health, with similar implications for future research and public policy.

Gallagher and Waite (1990, p. 47) similarly drew a parallel between the health risks of smoking and being single. Along with the warnings on cigarette packs, they suggested that the Surgeon General consider affixing a warning to divorce decrees that stated, "Not being married can be hazardous to your health."

As with the literature on children's well-being, the research literature on health outcomes is enormous. Accordingly, the review concentrates on general methodological developments. It also narrows its focus to physical health and mortality and leaves aside other important health outcomes such as psychological well-being and domestic violence.

Early research. The research linking marriage to health and mortality constitutes one of the oldest empirical literatures in the social sciences. Contributions from the nineteenth century included William Farr's analysis (referenced in Goldman 1993) of marriage and mortality in France and Emile Durkheim's (1897 [1951]) investigation of social isolation and suicide. In the United States, the Census offices—the predecessors to the permanent Bureau of the Census—began reporting differentials in mortality by marital status for the 1890 and 1900 censuses, and the Census Bureau continued to produce occasional reports through the 1920s (Klebba 1970). A larger and more complete analysis of mortality differentials was included in a 1933 research volume by Walter Willcox (referenced in Klebba 1970). Besides the reports from the Census Bureau, American researchers and policymakers during this period also appear to have recognized that widows were experiencing worse health and higher mortality than other women (see Berardo's 1968 review). Nevertheless, it is fair to say that with the exception of Durkheim's findings, early studies of the relationship between marital status and mortality went largely unnoticed by other social scientists and initially prompted little additional research.²⁴

Improvements in the quality and completeness of vital statistics data were catalysts for a new round of studies by U.S. health researchers in the middle of the century. Starting in 1933, mortality data were systematically reported for all states (Hetzl 1997). At the next census in 1940 when the mortality data could be combined with detailed population data, research volumes examined differentials by marital status in general mortality rates, specific causes of death, and cancer-related deaths (see Klebba 1970). Data refinements continued after the enactment of the Model State Vital Statistics Act in 1942 and after responsibility for preparing the statistics was

²⁴ An exception is William Fielding Ogburn (1927). Foreshadowing arguments that Waite and others would advance some 70 years later, Ogburn (p. 6) concluded that the available evidence led "to the presumption ... that marriage is a desirable state with regard to these social conditions of death, crime, and insanity."

transferred to the Public Health Service in 1946 (Hetzel 1997). The primary empirical methodology during this period consisted of comparing age-, gender-, and race-specific mortality rates for people in different marital statuses and reporting results in terms of relative mortality ratios (e.g., mortality rates of never married women relative to those of married women, widows relative to married women, etc.). Although the analyses relied on descriptive methods, the cross-tabulations were so detailed that the results could effectively be viewed as regression coefficients. Analyses of this sort by Berkson (1962) and Sheps (1961) showed that unmarried people had higher adjusted rates of mortality than married people.

Sociologists and family demographers returned to this research field in the late 1960s and early 1970s. Some research continued to rely on descriptive analyses, while other studies began to employ multivariate regressions. For instance, Berkman (1969) used data from a 1965 household survey from Alameda County, California to regress health outcomes for mothers against indicators for whether they were married, widowed or divorced/separated. He hypothesized that unmarried motherhood was a source of stress, which contributed to poor health outcomes. His findings of worse health among unmarried mothers were consistent with this hypothesis. In a later analysis of the same survey, Renne (1971) examined health outcomes for all ever-married respondents and further distinguished between people who were happily or unhappily married. She found that happily married people had the best health outcomes and that unhappily married people had the worst, below even those of divorced people.

A National Center for Health Statistics (NCHS) report by Klebba (1970) examined age-, race- and gender-specific mortality rates for people in different marital statuses using 1959-61 national data. Klebba found that unmarried women had age-adjusted mortality rates that were one third to three-quarters higher than those of married women, while unmarried men had rates that were one half to more than two times larger than those of married men. These general patterns applied to whites and non-whites, though mortality differentials for non-white widows were higher than those for white widows. Gove (1973) re-examined these data, focusing just on whites. Gove hypothesized that psychological well-being differed across men and women in different marital roles and that these differences translated into physical health and mortality differences. He interpreted results that showed particularly large differentials for deaths from suicide, homicide, accidents, and cirrhosis of the liver as supporting this explanation.

Like Gove, Verbrugge (1979) also summarized tabulations that had been reported earlier by the NCHS and Census Bureau. However, instead of mortality, she considered health and morbidity indicators including reports of acute conditions, physical limitations, disabilities, hospitalizations and medical visits. She framed her analysis in terms of a sociological model of illness behavior and found that divorced, separated and widowed people had worse health on average than married and never-married people across a broad range of age-adjusted indicators.

Other studies during the 1970 also generally applied either descriptive methods or regressions to cross-section data and interpreted results in terms of marriage reducing stress or improving psychological well-being. While selection was recognized as an alternative explanation for the positive association between marriage and health, studies did not directly control for this. Limited information in some of the data sets, such as mortality records, also meant that many studies were not able to control for important covariates like socioeconomic status that were associated with both health and marital status.

Later studies. An important methodological advance was the emergence of prospective, longitudinal, individual-level data in the early 1980s. With these data, researchers could temporally order outcomes and examine how marital status in one period affected health or mortality in another. House, Landis and Umberson (1987) reviewed results from several early longitudinal studies, which continued to show that marriage and other measures of social integration were positively associated with physical health. Although these studies addressed some potential problems associated with correlations among the unobserved contemporaneous determinants of marriage and health, they did not employ fixed effects methods that could have accounted for more general correlations arising from unobserved permanent characteristics.

Researchers also examined alternative hypotheses for protective factors associated with marriage. For instance, Umberson (1987) proposed a model of social control in which spouses act to promote the health and regulate the behavior of their mates. Researchers also used regression methods to test more detailed aspects of hypotheses. An example of this type of research is Anson's (1989) study of the effect of social isolation on health, which included indicators for the presence of other people and for social interaction along with measures for family structure. Adding these controls weakened the statistical relationship between marital status and health. For some outcomes, like chronic morbidity, the relationships were eliminated, while for other outcomes, like self-rated health and bed disability days, the relationships remained. Ross, Mirowsky and Goldstein (1990) reviewed other studies from the 1980s that considered explanations associated with isolation, social support and economic resources.

As research moved into the 1990s, specialized models, such as hazard and survival models, began to be used more routinely to analyze longitudinal mortality data. Studies by Coyne et al. (2001), Hemstrom (1996), Korenman, Goldman and Fu (1997), Lillard and Panis (1996), Lillard and Waite (1995), Murray (2000), Rogers (1996), Smith and Waitzman (1994), Smith and Zick (1994), Sorlie, Backlund and Keller (1995), Tower, Kasl and Darefsky (2002) and Zick and Smith (1991) employed hazard models. Almost of all of this research found that being married, being in a satisfying or close marriage, or being in a long-lasting marriage was associated with lower mortality.²⁵

Longitudinal data were also used to examine the relationship between family structure and health status. However, most studies (e.g., Goldman, Korenman and Weinstein 1995; Michael et al. 2001; Umberson 1992; and Waldron, Hughes and Brooks 1996) continued to employ standard OLS or binary choice models with controls for baseline health and marital status. Only a few researchers adopted more specialized longitudinal statistical models. In particular, Lillard and Panis (1996) used a random effects model to examine a categorical measure of health status; Wickrama et al. (1997) used a longitudinal structural equations (latent growth curve) model to examine an index of health problems; and Wickrama et al. (2001) used a logistic hazard to examine the onset of hypertension. These studies continued to find that marriage and marital quality were associated with better health.

An additional data-related contribution was the finding by Korenman, Goldman and Fu (1997) that marital status could be misclassified in some types of prospective longitudinal data.

²⁵ Some exceptions were findings by Lillard and Waite (1995) that widowed women had lower mortality than other women and Smith and Zick (1994) that divorced men had lower mortality than continuously married men.

Korenman et al. examined the relationship between widowhood and mortality in the 1984-90 Longitudinal Study of Aging (LSOA), a survey of people who were 70 years of age and older at baseline. They discovered that information on marital status for survey respondents who died between interviews was not updated after the last interview. Thus, the survey missed some transitions to widowhood or other marital statuses. The problem was especially severe in the LSOA because interviews occurred biennially and because the LSOA relied on proxy reports for many respondents. A comparison with death certificate data revealed that the LSOA misclassified marital status for 40 percent of men and 28 percent of women.

Besides these longitudinal developments, researchers in the 1990s also began to look more systematically at evidence regarding selection. Early in the decade, Goldman (1993) used simulation methods to demonstrate that selectivity could lead to many different types of biases in analyses of mortality. Goldman's analysis dispelled an argument advanced by some researchers, such as Gove (1973), that selectivity leads to easily recognizable patterns in the data. Her research instead showed that selectivity has to be addressed directly. Unfortunately, while selection subsequently received more attention in this literature, almost all of the studies of selection effects relied on a crude strategy of reversing the dependent and independent variables in their analyses and regressing marital status on health status. Studies by Booth and Johnson (1994), Fu and Goldman (1996), Murray (2000), and Waldron, Hughes and Brooks (1996) followed this approach. Although several of these analyses also reversed the temporal ordering of the variables (e.g., examined the relationship between health status at baseline on subsequent marriage), their methods still essentially amounted to regressing one selective or endogenous variable on another.

