Social Science Research 39 (2010) 801–813

Contents lists available at ScienceDirect

Social Science Research

journal homepage: www.elsevier.com/locate/ssresearch

Do wives' work hours hurt husbands' health? Reassessing the care work deficit thesis ☆

Kristen W. Springer *

Rutgers University, NJ, USA
Robert Wood Johnson Health & Society Scholars Program at Columbia University, NY, USA

ARTICLE INFO

Article history:
Received 19 August 2009
Available online 10 April 2010

Keywords:
Marriage
Health
Wives' work
Employment
Care work

ABSTRACT

Prior research suggests that wives' full-time employment harms husbands' health because employed wives have less time to promote their husbands' salubrious behavior ("care work deficit thesis" (CWDT)). In this manuscript, I analyze couple-level, longitudinal Health and Retirement Study data to assess whether evidence for the CWDT is robust to an array of sensitivity tests and correction of limitations from prior research. Specifically, I account for methodological/measurement limitations (i.e. proxy reporting), causal ordering/selection concerns (i.e. retirement and reverse causation), and conceptual issues (i.e. wives' income vs. wives' hours affecting husbands' health). The results provide strong and repeated evidence that prior support for the CWDT is due to conceptual and model misspecification. In other words, the results indicate that wives' full-time work hours do *not* harm husbands' health. However, further analyses suggest that wives' and husbands' income may differently affect husbands' health, underscoring the need for gendered analyses of income and health within marriage.

1. Introduction

As more married women participate in the paid labor market, there is growing interest in understanding how women's labor force participation shapes other social institutions – with particular interest in families. One aspect of this growing interest focuses on whether and how wives' paid employment affects husbands' health (McDonough et al., 1999; Ribet et al., 2001; Stolzenberg, 2001; Strogatz et al., 1998; Suarez and Barrett-Connor, 1984). A prominent explanation is that wives' full-time work hurts husbands' health because working wives have less time to manage husbands' health. Importantly, as previously articulated, this explanation (hereafter called the care work deficit thesis (CWDT)) is specifically gendered with wives' work harming husbands' health – but not the reverse (husbands' work does not harm wives' health). The argument for this asymmetry is that women are socialized to be homemakers and caretakers, thus "the average wife [is] far better prepared to manage her own health *and* her husband's health" (Stolzenberg, 2001, p. 66). The CWDT may make intuitive sense and resonates with the 1950s' historical ideal of men as breadwinners and women as caregivers (Becker, 1981). However, while vestiges of this "old-time" 1950s family dynamics persist today, assuming women's superior caretaking skills is unfounded for contemporary families (Coontz, 2005). Given the questionable assumptions underlying the CWDT, and the methodological and conceptual limitations of prior research on the CWDT, it is crucial to determine

☆ I would like to thank Molly Carnes, Deborah Carr, Myra Marx Ferree, Jeremy Freese, Robert M. Hauser, Dawne Mouzon and Elizabeth Thomson for their insightful comments. The research reported herein was supported in part by the Robert Wood Johnson Health & Society Scholars Program.

* Address: Department of Sociology, Rutgers University, 54 Joyce Kilmer Avenue, Piscataway, NJ 08854, USA.
E-mail address: kspringer@sociology.rutgers.edu

0049-089X/$ - see front matter © 2010 Elsevier Inc. All rights reserved.
doi:10.1016/j.ssresearch.2010.04.008

whether the proposed relationship between wives' hours worked and husbands' health is robust to sensitivity tests and corrections of limitations from prior research.

The care work deficit thesis was most notably articulated and assessed in Stolzenberg's (2001) influential article in the *American Journal of Sociology* (AJS) titled "It's about Time and Gender: Spousal Employment and Health." Prior to conducting any sensitivity analyses, I replicate Stolzenberg's (2001) core CWDT findings using a dataset (Health and Retirement Study) well designed to examine the effect of wives' hours worked on husbands' health. This initial replication allows me to be confident that any evidence from sensitivity analyses refuting the CWDT is not due to differences in the samples.

After replicating Stolzenberg's (2001) findings, I conduct 10 sensitivity tests (three methodological/measurement improvements; three causal ordering/selection enhancements; and four conceptual extensions) to assess the robustness of the CWDT findings. If wives' hours worked continue to negatively predict husbands' health through these analyses, this provides evidence that wives' paid employment negatively affects husbands' health. Furthermore, if results from the more rigorous models explored here mirror results from less-sophisticated models examined in previous research, this indicates that simpler models are sufficient to make accurate statements about employment and spousal health. In contrast, if wives' hours fail to predict husbands' health when more appropriate and rigorous models are used, these finding will suggest that wives' hours worked do not harm husbands' health. Finally, I move beyond directly examining the care work deficit thesis to theorize and test possible cross-spouse effects of income and wages on health outcomes. Taken together, the results in this project can help future researchers identify pathways connecting labor market activities with spousal health.

2. Wives' socioeconomic status and husbands' health

Early research exploring the effect of women's economic status on husbands' health focused on women's education – or as a 1984 article title suggests: "Is an educated wife hazardous to your health?" (Suarez and Barrett-Connor, 1984). Contemporary research on spousal health effects of women's social status has shifted to women's employment. Two of the most notable and well-designed studies of wives' work are by McDonough et al. (1999) and Stolzenberg (2001). McDonough et al. (1999) used Panel Study of Income Dynamics (PSID) data to examine the effect of income on spousal mortality risk controlling for education, age, and work disability – but not controlling for hours worked. McDonough et al. (1999) found that a man's higher income (five year annual average earnings) *negatively* predicted his wife's risk of dying, but a woman's income *positively* predicted her husband's risk of dying. McDonough et al. (1999) offered post-hoc explanations for the results, including the CWDT. Specifically, McDonough speculated that the negative relationship between wives' income and husbands' mortality risk could be because wives with higher income work more hours and therefore do not have time to care for their husbands' health.

Following McDonough's work, Stolzenberg (2001) argued that it is essential to disaggregate wives' earnings and wives' hours worked in order to understand how wives' employment affects husbands' health. Specifically, Stolzenberg (2001) used Americans' Changing Lives (ACL) data to assess the effect of wives' work hours on husbands' self-reported health – taking into account education, one measure of prior health status for each spouse, and family income. In these models, a wife's work hours – when they exceeded 40 h per week – were inversely related to her husband's general health controlling for *family* income. The primary argument offered was what I call the care work deficit thesis – when women work long hours outside of the home this limits their ability to manage their husbands' health. However, these analyses suffer from an array of limitations which call into question the validity of the CWDT. These limitations/issues can be categorized into three broad groups: methodological/measurement, selection/causal ordering, and conceptual extensions/refinements.

