[DRAFT NOTICE: this paper uses a publicly-available categorical variable to calculate distance – a key instrumental variable for the analysis. But we are requesting a restricted variable with a continuous measure of distance.]

DRAFT: Women, Work, and Caring for Aging Parents:
Modelling difference by education in the PSID

Melody Waring, A.M.
School of Social Work
University of Wisconsin-Madison
mwaring@wisc.edu

September 2018

Abstract

How does caring for an aging parent affect women's labor force outcomes? And do labor force effects vary by socioeconomic status? In this paper, I build on the existing literature by (1) using a sample of women of all ages in the 2013 Panel Study of Income Dynamics (N=5,296); (2) testing for a moderating effect by education. I also attempt a causal estimate of the effect of caregiving by instrumenting care with parental health and distance to parent. Preliminary associative results find caregivers are as likely to be employed as non-caregivers, but for lower wages. This pattern holds by education, except among caregivers over 51 with low education—who are slightly less likely to be employed, but among employed comparators work many more hours. Preliminary causal estimates suggest caregiving results in higher employment overall (even among women caring at least 10 hours/week), but lower employment for women with low education.

The largest source of long-term care services for aging Americans is unpaid care from family (CBO 2013). A conservative estimate puts the value of unpaid care in 2012 at $253 billion dollars—more than double the costs of institutional care at $134 billion (CBO 2013). Our patchwork system of long-term care in the U.S. relies on the availability of family caregivers, disproportionately women, even while Medicaid provides an entitlement to care for people who are sick enough and poor enough to qualify. What cost do family caregivers pay, in direct caregiving-related expenses and through foregone wages, missed promotions, and lost pensions? And what happens when the caregivers our system relies on are themselves too poor to afford leave from work, or lose their only source of income when caring for their parent conflicts with their employment? With this paper, I set out to examine the relationship between caregiving, labor force participation, and socioeconomic status among women.

I. BACKGROUND

Characteristics & Challenges of Caregiving

Several converging demographic shifts are creating a significant increase in demand for eldercare and aging services, at the same time that the supply of informal family care may be expected to decline. First, people are living longer, shifting the population distribution of working-age adults to seniors. Simultaneously, the first "baby boomer" will reach 85 in 2030, which, in conjunction with longer life expectancies, will lead to a massive swelling of living elders (Seltzer, Bianchi 2013; CBO 2013). Second, old models of caregiving relied on women at home, dependent on a male "breadwinner," to provide care to children and to the sick, elderly, or disabled (e.g., Fraser, Gordon 2013). But female labor-force participation has increased dramatically over the last half-century, from 35 percent of women in 1948, to a peak of 60 percent in 1990, and holding at 59 percent in 2015 (approaching the male labor-force participation of 70 percent in 2015; BLS 2016). As both dual-earner and female-headed households increase, there are fewer women available to fill time-intensive caregiving roles—though

historically rates of labor force participation have varied significantly by race (Allen 1993; Bianchi, Folbre, Wolf 2012).

Even with the swell of female labor-force participation, family members still provide the overwhelming majority of care to adults with disabilities and aging Americans (AARP 2015; CBO 2013). In 2015, more than 34 million Americans had provided unpaid care to an adult over age 50 in the previous 12 months (AARP 2015); 64 percent of family caregivers for aging adults (including spouses) were women in 2015—with no significant change in proportion since 1990 (Wolff et al 2017). The majority of unpaid family caregivers (61 percent) are employed and must balance earning income with providing unpaid labor for care (Earle & Heymann 2012).

While the Congressional Budget Office uses the median wage of paid caregivers to put the value of unpaid care in 2012 at $253 billion dollars (CBO 2013), a 2015 report from AARP used a meta-review of survey data to generate state-specific prevalence of caregiving and of average wages of caregivers, estimating that replacing unpaid family care with wages would cost $470 billion (Reinhard, Feinberg, Choula, Houser 2015). Though both of these estimates dwarf formal spending on long-term care in the U.S., neither attempts to estimate the individual costs borne by caregivers who forego wages, pass up promotions, take unpaid leave, or retire early in order to provide care. A 2011 report estimated the lifetime cost of adult daughters caregiving for their parents (in terms of lost wages, Social Security credits, and retirement pension) at $324,044 per woman (Metlife 2011).

People spend a significant amount of money on caregiving expenses, from direct expenditures on medical care or accommodations, to assuming the household needs of their parent, to paying for someone to provide care. An AARP mixed-methods study looked at expenses by income quartile (the lowest was less than $32,500, the highest was over $92,500) and found the range of how much people spend is relatively small across income quartiles, on average of $6,594 annually—from $5,114 for the lowest income quartile to $8,201 for the highest (Rainville, Skufca, Mehegan 2016). The weighted

average spending among all caregivers amounts to 20 percent of caregivers' annual income; but for the lowest income, the average is 44 percent. Notably, they found that the people who spend the most money on care also gave the highest number of hours providing care, suggesting that, on average, time spent caring and money spent on care expenses are complements, not substitutes.

Caregiving & Labor Force Participation

Does caregiving result in decreased labor force participation? Depends on which study you read. Underlying the lack of consensus in the literature on the relationship between caregiving and labor force participation is the problem of endogeneity when considering the effects of caregiving—what Heitmueller (2007) titles "the chicken or the egg": not knowing whether caregivers leave the labor force to begin providing care, or begin providing care because they were not in the labor force.

Lilly, Laporte, and Coyte (2007) conducted a systematic review of the literature on the relationship between caregiving and labor force participation (where participation is separately examined both as *any* employment and as hours of paid work). They include 27 studies using data from the US, 6 from the UK, one from Canada and one using comparative European data. The studies offer wildly varying estimates of how many caregivers are employed, ranging from a low estimate of 29 percent among American women in their 50s (Pavalko, Artis 1997) to a high of 82 percent among British men (Arber, Ginn 1995). They find that gender and age (particularly related to retirement) matter—and critique the U.S. literature (at the time) for insufficiently treating the question of retirement age. Using a weighted analysis of all studies, they find caregivers are less likely to be working than noncaregivers, but point out that individual studies find conflicting evidence (and the study by Pavalko and Artis, 1997, was the only in their review to systematically conclude caregiving decreases the likelihood of employment). The estimates on hours worked varied from no effect to an average decrease of six work hours per hour spent caregiving (Mutschler 1993), and overall they suggest the caregiving likely does result in a moderate reduction of hours of paid employment.

Van Houtven, Coe, and Skira (2013) offer an updated critique of the literature on caregiving and labor force participation, suggesting the lack of consensus may in part be because labor force participation can be defined on the extensive and intensive margins—and these actually may operate differently. Caregiving could reduce the extensive margin of labor force participation if people voluntarily quit their job in order to provide care, are discharged involuntarily (e.g., through objective or subjective assessments that caregiving disrupts their ability to perform work), or retire earlier than they would without caregiving. Caregiving could reduce the intensive margin of labor force participation if people are working less (e.g., reducing work hours, switching to part-time, taking leave). Van Houtven, Coe, and Skira (2013) fold into their discussion of the intensive margin a review of how people's income could decrease via wages, for instance if they take a lower-paying job, are passed over for a promotion, or decrease productivity.

There is evidence of significantly different patterns among subgroups of caregivers, for instance between men and women, or between more and less intensive caregiving. Lilly, Laporte and Coyte (2007), reviewing the literature at the time as a whole, find the caregiver's characteristics that emerge as predictive of reducing labor force participation (on the extensive or intensive margin), include being: a woman, in poor health, near retirement age, the primary caregiver (or providing "intensive" care, or caregiving for someone with greater disabilities), and being an immediate family member (particular a daughter or daughter-in-law). For two characteristics used as measures of socioeconomic status, income and education, Lilly, Laporte, and Coyte (2007) found mixed evidence. While most studies found that low-income caregivers were more likely to decrease labor force participation than high-income caregivers, Chang and White-Means (1995) found the opposite (as did Berecki-Gisolf et al 2008, published after their review). Likewise, most studies they reviewed found that caregivers with low educational attainment were more likely to decrease labor force participation, but Keating et al (1999), using Canadian data, found they were less likely to decrease. White-Means and Chollet (1996) attempt

to directly estimate the opportunity costs of caregiving by imputing wages for caregivers out of the labor force. In their regression of the effect of caregiving on employment, they included both opportunity costs (observed or imputed wages) and educational attainment as variables, controlling for other family income. They found that people with higher opportunity costs were *less* likely to decrease employment, but that people with higher education were *more* likely to decrease employment.