To my knowledge, only one observational study of marriage status and health has used formal statistical controls to account for selectivity. Lillard and Panis (1996) used 1968-90 data from the PSID to estimate hazard models of marriage formation, marriage dissolution and mortality along with longitudinal, ordered probit models of health status. These models incorporated Lillard's (1993) random effects approach to account for correlation in the unobserved determinants of the outcomes. Lillard and Panis found that health and marital status were each selective determinants of the other. They also found that marriage was negatively associated with mortality but could not reject the restriction that marriage was an exogenous determinant. Although the study had many strong features and advanced the methodological frontier, it had some notable shortcomings. One was that the much of the analysis was based on a short span of seven years (1984-90) during which health status data were available in the PSID. Another was that the analysis of the simultaneous relationship between marital status and health relied on questionable variable exclusion restrictions. For instance, family background measures for parents' education and poverty status were included in the health status model but not the marriage models. Years since leaving school entered the marriage formation equation but not the marriage dissolution or health equations.

Stronger evidence of a causal effect of marriage on health comes from a comprehensive review of separate "biopsychosocial" studies that link marital interactions to stress and stress to illness. Kiecolt-Glaser and Newton (2001) summarized evidence from 64 studies published during the 1990s and interpreted the results in terms of a detailed conceptual model. None of the studies that they considered tested all of the elements of the model; however, like Amato in his (1993) review of studies on children's adjustments to divorce, Kiecolt-Glaser and Newton

examined whether different individual results supported the model. To evaluate the link between marital interactions and stress, they described research that measured couples' blood pressure, heart rates and other outcomes during experimentally-induced conflict sessions. In these sessions, couples were either asked to discuss a topic that they identified as an area of disagreement or told to take opposing sides and resolve a hypothetical argument. Measurements were compared either to a baseline or to measurements from couples who were not put into conflict situations. These experiments generally found that conflict heightened stress. Some researchers also observed whether couples exhibited hostile or behaviors during the conflict sessions and found that poorer interactions led to larger physiological changes. Kiecolt-Glaser and Newton considered whether these experimentally-observed changes were large enough to affect physical health and concluded that links might exist with outcomes that are associated with stress like cardiovascular disease, immune function and wound healing.

Assessment. Studies that have examined the relationship between marriage and health have been remarkably consistent in their findings. Being married and being in a satisfying marriage are positively associated with health and negatively associated with mortality. The findings have been reported over time in the United States and across other countries as well (Hu and Goldman 1990). The introduction of longitudinal data and some advanced statistical techniques, such as were employed by Lillard and Panis (1996), have not overturned these long-standing findings.

The contributions by Lillard and Panis and near-uniform findings of the literature notwithstanding, the existing body of research on the effects of family structure on physical health and mortality still has to be judged to be methodologically weak. Although several recent studies have adopted more sophisticated statistical methods like hazard, longitudinal and structural equations models, research based on simple descriptive analyses (e.g., Luoma and Pearson 2002; Murphy, Glaser and Grundy 1997) continues to be published in peer-reviewed journals. Even when advanced techniques are used, the studies in this area generally do not address issues associated with selection and omitted variables bias. A recent review by Hummer, Rogers and Eberstein (1998) of the sociodemographic studies on mortality similarly concludes that the research in this area, while evolving, remains too descriptive. Because of the literature's methodological weaknesses, its findings have to be interpreted with a fair degree of caution.

The research on marital interactions, as opposed to marriage status and family structure, contains more solid evidence of a causal link to health outcomes. This research, which has included some experimental designs, has been able to show a link to stress. Although these experimental results have not been convincingly tied to larger health outcomes, they suggest that programs that improve relationship skills and marital interactions might improve health. This would seem to validate the Bush administration's emphasis on promoting "healthy" marriages.

Table 5. Selected Empirical Analyses of the Effects of Marriage on Physical Health

Study	Data	Dependent Variables	Key Explan. Vars.	Technique	Findings
<i>Early studies</i>					
Berkman (1969)	1965 survey from Alameda, CA (1,832 mothers 16+)	chronic conditions, functional disabilities, self-rated health status, bothered by serious problems	Married, widowed, divorced/separated	OLS	better health outcomes for married mothers; interprets results in terms of stress
Renne (1971)	1965 survey from Alameda, CA (ever married adults 16+)	chronic conditions or disabilities, self-rated health status	marital satisfaction (5 items); married, divorced/separated, remarried	descriptive analysis	marriage and marital happiness positively associated with health; unhappy marriages worse than divorces; remarriages worse than first marriages
Verbrugge (1979)	Tabulated data from 1960, 70 Census, HIS and HES	acute conditions, activity limitations, disabilities, medical visits, hospital days	married, spouse absent widowed, divorced, never married	descriptive analysis	marital status negatively associated with numerous health problems
<i>Later studies</i>					
Anson (1989)	1979 NHIS (25,542 white women, 18-55)	self-rated health status; chronic and acute morbidity; doctor visits	currently married, widowed, never married, or divorced	OLS	worse outcomes for women who were never married and divorced; results are weaker in models that control for proximate adults
Goldman, Korenman & Weinstein (1995)	1984-90 LSOA (people over 70 in 1984; 2,847 males, 4,631 females)	disability	married, widowed, divorced, never married	logit	widowed men at higher risk, divorced men and never married women at lower risk of disability in models that control for SES and social environment
Hahn (1993)	1987 NMES (9,356 women, 18-64)	general health status (5 item scale)	married, widowed, divorced, separated, never married	OLS	health status better for married women than other women; accounting for marital acquisitions reduces but does not eliminate association

Lillard & Panis (1996)	1968-90 PSID	self-rated health status	married, divorced, separated, widowed, never married, interactions with age	ordered probit with correlated random effects	positive effects of marriage on health
Michael, Berkman, Colditz & Kawachi (2001)	1992-96 Nurses' Health Study (women aged 60+)	change in physical function, vitality, and mental health	living alone, with a spouse or with others	logit	Less decline in vitality and mental functioning among women living alone
Murphy, Glaser & Grundy (1997)	1991 British Census	long-term illness	first marriage, remarried, divorced, widowed, never married	descriptive analysis	fewer long-term illnesses among men and women in first marriages; more among widowed and divorced men and women; more among never marrieds before age 75
Umberson (1987)	1974-5 survey with 1,826 usable ever-married observations	risk-taking scale, orderly lifestyle, alcohol problems and drug abuse	married, widowed, divorced	OLS	divorce positively associated with risk taking and drinking problems; divorce and being widowed negatively related to orderly lifestyle
Umberson (1992)	1986-89 Americans' Changing Lives survey (2,867 people 25+)	smoking, drinking, BMI, sleep, physical activity	married, widowed, divorced, never married	OLS	higher rates of smoking and drinking among divorced adults; lower BMI among divorced women; less physical activity among divorced men
Waldron, Hughes & Brooks (1996)	1968-88 NLS young women (covers ages 24-44)	health problems (17 item scale)	married, unmarried	OLS	initial and later marital status negatively associated with health problems
Wickrama, Lorenz, Conger & Elder (1997)	1989-94 survey from rural Iowa (364 couples married at least 14 years)	physical illness index (counts of illnesses and symptoms)	marital quality (7 items)	latent growth curve analysis	intercepts and slopes for marital quality and physical illness negatively correlated for men and women
Wickrama, Lorenz, Wallace, Peiris, Conger & Elder (2001)	1989-94 survey from rural Iowa (367 wives, 340 husbands married at least 14 years)	onset of hypertension	marital stress based on self-reports and observations	logistic hazard	higher initial marital stress speeds onset of hypertension for men and women

Table 6. Selected Empirical Analyses of the Effects of Marriage on Mortality

Study	Data	Dependent Variables	Key Explan. Vars.	Technique	Findings
<i>Early studies</i>					
Gove (1973)	1959-61 tabulated mortality rates (whites 25+)	age- and gender-specific mortality rates for different causes of death	married, widowed, divorced, never married	descriptive analysis	age-specific mortality rates higher for unmarried people than married people; differences larger for men than women
Klebba (1970)	1959-61 tabulated mortality rates (people 15+); compares with 1940 and 1949-51 data	age-, race- and gender-specific mortality rates for selected causes of death	married, widowed, divorced, never married	descriptive analysis	excess mortality for all unmarried groups; differentials higher for men than women; differentials for non-white widows higher than white widows; other differentials similar across races
<i>Later studies</i>					
Coyne et al. (2001)	1993-95 Congestive Heart Failure Patients in Michigan (193 men, 50 women)	mortality	relationship quality (based on self-reports and interviewer observations)	Cox proportional hazard	higher quality marriages associated with lower mortality
Goldman, Korenman & Weinstein (1995)	1984-90 LSOA (people over 70 in 1984; 2,847 males, 4,631 females)	mortality	married, widowed, divorced, never married	logit	no significant results
Hemstrom (1996)	1981-86 Swedish Census-Linked Deaths Registry (married people 20-49 in 1970 living in 1980)	age at death	still married, remarried, cohabiting, widowed, divorced, separated	hazard model	mortality higher for all dissolution groups; highest for people who were separated or divorced
Hu & Goldman (1990)	1940-85 mortality rates from 16 countries (ages 20-64)	age-, marriage-, and gender-specific mortality rates	married, widowed, divorced, never married	log linear models	divorced men have highest adjusted mortality; excess mortality of unmarried people increasing over time

