MARITAL STATUS AND MORTALITY: THE ROLE OF HEALTH*

LEE A. LILLARD AND CONSTANTIJN W.A. PANIS

Prior literature has shown that married men live longer than unmarried men. Possible explanations are that marriage protects its incumbents or that healthier men select themselves into marriage. Protective effects, however, introduce the possibility of adverse selection: Those in poor health have an incentive to marry. In this paper we explore the role of health in explaining mortality and marriage patterns, and distinguish protective effects from two types of selection effects. We find adverse *selection on the basis of health (unhealthy men tend to (re)marry sooner) and* positive *selection on the basis of unmeasured factors that both promote good health and encourage marriage.*

It is now well established that married persons have significantly lower mortality rates than unmarried persons. The relationship has been noted in numerous studies covering more than 100 years and many countries. The result is established for both men and women, but has been observed to be greater for men.

A potential explanation for married persons' mortality advantage is that marriage improves health status, which in turn reduces mortality risks. Marriage may have a protective effect on health by reducing risky behavior and by specialization and economies of scale in nutrition and caretaking. This paper assesses the role of health: first, as an intervening variable in the observed relationship between marital status and mortality, and second, as a determinant of marriage and dissolution behavior. We consider men only.

Observers have asked, at least since the time of Farr (1858), whether marriage has a direct protective effect and thus improves health and/or reduces the risk of mortality, or whether the differentials reflect a selection of healthier individuals into marriage. This question remains a topic of current debate, along with contemplation of the mechanisms that might generate the protective or selection effects. Virtually all the literature on marital status and mortality has focused on these two possibilities. If marriage has a protective effect, however, then persons in poor health and/or at higher risk of mortality have a greater incentive to marry and gain that protection. That is, *adverse* selection into marriage is theoretically possible, rather than *positive* selection, as is usually argued. We address these issues through a joint model of the relationships between health, mortality, and marriage formation and dissolution. We test whether health is a determinant in men's marriage and divorce behavior, and attempt to find the direction of this selection effect.

In the following section we discuss the literature. In subsequent sections we outline a theoretical framework illustrating the incentive to marry if marriage has a protective effect against poor health or mortality; describe the longitudinal data from the Panel Study of Income Dynamics (PSID), which allow us to disentangle causal from selection effects; explore gross relationships in the data; develop our empirical strategy for jointly modeling marriage formation and dissolution, health status, and mortality; and present the results of estimation and test hypotheses related to selection and direct effects. We end with a summary and concluding remarks.

Mechanisms Whereby Marriage Enhances Health and Reduces the Risk of Mortality

One of the most robust findings in demographic research is that married individuals' mortality rates are lower than those of their unmarried counterparts. The marriage advantage was noted as early as Farr (1858), and in both developed and developing countries (e.g., Hu and Goldman 1990; Kisker and Goldman 1987; Livi-Bacci 1984; Rahman 1993; Zick and Smith 1991). The differential generally has been found to be larger for men than for women (Gove 1973; Hu and Goldman 1990; Lillard and Waite 1995).

Several mechanisms have been suggested to explain why marriage might have a protective effect. One such mechanism operates by reducing stress and stress-related illness and through general and familial social integration (Kobrin and Hendershot 1977; Pearlin and Johnson 1977). Another mechanism works through caregiving in times of illness or poor health. Various medical studies have shown reduced morbidity, faster recovery, and lower mortality following medical problems among married patients. Marriage also may encourage healthy behaviors and discourage risky or unhealthy ones. Umberson (1987) found that marriage and parenthood reduce the occurrence of a number of health-threatening behaviors such as problem drinking, drinking and driving, substance abuse, and other forms of risk taking. Marriage also may result in greater material well-being stemming from the economies of scale achieved by combining resources, and from the specialization of tasks common in marriage (Becker 1981). These may lead (among other possibilities) to improved nutrition and to nurturing in case of an illness.

*Lee Lillard, Director, Center for the Study of Aging, RAND, 1700 Main Street, Santa Monica, CA 90407-2138; e-mail: lee lillard@rand.org. Constantijn W.A. Panis, Associate Economist, RAND. We thank Robert Marc and seminar participants at the National Bureau of Economic Research, Tilburg University, the Tinbergen Institute, and Stockholm University for their generous comments. We gratefully acknowledge support from the National Institute on Aging, Grants R01-AG11994 and R01-AG12420.

The literature on the relationship between marital status and health yields results similar to those on marital status and mortality, although the findings for health are less robust than for mortality. Waldron, Hughes, and Brooks (1996) found that married women had better health trends than their unmarried counterparts, but this was the case only among women who were not employed. Korenman, Goldman, and Hu (1990) found no striking differences by marital status in the ability to perform activities of daily living (ADL) in the elderly population. They found, however, that among people with disabilities in 1984, the married were more likely to be disability-free two years later.

Family Background and Other Long-Run Determinants of Health

Ample literature has documented the relationship between socioeconomic status and health status (Feinstein 1993; Preston and Taubman 1994). More education, for example, is generally associated with better health (Strauss et al. 1993). Researchers recently have begun to delineate the complex pathways by which poor socioeconomic status leads to poor health. Kuh and Wadsworth (1993) analyzed longitudinal data on a cohort of 3,000 36-year-old British adults followed since birth. They found that even after controlling for current socioeconomic status and health behaviors, socioeconomic factors associated with early childhood experiences remained significant predictors of health status in early adulthood. We carry this research a step further by linking childhood and adulthood socioeconomic factors to health status and mortality over the life cycle.

The Relationship Between Health and the Risk of Mortality

Health status is a widely accepted correlate of mortality risk. Idler and Kasl (1991) found that even after controlling for numerous other measures of health and illness (such as medical conditions, ADL score, smoking, weight, and social functioning), general perceptions of health measured as excellent, good, fair, poor, or bad remained a powerful predictor of mortality among the elderly. We use this type of general health indicator in our analysis.

Self-Selection into Marriage

The literature on marital status and mortality almost universally raises the possibility of positive selection into marriage; thus the effects of marriage on health and mortality might be overstated. This argument is so pervasive that the question is usually couched in terms of which mechanism—protection or selection—generates the observed advantage of the married. The argument is straightforward: Persons with observably poor health, and those with chronic conditions or dangerous or unhealthy lifestyles (unmeasured in our data), may find it more difficult to attract a spouse than do healthy, relatively settled individuals (Carter and Glick 1976; Sheps 1961). By a similar argument, those in good health may be better able to maintain a marital relationship and thus have lower dissolution rates. The possibility of selective marriage has been raised in almost every study on this subject since Farr (1858), but has rarely been tested.

Among the exceptions are Hu and Goldman (1990). In a study of a large number of developed countries, they found that a selection process operates for single and divorced persons such that the smaller the proportion of the population in these marital statuses, the higher their death rates in comparison with the married. Similarly, Kisker and Goldman (1987) related single persons' excess mortality to the proportion single in a (sub)population. The underlying hypothesis is that populations in which the great majority of individuals eventually marry should be characterized by greater selectivity effects among those who remain single than populations in which substantial proportions of people never marry. If such selection exists, the correlation should be apparent at an age by which the majority of persons (who eventually will marry) have done so. In agreement with Livi-Bacci (1984), Kisker and Goldman found evidence of selective marriage among healthier persons, especially women. Goldman (1993) returned to this argument and derived the underlying assumptions, which turned out to be quite stringent. By assuming various marriage rates in frail and in healthy subgroups, she showed that the relationship between the proportion single and excess mortality among single individuals may be negative, positive, or even nonmonotonic; thereby she cast doubt on the validity of the earlier results.

Fu and Goldman (1994) attempted to link health status directly to marriage behavior, and found very little predictive power of current health status among young adults (ages 14 to 33). They discovered, however, that unhealthy behaviors, such as alcohol and substance abuse and criminal activities, significantly delay entry into marriage; so do physical characteristics that typically are associated with poorer past and future health statuses, such as obesity and short stature. Mastekaasa (1992) related marriage rates to psychological well-being, and found that overall life satisfaction is associated positively with the hazard of marrying.

In a study based on women but addressing the same issues as ours, Waldron, Hughes, and Brooks (forthcoming) investigated the reciprocal relationships between marital status and health. They found no significant effects for employed women. Among women who did not have a job, they found weak evidence that women in better health were more likely to marry and less likely to experience marital dissolution in the five years following the interview. The authors used lagged health and marital status as explanatory variables, without accounting for selection into those states.

Employing a simultaneous-equations framework, we directly address the issue of endogeneity of marital status with respect to health. We incorporate the effects of marital status on health, of health on marital transitions (marriage formation and dissolution), and of both health and marital status on mortality.

BEHAVIORAL PERSPECTIVE

According to the basic paradigm explaining marriage behavior, decisions to enter or leave a marriage are based on a

comparison of utility in and outside the marriage union. Variations in the timing of marriage are governed by differentials in the time when the utility expected from marriage exceeds that of being single (Becker, Landes, and Michael 1977). These differentials may be affected by difficulties in assortative mating or in the partner search process (Lillard, Brien, and Waite 1995; Mortensen 1988).

