



DEMOGRAPHIC RESEARCH

A peer-reviewed, open-access journal of population sciences

DEMOGRAPHIC RESEARCH

VOLUME 37, ARTICLE 13, PAGES 363–416

PUBLISHED 16 AUGUST 2017

<http://www.demographic-research.org/Volumes/Vol37/13/>

DOI: 10.4054/DemRes.2017.37.13

Research Article

Birth spacing, human capital, and the motherhood penalty at midlife in the United States

Margaret Gough

© 2017 Margaret Gough.

This open-access work is published under the terms of the Creative Commons Attribution NonCommercial License 2.0 Germany, which permits use, reproduction, and distribution in any medium for noncommercial purposes, provided the original author(s) and source are given credit. See <http://creativecommons.org/licenses/by-nc/2.0/de/>

Contents

1	Introduction	364
2	Background	365
2.1	Labor force participation patterns of mothers in the United States	365
2.2	Policy context	366
2.3	Human capital theory	367
2.4	Selection	370
2.5	Contributions	371
3	Method	373
3.1	Data and sample	373
3.2	Measures	374
3.2.1	Dependent variables	374
3.2.2	Independent variables	375
3.2.3	Control variables	375
3.3	Analytic strategy	378
3.3.1	Estimation	380
4	Results	381
4.1	Predictors of birth spacing	384
4.2	Birth spacing and labor market outcomes (Hypothesis 1)	384
4.3	Age at birth, human capital, and labor market outcomes (Hypothesis 2)	385
4.4	Heterogeneous effects by age at first birth and education (Hypothesis 3)	387
4.5	Description of supplemental analyses	391
5	Discussion	392
6	Conclusions	394
7	Acknowledgements	397
	References	398
	Appendix	404

Birth spacing, human capital, and the motherhood penalty at midlife in the United States

Margaret Gough¹

Abstract

BACKGROUND

Researchers have examined how first-birth timing is related to motherhood wage penalties, but research that examines birth spacing is lacking. Furthermore, little research has examined the persistence of penalties across the life course.

OBJECTIVE

The objective is to estimate the effects of birth spacing on midlife labor market outcomes and assess the extent to which these effects vary by education and age at first birth.

METHODS

I use data from the United States from the 1979–2010 waves of the National Longitudinal Survey of Youth 1979 and dynamic inverse probability of treatment weighting to estimate the effects of different birth intervals on mothers' midlife cumulative work hours, cumulative earnings, and hourly wages. I examine how education and age at first birth moderate these effects.

RESULTS

Women with birth intervals longer than two years but no longer than six years have the smallest penalties for cumulative outcomes; in models interacting the birth interval with age at first birth, postponement of a first birth to at least age 30 appears to be more important for cumulative outcomes than birth spacing. College-educated women benefit more from a longer birth interval than less educated women.

CONCLUSIONS

Childbearing strategies that result in greater accumulation of human capital provide long-run labor market benefits to mothers, and results suggest that different birth-spacing patterns could play a small role in facilitating this accumulation, as theorized in past literature.

¹ University of La Verne, California, USA. E-Mail: mgough@laverne.edu.

CONTRIBUTION

I contribute to the demographic literature by testing the theory that birth spacing matters for mothers' labor market outcomes and by assessing the effects at midlife rather than immediately following a birth.

1. Introduction

Nearly 80% of mothers in the United States bear two or more children (Dye 2008), and a majority of these mothers are also employed (Bureau of Labor Statistics 2013). The labor market penalties experienced by mothers (the "motherhood penalty") are well known, and some evidence suggests that the labor market costs of a second child are even greater than those of a first child (Anderson, Binder, and Krause 2003; Budig and England 2001; Loughran and Zissimopoulos 2009). Yet, while we know women in the United States can mitigate the cost of a first child by postponing motherhood, we know virtually nothing about whether and how women can mitigate the costs of a second (or higher-order) child. In this study, I investigate whether birth spacing (the interval between the first and second births) plays a role in explaining the second-child penalty for US mothers. Specifically, I examine the second-child penalty accrued by age 45. By focusing on the second-child penalty at midlife, I provide insight into the long-run financial well-being of mothers with multiple children, especially how financial well-being may vary according to childbearing patterns.

The timing of first births and the spacing of subsequent children are intimately linked for biological reasons – the longer a woman postpones a first birth, the less time she has for bearing additional children. As such, the effects of birth timing and birth spacing may be confounded in existing estimates of the motherhood penalty. For example, if spacing one's children close together is good for women's labor market outcomes, as Ross (1974) and Mincer and Polachek (1974) have argued in their classic studies, and women who begin bearing children at later ages also tend to space their children close together (Troske and Voicu 2013), the reported financial benefits to postponing a first birth found in previous research will be overstated, since these benefits will derive from the spacing pattern as well as first-birth timing.

In this study I identify the effects of birth spacing on the accumulation of the motherhood penalty over the life course for mothers in the United States, and I test whether education or age at first birth moderates these effects. I hypothesize that longer birth intervals will be associated with smaller penalties and that the benefits of longer birth intervals will be amplified by having a college education or having postponed a first birth to at least age 30.

Understanding the relevance of birth spacing is important in the contemporary context because US women report that they want to have two children on average (Bongaarts 2002), but they also spend considerably longer in full-time, full-year work than previous generations of women, and they lack the social welfare supports that many other countries offer to reduce the conflict between paid labor and childbearing (Blau and Kahn 2013; Karimi 2014). Furthermore, their earnings comprise an increasingly large share of household income. While the median contribution of wives' earnings to family income was 26.7%, in 1980, the contribution has topped 36% since 2007, and the proportion of households headed by single mothers has also grown (Bureau of Labor Statistics 2013). Because birth-spacing patterns vary across different demographic characteristics and different social and policy contexts, differential birth spacing could be one contributor to inequalities in financial well-being across groups of women or across country contexts. In addition, spacing may be more influenced by family policies than first-birth timing, potentially suggesting a mechanism for policy change if certain spacing patterns are determined to be detrimental to women's long-run earnings or work experience.

2. Background

2.1 Labor force participation patterns of mothers in the United States

The majority of new mothers in the United States participate in the labor force within one year of birth; in 2008, 62% of US women with a birth in the last year were in the labor force (Laughlin 2011). While this percentage has increased over time, most of the increase in women working later into pregnancy and after a first birth in the United States came during the 1980s, during the period in which many of the women in the analytic sample for this study were having children (Laughlin 2011). In the 1980s and early 1990s, when most women in my sample were having children, about 65% of women were working during their first pregnancies, most of them full-time; there was a decline in the percentage quitting their jobs during pregnancy or after the birth; and there was an increase in the percentage taking paid leave offered by some employers (Laughlin 2011). Ultimately, for the period 1981–1995, 33%–42% of mothers were working within three months of a birth, 44%–53% were working within six months, and 53%–61% were working within 12 months; participation rates were higher for women who worked through pregnancy (Laughlin 2011). Thus, labor force participation rates have been high for mothers of infants in the United States for several decades, and this is especially true after first and second births compared to higher-order births (Han et al. 2008). Women with higher education, women whose first birth occurs at age 30 or

later, and married women are the most likely to have returned to the labor force within nine months of the first birth (Han et al. 2008), while women with low levels of resources and financial difficulties are more likely to return to the labor force within two months of the first birth (Han et al. 2008; National Partnership for Women and Families 2013).

2.2 Policy context

The policy context in which mothers in the United States make decisions about continued labor force participation differs considerably from the context in most other highly developed countries. For example, it was not until 1993 that the United States introduced the Family Medical Leave Act (FMLA), which mandated up to 12 weeks of unpaid leave for childbearing or family care over a 12-month period, and the act only applies to certain eligible employees (Laughlin 2011). There remains no federally mandated paid childbearing leave in the United States. As such, women taking leave under FMLA for a new child average only 58 days of leave (National Partnership for Women and Families 2013). Additionally, US women appear to experience wage penalties for motherhood, only part of which can be explained by reductions in work experience. This policy context contrasts considerably with countries such as Germany or Sweden. Germany, for example, provides maternity leave and maternity benefits (fully paid), parental leave of up to three years with job protection, and child-rearing benefits that are government transfers (Bergemann and Riphahn 2015). But labor force participation rates of mothers of young children are relatively low because the availability of childcare outside of the home is very limited (Brehm and Buchholz 2014), and researchers have estimated motherhood wage penalties of 18% (Gangl and Ziefle 2009). Sweden, on the other hand, combines a relatively long and generous parental leave period with widely available subsidized public childcare. As such, about 50% of Swedish mothers return to the labor market within three years of a first birth (Aisenbrey, Evertsson, and Grunow 2009), and some research indicates that the wage penalties for taking leave are minimal (Albrecht et al. 1999).

2.3 Human capital theory

Policy context aside, whether the length of time that passes between a first and second birth matters for women's labor market outcomes, especially in the long run, is an unresolved question, but existing theories in the motherhood penalty and labor market literatures can be leveraged to frame the question. In the US policy context, in the

motherhood penalty literature two of the common explanations for the penalty are most plausible as mechanisms through which birth spacing could affect women's outcomes: human capital accumulation and productivity, and selection. Lower investment in human capital (e.g., education, work experience, tenure) generally explains a portion of the motherhood penalty in the empirical literature (typically about 50%), but much of the penalty remains unexplained after accounting for human capital (Anderson, Binder, and Krause 2003; Budig and England 2001). Women may also suffer depreciation of their skills during periods out of the labor market (Mincer and Polachek 1974), which will also lower wages upon return to the labor force. Skill depreciation is greater for women with larger amounts of job- or employer-specific human capital compared to women with larger amounts of other types of human capital (Mincer and Polachek 1974).

It could also be the case that mothers are less productive in the workplace and this lower productivity results in the motherhood penalty. Becker (1985) argued that individuals have only a limited amount of effort to expend, and mothers will use more of their effort at home, leaving less available for paid work. This translates into lower on-the-job productivity, which results in lower wages, all else being constant. Unfortunately, productivity is largely untested in the literature and cannot be tested in this study either.

Women who return to the labor market after having a first child and before having a second child accumulate more work experience and tenure, which leads to higher wages over time (Troske and Voicu 2013). It is also likely to reduce the depreciation of human capital women experience from time out of the labor market for childbearing. From a productivity perspective, spacing children close together intensifies the early child-rearing period, which may make remaining attached to the labor force impossible, thereby hurting women's long-term economic growth. On the other hand, working part-time or shifting to the "mommy track" may halt women's career progress (Noonan and Corcoran 2004; Stone and Lovejoy 2004). Thus, if a longer birth interval means that women spend more time in part-time work, rather than returning to full-time work between births, the return to the labor market between children may be no more valuable than having two births close together and only returning to work (full-time) after the second child is born. For example, my informal calculations, based on Waldfogel's (1997) finding that part-time work carries a 10% penalty in the United States, indicate that working four years part-time would net the average woman only about 90% of the income of a comparable woman working two years full-time.

Education (as a measure of human capital) and age at first birth may moderate any effects of spacing in important ways. Women who postpone childbearing to later ages have the opportunity to accumulate more human capital in the form of education, work experience, and tenure, leading to higher relative wages (Blackburn, Bloom, and

Neumark 1993), greater market productivity (Troske and Voicu 2013), and a potential for higher wage growth. They may also experience less depreciation of human capital than earlier childbearers if depreciation costs decline with increased experience (Miller 2011). Greater human capital accumulation may also reduce the amount of time women take off after having a child (Herr 2012).

A recently published thesis chapter by Karimi (2014) is one of the few studies to address these issues as they relate to labor income and wages over the longer term, as I consider here. She studied income and wages 15 years after the second birth using Swedish register data and instrumenting for birth spacing using miscarriages. Karimi (2014) hypothesized that postponement of a first birth might not have the same benefits to women in Sweden as it does in the United States because Swedish family policies are universal and generous. She found that women who delay first births in Sweden tend to have the second child faster than those who did not delay. This is potentially problematic, because an increase in spacing between first and second births was associated with increased labor income over 15 years, an increased probability of returning to the labor market between births, increased long-run labor force participation for mothers, and an increase in monthly wages, especially for highly educated mothers (Karimi 2014). These findings are generally consistent with the theoretical predictions presented above, but the postponement evidence contrasts with studies from the United States. However, as discussed above, the Swedish welfare state context differs significantly from the US welfare state context, so it is difficult to know whether the results will hold in the US context.

Another study (Brehm and Buchholz 2014), looking at a sample of mothers in western Germany, found that only highly educated women could continue accumulating labor market prestige after childbearing, and only if they spaced their births close together. Mothers with low education experienced reduced prestige regardless of the spacing of their children. Furthermore, interspersing births with part-time employment was worse for women's careers than interrupting labor force participation completely. While not specifically focused on labor market earnings or labor force participation per se, Brehm and Buchholz's (2014) results contradict those of Karimi (2014) and suggest that the type of welfare state (in this case, a male breadwinner/female carer type) may have a substantial role to play in whether short or long spacing is better for women's long-run labor market outcomes.

In the United States, where job-protected leave is unpaid under FMLA, public childcare is rare, and the use of market-based solutions is encouraged, researchers have looked at birth spacing in relation only to labor force participation, rather than wages or income, with varying emphasis on human capital accumulation. For example, Troske and Voicu (2013) found that women with higher education had more to gain in terms of their return to full-time labor force participation by spacing children farther apart,

although they appear to more commonly space their children close together. By contrast, Peltola (2004) found that longer birth intervals were negatively associated with return to the labor market following a second birth. But Peltola's results were not disaggregated by prior labor force attachment, which could explain the differences in the results between the two studies.²

Human capital theory indicates the importance of minimizing time out of the labor force, and thus skill depreciation, for optimal labor market outcomes. Returning to work between births is one potential strategy to achieve this ideal, and evidence from Karimi (2014) and Troske and Voicu (2013) shows the potential benefits of spacing births farther apart to facilitate this between-birth labor force participation. Thus, I hypothesize that, because longer spacing may reduce time spent out of the labor market:

Hypothesis 1: The magnitude of the motherhood penalty for a second birth will decline as the birth interval increases.

Postponement of a first birth additionally allows women to temporarily exit the labor force around a birth with greater human capital accumulation and higher wages than an exit occurring at a younger age, and this could help explain the lower motherhood penalties for postponers found in the US literature (e.g., Amuedo-Dorantes and Kimmel 2005; Miller 2011), regardless of birth spacing. Thus, I hypothesize that:

Hypothesis 2a: Women who postpone a first birth to at least age 30 will have smaller motherhood penalties than those who have a first birth before age 30.

Hypothesis 2b: Women with college degrees will have smaller motherhood penalties than those without college degrees.

Finally, postponement and spacing may interact in such a way that longer intervals may amplify the benefits of postponement. If postponement is mainly a proxy for

² While these studies are useful for providing some insights into the role of birth spacing, they are considerably different than the current study. First, Troske and Voicu (2013) looked at the impact of birth timing and spacing on labor force participation (full-time work; full-time, part-year work; part-time work; and nonwork) only. They also included only married women and marital births in their study, which is limiting, and they used a different type of modeling strategy (Markov-chain Monte Carlo methods) than I use in this study. Peltola (2004) studied the timing of labor market reentry following a second birth, and she studied labor force participation five years after the second birth. She considered low part-time hours, high part-time hours, and full-time work. Her sample excluded women with twins and women who did not work in any of the years in the sample. Birth spacing was not a focal element of the study. Since Peltola was studying reentry, she used hazard models for estimation. Neither study focused on outcomes such as cumulative work hours, earnings, or wages, and neither study examined outcomes at midlife.

greater human capital accumulation, then longer intervals may amplify the benefits of human capital accumulation. I hypothesize that:

Hypothesis 3a: At longer birth intervals, women who postponed a first birth to at least age 30 will have smaller penalties than those who did not postpone to at least age 30.