Korenman, Goldman & Fu (1997)	1984-90 LSOA (people over 70 in 1984)	mortality, age at death	married, widowed in last year, widowed earlier, divorced, never married in baseline year and later	logit and hazard models	no significant effects of baseline marital status; some negative effects of being recently widowed
Lillard & Panis (1996)	1968-90 PSID	age at death	married, divorced, separated, widowed, never married, interactions with age	cont. time hazard with correlated random effects	negative effects of marriage on mortality; could not reject exogeneity of marital status
Lillard & Waite (1995)	1968-85 PSID (people 10+)	age at death	married, divorced, separated, widowed, never married, marital duration	cont. time (Gompertz) hazard	higher mortality for never married men and separated women; lower mortality for widowed women; mortality decreases with marital duration
Luoma & Pearson (2002)	1991-96 Multiple-Cause-of-Death files (people 20+)	age-, marriage-, race- and gender-specific suicide rates	married, widowed, divorced, never married	descriptive analysis	suicide lower among married adults; high rates among widows, especially young widows
Murray (2000)	Alumni records for Amherst college (born 1832-79)	age at death	marriage	hazard models	marriage decreases mortality
Rogers (1996)	1984-91 NHIS (15,938 people 55+)	age at death	married, previously married, never married interacted with other living arrangements	discrete logistic hazard	mortality higher for married people living with other relatives and previously married people living alone or with other relatives
Smith & Waitzman (1994)	1971-75 NHANES I with 1982-84 NHEFS (people 25-74 at baseline)	age at death; distinguish several causes of death	married, widowed, divorce/separated, never married; also interacted with poverty status	proportional hazards	among people initially 25-64, mortality higher for divorced/separated men and never married women; evidence of reinforcing interactions between poverty and marital status for men
Smith & Zick (1994)	1968-87 PSID (couples in first marriages in 1968 with men aged 35-64)	age at death	divorced	paired hazard	divorce increases risk of death for wives but decreases risk for husbands

Sorlie, Backlund & Keller (1995)	1979-89 National Longitudinal Mortality Survey (people 25+)	age at death	married, widowed, divorced, separated, never married at initial interview	Cox proportional hazard	marriage reduced mortality for men and women; results were only significant for young men and older women
Tower, Kasl & Darefsky (2002)	1982-88 Yale Health & Aging Project (married couples 65+)	age at death	marital closeness (name spouse as confidant or source of emotional support)	Cox proportional hazard	spouses in relationships in which the wife named the husband as source of support had lower mortality; spouses in relationships in which the husband named the wife had higher mortality
Zick & Smith (1991)	1971-84 PSID (401 female & 518 male decedents; 611 female and 460 male survivors in 1984)	age at death	married, widowed, divorced, never married two years earlier; transitions to married, widowed divorced in past year	discrete logistic hazard	no significant effects of marriage for women; higher mortality for men who were divorced, widowed, or became divorced

V. Conclusion

This review has examined quantitative research on the benefits of marriage on children's well-being, adults' earnings, and adults' physical health and mortality. It has considered theories of how marriage might directly affect these outcomes and offered alternative explanations of how the association between marriage and well-being might reflect selectivity from omitted variables and reverse causality. The review has discussed a general set of statistical techniques that have been developed to account for different forms of selectivity and examined how these techniques have been applied in empirical studies of marriage.

Consistent with the findings of previous reviews, there is evidence throughout the literature that marriage is associated with positive outcomes for adults and children. Most of this evidence comes from regression analyses that account for observed covariates but do not account for other sources of selectivity. More and more evidence, however, is coming from studies that employ other controls for observed and unobserved selectivity, such as longitudinal comparisons, instrumental variables methods, switching regression models and matching methods. Researchers have generally found that the use of such controls reduces the association between marriage and well-being, which validates initial concerns regarding selectivity. However, in many instances, they have also found that significant associations remain, which is consistent with causal explanations.

What the research from each domain can contribute to the others

The research in each domain has its strengths and weaknesses. This means that there are a number of promising opportunities for cross-pollination across the domains, with each domain contributing important results and useful insights that would benefit the others.

Statistical methods. Statistical methods for addressing selectivity have been applied in each of the three domains that were reviewed. However, the techniques have been and continue to be used more frequently in some areas than in others. The research on adults' earnings has been especially sensitive to the issue of selectivity. The published research in this area not only includes numerous individual applications of various approaches but also some multi-method studies (e.g., Korenman and Neumark 1992), which compare results from different techniques using the same data set. The use of statistical controls for selectivity has become standard practice in this literature, and current research focuses on refining these controls and testing more detailed aspects of the causal explanations associated with specialization and productivity.

While statistical techniques for addressing selectivity have also been applied in the research on children's well-being and adults' physical health and mortality, their use is not standard practice. Many studies continue to rely on methods that do not control for selectivity. Of the studies that do address selectivity, most just augment a standard model with a lagged value of the outcome variable. For example, researchers have specified models in which a child's behavior after being exposed to the risk of a family disruption depends on whether a disruption occurred, prior behavior, and other variables. Few studies in these domains have employed specialized methods, like fixed-effects models, for longitudinal data. Only a handful have adopted other approaches like instrumental variables, matching, or switching regressions.

Research on children's well-being and adults' health and mortality would be strengthened by a more careful consideration of the selectivity issue. This not only means adopting statistical techniques from other fields but also building on the experience of other researchers regarding the appropriate use of these techniques. As this review has shown, different approaches rely on different underlying assumptions and can give different answers. Unless they are careful, researchers may address one source of bias only to unmask another. While it is impossible to test every assumption that underlies a model, many can and should be tested. In addition, researchers should examine whether there are statistically meaningful differences in estimates across models, that is, formally test the robustness of their results.

Instrumental variables. Instrumental variables methods are attractive to researchers because they can address bias from several sources including omitted variables and reverse causality. Their application has been limited, however, by the practical problem of finding suitable variables that are correlated with marriage but otherwise unrelated to well-being. Results from several studies have been undercut by questionable choices regarding instruments.

The most promising source of instruments for current research on well-being outcomes in the United States appears to be state laws and policies. For example, researchers have found that unilateral, no-fault divorce laws are associated with worse outcomes for children (Gruber 2000) and a higher incidence of spousal homicide (Dee 2003). A natural interpretation of these results is that the laws have led to higher rates of marital dissolution and consequently to other negative outcomes. This is not the only interpretation, however, as the laws could have also affected a host of other marital outcomes including entry into marriage, marriage quality, bargaining within marriage, and divorce settlements (Mechoulan 2000; Peters 1986). Other changes in marriage policies that states have made or are considering, such as establishing covenant marriages and eliminating disincentives in tax and subsidy programs (Gardiner et al. 2002), may provide additional instruments. These could be used in conjunction with divorce laws to identify particular types of family structure and family process effects.

Certain types of experiments could provide additional sources of exogenous variation. Kiecolt-Glaser and Newton (2001) have described experiments involving couples' physiological reactions to small levels of induced conflict. These experiments could be extended to examine the effects of conflict on parenting behaviors or job functioning. The U.S. Department of Health and Human Services will soon undertake experimental evaluations of several types of programs designed to improve relationships and strengthen marriages. The three largest demonstrations will include programs to (a) provide services to unmarried new parents, (b) strengthen relationships among low-income married couples, and (c) assist communities in promoting healthy marriages. The Bush administration's welfare reauthorization proposal would provide resources for additional demonstration projects. Beyond the immediate analysis of these programs on marital functioning, the experimental demonstrations might provide a set of instrument conditions that could be used to examine the effects of marriage on other well-being outcomes (Gennetian et al. 2002). The variety of programs that may be examined as well as differences in the types of families that will be served may also be helpful in identifying different aspects of marriage.

While better instrumental variables may improve research, analysts still need to be mindful of other potential problems with this technique. Most importantly, researchers need to

be aware that the approach may not be appropriate in circumstances in which the effects of marriage differ across individuals and individuals take these effects into account in choosing to marry (Heckman et al. 1999).

Relationship quality. Numerous studies have found that children's well-being and adults' health and mortality are strongly related to both the existence and quality of a marital relationship. Some of the results regarding relationship quality from the health literature are particularly compelling because they come from experiments in which conflict between couples was artificially and randomly manipulated. Besides the policy implication that the Department of Health and Human Services is on the right track in promoting *healthy* marriages rather than marriages generally, the results also have some methodological implications. First, considerations of relationship quality, which have been almost entirely absent from the research on earnings, should be studied more carefully in this literature. Some sociological and family studies have examined how conflict relates to gender roles, specialization, and women's employment; however, these results have not found their way into wage studies. Second, as mentioned, marital quality is a dimension along which some degree of experimental manipulation is possible. Thus, it may be possible to sidestep selection issues altogether in this type of research. Third, upcoming experimental demonstrations of programs to build relationship skills and help couples prepare for marriage will provide an excellent opportunity to re-examine findings regarding marital quality. The demonstrations have the potential to make larger and longer-term changes in marital functioning than the laboratory experiments; these changes should make it easier to detect effects in other well-being outcomes. The demonstrations will also occur at a scale and in settings that will make it easier to generalize findings.

Direct and indirect effects. Much of the research on the association between family structure and well-being has included controls for mediating factors. For instance, the research on children's outcomes has considered factors such as economic resources, family conflict, parenting practices, community characteristics, and residential mobility. Studies of earnings have included controls for productivity, work habits, family specialization, and instrumental support. Research on adults' physical health and mortality has examined economic resources, conflict, stress, isolation, and social control. In each domain, the research has generally shown that the addition of these variables leads to smaller estimates of the direct association between family structure and well-being. While mediators have been valuable for examining the sensitivity of family structure to alternative specification assumptions, researchers have not been as careful in considering this aspect of their empirical models as others.

Many of the mediators that have been examined are behaviorally determined and therefore endogenous. A few studies, such as Gray's (1997) analysis of marriage and specialization, have accounted for this problem; however, most analyses have not. For instance, earnings and family incomes have been analyzed as outcome variables in numerous analyses but included as independent, explanatory variables in many others. Researchers need to take the endogeneity of these variables into account if they are going to test theories of *how* the impacts of family structure might be transmitted. Even for the more limited goal of examining the sensitivity of the association between family structure and well-being, there is still a concern because the use of endogenous controls may contaminate other parts of the specification.