2.1. Methodological/measurement issues

The first methodological/measurement limitation is that prior research has not tested whether the adverse effect of wives' hours worked on husbands' health is significantly different from the effect of husbands' hours worked on wives' health. This comparison is fundamental because the CWDT is about the adverse effect of women's (but not men's) paid employment. Specifically, the CWDT – as previously specified – is not supported if husbands' hours worked also negatively affects wives' health and/or if the negative cross-spouse effect of hours to health is not significantly different for men and women. Rather, these findings would point to the importance of spousal work in general, instead of suggesting women are particularly well-suited for care giving and that men's health is particularly affected by wives' work hours.

Second, all ACL data are reported directly by the respondent, not the spouse – including measures of the spouse's income, work, and health. Therefore, it is plausible that characteristics of the respondent could bias his or her reports of spousal measures. This bias could mean that previous findings of the association between wives' work hours and husbands' health are due to respondent characteristics that affect reporting – rather than a real relationship between work and health. For example, when including proxy reports by husbands it is not possible to know whether a negative relationship between wives' work hours and husbands' health is from diminished care giving or rather from sicker husbands over-reporting their wives' work hours – perhaps because they are more cognizant of their wives time not at home. Alternatively, a negative relationship between wives' work hours and husbands' health could be because women who work full time report poorer health for their husbands – perhaps because these women feel responsible for their husbands' poorer health given the ideological construction of wives as caretakers and/or because men with full-time working wives complain more about their health. Given the

likelihood of proxy report bias – regardless of the psychosocial motivation – it is crucial to determine whether CWDT results hold when data are directly reported by both the husband and wife.

2.2. Selection/causal ordering issues

Selection and causal ordering concerns are two major issues in any non-experimental research design. In assessing the evidence of the CWDT, one primary concern is that the causal ordering laid out by the CWDT – that wives' hours worked adversely affects husband's health – is backwards. In other words, it is possible that husbands' poorer health *causes* wives' increased work hours. This reverse scenario is plausible and has been documented in prior research (Smith, 1999). Specifically, in anticipation of, or in response to, husbands' poor health, wives increase their labor market commitment or delay retirement in order to provide more income to the family (O'Rand and Farkas, 2002). In order to best assure that evidence for the CWDT is not being driven by reverse causation it is essential to account for retirement and baseline health status.

Husbands' early retirement is a primary selection and causal ordering concern for assessing the validity of the CWDT – at least for midlife and older couples. For example, when men retire (or plan to retire) due to sickness or disability, their wives may increase their labor market participation to help offset the costs of their husbands' retirement *or* wives may decrease their work hours (or retire) to spend more time at home with their ailing spouse (McGarry, 2004; Siegel, 2006). Without accounting for retirement, it is impossible to know whether a relationship between wives' hours worked and husbands' health is due to wives increasing their work hours in response to their husbands' health-driven retirement decisions rather than wives' work hours adversely affecting husbands' health (as the CWDT suggests). Prior analyses of the CWDT were of a younger sample (~45 years old on average) than analyzed here (~55 years old on average) – so accounting for men's premature retirement may be more important when analyzing HRS data vs. ACL data.

In addition to examining the effect of husbands' retirement, it is also crucial to adequately control for baseline health status. Without rigorously accounting for prior health, it is not possible to know whether wives' employment affects husbands' health, or conversely, whether husbands' poor health leads to increases in wives' work hours. Stolzenberg (2001) included a baseline measure of general self-reported health for both spouses to help account for this possibility of reverse causation. However, it would be wise to assess whether evidence for the CWDT still holds with more extensive controls for baseline health.

2.3. Conceptual extensions/refinements

One fundamental test of the CWDT is also an opportunity for conceptual refinement of the relationship between wives' work and husbands' health. As previously indicated, Stolzenberg (2001) argued that wives' hours worked, not income, matter for husbands' health because wives' hours worked remained significant even controlling for *family* income. However, to truly test the effect of wives' income compared to wives' hours, it is necessary to disaggregate family income into husbands' income and wives' income. Indeed, McDonough's (1999) research indicates that husbands' income and wives' income have opposite effects on husbands' health. By examining family income rather than individual income it is impossible to say whether wives' hours worked, wives' income, neither, or both affect husbands' health. It is therefore necessary to simultaneously assess the effect of wives' income as well as wives' hours worked. Evidence that wives' hours worked does not adversely affect husbands' health when including wives' income provides strong evidence against the CWDT. However, if wives' income affects husbands' health this could suggest that something about wives' money – rather than wives' hours worked – is threatening for husbands' health.

It is also important to examine different ways of understanding and measuring the monetary aspects of wives' work. One issue is whether utilizing a natural log of income – rather than a direct measure of dollars – affects any of the results. Using a natural log accounts for the well-known skew of income variables and may be more meaningful because the coefficient of the natural log of income can be interpreted as a percentage increase affecting a change in the dependent variable. With income, it is usually the case that a percent change is substantively more relevant than a change in absolute dollars. Examining both a natural log of income as well as income measured in dollars will provide another test of past evidence for CWDT, because Stolzenberg (2001) used a non-logged measure of family income. In addition, assessing both forms of income can help illuminate whether actual or percent dollar changes in income affect husbands' health.

Finally, it is useful to assess whether wives' hours worked continues to affect husband's health with the inclusion of individual wage rates rather than individual income measures. Income is a measure that confounds labor supply and wage rate. However, a measure of wage rate captures only the effect of remuneration. Therefore, assessing wage rate and hours is another way to test the separate effect of employment time vs. remuneration on health.

3. Current project

The primary substantive goal of this manuscript is to assess the claim that wives' full-time work harms husbands' health. This substantive goal is achieved through methodological investigations of whether prior evidence supporting the CWDT is robust to an array of sensitivity tests. In order to best assess prior evidence, I first replicate Stolzenberg's (2001) CWDT analyses and findings. I structure the replication analyses – and all subsequent analyses – to be as comparable to Stozlenberg's

(2001) as possible. Specifically, all models include husband and wife measures of age, education, race, self-reported health at baseline, and hours worked (full-time and part-time vs. none). This comparability helps assure that any findings discounting the CWDT are not due to differences in the sample, data, or structure of HRS vs. ACL. After replicating these prior CWDT findings, I address the previously discussed array of limitations to examine the robustness of the CWDT. Specifically, I analyze couple-level data from the HRS to systematically and independently examine each of the ten possible issues detailed in Fig. 1.

