Center for Demography and Ecology

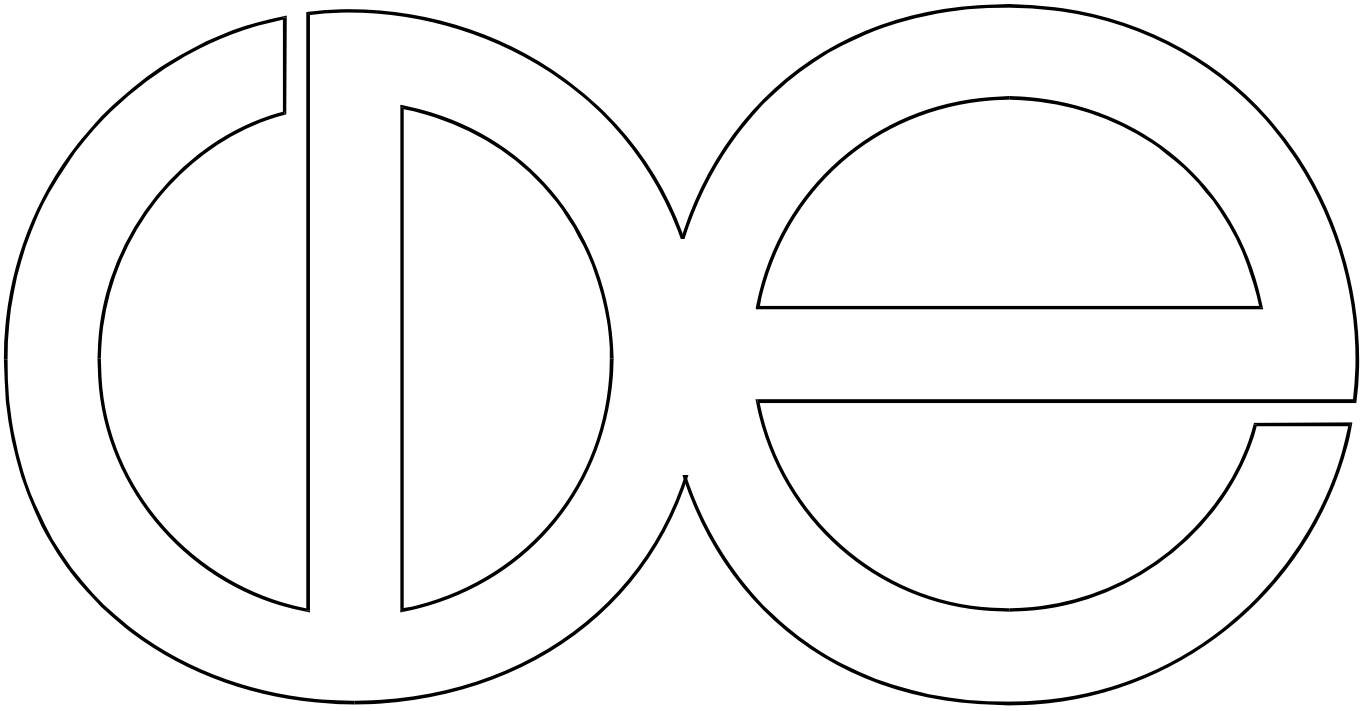
University of Wisconsin-Madison

Do Daughters Really Cause Divorce? Stress, Pregnancy and Family Composition

Amar Hamoudi

Jenna Nobles

CDE Working Paper No. 2014-02



Do Daughters Really Cause Divorce? Stress, Pregnancy and Family Composition

Amar Hamoudi
Duke University*

Jenna Nobles
University of Wisconsin at Madison†

This version: December 2013
First version: March 2013

Forthcoming in Demography

The ideas in this paper have benefited from comments of Elizabeth Ananat, Tim Bruckner, Ray Catalano, Jennifer Beam Dowd, Elizabeth Frankenberg, V. Joseph Hotz, Christopher McKelvey, Alberto Palloni, Elizabeth Peters, and Duncan Thomas, as well as three anonymous referees and our co-participants and attendees at session 161 ("Unions, Fertility, & Children") at the 2013 annual meeting of the Population Association of America. Research described in this paper was financially supported in part by the Center for Demography of Health and Aging at the University of Wisconsin-Madison. Authors are solely responsible for all content.

*Sanford School of Public Policy, Box 90312, Durham, NC 27708-0413. 919.613.9343. amar.hamoudi@duke.edu
†Department of Sociology, 1180 Observatory Drive, Madison, WI 53706. 608.262.4024. jnobles@ssc.wisc.edu

Abstract

Provocative studies have reported that in the United States, marriages producing firstborn daughters are more likely to divorce than those producing firstborn sons. The findings have been interpreted as contemporary evidence of fathers' son-preference. Our study explores the potential role of another set of dynamics that may drive these patterns–namely, selection into live birth. Epidemiological evidence indicates that the characteristic female survival advantage may begin before birth. If stress accompanying unstable marriages has biological effects on fecundity, a female survival advantage could generate an association between stability and the sex composition of offspring. Combining regression and simulation techniques to analyze real-world data, we ask: "*How much* of the observed association between sex of the firstborn child and risk of divorce could *plausibly* be accounted for by the joint effects of female survival advantage and reduced fecundity associated with unstable marriage?" Using data from the NLSY-79, we find that relationship conflict predicts the sex of children born *after* conflict was measured; conflict also predicts subsequent divorce. Conservative specification of parameters linking pregnancy characteristics, selection into live birth, and divorce are sufficient to generate a selection-driven association between offspring sex and divorce which is consequential in magnitude. Our findings illustrate the value of demographic accounting of processes which occur before birth– a period when many outcomes of central interest in the population sciences begin to take shape.

Provocative studies in economics and sociology, using data from the United States, have reported associations between offspring sex and marital stability (Ananat and Michaels, 2008; Bedard and Deschênes, 2005; Dahl and Moretti, 2008; Morgan et al., 1988; Mammen, 2008; Spanier and Glick, 1981). Several of these studies indicate that marriages producing firstborn daughters are more likely to divorce than marriages producing firstborn sons. This finding has generated considerable interest within the academic literature and beyond, influencing popular discourse (Kelly, 2010; Landsburg, 2003b,a; Belkin, 2010; Clark-Flory, 2010; Hutchison, 2010).

We, like many others, are intrigued by the findings, since they naturally raise the question, "whence the difference in divorce risk?" Some hypotheses build from the idea that fathers exhibit a "son preference"– for example, they might identify more closely with sons than daughters (Dahl and Moretti, 2008; Morgan et al., 1988), and so may be more firmly drawn into a family with sons than one with daughters. Others, including Lundberg (2005), have indicated reasons for caution in interpreting associations between sex of offspring and risk of divorce as evidence of son preference; among these are publication bias and deliberate sex-specific abortion. We investigate another possible reason– namely, the dynamics of selection into live birth. We are motivated by evidence that the sex of only a *subset* of pregnancies is observed in data, since almost nothing is known about all the pregnancies that do *not* end in a live birth.

Evidence in the epidemiological literature indicates that the unobserved subset of pregnancies may be large; as many as half or more of normal, otherwise viable pregnancies may end without a live birth (Benagiano et al., 2010; Grudzinskas and Nysenbaum, 1985; Macklon et al., 2002). There is also evidence that the female survival advantage observed throughout the human life course may begin before birth– manifesting in sex ratios among early pregnancies that are substantially more male than sex ratios among live births (Pergament et al., 2002).

In the presence of a prenatal female survival advantage, any factor affecting pregnancy survival could affect the sex ratio of a cohort of newborns. One such factor might be the biological effects of stressors like relationship conflict or other characteristics associated with unstable unions (Arck et al., 2008; Maconochie et al., 2006; Nakamura et al., 2008; Nepomnaschy et al., 2008). The dynamics we investigate are an example of the classic demographic concept of survival bias. It may be that unions producing male pregnancies are more stable than they *would have been*, had they produced female pregnancies; however, it may also (or, even, *instead*) be that more male pregnancies than female pregnancies in *already* unstable unions disappear before being observed in data. We explicate the distinction between the former ("counterfactual") types of hypotheses, and the latter ("selection-based") types, and discuss what the distinction means for both for policy and for population sciences.

The absolute magnitude of effects of relationship stress on population sex ratios at birth are likely to be small (Wilcox and Baird, 2011). Selection dynamics could not plausibly explain large sex-specific differences; for example, stressors would not be tenable to explain sex ratios observed in parts of contemporary Asia (Sen, 1992, 2003; Das Gupta, 2005). However, the magnitude of the observed association between offspring sex and divorce in the U.S. is *itself* small. Our empirical question is whether– given what is known in medicine, biology, and social science– selection dynamics could plausibly affect sex ratios at birth *enough* to warrant attention in this context. How *much* of the observed association between offspring sex and divorce could *plausibly* be accounted for by the joint effects of female survival advantage, and adverse conditions associated with marital instability?