Hypothesis 3b: At longer birth intervals, women with college degrees will have smaller penalties than those without college degrees.

2.4 Selection

Human capital accumulation is a plausible mechanism for explaining the motherhood penalty, but the motherhood penalty could also arise entirely because of selection. Women may choose to have children at times when their careers are stagnating or their wages are low because it reduces the cost of childbearing. Lundberg and Rose (2000) find some evidence for negative selection of motherhood on this basis. Yet a recent paper by Killewald and Gough (2013) finds no evidence of wage anticipation effects prior to birth in the United States. If anticipation effects exist, birth spacing will affect the motherhood penalty inasmuch as it is related to women's individual capacity for productivity or their lower labor market position. For example, if some women find that caring for children is so demanding that it requires leaving the labor force temporarily, a rise in a woman's wages will lead her to postpone having another child until the cost of leaving the labor force declines (Hotz, Klerman, and Willis 1997). This strategy poses a challenge for estimation of causal effects of birth-spacing intervals on future labor market outcomes because past labor market outcomes may influence both the birth-spacing interval and future outcomes. Dynamic models address this causal inference problem, which makes them an invaluable tool for my exploration of the causal effects of birth-spacing intervals.

Selection is also a challenge when studying the effects of birth spacing because a number of different mechanisms that may not be randomly distributed across the population may influence the length of a birth interval. For example, women (and their partners) may choose to space their births in a particular way to facilitate attachment to the labor force. Or they may choose to space their births in such a way as to reduce the family's expenditures on expensive childcare so that these women can justify continuing to work outside the home. This type of rational-actor story often implicitly underlies studies examining first-birth postponement and the motherhood penalty. To

the extent that such mechanisms are related to observed characteristics they can be accounted for indirectly in a statistical analysis.

The length of a birth interval may also result from unexpected circumstances. For example, either the first or the second birth may be mistimed, or women may have difficulty conceiving a second child, particularly if they postponed their childbearing to the point when their ability to conceive was already declining. These unexpected circumstances, though difficult to analyze, are beneficial from the researcher's perspective because they suggest that random variation exists (net of selection issues) and this is important for the identification of causal effects.

2.5 Contributions

I estimate the effect of birth spacing on the accumulation of the motherhood penalty for US women's midlife labor market outcomes. In doing so, I expand my focus beyond the standard outcomes in the motherhood penalty literature to examine cumulative labor market outcomes, in addition to hourly wage penalties. I examine whether penalties that may arise from different birth-spacing patterns are observed even over a long period of time, specifically around age 45. Past research suggests that motherhood penalties persist over time (Karimi 2014; Wilde, Batchelder, and Ellwood 2010), but to date persistence is an understudied issue.

Because of the inherent complications in comparing long-term outcomes among individuals with different times of entry into the labor market, different numbers of spells out of the labor force, and different labor market trajectories, I focus on the midlife outcomes at a single point in women's lives, at age 45, allowing elements of their previous histories to be absorbed into the net effect I observe at that age. I examine hourly wages for their theoretical importance and because they are the economic outcome of interest in the vast majority of motherhood penalty studies. However, I am also interested in women's cumulative work behavior and financial well-being. If some women accept a short period out of the labor force, during which they earn no income, in return for higher lifetime earnings than other women who spend more time in part-time work, it would be a mistake to believe that the women who take this time out of the labor force are disadvantaged in the long run. Therefore, I examine women's cumulative work experience and their cumulative earnings by age 45 to better understand potential long-run consequences.

One limitation of looking at outcomes many years after most childbearing has been completed is that women's wages and earnings may have recovered from any earlier career penalties. Estimating models with cumulative outcomes helps to address this issue. Furthermore, I begin by estimating standard pooled regression models of the

birth interval on the outcomes, using the same techniques as in other motherhood penalty research. This allows me to capture the possibility of penalties incurred across the life course, beyond any evidence I find for penalties measured at age 45.

I use a dynamic potential outcomes framework to estimate the effects of second births at different intervals on midlife outcomes. By introducing a dynamic framework, I can take into account the possibility that all second births do not have the same effect; in particular, having a second birth soon after a first birth may affect outcomes differently than having the births farther apart. I can also incorporate the influence of past labor market outcomes on the timing of second births to address the causal inference problem described earlier and account for selection in a time-varying fashion. Furthermore, this method allows me to produce unbiased estimates of the effect of birth spacing in the presence of heterogeneity by education and age at first birth, both of which are also predictors of birth-spacing patterns (Blundell, Dearden, and Sianesi 2005). I implement the framework using an inverse probability of treatment (IPT) weighting procedure.³ The purpose of the reweighting is to ensure that women's observed probability distribution of subsequently having a second birth is the same as the counterfactual distribution for the women who are observably similar but had a second birth if they had decided to wait for a second birth as well.

Finally, based on my hypotheses, I test whether the effect of birth spacing on mothers' outcomes is moderated by education or age at first birth. Both a later age at first birth and a college degree should be associated with smaller penalties regardless of birth interval length, but I expect the penalties to be smallest for women who postpone a first birth, or have a college degree, and have longer birth intervals.

³ Many studies in the motherhood penalty literature have used fixed-effects models to estimate the relationship between motherhood status and current wages. Such models have been invaluable. However, they are not well suited to the present research question because they compare within-person outcomes across periods. Since women do not experience more than one birth interval between their first and second children, fixed-effects methods are inappropriate here. Additionally, some researchers have used miscarriage as an instrument for the timing of a birth, including the timing of a second birth in comparison to a first (e.g., Buckles and Munnich 2012; Karimi 2014). There are a number of potential problems with this strategy, as discussed in Buckles and Munnich (2012) and Hoffman (1998), including the fact that miscarriages are underreported in the National Longitudinal Survey of Youth 1979 data (Hoffman 1998), and miscarriages may be associated with individual characteristics that are also related to labor market outcomes, such as race (Buckles and Munnich 2012).

3. Method

3.1 Data and sample

The data for this analysis comes from the National Longitudinal Survey of Youth 1979 (NLSY79), a long-term study conducted in the United States.⁴ The NLSY79 has been used in much of the literature assessing the motherhood penalty (e.g., Amuedo-Dorantes and Kimmel 2005; Budig and England 2001; Loughran and Zissimopoulos 2009). It is particularly appropriate for this research because it focuses on the experiences of young adults and captures nearly all of their work experiences until middle age.

Initiated in 1979 as a sample of 12,686 men and women aged 14–22, NLSY79 has surveyed respondents annually through 1994, and biennially thereafter. NLSY79 therefore provides a large sample of young women experiencing the transition to motherhood. I use data that covers 1979 to 2010. In 2010 the respondents were aged 45–53, and most women in the sample had completed their childbearing.

The analysis requires a number of sample restrictions. First, because I am interested in outcomes at age 45, I exclude all women who were no longer in the sample at age 45. This results in a loss of 2,558 women, more than half of whom come from the discontinued oversamples. Second, I exclude women who had a second birth prior to the first survey wave in 1979 (216 women). This is necessary for implementing the estimation strategy. Missing data on covariates results in the loss of 288 women; 95% of this missing data derives from missing data on gender role attitudes and Armed Force Qualification Test (AFQT) scores. After conducting a preliminary ordinary least squares (OLS) and fixed-effects analysis with all women to ensure that I could estimate a motherhood penalty for one or more children (Tables A-3 and A-4),⁵ consistent with past research, I drop women who are never observed to experience a birth (541 women). After comparing the outcomes of women with only one child to those of women with two or more children (Table A-3), I drop women who are never observed to have more than one child (734 women), so that I am always comparing the outcomes of women with at least two children based on their birth-spacing patterns. Note that

⁴ The initial response rate was 87%.

⁵ In the OLS models, looking across all pooled years of observations, penalties appear to accrue mainly for second and third children. This is true across all of the outcomes. For outcomes at age 45 there are similar patterns except there are no apparent wage penalties. This suggests that wages may recover by middle age, though cumulative outcomes may still suffer. The fixed-effects model results for outcomes pooled across years are somewhat different. As in the OLS models, one child is associated with more work hours and earnings, but unlike the OLS models one child (compared to none) is associated with large wage penalties. The second child is associated with greater work hours but a wage penalty, and unlike the OLS results, in the fixed-effects model a third child is associated with a penalty only for earnings. Thus, some of the penalties seem attributable to unobserved characteristics, as discussed extensively in the motherhood penalty literature.

these two births do not have to be to the same father, only the same mother. Supplemental analyses removing women with multipartnered fertility (MPF) are discussed in the Appendix. Because there is a greater failure to report work hours than earnings over the period, I limit the sample to women with available data on all outcome measures. This results in the elimination of 401 women. Twelve women are dropped because they did not respond to the survey for the several years surrounding their first and second births. The resulting analytic sample includes 36,104 person-year observations on 1,533 women, 79% of whom are observed in all 24 waves of data.

3.2 Measures

3.2.1 Dependent variables

There are three dependent variables, each measured at age 45.⁶ All three are logged, because the distributions of values are right-skewed. The first dependent variable is the log of cumulative work hours. Annual hours of work are calculated using the weekly labor force reports of actual hours worked in each week.⁷ The weekly reports are constructed by the NLSY79 survey team from the start and stop dates respondents report for jobs and the usual hours worked per week at that particular job. If the worker had periods in which they were not working at the focal job, those breaks are recorded as well. The details for the weekly accounting can be found in Appendix 18 of the NLSY79 Codebook Supplement.⁸ I calculate the respondent's annual hours (in 1,000s) worked over all of the years since the respondent entered the labor force. Greater cumulative work experience indicates less time out of the labor market in which skill depreciation could occur, may indicate greater investment in labor market human capital, and has implications for income because for many workers income is a product of work hours and hourly wages.

The second dependent variable is log cumulative earnings. Cumulative earnings are calculated by adding the respondent's annual income (in \$10,000s) over all of the years the respondent worked in the labor force. Earnings are important to study because

⁶ Due to the biennial nature of the data some respondents were not interviewed at exactly age 45. In those cases I use the value of the variable from the closest available interview to age 45 (either at age 44 or 46).

⁷ Two women in the sample are never observed to work before age 45. All other women have at least some history of paid labor.

⁸ While this method may introduce some measurement error over a method such as prospectively collecting weekly work hours or drawing from administrative data, it is likely to have less measurement error than methods that retrospectively ask respondents how many hours they worked each week between interviews or methods that simply multiply usual hours worked by 50 or 52 weeks, ignoring breaks or variation in usual hours across jobs.

they play a role in understanding mothers' long-term financial well-being. Greater earnings can arise from increased investment in human capital and reduced skill depreciation.

The third dependent variable is log hourly wage. This variable is measured for every year the respondent works. It is important because higher wages, like higher earnings, could arise from greater investment in human capital and reduced skill depreciation. Wages play an important role in understanding mothers' earning potential for long-term financial well-being (Pollak 2005). Models estimating log hourly wages at age 45 are limited to women with nonzero wages because, due to the implementation of the log scale, wages equal to zero are undefined. Supplementary analyses retaining these women in the wage models provide very similar results (available upon request). Wages and cumulative earnings are adjusted to 2011 dollars using the Consumer Price Index.

3.2.2 Independent variables

The main independent variables comprise a set of indicator variables for the length of the birth interval. I estimate models using two-year intervals, with the last (comparison) category defined as nine or more years. The intervals are calculated using the reported number of months between the first and second births.⁹ Further explanation of these categories is provided in the Analytic Strategy section.

3.2.3 Control variables

The OLS models used for comparison and the logit models used to calculate the propensity to have a second birth at a particular birth interval incorporate a set of control variables needed for the dynamic models. It is necessary to include all of the variables that potentially influence birth spacing. Based on the literature, one can identify a number of relevant demographic and labor market characteristics that are related to birth spacing or birth timing (which by extension may also be related to birth spacing) and labor market outcomes. These include variables such as age, education, and marital status.

To ensure that I accounted for all pertinent variables, I used a standard set of demographic characteristics that include education, race/ethnicity, marital status, age at

⁹ Supplemental analyses described at the end of the Results section examined birth interval as a continuous variable, but the continuous version obscures nonlinearities in the effects, so I proceed with indicators of categorical birth intervals.

first birth, residence in the South, urban residence, and year. I include both a linear term and a quadratic term in age at first birth to allow for the possibility of a curvilinear effect of age at first birth on the likelihood of a particular birth-spacing pattern and on labor market outcomes. Unless otherwise stated, each of these variables is time-varying.

Since the 1970s education has become increasingly linked with birth timing, with college-educated women more likely to delay childbearing than other women (Martin 2000; Rindfuss, Morgan, and Offutt 1996). Furthermore, those with greater education also have better labor market outcomes than those with less education. With regard to spacing, Yamaguchi and Ferguson (1995) found that women with low education had significantly shorter birth spacing than women with higher levels of education. Recently, Troske and Voicu (2013) also demonstrated education's relationship to birth spacing, making its inclusion in these models imperative. Education is a categorical variable with five levels: less than high school, high school (omitted category), some college, college, and more than college.

Race/ethnicity, marital status, and living in the South have all been associated with birth timing and labor market outcomes. Race and ethnicity are related to fertility delay and labor market outcomes; research by Bloom and Trussell (1984) indicated that black women were less likely to delay childbearing than nonblack women. Either through differential propensity to delay or through labor market experiences, race/ethnicity may influence birth spacing. Only three categories were provided in the survey data: nonblack, non-Hispanic; black; and Hispanic. Nonblack, non-Hispanic is the omitted category and is referred to as white for simplicity. There are strong effects of marital status on first-birth timing and the pace of subsequent fertility (Bumpass, Rindfuss, and Janosik 1977; Guzzo and Hayford 2011; Manning 1995). Marital status is a categorical variable with three categories: never married/widowed,¹⁰ married, and separated/divorced. Because the effect of marital status may vary by race/ethnicity, I also include an interaction between the two variables. Residence in the South and urban residence are each indicator variables. These account for important differences in fertility by region. For example, teenage birth rates are persistently higher in Southern states and rural areas (Centers for Disease Control 2015). Year is included as a linear measure.

In addition to demographic characteristics, I control for a set of individual factors that may be associated with both birth interval and labor market outcomes at midlife. I

¹⁰ There are very few cases of widowed respondents (less than 1% of the analytic sample), and analysis indicates that never-married and widowed respondents are not significantly different on any of the outcomes. This category includes women who may have been cohabiting. This is reasonable in this data because in a recent study using the same data set, Killewald and Gough (2013) found that the association between parenthood and wages did not differ for cohabiters compared to single people. Additionally, during the period in which these women are having first births, births to single, non-cohabiting mothers remained much more common than births to cohabiting mothers (Manning, Brown, and Stykes 2015).

control for potential experience (calculated as age minus years of education minus 5) as an indicator of the maximum length of experience a woman could have in the labor force, along with an interaction between potential experience and education, because the effect of potential experience may be nonparallel across educational groups (Heckman, Lochner, and Todd 2003). I also control for quartile of score on the AFQT, part of the Armed Services Vocational Aptitude Battery (ASVAB) administered to respondents in 1981. The AFQT is considered a proxy for skill level.