4. Data

The HRS is a nationally representative sample of Americans born between 1931 and 1941, as well as their spouses. Data were first collected in 1992 and there have been subsequent data collection rounds every two years up to 2004. In these analyses I focus on the time period 1992–1994 for two reasons. First, it most closely matches, in historical time and duration, Stolzenberg's (2001) three year window between 1986 and 1989, thereby enabling comparability. Second, the HRS respondents are at an age where retirement is quickly approaching – or has already occurred. This earliest time period captures the largest sample of employed midlife couples because fewer people have retired during this time period than any other. Alternative timeframes – such as examining couples during the four year period from 1992 to 1996 – would result in a sample that is more affected by retirement attrition, as well as a time period that is less comparable with prior research (Stolzenberg, 2001).

The first round of HRS data was collected by face-to-face, in-home interviews with each subsequent follow-up round of data collection conducted by phone. At each round, both spouses were asked detailed questions about their family life, occupation, and health status. If one person was unable or unwilling to respond, HRS collected data from a proxy respondent (usually the spouse), allowing me to assess the effect of proxy reporting. Some data, including household level data and income, were collected from one person.

Issue	Reason for Importance	Stolzenberg/ACL	Springer/HRS
Methods/Measurement (Table 2)			
W hours → H health vs. H hours →W health	CWDT predicts that wives' hours harm husbands' health – but not the reverse	No statistical test for gender difference	Statistical test for gender difference
Proxy reports in 1992	Bias in reporting hours	Spousal data all proxy	Restrict to self-report only
Proxy reports in 1994	Bias in reporting health	Spousal data all proxy	Restrict to self-report only
Selection/Causal Ordering (Table 3)			
Husbands' retirement in 1992	Risk of early retirement due to poor health leading to CWDT finding	No controls…but younger sample	Restrict to not completely retired in 1992
Husbands' retirement in 1994	Risk of early retirement due to poor health leading to CWDT finding	No controls…but younger sample	Restrict to not completely retired in 1994
Baseline health	Wives increase hours in response to, or in anticipation of, sick husbands	One measure of self-reported health for each spouse	H & W measures of: self-reported health, BMI, ADLs, smoking, health problems that limit work
Conceptual Extensions/Refinements (Table 4)			
Functional form of family income	Reduce skew of income variable; focus on percent change effects	Actual dollars	Natural log of couple income
Individual vs. family income	Assess whether wives' income and/or hours affect husbands' health	Family income	Individual income measures
Functional form of individual income	Reduce skew of income variable; focus on percent change effects	NA – no individual income measures	Natural log of individual income
Individual wages vs. individual income	Further disaggregate hours worked from money	NA – no individual wage measures	Natural log of wage rate

Fig. 1. Care work deficit thesis limitations/extensions. *Note:* This figure is organized by the three broad categories described in the text and includes references to the tables presenting empirical results.

I limit my sample to couples were who were married in 1992 and 1994, and who have complete data on variables of interest. For all analyses, I create a "constructed pair" dataset where all men (either respondents or spouses) are labeled as husbands and all women (either respondents or spouses) are labeled as wives.[1]

4.1. Dependent variable

The dependent variable, self-rated health in 1994, is dichotomized with excellent and very good health given a value of 1 and fair, poor, very poor health given a value of 0.[2] Stolzenberg (2001) used self-reported health and therefore it is the appropriate outcome to ensure comparability. Further, self-reported health has been shown to be predictive of mortality in an array of samples, across different countries, and measured in a variety of ways (Idler and Benyamini, 1997). In addition, subjective assessments of health have been shown to be valid and more predictive of mortality and disability than even physician-based evaluations (Ferraro and Farmer, 1999; Ferraro and Su, 2000).

4.2. Independent variables

All independent variables were measured in 1992, unless otherwise noted.

Employment measures. Employment is measured in time and money. For time, I combine measures of usual hours worked per week at a main job and a second job. I create three indicators of hours worked. Specifically, full-time work is more than 40 h per week, part-time work is 40 h or less per week, and no work is zero hours.[3]

Monetary compensation for employment is measured both by income and by wage rate. Income is reported by only one person – the best-informed financial respondent – about both spouses. The income measure includes wages, salary, bonuses, overtime pay, commissions, tips, military earnings, and professional practice or trade income in 1991. Missing values were imputed by HRS using a complex structure detailed elsewhere (St. Clair et al., 2004).

Finally, I assess the hourly wage rate as a sensitivity test. Using wage rate as a sensitivity test is conceptually important because it better disentangles money from time. With the HRS, including wage rate also provides a methodological check because HRS collects wage rate information from each spouse separately – therefore avoiding proxy report biases that may result when using income measures in the HRS. Hourly wage rate is based on the current job. For people who report currently not working, I set their hourly wage rate to zero. I imputed wage rate data for cases with missing data if they had income data from the preceding year, as well as complete data on hours and weeks they usually worked.[4]

Non-health controls. Age was measured in years taken at the mid-point of the interview. All respondents were, by design, born between 1931 and 1941 – but spouses' age range was wider. Education was measured as a count of years, top-coded at 17 years of education. These education data were taken from the 1992 survey when available; however, if the data are missing in 1992 but complete in later rounds, this later value is used. Race of both spouses is included in the models using a dichotomous measure of white/not-white.

Health controls. In all models, I include an ordinal scale of baseline (1992) general self-reported health for both spouses to be consistent with prior research and to help account for the possibility of reverse causation (Stolzenberg, 2001).[5] In causal order tests, I include this general self-reported health as well as four other baseline physical health measures for both spouses. Specifically, I include measures for activity limitations (ADLs), work limitations, body mass index (BMI), and current smoking. For ADLs, I include a baseline measure that indicates whether the respondent reported some or more difficulty with any of the five ADLs (bathing, eating, dressing, walking across the room, or getting into or out of bed). The work limitation variable is a dichotomous indicator measuring whether the respondent reported any health problem or impairment that limited the amount or kind of work he or she could do for pay. BMI is a measure of appropriate body weight for a given body size and is measured as kilograms divided by meters squared. Smoking is a dichotomous measure of whether the respondent reported currently smoking in 1992. Despite the large number of baseline health measures there are no problems with multicollinearity in the models.

[1] Of the 5262 HRS respondents who were coupled in 1992, 4886 spouses were also interviewed in 1992. Of these 4886 pairs, 4109 pairs were alive, still married, and agreed to participate in 1994. The sample truncation from 4109 to 4068 in the replication models is due to item non-response.