We take two approaches to answering this question. First, we use longitudinal data from the United States to identify time ordering between indicators of marital stability and offspring sex. If fathers' son preference alone explains the association, then sons born in the future would be unlikely to affect marital stability in the present. Second, we build a simulation tool that models links among pregnancy characteristics, survival to live birth, and marital stability. We use plausible parameter values based on existing evidence from biology and social science, and allow these parameters to vary reasonably; we find that even under conservative assumptions, these relationships combine to generate associations between offspring sex and divorce that are driven entirely by selection, and are of consequential magnitude.

Empirical social scientists drawing inferences about gender dynamics from the daughter/divorce association have relied on *implicit* assumptions about dynamics of selection into live birth (Dahl and Moretti, 2008; Morgan et al., 1988; Mammen, 2008; Spanier and Glick, 1981). This paper *explicates* some of those assumptions, and illustrates the importance of testing their validity. For this study, we choose to focus on the period beginning shortly after fertilization; we discuss that choice in our conclusion.

The association between sex of offspring and risk of divorce– and secular trends in that association– may in fact reveal important information about gender dynamics in families. Our results indicate that to learn that information, it is important to first account for selection into live birth. Innovations in measurement and modeling that integrate evidence and methods from biological and social sciences will be critical to that end. More generally, as we discuss in our conclusion, demographic analysis of selection into live birth may be valuable for advancing understanding of many questions in the population sciences.

1 Motivation: Sex Preference, Gender, and Parental Behavior

Population scientists have devoted substantial effort to identifying how parental investments in children are affected by social definitions and valuations of gender.[1] In the U.S., social scientists have explored evidence of parental sex preference by examining the relationships between offspring sex and multiple domains of family life. The conclusions from this work are mixed. Gallup data indicate that fathers (but not mothers) are more likely to express a preference for sons when presented with the hypothetical situation of having a single child (Dahl and Moretti, 2008). In contrast, patterns in fertility timing (Teachman and Schollaert, 1989; Morgan and Pollard, 2002) and self-reported preferences (Pebley and Westoff, 1982; Dahl et al., 2006) are more consistent with parental preference for mixed-sex sibships than for one sex over another.

Some parenting behaviors vary by the sex of the child, like child-care arrangements (Hiedemann et al., 2004), choice of playtime activities (Lytton and Romney, 1991), and fathers' time investments (Harris and Morgan, 1991; Lundberg et al., 2007; Tucker et al., 2003). Others do not– including investments of mothers' time (Crouter et al., 1993), either parent's verbal interaction (Lytton and Romney, 1991), prenatal care (Lhila and Simon, 2008), parenting styles (Shumow et al., 1998), parental availability (Hofferth and Anderson, 2003), and parental affection (Tucker et al., 2003).

Our study is relevant to all of those, but is directly motivated by provocative findings indicating relationships between offspring sex and marital stability. In some surveys, parents of sons report greater marital happiness than parents of daughters, and are more sanguine about the stability of their marriage (Cox et al., 1999; Katzev et al., 1994). Some studies report that unmarried couples with a son enter into marriage more quickly than those with a daughter (Lundberg and Rose, 2003), while others find no evidence of a difference (Bzostek et al., 2012). Especially provocative has been work indicating that sons are more likely than daughters to live with their fathers (Dahl and Moretti, 2008; Mott, 1994; Lundberg and Rose, 2003), in part because couples are more likely to divorce if they have daughters than if they have sons (Ananat and Michaels, 2008; Bedard and Deschênes, 2005; Dahl and Moretti, 2008; Morgan et al., 1988; Mammen, 2008; Spanier and Glick, 1981).

Others have noted that the reported relationship between offspring sex and divorce is small in practical, real-world terms (Lundberg, 2005; Raley and Bianchi, 2006). Table 1 summarizes published findings linking survival of a first marriage and the sex of its first child. The estimated risks in the top panel come from studies that compare divorce risk in each discrete time period after the birth of a daughter versus after the birth of a son. They indicate that couples with daughters stand about a 1-2 per 1000 higher chance of divorcing in

[1] For reviews, see Lundberg (2005) and Raley and Bianchi (2006).

each 8 month period after the birth of a daughter. Those in the bottom panel come from studies that sample parents at a a single point in time, and compare the likelihood that their first marriage has ended if they had a firstborn daughter versus a firstborn son. They indicate differences on the order of 5 per 1000. The differences are sufficiently small that they can only be discerned in large samples.

Others have counselled caution in interpreting associations between offspring sex and risk of divorce as evidence of son preference in the contemporary United States, in light of period changes in the associations (Morgan and Pollard, 2002), risk of publication bias (Lundberg, 2005), correlations between offspring sex and needs (Hiedemann et al., 2004; Raley and Bianchi, 2006), or even deliberate sex-selective abortion (Lundberg, 2005). We investigate an additional possible reason– selection into live birth. Given the small magnitude of the observed association between offspring sex and divorce, even weak selection effects might still warrant caution; whether such effects could plausibly be operating at sufficient scale to merit attention by population scientists is our empirical question.

2 Conceptual Framework

Associations between sex of offspring and divorce risk have been identified using standard discrete dependent variable regression approaches. These approaches begin by positing some latent characteristic, which we will call "marital fitness." The latent fitness of a marriage ($m^\star$) may be a function of any number of measured ($\mathbf{x}$) and unmeasured (η) characteristics. Among the measured characteristics might be an indication of whether or not the couple has a firstborn daughter (g), and whether or not they have a firstborn son (b):

$$m_i^\star = \mathbf{x}_i'\beta + \delta_1 g_i + \delta_2 b_i + \eta_i \tag{1}$$

Latent marital fitness is not observed in data. Rather, one observes only reflections of it, like whether the couple is still married ($m = 1$), or has divorced ($m = 0$). In that case, latent fitness would be defined such that those couples who are observed to be married represent the more fit subset:

$$m_i = \begin{cases} 1 & \text{if} \quad m_i^\star \geq 0 \\ \\ 0 & \text{if} \quad m_i^\star < 0 \end{cases}$$

Regression specifications like the probit or logit use assumptions about the probability distribution of η and observations of m, $\mathbf{x}$, g, and b to uncover the nature of relationships between marriage fitness and the sex

composition of offspring. In many cases, researchers using various implementations of this general approach have observed that marriages producing firstborn sons tend to be slightly more "fit" than marriages producing firstborn daughters– formally in terms of Eq. (1), they obtain estimates of δ_2 that are statistically significantly greater than δ_1.

One common interpretation of the observed patterns is *counterfactual* in nature– perhaps most succinctly expressed by Dahl and Moretti (2008, p.1090): "[Many] first-born daughters…would have had a father present in the household had they been first-born sons instead." This interpretation holds that $\left(\hat{\delta}_2 - \hat{\delta}_1\right) > 0$ because first-born sons *improve* the fitness of their parents' marriages– for example, because their presence draws fathers into a shared parental commitment to child-rearing (Morgan et al., 1988).

We aim to investigate the extent to which a *selection-based* mechanism– which is *not* counterfactual in nature– could also generate $\left(\hat{\delta}_2 - \hat{\delta}_1\right) > 0$. Only the sex of live births is observed in data, and not the sex of pregnancies overall. Even in population surveys that include questions about miscarriages, little is known about pregnancies that do *not* survive to live birth.