Socioeconomic Status

Relatively little attention has been given to the question of how caregiving varies by socioeconomic status, or what happens to the most economically vulnerable caregivers. As one study, attempting to break out of this pattern by looking at unpaid eldercare provision among welfare recipients, summarizes: the existing research on family eldercare is predominantly about "white, middle-class women" (Kneipp, Castleman, Gailor 2004, p. 25). When socioeconomic status is included in caregiving research, we have suggestive evidence that low-income families and people of color may be disproportionately likely to be providing unpaid eldercare and experience significantly more financial strain associated with eldercare (Rainville, Skufca, Mehegan 2016; Lee, Tan, Kim, Albert 2014). On the other hand, research focused on the transfer of both time and money between generations—not limited by the health or caregiving needs of parents—suggests economic status may make no difference on how much assistance adult children give their parents (e.g., Eggebeen, Hogan 1990; Klein Ikkink, Knipscheer 1999), or even that the financial constraints of being low-income depresses transfers to (and from) aging parents relative to higher-income families (Seltzer, Bianchi 2013; Hogan, Eggebeen, Clogg, 1993; Smits, van Gaalen, Mulder, 2010). Whether low-SES adult children are more or less likely to begin caregiving, we know people living paycheck-to-paycheck in low-wage jobs are the least likely to have access to flexible work places, leading to possibly job-threatening conflicts between their family needs and their employer's requirements (Lahaie, Earle, Heymann 2013; Swanberg, Pitt-Catsouphes, Drescher-Burke 2005; Covinsky et al 2001; Scharlach, Soberl, Roberts 1991).

Several papers examining the relationship between caregiving and labor force participation have used the Health and Retirement Study (HRS), which only includes people over age 50 (e.g., Johnson, Lo Sasso 2006; Wakabayashi, Donato 2006; Van Houtven, Coe, Skira 2013; Butrica, Karamchev 2014; Lee, Tang, Kim, Albert 2014; Gonzales, Lee, Brown 2017). Of these, the most methodologically sophisticated is Van Houtven, Coe and Skira's (2013) paper—which does not directly address variation by SES, though SES-related measures are included in some models as controls (e.g., asset quartile and home ownership). However, demographic research has documented that the accumulating health disadvantage of low-SES adults relative to high-SES is most pronounced in middle age, which could mean that the parents of low-SES children need assistance at relatively younger ages (e.g., House, Kessler, Regula Herzog 1990; McDonough, Duncan, Williams, House 1997). So while some of the most significant and sophisticated research in the U.S. has used the HRS (limiting analysis to caregivers over age 51), I expect that data including younger potential caregivers may be better suited to helping us understand the experience of low-SES caregivers—and thus findings on retirement-age caregivers may not be representative for low-SES caregivers.

If caregiving is endogenous to labor force participation—that is, if people begin providing care because they were not in the labor force to begin with—potential caregivers could be people with sufficient resources to prefer not working (e.g., retirees, spouses of earners) or people with very few resources unable to secure or maintain employment (e.g., low levels of education, high levels of disability, criminal history). In a world where people selecting into caregiving are not in the labor force and are more likely to have low socioeconomic status, the economic hardship they experience may not be additive from the responsibilities of caregiving at all, but only a reflection of their initial socioeconomic status. In this case, the policy implications for reducing economic hardship may be not specific for eldercare, but instead should focus on increasing labor force attachment and economic rewards for people at the bottom of the socioeconomic spectrum. (Though, notably, increasing the labor

force attachment of potential caregivers might be at odds with a policy initiative to increase the supply of unpaid caregivers.) In contrast, if there is an additive economic hardship experienced by caregivers, beyond the distribution of economic hardships by SES, the policy implications for reducing economic hardship should be to reduce disruption in labor force participation or directly subsidize the costs of care to alleviate economic hardship. Given how the policy implications vary by selection and causal mechanism, it is critical for informing policy that we use appropriate measures of socioeconomic status and economic hardship, and appropriately model the relationship of caregiving to economic hardship.

Research Question

In this research project, I will build on the existing literature to give more attention to socioeconomic status by (1) using a sample of all adult women, including under age 50; (2) allowing the relationship between caregiving and labor force participation to be moderated by SES. I am interested in whether the relationship between caregiving and labor force participation varies for economically vulnerable families, but defining economic vulnerability can be problematic, particularly if caregiving does cause people to leave the labor force. Many demographic researchers operationalize socioeconomic status in terms of an education gradient (e.g., Kalil, Ryan, Corey 2012; McLanahan, Jacobsen 2014). Education is both a proxy for socioeconomic status, in that educational attainment closely follows family's educational patterns and income brackets, and it is highly predictive of lifetime earnings potential (e.g., Draut, 2016). In this analysis, I will use education as my primary measure of SES; future efforts could investigate the robustness of my findings with alternative measures of SES.

II. METHODS

Data

I use the Panel Study of Income Dynamics (PSID). The PSID is a household-level longitudinal study that follows generations of families over time in the United States. The first wave of interviews in 1968 was based on a Census sample to achieve population-level representation, plus an oversample of

African American and low-income families. In 1997, a refresher sample of immigrant families was introduced to re-balance population-level representation. When used with weights, PSID findings are broadly representative of the US population as a whole in 2013.

Sample

I select my sample from the 2013 wave of the PSID, among people who were asked the Rosters & Transfers Module questions. The Roster & Transfers Module was administered to all respondents in the 2013 wave of interviews (N=9,063 households), though questions about transfers to parents were only asked of respondents under age 80 (i.e., assuming no living parents of respondents over age 80). I limit the sample to households with a woman (as the head of the house or the spouse of the head) with at least one living parent or parent-in-law in the 2013 wave of the PSID (N=5,817). Of that sample, I use listwise deletion to limit my sample to women not missing responses to my key variables, including time spent caregiving (N=180), education (N=74), employment status (N=55), parent's health (N=36), and distance to closest parent (N=212). Note that for time spent caregiving, I deleted respondents missing *any* report of time for a given parent unit, whereas for parent's health and distance to closest parent, I only deleted respondents missing reports for *all* parent units.

My final sample for analysis is N=5,296. For some outcomes, I only consider results among employed (N=4,237). For my final analysis, I consider subgroups by age: up to 50 N=4,179 (employed N=3,459), and age 51 or older N=1,117 (employed N=778).

By limiting my sample to people with parents currently living, I over-select households among the PSID with lower age-related mortality. Since age-related mortality is associated with socioeconomic status (e.g., House, Kessler, Regula Herzog 1990; McDonough, Duncan, Williams, House 1997), this introduces a confounder to my investigation of the relationship between SES, caregiving, and labor force participation. When looking at the whole sample, an education gradient is already apparent in the distribution of having at least one living parent—31 percent of respondents with a high school degree or

less do not have a living parent, versus 19 percent of respondents with some college and 16 percent of those with at least a college degree. This means my analysis is limited to a "select" set of low-education respondents who still have a living parent. Caution should be taken in interpreting and generalizing the findings beyond this select group.

Key variables

Predictor: Caregiving

My predictor is time spent caregiving for a parent (including step-parents and parent-in-laws) in the prior year. Respondents were asked separate reports about all living parents and parents-in-law (if applicable), resulting in a maximum of four parent response units plus respective spouses (i.e., father [and his spouse, if not mother], mother [and her spouse, if not father], father-in-law, if applicable [and his spouse, if not mother-in-law], and mother-in-law, if applicable [and her spouse, if not father-in-law]); where there are both biological and adoptive parents, respondents were constrained to answer about one or the other. I measure caregiving dichotomously as "Any Caregiving" (defined as at least 50 hours in 2012) and "Intense Caregiving" (at least 500 hours in 2012), plus a sensitivity test where I measure caregiving continuously as numbers of hours.[1] These thresholds are in line with several other studies in the literature (e.g., Barnett 2013; Butrica, Karamcheva 2014; Van Houtven, Coe, Skira 2013; and, for the "Intense" threshold, Carmichael, Charles 2003).

Since married respondents are asked who provided the majority of the hours given (asked qualitatively, i.e., "mostly head?"), I have pro-rated the count of hours so that married women get 90% of the hours they "mostly" provided, 50% of hours of "both equally", and 10% of hours that were "mostly" their spouse. I capped the number of hours per year at 8,736 (i.e., the total number of hours in

[1] The text question is as follows: "*Families sometimes help each other with activities such as errands, rides, chores, babysitting, or hands-on care. In 2012, did you (or your spouse, if applicable) spend time helping your parents? (If yes) About how many hours in 2012 did you (or your spouse) spend helping? Was that mostly (the head), mostly (spouse if applicable), or both equally?*"

a year), which affected less than N=10—I assume it is a reporting error based on how they were asked to report hours separately for each parental unit.

Outcomes: Labor Force Participation

I consider three labor force participation outcomes: a dichotomous indicator for "Whether employed" in 2012 (which I define as greater than $0 in wages, greater than $0 in income from labor, and greater than 0 hours working); plus two outcomes among the sample of employed women (N=4,216): a continuous measure of total hours worked in 2012 and a continuous measure of average hourly wage in 2012 (a variable constructed by the PSID that divides total income from labor by total hours worked) among employed. I created a "floor" on hourly wages so none are below the 2012 Federal Minimum Wage of $7.25 per hour, assuming wages below this weight are more likely to be artefacts of a mismatch in the reports of total earnings and total hours than to be sub-minimum wage.

Moderator: Education

I am interested in the moderating effect of socio-economic status (which I model here with educational achievement) on the relationship between caregiving and labor force outcomes. Education is defined categorically: high school or less (including GED), some college (including 2-year or technical degrees), and college or more (college defined as a 4-year degree).