[2] I am using a dichotomous rather than ordinal outcome in order to employ bivariate probit models. As will be described, these models are designed to account for correlated residuals, which is useful when conducting spousal models. Stolzenberg's (2001) used bivariate probit models when he had parallel spousal measures, but presented ordinal results when he did not have parallel spousal measures. However, he stated that all results were unchanged with ordinal and dichotomous measures. Therefore, replicating the care work deficit findings with a dichotomous outcome is appropriate.

[3] This breakdown is consistent with Stolzenberg (2001). However, the results are robust to an array of sensitivity tests including: using a more standard cut-off of 35+ hours as full-time work; using a continuous measure for hours worked; and including spousal interactions of multiple operationalizations of hours worked. Further, analyses using a larger array of dummy variables to assess nonlinearity (no work, <40 h, 40 h, >40 h and <50 h, and 50 or more hours per week) showed no significant difference in the effect of wives' working 40 h, >40 h and <50 h, and 50 or more hours per week, thereby validating the use of the "full-time, part-time, and no work" gradations used here.

[4] Including these few cases based on income does result in including some proxy reports. However, the results are not meaningfully different with or without these few cases.

[5] In addition, controlling for baseline health is standard practice in health research. Therefore, including baseline health permits scholars to more readily compare these findings with other related research.

4.3. Restricting variables

Proxy reports. I impose two restrictions on the sample to assess proxy report effects. I first restrict the sample to only self-reported cases in 1992 and then I restrict the sample to only self-reported cases in 1994.

Retirement. To assess the effect of retirement on the CWDT, I first restrict the sample to husbands who do not report being completely retired in 1992 and then to husbands who do not report being completely retired in 1994. These retirement questions were not asked of proxy respondents and therefore only self reporters could indicate they were not completely retired.

5. Statistical methods

I use bivariate probit models with maximum likelihood estimation for the analyses in this manuscript. Bivariate probit model are one type of simultaneous equation models that do not assume independence of disturbances, but rather allow for a nonzero correlation between disturbances in the equations. Statistical tests can be performed to assess whether the correlation of disturbances deviates significantly from zero. If the correlation is not significant, then running a bivariate probit model is no different than running two separate probit models.

Using simultaneous equation models is important in cases where one expects the unexplained portion of the equations to be correlated. Spouse models are a perfect example of a situation in which the residuals are likely to be correlated. Specifically, it is reasonable to assume that some of the unexplained variance in the health of husbands and wives has common causes not captured by the independent variables. Some examples might be shared ideological beliefs about visiting health care professionals or environmental toxins near the couples' home.

6. Results

6.1. Descriptive statistics

Table 1 presents descriptive statistics for the husband/wife constructed HRS sample. The average age of husbands is 57 years old and the average age of wives is slightly younger, at 53 years old. Both spouses have 12 years of education on average and approximately 85 percent of husbands and wives are white. In 1992, 38 percent of husbands worked more than 40 h per week compared to 15 percent of wives. Approximately 36 percent of husbands worked 40 h or less per week compared to 47 percent of wives. Mean income earned by husbands was $28,582 (median = $23,300) compared to just under $12,500 (median = $7000) for wives. In addition, husbands had a mean hourly wage rate of $13 (median = $10) compared to $7 (median = $5) for wives.

Around fifty percent of both husbands and wives reported excellent or very good general health in 1992 and 1994, with more wives than husbands reporting excellent or very good health at both time points, and with more husbands and wives reporting excellent or very good health in 1992 compared to 1994. On average, both husbands and wives were slightly overweight with a BMI of 27 (normal BMI = 18.5–24.9). Less than four percent of husbands and wives reported having at least some difficulty with any activity of daily living. In addition, 26 percent of husbands and 23 percent of wives currently smoked tobacco. Approximately 20 percent of husbands and 17 percent of wives reported having a health problem that limited the kind or amount of work they could do.

In terms of the restricting variables, over 8 percent of the husbands' reports were provided by a proxy in 1992 compared to only 2 percent of wives' reports. These numbers increased in 1994 with 12 percent of husbands and 2.5 percent of wives having a proxy reporter. In 1992, over 20 percent of the husbands reported being completely retired and in 1994, 27 percent of the husbands reported being completely retired. Wives' retirement is not included in the table because it is not a concern for this project. Mean total household income was $54,000 (median = $44,300).

6.2. Replication model

The first model was designed to replicate prior research and indeed I did find support for the CWDT, in that wives' full-time hours negatively predicted husbands' self-reported general health (see replication model in Column 1 of Table 2). The results also indicate that there is a significant correlation between spouse residuals (chi-square of rho (9.596) is significant at $p < 0.01$).

A wife's health was influenced by several of her own and her spouse's characteristics. Specifically, both wives' and husbands' education and prior health status positively affected wives' health. Being white also predicted better health for wives. A husband's health was affected by his own education and by both his own and his wife's prior health. In addition, a husband's full-time work positively affected his health, but did not significantly influence his wife's health. Taken together, these

Table 1
Descriptive statistics ($n = 4068$).

	Husband	Wife
Age	57.16	53.21
	(5.31)	(5.61)
Education	12.22	12.22
	(3.41)	(2.83)
White	85.10	85.37
	–	–
Full-time work	38.15	14.63
	–	–
Part-time work	35.77	47.07
	–	–
Annual earnings	28,581.6	12,499.17
	(35,302.31)	(16,864.74)
Wage rate	12.72	7.04
	(19.21)	(13.20)
Excellent/very good health (1994)	48.89	53.17
	–	–
Excellent/very good health (1992)	51.45	56.05
	–	–
BMI	27.32	26.70
	(4.10)	(5.49)
Any ADL	3.79	3.20
	–	–
Current smoker	25.69	22.79
	–	–
Work limitation	19.54	16.99
	–	–
Proxy in 1992	8.14	2.11
	–	–
Proxy in 1994	12.24	2.53
	–	–
Retired 1992	20.45	NA
	–	
Retired 1994	27.04	NA
	–	
Couple-level variable		
Household income	54,181.24	
	(48,353.53)	

Variables are measured in 1992 unless noted. Means are presented for ordinal/continuous variables with standard deviations in parentheses. Percentages are shown for dichotomous variables. Wage rate has fewer cases ($n = 4010$). Retirement measures have fewer cases because they were not asked of proxy reporters.

findings replicate previous support for the CWDT because wives' full-time work has a significant adverse effect on husbands' health, but husbands' full-time work does not significantly affect wives' health.

6.3. Methods/measurement issues

The first sensitivity check is to statistically test whether the effect of wives' full-time work on husbands' health is significantly different than the effect of husbands' full-time work on wives' health in the replication model (Column 1 of Table 2). I find that the effect of wives' full-time work on husband's health is *not* significantly different than the effect of husbands' full-time work on wives' health ($p = 0.233$). These results provide the first blow against the CWDT as specifically gendered, because there is no significant distinction between how wives' paid work hours affect husbands' health and how husbands' paid work hours affect wives' health.