As Cherlin (1999) has pointed out, we also need to be careful about how we interpret the results from these analyses. Researchers have tended to focus on the direct association between family structure and well-being and overlook the total association, which includes both the direct and indirect associations. It is possible that the total effect of marriage could be large even if the direct effect is small. Estimates of the direct effect are relevant if we consider policies that change the number or quality of marriages while keeping the other mediating factors constant (e.g., increasing marriage but taxing families' incomes so that their economic resources remain the same). However, it hardly seems reasonable to expect that policies would do this. In the more realistic case in which policies did not attempt to "undo" some of the effects of marriage, estimates of the total impact would be more appropriate and useful.

Appendix A. Overview of Statistical Techniques for More Complicated Models

More complicated descriptions of family structure

To discuss the issues that arise when multiple descriptors for family structure are used, the Appendix keeps the notation from the longitudinal models (as these can be used to describe the cross-section models as well) but respecifies the estimating equation so that it depends on a vector of family structure characteristics, $\mathbf{M}_i(t)$, instead of a single indicator. Thus, the empirical model for well-being outcomes becomes

$$Y_i(t) = A'\mathbf{M}_i(t) + B'\mathbf{X}_i(t) + \varepsilon_i(t) \quad (\text{A.1})$$

where A and B are vectors of coefficients to estimate. Each of the coefficients in A represents the association between a particular family structure characteristic and well-being holding the remaining family structure characteristics and observed variables constant. To provide a concrete example, suppose that $\mathbf{M}_i(t)$ includes two variables. As in the models considered in Section III, let the first variable be an indicator for whether someone is currently married. Let the second be a mutually exclusive indicator for whether the person is formerly married. Although the indicator for current marriage is defined the same way as in the earlier models, the addition of the second indicator changes the interpretation of the coefficient. Recall that when multiple, mutually exclusive indicators are used, each indicator is contrasted with the remaining, omitted category. In this case, the omitted category is never having been married. Thus, the coefficient on current marriage in this specification captures the difference in well-being between people who are married and never married instead of the more general difference between people who are married and unmarried. Similarly, the coefficient on the indicator for being formerly married captures the difference between that status and being never married.

Besides the shift in the interpretation, it is also important to note that if the binary indicator $M_i(t)$ is an element of $\mathbf{M}_i(t)$, then equation (3) is a special (restricted) version of equation (A.1). Thus, it is possible to compare the two equations directly and determine which provides a better statistical explanation for well-being. To the extent that equation (A.1) accounts for more family structure types and different effects, it should reduce problems associated with model misspecification, measurement error and omitted variables. While these are important potential improvements, other types of specification and omitted variables problems could remain.

In principle, all of the estimation methods that were reviewed in Section III can be extended to accommodate equation (A.1). In practice, however, some extensions are more difficult to implement than others. The complications associated with different estimation methods are discussed below.

Instrumental variables. As with the instrumental variable procedure for a single marital status indicator, the difficulties associated with multiple variables are practical, rather than computational. The instrumental variable technique uses predictions of each of the endogenous variables and their transformations. For example, a model that included variables for marital status and marital duration would need predictions for each term. Similarly, a model that

included marital duration and its square (a transformation that would allow the impact of marital duration to increase or decrease over time) would also need separate predictions. The predictions, in turn, need to contain some variation that is independent of the other explanatory variables as well as the other predictions. This usually means that additional instruments are needed for each prediction. Given the difficulties that researchers experience locating *any* suitable instruments for marital status, the prospects of finding multiple instruments seem bleak.

Matching methods. Imbens (2000) and Lechner (2001) have recently extended propensity score matching methods to examine differences in outcomes across multiple groups. One immediate complication in these methods involves estimating the propensity score for membership in the different groups. Lechner recommends following one of two approaches: either estimating a single multinomial choice model, such as a multinomial probit model, for all the comparison types or estimating a series of conditional bivariate logit models for each of the possible contrasts. The first approach is computationally difficult when there are large numbers of choices. Also, the multinomial models are very sensitive to specification errors—misspecification with respect to any pair of choices can contaminate the estimates for all choices. Moreover, it can be difficult to identify parameters in unrestricted models. The second approach presents less of a technical challenge and is more robust; however, estimating separate conditional models for large numbers of comparison groups is cumbersome. Another complication involves finding matches for all of the comparison groups. Recall that matching methods break down if observations with certain sets of measured characteristics are found in one comparison group but not all. The chances of this problem occurring increase as the groups are divided into finer and finer categories. A final complication is that causal interpretations of the differences in outcomes across matched groups require the researcher to account for the distribution of the propensity score in forming averages (Imbens 2000).

Selectivity controls. Easily implemented two-step procedures are available for models with two binary sources of selectivity (Tunali 1986) or a single multinomial source of selectivity (Lee 1983). The first type of model would be useful in an analysis that needed to account for selectivity from both marriage and parenthood. The second type of model might be applied in situations that involve multiple family structure categories such as currently, formerly and never married adults. Models that account for more sources of selectivity are substantially more complicated and difficult to compute. Maximum likelihood procedures involving joint estimation of the outcome and selectivity equations are tractable for models with up to three sources of selectivity. A drawback, however, is that these procedures are not standardized and must be developed or modified for each new application (see, e.g., Ribar 1992). If there are more than three sources of selectivity, regular maximum likelihood methods become infeasible. For these problems, researchers either adopt alternative estimation techniques, such as simulated maximum likelihood (Borsch-Supan and Hajivassiliou 1993), or place restrictions on the distribution of the unobserved variables.

Random effects and latent unobserved variables. One type of restriction that has been employed in maximum likelihood models with multiple sources of selectivity involves specifying the correlations among the unobserved variables to follow a factor-analytic structure. Longitudinal models with correlated random effects fall into this category. The same thing can also be accomplished by decomposing the error terms such that

$$\varepsilon_{ji} = \pi_{1j}\mu_{1i} + \pi_{2j}\mu_{2i} + \dots + \pi_{Kj}\mu_{Ki} + v_{ji} \quad (\text{A.2})$$

where ε_{ji} is the error term in the j th outcome or selectivity equation, v_{ji} is an equation-specific transitory term, $\mu_{1i}, \mu_{2i}, \dots, \mu_{Ki}$ are a series of independent, unobserved factors, and $\pi_{1j}, \pi_{2j}, \dots, \pi_{Kj}$ are equation-specific coefficients or weights on the factors. In this specification, the errors in each equation depend on linear combinations of unobserved factors that are common across the other equations. The weights and common factors lead to correlations among the errors. Correlated random effects and more general latent factor models can be evaluated using the methods developed by Butler and Moffitt (1982), Heckman and Singer (1984), and Mroz (1999). Implementation of these techniques is conceptually straightforward, though as with other maximum likelihood methods the researcher typically has to develop new software. A more serious shortcoming is that each additional factor increases the computation time by several orders of magnitude. Accordingly, researchers rarely incorporate more than three or four factors.

Nonlinear models

Binary outcome and ordered categorical models. Dependent variables that are characterized by binary, “yes/no” outcomes and by ordered categorical outcomes *can* be modeled using linear regression methods. The chief difficulties with using standard regression methods to model these outcomes are that the error terms in the estimating equations are heteroskedastic and the models might predict values outside the original range of the variables. Heteroskedasticity leads to incorrect standard errors and, thus, incorrect inferences; it also leads to inefficient estimates. While these are problems, techniques are available in most statistical packages to correct the standard errors. The second difficulty involving prediction may or may not be a concern depending on what the analyst needs the model results for.

An alternative approach, which addresses the heteroskedasticity and prediction problems, is to estimate a maximum likelihood logit or probit model (or ordered logit or ordered probit model) for these outcomes. Like other maximum likelihood estimators, these models require the researcher to specify the distribution of the error term—an extreme value distribution for the logit model or a normal distribution for the probit model. Unfortunately, the estimates may be biased if the distribution is not correctly specified. The methods are sensitive to other problems as well. For instance, heteroskedasticity, which only affects the standard errors and efficiency of the estimates in OLS, can lead to bias in the logit and probit models. Despite these shortcomings, logit and probit models are widely used.

Two challenges arise when these models include an endogenous variable. The first involves possible logical inconsistencies in the models. Unlike models involving continuous outcomes, a pair of equations with binary dependent variables cannot be specified to be fully simultaneous in the outcome measures. For example, a researcher examining the relationship between marriage and employment cannot jointly estimate equations in which marriage depends on employment outcomes while employment depends on marriage outcomes. Recursive models can be estimated (e.g., a model in which marriage depends on employment but employment does not depend on marriage), as can models that are simultaneous in terms of their underlying latent determinants (e.g., a model in which the net benefits of marriage depend on the net benefits of

employment and vice versa). However, models that are fully simultaneous in the outcome measures cannot be estimated (Maddala 1983).

The other challenge is that logit and probit models are nonlinear and, therefore, more difficult to estimate than linear regression models. Methods for addressing selectivity in probit and logit models are discussed briefly below.