To explicate potential selection dynamics, we add a term to the right hand side of Eq. (1), thereby capturing all *three* possible outcomes of a pregnancy, if it occurs. Using Eq. (1), consider couple i, who have had $P_i \geq 0$ pregnancies:

$$m_i^\star = \mathbf{x}_i'\beta + \delta_1 g_i + \delta_2 b_i + \delta_3 \sum_{p=0}^{P_i} \omega_p + \eta_i \tag{1a}$$

In this formulation, $\omega_p = 0$ if pregnancy number p ended in a live birth, and 1 otherwise. Consider a set of couples who have had at most one child (so, $b + g \leq 1$); since sex of a pregnancy is only observed if it ends in live birth, the sex of a pregnancy is only known for those couples for whom $\sum_{p=0}^{P_i} \omega_p < P_i$. In light of this sample selection process, we take conditional expectations of both sides of Eq. (1a). First, we consider a set of couples who are observed to have firstborn sons:

$$\mathrm{E}\left[m^\star \middle| \mathbf{x}, b = 1, \sum_{p=1}^{P} \omega_p < P \right] = \mathbf{x}'\beta + \delta_2 + \mathrm{E}\left[\eta \middle| \mathbf{x}, b = 1, \sum_{p=1}^{P} \omega_p < P \right]$$

Next, we consider another set of couples who are observed as otherwise identical, but who have firstborn daughters:

$$\mathrm{E}\left[m^\star \middle| \mathbf{x}, g = 1, \sum_{p=1}^{P} \omega_p < P \right] = \mathbf{x}'\beta + \delta_1 + \mathrm{E}\left[\eta \middle| \mathbf{x}, g = 1, \sum_{p=1}^{P} \omega_p < P \right]$$

Subtract the two to get the expected difference in underlying marriage "fitness" for the two sets of couples:

$$\begin{aligned} \mathrm{E}\left[\Delta m^{\star} \left| \mathbf{x}, \textstyle\sum_{p=1}^{P} \omega_p < P \right.\right] = \left(\delta_2 - \delta_1\right) + \\ \left\{ \mathrm{E}\left[\eta \left| \mathbf{x}, b = 1, \textstyle\sum_{p=1}^{P} \omega_p < P \right.\right] - \mathrm{E}\left[\eta \left| \mathbf{x}, g = 1, \textstyle\sum_{p=1}^{P} \omega_p < P \right.\right] \right\} \end{aligned} \tag{2}$$

In terms of Eq. (2), the difference $(\delta_2 - \delta_1)$ identifies the counterfactual marriage fitness impact of firstborn sons compared with firstborn daughters, *only if* the difference in curly braces is zero. The first term in that difference represents unobserved marital stability characteristics, on average, among couples that have a firstborn son; the second, among couples that have a firstborn daughter. We will discuss evidence from epidemiology, biology, and medicine implying that the difference between these two averages may not be zero; the fact a couple even *has* a firstborn child may contain *different information* about their circumstances, depending on that child's sex. (For continued mathematical discussion of these issues, see Appendix A.1).

A specific explication of the difference between what we call "selection based" versus "counterfactual" mechanisms may be warranted. Note that the sex of a pregnancy is determined by a characteristic of the sperm cell that fertilizes it– if the cell is of type "Y," the pregnancy is male; if of type "X," the pregnancy is female. **Counterfactual** interpretations hold that some divorced couples with a daughter would still be married if a "Y" type sperm had fertilized the pregnancy, instead of the "X" type sperm that actually did. By contrast, **selection based** interpretations hold that some of these couples would have divorced either way– no matter what type of sperm had fertilized. One *observes* more daughters among the divorced couples because those couples' male pregnancies tend not to reach live birth, and therefore tend not to show up in data.[2] These types of interpretations are not mutually exclusive. Some fraction of the observed differences in marriage fitness can be driven by counterfactual mechanisms (for example, sons draw their fathers into a shared parental commitment to child rearing), while another fraction can be driven by selection (for example, male pregnancies in adverse conditions never survive to be observed in data). In this paper, we ask– if *nothing but* selection mechanisms were at play– how big could the difference in divorce risk be, given what we know about the relevant processes?

[2]Dahl and Moretti (2008, Table 2) make an attempt at distinguishing selection-driven versus counterfactual channels. They look at births that happened at a time when there was still substantial variation in the use of ultrasound technology for prenatal care; they propose that this variation can be treated as a "natural experiment." It is important to note that the validity of that "natural experiment" rests on assumptions of no associations among the use of nonstandard prenatal medical intervention, parental circumstances, pregnancy survival, and the sex of the pregnancy. Those assumptions are nearly identical to assuming that the difference in curly braces in equation (2) is zero. Therefore, interpreting the ultrasound patterns as evidence of a specifically *counterfactual* effect of daughters on divorce represents the logical fallacy of "begging the question"– it is a test of the key assumption, which itself relies on the *very* assumption it is supposed to be testing.

3 Background: Biological Stress, Survival Bias, and Relationship Stability

3.1 Fecundity and Offspring Sex Composition

Throughout the paper, we maintain the assumption that the sex of pregnancies is assigned as if by coin flip; we discuss this assumption further in the conclusion. We emphasize that this assumption is *not* sufficient to justify a counterfactual interpretation of the difference $(\delta_2 - \delta_1)$ in Eq. (2). Rather, a counterfactual interpretation must also rely on *additional* assumptions about unobserved factors related both to marital stability and also to male and female prenatal survival probabilities. For example, in Appendix A.1 we carry forward the simple model of sample selection from section 2, assuming that the sex of pregnancies is randomly assigned, and explicate the additional specific assumption that underlies counterfactual interpretation of equations like (1). Having explicated it, we are not aware of any evidence that would justify it.

Additional assumptions necessary for counterfactual interpretation would be trivial if almost all pregnancies– regardless of sex– were to end in live birth, or if prenatal survival probabilities were about the same, regardless of the sex of the pregnancy. However, there is evidence in the epidemiological literature that neither of these is true. The evidence indicates that only a quarter to a half of pregnancies actually end in live birth (Wilcox et al., 1988; Wang et al., 2003; Nepomnaschy et al., 2004; Vitzthum et al., 2006; Roberts and Lowe, 1975; Benagiano et al., 2010; Macklon et al., 2002)[3], and that there may be a prenatal female survival advantage (Vatten and Skjaerven, 2004; Chahnazarian, 1988; Kellokumpu-Lehtinen and Pelliniemi, 1984; McMillen, 1979; Pergament et al., 2002).[4]

In the presence of a female survival advantage, a couple's fecundity– the probability that a fertilized egg will become a live birth– can end up reflected in the sex compostion of their offspring. In that case, assigning a purely counterfactual interpretation to $(\delta_2 - \delta_1)$ in Eq. (2) requires assumptions beyond merely that sex of pregnancies is assigned as if by coin flip. The simplest additional condition would be that no unobserved factor is associated both with fecundity and with marital stability. In the next section, we cite evidence to the contrary– biological stress may be associated both with fecundity and with marital stability.

3.2 Biological Stress, Relationship Stability, and Fecundity

"Biological stress" occurs when cells are exposed to corticosteroid hormones (McEwen and Wingfield, 2003). High levels of these hormones predict reduced fecundity (Ferin, 1999; Waffarn and Davis, 2012; Parker and

[3]More detailed discussion of the evidence is available in Appendix B.1.
[4]More detailed discussion of the evidence is available in Appendix B.2.

Douglas, 2010; Nepomnaschy et al., 2008; Arck et al., 2008; Nakamura et al., 2008). The biological system that regulates corticosteroid hormones overlaps substantially with the system that regulates progesterone– a hormone critical to healthy progress from fertilization to early pregnancy (Ferin, 1999; Parker and Douglas, 2010). At the other end of pregnancy, stress hormones are directly involved in finalizing organ development in the fetus, and inducing labor (Waffarn and Davis, 2012).

In a healthy person, levels of stress hormones rise sharply in the face of an acute challenge (a "stressor"), like a threat to personal safety. Major disasters like earthquakes or violent conflict are likely stressors; studies have reported associations between healthy pregnancy outcomes and these types of events (Torche, 2011; Lauderdale, 2006; Mansour and Rees, 2012). Complementing such studies, others have reported associations between a woman's self-reported emotional well-being and the probability that her pregnancy ends in live birth (Arck et al., 2008; Maconochie et al., 2006; Sugiura-Ogasawara et al., 2002; Nakamura et al., 2008).

Persistent exposure to stressors may eventually lead to high concentrations of stress hormones even without any immediately accompanying stressor or sense of arousal (Fremont and Bird, 2000; McEwen and Wingfield, 2003; Miller et al., 2007). Chronic hyperstimulation of the stress system is associated with adverse conditions in childhood and adolescence, including lower levels of parental support or availability (Lupien et al., 2009). These conditions are more common in the case of parental union dissolution which is predictive of lower-quality relationships in adulthood (Sassler et al., 2009). When exposed to a stressor, a person with a chronically hyperstimulated stress system faces greater risk of extended negative affect and externalizing behavior, in part due to effects of stress hormones on cells in the brain (Ruttle et al., 2011; Het et al., 2012; van Eck et al., 1996; McEwen, 2003; Gourley et al., 2013). Negative affect and externalizing behaviors may represent risk factors for divorce (Emery et al., 1999). Consistent with these patterns, psychologists have observed associations between relationship stability and biological stress responses to interpersonal conflict (Powers et al., 2006; Beck et al., 2013; Gunlicks-Stoessel and Powers, 2009; Laurent et al., 2013).