When restricting the sample to only cases that were self-reported in 1992, the effect of wives' full-time work on husbands' health is still negative and significant (Column 2 in Table 2), despite a reduction in sample size. However, this effect is still not significantly different than the effect of husbands' full-time work on wives' health. When excluding proxy reports in 1994, the effect of wives' full-time work on husbands' health weakens and becomes non-significant (Column 3 of Table 2). This could result, for example, from differential reporting of wives' hours by husbands, dependent on husbands' health status. Sicker husbands might over-report their wives' work hours because they are more cognizant of their wives' time not at home, thereby creating an association between wives' hours worked and husbands' health that is driven by proxy reports. Removing these proxy reports reduces the effect of wives' hours to non-significance. However, it is important to note that the reduced sample size (and correspondingly increased standard error) might have contributed to the reduction in significance.

Table 2
Bivariate probit results for excellent/very good health: replication model and method/measurement tests.

	1. Replication		2. No proxy in 1992		3. No proxy in 1994	
	H	W	H	W	H	W
Wife full-time	−0.15*	0.14+	−0.17*	0.13	−0.12	0.13
	(0.07)	(0.07)	(0.08)	(0.08)	(0.08)	(0.08)
Wife part-time	−0.03	0.06	−0.04	0.05	−0.03	0.03
	(0.05)	(0.05)	(0.06)	(0.05)	(0.06)	(0.06)
Husband full-time	0.22**	−0.03	0.23**	−0.01	0.21**	−0.03
	(0.07)	(0.07)	(0.07)	(0.07)	(0.07)	(0.07)
Husband part-time	0.12+	−0.06	0.15*	−0.03	0.14+	−0.05
	(0.06)	(0.06)	(0.07)	(0.07)	(0.07)	(0.07)
Household income (×1000)	0.001+	0.001+	0.001	0.001+	0.001+	0.001+
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Wife age	0.00	0.00	0.00	0.00	0.01	0.00
	(0.00)	(0.00)	(0.01)	(0.01)	(0.01)	(0.01)
Husband age	−0.01	0.00	0.00	0.00	−0.01	0.00
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Wife education	0.01	0.03**	0.01	0.03*	0.02	0.03**
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Husband education	0.05**	0.03**	0.05**	0.03**	0.05**	0.03**
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Husband health	0.71**	0.07**	0.73**	0.05*	0.74**	0.07**
	(0.02)	(0.02)	(0.03)	(0.02)	(0.03)	(0.02)
Wife health	0.06**	0.76**	0.06*	0.77**	0.05+	0.77**
	(0.02)	(0.03)	(0.02)	(0.03)	(0.02)	(0.03)
Wife white	−0.01	0.33*	0.01	0.33*	−0.11	0.22
	(0.14)	(0.14)	(0.01)	(0.15)	(0.16)	(0.15)
Husband white	0.09	−0.16	0.10	−0.17	0.18	−0.08
	(0.14)	(0.14)	(0.15)	(0.15)	(0.16)	(0.15)
Constant	−3.65**	−3.81**	−3.81*	−3.83**	−3.76*	−3.84**
	(0.35)	(0.35)	(0.37)	(0.36)	(0.38)	(0.38)
N	4068		3651		3471	
ρ	0.10		0.09		0.08	
$\rho(x^2)$	9.60**		6.90**		5.34*	
x^2	2323.57		2100.2		1999.85	
Log likelihood	−3977.22		−3555.75		−3386.68	

Standard errors are in parentheses; the coefficient of household income is multiplied by 1000 to see the coefficient magnitude; H at top of column indicates the model predicts husbands' health and W indicates the model predicts wives' health. Column 1 is the replication model, Column 2 restricts the data to self-reports in 1992, and Column 3 restricts to self-reports in 1994.
+ $p < 0.1$.
* $p < 0.05$.
** $p < 0.01$.

6.4. Selection/causal ordering issues

Turning to selection and causal ordering sensitivity tests, I first explore the effect of retirement. Table 3 contains these results in Columns 2 and 3, along with the replication results for comparison presented in Column 1.[6] Specifically, the model presented in Column 2 excludes husbands who reported being completely retired in 1992. By excluding the men who were retired in 1992, I have removed individuals who may have left the workforce early due to sickness. In this model the effect of wives' full-time work on husbands' health is not significant, suggesting that that retirement in 1992 has an important confounding effect. Column 3 in Table 3 controls for husbands' retirement in 1994. This sample therefore only has people who did not report being completely retired in 1994. As expected given the findings from Column 2, the effect of wives' full-time work on husbands' health is not significant and has a much smaller coefficient than in previous models. These retirement selection tests suggest that the evidence for the CWDT could have been due to the fact that wives work full time in order to compensate for their husbands' early retirement – possibly due to sickness. It is also possible that the reduced sample size is part of the reason for the non-significant effect of wives' hours worked on husbands' health.

After exploring the possibility of selection bias due to retirement, I include an array of initial health conditions (in addition to the basic 1992 general health measure included in all models) to reduce the possibility of reverse causation. Table 3 contains these results in Column 4. As expected, these baseline health measures are negatively related to excellent/very good health in 1994. However, the effect of women's full-time work on husbands' health diminishes and becomes non-significant when including an array of baseline health measures, suggesting that at least part of the original negative effect of wives'

[6] Tables 3 and 4 do not include the coefficients for the control variables to help streamline the results, and because the coefficients for controls are not meaningfully different from those presented in Table 2.

Table 3
Bivariate probit results for excellent/very good health: selection and causal ordering tests.