Work characteristics are also included to help characterize women's labor force attachment. Lagged labor market variables are included because, as described in the literature review, some past research has indicated that a rise in hourly wage, or other positive labor market experience, may induce postponement of a next birth. Although the evidence for this claim in the literature is mixed, I include lagged variables for sector and professional/managerial employment. In preliminary models I also included lagged hourly wage and *t*-2-lagged hourly wage, but neither was a significant predictor of second-birth timing. I use lagged variables so they are measured prior to the birth and not concurrently with the year of birth, as it is characteristics prior to the occurrence of the birth that may have influenced the timing of the birth. Sector can take three values: private industry (omitted), government, and self-employed/employed in a family business without pay. In addition, I control for family economic resources, which is calculated as total household income minus respondent income. This serves as a measure of additional economic resources available to the woman that may influence birth-timing and labor force decisions.

Finally, I control for two variables designed to elicit differences in women's preferences for work and family: the number of children women want and traditional gender role attitudes. Both of these characteristics are expected to influence the timing of childbearing and the likelihood of participating in the labor force while raising children. Women who want more children can be expected to space children more closely and may be less likely to participate in the labor market during child-rearing. Similarly, women with more traditional gender role attitudes may be less likely to see work as important to their identities and may also favor having more children and children similar in age. The number of children wanted is a continuous variable measured in 1979. The variable of gender role attitudes is developed using principal components analysis on a set of gender role attitudes questions asked in 1979, and I include the component that represents traditional gender role attitudes.

3.3 Analytic strategy

I estimate the effect of a second birth at a particular birth interval on women's midlife wages, along with cumulative labor market outcomes, to assess whether birth spacing moderates the motherhood penalty associated with a second birth.¹¹ To examine whether having a second child at different points in a woman's life affects her midlife labor market outcomes, I extend the counterfactual framework (described in great detail elsewhere: see, e.g., Rubin 1974) to the time-varying context, because the timing of a second birth is endogenous to labor market outcomes and cannot be controlled in a standard regression framework. The dynamic framework accounts for the possibility that all second births, regardless of timing, do not have equivalent effects on outcomes.¹²

For the purpose of implementing the framework, it is convenient to think of time as discrete (Sianesi 2004); in my analysis, mothers who are eligible for a second birth (the treatment group) at time t are those who still have not had a second child after t years following the first birth, irrespective of what happens after t . This is analogous to a survival model that includes duration-specific parameters. The comparison group for women who have a second birth at time t consists of all women who have experienced at least t years after the first birth who have not yet experienced a second birth but will experience a second birth at some point in the future. Thus, the comparison of the effect of a second birth at time t on labor market outcomes is with women who have reached t years after the first birth. For example, consider two women with first births occurring in 1980. Woman A has a second child in 1982, while Woman B has a second child in 1985. At $t=1$ (≤ 2 -year interval), women with a second child born within two years of the first, such as Woman A, are compared to women whose first child was also born two years before but whose second child will not be born until more than two years after the first, such as Woman B. Figure 1 shows this comparison across the different time points in the analysis, from $t=1$ (≤ 2 -year interval) to $t=4$ (7–8-year interval); the red lines identify the comparisons.¹³ The main analysis includes five categories: 1–2 years (intervals up to and including 24 months); 3–4 years (birth intervals of 25 to 48 months); 5–6 years (birth intervals of 49 to 72 months); 7–8 years (birth intervals of

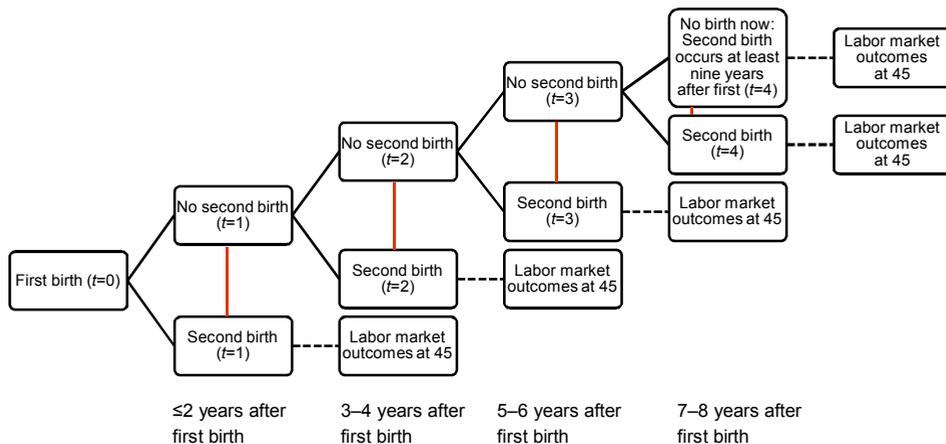
¹¹ There are limitations with regard to estimating a causal effect that are discussed in this section and further elaborated following the presentation of results. Nonetheless, the goal of the analysis is to achieve the best estimate of a causal effect, and the methods described in this section are designed for that purpose and implemented with that goal in mind.

¹² Similar methods have been used in sociology previously (see, e.g., Sharkey and Elwert 2011; Wodtke, Harding, and Elwert 2011).

¹³ One challenge with these models is that at longer intervals the comparison group potentially becomes more heterogeneous. Although the reweighting helps to address this, I provide supplemental analyses in the Appendix with different intervals to assess the robustness of the main results.

73–96 months); and 9+ years (always a comparison category only; birth intervals of 97 or more months).

Figure 1: Second-birth states from $t=0$ to $t=9+$ years after first birth



To identify the effect of a second birth at time t on labor market outcomes I use IPT weighting based on propensity scores, which is an alternative approach to propensity score matching. This method is useful – compared to OLS, for example – because it corrects for two sources of bias: differences in the supports of the variables and differences in the shapes of the distributions of variables in the region of common support (Heckman et al. 1998). In other words, the method increases the overlap in the values of variables for the two comparison groups (treatment and control groups), and it makes the shapes of these distributions of values more similar.

A third potential type of bias identified by Heckman et al. (1998) is selection bias. Obtaining the effects of second births at different intervals on women's labor force outcomes requires making assumptions about selection (also true in an OLS framework, not only in a dynamic potential outcomes framework). In particular, it requires making the conditional independence assumption (CIA), which states that treatment status (i.e., having or not having a second birth) is random conditional on a set of X observed covariates (Rosenbaum and Rubin 1983; Rubin 1978). The variables in X cannot be affected by the treatment, and X must include all of the variables that affect both the treatment and the outcomes.

In this dynamic context, this means that at each period the second birth is random conditional on observed covariates prior to that period, that these observed covariates are not affected by the second birth, and that the observed covariates included in the estimation comprise all of the variables affecting both the second birth at that time and the labor market outcomes.¹⁴ It is not possible to test whether the CIA is valid without conducting an experiment, but sensitivity analyses can help with assessment of validity.¹⁵ The variables in X are chosen based on theory and past research (Sianesi 2004). The dynamic CIA is plausible in this study because there is extensive data available longitudinally in the survey, and there is a large body of past research detailing the factors influencing timing of childbearing, previously described in Section 3.2.3, where I describe the included control variables.

3.3.1 Estimation

I estimate propensity scores and use them to create IPT weights, rather than matching on the propensity score directly (Busso, DiNardo, and McCrary 2014). To construct the IPT weights, I first estimate four logit models of selection into the different birth intervals and predict propensity scores. At the initial period of possible ‘treatment’ each woman is in the same treatment state – she has just had a first birth. This is followed by four two-year-long periods in which different treatment sequences could be realized (i.e., in each period a woman can have a second child or postpone a second child). Each logit model estimates the probability of a second birth in period t , conditional on X and on having reached time t without having already had a second birth. This is analogous in many ways to a hazard model, in which the population at risk is changing over time, as some women experience second births and some will not experience a second birth until some point in the future. Recently, Fitzenberger, Sommerfeld, and Steffes (2013) used this approach to study time to first birth in Germany. I provide estimates of the coefficients from the propensity score models in Table A-2. Examining the pseudo- R^2 values suggests that the models predict a second birth fairly well.

Once I have estimated the propensity scores for each period, I use the scores to generate IPT weights to be used in estimating the motherhood penalty. The most basic

¹⁴ Because there is a timing component involved (multiple years in which a second birth can occur), it is necessary for the CIA to hold in terms of future second births (dynamic CIA or DCIA). As such, the CIA does not have to hold in the conventional sense: having a second birth vs. never having a second birth. Instead it must hold at the margin: having a second birth at time t vs. postponing the second birth to at least $t+1$ (see, e.g., Sianesi 2004).

¹⁵ However, a formal sensitivity analysis along the lines of Robins (1999) requires assumptions about potential biases that are problematic in this context, including assumptions about direction, magnitude, and functional form (Sampson, Laub, and Wimer 2006).

form an IPT weight can take is $1/p$ where p indicates the propensity score derived from the logit model. In practice, stabilized or normalized weights are more commonly used, to guard against weights that are arbitrarily large. Busso, DiNardo, and McCrary (2014) show that IPT weighting is a good method for this type of analysis, but only if the weights are normalized.¹⁶

After I normalize the weights I then calculate the average treatment effect of a second birth at each time t on a mother's labor market outcomes to estimate the motherhood penalty (i.e., ATT). The main model is a simple regression of birth-at-time t on the labor force outcome, weighted by the IPT weight. Standard errors are calculated using Huber/White sandwich estimators to adjust for the two-stage estimation process. Finally, after obtaining the ATT for each time period, I follow the strategy of Abadie and Imbens (2011) and Fitzenberger, Sommerfeld, and Steffes (2013) and estimate ex post outcome regressions to examine effect heterogeneity. Specifically, I first add a control for age at first birth and separately a control for a college degree to test Hypothesis 2. I then add an interaction term to the simple regression on labor force outcomes, interacting birth-at-time t separately with age at first birth and then with education to test Hypothesis 3. I conduct Wald tests of the coefficients using the Bonferroni adjustment to address multiple comparisons in the pairwise tests. Rather than interacting the birth interval indicator with a continuous measure of age, I include age as a set of dummy variables to better account for possible nonlinearities in effects. There are three age indicators. The first indicates first births occurring between ages 18 and 29 (omitted category); the second indicates first births at ages less than 18; and the third indicates first births occurring at age 30 or older. I chose cutoffs of ages 18 and 30 because past research has indicated that childbearing before age 18 is distinct in a number of ways from childbearing at later ages and because in past literature a motherhood premium for postponement was found for women 30 and older at first birth (e.g., Amuedo-Dorantes and Kimmel 2005).

4. Results

Table 1 presents the sample-weighted descriptive statistics for the analytic sample.¹⁷ There are a couple of features to note. First, the average age at first birth is about 23.7

¹⁶ Sensitivity analyses discussed in the Results section also included estimates using nearest neighbor matching in line with Busso, DiNardo, and McCrary's (2014) suggestion that researchers examine estimates using a variety of approaches.

¹⁷ Table A-1 presents descriptive statistics for the analytic sample and the overall sample of women before restrictions were implemented. Because women in the oversamples were not followed past 1984 or 1990 (depending on the oversample) they contribute person-years at the lower end of the age spectrum. This is reflected in a slightly lower level of education in the original sample, fewer person-year observations as

years, which is consistent with the national average in the years in which most women in the sample had a first birth (1980–1989) (Hamilton and Martin 2013). Examination of the distribution of age at first birth in the sample (not shown) indicates that 25% of women had a first birth at age 19 or younger, while the oldest 25% of mothers in the sample had first births at age 26 or later. About 12% of first births occurred to women under the age of 18, while only about 1% of first births occurred to women over the age of 35. Thus, by measuring women’s outcomes at age 45, most women in the sample will have had ample time to return to work, following the birth of their second child. Indeed, at age 45 only 8% of the sample were not working in the labor force. Respondents who are working at age 45 have hourly wages averaging about \$18. Over the life course, women in the sample have worked nearly 39,000 hours on average, and they have accumulated an average of about \$415,000 in earnings. On average, by age 45 women in the sample have 2.6 children; an examination of the distribution indicates that a little over 50% of the sample have only two children by age 45, and the remaining women have three or more children, with the bulk of these women having three children.

Table 1: Sample-weighted descriptive statistics for the main analytic sample

Panel 1: Individuals	Mean (SD) ^a	Range
Number of children wanted (measured in 1979)	2.57 (1.35)	0.00–13.00
AFQT in 1980	48.36 (27.40)	1.00–99.00
Education at first birth		
Less than high school	13.71%	
High school	46.00%	
Some college	18.70%	
College	14.63%	
More than college	6.96%	
Race		
Nonblack, non-Hispanic	78.12%	
Hispanic	7.07%	
Black	14.82%	
Age at first birth	23.74 (5.03)	13.00–42.00
N	1,533	

married vs. unmarried, lower hourly wages, and lower cumulative hours. Nonetheless, the analytic sample has characteristics quite similar to the original sample overall, including age at first birth for those women experiencing a birth.

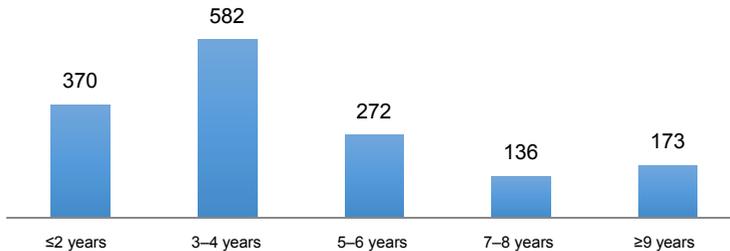
Table 1: (Continued)

Panel 2: Person-years		
Marital status		
Never married	24.92%	
Married	60.50%	
Separated/divorced	14.58%	
Occupation is professional/managerial	0.31	
Sector		
Private sector	78.05%	
Government	15.54%	
Self-employed/working without pay	6.42%	
Tenure with employer (years)	4.26 (5.21)	0.02–37.87
Part-time worker	0.24	
Not working in the labor force	0.20	
Family economic resources (household income minus respondent income)	36,766.54 (42,781.42)	0.00–252,740.20
Respondent resides in the South	0.34	
Respondent lives in urban area	0.73	
Hourly wage of respondent	12.22 (10.24)	0.00–49.07
Annual income of respondent	22,230.76 (21,948.26)	0.00–103,000.00
Cumulative hours of work experience	18,327.13 (15,976.87)	0.00–59,947.00
Cumulative earnings of respondent	202,352.50 (228,123.30)	0.00–2,042,437.00
N	36,103	
Panel 3: At age 45		
Hourly wage of respondent, if working	18.21 (10.99)	0.01–48.08
Cumulative earnings of respondent	415,025.70 (281,062.00)	0.00–1,836,467.00
Cumulative hours of work experience	38,642.29 (14,629.05)	0–59,947.00
Not working in the labor force	0.08	
Total number of children	2.63 (0.96)	2–11
N	1,533	

Note: ^a Standard deviations not provided for proportions, for which the entire distribution is provided.

Figure 2 shows the numbers of women in the sample with each birth interval. Second births are fairly well distributed across the intervals, with about 24% occurring within two years of the first birth, 38% occurring three to four years after the first birth, about 18% occurring five to six years after the first birth, and the remaining 20% of second births occurring at least seven years after the first birth.

Figure 2: Number of mothers by birth interval



4.1 Predictors of birth spacing

The results for the logit models predicting different birth-spacing intervals are shown in Table A-2. A few key predictors stand out. Age at first birth is a positive predictor of a birth at all intervals. This is consistent with the idea that women who are older at a first birth will have less time available for subsequent births. Being married, as opposed to being never married, is a positive predictor of a second birth for all intervals. Having at least some college education is a negative predictor of birth intervals of six or fewer years, while having less than high school education is a positive predictor of all intervals except those shorter than two years. This is consistent with the possibility raised in the literature that college-educated women might try to minimize the career penalties they will incur by having a second child by spacing their children farther apart. Past research in the United States and Sweden has indeed found that such behavior would be beneficial for these women (Karimi 2014; Troske and Voicu 2013).