Marriage may affect utility for many reasons including love, nurture and companionship, and the desire to create a home for (future) children. In this paper we focus on health aspects of marriage. First, health may affect the timing of marriage because healthy, attractive individuals may find themselves in high demand; thus health reduces the search costs. Second, if marriage protects an individual from poor health and high mortality risks, individuals in poor health (who stand to gain a great deal from marriage) have an incentive to actively seek a partner. These two arguments correspond to the two sides of the marriage market (demand and supply), and operate in opposite directions. The net effect is an empirical issue, which we address here. Unfortunately, our model is not a complete model of the marriage market and cannot distinguish supply from demand effects.

Decisions to enter and leave a marriage are based on a comparison of utility in and outside marriage. We extend this notion to account for future utility, and borrow from the literature on the life-cycle model of consumption with mortality risk (Hurd 1990; Lillard and Weiss forthcoming; Yaari 1965) and on health production and demand for health inputs (Grossman 1972). The resulting behavioral framework yields insight into the incentives to marry which the health and mortality benefits of marriage create.

The basic premise of the life-cycle model with uncertain mortality risk is that individuals attempt to maximize the expected present value of all future utility flows. In its standard formulation, utility at a given time depends on consumption expenditures at that time. In addition, the individual may derive utility from the knowledge that any bequeathable wealth remaining at the time of death goes to chosen beneficiaries, such as children born in a marriage. We argue that future utility is a function not only of consumption of goods that may be purchased directly, but also of the individual's health status and marital status. Individuals judge the expected value of future utility flows by taking their mortality risks into account; these risks also may depend on their current and future health and marital status. In each period, the individual maximizes the expected present value of all future utility flows by deciding on the optimal levels of consumption expenditures and expenditures on health care, and on the optimal marital status. These decisions are made within a monetary budget constraint and the constraints of the marriage market. Accordingly, one focus of this paper is the decision to enter or leave a marriage.

Marital status enters this framework in at least four ways. First, it affects utility directly through love and companionship. At the time of the wedding, both spouses are likely to have high expectations regarding the future direct utility benefits of marriage. Whether these expectations are realized depends (for example) on spousal compatibility. Thus the direct effect of marital status on utility is person- or couple-specific, and may change over time. Second, marriage may benefit one's health status, which in turn increases utility. Marriage thus may be viewed as an input to the production of health (Grossman 1972). Third, marriage may have a protective effect on the risk of dying, thereby increasing the expected value of future utility. Fourth, marriage may relax the budget constraint by increasing future income, possibly because marriage permits greater labor specialization (Becker 1981) or because it generates survivor benefits in Social Security and some pension plans. The second, third, and fourth mechanisms operate in favor of being married; the decision to enter or leave a marriage at any point in time will be based on the net result of these mechanisms, including (updated) expectations regarding utility from love and companionship.

To summarize, the literature suggests that marriage may improve health, reduce the risk of dying, and increase future income. All three effects increase future expected utility, providing an incentive to seek marriage to those individuals who are likely to gain the most from marriage. This argument is very similar to that made in the literature on health production and demand for health inputs (Grossman 1972): Marriage may be viewed as a health input, and is likely to be sought disproportionately by individuals who need the protection. In other words, economic incentives exist for *adverse* selection into marriage. Our empirical model explicitly accounts for this potential endogeneity of health status in marriage transition behavior, and of marital status in the health and mortality processes.

DATA

Our analysis is based on the Panel Study of Income Dynamics (PSID), a large national longitudinal data set that began in 1968 with approximately 5,500 households. The sample has been resurveyed each year since that time.[1] We use the panel through the 1990 survey; thus the respondents' longitudinal history spans as many as 23 survey waves. Each year the PSID collects detailed information on socioeconomic and demographic variables; in addition, special retrospective marital and fertility histories of both the head and the spouse were obtained in 1985 and have been updated annually since then.

Since 1984 the PSID has collected health data from household heads and their spouses. To be included in our sample, a respondent must have survived and remained in the panel through 1984. The resulting sample consists of 4,092 men. For each respondent, we have up to seven years (1984–1990) of general health measures. The question that was asked each year is, "Would you say your health is excel-

1. Each member of an original sample household, or any child born to a sample member, is followed when he or she exits a sample household. New households formed by sample members are also followed as sample units and are interviewed in the same way as original households. Thus the sample has grown over time with individuals' splitting off from the 1968 sample households.

TABLE 1. MEAN HEALTH STATUS (BY MARITAL STATUS)

Age	Single	Married
10–19	3.99	4.10
20–29	4.07	4.13
30–39	3.77	4.07
40–49	3.72	3.87
50–59	2.98	3.51
60–69	2.85	3.17
70–79	2.60	2.84
80–89	2.85	2.55
90+	2.27	1.90
Total	3.69	3.68

lent, very good, good, fair, or poor?" We recoded the ordinal responses on a scale from 1 to 5 such that better health corresponds to a higher score. Table 1 shows the mean health status report by age category and marital status. In this table the reported health measure is treated as if it represents a cardinal measure.[2]

Overall, mean reported health declines with age. Within age categories, married individuals tend to be in better health than their single (never-married, separated, divorced, widowed) counterparts. Unmarried and married males report approximately the same average health status because of relatively large numbers of unmarried males at low ages, when average health is very good.

Even though we use only 1984–1990 health reports and mortality history, we employ full retrospective marriage histories. For 96.0% of our sample, a retrospective survey was collected in 1985 and updated annually. For the other 4.0%, we constructed marriage histories using 1968–1990 panel information, including more highly detailed questions in 1968 (for the male head) and in 1976 (for the wife).

Because respondents must have been alive in 1984 to be included in the sample, we restrict the mortality analysis to 1984–1990.[3] Over this period, 279 males (6.8%) died.[4] At the time of death, 30 had never married, 182 were married, nine were separated, 20 were divorced, and 38 were widowed.

Multiple transitions in and out of marriage, unwelcome as they may be to the persons involved, aid in the econometric identification of heterogeneity due to unmeasured factors. We observe up to five marriages over the lifetime of any one male respondent; 721 men (17.4%) were married at least twice. Between 1984 and 1990, 645 men entered marriage, 370 separated, 323 converted a separation into a formal divorce, and 73 were widowed. At any point during that period, 3,001 men were married, 481 were in the separated state, 545 were in the divorced state, and 153 were widowers.

We converted the data on marital status into event-history format—that is, into spells from a young age (age 12) to first marriage, from wedding to separation date, from separation to divorce, and from divorce to remarriage. The spells may be censored because of widowhood, death, or attrition. We use them to model marital status as outcomes and as time-varying explanatory variables in models for health and mortality.

GROSS RELATIONSHIPS IN THE DATA

We begin by exploring the gross relationships between marital status, health, and mortality in the data. Then motivated by the potential endogeneity of marital status in health and mortality, we develop a simultaneous model of marital transitions, health, and mortality.

Gross Effects of Marital Status and Health on Mortality

First we show the effects of marital status and health on mortality without accounting for potential endogeneity. Mortality risk is modeled as a continuous-time hazard (failure-time) process:

$$\ln h^d(t) = \alpha_0 + \alpha_1' T(t) + \alpha_2' X(t) + \alpha_3' M(t) + \alpha_4 H(t), \quad (1)$$

where $T(t)$ denotes duration dependencies on age and calendar time (both piecewise-linear splines),[5] $X(t)$ represents exogenous demographic regressors, $M(t)$ is a vector of marital status indicator variables (where "married" is the omitted category), and $H(t)$ denotes current health status. The appendix describes maximum-likelihood estimation details. Table 2 shows the estimated coefficients on marital status and health.[6]

The first column does not include controls for current health status. It confirms the widely found result that married men face lower mortality risks than unmarried men. The effect is strongest for never-married and divorced men. In the second column we add current health status. This self-reported general health status ranges from 1 (poor) to 5 (excellent), and is treated here as a cardinal measure. As expected, better health is associated with lower mortality risks: When health is controlled, divorced men do not experience significantly higher mortality risks than married men. In other words, health appears to be the intervening variable that explains most of the observed difference in mortality risk between married and divorced men. This finding does not apply to never-married and widowed men; their excess

2. Our empirical models incorporate the ordinality of health reports.

3. Our analyses explicitly condition on selective survival to participate in the 1984 survey (see appendix).

4. In addition, 484 men (11.6%) left the sample. One may suspect that these individuals are at above-average risk of dying. The Institute for Social Research (ISR), however, which oversees the PSID data collection, employs a number of mechanisms to check on the mortality of individuals who leave through attrition.

5. The shapes of these piecewise-linear (piecewise-Gompertz) duration dependencies are identified individually because respondents are born at different dates; thus their "clocks" start ticking at different dates.