- (a) Instrumental variables. A two-step technique can be applied in which predicted values are substituted for endogenous explanatory variables; the main complication with this technique involves computing the correct standard errors (Murphy and Topel 1985). Other techniques such as non-linear instrumental variables estimators can also be applied (Greene 2003).
- (b) Matching methods. These methods can be applied directly.
- (c) Selectivity controls. Two-step procedures are available; however, they are relatively complicated and rarely used. For single sources of selection, a maximum likelihood procedure (conditional, bivariate probit) is commonly employed; this procedure is available in Stata and some other software packages. For multiple sources of selection, customized maximum likelihood routines are required.
- (d) Nonparametric bounds. This method actually becomes easier when binary outcomes are examined because the minimum and maximum values for the outcomes are known to be zero and one, respectively.
- (e) Fixed effects. A conditional logit technique is available to account for fixed effects in short panels (Chamberlain 1980). For longer panels and probit specifications, dummy variables have to be included for each individual; this greatly increases the computational time for models and leads to the incidental parameters problem (see the discussion in Maddala 1987).
- (f) Random effects. It is straightforward to extend the random effects approach to probit models; the main complication is the added computational burden.

Hazard and survival models. Hazard and survival models are used to examine how long someone stays in one state before transitioning to another state.²⁶ The models address censoring problems that arise in spell data (e.g., people reaching the end of a survey before making a transition). They do this by modeling the probability that someone will make a transition at a particular point in time *conditional* on them not having made the transition up until that point in time. Many hazard models impose distributional assumptions on the spell lengths and are estimated using maximum likelihood.²⁷ Like other maximum likelihood estimators, these models are very sensitive to specification errors.

Random variation arising from any source including omitted variables, imprecise measurement of spell lengths, or other errors presents one such specification issue. Maximum likelihood models fit the realized distribution of spells to the theorized distribution from the hazard function. If there are additional unobserved determinants of spell lengths, the variation from these determinants is confounded with the variation from the underlying hazard function

²⁶ Lancaster (1990) provides an extensive review of hazard and survival models.

²⁷ For instance, an exponential hazard model assumes that the spell length, s , follows an exponential density function $f(s) = \kappa \exp(-\kappa s)$. This model implies that the probability of making a transition at time s conditional on not having made a transition before s is constant ($= \kappa$) at every point in time. A generalization of this model is the Weibull hazard model, which assumes that the spell lengths follow a density, $f(s) = \kappa p (\kappa s)^{p-1} \exp[-(\kappa s)^p]$. This density allows the conditional probability of making a transition to increase or decrease with the length of the spell, depending on whether the parameter p is greater than or less than one. There are many other hazard specifications.

resulting in a new distribution. Unless the model accounts for both sources of variation, it ends up fitting the wrong distribution to the data. This can lead to biased estimates even if the unobserved determinants are independent of the observed explanatory variables in the model.

One way to address this problem is to model the source of unobserved heterogeneity and adjust the likelihood function accordingly. For instance, the unobserved determinants could be specified as a latent unobserved variable (i.e., as a random effect), and the resulting model could be estimated using the procedures developed by Butler and Moffitt (1982) or Heckman and Singer (1984). Another approach due to Cox (1972) is to assume that changes in the explanatory variables shift the entire realized distribution of spells proportionally up or down and then estimate these shifts conditional on an arbitrary (baseline) distribution of spells. The Cox procedure is useful in cases where researchers need to know how the explanatory variables shift the spell distribution but do not need to distinguish between the other sources of variation that contribute to that distribution.

There are relatively few methods for addressing selectivity and non-independent omitted variables in hazard models. Olsen and Farkas (1989) developed an instrumental variable approach for the special case where spell lengths could be estimated using regression methods. There is also a version of the Cox procedure that conditions on different baseline hazards for different groups. This “stratified” Cox procedure (available in Stata) accounts for omitted variables that are common across groups, like cohort- or community-specific effects, but cannot be applied to address person-specific effects (unless the data contain many spells for each person). Maximum likelihood models that incorporate correlated random effects and latent unobserved factors (Lillard 1993) have also proven useful. These models specify joint sets of primary estimating equations for the outcomes of interest and secondary equations for the endogenous variables. All of the estimating equations include a common set of unobserved latent variables, which in turn allow for correlations between the unobserved determinants for each outcome. The aML software package contains procedures to estimate these models.

Structural equations, path, and covariance structure models. Structural equations and path models are used in situations where there are multiple indicators for an outcome variable, the explanatory variables, or all of these.

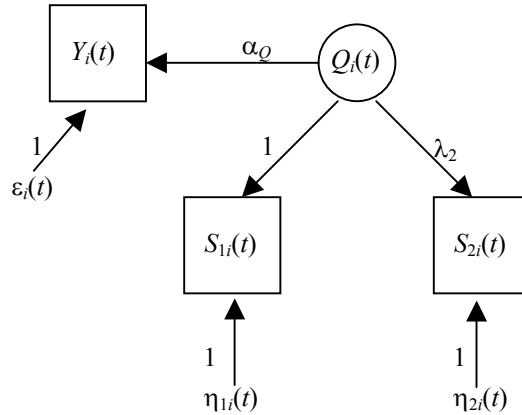
As an illustration, suppose that we wish to examine the relationship between well-being, $Y_i(t)$, and marriage quality, $Q_i(t)$. For simplicity, assume that there are no other observed explanatory variables, so that we could rewrite equation (A.1) as

$$Y_i(t) = \alpha_Q Q_i(t) + \varepsilon_i(t) \tag{A.3}$$

where α_Q is a coefficient to estimate and $\varepsilon_i(t)$ is an error term. If $Q_i(t)$ were directly observed, equation (A.3) could be estimated using standard regression methods. Suppose, however, that $Q_i(t)$ is not observed. Instead, we observe two indicators, $S_{1i}(t)$ and $S_{2i}(t)$, that are related to $Q_i(t)$ as follows

$$S_{1i}(t) = Q_i(t) + \eta_{1i}(t) \quad \text{and} \quad S_{2i}(t) = \lambda_2 Q_i(t) + \eta_{2i}(t) \tag{A.4}$$

where λ_2 is a coefficient and $\eta_{1i}(t)$ and $\eta_{2i}(t)$ are random terms, which are distributed independently of one another and of $\varepsilon_i(t)$. The indicators could be answers to two survey questions, such as “Are you happy with your marriage?” and “Do you quarrel a lot with your spouse?” or two other variables related to marital quality. In this specification, equation (A.3) represents the structural model, and (A.4) describes a measurement model. An alternative representation of these relationships is given by the following path diagram.



In the diagram, the terms in the square boxes represent variables that are observed; arrows represent directions of effects, and labels on the arrows indicate magnitudes of effects. The diagram shows that quality affects well-being and people’s reports of $S_{1i}(t)$ and $S_{2i}(t)$. Well-being also has a random component ($\varepsilon_i(t)$) as do the two, indirect quality measures.

Although this particular model is fairly simple, there are challenges to estimating the coefficients. Consider just the coefficient, α_Q . A simple regression of $Y_i(t)$ on $S_{1i}(t)$ will give a biased estimate because $S_{1i}(t)$ is an imperfect measure of $Q_i(t)$. A regression of $Y_i(t)$ on $S_{2i}(t)$ is even more problematic because of the coefficient λ_2 in the measurement model. Moreover, even if an unbiased estimate could be obtained from one of these regressions, there would still be concerns about the information that would be lost from not using both indicators for $Q_i(t)$. This suggests a third approach, which is to regress well-being on both of the indirect measures. The shortcoming with this approach is that it would be difficult to interpret the results, as neither regression coefficient would indicate the direct effect of quality, α_Q . Even if the problems of interpretation were put aside, such a regression would still suffer from measurement error in the individual $S_{1i}(t)$ and $S_{2i}(t)$ measures as well as multicollinearity between the measures.

Structural equations methods address these problems. The methods usually adopt one of three general estimation strategies, reformulating the model as either a maximum likelihood, generalized method of moments or minimum distance specification. The methods can be applied to the example above as well as to substantially more complicated models involving several latent variables and numerous indicators for each variable (Bollen 1989).

References

- Amato, Paul. "Children's Adjustment to Divorce: Theories, Hypotheses, and Empirical Support." *Journal of Marriage and the Family* 55 (February 1993): 23-38.
- Amato, Paul R. "The Consequences of Divorce for Adults and Children." *Journal of Marriage and Family* 62 (November 2000): 1269–1287.
- Amato, Paul R., and Alan Booth. "A Prospective Study of Divorce and Parent-Child Relationships." *Journal of Marriage and the Family* 58 (May 1996): 356-365.
- Amato, Paul R., Laura Spencer Loomis, and Alan Booth. "Parental Divorce, Marital Conflict, and Offspring Well-being during Early Adulthood." *Social Forces* 73 (March 1995): 895-915.
- Anson, Ofra. "Marital Status and Women's Health Revisited: The Importance of a Proximate Adult." *Journal of Marriage and the Family* 51 (February 1989): 185-194.
- Astone, Nan Marie, and Sara S. McLanahan. "Family Structure, Parental Practices and High School Completion." *American Sociological Review* 56 (June 1991): 309-320.
- Astone, Nan Marie, and Sara S. McLanahan. "Family Structure, Residential Mobility, and School Dropout: A Research Note." *Demography* 31 (November 1994): 575-584.
- Bauman, Kurt J. "Shifting Family Definitions: The Effect of Cohabitation and Other Nonfamily Household Relationships on Measures of Poverty." *Demography* 36 (August 1999): 315-325.
- Becker, Gary S. "A Theory of the Allocation of Time." *Economic Journal* 75 (September 1965): 493-517.
- Becker, Gary S. *A Treatise on the Family*. Cambridge, MA: Harvard University Press, 1981.
- Becker, Gary S. *Human Capital: A Theoretical and Empirical Analysis with Special Reference to Education*, 3rd edition. Chicago: University of Chicago Press, 1993.
- Becker, Gary S., Elisabeth M. Landes and Robert T. Michael. "An Economic Analysis of Marital Instability." *Journal of Political Economy* 85 (December 1977): 1141-1188.
- Bellas, Marcia. "The Effects of Marital Status and Wives' Employment on the Salaries of Faculty Men: The (House) Wife Bonus." *Gender and Society* 6 (December 1992): 609-622.
- Berardo, Felix. "Widowhood Status in the United States: Perspective on a Neglected Aspect of the Family Life-Cycle." *Family Coordinator* 17 (July 1968): 191-203.