Instrumental Variables

If labor force participation is endogenous to caregiving, my OLS estimates will be biased. An estimation strategy to get around the problem of endogeneity is to find an instrumental variable—something correlated with the explanatory variable, and only correlated with the outcome via the explanatory variable (Wooldridge 2013). So for instance, we could instrument caregiving with parental health—as Van Houtven, Coe, and Skira point out: "variation in the health of a parent… should directly vary the demand for informal care, but not directly affect work behavior of an adult child other than through the informal care path" (2013, p. 243). This estimation method requires a two-stage least

squares model, where we first predict the likelihood of caregiving (the instrumented variable) by parental health, then we regress labor force participation on the predicted value for caregiving.

I tested several instruments in combination, including whether any parent's health is Fair/Poor (versus Excellent/Very Good/Good),[2] whether any parent is unmarried, the number of living siblings (summed between focal woman and spouse if applicable, but not necessarily including step-siblings if her parent has re-partnered), and the distance to the closest parent (defined categorically, including co-resident, within 30 miles, 30-200 miles, 200-500 miles, and greater than 500 miles). The exogeneity of the distance measure is in question if we expect a person struggling with employment would decide to move back in with their parents (i.e., it may not be exogenous) but make the assumption that the other values for distance are exogenous to labor force participation but correlated with caregiving (and thus meet the requirements for instrumental variables). I exclude coresident respondents in my instrumental variable analysis models (N=324). I discuss my assessment of the instruments below in the analysis section.

Demographic Characteristics

I control for a set of characteristics that may affect the relationship between caregiving and labor force participation in order to generate more precise estimates. First, I include several controls related to the focal woman's parents and family, including the number of living parents (minimum of 1, maximum of 8 parents—if the focal woman is married, and both her and her spouse's parents have re-partnered with new spouses); the age of the oldest parent (in categories, from under 65, 65-75, 75-84, to over 85); [3] and age of mother at focal woman's birth. For variables I tested as instruments and decided were not strong enough (i.e., any parent unmarried and number of living siblings), I include these as demographic control variables.

[2] For parents who were married to someone other than a biological or adoptive parent, only marital status was included (i.e., no "step"-parent's characteristics are included). So any parent in poor health excludes step-parents.
[3] Age of oldest parent excludes step-parents.

The demographic characteristics of the focal woman include race/ethnicity (categorized as Non-Hispanic Black, Hispanic, Non-Hispanic White, and Non-Hispanic Other—including N=9 missing race), age of focal woman (in categories), current marital status of focal woman (single, cohabiting, and married), and whether the focal woman's health is Fair/Poor (versus Excellent/Very Good/Good). Race was constructed by first taking the response to a question about Hispanic or Latino ethnicity, then taking the variable for the first mention of race (i.e., a person with multiple racial identities is categorized by the first racial identity mentioned). Age categories distinguish eligibility for employment established by the Social Security Administration, which prior research has found significant (e.g., Van Houtven, Coe, Skira 2013); the early retirement age (ERA) for everyone is age 62, but the Federal Retirement Age (FRA) depends on birth year: for women born before 1943 (i.e., over age 70 in the 2013 wave of the PSID), the FRA is 65, for women born between 1943-1959, the FRA is 66, and for women born since 1960, the federal retirement age is 67.

I include two household-level controls: the number of people in the household, and an indicator for whether any child under the age of 18 lives in the household. Since other research on time use or domestic labor finds that self-reports and proxy reports vary (e.g., Bryant, Kang, Zick, Chan 2003), I include an indicator for whether the focal woman was the survey respondent or whether it was a proxy (which can include her spouse or another adult in the household).

I include additional controls only relevant to employed women: an indicator for whether her main job in 2013 is a salaried position (i.e., versus hourly or contract), which is an imperfect proxy since the employment outcomes are from 2012, but I expect to be strongly correlated with whether her 2012 job was salaried; the number of years she has worked since age 18 (this is a self-reported variable taken in the first wave she is in the PSID survey, then automatically updated every subsequent wave based on her employment status), plus its square (i.e., to accommodate non-linear effects of work experience);

and her tenure in years, plus its square, at her main job in 2013, which again is an imperfect proxy for her 2012 labor force outcomes.

Missing values

I include focal women missing values on the demographic characteristics, using an indicator for a missing value in my regression analyses. For continuous variables, I set the missing value equal to the mean of the non-missing; for categorical variables, I created a category for missing values; and for dichotomous indicators, I set the missing to 0. Of the whole sample, 10 percent are missing mother's age at focal woman's birth, and 5 percent are missing *at least one* report of whether any parent is unmarried, and less than 5 percent are missing all reports of parents' age and number of living siblings. Of the employed sample, 12% are missing the indicator for whether 2013 job is salaried and 5% are missing work experience since age 18.

Note that while an inclusion criterion for my sample was having non-missing values for whether any parent is in poor health, I did not require a non-missing value for every parent—just *at least one non-missing* report.

Analysis

I first present characteristics of the sample, which I split by caregiving status (including chi-square or t-tests to test for significant differences between caregivers and non-caregivers). For the key variables, I show categories of hours spent caring for more detailed comparison, and bivariate results for the relationship between caregiving and labor force outcomes, before controlling for confounders or endogeneity.

For my first set of multivariate regression models, I predict each of the three labor force participation outcomes, controlling for demographic characteristics—the "whether employed" outcome among the whole sample, and the "total hours worked" and "hourly wages" only among the employed. I present results for all three analyses as OLS models, to facilitate easy comparisons across models and

with the instrumental variable 2-stage least squares model (I conduct a logit analysis of the binary outcome "whether employed," though results only noted if differ in direction and statistical significance):

Base Model (OLS): $Y_i = \beta + \beta_1 CG_i + \beta_2 Ed_i + \beta_3 \boldsymbol{X_i} + \epsilon_c$

where Y is the predicted outcome (whether employed, total hours worked, and hourly wages) and i indexes individuals. β_1 is the coefficient for caregiving (CG), β_2 is the coefficient for education (Ed), and $\boldsymbol{X}$ is a vector of demographic and household characteristics. The set of controls vary slightly by outcome—in the two models among employed ("total hours worked" and "hourly wages"), I add a control for whether or not the respondent had a salaried position. For the "hourly wages" model, following Van Houtven, Coe, and Skira (2013), I include measures of years of work experience since 18 (and its square), and tenure at current job (and its square)—controls derived from the Mincer Equation formula originally developed to model the marginal returns of schooling to wages but commonly used now to more precisely model effects of other determinants on wage (e.g., see Björklund, Kjellström 2002). Note that the functional form of the Mincer Equation requires the log of the hourly wage; I test this as a sensitivity test, but when I tested the functional form for hourly wages, the linear term is preferred by both the AIC and BIC and ocular review suggests less dispersion in the residuals for the linear term.

Because the PSID is a generational study, the sample follows children of the original households over time as they move out and form their own households. As such, there are siblings and multiple generations represented in the Rosters & Transfers Module, so errors are likely to be correlated between families. ε_c is a clustered standard error term where all members of the same family tree are assigned one identifier (N=1,350 have at least one family member in the sample). Each outcome will be regressed on both measures of caregiving (any and intensive).

The second set of models use the same outcomes as my base models, but I use a two-stage least squares (2LS) instrumental variable model to instrument for hours of caregiving. I discuss the strength of my instruments below.

My third set of models test whether the relationship between caregiving and labor force participation vary by education. I introduce an interaction term between education and caregiving, first in the base OLS models and, in my fourth set of models, in the 2LS analysis. For this analysis, I only use moderated models (i.e., with one interaction term), constraining the relationship between all other control variables to be linear across education. Future checks could investigate analysis by educational subgroup, correcting for multiple comparisons. The advantage to not using subgroups in this analysis is the sample size is large enough to accommodate instruments.

In order to compare my results to the body of literature that uses data on people over age 51, I split my OLS models in two subgroups by age (under age 51, and 51 or over). The samples were too small to include the 2LS instrumental variable models.

Instrumental Variables

All analyses were conducted using the user-added "ivreg2" command in Stata 15.1 (Baum, Schaffer, Stillman 2010). Interpretation of the results, below, rely extensively on the technical notes provided by the "ivreg2" authors.

I considered several variables for my excluded instruments, including whether any parent was in poor health, the distance to closest parent, whether any parent is unmarried, and the number of living siblings. I tested these variables jointly and separately to determine the strongest instruments. Number of living siblings was not predictive, so I only use it as a control to get a more precise estimate. Any parent unmarried was weakly predictive, but the tests performed better when I did not use any parent unmarried as an instrument, so I only use it as a control. Distance to closest parent (excluding co-resident) was strongly predictive. Any parent in poor health met all assumptions, though it is not a very

strong predictor of caregiving on its own, so I only show results with both parent in poor health and distance from closest parents.

For the interacted model, I assume the interaction between caregiving and education suffers from the same endogeneity with labor force participation as caregiving. So I instrument for the interaction term, creating an interaction term between any parent in poor health and the indicator for high school education, then instrument for both caregiving and the interaction between caregiving and high school with the interacted parent health variable (Bun, Harrison 2014). Since in the interacted models I have multiple endogenous regressors, I use the Sanderson-Windmeijer multivariate F test appropriate for multiple excluded instruments instead of the basic joint F test in the first step.