6. In addition to age, calendar time, health, and marital status, the regressions control for race and education. The full set of estimates of this and all following models is available on request.

TABLE 2. GROSS EFFECTS OF MARITAL STATUS AND HEALTH ON MORTALITY

	(1)	(2)
Never Married	0.7898***	0.9585***
	(0.2276)	(0.2272)
Separated	−0.0010	−0.3711
	(0.5551)	(0.5642)
Divorced	0.5888***	0.3157
	(0.2266)	(0.2416)
Widowed	0.3775**	0.4461 ***
	(0.1631)	(0.1715)
Health Status $H(t)$		−0.8530***
		(0.0544)

**$p < .05$; ***$p < .01$

Note: Standard errors in parentheses.

mortality must be due to something other than health as measured in the PSID. Naturally our self-reported general health measure is only one of many possible health measures and must be interpreted with this caution in mind.

Gross Effects of Marital Status on Health

The results in the previous subsection suggest that divorced men tend to be in poorer health than married men, but that no significant health differences exist between never-married, widowed, and married men. We estimate a model to explain health differences across marital statuses. Health status is measured by up to seven annual observations (1984–1990) of an ordinal qualitative health indicator. Our ordered probit model of health status explicitly incorporates this ordinal property of the reported general health indicators, $H(t)$. Long-run health status is assumed to be a function of age, marital status, and demographic factors:

$$H^{*}(t) = \eta_0 + \eta_1' A(t)\left[1 + \eta_2' M(t)\right] + \eta_3' M(t) + \eta_4' X(t) + \varepsilon_h, \quad (2)$$

where $A(t)$ represents a piecewise-linear spline transformation of age, $M(t)$ is a vector of marital status indicator variables, $X(t)$ denotes exogenous demographic regressors (race, education, and several instruments that are described below), and ε_h represents respondent-specific variation (heterogeneity) in long-run health status. The reference person is a married man age 50. Marital status both affects the intercept ($\eta_3' M(t)$) and tilts the age profile around reference age 50 ($\eta_2' M(t)$). The person-specific random effect, ε_h, captures determinants of health that are unmeasured in our data, such as innate frailty, health habits, and preferences for regular exercise. This effect is assumed to be distributed normally with mean 0 and standard deviation σ_{ε_h}, and is identified by the repeated (up to seven) health measures for each respondent.

At each survey wave, respondents are assumed to report a health category on the basis of an underlying continuous health value $H^{**}(t)$, which consists of the long-run health value and a transitory residual term:

$$H^{**}(t) = H^{*}(t) + u(t) \quad (3)$$

where $u(t)$ is normalized to be distributed standard normally. The correspondence between the continuous health value and the ordered categories is denoted by

$$H(t) = \begin{cases} \text{poor (1)} & \text{if } H^{**}(t) < 0 \\ \text{fair (2)} & \text{if } 0 \le H^{**}(t) < \tau_2 \\ \text{good (3)} & \text{if } \tau_2 \le H^{**}(t) < \tau_3 \\ \text{very good (4)} & \text{if } \tau_3 \le H^{**}(t) < \tau_4 \\ \text{excellent (5)} & \text{if } \tau_4 \le H^{**}(t), \end{cases}$$

where the three thresholds τ_2, τ_3, and τ_4 are common across years and are parameters[7] to be estimated, and $t = 84,\ldots,90$ are survey dates. The underlying health value is given by the probit index function discussed above. Details of model specification and estimation are presented in the appendix.

Table 3 presents the subset of coefficient estimates that relate to age and marital status. All age coefficients are negative, indicating that, on average, health deteriorates with age. We estimated the first column without interactions between age and marital status—that is, without the $\eta_2' M(t)$ term. Separated and divorced men report worse health than their married counterparts, whereas never-married and widowed men are, on average, approximately as healthy as married men. In the second column we add the interaction of age and marital status; intercept differences apply to the reference age, 50. The interaction terms show that never-married, divorced, and widowed men have significantly different age profiles: The health of never-married and divorced men deteriorates with age about 14% and 16% faster, respectively, than that of married men. On average, widowers report about the same level of health as married men (Column 1), but we find that widowers' health deteriorates 33% more slowly than that of married men. At age 50 widowers are in worse health, but the difference reverses at older ages.

These results are consistent with our prior finding that health appears to be the intervening variable which explains excess mortality among divorcés, but not among bachelors and widowers. These gross relationships, however, may be due partially to selective marriage or divorce behavior. Marital status may be endogenous to health and mortality, so that our gross estimates are biased.

Gross Effects of Health on Marital Status

Researchers ask repeatedly whether healthier persons are more likely to marry and remain married. To address this issue of reverse causality, we analyze the effects of health on marriage formation and dissolution behavior. Formally we consider a marriage dissolved at the date of the separation, not at the legal divorce date, and we do not model the transition from separation to divorce. We use a continuous-time

7. The first threshold, τ_1, is normalized to 0 to set the origin of the index.

TABLE 3. GROSS EFFECTS OF MARITAL STATUS ON HEALTH

	(1)	(2)
Intercept Differences at Age 50 $(\eta_3' M(t))$		
Never married	−0.0508	−0.1531 **
	(0.0407)	(0.0633)
Separated	−0.2022 ***	−0.1567 **
	(0.0557)	(0.0685)
Divorced	−0.1942 ***	−0.2390 ***
	(0.0344)	(0.0432)
Widowed	0.0763	−0.1915 **
	(0.0577)	(0.0958)
Age Spline (Relative to Age 50)		
< 40 years	−0.0374 ***	−0.0367 ***
	(0.0033)	(0.0033)
40–50 years	−0.0447 ***	−0.0424 ***
	(0.0045)	(0.0045)
51–65 years	−0.0507 ***	−0.0501 ***
	(0.0035)	(0.0035)
66–80 years	−0.0464 ***	−0.0476 ***
	(0.0044)	(0.0044)
> 80 years	−0.0878 ***	−0.1022 ***
	(0.0132)	(0.0140)
Age Interactions $(\eta_2' M(t))$		
Never married		0.1391 **
		(0.0703)
Separated		−0.0806
		(0.1041)
Divorced		0.1578 **
		(0.0739)
Widowed		−0.3315 ***
		(0.0797)

**$p < .05$; ***$p < .01$

Note: Standard errors in parentheses.

hazard formulation for both marriage formation and dissolution. The models are similar to the mortality model described above, except that we add heterogeneity terms to allow for the possibility that men differ, because of unmeasured factors, in their propensities to marry or separate. We assume these factors apply equally to all marriages of any particular man.[8] The log-hazards of marriage formation and dissolution are given by

$$\ln h_j^m(t) = \beta_0 + \beta_1' T(t) + \beta_2' X(t) + \beta_3 H^*(t) + \varepsilon_m, \quad (4)$$

$$\ln h_j^s(t) = \gamma_0 + \gamma_1' T(t) + \gamma_2' X(t) + \gamma_3 H^*(t) + \varepsilon_s, \quad (5)$$

where j is the (upcoming or current) marriage number, m indicates marrying and s indicates separating, $T(t)$ is a vector of piecewise-linear duration dependencies, $X(t)$ represents exogenous demographic variables, $H^*(t)$ is a measure of health, and ε_m and ε_s are heterogeneity terms reflecting person-specific unmeasured factors. Duration dependencies $T(t)$ in the marriage formation equation include respondent's age, calendar time, duration since leaving school, and (for divorced and widowed men) duration since becoming eligible for remarriage. The hazard of separation depends on duration since the wedding, respondent's age, and calendar time. Characteristics $X(t)$ include race, education, number of children from current or previous marriage(s), indicators for being divorced and widowed (in the marriage equation), spousal characteristics, and strictness of divorce laws (by state and year, in the separation equation). We assume that heterogeneity components ε_m and ε_s are distributed normally with mean 0 and standard deviations σ_{ε_m} and σ_{ε_s}, respectively. The appendix contains details of the estimation procedure.

We explore three different health measures and estimate their effects on the full marriage histories, including transitions that took place before 1984. Table 4 shows the subset of coefficient estimates in marriage formation (Column 1) and dissolution (Column 2) related to the three health measures. In the first row, the average reported health status is substituted for $H^*(t)$. This specification ignores the ordinal character of reported health; we take the simple average over (up to) seven health reports. Healthier men appear to marry younger, remarry sooner, and remain married longer. In other words, there may be positive selection into marriage, which may account in part for married men's observed health advantage.

In the second row we use a variance component from the ordered probit model of health discussed earlier, but without any covariates. This variance component may be interpreted as a person-specific deviation from the overall reported mean health status, averaged over all his reports (1984–1990). The results are the same except for scale, an indication that using a variance component (as we do again below) is not the source of our results relative to studies that use a raw reported health score.

These average health measures suggest that better health is associated with a greater hazard of marriage and a lower hazard of marital dissolution. Both of these findings represent the commonly accepted relationships. The significance of these health measures indicates the durability of measured health in its ability to predict changes in marital status over the whole life cycle.