Finally, in the opposite causal direction, higher quality relationships may buffer biological response to stressors– in effect, a healthy marriage may *substitute* for corticosteroid hormones (Hennessy et al., 2009; Meuwly et al., 2012; Friedman et al., 2012).

Linking chronic stress to pregnancy outcomes, emerging research indicates evidence of a negative association between chronic hyperstimulation of a woman's stress system and the likelihood of having a live birth. For example, Nepomnaschy et al. (2004) collected biological samples from noncontracepting married women volunteers, and found a negative association between the probability that a woman's pregnancy survived and the urinary concentration of cortisol– an important stress hormone– during the first few weeks after

fertilization. Others have observed negative associations between the likelihood of pregnancy success and other indicators of chronic hyperstimulation (Nakamura et al., 2008; Sugiura-Ogasawara et al., 2002).

Dynamics linking divorce to elevated stress hormone levels, linking elevated stress hormones to adverse conditions for pregnancies, and linking the sex of a pregnancy to its likelihood of reaching live birth could interact to generate a selection-driven daughter/divorce association. In that case, it remains an open empirical question whether the relevant associations are so strong that they matter. Direct investigation of that question would require high-quality, prospective measurement of stress hormone levels, reproductive hormone levels, fertility behavior, marital status and relationship quality, and menstrual, ovulatory, and pregnancy-related outcomes for a large, population-representative sample of women. It would involve stratifying chromosomally normal, viable fertilized eggs by sex as early as possible, and following them prospectively to determine whether they become clinical pregnancies, and then live births. One could apply standard demographic methods to assess the selection processes. For the most part, such sampling and measurement would be technologically possible, but expensive. To our knowledge, it has never been done. Given the incompleteness of sampling and measurement, in the next section we lay out a second-best approach.

4 Data and Methods

4.1 Adverse Conditions and Sex of Subsequent Births

We begin by revisiting the association between marital instability and offspring sex, using marital histories and indicators of stress from a longitudinal population survey. *Counterfactual* explanations of the association posit that *after* they are born, girls place greater strains on marriages than do boys; a *selection-driven* explanation posits that girls are more likely than boys to be born into marriages that were *already* strained. We use this clear distinction in time-ordering to evaluate these hypotheses, by testing whether marital fitness correlates with the sex of births that occur *after* strain was measured.

We analyze data from the 1979 National Longitudinal Study of Youth (NLSY79), a nationally representative sample of U.S. residents born 1957-1965 (that is, aged 14-22 years in 1979). Respondents were followed annually until 1994 and biennially thereafter. The followup period for our analyses ends in 2010.[5]

Following previous studies on offspring sex and divorce, we limit the sample to white women without premarital births. Given our interest in the time-ordering of marital discord and the gender composition of children,

[5]For more details on the study, including documentation of sampling and attrition, see `http://www.bls.gov/nls/nlsy79.htm`

we further limit the sample to women who have at least one birth in their first marriage. The sample meeting these criteria consists of 1314 women.

We use two measures to capture indicators of biological stress as well as marital "fitness." The first is a marital conflict scale. Married female respondents are asked about the frequency with which they argue with husbands in 9 domains. Responses range from 0 ("never") to 3 ("all the time"). We sum these responses, producing a "marital conflict scale" that ranges 0-24 with a mean of 7.8 and a standard deviation of 4.4 ($\alpha = 0.76$). The second is a psychometric instrument, the Center for Epidemiological Studies Depression (CES-D) Scale (Radloff, 1977), the only item capturing emotional health in the NLSY79 prior to 2000. The scale is generated by summing responses to questions about feelings or behaviors indicating chronic depression risk. Scores range 0-60; higher values indicate higher levels of depression risk. Depression is associated with biological stress (McEwen, 2003; Rawdin et al., 2013). The mean CES-D score in our sample was 8.1, with a standard deviation of 8.2. Both measures were first asked of childless women in 1992. Depression risk was not measured again until respondents reached age 40, but marital conflict was measured biennially after 1992.

A "first-best" test of time-ordering of marital strain and offspring sex would use data collected throughout marriages. By 1992, when these measures were first collected, NLSY79 respondents in this sample were 27-35 years of age, and had already been married for 8.8 years on average. Nevertheless, because many couples had not yet completed their fertility by 1992, a "second-best" approach is possible. We focus on those marriages that survived until 1992 and that went on to have at least one more child between 1992 and 2010 in order assess whether marital conflict and/or emotional health *predict* the birth of daughters. The NLSY79 remains the only longitudinal U.S. data set that has collected measures of emotional health and marital conflict during reproductive years with sufficiently long followup that fertility and divorce can be observed.

We begin by describing the association between these measures and subsequent risk of divorce, using linear probability models (results are not sensitive to functional form). In each regression, our dependent variable is a dichotomous indicator of whether a respondent who was in her first marriage in 1992 divorced between 1993 and 2010. Covariates of interest– investigated in two separate regressions– are the respondent's 1992 CES-D score and 1992 marital conflict score. In both regressions additional covariates include age at marriage and total number of births.

We next identify the association between the distress measures and the sex of offspring. We pool all first, second, and third births occurring within a woman's first marriage during the follow-up period. For each birth, we regress a dichotomous indicator for whether it is male against its mother's 1992 scores on the

conflict and CES-D instruments. Results are the same using the conflict score measured in the year most closely preceding the birth. Additional covariates include age at first birth, birth order, and total births after 1992; results do not change if we add indicators for educational attainment and urban/rural residence. Standard errors are estimated in a manner that accounts for possible statistical non-independence between observations of different births to the same mother (which might arise, for example, because the psychometric scores are repeated in every observation of the same mother).

All births in our regression occurred *after* distress was measured. Sons are unlikely to have improved the quality of their parents' marriages before they were born; any association observed through these analyses is more likely to reflect selection into live birth. If sex selective medical abortions were sufficently common and sufficiently correlated with marital strain, interpretation would be more complicated. In our simulation exercise described in the next section, we explicitly rule out any direct relationship between prenatal sex selection and stress.

4.2 Simulation: Stress, Fecundity, Sex of First Birth, and Divorce

In the analysis described in section 4.1, we use survey data on characteristics that are *associated* with biological stress. Direct measures of biological stress are increasingly available in population representative data (Seeman et al., 2010), though no major U.S. cohort studies contain these indicators measured simultaneously with fertility behavior. Furthermore, in population surveys, little is observed about pregnancies that do *not* end in a live birth. In some surveys, respondents are asked to report past miscarriages; these almost certainly represent a small fraction of lost pregnancies. Furthermore, genetic characteristics indicating the viability and sex of a miscarriage are almost never observed in these data. In retrospective data on miscarriages, it is also difficult to reliably observe biological stress at the time of that pregnancy. In light of these gaps in measurement, we build a simulation that draws on information taken from vital statistics, from population surveys, and from published studies in biology and medicine, as well as our findings from section 4.1.

The description in this section is summarized in table 3 and figure 1. Further detail is available in Appendix A.2.

4.2.1 Inputs

The simulation tool combines four pieces of information on biological and family processes, together with our findings from the analyses described in section 4.1. Where appropriate, we test a range of values.

1. The time path of divorce risk for married couples. This path takes the form of 15 period-specific risks, where each period represents 8 months (results are not sensitive to the length of these intervals). We have explored the implications of two sets of risks; one is from the June 1980 Current Population Survey (CPS80), as reported in Morgan et al. (1988), the other from NLSY79. Since results were similar for all three sets of risks, we report only those using the CPS80 series here.

2. Sex ratios among early pregnancies and among live births.

   We fix the sex ratio of live births at 105.88 boys per 100 girls, which corresponds to the ratio observed in U.S. vital statistics (registered first births to married white women between 1990 and 1995).

   We explore the implications of three different assumptions regarding the sex ratio in early pregnancy, all based on information in the epidemiological literature. Estimates in the literature vary from 110 to 170 early male pregancies per 100 female (Pergament et al., 2002). Therefore, for our base case, we assume a sex ratio of 140 male pregnancies per 100 female; we also explore the implications of the ends of this range.

   The extent of the prenatal female survival advantage is determined mechanically by the difference between the sex ratio early in pregnancy versus the ratio at live birth. Sex does not change in the course of a pregnancy, so a falling sex ratio implies a female survival advantage.

3. The probability that a chromosomally viable pregnancy will survive to live birth.

   We explore the implications of three different assumptions about this probability, all based on information in the epidemiological literature. Estimates in the literature vary from 30% to 70% (Benagiano et al., 2010). Therefore, we explore the implications of assuming fertilized eggs have a 30%, 50%, or 70% probability of surviving to live birth.