For each model, I first test if there is evidence of endogeneity in the regressor (the Hausman test, only possible with more than one instrument—which I use in all models). I only show results where I could reject the null that the relationship is endogenous. However, I note that in a heterogenous effects world, where the exogenous variables may be operating independently from each other, the null on the Hausman test for endogeneity does not necessarily indicate the regressor is endogenous. Future tests could further examine whether there is evidence of heterogeneity in the effects of the excluded instruments.

I test the validity and strength of the instruments using several empirical tests: the Sargan-Hansen J test for overidentification, where the null is that the instruments are uncorrelated with error term and correlated with endogenous regressor (i.e., the instrument is valid if you cannot reject the null); the joint-F statistic from the first stage results, where the threshold for weak instruments is generally considered to be F>=10; and two tests for weak and underidentified instruments, using the Stock and Yogo thresholds for the special F tests, which are adjusted by sample size and tell us the F value necessary for varying levels of relative bias and sample size (i.e., since instrumental variable models have much larger standard errors than OLS analyses). Since my analysis includes clustering

standard errors for related family members within the PSID, I use the Kleibergen-Paap Wald F-Statistic instead of the conventional Cragg-Donaldson statistic to test for weak instruments.

Results of the tests are summarized below in Table 1; only the base and interacted models for "whether employed" (by any and intense caregiving for the base, and only by any caregiving for the interaction) had both evidence of endogeneity and had strong and valid instruments. Future efforts could examine whether the lack of evidence of endogeneity reflects hetoregenous effects of the instruments (i.e., that "cancel out" the evidence of endogeneity) or reflects sample sizes too small for the multiple parameters (e.g., the test of endogeneity for hourly wage by both measures of caregiving was marginally significant at $p<0.1$).

Table 1. Summary of results for tests for endogeneity and strength/validity of preferred excluded instruments (i.e., any parent in poor health and distance to closest parent)

	Whether Employed		Total Hours worked		Hourly Wage	
	Endogenous?	Strong / Valid?	Endogenous?	Strong / Valid?	Endogenous?	Strong / Valid?
Base IV - Any Caregiving	**Y**	**Y**	N	Y	N	Y
Base IV - Intense Caregiving	**Y**	**Y**	N	~Y	N	Y
Interacted IV - Any Caregiving (plus interaction with high school)	**Y**	**Y**	N	Y	N	N
Interacted IV - Intense Caregiving (plus interaction with high school)	Y	N	N	N	N	N

Key: Only results in shaded cells (i.e., where evidence of endogeneity and excluded instrument passes strength and validity tests) are presented.

III. RESULTS

Demographic characteristics of the sample are presented in Table 2, summarized for the whole sample and by caregiving status (where "Any caregiving" is at least 50 hours in 2012). Comparing the total sample characteristics to US Census data, the sample fairly closely matches the 2015 national total that about one third of women have at least a college education, and just over one third have a high school degree or less (Ryan, Bauman 2016).The distribution by race/ethnicity is skewed, with an overrepresentation of Non-Hispanic Black respondents (an artefact of the original oversample of African Americans, and using the PSID without weights)—at a third of my sample, versus making up about 13% of the population in the 2010 US Census—and an underrepresentation of Non-Hispanic Whites and

possibly Hispanics (at 56% versus 64% of the population, and 9% versus 16% respectively, though the 16% includes Hispanic ethnicities of any race) (Humes, Jones, Ramirez 2011).

Caregivers are more likely to have a number of markers associated with lower incomes, including low education, women of color, being single, and poor health. Women with a high school degree or less are overrepresented at 40% of caregivers (versus 33% of non-caregivers) and vice versa women with at least college degree are underrepresented at 25% of caregivers (versus 37% of non-caregivers) (p<.001); of all women with a high school degree or less, 30% are caregivers, versus 20% of all women with at least a college degree (not shown) (p<.001). An even wider spread is seen by race, where 45% of caregivers are Non-Hispanic Black women (versus 29% of non-caregivers), and 47% of caregivers are Non-Hispanic White women (versus 59.4% of non-caregivers) (p<.001); 36% of all Non-Hispanic Black women are caregivers, versus only 22% of Non-Hispanic White women (not shown) (p<.001). Single women are significantly overrepresented, making up 47% of caregivers (versus 27% of non-caregivers)—though there is no significant difference in caregiving status for the 11% of the sample that are cohabitating. Women in poor health are somewhat overrepresented among caregivers, comprising 15% of caregivers and only 13% of non-caregivers (p<.05). This difference may reflect an underlying endogeneity problem of the relationship between caregiving and labor force participation. That is, women in poor health may be less likely to be employed, and women who are not employed may be more likely to begin caregiving.

Surprisingly, there is no significant difference in the percent of caregivers and noncaregivers when the oldest parent under 65—I would expect there would be relatively few caregivers caring for a parent under 65. (Relatedly, there are significantly more focal women under age 24 who are caring for a parent.) This might suggest there are two different types of caregiving needs—needs related to general aging, and needs related to a disability or chronic illness with an onset before age 65. There are

significantly more caregivers with a parent over age 85 (p<.001), which we would expect for eldercare. There is no difference in mother's age at the focal woman's birth.

There is also no significant difference in caregiving by having a child under 18 in the household. This result is notable, since coresident children have been considered one of the "competing demands" that lead to decreased caregiving for aging parents (e.g., Barnett, 2013).

Among employed women, there is no difference by caregiving status for the number of years of work experience or tenure at current job; salaried employees, however, are underrepresented among caregivers (at 29% versus 37% of non-caregivers), which could reflect another measure of caregivers' lower socioeconomic status, or, if caregiving is endogenous to labor force participation, could indicate that salaried positions are more difficult to incorporate caregiving responsibilities with.

Table 2. Sample Characteristics by Caregiving Status

	Variable	N = Total 5,296 % / Mean (SD)	No caregiving 3,889 % / Mean (SD)	Caregiving 1,407 % / Mean (SD)	[a]
Moderating Variables	Education				
	High School or less	35.2	33.3	40.4	***
	Some college	31.2	29.8	34.5	**
	College or more	33.8	36.9	25.1	***
Controls (Among Employed)	2013 job is salaried [b]	35.2	37.3	29.1	***
	Years of work experience since 18 [c]	9.3	9.2	9.6	
	(SD)	(7.7)	(7.4)	(8.3)	
	Tenure at 2013 main job	4.7	4.6	4.8	
	(SD)	(7.1)	(6.9)	(7.6)	
Controls (All Models)	# Living parents	2.5	2.6	2.2	***
	(SD)	(1.1)	(1.1)	(1.0)	
	Any parent unmarried [d]	60.7	58.0	68.3	***
	# Living siblings [e]	4.6	4.7	4.5	*
	(SD)	(3.2)	(3.2)	(3.1)	
	Age of oldest parent: [f]				
	Oldest parent under 65	43.9	46.9	44.7	
	Oldest parent 65-74	24.8	26.1	21.0	***
	Oldest parent 75-84	19.5	19.9	18.6	
	Oldest parent over 85	11.0	10.1	13.5	***
	Age of mother at birth [g]	25	25	25	
	(SD)	(5.5)	(5.4)	(5.6)	
	Race/Ethnicity of focal woman				
	Non-Hispanic Black	33.0	28.8	44.7	***
	Hispanic	8.6	8.9	7.6	
	Non-Hispanic White	56.0	59.4	46.5	***
	Non-Hispanic Other [h]	2.4	2.9	1.2	**
	Age of focal woman				
	<24	6.9	6.4	8.4	*
	>=24 and <31	22.7	22.4	23.5	
	>=31 and < 50	49.3	50.9	44.9	***
	>= 50 and < EEA	17.1	16.7	18.2	
	>=Early eligibility age (EEA) and < FRA	2.8	2.7	3.0	
	>=Federal Retirement Age (FRA)	1.2	0.9	2.1	**
	Marital status of focal woman				
	Single	32.1	26.8	46.6	***
	Cohab	11.1	11.4	10.3	
	Married	56.8	61.7	43.1	***
	Focal woman in poor health	13.2	12.5	15.1	*
	Household size	3.3	3.3	3.3	
	(SD)	(1.6)	(1.6)	(1.6)	
	Any child under 18 in household	56.9	57.3	55.9	
	Response by proxy, not focal woman	30.5	32.7	24.5	***

[a] Test of significant difference between caregiving (>=50 hours in 2012) and no caregiving (<50 hours)

[b] Salary indicator missing among employed for N=492, no difference in missing by caregiving status (shown among non-missing)

[c] Years work experience missing among employed for N=201, with significantly fewer missing among caregivers (p<.05) (calculated among non-missing)

[d] Missing parent marital status for N=23, no difference in missing by caregiving status (shown among non-missing)

[e] Missing number of living siblings for N=2, no difference in missing by caregiving status (calculated among non-missing)

[f] Missing age of oldest parent for N=54, with significantly fewer missing among caregivers (p<.05) (calculated among non-missing)

[g] Missing age of mother at focal woman's birth for N=517, with significantly fewer missing among caregivers (p<.001) (calculated among non-missing)

[h] "Non-Hispanic Other" includes N=9 missing race/ethnicity

**p ≤ .05 **p ≤ .01 ***p ≤ .001*

Tables 3 and 4 summarize the relationships between the key variables of caregiving and labor force participation—Table 3 reports the coefficients and standard errors from bivariate analyses, and Table 4 uses finer measures of hours spent caregiving. As expected, caregivers are more likely to have a parent in poor health, and more likely to live near a parent (part of the basic requirement for these to be valid instrumental variables is they must be associated with the endogenous regressor).