Next we enrich the health measure to include a piecewise-linear spline in age and a number of instruments (constant over the life cycle) in addition to a health variance component (now a residual). The instruments predict health but are assumed not to directly affect changes in marital status (see below). The specification may be viewed as following from a two-step procedure. First, we estimate an ordered

8. Conceptually, mortality risk also may be heterogeneous. Yet because "you only live once," such a term would be difficult to identify separately from the age/duration pattern in the mortality equation; thus we omit it. By contrast, we observe up to five marriages for any one man in the data.

TABLE 4. GROSS EFFECTS OF HEALTH ON MARRIAGE FORMATION AND DISSOLUTION

	Marriage Formation	Marriage Dissolution
Average Health Indicator (Raw Scale 1–5)	0.0676 *** (0.0149)	–0.1420 *** (0.0296)
Health Variance Component	0.0439 *** (0.0104)	–0.0889 *** (0.0200)
Predicted Health Index (Based on Instruments)	–0.4732 *** (0.0478)	–0.1995 ** (0.0933)
Residual Variance Component	0.0642 *** (0.0108)	–0.0938 *** (0.0207)

$^{**}p < .05$; $^{***}p < .01$

Note: Standard errors in parentheses.

probit function of health, similar to Eq. (2). Second, for each individual respondent, using the ordered probit estimates, we predict the part of the health index due to the measured covariates and the person-specific part due to unmeasured factors (ε_h). The residual variance component representing unmeasured factors may be viewed as the person-specific deviation from the predictions, averaged over all health reports (1984–1990).[9] This formulation allows the prediction of a health index over the life cycle (including ages before 1984) in addition to the variance component, which is determined entirely by the 1984–1990 reports.

The striking result in Table 4 is the negative sign on the predicted health coefficient in the marriage formation equation (–0.4732): Healthier men apparently are *less* likely to (re)marry. The negative sign indicates *adverse* selection into marriage on the basis of measured covariates predicting long-run health, and hints that two types of selectivity may be at work: one adverse on the basis of health (as measured in the PSID), and the other positive on the basis of unmeasured characteristics that affect both health and marrying. The residual component (representing 1984–1990) has a significant positive sign, indicating that unmeasured factors affecting both health and the hazard of marriage are correlated positively.

The instruments that identify the health effect measure whether anyone in the respondent's 1968–1969 household smoked, parents' education, whether the respondent grew up in poverty, and whether or not the respondent grew up on a farm or in a small town (see Table 5 below). The assumption that these items affect transitions in marital status only through health may be somewhat arbitrary. Adding a subset of these instruments to the marriage or separation hazard equations did not substantially change the result over a large number of alternatives.

The gross relationships that we have discussed here suggest that health status affects transitions in marital status. The implication is that it may be inappropriate to treat marital status as exogenous, as we did in the health and mortality equations above. The estimated effects of marital status may be biased by reverse causality. Therefore we estimate the above equations jointly, allowing all heterogeneity components to be correlated freely.

A SIMULTANEOUS MODEL OF MARITAL STATUS, HEALTH, AND MORTALITY

We address the potential endogeneity of marital status for health and mortality outcomes by estimating a system of simultaneous equations involving mortality, health, marriage formation, and marriage dissolution. Endogeneity bias may arise from the presence of unmeasured factors that affect both health and transitions in marital status, thus inducing a spurious correlation of an explanatory variable (marital status, health) with the residual portion of the outcome (health, marital status). A joint model explicitly makes these sources of correlation a part of the model, thereby eliminating the bias. We account for correlation due to person-specific unmeasured factors that are identified through repeated observations of a particular outcome: up to seven health reports, up to five marriages, and up to five separations.

Our method does not account for any bias introduced by a correlation of marital status with a transitory residual, such as a permanent shock in health.[10] As we see below, however, 64% of the total residual variance in health ($\varepsilon_h + u(t)$) is due to heterogeneity component ε_h.

Health is an explanatory variable in the equations for the risk of marriage formation and marriage dissolution. Recall that the log hazards of marriage formation and marriage dissolution are given respectively, by

$$\ln h_i^m(t) = \beta_0 + \beta_1' T(t) + \beta_2' X(t) + \beta_3 H^*(t) + \varepsilon_m , \qquad (4)$$

$$\ln h_j^s(t) = \gamma_0 + \gamma_1' T(t) + \gamma_2' X(t) + \gamma_3 H^*(t) + \varepsilon_s , \qquad (5)$$

where $H^*(t)$ represents health. Each of these equations includes both a predicted part, depending on time-varying measured covariates, and a residual part, reflecting the effects of person-specific unmeasured factors that influence marital change, which is identified from multiple changes of marital status for many respondents.

Health is measured as long-run health status, given by the health equation

9. This is a generalization of the two-stage least squares (2SLS) approach to a context with ordered probit and failure-time models. Even though the health index is estimated by using an ordered probit model, it is a continuous variable. Its prediction enters a hazard specification (of marriage formation or dissolution). The health index is a piecewise-linear spline function of age, which also determines the baseline hazard pattern. The health index prediction thus adds a second duration dependency on age in the hazard models. Our ordered probit health model includes a variance component to capture person-specific unmeasured effects, which enters as a second term in the two-stage predicted value.

10. In our data, health status is measured on an ordinal scale and is observed only in snapshots at one-year intervals. Controlling for lagged health status, so that health shocks may have permanent effects, thus would be logically incoherent. Such a model would require a more continuous, and more frequently measured health measure.

$$H^*(t) = \eta_0 + \eta_1' A(t)\left[1 + \eta_2' M(t)\right] + \eta_3' M(t) + \eta_4' X(t) + \varepsilon_h \quad (2)$$

which itself depends on current marital status, represented by $M(t)$. This measure of health includes both a predicted part, based on observed characteristics, and a residual part reflecting person-specific unmeasured factors, which is identified from multiple health reports. The procedure we followed in our discussion of gross effects of health on marital status treated those two parts separately; the fully joint model accounts for any correlation in unmeasured factors affecting more than one outcome directly as part of the residual structure ($\rho_{\varepsilon_m\varepsilon_h}$ and $\rho_{\varepsilon_s\varepsilon_h}$). The risk of mortality is a function of both marital status and long-run health, in addition to other covariates:

$$\ln h^d(t) = \alpha_0 + \alpha_1' T(t) + \alpha_2' X(t) + \alpha_3' M(t) + \alpha_4 H^*(t) + \lambda'\varepsilon \quad . \quad (6)$$

We also allow residual heterogeneity components from marriage formation, marriage dissolution, and health processes to affect the risk of mortality,[11] $\lambda'\varepsilon = \lambda_m\varepsilon_m + \lambda_s\varepsilon_s + \lambda_h\varepsilon_h$. The full model is given by Eqs. (2), (4), (5), and (6). Heterogeneity components ε_m, ε_s, and ε_h are allowed to be correlated freely.

To identify the effects of health on the risks of marriage formation and dissolution and on the risk of mortality, at least one nontrivial identifying covariate is needed: one which significantly affects health but does not directly affect the other outcome.[12] Similarly, to identify the effects of marital status on health and the risk of mortality, it is desirable[13] to use covariates that affect marital status transitions but do not directly affect health and mortality risk. Here we discuss our choice of instruments and the sensitivity of the estimates to those choices.

Table 5 provides an overview of the model specifications and exclusion restrictions. Health is a function of marital status, age (to capture the decline in health over one's lifetime), race, and education (related to income and thus to the ability to purchase medical care and other health inputs, and capturing access to information on efficient production of health). The effects of health on other outcomes are identified by the presence of variables presumed to affect health but not the other outcomes. These include coresidence with smokers in 1968–1969 (a negative health input), parental education (related to efficiency of health production in childhood), childhood poverty (related to access to health care and nutrition), and an indicator for whether or not the respondent grew up on a farm or in a small town (capturing density of health care providers in childhood). Formally only one covariate is needed if it appears only in the health equation; thus the set of exclusion restrictions represented in Table 5 substantially overidentifies the effect of health, but is our preferred specification.

We tested the robustness of the results of the joint model to the choice of instruments and found that the sign, magnitude, and significance of the health effects are affected only mildly. Adding a subset of the instruments to the hazard equations (removing them as instruments) does not substantially change the result, regardless of which subset is removed (for all possible subsets, including leaving in only one covariate). The result thus is robust.

Marriage formation is a function of current marital status, age, and the duration for which one has been eligible for (re)marriage, health status, calendar time (capturing secular trends in marriage rates), race, and education and enrollment (to account for the competition between the demands of schooling and access to resources needed to support a family). Covariates affecting only marriage include the duration since leaving school and the number of children by prior partners (which may be reared at lower cost inside marriage). Marriage dissolution is a function of marriage duration (related to information on spousal compatibility and to investment in the relationship), health, age, calendar time, race, own and wife's education (possibly reflecting ability to communicate), and age and race differences between husband and wife (measuring spousal compatibility). Covariates affecting only marriage dissolution include children by prior partners (related to costs to the spouse of remaining in the marriage), children by the current partner (reflecting investments in the marriage), and the strictness of divorce laws (by state and year, capturing costs of divorce).