4. The time path of pregnancy risk for married couples.

   Since no population survey identifies early pregnancies, we cannot draw on any real-world study for this time path. However, population studies do report *birth* risks for married couples; we therefore combine those with the probability of survival to live birth (see input 3, above). For example, in simulation runs specifying that pregnancies have a 50% probability of surviving to live birth, we also specify that for each live birth that occurs, two pregnancies must have begun in the previous period.

   We have explored the implications of using two different sets of birth risks; one is from the CPS80, as reported in Morgan et al. (1988), and the other from the NLSY79. Since results were similar for both sets of risks, we report only those using the CPS80 here.

We use the simulation to explore implications of these factors the divorce/offspring sex association in a cohort of 250,000 hypothetical couples during the ten years following their formation.

4.2.2 The Operation of the Simulation Tool

The simulation tool runs in discrete time; there are 15 time periods. In each time period, it uses the inputs in section 4.2.1 to assign new pregnancies to couples who are still married and not currently pregnant, new live births to couples who are currently pregnant, and new divorces to couples who are still married.

In keeping with the maintained assumption that the sex of pregnancies is randomly assigned, every new pregnancy is given equal probability of being male, as specified in input 2. A subset of pregnant couples is chosen to experience a live birth; less stressed couples, and couples carrying female pregnancies, are more likely to be chosen. We do not allow stress to *directly* generate sex-specific mortality. Instead, since it operates against a background of female survival advantage, it indirectly influences sex ratios across the stress distribution– this is because all pregnancies in more stressed couples suffer a survival disadvantage, but *male* pregnancies in more stressed couples bear *two* disadvantages (first, they lack the female survival advantage, and second, there is the direct effect of stress on fecundity).

At the end of each period, the simulation tool chooses a subset of still married couples to divorce. More stressed couples are at higher risk of being chosen. Intentionally, we allow no counterfactual effects on divorce risk, either of the presence of children, or of their sex composition. This is in order to investigate the magnitude of *purely selection-driven* daughter-divorce associations that could be generated in plausible scenarios. If these associations are sufficiently small relative to the associations reported in the literature, then a counterfactual effect would likely be filling the difference.

Finally, we note that the simulation explores only "mechanical" relationships between probabilistic outcomes; that is, it does not allow for any forward looking or strategic behavior. In particular, we do not build in an opportunity for simulated couples to take into account their underlying marital stability when planning their fertility. In order to account for these sorts of dynamics, it assigns to each couple a fertility behavior parameter, and also a marital stability parameter. We examine the implications of correlations between these parameters of -0.4, 0, and 0.4. The simulation rules out any form of deliberate prenatal sex selection or sex-specific parity progression.

More detail on the operation of the simulation tool is available in Appendix A.2.

4.2.3 Outputs

At the end of a simulation run, we have, for the entire hypothetical cohort, a panel dataset containing:

- whether and when each couple divorced (D_t for $t \in \{1, \ldots, 15\}$)
- the number of children (0, 1, 2, or 3 or more) born so far to each couple, which we convert to an indicator, $C_t = 1$ if the couple has had a child, and 0 otherwise.
- the sex of each child born, which we convert to an indicator, $S_t = 1$ if the child has had a firstborn son, and 0 otherwise.

Using this panel dataset, we can model relationships between fertility, offspring sex, and marital status using standard regression equations analogous to those in the literature on offspring sex and divorce.

Choosing Magnitudes for Stress Associations As highlighted in the description in section 4.2.2, the simulation is more likely to subject a couple to divorce if it is more stressed, and is less likely to choose a pregnancy to end in live birth if the couple is more stressed. In order for our simulation exercise to be useful, it is critical that we choose plausible magnitudes for these stress associations. To that end, we calibrate these magnitudes in part by drawing on the results of the analyses described in section 4.1. First, we run the simulation tool using putative magnitudes for stress associations. At the end of the simulation run, we rescale our stress parameter to have the same standard deviation as the CES-D score observed in the NLSY79 sample. Then, we mimic the analyses described in section 4.1 on the simulated histories. We compare the stress-divorce and stress-offspring sex associations in the simulated histories against those we observed in the NLSY79. Then, we adjust putative magnitudes of the stress associations, and repeat the process. In order to be conservative, we calibrate putative magnitudes so that they generate associations in our NLSY79-*like* analyses that are *half* as strong as the associations we actually observed in the NLSY79. We also test the implications of setting relationships that generate associations of *equal* strength to those we observed in the real world data, and of setting relationships that generate associations of only *one-tenth* the magnitude. Is our putative stress-termination association plausible? There is little direct real-world evidence about the magnitude of associations between early pregnancy loss and biological stress. In one study that has shed light on this, Nepomnaschy et al. (2008, Table 1) followed a cohort of pregnancies that was selected to be typical, and observed a 2.7 times higher risk of non-live termination for the pregnancies that were exposed to more than

the cohort median level of stress hormones. The analogous association in our base case of the simulation is 2.0, more conservative than the 2.7 they report.[6]

4.2.4 Analysis of Outputs

Prospective Analyses One set of regression equations mimics the sort used in studies like those in the top panel of table 1. Specifically, we drop a couple from follow-up after it divorces, and then use a logistic regression specification to estmate discrete-time, proportional hazards regressions of the following form:

$$d_{i,t} = \beta_0 + \beta_1 t + \beta_2 t^2 + \delta_1 C_{i,t} + \delta_2 S_{i,t} + \beta_3 C_{i,t} t + \beta_4 C_{i,t} t^2 + \varepsilon_{i,t} \tag{3}$$

Where $d_{i,t}$ represents the log-odds that couple i will divorce in period t. The parameter of interest is the apparent son "effect" (δ_2), which we convert into percentage point terms by computing marginal effects holding time constant at the period of peak divorce risk.

Cross-sectional Analyses A second set of regression equations mimics the sort used in studies like those in the bottom panel of table 1. Specifically, we mimic the process of drawing parents from a cross-sectional dataset like the U.S. Census. We keep only couple-periods that have children present. We then randomly draw a single period for each couple, and estimate linear probability regressions of the following form:

$$D_i = \delta_0 + \tilde{\delta}_2 \tilde{S}_i + \tilde{\varepsilon}_i \tag{4}$$

Where D_i takes the value 1 if the couple is observed to be divorced, and 0 otherwise. As with the prospective analyses, the parameter of interest is the apparent son "effect" ($\tilde{\delta}_2$), in percentage point terms.

Interpretation The objective of the simulation is to identify how much of the observed associations between offspring sex and divorce could plausibly be driven by selection into live birth. To that end, we compare estimates of δ_2 and $\tilde{\delta}_2$ to the real-world associations quoted in table 1. Since the simulation *only* allows for selection-driven associations, if these estimates are similar then there may be need for caution regarding counterfactual interpretations of the real-world associations.

[6]The analogy is not perfect– for example, the study design in Nepomnaschy et al. (2008) probably captures acute stress but not chronic hyperstimulation of the stress system. Nonetheless, the comparability in these relative risks arguably indicates that our putative association is plausible.

5 Results and Discussion

5.1 Time-ordering: Reported Stressors and Offspring Sex

Table 2 contains coefficient estimates from the linear proability regressions using the NLSY79 data. Column A reports the results of two regressions (a separate one for each row) that describe associations between prospective risk of divorce and each of the two distress measures.

The top row indicates that women who were one standard deviation below the mean conflict score were 13 percentage points less likely to divorce than those who were one standard deviation above the mean.[7] The second row describes the association between the CES-D score and the subsequent 18-year risk of divorce. Women scoring one standard deviation below the mean on the CES-D instrument were, on average, 13 percentage points less likely to divorce over the subsequent 18 years than those who scored one standard deviation above the mean.[8] These findings are consistent with evidence in the sociological literature that depression and marital stress predict divorce (Wade and Pevalin, 2004; Amato, 2010).

The results in column A are unlikely to reflect a causal association; the objective of the analyses is purely descriptive. Many characteristics are likely correlated with both distress and divorce, but if distress, *or any associated characteristics* are also correlated with fecundity, then the female prenatal survival advantage can generate selection-driven associations between divorce and sex of offpsring.