4.2 Birth spacing and labor market outcomes (Hypothesis 1)

The results of the focal models of the role of birth spacing for women's midlife labor market outcomes are shown in Table 2. Hypothesis 1 stated that the magnitude of the motherhood penalty for a second birth would decline as the birth interval increases. If the hypothesis were supported we would expect to see that the penalties would become smaller as we look down the columns at each successive birth interval. The results in Table 2 indicate that this may be true for some women, but the pattern is not monotonic. Women with birth intervals shorter than two years do appear to experience

significant penalties compared to mothers with longer birth intervals: 14%¹⁸ lower cumulative work hours by age 45 ($p < .05$); 27% lower cumulative earnings by age 45 ($p < .001$); and 11% lower wages at age 45 ($p < .10$). By comparison, women with 3–4- and 5–6-year birth intervals do not appear to experience significant penalties compared to mothers with longer intervals, the exception being wages at age 45 for women with 3–4-year birth intervals. But significant penalties for cumulative outcomes return for women with 7–8-year birth intervals compared to women with the longest interval; these women have 20% fewer work hours and 30% lower cumulative earnings by age 45. Thus, the results for Hypothesis 1 are mixed. There is some evidence that longer intervals may be associated with smaller second-child penalties by midlife, but the pattern is not consistent across all of the intervals. This begs the question whether there may be unobserved heterogeneity at play. I return to this possibility of heterogeneity in Section 4.4.

Table 2: Inverse probability of treatment-weighted estimates of the average effect of a second birth on log cumulative work hours (in 1,000s), log cumulative earnings (in \$10,000s), and log hourly wages, at age 45, by two-year birth intervals

	Log cumulative work hours	Log cumulative earnings	Log wages
1–2-year birth interval ($N=1,533, 1,410$)	-0.15 (0.06)*	-0.31 (0.09)***	-0.12 (0.06) [†]
3–4-year birth interval ($N=1,163, 1,071$)	-0.02 (0.07)	-0.16 (0.11)	-0.22 (0.08)**
5–6-year birth interval ($N=581, 527$)	0.04 (0.07)	-0.12 (0.10)	-0.05 (0.07)
7–8-year birth interval ($N=309, 279$)	-0.22 (0.12) [†]	-0.36 (0.17)*	0.14 (0.10)

Notes: [†] $p < .10$; * $p < .05$; ** $p < .01$; *** $p < .001$. Cumulative models include women not working at age 44 or 45. Robust standard errors, clustered at the individual level, are estimated. Log wage model includes only women working at age 44 or 45, resulting in a smaller sample size.

4.3 Age at birth, human capital, and labor market outcomes (Hypothesis 2)

The results presented in Table 2 provide average effects of a second birth at each time interval t for women with a second birth at time t , as opposed to postponing the second birth to a later time. The mixed results suggest there could be unobserved heterogeneity that needs examination. Before turning attention to this possible heterogeneity, I examine the evidence for Hypotheses 2a and 2b to determine whether age at first birth and education could plausibly act as moderating factors in this sample. Hypothesis 2a stated that women who postpone a first birth to at least age 30 would have smaller motherhood penalties than those who do not postpone to at least age 30. If this

¹⁸ Percentages calculated from exponentiating the coefficients.

hypothesis is supported we should see smaller penalties for women age 30 and older at first birth compared to women who were younger at first birth. Evidence for Hypothesis 2a can be found in Panel 1 of Table 3. Looking at all three outcomes, the hypothesis is supported almost across the board. Only for log wages in the model of women with birth intervals of seven or more years is postponement not statistically significant. Exponentiating the coefficients on being age 30 or older at first birth, and averaging across models, indicate that, compared to women aged 18–29 at first birth, women who postpone achieve about 22% greater work hours by age 45; more than twice the cumulative earnings; and wages at age 45 that are roughly 25%–40% higher depending on the model. Thus, there is robust support for Hypothesis 2a that women who postpone a first birth have smaller penalties than those who do not postpone.

Hypothesis 2b stated that women with a college degree would have smaller motherhood penalties than those without a college degree. If this hypothesis is supported we should see smaller penalties for women with a college degree than we do for women with less education. Evidence for Hypothesis 2b can be found in Panel 2 of Table 3. An examination of the coefficients on college degree suggests strong, though not universal, support for this hypothesis. For cumulative work hours, in all but the first model college-educated women have work hours that are 22%–40% greater by age 45 than those of women without a college degree. Similarly, for cumulative earnings, and much like the postponement of a first birth, these women have twice the earnings at age 45 of women without a college degree. And finally, in three of the four models college-educated women have hourly wages at age 45 that are 25%–60% higher than those of women without college degrees, depending on the model. Thus, as with Hypothesis 2a, Hypothesis 2b, that women with college degrees have smaller penalties than those without college degrees, is well supported.

Table 3: Inverse probability of treatment-weighted estimates of the average effect of a second birth on log cumulative work hours (in 1,000s), log cumulative earnings (in \$10,000s), and log hourly wages, at age 45, by two-year birth intervals, accounting for age at first birth (Panel 1) and college degree (Panel 2)

	Log cumulative work hours	Log cumulative earnings	Log wages
Panel 1: Age at first birth			
1–2-year birth interval (N=1,533, 1,410)	–0.15 (0.06)**	–0.33 (0.08)***	–0.12 (0.06) [†]
<18 at first birth	–0.09 (0.07)	–0.45 (0.12)***	–0.10 (0.08)
≥30 at first birth	0.25 (0.05)***	0.77 (0.08)***	0.33 (0.09)***
3–4-year birth interval (N=1,163, 1,071)	–0.01 (0.06)	–0.11 (0.10)	–0.20 (0.08)*
<18 at first birth	–0.16 (0.16)	–0.43 (0.20)*	–0.13 (0.08)
≥30 at first birth	0.20 (0.05)***	0.71 (0.08)***	0.25 (0.08)**
5–6-year birth interval (N=581, 527)	0.05 (0.07)	–0.10 (0.10)	–0.04 (0.07)
<18 at first birth	–0.04 (0.08)	–0.18 (0.12)	–0.06 (0.08)
≥30 at first birth	0.20 (0.06)**	0.71 (0.11)***	0.22 (0.09)*
7–8-year birth interval (N=309, 279)	–0.21 (0.12) [†]	–0.33 (0.15) [†]	0.14 (0.10)
<18 at first birth	–0.26 (0.21)	–0.31 (0.27)	–0.02 (0.13)
≥30 at first birth	0.24 (0.08)**	0.86 (0.10)***	0.17 (0.12)
Panel 2: College degree			
1–2-year birth interval (N=1,533, 1,410)	–0.15 (0.06)*	–0.30 (0.09)**	–0.11 (0.06) [†]
College degree	0.05 (0.14)	0.70 (0.18)***	0.22 (0.12) [†]
3–4-year birth interval (N=1,163, 1,071)	–0.02 (0.07)	–0.14 (0.10)	–0.21 (0.08)**
College degree	0.26 (0.05)***	0.88 (0.07)***	0.24 (0.10)*
5–6-year birth interval (N=581, 527)	0.04 (0.08)	–0.12 (0.10)	–0.05 (0.07)
College degree	0.21 (0.06)**	0.74 (0.09)***	0.10 (0.16)
7–8-year birth interval (N=309, 279)	–0.23 (0.12) [†]	–0.39 (0.16) [†]	0.12 (0.10)
College degree	0.33 (0.11)**	0.85 (0.17)***	0.48 (0.14)**

Notes: [†] p<.10; * p<.05; ** p<.01; *** p<.001. Cumulative models include women not working at age 44 or 45. Robust standard errors, clustered at the individual level, are estimated. Log wage model includes only women working at age 44 or 45, resulting in a smaller sample size.

4.4 Heterogeneous effects by age at first birth and education (Hypothesis 3)

In the previous section I demonstrated that age at first birth and education affect midlife labor market outcomes in the ways previously hypothesized. But earlier in the paper I also hypothesized that education and age at first birth would moderate the effects of birth spacing because of their influence on human capital accumulation before and after a first birth. Specifically, I hypothesized that at longer birth intervals women who

postponed a first birth to at least age 30 would have smaller penalties than those who did not postpone to at least age 30 (Hypothesis 3a). If this hypothesis is supported we should see nonsignificant differences with regard to postponement for the shortest birth interval but a growing importance of postponement and significant effects at longer intervals. To make the interpretation of the interactions more concrete, I provide estimated marginal means, rather than coefficients, in Table 4. Each panel of the table indicates the estimated marginal mean with a second birth at that interval, broken down by age at first birth. Note that these marginal means are still logged; to calculate actual mean cumulative work hours, for example, one would need to exponentiate the appropriate marginal mean value. As an example, for the 1–2-year birth interval and cumulative work hours, the estimated means are 28,503 for women aged 18–29 at first birth ($(e^{3.35}) \times 1,000$); 26,311 for women under age 18 at first birth; and 39,646 for women at least age 30 at first birth.

Table 4: Estimated marginal means derived from inverse probability of treatment-weighted estimates of the average effect of a second birth on log cumulative work hours (in 1,000s), log cumulative earnings (in \$10,000s), and log hourly wages, at age 45, by two-year birth intervals, interacted with age at first birth

	Log cumulative work hours	Log cumulative earnings	Log wages
Panel 1: 1–2-year birth interval			
18–29 at first birth	3.35 (0.07)	2.97 (0.10)	2.49 (0.07)
<18 at first birth	3.27 (0.10)	2.42 (0.19)	2.34 (0.12)
≥30 at first birth	3.68 (0.05)	3.88 (0.10)	2.91 (0.12)
Panel 2: 3–4-year birth interval			
18–29 at first birth	3.51 (0.04)	3.21 (0.09)	2.50 (0.10)
<18 at first birth	3.28 (0.25)	2.66 (0.29)	2.37 (0.08)
≥30 at first birth	3.68 (0.05)	3.89 (0.10)	2.78 (0.10)
Panel 3: 5–6-year birth interval			
18–29 at first birth	3.57 (0.07)	3.18 (0.08)	2.62 (0.05)
<18 at first birth	3.43 (0.09)	2.81 (0.17)	2.64 (0.09)
≥30 at first birth	3.71 (0.08)	4.02 (0.11)	2.91 (0.11)
Panel 4: 7–8-year birth interval			
18–29 at first birth	3.42 (0.10)	3.05 (0.12)	2.74 (0.08)
<18 at first birth	2.72 (0.32)	2.18 (0.43)	2.76 (0.12)
≥30 at first birth	3.72 (0.10)	3.96 (0.12)	3.00 (0.09)

Notes: Estimated marginal means shown with delta-method standard errors.

Looking first at cumulative work hours in Column 1, it appears that women who postponed a first birth to at least age 30 always have higher work hours, while women experiencing a first birth prior to age 18 have the lowest. An examination of the Wald test results indicate that at the 5–6-year ($p < .05$) and 7–8-year ($p < .01$) birth intervals only the comparison between mothers younger than age 18 at first birth and women who postponed the first birth to at least age 30 is statistically significant. Thus, for cumulative work hours, the hypothesis does not hold in general, but rather only in comparison to the most disadvantaged mothers.

Turning to cumulative earnings in Column 2, we see the same pattern, whereby mothers under age 18 at first birth have the lowest earnings, mothers aged 18–29 at first birth have higher earnings, and mothers at least age 30 at first birth have the highest earnings. In contrast to cumulative work hours, Wald test results indicate that women who postponed a first birth to at least age 30 always have significantly better outcomes than women who did not postpone a first birth to age 30 ($p < .001$ in most cases). This result is partially in conflict with Hypothesis 3a, since women who postponed a first birth and have shorter birth intervals are also seeing benefits, not only women with longer birth intervals, and the magnitudes of the differences between the postponers and nonpostponers are fairly stable across the intervals.

Turning finally to log wages in Column 3, in models of working women only, we see limited evidence for the hypothesis. According to the results of the Wald tests, women who postponed a first birth to at least age 30 have significantly higher wages than women with a first birth prior to age 18 in models of the 1–2-year ($p < .01$) and 3–4-year ($p < .01$) birth intervals. They have significantly higher wages than women aged 18–29 at first birth in the 1–2-year ($p < .01$) and 5–6-year ($p < .05$) models. But women who postponed a first birth to at least age 30 do not have higher wages at age 45 when they have longer birth intervals, which is inconsistent with Hypothesis 3a.

Overall, there is very good evidence that postponement of a first birth to at least age 30 has a positive effect on midlife labor market outcomes. There is less consistent evidence that it meaningfully modifies the effects of birth spacing on outcomes, though in the case of log wages, age at first birth modifies the effects of birth spacing at the shorter (and mid-length) intervals, but not at the 7–8-year birth interval as hypothesized.¹⁹

I also hypothesized that at longer birth intervals, women with college degrees would have smaller penalties than those without college degrees (Hypothesis 3b). If this hypothesis is supported we should see nonsignificant differences with regard to education for the shortest birth interval but growing importance of education and

¹⁹ As discussed in the Appendix, I examined whether these postponement effects are also seen when looking in the short run – five, ten, and 15 years after first birth. The results of these models are consistent with the results shown in Table 4. Results are shown in Table A-8.

significant effects at longer intervals. As with the results for Hypothesis 3a, I provide estimated marginal means, rather than coefficients, in Table 5 for ease of interpretation. Each panel of the table indicates the estimated marginal mean with a second birth at that interval, broken down by whether the respondent has a college degree. All marginal means are still logged as noted previously.

Table 5: Estimated marginal means derived from inverse probability of treatment-weighted estimates of the average effect of a second birth on log cumulative work hours (in 1,000s), log cumulative earnings (in \$10,000s), and log hourly wages, at age 45, by two-year birth intervals, interacted with college degree

	Log cumulative work hours	Log cumulative earnings	Log wages
Panel 1: 1–2-year birth interval			
No college degree	3.40 (0.05)	2.95 (0.08)	2.49 (0.06)
College degree	3.27 (0.26)	3.54 (0.34)	2.79 (0.21)
Panel 2: 3–4-year birth interval			
No college degree	3.44 (0.06)	3.03 (0.09)	2.46 (0.08)
College degree	3.74 (0.04)	4.02 (0.05)	2.76 (0.14)
Panel 3: 5–6-year birth interval			
No college degree	3.54 (0.07)	3.10 (0.07)	2.65 (0.04)
College degree	3.72 (0.05)	3.96 (0.09)	2.57 (0.27)
Panel 4: 7–8-year birth interval			
No college degree	3.28 (0.11)	2.85 (0.13)	2.70 (0.06)
College degree	3.76 (0.09)	3.94 (0.16)	3.21 (0.19)

Notes: Estimated marginal means shown with delta-method standard errors.

Looking first at cumulative work hours in Column 1, it appears that women with a college degree have higher work hours than women without a college degree, except for women with a short birth interval. An examination of the Wald test results indicates that for all but the 1–2-year birth interval the differences between these groups are statistically significant ($p < .001$). Thus, for cumulative work hours, the hypothesis holds; college-educated women with short birth intervals do not have an advantage over women with the same birth interval and no college degree, but at the longer birth intervals having a college degree is advantageous.

Turning to cumulative earnings in Column 2, we see that college-educated mothers appear to always have higher cumulative earnings, regardless of birth interval. But an examination of the Wald test results indicates that among women with a birth interval of 1–2 years this difference is not statistically significant. For the other birth intervals, college-educated women have statistically significant advantages ($p < .001$ for 3–4 years;

$p < .05$ for 5–6 years; $p < .01$ for 7–8 years). Thus, the results for cumulative earnings also support the hypothesis: Women with short intervals have similar outcomes regardless of education, whereas at longer birth intervals the college-educated women are advantaged.