- Bergstrom, Theodore C. "Economics in a Family Way." *Journal of Economic Literature* 34 (December 1996): 1903-1934.
- Bergstrom, Theodore C., and Robert F. Schoeni. "Income prospects and age-at-marriage." *Journal of Population Economics* 9 (1996): 115-130.
- Berkman, Paul. "Spouseless Motherhood, Psychological Stress, and Physical Morbidity." *Journal of Health and Social Behavior* 10 (December 1969): 323-334.
- Biblarz, Timothy J., and Adrian E. Raftery. "Family Structure, Educational Attainment, and Socioeconomic Success: Rethinking the 'Pathology of Matriarchy.'" *American Journal of Sociology* 105 (September 1999): 321-365.
- Blackburn, McKinley, and Sanders Korenman. "The Declining Marital-Status Earnings Differential." *Journal of Population Economics* 7 (July 1994): 247-70.
- Bollen, Kenneth. *Structural Equations with Latent Variables*. New York: John Wiley & Sons, Inc., 1989.
- Booth, Alan, and David Johnson. "Declining Health and Marital Quality." *Journal of Marriage and the Family* 56 (February 1994): 218-223.
- Borsch-Supan, Axel, and Vassilis Hajivassiliou. "Smooth Unbiased Multivariate Probability Simulators for Maximum Likelihood Estimation of Limited Dependent Variable Models." *Journal of Econometrics* 58 (August 1993): 347-368.
- Budig, Michelle and Paula England. "The Wage Penalty for Motherhood." *American Sociological Review* 66 (April 2001): 204-225.
- Butler, J.S., and Robert Moffitt. "A Computationally Efficient Quadrature Procedure for the One-Factor Multinomial Probit Model." *Econometrica* 50 (May 1982): 761-764.
- Cain, Glen. "The Economic Analysis of Labor Market Discrimination: A Survey," in *Handbook of Labor Economics*, volume 1. O. Ashenfelter and R. Layard (eds.). Amsterdam: Elsevier Science, 1986.
- Card, David. "The Causal Effect of Education on Earnings," in *Handbook of Labor Economics* volume 3A. O. Ashenfelter and D. Card (eds.). Amsterdam: Elsevier Science, 1999.
- Chamberlain, Gary. "Analysis of Covariance with Qualitative Data." *Review of Economic Studies* 47 (January 1980): 225-238.
- Chase-Lansdale, P. Lindsay. "How Developmental Psychologists Think About Family Process and Child Development in Low Income Families," mimeo, University of Chicago (May 1998).
- Cherlin, Andrew. "Going to Extremes: Family Structure, Children's Well-Being, and Social Science." *Demography* 36 (November 1999): 421-428.

- Cherlin, Andrew J., Frank F. Furstenberg, Jr., P. Lindsay Chase-Linsdale, Kathleen E. Kiernan, Philip K. Robins, Donna Ruane Morrison, and Julien O. Teitler. "Longitudinal Studies of Effects of Divorce on Children in Great Britain and the United States." *Science* 252 (June 7, 1991): 1386-1389.
- Cherlin, Andrew J., Kathleen E. Kiernan, and P. Lindsay Chase-Lansdale. "Parental Divorce in Childhood and Demographic Outcomes in Young Adulthood." *Demography* 32 (August 1995): 299-318.
- Chiodo, Abbigail, and Michael Owyang. "Marriage, Motherhood and Money: How Do Women's Life Decisions Influence Their Wages?" *The Regional Economist* (April 2003): 5-9.
- Chiodo, Abbigail, and Michael Owyang. "For Love or Money: Why Married Men Make More." *The Regional Economist* (April 2002): 10-11.
- Chun, Hyunbae and Injae Lee. "Why Do Married Men Earn More: Productivity or Marriage Selection?" *Economic Inquiry* 39 (April 2001): 307-319.
- Cohen, Philip, N. "Cohabitation and the Declining Marriage Premium for Men." *Work and Occupation* 29 (August 2002): 346-363.
- Coleman, Marilyn, Lawrence Ganong, and Mark Fine. "Reinvestigating Remarriage: Another Decade of Progress." *Journal of Marriage and Family* 62 (November 2000): 1288-1307.
- Corak, Miles. "Death and Divorce: The Long-Term Consequences of Parental Loss on Adolescents." *Journal of Labor Economics* 19 (July 2001): 682-715.
- Corman, Hope, and Robert Kaestner. "The Effects of Child Health on Marital Status and Family Structure." *Demography* 29 (August 1992): 389-408.
- Cornwell, Christopher and Peter Rupert. "Unobservable Individual Effects, Marriage and the Earnings of Young Men." *Economic Inquiry* 35 (April 1997): 285-294.
- Coyne, James, Michael Rohrbaugh, Varda Shohan, John Sonnega, John Nicklas, and James Cranford. "Prognostic Importance of Marital Quality for Survival of Congestive Heart Failure." *American Journal of Cardiology* 88 (1 September 2001): 526-529.
- Cox, D. R. "Regression Models and Life-Tables" (with discussion). *Journal of the Royal Statistical Society, Series B* 34 (1972): 187-220.
- Daniel, Kermit. "The Marriage Premium," In *The New Economics of Human Behavior*, Mariano Tommasi and Kathryn Ierulli (eds.). New York: Cambridge University Press, 1995.
- Dawson, Deborah A. "Family Structure and Children's Health and Well-Being: Data from the 1988 National Health Interview Survey on Child Health." *Journal of Marriage and the Family* 53 (August 1991): 573-584.

- Dee, Thomas S. "Until Death Do You Part: The Effects of Unilateral Divorce on Spousal Homicides." *Economic Inquiry* 41 (January 2003): 163-182.
- Deleire, Thomas, and Ariel Kalil. "How Do Cohabiting Couples with Children Spend Their Money?" Unpublished manuscript, Chicago: University of Chicago, April 2002a.
- Deleire, Thomas, and Ariel Kalil. "Good Things Come in Threes: Single-Parent Multigenerational Family Structure and Adolescent Adjustment" *Demography* 39 (May 2002b): 393-413.
- Edin, Kathryn. "What Do Low-Income Single Mothers Say about Marriage." *Social Problems* 47 (February 2000): 112-133.
- Ehrle, Jennifer, Katherine Kortenkamp, and Matthew Stagner. "Family Structure and Outcomes fo Children of Long-Term Welfare Recipients." Unpublished manuscript, Washington, DC: Urban Institute, 2003.
- Ermisch, John F., and Marco Francesconi. "Family Structure and Children's Achievements." *Journal of Population Economics* 14 (May 2001): 249-270.
- Freedman, David. "Statistical Models and Shoe Leather." *Sociological Methodology* 21 (1991): 291-313.
- Fronstin, Paul, David H. Greenberg, and Philip K. Robins. "Parental Disruption and the Labour Market Performance of Children When They Reach Adulthood." *Journal of Population Economics* 14 (February 2001): 137-72.
- Fu, Haishan, and Noreen Goldman. "Incorporating Health into Models of Marriage Choice: Demographic and Sociological Perspectives." *Journal of Marriage and the Family* 58 (August 1996): 740-758.
- Gardiner, Karen N, Michael E. Fishman, Plamen Nikolov and Stephanie Laud. *State Policies to Support Marriage: Preliminary Report*, U.S. Department of Health and Human Services, Assistant Secretary for Planning and Evaluation, Contract No. HHS-100-99-0008. Washington, 2002.
- Gennetian, Lisa A., Johannes M. Bos and Pamela A. Morris. "Using Instrumental Variables Analysis to Learn More from Social Policy Experiments," mimeo, Manpower Demonstration Research Corporation (October 2002).
- Ginther, Donna and Robert Pollak. "Does Family Structure Affect Children's Educational Outcomes?" Cambridge, MA: National Bureau of Economic Research Working Paper No. w9628, April 2003.
- Ginther, Donna and Madeline Zavodny. "Is the Male Marriage Premium Due to Selection? The Effect of Shotgun Weddings on the Return to Marriage." *Journal of Population Economics* 14 (May 2001): 313-328.