The top row in column B reports coefficient estimates from regressions of the sex of children against their mothers' marital conflict score in 1992; the bottom row, against their mothers' CES-D score. Column B includes only births after 1992. If the sex of live births is assigned as if by coin flip, there should be no association in this column. However, the associations are large and negative. Women reporting levels of conflict in 1992 that were one standard deviation above the mean were 7.7 percentage points less likely[9] to subsequently give birth to boys than women who had similar fertility but who who had reported levels of conflict one standard deviation below the mean. This magnitude is large. For example, race is a well-known predictor of child sex (Pergament et al., 2002; Chahnazarian, 1988; Davis et al., 2007; Marcus et al., 1998); black mothers are about 0.6 percentage points less likely to have a son than white mothers. The association between the likelihood of having a son and a two-standard-devation difference in marital conflict score is ten times that size. Given the lack of similar information in other datasets, we cannot determine if this magnitude

[7] $0.015 \times 4.4 \times 2 = 13.2\%$
[8] $0.008 \times 8.2 \times 2 = 13.1\%$
[9] $-0.010 \times 4.4 \times 2 = -7.7\%$

is anomalous. Lending credence to the result is the fact that if we restrict the follow-up period to include only births within 5 years after psychometric measurement, we estimate stronger associations; furthermore, restricting only to *first* births greatly reduces the sample size but has no effect on the magnitude of the associations. Nonetheless, as reliable longitudinal measures of biological stress, divorce, and fertility outcomes become available in other population-representative surveys, analyses like these should be replicated. For example, the 1997 cohort of the NLSY, and the National Longitudinal Study of Adolescent Health are following these outcomes; in about a decade, these cohorts will be sufficiently far along in their marriage and fertility to facilitate such analysis. In the meantime, we treat these results conservatively by using relationships in our simulation that are half and one-tenth as strong (and that are separately benchmarked against pregnancy survival studies, as described in section 4.2.2).

5.2 Simulation

Table 3 and figure 3 detail the parameters used in the simulation.

5.2.1 Base Case

We begin by executing the simulation 1000 times, each time using the parameter vector in column 3 of table 3, but a different random number seed (and therefore, different values for all stochastic terms). At the end of each simulation run, we estimate regression equations (3) and (4).

Figure 2 is a histogram of the 1000 estimates of δ_2 from equation (3) (one for each simulation run), converted into percentage point marginal effects for the highest risk period for couples with children. The association between firstborn son and divorce risk is negative in all but 10 of the 1000 runs. On average over the 1000 runs, the effect was to appear to reduce the probability of divorce in each 8-month period by about 0.094 percentage points, which represents between 55% and 23,500% of the effect sizes reported in the studies in the top panel of table 1. In 561 of the 1000 runs, the selection-driven son association exceeded 50% of the effect size reported by Morgan et al. (1988).

Figure 3 is a histogram of the 1000 estimates of $\tilde{\delta}_2$ from Eq. (4) (one for each simulation run). On average over the 1000 runs, the association of first-born son and divorce risk was -0.43 percentage points. This association represents about 40-100% of the effect sizes reported in the studies in the bottom panel of table 1. In about half the cases, the selection-driven son association was equal to or greater than the effect size reported by Dahl and Moretti (2008).

5.2.2 Sensitivity Tests

Table 4 summarizes results using alternative values of our key parameters. For each alternative set of values, we executed the simulation 250 times, using a different random number seed each time. At the end of each run, we estimated regression equations (3) and (4). In the table, we report the mean and standard deviation of the selection-driven son associations across the 250 runs.

Rows 2 and 3 in the table indicate the impact of alternative assumptions about the sex ratio of pregnancies. If the female survival advantage is at the top of the plausible range (i.e., if sex ratios fall by 60% during pregnancy), then selection-driven effects could account for *all* of the associations reported in most of the studies cited in table 1. On the other hand, if that survival advantage is low, then selection-driven effects are likely to be trivial in size.

The next two rows illustrate the impact of alternative assumptions about the likelihood that a pregnancy will survive to live birth. The results indicate that lower prenatal mortality risk implies larger selection-driven son associations, compared with the base case, especially in the cross-sectional analysis (column 2). If survival probabilities are high, then selection-driven effects could fully account for all of the associations reported in most of the studies cited in the bottom panel of table 1.

Row 6 illustrates the implications of stress-divorce and stress-fecundity relationships of the strength implied by the NLSY79 results (table 2). The selection-driven son association matches or exceeds those reported in every one of the studies cited in table 1. Row 7 illustrates that if the stress-fecundity relationship is weak, then selection-driven effects are likely to be trivial.

In row 8, we explore an alternative functional form in the relationship between stress and divorce. In the NLSY79 cohort, respondents with scores indicating highest risk of depression or high levels of relationship conflict were substantially more likely to divorce compared with those in the middle of the distribution on these measures. However, those with low scores and those with scores in the middle of the distribution were nearly equally likely to divorce. In light of that pattern, we hold divorce risk constant for the bottom three quartiles of the stress distribution, and then it then jumps to a new, higher level. Results are similar between this alternative functional form and the base case, suggesting that selection-driven son associations may confound interpretation of results like those reported in table 1, *even if* the relationship between biological stress and divorce is more complicated than we specified in our base case, as long as the best linear representation of that relationship is as specified in our base case.

The last pair of rows indicates the implications of alternative assumptions about the relationship between marital "fitness" and fertility behavior. In the penultimate row, couples that get a high draw in terms of

marital "fitness" are also likely to get a high draw in terms of the likelihood of getting pregnant in any period. In that case, stable couples would be overrepresented in any cohort of pregnancies. In the last row, the correlation is specified in the opposite direction, so that *unstable* couples would be overrepresented in any cohort of pregnancies. The results suggest that if more unstable couples try harder than more stable couples to have children, this behavior may offset the effects of stress on fecundity and thereby reduce the selection-driven son association. This pattern further highlights important possible interrelationships between family processes that might be described as "primarily social" and those that might be described as "primarily biological."

6 Conclusion

Son-preference exists in the modern world; it almost certainly affects some population and family processes. In this study, we investigate a tool social scientists have used in efforts to uncover some of those effects– namely, the treatment of offspring sex as a "natural experiment" to approximate counterfactual comparisons. We draw on evidence in the social and natural science literatures, and find that this approach may not be well-suited to its purpose when associations are small in terms of absolute magnitude. Specifically, we find that selection into live birth may account for a large proportion of the observed association between offspring sex and divorce in the United States.

Using data from a large cohort study, we find evidence that indicators of stress predict the sex of future live births, and also predict future divorce. Then, using the current state of knowledge regarding patterns of fertility, marriage, and pregnancy outcomes, we investigate plausible implications of these patterns when they interact. The results indicate that female survival advantage, operating against a background of prenatal mortality risk that is correlated with marital instability, could plausibly give rise to at least half of the observed correlation between offspring sex and marital stability.

Our findings add one more cause for caution– in addition to those that have already been noted (for example, in Lundberg (2005)) – regarding the use of the association to make claims about social norms, or to make claims about secular changes in those norms, or as a source of quasirandom variation in divorce risk.

Our analysis was designed to speak to current research on offspring sex and *divorce*. Since marriage is no longer the dominant environment of fertility and child rearing in the United States (Kennedy and Bumpass, 2007), it may be worth extending this type of inquiry to outcomes other than divorce. Extending our approach

to apply to non-marital unions, or to investigate outcomes other than union stability– including, for example, sex-specific child support payment (Mammen, 2008) – may be instructive.

Evidence that offspring sex might reflect unmeasured characteristics of families could also have implications that extend to other lines of inquiry in demography. Research across the empirical social sciences investigates ways in which family demographic composition and living arrangements affect and are affected by family members' individual outcomes as well as conditions in the family's social and economic environment.[10] In some of these cases, nonrandom selection into live birth is likely to be trivial to the main inferences; in others, it might be more consequential. In order to know when we should care, it will be important to continue demographic investigation of the dynamics of selection into live birth, in a manner which is integrated with existing research on fertility behavior and other family processes.

The strength of our conclusions is naturally limited by the degree of uncertainty about the relevant parameters. Many of these parameters are difficult or expensive to measure. For the most part, our main conclusions are robust to values across ranges that are considered plausible based on the current state of knowledge, although our sensitivity tests have specifically highlighted potential value of identifying the true size of the prenatal female survival advantage and the true fraction of fertilized eggs that become live births, and of better pinning down relationships between fecundability, fecundity, and fertility behavior.

Throughout this paper, we have maintained the assumption that the sex of *early* pregnancies is randomly assigned. Our analysis begins in the days after an embryo has formed, when it appears that sex ratios are male-skewed. However, there may be relevant selection dynamics even earlier than that; evidence suggests that sex ratios become increasingly male as a cohort of fertilized eggs develop into embryos, after which the female survival advantage begins (Boklage, 2005); unanswered questions remain around prenatal and neonatal female survival advantage (Steinsaltz, 2013; Lawn et al., 2013). Tracing the time path of population sex ratios at as high a temporal resolution as possible, beginning as early as possible (perhaps even before fertilization), in as many well-characterized subpopulations as possible, is an important demographic exercise.