Turning finally to log wages in Column 3, in models of working women only, we see limited evidence for the hypothesis. Looking at the estimated marginal means descriptively, we can see that college-educated women still appear to be advantaged at most birth intervals, but except for the 7–8-year birth interval the Wald test results indicate that these are not statistically significant differences. For the 7–8-year birth interval college-educated women have statistically significantly higher log wages at age 45 than those without a college degree ($p < .05$). Thus, the results for log wages also support the hypothesis, though it is only at the 7–8-year birth interval where the interaction between spacing and a college education is significant.

Overall, there is very good evidence that having a college degree has a positive effect on midlife labor market outcomes, and there is consistent evidence that it meaningfully modifies the effects of birth spacing on outcomes. For women with the shortest birth intervals, postponement of a first birth to at least age 30 helped to reduce the midlife penalties of having a second child, but having a college degree does not. On the other hand, at longer intervals having a college degree reduces penalties more consistently than postponement of a first birth. These results suggest that postponement and education work differently to attenuate the motherhood penalty.

4.5 Description of supplemental analyses

To assess the robustness of the results, I conducted a number of sensitivity analyses. I estimated models with alternative birth interval measures, and I also estimated models with additional controls or slightly modified samples. While patterns of results tend to be consistent across specifications, there are some differences. For example, when the sample is limited to women with no more than two children the cumulative penalties seen in the main models are significantly reduced in magnitude and are no longer statistically significant, suggesting that the cumulative penalties may be driven primarily by women with more than two spells out of the labor force for childbearing. Additional discussion of the sensitivity analysis can be found in the Appendix.

5. Discussion

In this study I set out to examine whether different birth-spacing patterns were associated with the accumulation of motherhood penalties in the labor market at midlife for US women. While spacing is mentioned in the early economic research linking childbearing and labor force participation for mothers, it has been understudied. Informed by these early theories and contemporary research, I tested three hypotheses. Hypothesis 1 states that the magnitude of the motherhood penalty for a second birth will decline as the birth interval increases. The results of models testing Hypothesis 1 provide evidence of a U-shape rather than a linear decline. For cumulative work hours and cumulative earnings, women with the 1–2-year and 7–8-year birth intervals have the largest penalties. For log wages, women with birth intervals of 1–2 and 3–4 years have larger penalties than women with longer birth intervals. Thus, Hypothesis 1 was not supported for cumulative outcomes but is somewhat supported for log wages.

Because the literature has emphasized the importance of human capital accumulation for reducing women's motherhood penalties in the labor force, I considered the possibility that the results for Hypothesis 1 might obscure heterogeneity on the basis of human capital. Before testing for moderation, I first tested Hypothesis 2. Hypothesis 2a states that women who postpone a first birth to at least age 30 will have smaller motherhood penalties than those who have first births before age 30. Hypothesis 2b states that women with college degrees will have smaller motherhood penalties than those without college degrees. There was strong evidence for both parts of Hypothesis 2. As theorized in the literature and demonstrated in empirical studies of US samples that did not assess the role of birth spacing or examine midlife outcomes, postponement of a first birth to age 30 or later is positively associated with labor market outcomes at age 45. This is most likely because women who postpone the first birth have ample time to accumulate general human capital through education and/or job- or employer-specific human capital in the labor market prior to childbearing. Both types of human capital have the ability to improve women's employment position on return to the labor market following the birth of a child, compared to mothers with less human capital. Similarly, women with at least a college degree at the time of the first birth had significantly better labor market outcomes at midlife than women without a college degree, providing further evidence of the importance of human capital, all of which is consistent with past research on the motherhood penalty. Notably, these results, while consistent with those from other studies using US samples, contrast with the findings of Karimi (2014) for Sweden, suggesting that welfare state context and family policy may play an important role in how birth-timing and birth-spacing patterns affect women's midlife labor market outcomes.

Having established that the postponement and education patterns seen in past literature persist for midlife outcomes, I next tested Hypothesis 3 to assess heterogeneity in motherhood penalties related to birth spacing. Hypothesis 3a states that at longer birth intervals women who postponed a first birth to at least age 30 will have smaller penalties than those who did not postpone to at least age 30. Hypothesis 3b states that at longer birth intervals women with college degrees will have smaller penalties than those without college degrees. The evidence for Hypothesis 3 is mixed. With regard to postponement, although there is good evidence that postponement of a first birth to at least age 30 has positive effects on midlife labor market outcomes, as noted in the discussion of Hypothesis 2, there is little evidence that it modifies the effect of birth spacing on outcomes in the way hypothesized. Postponement does help to reduce midlife labor market penalties of having a second child, but this appears to be true for women with shorter birth intervals more than for women with longer birth intervals. In this way, rather than amplifying the possible human capital benefits of a longer birth interval, postponement seems to help protect women with shorter birth intervals that might otherwise be disadvantageous. This contrasts with Karimi's (2014) findings for Sweden. But there is evidence consistent with Hypothesis 3 for education. Education has a positive effect on midlife labor market outcomes, and at longer birth intervals having a college degree reduces penalties more consistently than postponement of a first birth. For Sweden, Karimi (2014) found similar interactive results – longer birth intervals were especially beneficial for women with higher education.

There are limitations to the analysis, including those related to causal inference. The dynamic IPT weighting method simulates random assignment, but second births are still not randomly assigned to occur at different intervals. This is a limitation of the literature more generally, as one cannot assign women to give birth at particular times. Researchers have sometimes tried to address the nonrandomness of births using instrumental variable methods – for example, instrumenting using miscarriages – but these methods also have important limitations (Wilde, Batchelder, and Ellwood 2010). Furthermore, there is no common wisdom about which spacing intervals are better. Indeed, the literature suggests this may differ for different types of women. As such, instrumental variables would not obviously improve upon the IPT-weighted estimates. However, we cannot abandon the search for causal estimates because of these limitations. Rather, we must employ the best methods we have for causal inference and examine the sensitivity of results to different specifications. I have done this in the sensitivity analyses described in the Appendix, but in the case of observational data it is impossible to prove that the dynamic CIA always holds.

Although the analysis presented is quite comprehensive, the study has another important limitation. The effective sample size after implementing the necessary

restrictions is quite modest for estimating the effects of a second birth occurring at several possible time points and allowing for heterogeneity by age at first birth and education. Because of this limitation, the coefficients may be imprecisely estimated, in some cases masking what would otherwise be statistically significant results. Nevertheless, the main results are quite consistent across specifications.

Finally, there are limitations in terms of the variables available with which to predict a second birth. A number of other variables could influence the timing of a second birth, such as negotiations between partners, unobserved changes in job-related duties, and the “quality” of the first child. That is to say, if the first child is difficult to parent or is frequently ill mothers may delay having a second child, while if the first child is easy to parent mothers might decide to have a second child sooner than they had originally anticipated. If future researchers can leverage such variables, they may gain additional insight into the determinants of birth spacing. Additionally, “unexpected” or “unintended” birth intervals – arising, for example, from a mistimed pregnancy – may have different consequences for women’s economic outcomes than intended or planned birth intervals, and in this paper I cannot address that possibility. As such, I focused primarily on determining the individual economic consequences of different birth intervals regardless of the underlying mechanisms. Thus, this study serves as a first step in examining the underexplored nature of the relationship between birth spacing and women’s labor market outcomes.

6. Conclusions

The key research question of this paper is whether birth spacing affects the motherhood penalty. The answer to this question is complex. Overall, there is some evidence that, net of the factors associated with the likelihood of having children at particular birth intervals, women’s labor force outcomes at midlife are affected by the timing of the birth of the second child. In the overall models, it is the women with mid-range birth intervals who have the smallest penalties for cumulative outcomes, though some of them (3–4-year birth interval) have poor wage outcomes at age 45. This is somewhat consistent with Peltola’s (2004) findings for labor force participation in that she found long intervals to be negatively associated with return to the labor market. Results for the models interacted with postponement of a first birth to at least age 30 suggest that postponement is a more important factor for cumulative outcomes than birth spacing, though postponement does appear to help attenuate wage penalties for mothers with shorter birth intervals. The benefits of a college education, by contrast, seem to be greater for women with longer birth intervals compared to women with less education than they are for women with the shortest birth intervals. This is consistent with Troske

and Voicu's (2013) findings for full-time labor force participation in their more selective NLSY79 sample.

Based on the classical family economics literature, I hypothesized that penalties would be smaller for mothers with longer birth intervals, especially mothers who were older at the time of the first birth or who had greater human capital as indicated by education. This is because such patterns would minimize the loss of human capital for childbearing, especially job- or employer-specific capital. Postponing a first birth allows women greater opportunity to accumulate human capital before having children (Blackburn, Bloom, and Neumark 1993), and spacing births farther apart may facilitate a return to the labor market between births, thus reducing the depreciation of skills many women experience during their period(s) out of the labor market for childbearing (Mincer and Polachek 1974). While theories linking human capital accumulation and birth timing and spacing are fairly widely cited, the spacing component has rarely been tested.

Consistent with human capital theory, I find evidence that women who postpone first births to age 30 or older work significantly more by age 45 and have significantly higher cumulative earnings than women who did not postpone the first birth to at least age 30. Waiting until age 30 for a first birth allows the mother to accumulate further human capital and invest in wage growth before any exits from the labor market for childbearing. But spacing interacts with postponement only minimally, except for wages at age 45; for women at least age 30 at first birth, both long and short intervals are generally associated with positive outcomes, suggesting that postponement is the main factor, and not birth interval. Postponement is significantly related to wages at age 45 for women with birth intervals shorter than seven years. Overall, postponement of a first birth to at least age 30 appears to play a more significant role in positive labor market outcomes than birth spacing. This contrasts with Karimi's (2014) results from Sweden, which indicated that postponement was detrimental to women's long-run outcomes because it tended to result in shorter birth spacing, which reduced labor force participation and wages, especially for highly educated women. This suggests an important role for family policy and the welfare state context in women's fertility timing decisions and long-run labor market outcomes.

I further hypothesized, based on the modern motherhood penalty literature, that women with a college degree would experience smaller penalties than women without a college degree and that education would interact with birth spacing. The assumption was that the protective effect of a college degree for mothers' labor market outcomes (e.g., Amuedo-Dorantes and Kimmel 2005; Anderson, Binder, and Krause 2003) would extend beyond the first birth to also be protective for a second birth, regardless of when that second birth occurred in relation to the first. I found evidence in support of these hypotheses. Women with college degrees generally have better cumulative outcomes

than women without, and an examination of these relationships by birth-spacing interval indicates that the significant gaps between college-educated and non-college-educated women tend to be more evident for women with longer birth intervals. This is particularly true for log wages at age 45. The benefit of a long birth interval seen for highly educated women is consistent with findings from Karimi's (2014) recent research from Sweden, though in contrast to Brehm and Buchholz's (2014) results for western Germany. Highly educated women may be more likely to hold the types of jobs for which significant career benefits accrue by returning to the labor market between births, and spacing children farther apart allows these women to build their salaries for several years after the first birth before taking another temporary (or permanent) step away from the labor market. This appears to be possible in the United States in a way that is not possible in western Germany, due to either welfare state policies or cultural norms around child-rearing. Education seems to be more important as a moderator of birth-spacing effects on long-run outcomes than age at first birth, but overall the evidence of an important role for birth spacing for mothers' midlife labor market outcomes is somewhat limited and nonlinear.

The best outcomes are seen for women with a college degree or who postponed a first birth to at least age 30; for women without a college degree or with a first birth occurring at a younger age, having a second birth between three and six years after the first appears to produce the best outcomes by midlife. While the reasons for worse outcomes at shorter birth intervals may seem self-evident (e.g., having two very small children is demanding and may require a reduction in productivity or exiting the labor force), the reasons for worse outcomes at the 7–8-year birth interval compared to intervals longer than eight years are not obvious. Examining their outcomes descriptively, these women simply appear to work less than women with longer birth intervals. It is possible that this is because women with the longest birth intervals have less choice about how much to work; overall, women with birth intervals of nine or more years are more disadvantaged, with less education, a much younger age at first birth, and a more unstable work history. Perhaps the limited prospects of the mothers with the longest birth intervals are the reason why mothers with 7–8-year birth intervals (without a college education or postponement of a first birth to at least age 30) have fewer work hours and lower cumulative earnings (as a result of fewer work hours) by age 45 but do not have lower wages.

Future research should examine in more detail the mechanisms underlying the U-shaped pattern of outcomes seen for women without a college degree or postponement of a first birth. It may be particularly important to examine more closely the characteristics of women with long birth intervals. To what extent are their second births planned? Are they more likely to have children with more than one man? To what extent do their partners' characteristics potentially explain their midlife labor

market outcomes? Answering these types of questions may help us to understand why longer birth intervals do not appear to be universally advantageous for women, in contrast with theoretical predictions in the economics literature, given the US labor market context and welfare state policies. Future research should also examine the mechanisms underlying the postponement and education benefits that generally seem to relate to birth spacing in different ways for mothers' long-run outcomes. The results suggest that to best understand midlife labor market penalties it may be necessary to consider a variety of different mechanisms that may influence these outcomes net of one's propensity for a certain birth interval. Researchers studying the motherhood penalty should continue to consider the possibility of long-term labor market consequences, whether negative or positive, for women's well-being. It is important not to lose sight of the life course context, and the social and economic context, in which women make decisions about family formation and labor force participation, even though this can be challenging methodologically.

7. Acknowledgments

An earlier version of this article was presented at the 2010 Annual Meeting of the Population Association of America in Dallas, Texas. The author acknowledges support from a National Institute of Child Health and Human Development (NICHD) training grant to the Population Studies Center (T32 HD007339), a fellowship from the Institute for Research on Women and Gender, both at the University of Michigan, and support from a National Science Foundation (NSF)-American Sociological Association Postdoctoral Fellowship. NICHD and NSF had no involvement in the analysis or interpretation of the data presented in this paper. Yu Xie, Jeffrey Smith, Pamela Smock, Karin Martin, Brenda Volling, Alexandra Killewald, Elyse Jennings, Qing Lai, Steven Martin, Luke Miratrix, Kanya Godde, and members of the Quantitative Sociology Seminar at the University of Michigan provided helpful comments on previous versions of this article. The author would like to express gratitude to Cassandra Dorius for generously allowing her to use the multipartner fertility variable created by Dorius for her dissertation.