	1. Replication		2. Not retired in 1992		3. Not retired in 1994		4. Baseline health	
	H	W	H	W	H	W	H	W
Wife full-time	−0.15*	0.14+	−0.12	0.15+	−0.12	0.15+	−0.12	0.10
	(0.07)	(0.07)	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)
Wife part-time	−0.03	0.06	−0.01	0.09	−0.02	0.10+	0.00	0.02
	(0.05)	(0.05)	(0.08)	(0.06)	(0.06)	(0.06)	(0.05)	(0.05)
Husband full-time	0.22**	−0.03	0.38**	0.09	0.29**	0.02	0.10	−0.05
	(0.07)	(0.07)	(0.10)	(0.10)	(0.09)	(0.09)	(0.07)	(0.07)
Husband part-time	0.12+	−0.06	0.28**	0.04	0.22*	0.02	0.02	−0.07
	(0.06)	(0.06)	(0.10)	(0.09)	(0.09)	(0.09)	(0.07)	(0.07)
Household income (×1000)	0.001+	0.001+	0.001+	0.001	0.001+	0.001	0.001	0.001
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Husband health	0.71**	0.07**	0.69**	0.06*	0.69**	0.05+	0.66**	0.05*
	(0.02)	(0.02)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.02)
Wife health	0.06**	0.76**	0.04	0.74**	0.04	0.75**	0.07**	0.71**
	(0.02)	(0.03)	(0.02)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
Husband work limitation	–	–	–	–	–	–	−0.38**	0.00
	–	–	–	–	–	–	(0.07)	(0.07)
Wife work limitation	–	–	–	–	–	–	0.11	−0.37*
	–	–	–	–	–	–	(0.07)	(0.07)
Husband smoke	–	–	–	–	–	–	−0.16**	0.00
	–	–	–	–	–	–	(0.06)	(0.06)
Wife smoke	–	–	–	–	–	–	−0.01	−0.06
	–	–	–	–	–	–	(0.06)	(0.06)
Husband BMI	–	–	–	–	–	–	−0.01+	0.00
							(0.01)	(0.01)
Wife BMI	–	–	–	–	–	–	−0.01**	−0.03**
	–	–	–	–	–	–	(0.00)	(0.00)
Husband any ADL	–	–	–	–	–	–	−0.39**	−0.27+
	–	–	–	–	–	–	(0.19)	(0.14)
Wife any ADL	–	–	–	–	–	–	0.17	0.06
	–	–	–	–	–	–	(0.14)	(0.19)
Constant	−3.65**	−3.81**	−3.50**	−3.52**	−3.42**	−3.58**	−2.48**	−2.61**
	(0.35)	(0.35)	(0.38)	(0.39)	(0.40)	(0.41)	(0.44)	(0.44)
N	4068		3302		3102		4068	
ρ	0.10		0.13		0.12		0.10	
$\rho(x^2)$	9.60**		12.91**		10.95**		8.60**	
	2323.57		1811.65		1712.73		2320.96	
Log likelihood	−3977.22		−3297.53		−3080.37		−3912.91	

Standard errors are in parentheses; the coefficient of household income is multiplied by 1000 to see the coefficient magnitude; all models control for husbands' and wives' race, age, and education in 1992. H at top of column indicates that the model predicts husbands' health and W indicates the model predicts wives' health; Column 1 is the replication model; Column 2 restricts the sample to husbands not completely retired in 1992; Column 3 restricts the sample to husbands not completely retired in 1994; Column 4 includes multiple baseline health variables to test for reverse causation.

+ $p < 0.1$.

* $p < 0.05$.

** $p < 0.01$.

hours worked was due to women working more when their husbands are not healthy. Importantly, the sample is exactly the same as in the replication model indicating that the attenuation of wives' hours worked is due only to the inclusion of the baseline health measures.

6.5. Conceptual extensions/refinements

I next examine how the functional form of income affects health outcomes. Column 1 in Table 4 is again the replication model. Column 2 in Table 4 displays results of the baseline replication model that includes a natural log of household income. This procedure does not erase the effect of wives' hours on husbands' health; rather, it slightly increases the coefficient, suggesting that the functional form of household income is not the root cause of prior findings. However, statistical tests again demonstrate that the effect of wives' full-time work on husbands' health does not differ significantly from the effect of husbands' full-time work on wives' health.

The next income test is to assess the effect of including separate measures of income for husbands and wives, rather than a combined household income variable. As discussed previously, to rigorously assess the effect of specific aspects of employment on spousal health, it is important to include hours and income for both spouses. In addition, examining the influence of individual income measures can illuminate whether the source of income matters for health outcomes, or whether all marital dollars affect health in the same way.

Table 4
Bivariate probit results for excellent/very good health extensions of the carework deficit thesis.

	1. Replication		2. Log family income		3. Individual income		4. Log individual income		5. Log individual wages	
	H	W	H	W	H	W	H	W	H	W
Wife full-time	−0.15*	0.14+	−0.16*	0.12+	−0.09	0.10	−0.02	0.07	−0.06	0.08
	(0.07)	(0.07)	(0.07)	(0.07)	(0.08)	(0.08)	(0.09)	(0.09)	(0.13)	(0.13)
Wife part-time	−0.03	0.06	−0.04	0.05	0.00	0.03	0.07	0.00	0.03	0.01
	(0.05)	(0.05)	(0.05)	(0.05)	(0.06)	(0.06)	(0.07)	(0.07)	(0.12)	(0.12)
Husband full-time	0.22**	−0.03	0.21**	−0.04	0.19	−0.03	0.14+	−0.03	−0.20+	−0.14
	(0.07)	(0.07)	(0.07)	(0.07)	(0.07)	(0.07)	(0.08)	(0.08)	(0.12)	(0.12)
Husband part-time	0.12+	−0.06	0.12+	−0.07	0.10	−0.07	0.05	−0.07	−0.31**	−0.18
	(0.06)	(0.06)	(0.07)	(0.06)	(0.07)	(0.07)	(0.07)	(0.07)	(0.12)	(0.12)
Husband health	0.71**	0.07**	0.71**	0.07**	0.71**	0.07**	0.71**	0.07**	0.71**	0.07**
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Wife health	0.06**	0.76**	0.06*	0.76**	0.06**	0.76**	0.06**	0.76**	0.06**	0.77**
	(0.02)	(0.03)	(0.02)	(0.03)	(0.02)	(0.03)	(0.02)	(0.03)	(0.02)	(0.03)
Household income (×1000)	0.001+	0.001+	–	–	–	–	–	–	–	–
	(0.00)	(0.00)	–	–	–	–	–	–	–	–
Log household income	–	–	0.07	0.09**	–	–	–	–	–	–
	–	–	(0.04)	(0.04)	–	–	–	–	–	–
Husband earnings (×1000)	–	–	–	–	0.002*	0.001	–	–	–	–
	–	–	–	–	(0.00)	(0.00)	–	–	–	–
Wife earnings (×1000)	–	–	–	–	−0.001	0.003+	–	–	–	–
	–	–	–	–	(0.00)	(0.00)	–	–	–	–
Log husband income	–	–	–	–	–	–	0.051*	0.012	–	–
	–	–	–	–	–	–	(0.02)	(0.02)	–	–
Log wife income	–	–	–	–	–	–	−0.045+	0.033	–	–
	–	–	–	–	–	–	(0.02)	(0.02)	–	–
Log husband wage	–	–	–	–	–	–	–	–	0.17**	0.06
	–	–	–	–	–	–	–	–	(0.04)	(0.04)
Log wife wage	–	–	–	–	–	–	–	–	−0.02	0.03
	–	–	–	–	–	–	–	–	(0.05)	(0.05)
Constant	−3.65**	−3.81**	−4.31**	−4.65**	−3.72*	−3.83**	−3.88+	−4.19+	−3.69*	−3.89**
	(0.35)	(0.35)	(0.47)	(0.46)	(0.35)	(0.35)	(0.42)	(0.42)	(0.35)	(0.35)
N	4068		4068		4068		4068		4010	
ρ	0.10		0.10		0.10		0.11		0.09	
$\rho(x^2)$	9.60**		9.61**		9.83**		10.23**		7.75**	
x^2	2323.57		2323.05		2324.23		2323.38		2312.53	
Log likelihood	−3977.22		−3975.60		−3974.77		−3974.72		−3902.55	