Valuable work has been done in biology, medicine, and epidemiology over the past half-century to investigate sex-specific selection into live birth. However, measurement and analysis of selective mortality, group-specific survival advantage, and cohort composition are squarely in the purview of population studies (Vaupel and Yashin, 1985; Manton et al., 1995; Crimmins et al., 2008). Standard components in the toolkits of empirical social scientists– including the tools of sampling design, formal demography, and behavioral modeling–

[10]For example, these might include studies like those cited in Raley and Bianchi (2006), Snyder (1998), Steelman et al. (2002), or Conley et al. (2007).

offer much promise toward integrating reproductive biology with a broader understanding of social and population dynamics.

Table 1: Previous Studies on Divorce and Sex Composition of Offspring

Longitudinal or Quasilongitudinal Analyses (Outcome: Risk of divorce per t month period)

Authors (year)	*Comparison*	$t =$	Δ *(pctg points)*	*Sample size*	*Data source*
Morgan, Lye, & Condran (1988)	Couples with 1 daughter vs 1 son	7	0.17[a]	178,310	CPS[b], 1980
Morgan & Pollard (2002)	Couples with 1st-born daughters vs sons	12	0.12[c]	261,759	CPS, 1985, 90, 95
Diekmann & Schmidheiney (2004)	Women with 1 daughter vs. 1 son	1	0.0004[d]	5396	FFS[e], 1995

Cross-sectional Analyses (Outcome: At time of survey, woman has divorced from 1st husband)

Authors (year)	*Comparison*	Δ *(pctg points)*	*Sample size*	*Data source*
Dahl & Moretti (2008)	Ever-married adults aged 18-40 coresiding with ≥ 1 child aged 0-12: 1st-born daughters vs sons	1.03	96,859	CPS, 1980, 85, 90, 95
		0.45	1,932,964	Census, 1960-1980
Ananat & Michaels (2008)	White women with a ≤ 17 yr old 1st-born daughter vs son	0.63	619,499	Census, 1980
Bedard & Deschênes (2008)	U.S. born ever-married white women with birth ≤ 5 yrs into marriage: 1st-born daughter vs. son	0.80	465,595	CPS, 1980, 85, 90

Notes for table 1: See also the 11 studies cited in Morgan and Pollard (2002, Table 1A).
a - Odds ratio of 1.071, on base risk that ranges 1.7-2.5%, depending on time since marriage
b - "CPS" is an abbreviation for the June Marriage and Fertility Supplement to the Current Population Survey
c - Odds ratio of 1.03, but base risk not reported in paper. Using different data, we estimated the base risk around 1-4%, depending on time since marriage.
d - Odds ratio of $\frac{1}{0.88}$ on base risk that ranges 0.28% to 0.34%, depending on time since marriage
e - "FFS" is an abbreviation for the Fertility and Family Survey

Table 2: Reported stressors & prospective risks of divorce & male birth

Dependent variables: As given in column header. First marriages, intact in 1992.	[A] Divorced between 1992 & 2010 (1=Divorced, 0=Not)	[B] Sex of births 1992-2010 (1=Male; 0=Female)
Mother-reported characteristics in 1992		
Marital conflict	0.015**	-0.010†
($\mu = 7.8$, $\sigma = 4.4$)[a]	[0.003]	[0.005]
CES-D[b] Score	0.008**	-0.006*
($\mu = 8.1$, $\sigma = 8.2$)[a]	[0.002]	[0.003]
Fraction with $y = 1$:	23.7%	50.3%
Other covariates[a]		
Age at marriage	$\mu = 22.4$, $\sigma = 3.8$	Not included
Total births in this marriage	$\mu = 2.3$, $\sigma = 1.0$	Not included
Birth order	Not included	$\mu = 2.0$, $\sigma = 0.7$
Age at first birth	Not included	$\mu = 34.0$, $\sigma = 2.9$
Total births in this marriage after 1992	Not included	$\mu = 1.7$, $\sigma = 0.7$
Sample sizes[d]:		
# of mothers	1314	
# of births		587

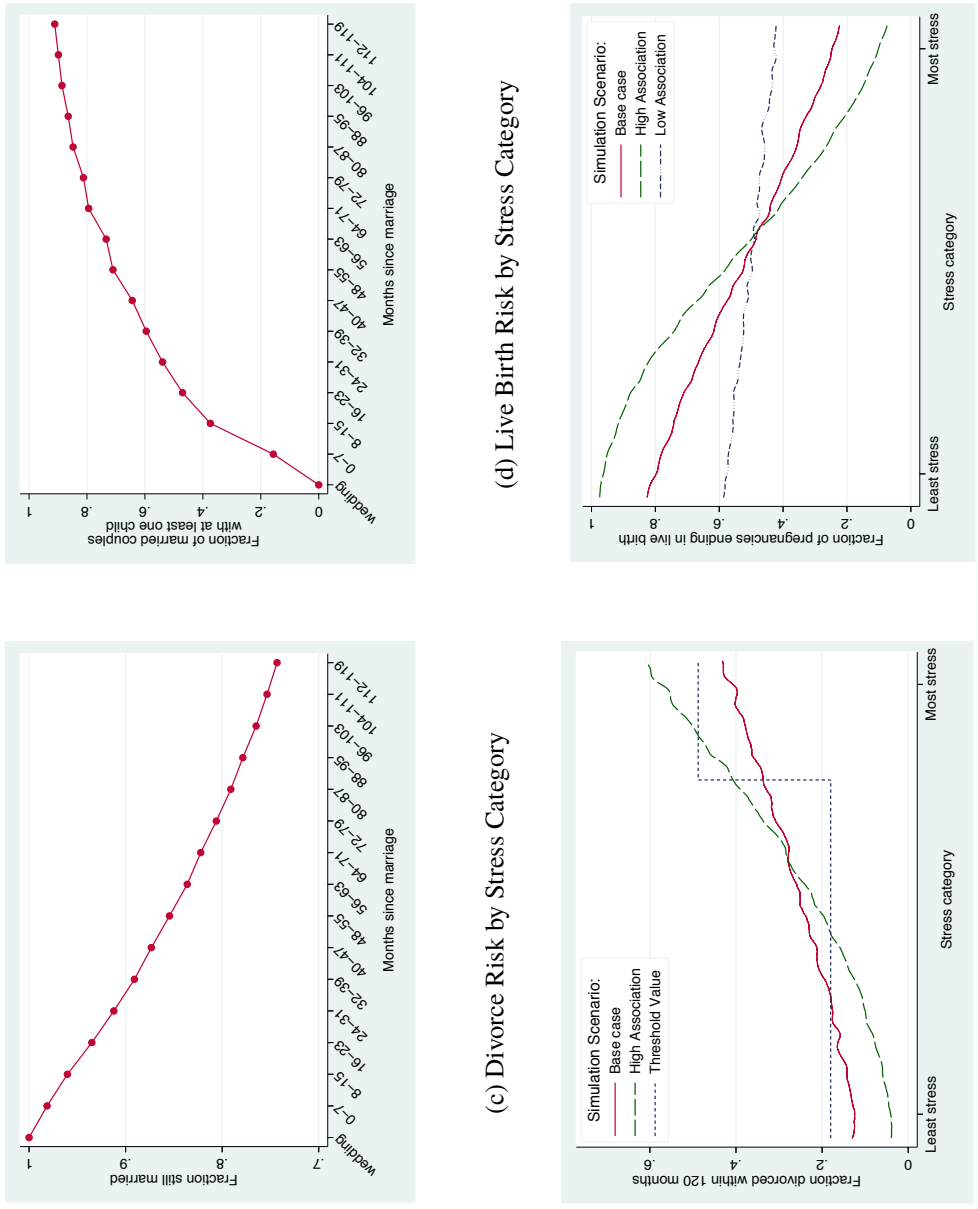
Notes for table 2: Data are from the National Longitudinal Survey of Youth, 1979 cohort. For all statistics except observation counts, observations are weighted to represent the population from which the regression sample was drawn.
a - μ: mean of characteristic; σ: standard deviation.
b - Center for Epidemiologic Studies- Depression. A standard psychometric indicator of depression risk (see Radloff (1977)).
c - Observation counts are unweighted.
†- significantly different from zero at a 10% or larger test size
$\star$ - significantly different from zero at 5% or larger test size
$\star\star$ - significantly different from zero at 1% or larger test size.