Mortality, finally, is a function of marital status, health, age, calendar time, race, and education (which capture factors similar to those affecting health). As mentioned above, the exclusion of family background variables from the marriage formation and dissolution processes may be somewhat arbitrary, but extensive tests showed that the effects of health are very robust to omission of subsets of instruments.

The full model includes both continuous-time hazard models and ordered probit outcomes. The appendix contains a detailed discussion of the estimation procedure. Also see Lillard (1993) for simultaneous equations in hazard models, and Panis and Lillard (1994) for simultaneous equations in hazard and probit models.

RESULTS

Unless noted otherwise, the following results are based on the fully joint model estimates for marriage formation and dissolution, health, and mortality. The parameter values discussed here thus account for the simultaneity of the relationships involved, including the effects of unmeasured factors affecting more than one outcome. We explore results related to the following questions and issues: Does marriage have a direct benefit on health, or is there only self-selection into marriage on the basis of health? In what direction is the relationship? Is the selection positive or adverse? Is self-selection present among those who remain married? Are estimates

11. We do not include a separate heterogencity term in the mortality equation because only one mortality spell per respondent is observed.

12. More formally, these are zero-coefficient restrictions in the other outcome equations.

13. In this way, identification is not based solely on fixed effects.

TABLE 5. VARIABLE DEFINITIONS AND MODEL SPECIFICATIONS

	Outcome Equation Set			
Explanatory Variable	Health	Marriage	Dissolution	Mortality
Marital Status[a]				
Indicators for unmarried	X	X		X
Duration since last married		X		
Marriage duration			X	
Health				
Latent health index $H^*(t)$		X	X	X
Exogenous Regressors				
Age $A(t)$	X	X	X	X
Calendar time $T(t)$		X	X	X
Race (black/nonblack)[b]	X	X	X	X
Education	X	X	X	X
Spouse's education			X	
Age husband minus age wife			X	
Husband and wife same race			X	
Out of school and how long		X		
How long unmarried		X		
Children with prior partners		X	X	
Children with current partner			X	
Strictness of divorce laws			X	
Any smokers in household, 1968–69	X			
Parents' education	X			
Grew up in poverty	X			
Grew up on farm or in small town	X			

[a]Relative to married.

[b]Nonblack is termed white.

of the direct protective effect of marriage on health and mortality affected substantially by accounting for self-selection?

Effects of Health on Marriage Formation and Dissolution

Long-run health, defined in Eq. (2), is a predicted index based both on observed characteristics and on the heterogeneity term ε_h. Table 6 shows the direct effect of long-run health $H^*(t)$ on the hazards of (re)marriage and dissolution, as well as the estimated correlations between the heterogeneity components in health, on the one hand, and marriage formation ($\rho_{\varepsilon_m \varepsilon_h}$) and dissolution ($\rho_{\varepsilon_s \varepsilon_h}$), on the other. Other covariates are controlled as observed in Table 5.

After accounting for unmeasured factors, and contrary to conventional wisdom, we find that healthier men are actually less likely to marry; that is, healthier men marry later and postpone remarriage. Relatively unhealthy men tend to (re)marry early; thus there is *adverse* selection into marriage on the basis of health. At the same time, we find *positive* selection into marriage on the basis of unmeasured factors:[14] A significant and positive correlation exists, induced by unmeasured factors that affect both health and the hazard of marrying. In other words, unmeasured factors that promote good health also tend to encourage marriage, and unmeasured factors that harm health also tend to discourage marriage. This strong positive correlation in unmeasured factors implies that men who marry early tend to be healthier than men who remain unmarried, conditional on the measured factors that affect general health; this point in turn implies *positive* selection into marriage. This selection is not based on general health as measured in the PSID data; it is based on

14. Adverse selection into marriage on the basis of health is not governed by any structure placed on the joint model. We also found it in the exploratory specification of the hazard of marriage formation, where we distinguished between the effects due to predicted long-run health on the basis of measured and unmeasured factors. The coefficient on the measured part was negative, indicating adverse selection, whereas the unmeasured portion of health tended to encourage marriage (Table 4).

TABLE 6. DIRECT EFFECTS OF HEALTH ON MARRIAGE FORMATION AND DISSOLUTION, AND RESIDUAL HETEROGENEITY CORRELATIONS

	Marriage Formation	Marriage Dissolution
Fully Joint Estimates		
Direct effect of long-run health $H^*(t)$	−0.4190 *** (0.0652)	−0.3110 *** (0.0942)
Residual correlation with health $\rho_{\varepsilon_m\varepsilon_h}$, $\rho_{\varepsilon_s\varepsilon_h}$	0.5827 *** (0.0571)	0.6757 *** (0.1067)
Pairwise Joint Estimates with Health		
Direct effect of long-run health $H^*(t)$	−0.7120 *** (0.0648)	0.0779 (0.0498)
Residual correlation with health $\rho_{\varepsilon_m\varepsilon_h}$, $\rho_{\varepsilon_s\varepsilon_h}$	0.7691 *** (0.0248)	−0.4256 *** (0.0661)

**$p < .05$; ***$p < .01$

Note: Standard errors in parentheses.

unmeasured factors that are correlated positively with that health measure.[15] As we see below, each of the two selection mechanisms (measured and unmeasured factors) may dominate under different circumstances.

Although this finding is novel, it is consistent with the argument unhealthy men have an incentive to seek the health protection offered by marriage. A similar argument has been well established in the health economics literature (Grossman 1972). For example, a visit to a doctor tends to benefit one's health (net effect). Doctors, however, tend to be visited disproportionately by unhealthy people (adverse selection). The combination of the net beneficial effect and the adverse selection may result in either a positive or a negative gross relationship between doctor visits and health, depending on which dominates. In our application, marriage may be viewed as a health input that is sought disproportionately by unhealthy men.

In addition to the net beneficial effect of marriage and the adverse selection based on health, we find a third confounding factor, namely a positive correlation in unmeasured factors related both to health and to the propensity to marry. The gross effect of health on marriage is positive (Table 3 above); that is, adverse selection based on health is dominated by the protective effects of marriage and by the positive selection on the basis of unmeasured factors.

After unmeasured factors are accounted for, healthier men have a reduced hazard of marriage dissolution; that is, they tend to remain married longer. This finding is consistent with the gross relationships explored above. Yet we also find a significant selection effect on the basis of unmeasured factors: Men who (for unmeasured reasons) tend to be in good health also tend to have a high risk of divorce (adverse selection). This result differs both from the findings discussed above and from results based on pairwise joint estimation involving health and marriage dissolution, but omitting the marriage equation altogether (see Table 6, bottom panel). We arrive at this finding in the joint model because we now also account for nonrandom entry into marriage. Men who are at risk of a divorce must have married in the past; this implies that they were self-selected into marriage. The correlation between heterogeneity components in the marriage formation and the dissolution equations is very close to 1 ($\rho_{\varepsilon_s\varepsilon_h} = 0.9638$), and the gross relationship between health and marriage dissolution is due to nonrandom entry into marriage.

Effects of Marital Status on Health

We showed above that health partially determines marriage formation and dissolution. Here we demonstrate the net direct effects of marital status on health. The first column of Table 7 shows estimates of marital status coefficients where marital status is treated exogenously (taken from Table 3). The second column is taken from the fully joint model of marriage formation, dissolution, health, and mortality. As discussed above, the estimates assuming exogenous marital status indicate that never-married men are less healthy than married men. The second column, however, shows no significant differences between never-married and currently married men. The gross health difference thus is due to selection on the basis of unmeasured characteristics. Never-married men with unmeasured characteristics (habits, preferences) that promote their good health tend to move into marriage early. The decision to marry for the first time, by itself, does not generate any benefits as measured by self-reported general health status.

Divorced men, by contrast, may expect a health benefit from remarriage. For divorced men over age 50, the net benefit is even larger than the gross benefit, an indication that the negative direct health effect (rather than the effect of unmeasured factors) dominates for this group. At reference age 50, both specifications show that divorced men's underlying health index is about 0.23 point lower than that of married men.[16] The net age pattern, however, is inclined downward more steeply than indicated by the gross estimates. Divorced men's health deteriorates 28% faster than that of married men, whereas the gross age trend is only 16% steeper. In other words, after age 50 the net health gap between married

15. Because of their very nature, the interpretation of unmeasured factors is speculative. One potential interpretation is found in men's preferences regarding lifestyle. Adventurous types may expose themselves to relatively large health risks; at the same time, they may have little desire (or ability) to commit to a long-term relationship. On the other end of the spectrum may be risk-averse men who prefer the safety of a family. Unmeasured factors also may be related to potential spouses' preferences. For example, heavy drinking and smoking have a negative influence on health and may discourage women from entering a relationship. The text provides yet another interpretation, namely through unmeasured differences in preferences for social contact. Finally, our health measure refers to self-reported general health; other dimensions of health are unmeasured.