In terms of the extensive margin of employment (i.e., whether or not the caregiver is employed at all), there is generally no difference in employment rate. However, intense caregivers (at least 500 hours a year, roughly equivalent to 10 hours per week) have a significantly lower employment rate than non-intense caregivers (73%, versus the overall sample employment of 80%, $p<.01$).

On the intensive margin (i.e., how much employed people work), there is no significant difference in average number of hours worked per week. This null is an interesting finding, since it means (for instance among intense caregivers) the 10 hours per week spent caregiving must be traded off with hours of time use other than paid employment. Likewise, there is no significant difference in total hours worked or number of weeks worked between caregivers and no caregivers; intense caregivers worked just over 100 hours less in the whole year than the overall sample total of 1,727 hours in 2012, equivalent to about 3 weeks less in the whole year than the overall sample total of 44.1 weeks in 2012. Caregivers make significantly less per hour than non-caregivers, with an average difference of about $3 less per hour for any caregiving, and $6 less per hour for intense caregiving.

Table 3. Coefficient on measures of caregiving from bivariate models predicting labor force participation outcomes

(Models for hours worked and hourly wages only among employed focal women)

	Whether employed			Total hours worked			Hourly wages		
	β	*	SE	β	*	SE	β	*	SE
Any caregiving	-0.02		(0.01)	-32.1		(26.59)	-3.01	***	(0.61)
Intense caregiving	-0.07	**	(0.02)	-132.3	**	(50.29)	-6.53	***	(0.76)
# hours (in 100s) caregiving	0.00	**	(0.00)	-4.7		(2.70)	-0.25	***	(0.04)
N =	*5,296*			*4,237*					

**p ≤ .05 **p ≤ .01 ***p ≤ .001*

Table 4. Distribution of key variables by Caregiving Status

		No Caregiving		Any Caregiving				
						Intense Caregiving		
Hours spent helping a parent:	TOTAL	0 hours	>0-<50 hours	50-<208 hours	208-<500 hours	>=500 hours	[b]	[a]
N:	5,296	2,836	1,053	716	323	368		
Instrument for Caregiving:								
Any parent in poor health	49.4	45.7	51.5	55.2	56.4	54.9	*	***
Distance to nearest parent household								
Co-resident (excluded from IV analysis)	6.2	3.7	3.6	7.7	13.0	23.6	***	***
w/in 30 miles	67.5	60.5	73.5	80.0	78.0	70.1		***
>30-200 miles	12.9	15.6	13.8	8.5	5.9	5.2	***	***
>200-500 miles	6.0	9.0	4.3	1.7	0.6	0.5	***	***
>500 miles	7.5	11.3	4.8	2.1	2.5	0.5	***	***
Outcome Variables:								
% Employed	80.0	79.2	83.8	81.7	79.0	73.1	**	
Among employed								
Avg hrs/week	37.9	38.0	37.7	38.4	36.7	37.7		
(SD)	(11.9)	(11.7)	(12.3)	(11.9)	(11.8)	(13.0)		
Total hours worked 2012	1,727	1,729	1,750	1,769	1,657	1,603	**	
(SD)	(736)	(736)	(711)	(731)	(745)	(803)		
Total weeks worked 2012	44.1	44.0	45.0	44.8	43.1	41.1	***	
(SD)	(11.9)	(12.0)	(10.6)	(11.3)	(13.0)	(14.2)		
Hourly wage	20.6	20.6	23.2	20.7	17.1	14.4	***	***
(SD)	(21.0)	(17.7)	(31.7)	(17.3)	(13.0)	(11.1)		

[a] Test of significant difference between caregiving (>=50 hours in 2012) and no caregiving (<50 hours)

[b] Test of significant difference between intense caregiving (>=500 hours in 2012) and not-intense caregiving (<500 hours)

**p ≤ .05 **p ≤ .01 ***p ≤ .001*

Tables 5 presents the results of the OLS analyses regressing labor force participation outcomes on caregiving, adjusting for demographic characteristics. The main panel shows the models for any caregiving and associated controls; the final panel shows just the coefficient from the model for intense caregiving (the coefficients on the controls followed a very similar pattern). Even after controlling for demographic characteristics, there is still no difference in employment rates between caregivers and non-caregivers, or between intense caregivers and non-intense/non-caregivers. Among employed focal women, holding all else equal, there is no difference in total hours worked by either measure of caregiving. Again, this null is quite surprising, since it means that even with additional hours spent providing care, caregivers are not reducing their hours more or taking more time off than non-caregivers. The coefficient on hourly wage is smaller than in the bivariate analysis, but still significant, where employed caregivers make about $1 less per hour that non-caregivers (p<.05), and employed intense caregivers make about $2 less per hour ($p$<.001).

Table 5. Relationship between measures of caregiving and labor force participation outcomes adjusted by demographic controls (coefficients on controls only shown for Any Caregiving model)
(Models for hours worked and hourly wages only among employed focal women)

	Whether Employed			Total hours worked			Hourly wages		
Variable	β	*	SE	β	*	SE	β	*	SE
Any caregiving	0.00		(0.01)	-9.99		(25.29)	-1.17	*	(0.69)
Education									
High School or less	-0.12	***	(0.01)	14.65		(28.77)	-9.38	***	(1.16)
Some college	-0.04	***	(0.01)	39.40		(26.51)	-7.42	***	(0.93)
College or more (omitted)									
2013 job is salaried [a]				360.70	***	(24.16)	5.32	***	(0.78)
Years of work experience since 18 [b]							-0.05		(0.36)
Years of work experience since 18 squared							0.01		(0.02)
Tenure at 2013 main job							0.33	***	(0.07)
Tenure at 2013 main job squared							0.00	***	(0.00)
# living parents	0.00		(0.01)	-8.11		(14.18)	-0.68	*	(0.38)
Any parents unmarried [c]	-0.01		(0.01)	-10.70		(22.12)	0.22		(0.61)
# Living siblings [d]	0.00	**	(0.00)	3.66		(3.92)	-0.30	***	(0.09)
Age of oldest parent: [e]									
Oldest parent under 65	0.00		(0.01)	29.30		(29.46)	-0.14		(0.66)
Oldest parent 65-74 (omitted)									
Oldest parent 75-84	0.00		(0.02)	6.44		(34.33)	2.28	*	(1.16)
Oldest parent over 85	-0.03		(0.03)	20.87		(51.02)	3.73		(2.49)
Age of mother at birth [f]	0.00		(0.00)	-2.04		(2.22)	-0.03		(0.05)
Race/Ethnicity of focal woman									
Non-Hispanic Black	0.01		(0.01)	99.01	***	(27.43)	-2.39	***	(0.63)
Hispanic	-0.01		(0.02)	89.72	**	(42.26)	-0.83		(0.79)
Non-Hispanic White (omitted)									
Non-Hispanic Other [g]	-0.04		(0.04)	160.80	**	(77.13)	-2.94	**	(1.31)
Age of focal woman									
<24	-0.04	*	(0.02)	-243.70	***	(52.34)	-3.95	***	(1.38)
>=24 and <31	-0.01		(0.01)	-53.19	*	(31.23)	-2.50	***	(0.72)
>=31 and < 50 (omitted)									
>= 50 and < EEA	-0.07	***	(0.02)	-114.10	***	(40.50)	-3.09	***	(1.18)
>=Early eligibility age (EEA) and < FRA	-0.24	***	(0.05)	-210.00	**	(90.44)	-3.15		(3.28)
>=Federal Retirement Age (FRA)	-0.44	***	(0.07)	-219.80		(154.30)	-2.97		(7.49)
Marital status of focal woman									
Single	0.01		(0.02)	60.20	*	(35.49)	-1.96	**	(0.93)
Cohab	0.02		(0.02)	45.60		(35.90)	-1.62	**	(0.82)
Married (omitted)									
Focal woman in poor health	-0.18	***	(0.02)	-131.60	***	(38.12)	-1.90	**	(0.80)
Household size	-0.02	***	(0.01)	-30.46	***	(9.58)	-0.51	***	(0.19)
Any child under 18 in household	-0.03	**	(0.01)	-120.70	***	(27.39)	1.46	**	(0.66)
Response by proxy, not focal woman	-0.05	***	(0.01)	10.75		(25.38)	1.41	**	(0.65)
Any Caregiving model Adjusted R^2		0.10			0.17			0.13	
Intense caregiving [h]	-0.04		(0.02)	-47.60		(47.94)	-2.27	***	(0.85)
Intense Caregiving model Adjusted R^2		0.10			0.17			0.13	
N =		*5,296*				*4,237*			

[a-f] Not shown but included in model, indicators for:

[a] missing salary (salary set to 0 where missing) (N=492)

[b] missing work experience (set to mean where missing, experience squared set to mean of experience squared) (N=201)

[c] missing parent's marital status (set to 0 where missing) (N=23)

[d] missing number of siblings (set to mean where missing) (N=2)

[e] missing age of oldest parents (categorical variable for missing) (N=54)

[f] missing age of mother at focal woman's birth (set to mean where missing) (N=517)

[g] "Non-Hispanic Other" includes N=9 missing race/ethnicity

[h] Coefficients demographic control variables follow similar pattern for Intense Caregiving as for Any Caregiving, so not shown separately

**p ≤ .05 **p ≤ .01 ***p ≤ .001*

Women with a high school education or less have an employment rate 12 percentage points lower than women with at least a college education (p<.001), however among employed women there

is no difference in hours worked by education. There are marked differences in wages by educational achievement (as we expect, since this is functioning as a proxy for socioeconomic status): women with a high school education or less make $9 less per hour than women with a college education or more, and women with some college education make $7 less ($p$<.001).