References

- Abadie, A. and Imbens, G.W. (2011). Bias-corrected matching estimators for average treatment effects. *Journal of Business and Economic Statistics* 29(1): 1–11. doi:10.1198/jbes.2009.07333.
- Aisenbrey, S., Evertsson, M., and Grunow, D. (2009). Is there a career penalty for mothers' time out? A comparison of Germany, Sweden, and the United States. *Social Forces* 88(2): 573–605. doi:10.1353/sof.0.0252.
- Albrecht, J.W., Edin, P., Sundström, M., and Vroman, S.B. (1999). Career interruptions and subsequent earnings: A reexamination using Swedish data. *The Journal of Human Resources* 34(2): 294–311. doi:10.2307/146347.
- Amuedo-Dorantes, C. and Kimmel, J. (2005). The motherhood wage gap for women in the United States: The importance of college and fertility delay. *Review of Economics of the Household* 3(1): 17–48. doi:10.1007/s11150-004-0978-9.
- Anderson, D.J., Binder, M., and Krause, K. (2003). The motherhood wage penalty revisited: Experience, heterogeneity, work effort, and work schedule flexibility. *Industrial and Labor Relations Review* 56(2): 273–294. doi:10.1177/001979390305600204.
- Becker, G.S. (1985). Human capital, effort, and the sexual division of labor. *Journal of Labor Economics* 3(1): S33–S58. doi:10.1086/298075.
- Bergemann, A. and Riphahn, R.T. (2015). Maternal employment effects of paid parental leave. Bonn: IZA Institute of Labor Economics (IZA discussion paper no. 9073).
- Blackburn, M.L., Bloom, D.E., and Neumark, D. (1993). Fertility timing, wages, and human capital. *Journal of Population Economics* 6(1): 1–30. doi:10.1007/BF00164336.
- Blau, F.D. and Kahn, L.M. (2013). Female labor supply: Why is the US falling behind? Bonn: IZA Institute of Labor Economics (IZA discussion paper no. 7140).
- Bloom, D.E. and Trussell, J. (1984). What are the determinants of delayed childbearing and permanent childlessness in the United States? *Demography* 21(4): 591–611. doi:10.2307/2060917.
- Blundell, R., Dearden, L., and Sianesi, B. (2005). Evaluating the effect of education on earnings: Models, methods, and results from the National Child Development

- Study. *Journal of the Royal Statistical Society* 168(3): 473–512. doi:10.1111/j.1467-985X.2004.00360.x.
- Bongaarts, J. (2002). The end of the fertility transition in the developed world. *Population and Development Review* 28(3): 419–443. doi:10.1111/j.1728-4457.2002.00419.x
- Brehm, U. and Buchholz, S. (2014). Is there a wrong time for a right decision? The impact of the timing of first births and the spacing of second births on women's careers. *Journal of Family Research* 26(3): 269–301. doi:10.3224/zff.v26i3.18989.
- Buckles, K.S. and Munnich, E.L. (2012). Birth spacing and sibling outcomes. *Journal of Human Resources* 47(3): 613–642. doi:10.1353/jhr.2012.0019.
- Budig, M.J. and England, P. (2001). The wage penalty for motherhood. *American Sociological Review* 66(2): 204–225. doi:10.2307/2657415.
- Bumpass, L.L., Rindfuss, R.R., and Janosik, R.B. (1977). Age and marital status at first birth and the pace of subsequent fertility. Madison: Center for Demography and Ecology (Center for Demography and Ecology working paper no. 77–1).
- Bureau of Labor Statistics (2013). Women in the labor force: A datebook. Washington, D.C.: United States Department of Labor (BLS Report 1040). www.bls.gov/cps/wlf-datebook-2012.pdf.
- Busso, M., DiNardo, J., and McCrary, J. (2014). New evidence on the finite sample properties of propensity score reweighting and matching estimators. *Review of Economics and Statistics* 96(5): 885–897. doi:10.1162/REST_a_00431.
- Centers for Disease Control (2015). Social determinants and eliminating disparities in teen pregnancy [electronic resource]. Atlanta: U.S. Department of Health and Human Services. <http://www.cdc.gov/teenpregnancy/prevent-teen-pregnancy/social-determinants-disparities-teen-pregnancy.htm>.
- Dorius, C. (2010). Does multipartnered fertility harm women over the long run? [PhD Thesis]. State College: The Pennsylvania State University.
- Dorius, C. (2012). New approaches to measuring multipartnered fertility over the life course. Ann Arbor: Population Studies Center, University of Michigan (PSC Research Report 12-769).
- Dye, J.L. (2008). Fertility of American women: 2006 [electronic resource]. Washington, D.C.: US Census Bureau. <http://www.census.gov/prod/2008pubs/p20-558.pdf>.

- Fitzenberger, B., Sommerfeld, K., and Steffes, S. (2013). Causal effects on employment after first birth: A dynamic treatment approach. Bonn: IZA Institute of Labor Economics (IZA discussion paper no. 7438).
- Gangl, M. and Ziefle, A. (2009). Motherhood, labor force behavior, and women's careers: An empirical assessment of the wage penalty for motherhood in Britain, Germany, and the United States. *Demography* 46(2): 341–369. doi:10.1353/dem.0.0056.
- Guzzo, K.B. and Hayford, S. (2011). Fertility following an unintended first birth. *Demography* 48(4): 1493–1516. doi:10.1007/s13524-011-0059-7.
- Hamilton, B.E. and Martin, J.A. (2013). *Historical and recent trends in childbirth in the United States*. Paper presented at the Institute on Medicine and National Research Council workshop on research issues in the assessment of birth settings, Washington, D.C., March 6–7, 2013.
- Han, W., Ruhm, C.J., Waldfogel, J., and Washbrook, E. (2008). The timing of mothers' employment after childbirth. *Monthly Labor Review* 131(6): 15–27.
- Heckman, J.J., Ichimura, H., Smith, J., and Todd, P.E. (1998). Characterizing selection bias using experimental data. *Econometrica* 66(5): 1017–1098. doi:10.3386/w6699.
- Heckman, J.J., Lochner, L.J., and Todd, P.E. (2003). Fifty years of Mincer earnings regressions. Cambridge: National Bureau of Economic Research (NBER working paper no. 9732). doi:10.3386/w9732.
- Herr, J.L. (2012). Measuring the effect of the timing of first birth [unpublished manuscript]. Berkeley: University of California, Berkeley http://emlab.berkeley.edu/users/webfac/moretti/e251_fl2/herr.pdf.
- Hoffman, S.D. (1998). Teenage childbearing is not so bad after all... or is it? A review of the new literature. *Family Planning Perspectives* 30(5): 236–243. doi:10.2307/2991610.
- Hotz, V.J., Klerman, J.A., and Willis, R.J. (1997). The economics of fertility in developed countries. In: Rosenzweig, M.R. and Stark, O. (eds.). *Handbook of population and family economics*. Amsterdam: Elsevier Science: 275–347. doi:10.1016/S1574-003X(97)80024-4.
- Karimi, A. (2014). The spacing of births and women's subsequent earnings: Evidence from a natural experiment. Uppsala: IFAU Institute for Evaluation of Labour Market and Education Policy (IFAU working paper 2014: 18).

- Killewald, A. and Gough, M. (2013). Does specialization explain marriage and parenthood penalties and premiums? *American Sociological Review* 78(3): 477–502. doi:10.1177/0003122413484151.
- Laughlin, L. (2011). Maternity leave and employment patterns, 1961–2008. Washington, D.C.: US Census Bureau (Current population reports P70–128).
- Lechner, M. and Wiehler, S. (2011). Kids or courses? Gender differences in the effects of active labor market policies. *Journal of Population Economics* 24(3): 783–812. doi:10.1007/s00148-009-0267-2.
- Loughran, D.S. and Zissimopoulos, J. (2009). Why wait? The effect of marriage and childbearing on the wages of men and women. *Journal of Human Resources* 44(2): 326–349. doi:10.1353/jhr.2009.0032.
- Lundberg, S. and Rose, E. (2000). Parenthood and the earnings of married men and women. *Labour Economics* 7(6): 689–710. doi:10.1016/S0927-5371(00)00020-8.
- Manning, W.D. (1995). Cohabitation, marriage, and entry into motherhood. *Journal of Marriage and Family* 57(1): 191–200. doi:10.2307/353827.
- Manning, W.D., Brown, S.L., and Stykes, B. (2015). Trends in births to single and cohabiting mothers, 1980–2013. Bowling Green: National Center for Family and Marriage Research (NCFMR Family Profiles FP-15-03).
- Martin, S.P. (2000). Diverging fertility among U.S. women who delay childbearing past age 30. *Demography* 37(4): 523–533. doi:10.1353/dem.2000.0007.
- Miller, A.R. (2011). The effects of motherhood timing on career path. *Journal of Population Economics* 24(3): 1071–1100. doi:10.1007/s00148-009-0296-x.
- Mincer, J. and Polachek, S. (1974). Family investments in human capital: Earnings of women. In: Schultz, T.W. (ed.). *Marriage, family, human capital, and fertility*. Ann Arbor: UMI: 76–110.
- National Partnership for Women and Families (2013). A look at the U.S. Department of Labor’s 2012 Family and Medical Leave Act employee and worksite surveys [electronic resource]. Washington, D.C.: National Partnership for Women and Families. <http://www.nationalpartnership.org/research-library/work-family/fmla/dol-fmla-survey-key-findings-2012.pdf>.
- Noonan, M. and Corcoran, M. (2004). The mommy track and partnership: Temporary delay or dead end? *The Annals of the American Academy of Political and Social Science* 596(1): 130–150. doi:10.1177/000271620459600106.

- Peltola, P.K. (2004). Mothers' level of attachment to the labor market following the birth of a second child. [PhD thesis]. College Park: University of Maryland, Department of Sociology.
- Pollak, R.A. (2005). Bargaining power in marriage: Earnings, wage rates, and household production. Cambridge: National Bureau of Economic Research (NBER working paper no. 11239).
- Rindfuss, R.R., Morgan, S.P., and Offutt, K. (1996). Education and the changing age pattern of American fertility, 1963–1989. *Demography* 33(3): 277–290. doi:10.2307/2061761.
- Robins, J. (1999). Association, causation, and marginal structural models. *Synthese* 121(1–2): 151–179. doi:10.1023/A:1005285815569.
- Rosenbaum, P.R. and Rubin, D.B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika* 70(1): 41–55. doi:10.1093/biomet/70.1.41.
- Ross, S.G. (1974). The timing and spacing of births and women's labor force participation: An economic analysis. Cambridge: National Bureau of Economic Research (NBER working paper no. 30). doi:10.3386/w0030.
- Rubin, D.B. (1974). Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of Educational Psychology* 66(5): 688–701. doi:10.1037/h0037350.
- Rubin, D.B. (1978). Bayesian inference for causal effects: The role of randomization. *Annals of Statistics* 6(1): 34–58. doi:10.1214/aos/1176344064.
- Sampson, R.J., Laub, J.H., and Wimer, C. (2006). Does marriage reduce crime? A counterfactual approach to within-individual causal effects. *Criminology* 44(3): 465–508. doi:10.1111/j.1745-9125.2006.00055.x.
- Sharkey, P. and Elwert, F. (2011). The legacy of disadvantage: Multigenerational neighborhood effects on cognitive ability. *American Journal of Sociology* 116(6): 1934–1981. doi:10.1086/660009.
- Sianesi, B. (2004). An evaluation of the Swedish system of active labor market programs in the 1990s. *The Review of Economics and Statistics* 86(1): 133–155. doi:10.1162/003465304323023723.
- Stone, P. and Lovejoy, M. (2004). Fast-track women and the 'choice' to stay home. *The Annals of the American Academy of Political and Social Science* 596(1): 62–83. doi:10.1177/0002716204268552.

- Troske, K.R. and Voicu, A. (2013). The effect of the timing and spacing of births on the level of labor market involvement of married women. *Empirical Economics* 45(1): 483–521. doi:[10.1007/s00181-012-0620-2](https://doi.org/10.1007/s00181-012-0620-2).
- Waldfogel, J. (1997). The effect of children on women's wages. *American Sociological Review* 62(2): 209–217. doi:[10.2307/2657300](https://doi.org/10.2307/2657300).
- Wilde, E.T., Batchelder, L., and Ellwood, D. (2010). The mommy track divides: The impact of childbearing on wages of women of differing skill levels. Cambridge: National Bureau of Economic Research (NBER working paper no. 16582).
- Wodtke, G.T., Harding, D.J., and Elwert, F. (2011). Neighborhood effects in temporal perspective. *American Sociological Review* 76(5): 713–736. doi:[10.1177/0003122411420816](https://doi.org/10.1177/0003122411420816).
- Yamaguchi, K. and Ferguson, L.R. (1995). The stopping and spacing of childbirths and their birth-history predictors: Rational-choice theory and event-history analysis. *American Sociological Review* 60(2): 272–298.

Appendix: Sensitivity analysis

I conducted several sets of sensitivity analyses. First, I estimated models with alternative intervals. For alternative intervals, I estimated models with intervals of the following patterns: a) 1, 2, 3–4, 5–6, 7+ years; b) 1–2, 3–4, 5+ years (i.e., a truncated version of the main model); c) 1–2 years vs. 3–4, 5–6, and 7+ years using propensity score matching; and d) continuous birth interval. Estimates are generally consistent with the main models (results for a, c, and d are shown in Table A-7). Separating the 1-year and 2-year birth intervals indicates that the negative results are driven primarily by women with 2-year birth intervals, though the 1-year birth interval is negative and statistically significant for cumulative earnings. Directly comparing women with 1–2-year birth intervals to those with 3–4-, 5–6-, or 7–8-year birth intervals indicates that for all cumulative outcomes women with birth intervals of 1–2 years appear to have worse outcomes than women with longer intervals. Finally, using a continuous measure of birth interval, results indicate that for all outcomes a longer birth interval is positively associated with the outcome. While this result is consistent with the main models and human capital explanations, these models assume the relationship between birth interval and outcomes is linear, which is a strong assumption and one that the main analysis suggests may be faulty.

Second, I estimated models with additional controls or modified samples. I estimated models of the following: a) controlling for MPF²⁰; b) excluding women with MPF; c) controlling for actual work experience; d) controlling for potential work experience; e) outcomes at age 30, 35, and 40; f) outcomes five, ten, and 15 years after the first birth; g) controlling for number of siblings; h) limiting the sample to women with only two births; i) controlling for the number of children; j) restricting the comparison to the next birth interval (e.g., 1–2 vs. 3–4, 3–4 vs. 5–6), and k) blacks only and Hispanics only. Results for a, c, e, f, h, and k are shown in Table A-7. Results for f that include the interaction for age at first birth are shown in Table A-8. Controlling for or excluding women with MPF results in findings that are consistent with the main models. This is also true for models including work experience and models that control for the number of siblings a woman has. In models estimated with outcomes at ages 30, 35, and 40, while the statistical significance sometimes varies from the main models, the patterns and magnitudes of the coefficients are quite consistent.

In models estimated with outcomes five years after the first birth, there are large penalties to cumulative work hours and cumulative earnings for women with birth intervals of four or fewer years; there are also wage penalties for the 3–4-year interval.

²⁰ This variable was provided by Cassandra Dorius (personal communication) who created it as part of her dissertation (Dorius 2010). Further details about how she created the variable can be found in a separate paper (Dorius 2012).

This is not unexpected, since these women would have had significant time out of the labor market in this five-year period. These penalties remain ten years after the first birth, but by 15 years after the first birth the cumulative penalties are only significant for women with a 1–2-year birth interval, though wage penalties are seen for women with 3–4-year and 5–6-year birth intervals. Thus, penalties, particularly for cumulative outcomes, appear to develop early and persist over time, especially for women with 1–2-year birth intervals. An examination of the models with interactions for age at first birth in Table A-8 indicates that the patterns in the main models hold even at five, ten, and 15 years after first birth with one exception. At ten years after the first birth, women with birth intervals of 5–6 years who postponed the first birth to at least age 30 have higher wages than women with the same interval who had the first child before age 30, which is a result not seen in the main models. Overall, the results indicate that the role of postponement of a first birth to at least age 30 is present even five years after a first birth.

Women who have more than two children may differ in important ways from women with only two children, and it could be the case that these differences may not be fully accounted for in the main models. In comparison to the results shown in the main models in Table 2, after controlling for the number of children the woman has at age 45, the coefficients are nearly identical and the statistically significant coefficients in Table 2 remain significant. However, when women with more than two children are excluded from the analysis, only the negative effect on wages for the 3–4-year birth interval remains significant. The interpretation of this result must be approached cautiously, however, since limited sample size remains an issue for statistical power; in the models limited to women with only two births, nearly half of the analytic sample is lost. A close examination of the coefficients and standard errors suggests that statistical power is likely to be a concern in models of wages, but for cumulative outcomes there is a sizeable reduction in the magnitude of the coefficients when the sample is restricted to women with only two children. In a sense this is not surprising: while these women have had two child-related exits from the labor force, the excluded women have had three or more child-related exits from the labor force and should thus be expected to have lower cumulative work hours and earnings. Comparing the outcomes of a woman with one of the birth intervals only to women with births in the next birth interval (i.e., 1–2 vs. 3–4, 3–4 vs. 5–6, 5–6 vs. 7–8) produces results consistent with the main models, whereby shorter birth intervals appear more detrimental to women's labor market outcomes.