Standard errors are in parentheses; income coefficients are multiplied by 1000 in order to see the coefficient magnitude; all models control for husbands' and wives' race, age, and in 1992. H at top of column indicates that the model predicts husbands' health and W indicates the model predicts wives' health; Column 1 is the replication model; Column 2 includes logged household income; Column 3 includes income measures for husbands and wives; Column 4 includes logged income measures for husbands and wives; Column 5 includes wage measures for husbands and wives.

+ $p < 0.1$.

* $p < 0.05$.

** $p < 0.01$.

Columns 3 and 4 in Table 4 display the models with individual income measures.[7] Column 3 presents individual income measured in dollars and Column 4 presents the natural log of individual income. The sample size is identical to the replication model and therefore any change in the effect of wives' hours worked is not due to sample issues.

As displayed in both Column 3 and Column 4, the results demonstrate that regardless of the functional form of income, including individual income measures renders the effect of women's work hours on husbands' health non-significant. For example, the coefficient for wives' full-time work on husbands' health is dramatically smaller in the model that includes log transformed individual earnings (Column 4) compared to the replication model (Column 1) with household income.[8] In addition, the effects of husbands' and wives' incomes on husbands' health are similar in size, but opposite in sign (Column 4). Husbands' earnings slightly improve their own health, whereas wives' earnings slightly decrease husbands' health. Importantly, these coefficients are significantly different from each other – indicating that the effect of wives' income on husbands' health is significantly different than the effect of husbands' income on husbands' health. This result underscores the importance of using individual measures of income, rather than collapsing both spouses' income into one household income measure.

[7] Sensitivity tests including an interaction of spousal income measures produced essentially the same results as presented in Table 4.

[8] The replication model in Table 4 is based on Stolzenberg and therefore includes household income, which does include income not just from husbands and wives (marital income). However, the results are virtually identical when rerunning the model in Column 1 to include only income from husbands and wives. This suggests that changes in the models when going from household to individual earnings is not due to the fact that household earnings contain slightly more money than looking only at marital income.

Issue	Support for the Care Work Deficit Thesis
Methods/Measurement (Table 2)	
W hours → H health vs. H hours →W health	No
Proxy reports in 1992	Yes
Proxy reports in 1994	No
Selection/Causal Ordering (Table 3)	
Husbands' retirement status in 1992	No
Husbands' retirement status in 1994	No
Baseline health	No
Conceptual Extensions/Refinements (Table 4)	
Functional form of family income	Yes
Individual vs. family income	No
Functional form of individual income	No
Individual wages vs. individual income	No

Fig. 2. Care work deficit thesis limitation/extension results.

Nonetheless, looking at model fit statistics suggests that the replication model in Column 1 fits just as well as the individual income models (Column 4, for example). This would seem to suggest that including individual income adds no explanatory power to understanding health outcomes. However, because these are bivariate probit models, the model fit statistics correspond to how well the models fit for both husbands' and wives' equations. It is possible therefore, that adding individual income measures rather than household income does indeed improve the model fit for husbands' health – but not for wives' health. To further explore this possibility, I estimated probit models corresponding to Column 1 and Column 4 for husbands' health to examine changes in model fit for just husbands' health. In these models, the coefficients were nearly identical to those in the bivariate probit model, yet model fit statistics indicated that including individual income measures, compared to including marital income, improved prediction of husbands' health.[9]

The effect of including separate measures of wage rate is similar – the influence of wives' hours on husbands' health is reduced to non-significance (Column 5 in Table 4). In addition, the effects of each spouses' wage rate on husbands' health are in opposite directions and significantly different from each other. However, the effect of wives' wage rate on husbands' health is not significant.[10] These results suggest that the CWDT is not supported regardless of how remuneration is operationalized.

7. Discussion

The primary purpose of this paper was to assess the claim that wives' full-time hours worked harmed husbands' health – presumably because wives who work full time have less time to promote their husbands' salubrious behavior (Stolzenberg, 2001). Specifically, I examined the robustness of evidence for the CWDT when subjected to an array of sensitivity tests and corrections for limitations in prior research. As summarized in Fig. 2, the CWDT was only robust to two out of ten sensitivity analyses. However, even within these two sensitivity tests that seemingly support the CWDT – the effect of wives' hours worked on husbands' health was *never* significantly different than the effect of husbands' hours worked on wives' health. In other words, the results presented here provide strong and repeated evidence against the thesis that wives' paid work adversely affects husbands' health because it limits wives' time to support husbands' health. Rather, the current findings indicate that prior support for the CWDT results from model misspecification, not sufficiently accounting for reverse causation, and conceptual limitations such as examining family income rather than individual income. The results also provide evidence that wives' income and husbands' income do not have the same effect on husbands' health. The suggestive finding that wives' income may be negatively related to husbands' health underscores the need for further gendered analyses of employment and health within marriage.

Although I began by replicating prior findings showing a negative association between wives' full-time work and husbands' health, the effect of wives' full-time work on husbands' health was never significantly different than the effect of husbands' full-time work on wives' health (Stolzenberg, 2001). This finding dealt the first blow against the CWDT. Substantively,

[9] The chi-square statistic for the model with household income (df = 13) is 1610.34 compared to a chi-square of 1616.25 (df = 14) for the model with logged individual income.

[10] Note that the sample including wages is slightly smaller.

these non-significant findings suggest that there is no gender difference in how work hours affect spousal health. This non-significance is an important contrast with the 1950s ideology of the female homemaker/caretaker that underlies the previous articulation of the CWDT. In other words, these findings suggest that if "excessive" work hours harm spousal health due to diminished care giving time – both men and women are equally affected. These results point toward the possibility of recasting the concept of the CWDT as gender-neutral, rather than specifically about women's work and husbands' health. This gender-neutral revisioning of the CWDT corresponds with the demographic reality of increasingly isolated aging couples who rely primarily on each other for support and care.