As we would expect, women age 24-30 and 31-50 have the highest employment rates, holding all else equal, with a very sizeable dropoff in employment for women over the Federal Retirement Age (44 percentage points lower than women ages 31-50, p<.001). Employed women ages 31-50 also work the most hours and have the highest hourly wage (though the difference is not significant with women over the early eligibility age or retirement age). Holding age of focal woman constant, there is essentially no significant difference by age of oldest parent or age of mother at birth.

Table 6. Instrumental Variable Analysis of whether employed, instrumenting measures of caregiving with parent health and distance from closest parent (demographic controls not shown)
(Model excludes co-resident, which may be endogenous to labor force participation)

	Whether Employed					
	Any Caregiving			Intense Caregiving		
	β	*	SE	β	*	SE
Caregiving (2SLS)	0.23	***	(0.06)	1.01	***	(0.28)
First-Stage predicting caregiving:						
Any parent in poor health	0.08	***	(0.01)	0.01	*	(0.01)
Distance to nearest parent household						
w/in 30 miles (omitted)						
>30-200 miles	-0.14	***	(0.02)	-0.03	***	(0.01)
>200-500 miles	-0.25	***	(0.02)	-0.06	***	(0.01)
>500 miles	-0.24	***	(0.02)	-0.06	***	(0.01)
Joint-F on excluded instruments for any caregiving [a]	109.78	***		27.61	***	
Regressor is endogenous (Hausman) [b]	16.97	***		16.83	***	
Underidentification (Kleibergen-Paap rk LM statistic) [c]	327.13	***		99.92	***	
Weak identification (Kleibergen-Paap Wald rk F statistic) [d]	109.78	[f]		27.61	[f]	
Overidentification (Hansen-J) [e]	2.14	n/s		1.78	n/s	
N =	*4,970*					

[a] Threshold for weak instruments is generally considered 10, so Joint-F >10 is strong enough
[b] Test of endogeneity only possible with >1 instrument—significant finding usually taken to suggest evidence of endogeneity, though in a heterogeneous effects model non-significance not conclusive of exogeneity. For these models, I only show results if I could reject the null of the Hausman test (i.e., assuming homogenous effects).
[c] Null is that the excluded instruments are not correlated with the endogenous variable; rejection of the null indicates excluded instruments are correlated with endogenous variable (a requirement to be valid instruments).
[d] Assessing weak identification requires thresholds based on sample size (see footnote f below); when conventional standard errors are used, this threshold is conducted with the Cragg-Donaldson statistic. But since I am clustering standard errors on the family ID, I use the Kleibergen-Paap Wald F-Statistic.
[e] Null is that the instruments are uncorrelated with error term and correlated with endogenous regressor—to indicate valid instruments, do not reject null.
[f] Stock-Yogo critical value maximal IV relative bias threshold: 16.85. Stock-Yogo critical value maximal IV size threshold: 24.58
**p ≤ .05 **p ≤ .01 ***p ≤ .001*

Table 6 shows the results of the instrumental variable analysis for whether employed in two columns, for any caregiving and intense caregiving. (Total hours worked and hourly wages not shown because the Hausman test for endogeneity was not significant.) The incremental caregiving induced by parent health and distance to parents is associated with a significant higher employment rate of about 23 percentage points—even higher for intense caregivers (though the instruments are weaker, calling into question the point estimate produced by the 2SLS). This suggests that, holding all else equal, women whose parents are in poor health or who live near their parents are *more* likely to be employed than their peers.

These results are surprising, since, as with the discussion of hours worked above, we might expect that people who begin caregiving would reduce their employment in order to manage their responsibilities—or at least that there is no significant difference, as the OLS results indicated in the cross-section. The significant results of the instrumental variable model also suggest that endogeneity is a problem in the OLS model—that is, the OLS is downwardly biased (toward zero).

Table 7 presents the coefficients from the interacted model to test the moderating effect of education on the relationship between caregiving and labor force participation (all models were run with the same demographic controls as in the base OLS models). The only significant interaction term across all the models is between intense caregiving and high school in the model predicting any employment, suggesting that intense caregivers with a high school education or less take an additional dip in employment relative to other caregivers.

Table 7. Education moderating OLS analysis of Caregiving (separate models for any/intense) and labor force participation outcomes (demographic controls not shown)

(Models for hours worked and hourly wages only among employed focal women)

	Whether Employed			Total hours worked			Hourly wages		
	β	*	SE	β	*	SE	β	*	SE
Any Caregiving	0.02		(0.02)	-9.2		(41.8)	-2.6		(1.6)
HS (vs College)	-0.11	***	(0.02)	9.4		(32.6)	-9.8	***	(1.4)
HS + Caregiving	-0.05		(0.03)	16.2		(59.4)	1.8		(1.5)
Intense Caregiving	0.06		(0.03)	-103.1		(95.7)	-3.8		(2.9)
HS (vs College)	-0.11	***	(0.01)	7.0		(29.5)	-9.5	***	(1.2)
HS + Intense Caregiving	-0.10	*	(0.05)	124.5		(117.8)	2.3		(2.9)
N =	*5,296*			*4,237*					

**p ≤ .05 **p ≤ .01 ***p ≤ .001*

However, while no other coefficients meet the threshold of p<.05, the coefficients for any employment and for hourly wages are all nearing significance, and might be expected to be significant if we had a large enough sample to accommodate the small cells. The general pattern is that college educated-caregivers have little difference from college educated-non caregivers; people with a high school education or less are significantly worse off than people with a college education or more; and caregivers with a high school education or less may be slightly less likely to be employed but have a slightly better (or no different) wage than high-school educated non-caregivers. Taken together, this is suggestive evidence that a measure of SES indicates greater hardship for low-SES caregivers. But results are not conclusive, since the coefficients may be insignificant either due to relatively small sample sizes or because caregiving is not a salient difference by SES. Future research could attempt this analysis with larger samples (i.e., using different data or with fewer exclusions than I used in my sample construction) or with another measure of SES.

Table 8 shows the results of adding the instrumented interaction term to the instrumental variable analysis. Note that while the results of the Hausman test for endogeneity were significant and the model passes several tests of strength, the test for weak identification (Kleibergen-Paap Wald F statistic, among analyses with clustered sample errors) does not meet the Stock and Yogo critical values

for the 5% relative bias or 10% maximal size. So the actual estimates from the model should be interpreted cautiously, since they are subject to sample size bias.

With that caveat of caution, we see again a positive relationship between the induced caregiving associated with parent's health and distance to parent and the extensive margin of labor force participation. However, instrumenting here for the interaction between caregiving and low education, we see a large reduction in employment rates.