16. To place this difference in perspective, the difference between the midpoints of average and of good health is about 1.5 points.

TABLE 7. GROSS AND NET EFFECTS OF MARITAL STATUS ON HEALTH

	Gross Effects	Net Effects
Intercept Differences at Age 50 $(\eta_3' M(t))$		
Never married	−0.1531 **	−0.0962
	(0.0633)	(0.1005)
Separated	−0.1567 **	−0.1920 ***
	(0.0685)	(0.0732)
Divorced	−0.2390 ***	−0.2317 ***
	(0.0432)	(0.0563)
Widowed	−0.1915 **	−0.2592 **
	(0.0958)	(0.1149)
Age Interactions $(\eta_2' M(t))$		
Never married	0.1391 **	0.0843
	(0.0703)	(0.0968)
Separated	−0.0806	0.0225
	(0.1041)	(0.1020)
Divorced	0.1578 **	0.2831 ***
	(0.0739)	(0.0832)
Widowed	−0.3315 ***	−0.3434 ***
	(0.0797)	(0.0738)

**$p < .05$; ***$p < .01$

Notes: Reference person is married, age 50. Standard errors in parentheses.

and divorced men is larger than the gross gap. This bias is due to the dominant negative direct effect of health on remarriage: Relatively unhealthy divorcés tend to remarry soon, thereby removing themselves from the divorced population. For older divorcés, this adverse selection on the basis of health dominates the positive selection on the basis of unmeasured characteristics.

Why does marriage protect against bad health? If the protection arises from better nutrition and availability of care in times of illness, one expects the marriage gain to increase with marriage duration. We tested this possibility but could not reject the hypothesis that the gain is constant over the entire marriage.

Effects of Marital Status and Health on Mortality

Table 8 shows the effects of long-run health and marital status on the hazard of mortality. The first column is estimated independently of other processes (from Table 2). The second column is based on the fully joint model. We do not gain any new substantive insights by modeling mortality jointly with health and marital transitions. Unmeasured factors affecting marriage formation, dissolution, and health enter the log hazard of mortality as $\lambda'\varepsilon = \lambda_m\varepsilon_m + \lambda_s\varepsilon_s + \lambda_h\varepsilon_h$, but we could not reject the joint hypothesis that all of the lambdas were 0. In other words, we found no evidence of endogeneity of marital status or health in the mortality equation, provided that one controls for both.

TABLE 8. GROSS AND NET EFFECTS OF MARITAL STATUS AND HEALTH ON MORTALITY

	Gross	Net
Never Married	0.9585 ***	1.1297 ***
	(0.2272)	(0.2374)
Separated	−0.3711	−0.5243
	(0.5642)	(0.5821)
Divorced	0.3157	0.2810
	(0.2416)	(0.2478)
Widowed	0.4461 ***	0.3903 **
	(0.1715)	(0.1810)
Health Status $H(t)$, $H^*(t)$	−0.8530 ***	−0.5546 ***
	(0.0544)	(0.0504)

**$p < .05$; ***$p < .01$

Note: Standard errors in parentheses.

Good health, not surprisingly, is related strongly and negatively to the risk of mortality.[17] As we saw before, never-married, divorced, and widowed men are at higher risk of mortality than married men. For divorced men, this higher mortality risk is explained largely by health as an intervening variable. Excess mortality among bachelors and widowers, however, is due to factors other than self-reported general health status.

The finding that excess mortality among never-married men is not due to health differentials raises the question of the underlying mechanism. Unmeasured factors possibly include a degree of preference for social contact. Lillard and Waite (1995) found that men's marriage advantage is due partly to coresidence with a partner or other adults. Never-married men may prefer to live alone, thus forgoing the benefits of social integration. Unfortunately our model permits the inclusion only of covariates that are known since age 12, so we cannot resolve this issue here.

SUMMARY AND CONCLUSION

Numerous studies have shown that marriage is associated with lower mortality rates and generally with better health. In this paper we investigate, first, the extent to which health as intervening variable explains the marriage advantage in mortality, and, second, the role of health in men's marriage behavior. Although it has often been suggested that selection mechanisms may account for at least part of the marriage advantage, very little empirical work has been done to determine the role of selection into and out of marriage. In this paper we attempt to fill that gap.

The theoretical debate has centered around "protection" versus "selection": That is, the extent to which marriage protects its incumbents against poor health and mortality risks,

17. The scales of the health measures in Columns 1 and 2 are different. Health in Column 1 is measured as 1 = poor through 5 = excellent. In Column 2 we use the long-run health index (Eq. 2) ranging from below $\tau_1 = 0$ (poor) through above $\tau_1 - 4.5522$ (excellent).

versus the extent to which the relationship is due to the possibility that healthier individuals tend to be married more often—for example, because they are better able to find a partner and to maintain a long-term relationship. Exploration of this question points to positive selection into marriage.

Yet if marriage does enhance health and protect against mortality risks, an argument can be made for adverse selection into marriage. The theory of economic choice suggests that persons who benefit more from marriage in terms of better health and/or reduction of mortality risk are more likely to marry and less likely to leave the marriage. The argument alternatively may be phrased in terms of health production. Inputs that improve health, such as marriage, are most likely to be "purchased" by those who need its protection most. This analogy is not complete, however: Men in poor health may be considered unattractive, so that they cannot "purchase" protection through marriage.

We explore these relationships empirically, using a model that explicitly incorporates the possibility of reverse causality: Long-run general health status may affect marriage decisions. In addition, marital status may affect health, and both may affect mortality. Underlying the system is a structure that explicitly incorporates the effects of unmeasured factors which may affect multiple outcomes.

We find support for the hypothesis that men base their marriage decisions in part on the potential health benefits. Relatively unhealthy men tend to (re)marry early and to remain married longer; that is, we find adverse selection into marriage based on self-perceived general health. Our results also show evidence of positive selection into marriage based on unmeasured factors. Such positive selection dominates for never-married men. The unmeasured factors may include the degree to which one's lifestyle includes adventure, risk, and stability, or the degree of preference for social contact. This positive selection may be due to any habit or preference that tends to both promote good (bad) health and encourage (discourage) marriage. Adverse selection with regard to general health dominates for divorced men over age 50. Overall the positive selection dominates.

Thus we find evidence for both positive and adverse selection, and show that each dominates in different circumstances. Health models that ignore the endogeneity of marital status thus yield estimates of the beneficial effects of marriage that are biased upward or downward, depending on marital status. The net health benefit of entry into a first marriage is zero, but a significant health gain may be expected from remarriage by divorced men. *Health* throughout this paper refers to self-reported general health status; the results may vary for other types of health measures.

In this study, joint estimation of a system of simultaneous equations reveals more than do the more commonly used methods of accounting for endogeneity, such as the two-stage instrumental variable method. Both types of method are capable of detecting adverse selection into marriage based on of health, but only joint estimation provides estimates of correlations induced by unmeasured factors. These correlations often have a substantive interpretation, such as positive selection into marriage based on characteristics that are correlated with health.

In keeping with our findings on two types of selectivity, we show that high mortality rates among divorced men are explained largely by their poorer health, relative to married men. For never-married and widowed men, however, the higher mortality rates are explained by something other than self-reported general health status. Further research is needed to determine the nature of the intervening variable(s) for never-married and widowed men.

APPENDIX. JOINT ESTIMATION OF MODEL PARAMETERS

In this appendix we address issues related to estimation and identification of parameters. We begin by summarizing our overall approach to estimation and then discuss technical details. Because three of the four processes are similar hazard processes, we discuss a generic version of the hazard model we use. We then discuss the ordinal health indicator. Finally we draw together these results for the individual processes to discuss joint estimation and conditioning on survival to the initial interview in 1984.

Overview of Estimation

We estimate parameters by full-information maximum likelihood for all four processes combined, using analytic first derivatives and the BHHH search algorithm (Berndt, Hall, Hall, and Hausman 1974). We jointly model the event history of all four types of outcomes for each person. The joint likelihood function is the probability of the person's joint observed event histories conditional on all exogenous covariates and model parameters, including both the regression parameters of the hazard and index functions and the stochastic specifications. The processes are linked together directly by their interdependencies in terms of direct effects on each other, and indirectly through their common dependence on unmeasured factors, which affect all of them. Each process includes both stochastic variation unique to a particular outcome (e.g., a marriage) and a vector of heterogeneity components constant over decisions within a process (e.g., all marriages). These heterogeneity components link processes together and thus may be correlated across processes. They are assumed to be jointly normally distributed.