To determine whether the results hold if performed separately for ethnic population subgroups, I estimated models separately for blacks and Hispanics and compared the results to the main models. Overall, the results for blacks are consistent with the main models, but the magnitudes of the coefficients are larger, indicating that

black women experience greater penalties than the sample as a whole. The interaction model results (not shown) are also consistent except for the 5–6-year birth interval, where, for example, women with first births at age 30 or later experience cumulative hours penalties and wage penalties. The results for Hispanic women are generally consistent with the main models, but the standard errors are larger. The interaction model results (not shown) indicate that Hispanic women may see penalties for having a short birth interval if they postponed their first birth to at least age 30 compared to other women with short birth intervals who had their first child before age 30.

Third, I conducted sensitivity analyses similar to Fitzenberger, Sommerfeld, and Steffes (2013). I estimated OLS models to compare the results to the ATT estimates. The point of this comparison was to examine whether the main ATT results were sensitive to the estimation strategy employed. Results are shown in Table A-5. The patterns of coefficients are similar across models, though in most cases the magnitudes of the coefficients for the IPT-weighted models are larger than the magnitudes of the OLS coefficients. Because the effect of a second birth at a particular interval varies by some of the characteristics X controlled in the models and used to predict the propensity scores (age at first birth and education), I rely on the IPT-weighted estimates.

Fourth, I examined the possibility that becoming unemployed may induce fertility, in line with the findings of Lechner and Wiehler (2011) for Austria. I find no evidence that being currently unemployed or being unemployed in the past year is positively associated with a second birth. In the few cases where being unemployed is significantly related to having a second birth the association is negative.

Finally, I reestimated the models using propensity score matching (paired and four nearest neighbors (4NN) as opposed to IPT weighting.²¹ The results of these models are shown in Table A-6. In these models, the propensity score-matched estimates are fairly similar to the IPT-weighted estimates though somewhat more conservative. Focusing on the 4NN propensity score-matched models indicates that women with birth intervals of less than two years work about 10% less by age 45 ($p < .01$; compared to 14% and significant at $p < .05$ in the IPT model) than women with longer birth intervals. They also have about 20% lower cumulative earnings ($p < .001$; compared to 27% and significant at $p < .001$ in the IPT model). However, the results for wages at age 45 remain substantively unchanged, with the estimates in the matched models being slightly smaller in magnitude than those from the IPT models.

²¹ Busso, DiNardo, and McCrary (2014) argue that when overlap is good, reweighting is more effective than matching and that in finite samples the reweighting estimator typically outperforms or is equally competitive with matching estimators. I examined the overlap plots and determined that overlap was good. However, Busso, DiNardo, and McCrary (2014) also recommend testing multiple estimators, so I estimated matching models. But Sampson, Laub, and Wimer (2006) caution that propensity score matching is not appropriate in the context of time-varying confounding, which further strengthens the case for using reweighting in lieu of matching.

Table A-1: Sample-weighted descriptive statistics comparing the analytic sample to the sample before restrictions

	Analytic sample mean (SD) ^a	Overall sample ^b mean (SD) ^a
Panel 1: Individuals		
Number of children wanted (measured in 1979)	2.57 (1.35)	2.53 (1.51)
AFQT in 1980	48.36 (27.40)	47.51 (27.84)
Race		
Nonblack, non-Hispanic	78.12%	79.67%
Hispanic	7.07%	6.32%
Black	14.82%	14.01%
Age at first birth	23.74 (5.03)	23.99 (5.67)
N	1,533	6,283
Panel 2: Person-years		
Education		
Less than high school	11.95%	15.07%
High school	45.27%	41.62%
Some college	24.29%	24.31%
College	11.78%	12.06%
More than college	6.70%	6.93%
Marital status		
Never married	24.92%	33.27%
Married	60.50%	51.78%
Separated/divorced	14.58%	14.95%
Occupation is professional/managerial	0.31	0.30
Sector		
Private sector	78.05%	78.35%
Government	15.54%	14.68%
Self-employed/working without pay	6.42%	6.97%
Tenure with employer (years)	4.26 (5.21)	3.89 (4.95)
Part-time worker	0.24	0.23
Not working in the labor force	0.20	0.25
Family economic resources (household income minus respondent income)	36,766.54 (42,781.42)	32,433.92 (42,926.29)
Respondent resides in the South	0.34	0.36
Respondent lives in urban area	0.73	0.73
Hourly wage of respondent, if working	15.31 (9.17)	12.15 (10.33)
Annual income of respondent, if working	26,355.72 (21,584.29)	24,024.87 (22,096.63)
Cumulative hours of work experience	18,327.13 (15,976.87)	16,735.52 (15,701.32)
Cumulative earnings of respondent	202,352.50 (228,123.30)	279,455.30 (352,291.10)
N	36,103	119,480
Panel 3: At age 45		
Hourly wage of respondent, if working	18.21 (10.99)	18.21 (11.83)
Cumulative earnings of respondent	415,257.00 (281,062.00)	662,016.10 (457,545.30)
Cumulative hours of work experience	386,42.29 (14,629.05)	28,939.26 (11,692.44)
Not working in the labor force	0.08	0.05
N^c	1,533	3,736

Notes: ^a Standard deviations not provided for proportions, for which the entire distribution is provided. ^b Sample sizes for the overall sample are variable across statistics because this sample includes observations on individuals with missing data on variables of interest. ^c Valid sample size for cumulative work hours in the original sample is 3,163.

Table A-2: Coefficients from logit models predicting the propensity to have a second birth at a given interval, by two-year birth intervals

	1-2-yr. birth interval	3-4-yr. birth interval	5-6-yr. birth interval	7-8-yr. birth interval
Education (omitting high school)				
Less than high school	0.48 (0.31)	1.69 (0.37)***	1.98 (0.59)**	1.46 (0.86) [†]
Some college	-1.04 (0.39)**	-2.39 (0.46)***	-1.75 (0.60)**	-2.16 (1.17) [†]
College	-2.50 (0.62)***	-4.80 (0.77)***	-3.24 (1.12)**	-2.58 (2.46)
More than college	-3.15 (0.84)***	-7.51 (1.08)***	-5.98 (1.54)***	-4.22 (3.11)
Potential experience	-0.33 (0.08)***	-1.08 (0.11)***	-0.90 (0.16)***	-0.93 (0.25)***
Education × potential experience	0.04 (0.02) [†]	0.04 (0.02) [†]	0.01 (0.03)	0.01 (0.05)
Lagged sector (omitting private industry)				
Government	-0.31 (0.29)	0.43 (0.30)	-0.81 (0.43) [†]	0.25 (0.54)
Self-empl./working without pay	-0.30 (0.42)	0.27 (0.48)	-1.90 (0.86) [†]	0.06 (1.05)
Lagged professional/managerial occupation	0.30 (0.24)	-0.04 (0.26)	0.44 (0.35)	0.85 (0.56)
AFQT quartile 1	0.03 (0.03)	-0.03 (0.03)	-0.01 (0.04)	-0.13 (0.05) [†]
AFQT quartile 2	-0.001 (0.02)	0.01 (0.02)	-0.01 (0.03)	-0.05 (0.05)
AFQT quartile 3	-0.03 (0.01) [†]	-0.01 (0.01)	-0.01 (0.02)	0.08 (0.03)**
AFQT quartile 4	0.02 (0.01) [†]	-0.004 (0.01)	0.03 (0.02)	-0.09 (0.04) [†]
Age at first birth	0.48 (0.18)**	1.46 (0.23)***	0.82 (0.26)**	1.08 (0.42) [†]
(Age at first birth) ²	-0.002 (0.003)	-0.007 (0.004)	0.004 (0.005)	-0.001 (0.01)
South	0.09 (0.17)	-0.45 (0.19)*	-0.22 (0.26)	-0.98 (0.45) [†]
Urban residence	-0.13 (0.19)	-0.24 (0.21)	0.11 (0.31)	0.16 (0.51)
Marital status (omitting never married/widowed)				
Married	0.95 (0.43) [†]	0.67 (0.36) [†]	1.42 (0.76) [†]	1.65 (0.69) [†]
Divorced	0.11 (0.65)	-1.64 (0.55)**	-0.51 (0.88)	-0.002 (0.82)
Race (omitting nonblack, non-Hispanic)				
Hispanic	0.96 (0.53) [†]	-0.69 (0.54)	-0.34 (0.95)	0.07 (0.99)
Black	0.66 (0.48)	-0.46 (0.41)	0.12 (0.78)	0.52 (0.78)
Marital status × race				
Married × Hispanic	-1.37 0.56)*	0.68 (0.58)	0.49 (1.03)	-0.33 (1.11)
Married × Black	-0.94 (0.51) [†]	0.06 (0.46)	-0.41 (0.81)	0.87 (0.89)
Divorced × Hispanic	-0.94 (0.97)	2.57 (0.86)**	-1.09 (1.50)	-1.93 (1.82)
Divorced × Black	-1.33 (1.08)	1.50 (0.70)*	1.39 (1.04)	-0.18 (1.16)
Number of children wanted (1979)	0.23 (0.05)***	0.01 (0.06)	0.06 (0.09)	0.35 (0.14) [†]
Traditional gender role attitudes (factor)	0.35 (0.10)**	-0.11 (0.12)	-0.06 (0.18)	0.25 (0.23)
Family economic resources (h'hold income minus resp. income)	-0.01 (0.03)	-0.01 (0.03)	-0.03 (0.04)	-0.10 (0.07)
Year	-0.04 (0.04)	0.01 (0.04)	-0.07 (0.06)	-0.02 (0.08)
Pseudo R ²	0.0933	.2429	.2522	.3265
Median propensity score	0.22	0.48	0.45	0.40
Mean propensity score	0.24	0.50	0.47	0.43
N	1,533	1,163	581	309

Notes: [†] p<.10; * p<.05; ** p<.01; *** p<.001.

Table A-3: Motherhood penalty pooled across years and at age 45, by parity, OLS estimates of log cumulative work hours (in 1,000s), log cumulative earnings (in \$10,000s), and log wages

	Any parity (including no births) OLS with covariates	At least one birth OLS with covariates	At least two births OLS with covariates
Panel 1: Outcomes pooled across years			
Log cumulative work hours			
First child	0.17 (0.02)***		
Second child	-0.07 (0.02)**	-0.004 (0.02)	
Third child	-0.21 (0.03)***	-0.19 (0.03)***	-0.12 (0.03)***
Log cumulative earnings			
First child	0.13 (0.03)***		
Second child	-0.04 (0.03)	0.03 (0.03)	
Third+ child	-0.31 (0.05)***	-0.27 (0.05)***	-0.21 (0.05)***
Log wages, if working			
First child	0.02 (0.02)		
Second child	-0.05 (0.02)*	-0.04 (0.02)*	
Third+ child	-0.08 (0.03)**	-0.08 (0.03)**	-0.09(0.03)**
Panel 2: Outcomes at age 45			
Log cumulative work hours			
First child	0.19 (0.22)		
Second child	-0.07 (0.03)*	-0.07 (0.03)**	
Third+ child	-0.22 (0.03)***	-0.22 (0.03)***	-0.17 (0.03)***
Log cumulative earnings			
First child	-0.15 (0.35)		
Second child	-0.01 (0.05)	-0.02 (0.05)	
Third+ child	-0.27 (0.05)***	-0.27 (0.05)***	-0.21 (0.05)***
Log wages, if working			
First child	0.05 (0.44)		
Second child	-0.02 (0.06)	-0.02 (0.06)	
Third+ child	-0.10 (0.06)	-0.10 (0.06)	-0.09 (0.06)

Notes: * p<.10; * p<.05; ** p<.01; *** p<.001. Models estimated with standard errors clustered at the individual level. Models control for education, potential experience, age at first birth and its square, marital status, race/ethnicity, lagged sector, residence in the South, family economic resources (total household income minus respondent income), quartiles of AFQT score, lagged professional/managerial status, number of children desired by women in 1979, a factor variable indicating traditional gender role attitudes, urban residence, year, an interaction between education and potential experience, and an interaction between marital status and race. For models of any parity, pooled, N=62,867 for log cumulative earnings, N=59,276 for log cumulative work hours, and N=47,244 for log wages. For models of any parity at age 45, N=2,684 for log cumulative earnings, N=2,671 for log cumulative work hours, and N=2,168 for log wages. For models of women with at least one birth, pooled, N=52,650 for log cumulative earnings, N=49,603 for log cumulative work hours, and N=38,865 for log wages. For models of women with at least one birth at age 45, N=2,248 for log cumulative earnings, N=2,237 for log cumulative work hours, and N=1,808 for log wages. For models of women with at least two births, pooled, N=36,100 for log cumulative earnings, N=34,329 for log cumulative work hours, and N=27,637 for log wages. For models of women with at least two births at age 45, N=1,533 for log cumulative earnings and log cumulative work hours, and N=1,410 for log wages.

Table A-4: Preliminary fixed-effects models including women without children, estimating the average effect of a birth on log cumulative work hours (in 1000s), log cumulative earnings (in \$10000s), and log hourly wages, pooled across years

	Log cumulative work hours	Log cumulative earnings	Log hourly wages
First child	0.37 (0.02)***	0.43 (0.02)***	-0.19 (0.02)***
Second child	0.05 (0.02)**	0.04 (0.02) [†]	-0.14 (0.02)***
Third+ child	-0.02 (0.03)	-0.12 (0.03) [†]	-0.001 (0.03)
Pseudo R ²	0.4824	0.4737	0.3574
Rho	0.75	0.67	0.26
N	2,670	2,683	2,683

Notes: [†] p<.10; * p<.05; ** p<.01; *** p<.001. Models control for potential experience, lagged sector, residence in the South, family economic resources, lagged professional/managerial status, urban residence, year, an interaction between education and potential experience, and interactions between marital status and race.

Table A-5: OLS and inverse probability of treatment-weighted estimates of the average effect of a second birth on log cumulative work hours (in 1000s), log cumulative earnings (in \$10000s), and log hourly wages, at age 45

	IPT-weighted models			OLS models		
	Log cumulative work hours	Log cumulative earnings	Log wages	Log cumulative work hours	Log cumulative earnings	Log wages
1–2-year birth interval (N=1,533, 1,410)						
	-0.15 (0.06)*	-0.31 (0.09)***	-0.12 (0.06) [†]	-0.09 (0.04)*	-0.13 (0.05)**	-0.09 (0.08)
3–4-year birth interval (N=1,163, 1,071)						
	-0.02 (0.07)	-0.16 (0.11)	-0.22 (0.08)**	0.002 (0.03)	-0.02 (0.05)	-0.14 (0.07) [†]
5–6-year birth interval (N=581, 527)						
	0.04 (0.07)	-0.12 (0.10)	-0.05 (0.07)	-0.03 (0.05)	-0.09 (0.08)	-0.03 (0.09)
7–8-year birth interval (N=309, 279)						
	-0.22 (0.12) [†]	-0.36 (0.17)*	0.14 (0.10)	-0.07 (0.08)	-0.10 (0.11)	-0.01 (0.11)

Notes: [†] p<.10; * p<.05; ** p<.01; *** p<.001. Models include women not working at age 44 or 45 whose wages take a value of zero. Robust standard errors, clustered at the individual level, are estimated for IPT-weighted models. OLS models estimated with standard errors clustered at the individual level. OLS models control for education, potential experience, age at first birth and its square, marital status, race/ethnicity, lagged sector, residence in the South, family economic resources (total family income minus respondent income), quartiles of AFQT score, lagged professional/managerial status, number of children desired by women in 1979, a factor variable indicating traditional gender role attitudes, urban residence, year, an interaction between education and potential experience, and an interaction between marital status and race. Log wage model includes only women working at age 44 or 45, resulting in a smaller sample size.