Table 8. Education moderating IV analysis of any caregiving and labor force participation outcomes (demographic controls not shown), instrumenting caregiving and interaction term
(Model excludes co-resident, which may be endogenous to labor force participation)

	Whether Employed		
Variable	β	*	SE
Any caregiving (2SLS)	0.61	***	(0.16)
HS (vs College) (2SLS)	0.18		(0.12)
HS + Caregiving (2SLS)	-1.18	*	0.47
First-Stage predicting Any Caregiving			
Any parent in poor health	0.08	***	(0.01)
HS + Any parent in poor health	0.00		(0.03)
Distance to nearest parent household			
w/in 30 miles (omitted)			
>30-200 miles	-0.14	***	(0.02)
>200-500 miles	-0.25	***	(0.01)
>500 miles	-0.24	***	(0.02)
Joint-F on excluded instruments for any caregiving [a]	88.71	***	
First-Stage predicting HS + Caregiving			
Any parent in poor health	0.01	**	(0.00)
HS + Any parent in poor health	0.06	**	(0.02)
Distance to nearest parent household			
w/in 30 miles (omitted)			
>30-200 miles	-0.04	***	(0.01)
>200-500 miles	-0.08	***	(0.01)
>500 miles	-0.08	***	(0.01)
Joint-F on excluded instruments for any caregiving [a]	27.20	***	
Regressor is endogenous (Hausman) [b]	25.65	***	
Underidentification (Kleibergen-Paap rk LM statistic) [c]	20.25	***	
Weak identification (Kleibergen-Paap Wald rk F statistic) [d]	4.06	[f]	
Overidentification (Hansen-J) [e]	0.73	n/s	
N =	*4,970*		

[a] Threshold for weak instruments is generally considered 10, so Joint-F >10 is strong enough
[b] Test of endogeneity only possible with >1 instrument—significant finding usually taken to suggest evidence of endogeneity, though in a heterogeneous effects model non-significance not conclusive of exogeneity. For these models, I only show results if I could reject the null of the Hausman test (i.e., assuming homogenous effects).
[c] Null is that the excluded instruments are not correlated with the endogenous variable; rejection of the null indicates excluded instruments are correlated with endogenous variable (a requirement to be valid instruments).
[d] Assessing weak identification requires thresholds based on sample size (see footnote f below); when conventional standard errors are used, this threshold is conducted with the Cragg-Donaldson statistic. But since I am clustering standard errors on the family ID, I use the Kleibergen-Paap Wald F-Statistic.
[e] Null is that the instruments are uncorrelated with error term and correlated with endogenous regressor—to indicate valid instruments, do not reject null.
[f] Stock-Yogo critical value maximal IV relative bias threshold: 13.97. Stock-Yogo critical value maximal IV size threshold: 19.45
**p ≤ .05 **p ≤ .01 ***p ≤ .001*

Table 9 summarizes a series of models that subgroup the sample by age, showing coefficients for women under age 51 and 51 and over—the age of the highly influential HRS sample that other research has relied on. Subgrouping the sample gives us some insight into how the results I have found might hold for the methodology using the HRS if their samples included people under age 51. I only show results for the measure of caregiving for "any" (not "intense"), because the sample is too small.

On the extensive margin, an important difference emerges in the interacted model, where the interaction between education and caregiving is only significant for the subgroup under age 51. Part of my rationale for using the PSID sample for examining differential effects by SES is the younger age of onset for disabilities in low-SES individuals. The results among younger provide suggestive evidence that differences by education should, indeed, be investigated among younger potential caregivers. Future research could consider why the instrument for any caregiving holds for younger women but not older—perhaps because more parents of older women are in poor health, so there is a different underlying mechanism inducing people to begin caregiving? On the intensive margin, the point estimates for total hours worked are markedly different by age group (an average decrease of 18 hours for women under 51, versus an average increase of 26 hours for women 51 or older), though with large standard errors the coefficients are not significantly different from zero. In the interacted model for total hours worked, though, there is a very large coefficient on the interaction term between high school and caregiving (an increase of 254 hours per year, or the equivalent of nearly 5 more weeks of full-time work in 2012, $p<.05$). This finding suggests that, while the differential effects of socioeconomic status seem to vary by age, there is still further work to be done even in the HRS sample of people over age 51 to understand how vulnerable people (i.e., with low education or another proxy of low SES) fare. The pattern of results for hourly wages is qualitatively similar by age.

Table 9. Subgroup analyses by age (up to age 50, and 51 or older): OLS, IV, and education moderator models (demographic controls not shown)

(Models for hours worked and hourly wages only among employed focal women; IV model excludes co-resident, which may be endogenous to labor force participation)

		Whether Employed			Total hours worked			Hourly wages		
		β	*	SE	β	*	SE	β	*	SE
OLS Any caregiving (No Interaction)	Less than 51	0.00		(0.01)	-17.58		(28.72)	-0.48		(0.61)
	Age 51 or over	-0.04		(0.03)	25.91		(55.61)	-3.92		(2.90)
IV Any caregiving (No Interaction)	Less than 51	0.21	**	(0.07)						
	Age 51 or over		WEAK							
OLS Any Caregiving (Interaction)	Less than 51									
	Any caregiving	0.03		(0.02)	11.70		(46.85)	-0.94		(1.36)
	HS (vs college)	-0.10	***	(0.02)	-10.62		(36.59)	-7.99	***	(0.82)
	HS + caregiving	-0.07	*	(0.03)	-47.67		(67.27)	0.33		(1.49)
	Age 51 or over									
	Any caregiving	-0.08		0.05	-83.60		88.78	-8.68		(5.67)
	HS (vs college)	-0.19	***	0.04	63.20		(76.18)	-14.45	**	(5.21)
	HS + caregiving	0.06		0.07	253.52	*	(119.40)	7.94		(5.49)
	N Age Less than 51 =		*4,275*			*3,531*				
	N Age 51 or over =		*1,176*			*818*				
	N for IV analysis =		*3,901*							

**p ≤ .05 **p ≤ .01 ***p ≤ .001*

IV. DISCUSSION

Taken as a whole, this analysis has several limitations that should be considered to contextualize the results. First, I am only using one wave of the PSID, because the measure of caregiving was only asked 2013. We do not know when caregiving began, so we cannot follow individual changes over time, comparing a woman to herself before and after beginning caregiving. Future analyses could attempt to use the very rich longitudinal data in the PSID to construct instrumented or reduced form analyses, even without annual measures of caregiving. As it is, though, we know the OLS estimates have the potential to be quite substantially biased due to the endogeneity of caregiving to labor force participation. While I identified some strong instruments and found evidence of endogeneity (when I used more than one excluded instrument, so I could invoke the Hausman test for endogeneity), the results were not robust enough to allow a full comparison across all outcomes or by age subgroups.

However, given this stringent limitation, it is noteworthy how actively caregivers participate in the labor force—employed caregivers working as many hours as employed non-caregivers, but at lower hourly wages. There is inconsistent evidence of the salience of SES—whether vulnerable potential

caregivers experience a differentially disruptive relationship to labor force participation—but some signs of this, particularly on the extensive margin when we look at intense caregivers or at the instrumental variable for any caregiving, and on the intensive margin for women over 51. The inconsistent evidence could be partly due to small cell sizes in my analysis, which could be addressed in future research by selecting a larger sample (i.e., either through a different data source or by using less restrictive sample eligibility criteria). Or it could be that education is not the most relevant proxy of socioeconomic status for this sample—for instance, as I discussed in the sample characteristics, there is a very wide spread in the proportion of caregivers by race. Women of color face many socioeconomic status challenges that hold even when we control for education. Future research should investigate the most appropriate measure of SES affecting potential caregivers. It is also possible that SES is not a particularly salient point of difference in the relationship between caregiving and labor force participation, but there is more to investigate.

Finally, future research using the PSID of caregivers' experience should also include the experiences of caregivers who are men. While the majority of caregivers are still women, there is a significant number of men involved in eldercare, and we might expect that proportion to increase with secular demographic changes.

References

AARP. (2015). Caregiving in the US [Report]. Available at: http://www.aarp.org/content/dam/aarp/ppi/2015/caregiving-in-the-united-states-2015-report-revised.pdf

Allen, J. (1993). Caring work and gender equity in an aging society. In J. Allen & A. Pifer (eds.), *Women on the front lines: meeting the challenge of an aging America* (pp. 221-239). Washington, D.C.: The Urban Institute Press.

Arber, S., and J. Ginn. (1995). Gender differences in the relationship between paid employment and informal care. *Work, Employment and Society 9*(3):445–71.

Barnett, A. (2013). Pathways of adult children providing care to older parents. *Journal of Marriage and Family, 75*(1), 178-190.

Baum, C.F., Schaffer, M.E., Stillman, S. 2010. *ivreg2: Stata module for extended instrumental variables/2SLS, GMM and AC/HAC, LIML and k-class regression*. Available at: http://ideas.repec.org/c/boc/bocode/s425401.html. Technical notes available at: http://www.repec.org/bocode/i/ivreg2.html

Berecki-Gisolf, J., Lucke, J., Hockey, R., & Dobson, A. (2008). Transitions into informal caregiving and out of paid employment of women in their 50s. *Social Science & Medicine, 67*(1), 122-127.

Bianchi, S., Folbre, N., Wolf, D. (2012). Unpaid care work. In N. Folbre (ed.), *For Love and Money: Care Provision in the United States* (pp. 40-64). New York: Russell Sage Foundation.

Björklund, A., & Kjellström, C. (2002). Estimating the return to investments in education: how useful is the standard Mincer equation? *Economics of Education Review*, *21*(3), 195-210.

BLS (The U.S. Bureau of Labor Statistics). (2016 April). Women in the labor force: a databook. Retrieved from https://www.bls.gov/opub/reports/womens-databook/2016/home.htm on November 11, 2017.

Bryant, W., Kang, H., Zick, C., & Chan, A. (2003). His and Hers: Evaluating Husbands' Reports of Wives' Housework. *Family and Consumer Sciences Research Journal, 32*(1), 8-26.

Bun, M. J., & Harrison, T. D. (2014). *OLS and IV estimation of regression models including endogenous interaction terms*. [Discussion paper] UvA Econometrics. Available at: https://editorialexpress.com/cgi-bin/conference/download.cgi?db_name=IAAE2015&paper_id=140

Butrica, B., & Karamcheva, N. (2014 October). *The impact of informal caregiving on older adults' labor supply and economic resources* [Report]. Washington, DC: The Urban Institute. Available at: https://www.dol.gov/sites/default/files/ebsa/researchers/analysis/retirement/impactofinformalcaregivingonolderadults.pdf.