For any given person, conditional on his vector of heterogeneity components (as if they were known), the stochastic elements are independent both within and across processes; thus the joint probability of all outcomes is the product of the probabilities of the individual outcomes. Each independent conditional probability is termed a *conditional likelihood*; the product is termed the *joint conditional likelihood*. The heterogeneity components, however, are not known, and are "nuisance" parameters that must be integrated out. Put differently, the likelihood is a weighted average of all possible conditional likelihoods, where the weights are dictated by the distribution of the heterogeneity components. Given that their joint distribution is multivariate normal, the joint probability of all observed outcomes (the like-

lihood) may be written as the joint conditional likelihood times the distribution of heterogeneity components integrated over the full range of the heterogeneity components. The result is the *marginal likelihood* of the observed outcomes. This is the likelihood to be maximized.

Conditional Likelihood for a Panel Ordinal Health Indicator Model

Conditional on the measured covariates, $\underline{X}'(t) = \{T(t), M(t), X(t)\}$, and the person-specific components in the health index function (ε_h), the probabilities of the responses each year ($H(t)$) are independent. The conditional probabilities are given by the ordered probit probabilities:

l	$P(H(t) = l \mid \underline{X}(t), \varepsilon_h)$
1	$\Phi\left(-\left(\eta'\underline{X}(t) + \varepsilon_h\right)\right)$
2	$\Phi\left(\tau_2 - \left(\eta'\underline{X}(t) + \varepsilon_h\right)\right) - \Phi\left(-\left(\eta'\underline{X}(t) + \varepsilon_h\right)\right)$
3	$\Phi\left(\tau_3 - \left(\eta'\underline{X}(t) + \varepsilon_h\right)\right) - \Phi\left(\tau_2 - \left(\eta'\underline{X}(t) + \varepsilon_h\right)\right)$
4	$\Phi\left(\tau_4 - \left(\eta'\underline{X}(t) + \varepsilon_h\right)\right) - \Phi\left(\tau_3 - \left(\eta'\underline{X}(t) + \varepsilon_h\right)\right)$
5	$1 - \Phi\left(\tau_4 - \left(\eta'\underline{X}(t) + \varepsilon_h\right)\right)$

The conditional likelihood of an observed sequence of health reports values from 1984 to the final survey year T_h (1990, year before death, or year before leaving the sample) is given by

$$\ell_h\left(\left(H(t), t = 84, T_h\right) \mid \underline{X}, \varepsilon_h\right) = \prod_{t=84}^{T_h} P\left(H(t) = l \mid \underline{X}(t), \varepsilon_h\right).$$

Reduced-Form Hazard Equations

The latent health status measures ($H^*(t)$) are not directly observable. One must substitute for them in the equation for the hazards of marriage formation, dissolution, and mortality, yielding the reduced form hazard equations

$$\ln h_j^m(t) = \pi_{m0} + \pi'_{m1}T(t) + \pi'_{m2}X_j(t) + \varepsilon_m + \pi_{m3}\varepsilon_h$$

$$\ln h_j^s(t) = \pi_{s0} + \pi'_{s1}T(t) + \pi'_{s2}X(t) + \varepsilon_s + \pi_{s3}\varepsilon_h$$

$$\ln h^d(t) = \pi_{d0} + \pi'_{d1}T(t) + \pi'_{d2}X(t) + \varepsilon_d + \pi_{d3}\varepsilon_h \ ,$$

where the π_m, π_s, and π_d values are linear combinations of βs and ηs, γs and ηs, and αs and ηs, respectively. The equations are written in reduced-form for notational convenience, but all structural parameters are estimated directly. Distinguishing reduced-form from structural parameters requires an identifying variable affecting long-run health but not *directly* affecting marrying, separating, or mortality. Also, because ε_h is part of all reduced-form heterogeneity components, it will induce correlation between those components and with health status.

Because "you only live once," heterogeneity in the ortality equation is hard to identify from a single occurrence. Our model assumes that any heterogeneity in mortality is reflected in correlation with the heterogeneity components of the marriage, dissolution, and health processes. That is, these correlations are given by the following regression equation:

$$\varepsilon_d = \lambda_m\varepsilon_m + \lambda_s\varepsilon_s + \lambda_h\varepsilon_h.$$

Individual or joint tests of hypotheses that $\lambda_1 = 0$ provide tests for exogeneity with regard to the mortality process and to nonrandom censoring of marriage duration and unmarried spells and health status by mortality. We could not reject the restriction that all λs are individually or jointly equal to 0, so ε_d was omitted from our final specification.

Conditional Likelihood for a Generic Duration Process with Heterogeneity

To minimize repetition of model development, consider a generic log-hazard equation for episode e of process p (*m, s,* or *d)*:

$$\ln h_e^p(t) = \pi_{p0} + \pi'_{p1}T(t) + \pi'_{p2}X(t) + \pi'_{p3}\varepsilon \ ,$$

where π_p is π_m, π_s, or π_d. The sum of all forms of duration dependence, including any endogenous durations, is combined into $T(t)$. All duration dependences are captured in piecewise-linear spline transformations of time, so that their sum, again is a piecewise-linear spline. Covariates, including both exogenous and endogenous outcomes of other processes, are represented by the vector $X(t)$. Heterogeneity components are captured by the linear combination $\pi'_{p3}\varepsilon$.

The baseline survivor function is determined by the combined effect of all duration variables, which covary perfectly with time. That is,

$$S_0^p(t) = \exp\left\{-\int_{t_0}^{t} e^{\pi_{p0}+\pi'_{p1}T(\upsilon)}d\upsilon\right\} \ ,$$

where t_0 is the moment at which the event came to be at risk of occurrence.

The constant and time-varying covariates (constant within intervals) combine with the heterogeneity components to proportionally shift the baseline hazard. The resulting "conditional" (on heterogeneity) survivor function is given by

$$S^p(t, \chi(t), \varepsilon) = \prod_{i=1}^{I}\left[\frac{S_0^p(t_{i+1})}{S_0^p(t_i)}\right]^{\exp\left\{\pi'_{p2}X(t_i)+\pi'_{p3}\varepsilon\right\}} \ ,$$

where I is the number of subintervals within which covariates $X(t)$ are constant, and $t_{I+1} = t$. The conditional (on ε) density function for completed duration t_e^p is given by

$$f^p\left(t_e^p, \chi\left(t_e^p\right), \varepsilon\right) = S^p\left(t_e^p, \chi\left(t_e^p\right), \varepsilon\right)h\left(t_e^p, \varepsilon\right).$$

The survivor and density functions for each episode (both censored and uncensored) of process p may be combined into the joint conditional likelihood, which is given by

$$\ell_e^p\left(\left(t_e^p, D_e^p, e=1, N_p\right)\middle|\underline{\chi}(t), \varepsilon\right) = \prod_{e=1}^{N_p} S^p\left(t_{e\cdot}^p, \chi\left(t_{e\cdot}^p\right), \varepsilon\right)\left[h\left(t_{e\cdot}^p, \varepsilon\right)\right]^{D_e^p},$$

where N_p is the number of episodes observed for process p, D_e^p is 0 if duration e is censored and 1 if it is not, and $\underline{\chi}(t)$ denotes the full history of all time-varying covariates over the periods covered by all episodes of the process. Conditional on the heterogeneity component, the probabilities of the observed outcomes are independent and the conditional likelihood is the product of their probabilities.

Joint Marginal Likelihood over All Processes

The four equations contain a vector $\underline{\varepsilon}$ of three heterogeneity components: ε_h in health, ε_m in marriage formation, and e_s in marriage dissolution. The three components are assumed multivariate normal, so that $\underline{\varepsilon} \sim N(\underline{0}, \Sigma_{\underline{\varepsilon}\underline{\varepsilon}})$.

The full joint likelihood of all relevant life-cycle events, from age 12 until either the time of death or the time when a person's event history is censored by the final survey or by attrition, is given by the marginal likelihood obtained by integrating the product of independent conditional likelihoods over the full range of the three heterogeneity components weighted by their joint distribution. That is,

$$\begin{aligned} &L\left(H(\vartheta), \vartheta = 84, \ldots, T_h;\right. \\ &\left.\left(\left(t_{e\cdot}^p, D_e^p, e=1, \ldots, N_p\right), p=m,s,d\middle|\underline{X}, \underline{\chi}\left(t^*\right)\right)\right) \\ &= \int_{\varepsilon_h}\int_{\varepsilon_m}\int_{\varepsilon_s} f_n\left(\underline{\varepsilon}\middle|\Sigma_{\underline{\varepsilon}\underline{\varepsilon}}\right)\prod_{t=84}^{T_h} P\left(H(t)=l_t\middle|\underline{X}(t), \varepsilon_h\right)\prod_{j=1}^{N_m} S^m\left(t_{j\cdot}^m, \chi\left(t_{j\cdot}^m\right), \varepsilon_m\right) \\ &\left[h\left(t_{j\cdot}^m, \varepsilon_m\right)\right]^{D_j^m}\prod_{j=1}^{N_s} S^s\left(t_{j\cdot}^s, \chi\left(t_{j\cdot}^s\right), \varepsilon_s\right)\left[h\left(t_{j\cdot}^s, \varepsilon_s\right)\right]^{D_j^s} \\ &S^d\left(t_{\cdot}^d, \chi\left(t_{\cdot}^d\right), \varepsilon_d\right)\left[h\left(t_{\cdot}^d, \varepsilon_d\right)\right]^{D^d} d\varepsilon_h d\varepsilon_m d\varepsilon_s, \end{aligned}$$

where $f_n(\cdot)$ denotes the normal density function, $\underline{X}$ denotes the values of covariates in the health index function 1984–T_h, and $\underline{\chi}(t^*)$ denotes the full history of covariates relevant to marriage, dissolution, and mortality from age 12 until the final observation date, t^*.