- Glenn, Norval. "Quantitative Research on Marital Quality in the 1980s: A Critical Review." *Journal of Marriage and the Family* 52 (November 1990): 818-831.
- Glenn, Norval, et al. *Why Marriage Matters: Twenty-One Conclusions from the Social Sciences*. New York: Center of the American Experiment, Coalition of Marriage, Family and Couples Education, 2002.
- Godwin, Deborah. "Causal Modeling in Family Research." *Journal of Marriage and the Family* 50 (November 1988): 917-927.
- Goldin, Claudia. *Understanding the Gender Gap: An Economic History of American Women*. New York: Oxford University Press, 1990.
- Goldman, Noreen. "Marriage Selection and Mortality Patterns: Inferences and Fallacies." *Demography* 30 (May 1993): 189-208.
- Goldman, Noreen, Sanders Korenman, and Rachel Weinstein. "Marital status and health among the elderly." *Social Science and Medicine* 40 (June 1995): 1717-1730.
- Gove, Walter. "Sex, Marital Status, and Mortality." *American Journal of Sociology* 79 (July 1973): 45-67.
- Gray, Jeffrey S. "The Fall in Men's Return to Marriage: Declining Productivity Effects Or Changing Selection?" *Journal of Human Resources* 32 (Summer 1997): 481-504.
- Greene, William. *Econometric Analysis* 5th edition. Upper Saddle River, NJ: Prentice Hall, 2003.
- Griliches, Zvi. "Sibling Models and Data in Economics: Beginnings of a Survey." *Journal of Political Economy* 87 (October 1979, part 2): S37-S64.
- Grossman, Michael. "On the Concept of Health Capital and the Demand for Health." *Journal of Political Economy* 80 (March/April 1972): 223-255.
- Gruber, Jonathan. "Is Making Divorce Easier Bad for Children? The Long Run Implications of Unilateral Divorce." Cambridge, MA: National Bureau of Economic Research Working Paper No. w7968, October 2000.
- Hahn, Beth. "Marital Status and Women's Health: The Effect of Economic Marital Acquisitions." *Journal of Marriage and the Family* 55 (May 1993): 495-504.
- Hanson, Thomas. "Does Parental Conflict Explain Why Divorce is Negatively Associated with Child Welfare?" *Social Forces* (June 1999): 1283-1316.
- Harris, Judith Rich. *The Nurture Assumption: Why Children Turn Out the Way They Do*. New York: Simon and Schuster, Inc., 1998.
- Haveman, Robert, and Barbara Wolfe. *Succeeding Generations: On the Effects of Investments in Children*. New York: Russell Sage Foundation, 1994.

- Haveman, Robert, and Barbara Wolfe. "The Determinants of Children's Attainments: A Review of Methods and Findings." *Journal of Economic Literature* 33 (December 1995): 1829-1878.
- Hechter, Michael, and Satoshi Kanazawa. "Sociological Rational Choice Theory." *Annual Review of Sociology* 23 (1997): 191-214.
- Heckman, James J. "Dummy Endogenous Variables in a Simultaneous Equations System." *Econometrica* 46 (July 1978): 931-959.
- Heckman, James J. "Sample Selection Bias as a Specification Error." *Econometrica* 47 (January 1979): 153-161.
- Heckman, James J. "Randomization as an Instrumental Variable." *Review of Economics and Statistics* 78 (May 1996): 336-341.
- Heckman, James J., Robert J. LaLonde and Jeffrey A. Smith. "The Economics and Econometrics of Active Labor Market Programs," in *Handbook of Labor Economics* volume 3A, O. Ashenfelter and D. Card (eds.). Amsterdam: Elsevier Science, 1999.
- Heckman, James J., and Richard Robb. "Alternative methods for evaluating the impact of interventions," in *Longitudinal Analysis of Labor Market Data*, J. Heckman & B. Singer (eds.). New York: Cambridge University Press, 1985.
- Heckman, James J., and Burton Singer. "A Method for Minimizing the Impact of Distributional Assumptions in Econometric Models for Duration Data." *Econometrica* 52 (March 1984): 271-320.
- Hemstrom, Orjan. "Is Marriage Dissolution Linked to Differences in Mortality Risks for Men and Women?" *Journal of Marriage and the Family* 58 (May 1996): 366-378.
- Hersch, Joni and Leslie Stratton. "Household Specialization and the Male Marriage Wage Premium." *Industrial and Labor Relations Review* 54 (October 2000): 78-94.
- Herzog, Elizabeth, and Cecelia E. Sudia, C. "Children in Fatherless Families." In B. Caldwell and H. Ricciuti (Eds.), *Review of Child Development Research Vol. 3*. Chicago, IL: University of Chicago Press, 1973, 141-232.
- Hewitt, Belinda, Mark Western and Janeen Baxter. "Marriage and Money: The Impact of Marriage on Men's and Women's Earnings." Discussion Paper DP-007, The University of Queensland, July 2002.
- Hill, Martha. "The Wage Effects of Marital Status and Children," *Journal of Human Resources* 14 (Autumn 1979): 579-594.
- Hines, Alice M. "Divorce-Related Transitions, Adolescent Development, and the Role of the Parent-Child Relationship: A Review of the Literature." *Journal of Marriage and the Family* 59 (May 1997): 375-388.

- House, James, Karl Landis, and Debra Umberson. "Social Relationships and Health." *Science* 241 (July 29, 1988): 540-545.
- Hu, Yuanreng, and Noreen Goldman. "Mortality Differentials by Marital Status: An International Comparison." *Demography* 27 (May 1990): 233-250.
- Hummer, Robert, Richard Rogers, and Isaac Eberstein. "Sociodemographic Differentials in Adult Mortality: A Review of Analytic Approaches." *Population and Development Review* 24 (September 1998): 553-578.
- Imbens, Guido. "The Role of the Propensity Score in Estimating Dose-Response Functions." *Biometrika* 87 (September 2000): 706-710.
- Jekielek, Susan M. "Parental Conflict, Marital Disruption and Children's Emotional Well-Being." *Social Forces* 76 (March 1998): 905-936.
- Katzev, Aphra R., Rebecca L. Warner, and Alan C. Acock. "Girls or Boys? Relationship of Child Gender to Marital Instability." *Journal of Marriage and the Family* 56 (February 1994): 89-100.
- Kenny, Lawrence G. "The Accumulation of Human Capital During Marriage by Males." *Economic Inquiry* 21 (April 1983): 223-231.
- Kiecolt-Glaser, Janice K., and Tamara Newton. "Marriage and Health: His and Hers." *Psychological Bulletin* 127 (July 2001): 472-503.
- Klebba, A. Joan. *Mortality from Selected Causes by Marital Status: United States* (parts A and B), Public Health Service Publication No. 1000 – Series 20, Nos. 8a and 8b. Rockville, MD: U.S. Department of Health, Education, and Welfare, December 1970.
- Korenman, Sanders, Noreen Goldman, and Haishan Fu. "Misclassification Bias in Estimates of Bereavement Effects." *American Journal of Epidemiology* 145 (June 1997): 995-1002.
- Korenman, Sanders and David Neumark. "Marriage, Motherhood, and Wages." *Journal of Human Resources* 27 (Spring 1992): 233-255.
- Korenman, Sanders and David Neumark. "Does Marriage Really Make Men More Productive?" *Journal of Human Resources* 26 (Spring 1991): 282-307.
- Lancaster, Tony. *The Econometric Analysis of Transition Data*. New York: Cambridge University Press, 1990.
- Lang, Kevin, and Jay L. Zagorsky. "Does Growing Up with a Parent Absent Really Hurt?" *Journal of Human Resources* 36 (Spring 2001): 253-73.
- Lechner, Michael. "Program Heterogeneity and Propensity Score Matching: An Application to the Evaluation of Active Labor Market Policies." St. Gallen, Switzerland: University of St. Gallen working paper, 2001.

- Lee, Lung-Fei. "Generalized Econometric Models with Selectivity." *Econometrica* 51 (March 1983): 507-512.
- Lerman, Robert. "The Impact of the Changing US Family Structure on Child Poverty and Income Inequality." *Economica* 63 (Supplement 1996): S119-S139.
- Lerman, Robert. "How do Marriage, Cohabitation, and Single Parenthood affect the Material Hardships of Families with Children?" Washington, D.C.: Urban Institute, July 2002a.
- Lerman, Robert. "Impacts of Marital Status and Parental Presence on the Material Hardship of Families with Children," Washington, D.C.: Urban Institute, July 2002b.
- Lerman, Robert. "Marriage and the Economic Well-Being of Families with Children." Washington, DC: Urban Institute, 2002c.
- Lerman, Robert. "Married and Unmarried Parenthood and Economic Well-Being: A Dynamic Analysis of a Recent Cohort." Washington, DC: Urban Institute, 2002d.
- Lillard, Lee. "Simultaneous Equations for Hazards: Marriage Duration and Fertility Timing." *Journal of Econometrics* 56 (March 1993): 189-217.
- Lillard, Lee, and Constantijn Panis. "Marital Status and Mortality: The Role of Health." *Demography* 33 (August 1996): 313-327.
- Lillard, Lee A., and Linda J. Waite. "A Joint Model of Marital Childbearing and Marital Disruption." *Demography* 30 (November 1993): 653-681.
- Lillard, Lee, and Linda Waite. "'Til Death Do Us Part: Marital Disruption and Mortality." *American Journal of Sociology* 100 (March 1995): 1131-1156.
- Loh, Eng Seng. "Productivity Differences and the Marriage Wage Premium for White Males." *Journal of Human Resources* 31 (Summer 1996): 566-589.
- Loughran, David S., and Julie M. Zissimopoulos. "Are There Gains to Delaying Marriage? The Effect of Age at First Marriage on Career Development and Wages." [Preliminary]. Unpublished manuscript, Santa Monica, CA: RAND Corporation, April 2003.
- Luoma, Jason, and Jane Pearson. "Suicide and Marital Status in the United States, 1991-1996: Is Widowhood a Risk Factor?" *American Journal of Public Health* 92 (September 2002): 1518-1522.
- Maddala, G. S. *Limited Dependent and Qualitative Variables in Econometrics*. New York: Cambridge University Press, 1983.
- Maddala, G. S. "Limited Dependent Variable Models Using Panel Data." *Journal of Human Resources* 22 (Summer 1987): 307-338.