Carmichael, F. & Charles, S. (2003). The opportunity costs of informal care: Does gender matter? *Journal of Health Economics, 22*(5), 781-803.

CBO. (2013 June). *Rising demand for long-term services and supports for elderly people* [Report]. Congress of the United States Congressional Budget Office. Available at: https://www.cbo.gov/sites/default/files/113th-congress-2013-2014/reports/44363-ltc.pdf

Chang, C. F., & White-Means, S. I. (1995). Labour supply of informal caregivers. *International Review of Applied Economics, 9*(2), 192-205.

Covinsky, K. E., Eng, C., Lui, L., Sands, L. P., Sehgal, A. R., Walter, L. C., . . . Yaffe, Kristine. (2001). Reduced employment in caregivers of frail elders: Impact of ethnicity, patient clinical characteristics, and caregiver characteristics. *The Journals of Gerontology: Medical Sciences 56A*(11), M707-M713.

Draut, T. (2016). *Sleeping giant: how the new working class will transform America*. New York: Doubleday.

Earle, A., & Heymann, J. (2012). The cost of caregiving: Wage loss among caregivers of elderly and disabled adults and children with special needs. *Community, Work & Family, 15*(3), 357-375.

Eggebeen, D., & Hogan, J. (1990). Giving between generations in American families. *Human Nature, 1*(3), 211-232.

Ettner, S. (1995). The impact of "parent care" on female labor supply decisions. *Demography, 32*(1), 63-80.

Ettner, S. (1996). The opportunity costs of elder care. *Journal of Human Resources, XXXI*(1), 189-205.

Fraser, N., & Gordon, L. (2013). A genealogy of 'dependency': tracing a keyword of the US welfare state. In N. Fraser (ed.), *Fortunes of feminism: from state-managed capitalism to neoliberal crisis* (pp. 83-110). Brooklyn, NY: Verso Books.

Gonzales, E., Lee, Y., & Brown, C. (2017). Back to work? Not everyone. Examining the longitudinal relationships between informal caregiving and paid work after formal retirement. *The Journals of Gerontology: Series B*, *72*(3), 532-539.

Heitmueller, A. (2007). The chicken or the egg?: Endogeneity in labour market participation of informal carers in England. *Journal of Health Economics, 26*(3), 536-559.

Henz, U. (2006). Informal caregiving at working age: effects of job characteristics and family configuration. *Journal of Marriage and Family, 68*(2), 411-429.

Hogan, D., Eggebeen, D., & Clogg, C. (1993). The structure of intergenerational exchanges in American families. *American Journal of Sociology, 98*(6), 1428-1458.

House, J. S., Kessler, R. C., & Regula Herzog, A. (1990). Age, socioeconomic status, and health. *The Milbank Quarterly*, *68*(3), 383-411.

Humes, K. R., Jones, N. A., & Ramirez, R. R. (2011, March). *Overview of race and Hispanic origin 2010: 2010 Census Briefs.* [Report] U. S. Census Bureau. Available at: https://www.census.gov/prod/cen2010/briefs/c2010br-02.pdf

Johnson, R., & Lo Sasso, A. (2006). The impact of elder care on women's labor supply. *INQUIRY: The Journal of Health Care Organization, Provision, and Financing, 43*(3), 195-210.

Kalil, A., Ryan, R., & Corey, M. (2012). Diverging destinies: Maternal education and the developmental gradient in time with children. *Demography, 49*(4), 1361-1383.

Keating, N., J. Fast, J. Frederick, K. Cranswick, and C. Perrier. 1999. *Eldercare in Canada: Context, Content and Consequences*. [Report]. Report no. 89-570-XPE. Ottawa.

Klein Ikkink, C.E., & Knipscheer, C.P.M. (1999). Perceived instrumental support exchanges in relationships between elderly parents and their adult children: Normative and structural explanations. *Journal of Marriage and Family, 61*, 831.

Kneipp, S. M., Castleman, J. B., & Gailor, N. (2004). Informal caregiving burden: an overlooked aspect of the lives and health of women transitioning from welfare to employment? *Public Health Nursing, 21*(1), 24-31.

Lahaie, C., Earle, A., & Heymann, J. (2013). An uneven burden: Social disparities in adult caregiving responsibilities, working conditions, and caregiver outcomes. *Research on Aging, 35*(3), 243-274.

Lee, Y., Tang, F., Kim, K., & Albert, S. (2015). The vicious cycle of parental caregiving and financial well-being: a longitudinal study of women. *Journals of Gerontology Series B: Psychological Sciences and Social Sciences, 70*(3), 425-431.

Lilly, M., Laporte, A., & Coyte, P. (2007). Labor market work and home care's unpaid caregivers: a systematic review of labor force participation rates, predictors of labor market withdrawal, and hours of work. *The Milbank Quarterly, 85*(4), 641-690.

McDonough, P., Duncan, G. J., Williams, D., & House, J. (1997). Income dynamics and adult mortality in the United States, 1972 through 1989. *American journal of public health, 87*(9), 1476-1483.

McLanahan, S., & Jacobsen, W. (2014). Diverging destinies revisited. In P. R. Amato, A. Booth, S. M. McHale & J. Van Hook (eds.) *Families in an Era of Increasing Inequality: Diverging Destinies* (pp. 3-23). Springer.

Metlife. (2011 June). *The Metlife study of caregiving costs to working caregivers: double jeopardy for baby boomers caring for their parents* [Report]. Available at: https://www.metlife.com/assets/cao/mmi/publications/studies/2011/Caregiving-Costs-to-Working-Caregivers.pdf

Mutschler, P. (1994). From executive suite to production line. *Research on Aging, 16*(1), 7-26.

Pavalko, E., & Artis, J. (1997). Women's caregiving and paid work: causal relationships in late midlife. *The Journals of Gerontology. Series B, Psychological Sciences and Social Sciences, 52*(4), S170-9.

Pinquart, M., & Sörensen, S. (2007). Correlates of physical health of informal caregivers: a meta-analysis. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences, 62*(2), P126-P137.

Rainville, C., Skufca, L., & Mehegan, L. (2016 November). *Family caregiving and out-of-pocket costs: 2016 report*. Washington, DC: AARP. Available at: https://www.aarp.org/research/topics/care/info-2016/family-caregivers-cost-survey.html.

Reinhard, S. C., Feinberg, L. F., Choula, R. & Houser, A. (2015 July). *Valuing the invaluable: 2015 update. Undeniable progress, but big gaps remain*. Washington, DC: AARP. Available at: https://www.aarp.org/content/dam/aarp/ppi/2015/valuing-the-invaluable-2015-update-new.pdf.

Ryan, C. L., & Bauman, K. (2016, March). *Educational attainment in the United States: 2015 population characteristics.* [Report] Current Population Reports, P20-578. Available at: https://www.census.gov/content/dam/Census/library/publications/2016/demo/p20-578.pdf

Scharlach, A. E., Sobel, E. L., & Roberts, R. E. (1991). Employment and caregiver strain: An integrative model. *The Gerontologist, 31*(6), 778-787.

Schulz, R., & Eden, J. (Eds.). (2016). *Families caring for an aging America.* Washington, DC: The National Academies Press. Available for download at: https://www.nap.edu/read/23606/chapter/1.

Seltzer, J. A., & Bianchi, S. M. (2013). Demographic change and parent-child relationships in adulthood. *Annual Review of Sociology, 39*, 275-290.

Smits, A., Van Gaalen, R., & Mulder, C. (2010). Parent–child coresidence: who moves in with whom and for whose needs? *Journal of Marriage and Family, 72*(4), 1022-1033.

Swanberg, J. E., Pitt-Catsouphes, M., & Drescher-Burke, K. (2005). A question of justice: Disparities in employees' access to flexible schedule arrangements. *Journal of Family Issues, 26*(6), 866-895.

Van Houtven, C., Coe, N. B., & Skira, M. M. (2013). The effect of informal care on work and wages. *Journal of Health Economics, 32*(1), 240-252.

Wakabayashi, C., & Donato, K. (2006). Does caregiving increase poverty among women in later life? Evidence from the Health and Retirement Survey. *Journal of Health and Social Behavior, 47*(3), 258-274.

White-Means, S. & Chollet, D. (1996). Opportunity wages and workforce adjustments: Understanding the cost of in-home elder care. *The Journals of Gerontology, Series B, 51*(2), S82-S90.

Wolff, J. L., Mulcahy, J., Huang, J., Roth, D. L., Covinsky, K., & Kasper, J. D. (2017). Family caregivers of older adults, 1999–2015: Trends in characteristics, circumstances, and role-related appraisal. *The Gerontologist,* 16 June 2017. https://doi.org/10.1093/geront/gnx093.

Wolff, J., C. Spillman, B., A. Freedman, V., & D. Kasper, J. (2016). A national profile of family and unpaid caregivers who assist older adults with health care activities. *JAMA Intern Med. 176*(3), 372-379.