This is the likelihood function when an individual's full life history is observed. Our sample consists of persons alive to participate in the 1984 wave of the PSID when the general health sequence began. Therefore the likelihood of the observed information must be made conditional on survival to that survey date. The pre-1984 period could be ignored if the mortality equation did not include heterogeneity; the probability of survival through 1984 then would cancel in numerator and denominator. In the current application including health, however, we must account explicitly for survival. The probability is

$$S^d\left(t_{84}\middle|\chi\left(t_{84}\right)\right) = \int_{\varepsilon_h} f_n\left(\varepsilon_h\middle|\sigma_{\varepsilon_h}^2\right) S^d\left(t_{\cdot}^d, \chi\left(t_{\cdot}^d\right), \pi_{d3}\varepsilon_h\right) d\varepsilon_h,$$

where t_0 is the time at which the respondent came to be at risk of dying (twelfth birthday) and t_{84} is the 1984 survey date.

Conditional on survival through 1984, the likelihood is given by

$$L = \frac{L\left(H(\vartheta), \vartheta = 84, \ldots, T_h; \left(\left(t_{e\cdot}^p, D_e^p, e=1, \ldots, N_p\right), p=m,s,d\middle|\underline{X}, \underline{\chi}\left(t^*\right)\right)\right)}{S^d\left(t_{84}\middle|\chi\left(t_{84}\right)\right)}$$

The ability to use this formulation of the model depends critically on possessing information on the full marital history before 1984. (Because we do not observe anyone who died before 1984, we cannot compute the secular time trend in mortality before that date. We use data from Vital Statistics (1900–1990), by gender and race, to estimate mortality age and time trends, and we estimate age and time in the PSID in deviations from Vital Statistics estimates.) Although no information on health is available before 1984, the predicted values $H^*(t)$ can be computed at any point over the life cycle.

We estimate all parameters jointly, using full-information maximum likelihood. The heterogeneity components are integrated out numerically with a multivariate generalization of Gauss-Hermite approximation (Davis and Rabinovitz 1967; Naylor and Smith 1982; Panis 1992). We compute all first derivatives analytically; we approximate second derivatives (and standard errors) by the negative of the outer product of first derivatives (Berndt et al. 1974).

REFERENCES

Becker, G.S. 1981. *A Treatise on the Family*. Cambridge, MA: Harvard University Press.

Becker, G.S., E.M. Landes, and R.T. Michael. 1977. "An Economic Analysis of Marital Instability." *Journal of Political Economy* 85:1141–87.

Berndt, E.R., B.H. Hall, R.E. Hall, and J.A. Hausman. 1974. "Estimation and Inference in Non-Linear Structural Models." *Annals of Economic and Social Measurement* 3:653–65.

Carter, H. and P. Glick. 1976. *Marriage and Divorce: A Social and Economic Study*. Cambridge, MA: Harvard University Press.

Davis, P.J. and P. Rabinowitz. 1967. *Numerical Integration*. Waltham, MA: Blaisdell.

Farr, W. 1858. "The Influence of Marriage on the Mortality of the French People." Pp. 504–13 in *Transactions of the National Association for the Promotion of Social Sciences*, edited by G.W. Hastings. London: John W. Parker and Son.

Feinstein, J. 1993. "The Relationship between Socioeconomic Status and Health: A Review of the Literature." *Milbank Quarterly* 71:279–322.

Fu, H. and N. Goldman. 1994. "Are Healthier People More Likely to Marry? An Event History Analysis Based on the NLSY." Working Paper 94–5. Office of Population Research.

Goldman, N. 1993. "Marriage Selection and Mortality Patterns: Inferences and Fallacies." *Demography* 30:189–208.

Gove, W.R. 1973. "Sex, Marital Status, and Mortality." *American Journal of Sociology* 70:45–67.

Grossman, M. 1972. "On the Concept of Health Capital and the

Demand for Health." *Journal of Political Economy* 80:223–55.
Hu, Y. and N. Goldman. 1990. "Mortality Differentials by Marital Status: An International Comparison." *Demography* 27:233–50.
Hurd, M.D. 1990. "Research on the Elderly: Economic Status, Retirement, and Consumption and Saving." *Journal of Economic Literature* 28:565–637.
Idler, E.L. and S. Kasl. 1991. "Health Perceptions and Survival: Do Global Evaluations of Health Status Really Predict Mortality?" *Journal of Gerontology* 46(2):S55–65.
Kisker, E.E. and N. Goldman. 1987. "Perils of Single Life and Benefits of Marriage." *Social Biology* 34:135–52.
Kobrin, F.E. and G.E. Hendershot. 1977. "Do Family Ties Reduce Mortality? Evidence from the United States, 1966–1968." *Journal of Marriage and the Family* 39:737–45.
Korenman, S., N. Goldman, and Y. Hu. 1990. "Health and Mortality Differentials by Marital Status at Older Ages: Economics and Gender." Mimeographed document, Princeton University.
Kuh, D.J.L. and M.E.J. Wadsworth. 1993. "Physical Health Status at 36 Years in a British National Birth Cohort." *Social Science and Medicine* 37:905–16.
Lillard, L.A. 1993. "Simultaneous Equations for Hazards: Marriage Duration and Fertility Timing." *Journal of Econometrics* 56: 189–217.
Lillard, L.A., M.J. Brien, and L.J. Waite. 1995. "Premarital Cohabitation and Subsequent Marital Dissolution: A Matter of Self-Selection?" *Demography* 32:437–57.
Lillard, L.A. and L.J. Waite. 1995. "'Til Death Do Us Part: Marital Disruption and Mortality." *American Journal of Sociology* 100: 1131–56.
Lillard, L.A. and Y. Weiss. Forthcoming. "Uncertain Health and Survival: Effects on End of Life Consumption." *Journal of Business and Economic Statistics.*
Livi-Bacci, M. 1984. "Selectivity of Marriage and Mortality: Notes for Future Research." Pp. 99–108 in *Population and Biology*, edited by N. Keyfitz. Liege, Belgium: International Union for the Scientific Study of Population.
Mastekaasa, A. 1992. "Marriage and Psychological Well-Being: Some Evidence on Selection into Marriage." *Journal of Marriage and the Family* 54:901–11.
Mortensen, D.T. 1988. "Matching: Finding a Partner for Life or Otherwise." *American Journal of Sociology* 94:S215–40.
Naylor, J.C. and A.F.M. Smith. 1982. "Applications of a Method for the Efficient Computation of Posterior Distributions." *Applied Statistics* 31:214–25.
Panis, C.W.A. 1992. "A Simultaneous Treatment of Fertility Timing, Child Mortality and Health Inputs in Malaysia." Doctoral dissertation, Department of Economics, University of Southern California.
Panis, C.W.A. and L.A. Lillard. 1994. "Health Inputs and Child Mortality: Malaysia." *Journal of Health Economics* 13:455–89.
Pearlin, L. and J. Johnson. 1977. "Marital Status, Life Strains, and Depression." *American Sociological Review* 42:704–15.
Preston, S.H. and P. Taubman 1994. "Socioeconomic Differences in Adult Mortality and Health Status." *Demography of Aging*, edited by L. Martin and S. Preston. Washington, DC: National Research Council.
Rahman, O. 1993. "Excess Mortality for the Non-Married in Rural Bangladesh." *International Journal of Epidemiology* 22:445–56.
Sheps, M.C. 1961. "Marriage and Mortality." *American Journal of Public Health* 51:547–55.
Strauss, J., P. Gertler, O. Rahman, and K. Fox. 1993. "Gender and Life-Cycle Differentials in the Patterns and Determinants of Adult Health." *Journal of Human Resources* 28:791–837.
Umberson, D. 1987. "Family Status and Health Behaviors: Social Control as a Dimension of Social Integration." *Journal of Health and Social Behavior* 28:306–19.
Waldron, I., M.E. Hughes, and T.L. Brooks. Forthcoming. "Marriage Protection and Marriage Selection: Prospective Evidence for Reciprocal Effects of Marital Status and Health." *Social Science and Medicine.*
Yaari, M.E. 1965. "Uncertain Lifetime, Life Insurance, and the Theory of the Consumer." *Review of Economic Studies* 32:137–50.
Zick, C.D. and K.R. Smith. 1991. "Marital Transitions, Poverty, and Gender Differences in Mortality." *Journal of Marriage and the Family* 53:327–36.