- Manski, Charles F. *The Identification Problem in the Social Sciences*. Cambridge, MA: Harvard University Press, 1995.
- Manski, Charles F., Gary D. Sandefur, Sara McLanahan, and Daniel Powers. "Alternative Estimates of the Effect of Family Structure During Adolescence on High School Graduation." *Journal of the American Statistical Association* 87 (March 1992): 25-37.
- Marini, Margaret M., and Burton Singer. "Causality in the Social Sciences." *Sociological Methodology* 18 (1988): 347-409.
- Mauldon, Jane. "The Effect of Marital Disruption on Children's Health." *Demography* 27 (August 1990): 431-446.
- Mauldon, Jane. "Children's Risks of Experiencing Divorce and Remarriage: Do Disabled Children Destabilize Marriages?" *Population Studies* 46 (July 1992): 349-362.
- McLanahan, Sara, and Gary Sandefur. *Growing Up with a Single Parent: What Hurts, What Helps*. Cambridge, MA: Harvard University Press, 1994.
- Mechoulan, Stéphane. "Divorce Laws & the Structure of the American Family." Unpublished manuscript, Evanston, IL: Northwestern University, October 2000.
- Michael, Yvonne, Lisa Berkman, Graham Colditz, and Ichiro Kawachi. "Living Arrangements, Social Integration, and Change in Functional Health Status." *American Journal of Epidemiology* 153 (15 January 2001): 123-131.
- Morrison, Donna Ruane, and Andrew J. Cherlin. "The Divorce Process and Young Children's Well-Being: A Prospective Analysis." *Journal of Marriage and the Family* 57 (August 1995): 800-812.
- Morrison, Donna Ruane, and Mary Jo Coiro. "Parental Conflict and Marital Disruption: Do Children Benefit When High-Conflict Marriages Are Dissolved?" *Journal of Marriage and the Family* 61 (August 1999): 626-637.
- Morrison, Donna Ruane, and Amy Ritualo. "Routes to Children's Economic Recovery after Divorce: Are Cohabitation and Remarriage Equivalent?" *American Sociological Review* 65 (August 2000): 560-580.
- Murphy, Kevin, Robert Topel. "Estimation and Inference in Two-Step Econometric Models." *Journal of Business and Economic Statistics* 3 (October 1985): 370-379.
- Murphy, Mike; Karen Glaser, and Emily Grundy. "Marital Status and Long-Term Illness in Great Britain." *Journal of Marriage and the Family* 59 (February 1997): 156-164.
- Murray, John. "Marital Protection and Marital Selection: Evidence from a Historical-Prospective Sample of American Men." *Demography* 37 (November 2000): 511-521.

- Mroz, Thomas A. "Discrete Factor Approximations in Simultaneous Equation Models: Estimating the Impact of a Dummy Endogenous Variable on a Continuous Outcomes." *Journal of Econometrics* 92 (October 1999): 233-74.
- Nakosteen, Robert and Michael Zimmer. "Marital Status and Earnings of Young Men: A Model with Endogenous Selection." *Journal of Human Resources* 22 (Spring 1987): 248-268.
- Nakosteen, Robert and Michael Zimmer. "Men, Money, and Marriage: Are High Earners More Prone than Low Earners to Marry?" *Social Science Quarterly* 78 (March 1997): 66-82.
- Ogburn, William Fielding. "Eleven Questions Concerning American Marriages." *Social Forces* 6 (September 1927): 5-12.
- Painter, Gary, and David I. Levine. "Family Structure and Youths' Outcomes: Which Correlations are Causal?" *Journal of Human Resources* 35 (Summer 2000): 524-549.
- Peters, H. Elizabeth. "Marriage and Divorce: Informational Constraints and Private Contracting." *American Economic Review* 76 (June 1986): 437-454.
- Renne, Karen. "Health and Marital Experience in an Urban Population." *Journal of Marriage and the Family* 33 (May 1971): 338-350.
- Ribar, David. "Child Care and the Labor Supply of Married Women: Reduced Form Evidence." *Journal of Human Resources* 27 (Winter 1992): 134-165.
- Richardson, Katarina. "The Evolution of the Marriage Premium in the Swedish Labor Market 1968-1991." Unpublished manuscript, Labor Market Policy Evaluation, Uppsala, Sweden, June 2000.
- Rogers, Richard. "The Effects of Family Composition, Health, and Social Support Linkages on Mortality." *Journal of Health and Social Behavior* 37 (December 1996): 326-338.
- Rosenbaum, Paul, and Donald Rubin. "The Central Role of the Propensity Score in Observational Studies for Causal Effects." *Biometrika* 70 (April 1983): 43-55.
- Ross, Catherine; John Mirowsky, and Karen Goldsteen. "The Impact of the Family on Health: The Decade in Review." *Journal of Marriage and the Family* 52 (November 1990): 1059-1078.
- Rubin, Donald. "Using Multivariate Matched Sampling and Regression Adjustment to Control Bias in Observational Studies." *Journal of the American Statistical Association* 74 (June 1979): 318-328.
- Sandefur, Gary D., and Thomas Wells. "Does Family Structure Really Influence Educational Attainment?" *Social Science Research* 28 (December 1999): 331-357.
- Seltzer, Judith A. "Consequences of Marital Dissolution for Children." *Annual Review of Sociology* 20 (1994): 235-266.

- Seltzer, Judith A. "Families Formed Outside of Marriage." *Journal of Marriage and Family* 62 (November 2000): 1247–1268.
- Schoeni, Robert F. "Marital Status and Earnings in Developed Countries." *Journal of Population Economics* 8 (November 1995): 351-59.
- Sigle-Rushton, Wendy, and Sara McLanahan. "For Richer or Poorer? Marriage as an Anti-Poverty Strategy in the United States." Princeton, NJ: Center for Research on Child Wellbeing Working Paper, June 2002a
- Sigle-Rushton, Wendy, and Sara McLanahan. "Father Absence and Child Well-being: A Critical Review." Princeton, NJ: Center for Research on Child Wellbeing Working Paper #02-20, November 2002b.
- Smith, Ken, and Norman Waitzman. "Double Jeopardy: Interaction Effects of Marital and Poverty Status on the Risk of Mortality." *Demography* 31 (August 1994): 487-507.
- Smith, Ken, and Cathleen Zick. "Linked Lives, Dependent Demise? Survival Analysis of Husbands and Wives." *Demography* 31 (February 1994): 81-93.
- Sorlie, Paul, Eric Backlund, and Jacob Keller. "US Mortality by Economic, Demographic, and Social Characteristics: The National Longitudinal Mortality Study." *American Journal of Public Health* 85 (July 1995): 949-956.
- Stratton, Leslie S. "Examining the Wage Differential for Married and Cohabiting Men." *Economic Inquiry* 40 (April 2002): 199-212.
- Thomas, Adam, and Isabel Sawhill. "For Richer or for Poorer: Marriage as an Antipoverty Strategy." *Journal of Public Policy Analysis and Management* 21 (Fall 2002): 587-599.
- Thornton, Arland (ed.). *The Well-Being of Children and Families*. Ann Arbor, MI: University of Michigan Press, 2001.
- Tower, Roni Beth, Stanislav V. Kasl, and Amy S. Darefsky. "Types of Marital Closeness and Mortality Risk in Older Couples." *Psychosomatic Medicine* 64 (July/August 2002): 644-659
- Tunali, Insan. "A General Structure for Models of Double-Selection and an Application to a Joint Migration/Earnings Process with Remigration," in *Research in Labor Economics* vol. 8 (part B), R. Ehrenberg (ed.). Greenwich, CT: JAI Press, 1986.
- Vella, Francis. "Estimating Models with Sample Selection Bias." *Journal of Human Resources* 33 (Winter 1998), 127-169.
- Umberson, Debra. "Family Status and Health Behaviors: Social Control as a Dimension of Social Integration." *Journal of Health and Social Behavior* 28 (September 1987): 306-319.

- Umberson, Debra. "Gender, marital status and the social control of health behavior." *Social Science and Medicine* 34 (April 1992): 907-917.
- Verbrugge, Lois. "Marital Status and Health." *Journal of Marriage and the Family* 41 (May 1979): 267-285.
- Waite, Linda. "Does Marriage Matter?" *Demography* 32 (November 1995): 483-507.
- Waite, Linda J., and Maggie Gallagher. *The Case for Marriage: Why People Are Happier, Healthier, and Better Off Financially*. New York: Broadway Books, 2000.
- Waite, Linda J., and Lee A. Lillard. "Children and Marital Disruption." *American Journal of Sociology* 96 (January 1991): 930-953.
- Waldfoegel, Jane. "The Effect of Children on Women's Wages." *American Sociological Review* 62 (April 1997): 209-217.
- Waldron, Ingrid, Mary Elizabeth Hughes, and Tracy Brooks. "Marriage protection and marriage selection—prospective evidence for reciprocal effects of marital status and health." *Social Science and Medicine* 43 (July 1996): 113-123.
- Weiss, Yoram, and Robert Willis. "Children as Collective Goods and Divorce Settlements." *Journal of Labor Economics* 3 (July 1985): 268-292.
- Wickrama, K.A.S.; Frederick Lorenz, Rand Conger, and Glen Elder, Jr. "Marital Quality and Physical Illness: A Latent Growth Curve Analysis." *Journal of Marriage and the Family* 59 (February 1997): 143-155.
- Wickrama, K.A.S., Frederick Lorenz, Lora Ebert Wallace, Laknath Peiris, Rand Conger, and Glen Elder, Jr. "Family Influence on Physical Health During the Middle Years: The Case of Onset of Hypertension." *Journal of Marriage and the Family* 63 (May 2001): 527-539.
- Wu, Lawrence L. "Effects of Family Instability, Income, and Income Instability on the Risk of a Premarital Birth." *American Sociological Review* 61 (June 1996): 386-406.
- Wu, Lawrence L., and Brian C. Martinson. "Family Structure and the Risk of a Premarital Birth." *American Sociological Review* 58 (April 1993): 210-32.
- Zick, Cathleen, and Ken Smith. "Marital Transitions, Poverty, and Gender Differences in Mortality." *Journal of Marriage and the Family* 53 (May 1991): 327-336.