Table A-6: Inverse probability of treatment-weighted estimates and propensity score-matched estimates of the average effect of a second birth on log cumulative work hours(in 1000s), log cumulative earnings (in \$10000s), and log hourly wages, at age 45

Log cumulative work hours			Log cumulative earnings			Log wages		
IPT-weighted	PSM-paired	PSM-4NN	IPT-weighted	PSM-paired	PSM-4NN	IPT-weighted	PSM-paired	PSM-4NN
1–2-year birth interval (N=1,533, 1,410)								
–0.15 (0.06)*	–0.16 (0.05)**	–0.11 (0.04)**	–0.31 (0.09)***	–0.26 (0.06)***	–0.22 (0.05)***	–0.12 (0.06)*	–0.04 (0.08)	–0.09 (0.06)
3–4-year birth interval (N=1,163, 1,071)								
–0.02 (0.07)	0.05 (0.06)	0.03 (0.05)	–0.16 (0.11)	0.01 (0.10)	–0.01 (0.08)	–0.22 (0.08)**	–0.15 (0.06)*	–0.13 (0.06)*
5–6-year birth interval (N=581, 527)								
0.04 (0.07)	–0.05 (0.06)	–0.04 (0.05)	–0.12 (0.10)	–0.17 (0.10)	–0.13 (0.08)	–0.05 (0.07)	0.18 (0.13)	–0.01 (0.08)
7–8-year birth interval (N=309, 279)								
–0.22 (0.12)*	–0.01 (0.17)	0.26 (0.26)	–0.36 (0.17)*	–0.14 (0.15)	0.01 (0.07)	0.14 (0.10)	0.17 (0.19)	0.19 (0.19)

Notes: † p<.10; * p<.05; ** p<.01; *** p<.001. Robust standard errors are estimated for IPT-weighted models. Propensity score-matched models are estimated using teffects in Stata. PSM-paired is simple nearest neighbor matching. PSM-4NN matches with the four nearest neighbors. The teffects program adjusts the standard errors to account for two-stage estimation. Log wage model includes only women working at age 44 or 45, resulting in a smaller sample size.

Table A-7: Results from selected supplementary analyses described in the Appendix

	Log cumulative work hours	Log cumulative earnings	Log wages
Panel 1: Main results			
1–2-year birth interval (N=1,533, 1,410)	–0.15 (0.06)*	–0.31 (0.09)***	–0.12 (0.06)*
3–4-year birth interval (N=1,163, 1,071)	–0.02 (0.07)	–0.16 (0.11)	–0.22 (0.08)**
5–6-year birth interval (N=581, 527)	0.04 (0.07)	–0.12 (0.10)	–0.05 (0.07)
7–8-year birth interval (N=309, 279)	–0.22 (0.12)*	–0.36 (0.17)*	0.14 (0.10)
Panel 2: Alternative intervals			
1-year birth interval (N=1,527, 1,404)	–0.08 (0.10)	–0.38 (0.16)*	–0.02 (0.12)
2-year birth interval (N=1,482, 1,362)	–0.13 (0.05)*	–0.32 (0.09)***	–0.10 (0.06)*
3–4-year birth interval (N=1,163, 1,071)	–0.02 (0.07)	–0.16 (0.11)	–0.22 (0.08)**
5–6-year birth interval (N=581, 527)	0.04 (0.07)	–0.12 (0.10)	–0.05 (0.07)
7–8-year birth interval (N=309, 279)	–0.22 (0.12)*	–0.36 (0.17)*	0.14 (0.10)

Table A-7: (Continued)

	Log cumulative work hours	Log cumulative earnings	Log wages
Panel 3: Alternative comparisons v. 1			
1–2-year birth interval vs. 3–4 year birth interval (N=951, 883)	-0.17 (0.06)**	-0.37 (0.09)***	-0.07 (0.07)
1–2-year birth interval vs. 5–6 year birth interval (N=642, 586)	-0.15 (0.07)*	-0.33 (0.10)**	-0.19 (0.07)*
Panel 4: Alternative comparisons v. 2			
1–2-year birth interval vs. 3–4 year birth interval (N=951, 883)	-0.17 (0.06)**	-0.37 (0.09)***	-0.07 (0.07)
3–4-year birth interval vs. 5–6 year birth interval (N=854, 792)	-0.06 (0.07)	-0.28 (0.11)*	-0.26 (0.08)**
5–6-year birth interval vs. 7–8 year birth interval (N=408, 369)	0.08 (0.09)	-0.10 (0.13)	-0.09 (0.09)
Panel 5: With control for multipartnered fertility			
1–2-year birth interval (N=1,501, 1,381)	-0.15 (0.06)**	-0.34 (0.09)***	-0.12 (0.06) [†]
3–4-year birth interval (N=1,142, 1,052)	-0.04 (0.07)	-0.19 (0.11) [†]	-0.22 (0.08)**
5–6-year birth interval (N=567, 514)	0.04 (0.08)	-0.12 (0.09)	-0.05 (0.06)
7–8-year birth interval (N=300, 270)	-0.24 (0.12) [†]	-0.39 (0.16)*	0.13 (0.10)
Panel 6: With control for actual work experience			
1–2-year birth interval (N=1,460, 1,343)	-0.14 (0.06)*	-0.28 (0.07)***	-0.11 (0.06) [†]
3–4-year birth interval (N=1,100, 1,013)	0.04 (0.07)	-0.04 (0.09)	-0.17 (0.07)*
5–6-year birth interval (N=491, 514)	0.004 (0.06)	-0.06 (0.09)	0.005 (0.07)
7–8-year birth interval (N=252, 270)	-0.15 (0.11)	-0.18 (0.13)	0.15 (0.10)
Panel 7: With outcomes measured at age 30			
1–2-year birth interval (N=1,523, 1,180)	-0.29 (0.08)***	-0.41 (0.11)***	-0.06 (0.05)
3–4-year birth interval (N=1,158, 913)	-0.18 (0.14)	-0.29 (0.19)	-0.12 (0.06)*
5–6-year birth interval (N=579, 459)	-0.005 (0.13)	-0.19 (0.14)	-0.15 (0.08) [†]
7–8-year birth interval (N=309, 246)	-0.35 (0.14)*	-0.45 (0.21)*	-0.12 (0.11)
Panel 8: With outcomes measured at age 35			
1–2-year birth interval (N=1,470, 1,242)	-0.21 (0.07)**	-0.39 (0.11)***	-0.09 (0.04)*
3–4-year birth interval (N=1,116, 956)	-0.09 (0.09)	-0.23 (0.14) [†]	-0.21 (0.06)**
5–6-year birth interval (N=559, 481)	0.04 (0.10)	-0.15 (0.11)	-0.06 (0.11)
7–8-year birth interval (N=295, 250)	-0.30 (0.14)*	-0.37 (0.21) [†]	-0.14 (0.12)
Panel 9: With outcomes measured at age 40			
1–2-year birth interval (N=1,361, 1,214)	-0.18 (0.07)**	-0.35 (0.10)**	-0.05 (0.05)
3–4-year birth interval (N=1,035, 931)	-0.06 (0.08)	-0.21 (0.13)	-0.11 (0.07)
5–6-year birth interval (N=509, 462)	0.06 (0.10)	-0.14 (0.11)	-0.01 (0.08)
7–8-year birth interval (N=273, 244)	-0.30 (0.14)*	-0.42 (0.20)*	-0.02 (0.14)
Panel 10: Regression estimates with predictor of continuous birth interval in years			
Birth interval in years (N=1,534, 1,411)	0.01 (0.007)*	0.03 (0.01)**	0.03 (0.01)*
Panel 11: With outcomes measured five years after first birth			
1–2-year birth interval (N=1,528, 1,076)	-0.25 (0.10)*	-0.42 (0.13)**	-0.06 (0.05)
3–4-year birth interval (N=1,161, 848)	-0.43 (0.18)*	-0.60 (0.23)**	-0.21 (0.10)*

Table A-7: (Continued)

	Log cumulative work hours	Log cumulative earnings	Log wages
Panel 12: With outcomes measured ten years after first birth			
1–2-year birth interval (N=1,506, 1,183)	–0.27 (0.10)**	–0.43 (0.12)**	–0.07 (0.05)
3–4-year birth interval (N=1,149, 910)	–0.25 (0.11)*	–0.36 (0.17)*	–0.15 (0.07)*
5–6-year birth interval (N=577, 444)	–0.04 (0.14)	–0.22 (0.15)	–0.15 (0.11)
7–8-year birth interval (N=309, 240)	–0.30 (0.18)	–0.40 (0.26)	0.13 (0.12)
Panel 13: With outcomes measured 15 years after first birth			
1–2-year birth interval (N=1,445, 1,244)	–0.21 (0.08)*	–0.38 (0.11)**	–0.03 (0.07)
3–4-year birth interval (N=1,103, 952)	–0.13 (0.10)	–0.23 (0.14) ⁺	–0.17 (0.07)**
5–6-year birth interval (N=560, 470)	0.04 (0.12)	–0.15 (0.13)	–0.15 (0.07)*
7–8-year birth interval (N=297, 244)	–0.29 (0.17) ⁺	–0.36 (0.23)	0.06 (0.10)
Panel 14: Restricted to women with no more than two children			
1–2-year birth interval (N=823, 761)	–0.02 (0.05)	–0.03 (0.10)	–0.10 (0.12)
3–4-year birth interval (N=675, 627)	–0.07 (0.06)	–0.22 (0.14)	–0.23 (0.12) ⁺
5–6-year birth interval (N=359, 332)	0.04 (0.05)	0.08 (0.11)	–0.02 (0.08)
7–8-year birth interval (N=200, 183)	–0.03 (0.08)	–0.24 (0.19)	0.10 (0.12)
Panel 15: Restricted to black women			
1–2-year birth interval (N=320, 293)	–0.40 (0.19)*	–0.68 (0.23)**	–0.12 (0.09)
3–4-year birth interval (N=243, 225)	0.04 (0.14)	–0.32 (0.24)	–0.54 (0.17)**
5–6-year birth interval (N=125, 111)	0.18 (0.20)	–0.32 (0.22)	–0.24 (0.13) ⁺
7–8-year birth interval (N=73, 62)	–0.63 (0.37) ⁺	–0.73 (0.37) ⁺	–0.01 (0.15)
Panel 16: Restricted to Hispanic women			
1–2-year birth interval (N=462, 424)	–0.04 (0.08)	–0.23 (0.16)	–0.20 (0.12) ⁺
3–4-year birth interval (N=367, 335)	–0.04 (0.17)	–0.12 (0.21)	–0.08 (0.08)
5–6-year birth interval (N=228, 207)	0.03 (0.10)	–0.03 (0.16)	0.02 (0.08)
7–8-year birth interval (N=128, 118)	–0.16 (0.14)	–0.29 (0.26)	0.29 (0.18)

Notes: ⁺ p < .10; * p < .05; ** p < .01; *** p < .001. Cumulative models include women not working at age 44 or 45. Robust standard errors, clustered at the individual level, are estimated. Log wage model includes only women working at age 44 or 45, resulting in a smaller sample size.

Table A-8: Estimated marginal means derived from inverse probability of treatment-weighted estimates of the average effect of a second birth on log cumulative work hours (in 1,000s), log cumulative earnings (in \$10,000s), and log hourly wages, at age 45, by two-year birth intervals, accounting for age at first birth for five, ten, and 15 years after first birth

	Log cumulative work hours	Log cumulative earnings	Log wages
Five years after first birth			
Panel 1: 1–2-year birth interval			
18–29 at first birth	1.73 (0.09)	1.35 (0.13)	2.41 (0.02)
<18 at first birth	0.13 (0.22)	–0.65 (0.29)	1.90 (0.07)
≥30 at first birth	3.34 (0.05)	3.59 (0.11)	2.94 (0.09)
Panel 2: 3–4-year birth interval			
18–29 at first birth	1.76 (0.14)	1.52 (0.17)	2.44 (0.03)
<18 at first birth	–0.11 (0.29)	–0.99 (0.39)	1.65 (0.36)
≥30 at first birth	3.40 (0.05)	3.65 (0.09)	2.93 (0.13)
Ten years after first birth			
Panel 1: 1–2-year birth interval			
18–29 at first birth	2.31 (0.09)	1.99 (0.13)	2.46 (0.05)
<18 at first birth	0.94 (0.35)	0.44 (0.31)	2.24 (0.07)
≥30 at first birth	3.58 (0.05)	3.84 (0.10)	2.82 (0.12)
Panel 2: 3–4-year birth interval			
18–29 at first birth	2.41 (0.11)	2.12 (0.19)	2.43 (0.07)
<18 at first birth	1.31 (0.21)	0.99 (0.22)	2.36 (0.07)
≥30 at first birth	3.58 (0.05)	3.80 (0.10)	2.80 (0.15)
Panel 3: 5–6-year birth interval			
18–29 at first birth	2.62 (0.12)	2.32 (0.12)	2.39 (0.11)
<18 at first birth	1.25 (0.19)	0.69 (0.23)	2.08 (0.08)
≥30 at first birth	3.64 (0.08)	3.94 (0.11)	3.08 (0.11)
Panel 4: 7–8-year birth interval			
18–29 at first birth	2.45 (0.11)	2.23 (0.14)	2.55 (0.08)
<18 at first birth	0.91 (0.39)	0.14 (0.65)	2.23 (0.08)
≥30 at first birth	3.61 (0.10)	3.87 (0.14)	3.04 (0.15)
15 years after first birth			
Panel 1: 1–2-year birth interval			
18–29 at first birth	2.76 (0.08)	2.41 (0.12)	2.51 (0.06)
<18 at first birth	1.90 (0.20)	1.23 (0.27)	2.29 (0.32)
≥30 at first birth	3.72 (0.05)	3.92 (0.11)	2.88 (0.19)
Panel 2: 3–4-year birth interval			
18–29 at first birth	2.92 (0.06)	2.66 (0.12)	2.46 (0.06)
<18 at first birth	2.04 (0.28)	1.67 (0.32)	2.21 (0.09)
≥30 at first birth	3.73 (0.06)	3.92 (0.10)	2.79 (0.18)

Table A-8: (Continued)

	Log cumulative work hours	Log cumulative earnings	Log wages
Panel 3: 5–6-year birth interval			
18–29 at first birth	3.06 (0.10)	2.75 (0.09)	2.53 (0.06)
<18 at first birth	2.16 (0.13)	1.46 (0.22)	2.31 (0.07)
≥30 at first birth	3.77 (0.09)	4.16 (0.11)	2.77 (0.26)
Panel 4: 7–8-year birth interval			
18–29 at first birth	2.79 (0.12)	2.55 (0.14)	2.61 (0.10)
<18 at first birth	1.64 (0.37)	1.13 (0.53)	2.43 (0.18)
≥30 at first birth	3.74 (0.10)	3.99 (0.14)	3.07 (0.14)

Notes: Estimated marginal means shown with delta-method standard errors.