Causal ordering and selection checks also indicate that prior support for the CWDT is an artifact of methodological and conceptual limitations. For example, accounting for reverse causation by including an array of baseline health measures erased the significant effect of wives' full-time work on husbands' health. These results suggest that prior findings may be due, in part, to wives working more hours *because* their husbands are sick. This is the opposite causal order suggested by the CWDT. Substantively, these findings point to the responsive and adaptive work choices by spouses, whereby one spouse increases work hours in order to compensate for the other spouse's health problems. Viewed through this lens, working wives may be an important pathway to ensure at least a minimal standard of living for unhealthy pre-retirement age men. Wives' work may become even more important in the future as the population continues to age during a time of economic uncertainty.

The negative effect of wives' full-time work on husbands' health was most attenuated by the inclusion of individual measures of logged income. This result indicates several things about this line of work. As Stolzenberg (2001) argued, it is necessary to look at both income and hours to truly understand gendered effects of employment on spousal health. However, the results of this project also suggest that it is crucial to examine the effect of each person's income separately – rather than using a measure of household income. Indeed, wives' income had a significantly different effect on husbands' health than the effect of husbands' income on husbands' health. In other words, researchers cannot treat all money in marriage equally; rather, the source of income appears to matter – at least for husbands' health.

An important next stage of research is to directly examine the gendered effect of income by examining the influence of relative earnings on husbands' health. Prior analyses on marital outcomes such as housework suggest that marital relations and well-being are influenced by the percent contribution of wives' and husbands' earnings – rather than by separate earning amounts (Brines, 1994). The fact that wives' income and husbands' income do not have the same effect on husbands' health suggests that there is something qualitatively different about money earned by wives compared to husbands. For example, it could be that wives' increasing income contributions destabilize gendered expectations of marriage resulting in marital distress that in turn harms husbands' health. However, whether this relative income effect translates into understanding men's health remains an open question and warrants further investigation.

In addition, future research should focus on examining the processes and decisions that drive any connection between paid employment, care giving, and spousal health. For example, I do not actually measure spousal care giving. Future quantitative research could further assess the (in)validity of the CWDT by directly examining wives' time spent care giving as well as hours worked in paid employment. In addition, qualitative research could illuminate how couples balance paid employment and care giving when a spouse is ill. If these future studies provide further support against the CWDT, this will help alleviate concerns that the non-significance of wives' hours in the sensitivity analyses here is simply due to the statistical artifact of reduced samples and/or adding variables correlated with wives' work. Finally, it would be useful to examine whether husbands' and/or wives' hours worked affect health outcomes not studied here – such as chronic conditions or functional limitations.

Despite these limitations, the findings from this project have important substantive implications for research on work, gender, families, SES, and health. For example, the results suggest that to best understand the health effects of income in marriage – at least for men's health – it is essential to examine individual measures of income within marriage, rather than examining a composite household income measure. This finding that wives' income and husbands' income have different effects on husbands' health calls into question the long-standing practice in health research of treating all marital money as fungible.

The results also contribute to the growing body of literature demonstrating that the 1950s' notion of women as uniquely qualified caretakers and men as specialized workers is obsolete. Not only did the results provide no support for the idea that wives' paid hours harm husbands' health, the findings pointed to the complex interplay of tradeoffs in modern marriage and suggested that wives' longer work hours are sometimes a strategy to improve the couples' financial well-being when husbands are ill. In sum, the results of this project underscore the importance of more nuanced conceptualization and measurement of work and health in the context of midlife marriages.

References

Becker, Gary, 1981. A Treatise on the Family. Harvard University Press, Cambridge, MA.
Brines, Julie, 1994. Economic dependency, gender, and the division of labor at home. American Journal of Sociology 100 (3), 652–688.
Coontz, Stephanie, 2005. Marriage, A History: From Obedience to Intimacy or How Love Conquered Marriage. Viking, New York, NY.
Ferraro, Kenneth F., Farmer, Melissa M., 1999. Utility of health data from social surveys: Is there a gold standard for measuring morbidity? American Sociological Review 64 (2), 303–315.
Ferraro, Kenneth F., Su, Ya-ping, 2000. Physician-evaluated and self-reported morbidity for predicting disability. American Journal of Public Health 90 (1), 103–108.

Idler, Ellen L., Benyamini, Yael, 1997. Self-rated health and mortality: A review of twenty-seven community studies. Journal of Health and Social Behavior 38, 21–37.
McDonough, Peggy, Williams, David R., House, James S., Duncan, Greg J., 1999. Gender and the socioeconomic gradient in mortality. Journal of Health and Social Behavior 40, 17–31.
McGarry, Kathleen, 2004. Health and retirement: Do changes in health affect retirement expectations? Journal of Human Resources 3, 624–648.
O'Rand, Angela M., Farkas, Janice I., 2002. Couples' retirement timing the United States in the 1990s: The impact of market and family role demands on joint work exits. International Journal of Sociology 32, 11–29.
Ribet, C., Land, T., Zins, M., Bingham, A., Ferrieres, J., Arveiler, D., Amouyel, P., Bonenfant, S., Chastang, J.F., Golberg, M., Ducimetiere, P., 2001. Do cardiovascular risk factors in men depend on their spouses' occupational category? European Journal of Epidemiology 17 (4), 347–356.
Siegel, Michelle J., 2006. Measuring the effect of husband's health on wife's labor supply. Health Economics 15, 579–601.
Smith, James P., 1999. Health bodies and thick wallets: The dual relation between health and economic status. Journal of Economic Perspectives 13 (2), 145–166.
St. Clair, Patricia, Bugliari, Delia, Chien, Sandy, Haider, Steven, Hayden, Orla, Hurd, Michael, Ilchuk, Serhii, Lopez, Gabriela, Loughram, David, Panis, Constantijn, Pantoja, Philip, Reti, Monika, Zissimopoulos, Julie, 2004. RAND HRS Data Documentation, Version D. Labor & Population Program, RAND, Santa Monica, CA.
Stolzenberg, Ross M., 2001. It's about time and gender: spousal employment and health. The American Journal of Sociology 107 (1), 61–100.
Strogatz, David S., Siscovick, David S., Weiss, Noel S., Rennert, Gad, 1998. Wife's level of education and husband's risk of primary cardiac arrest. American Journal of Public Health 78 (11), 1491–1493.
Suarez, Lucina, Barrett-Connor, Elizabeth, 1984. Is an educated wife hazardous to your health? American Journal of Epidemiology 119 (2), 244–249.