Table 3: Simulation inputs

Parameter Definition	Base value	High value	Low value	Reference
Population Sex Ratios				
Male pregnancies per 100 female pregnancies	140	170	110	Pergament et al. (2002)
Male births per 100 female births	105.88	105.88	105.88	US Vital Statistics (1990-95)
Fertility/Fecundity				
Pr {$\geq$ 1child | time since marriage}		See figure 1b		Table A1 of MLC[a]
Overall fraction of pregnancies reaching live birth	50%	70%	30%	Benagiano et al. (2010)
Pr {live birth | stress}		See figure 1d		Relationships in table 2.[b]
Divorce Risk				
Pr {divorce | time since marriage}		See figure 1a		Table A1 of MLC[a]
Pr {divorce | stress}		See figure 1c		Relationships in table 2.[b]
Other factors				
Correlation: fertility behavior/fecundability & marriage "fitness"	0	0.4	-0.4	Authors' conjectures

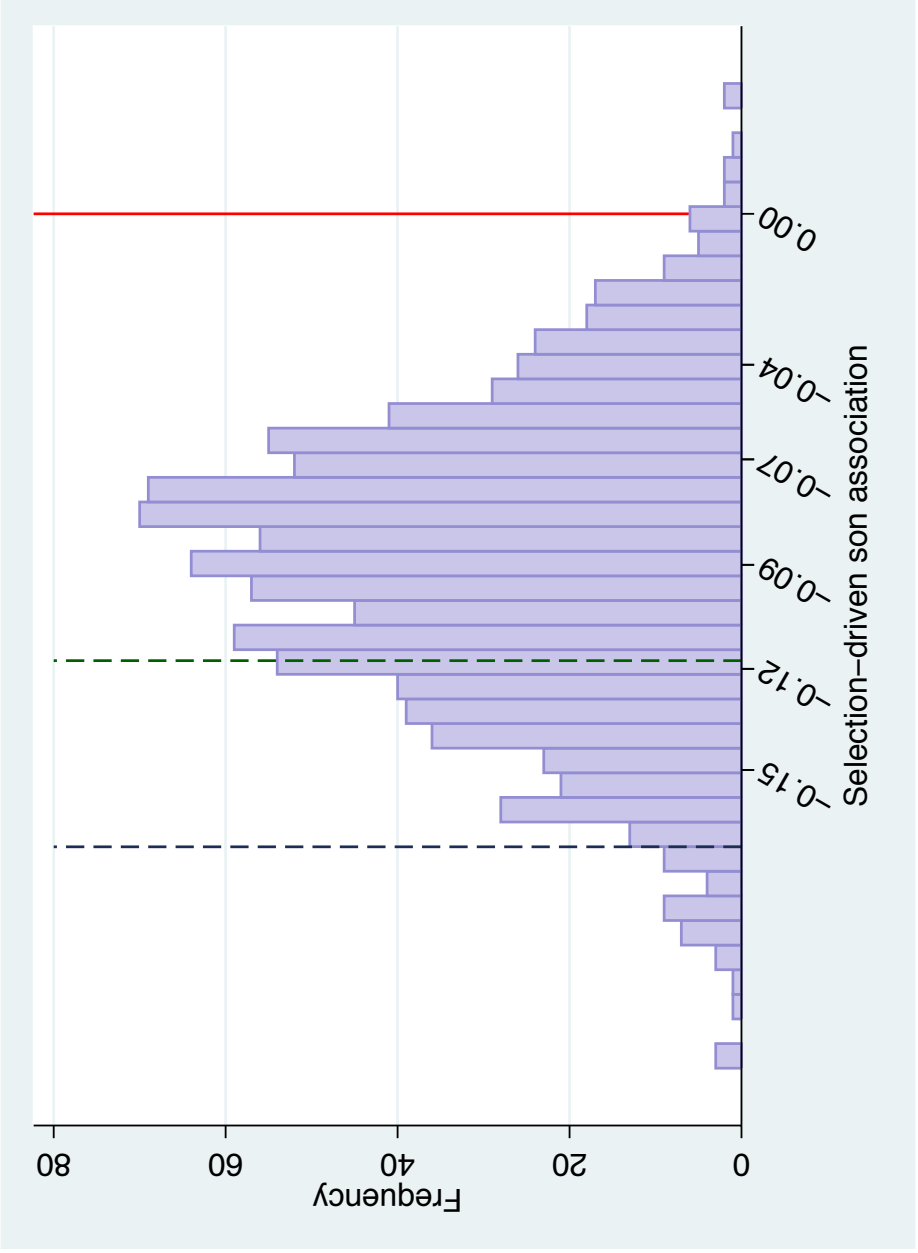
Notes for table 3: These parameter values comprise those used in all 11 scenarios presented in table 4.
a - MLC refers to Morgan, Lye, and Condran (1988).
b- For details on how we used the results in table 2 to arrive at the parameter values shown, see section 4.2.

Figure 1: Simulation Inputs: Divorce and Live Birth Probabilities



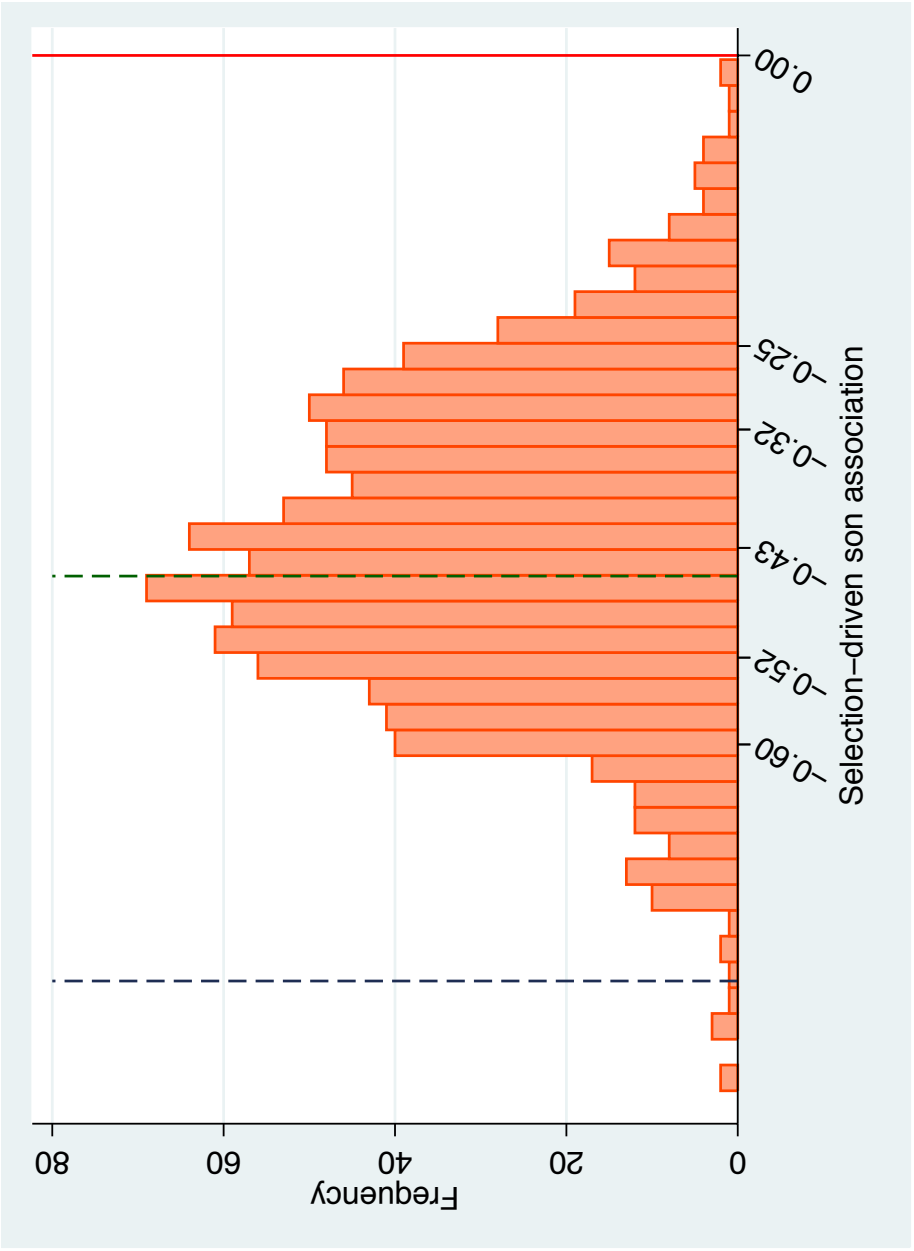
Notes for figure 1: Graphical representations of parameter values referenced in table 3.

Figure 2: Simulation results: Longitudinal analyses, base parameter values



Notes for figure 2: We executed the simulation tool 1000 times, using the base parameter vector outlined in table 3 (and in the series traced by the solid red lines in figure 1) each time. At the end of each execution, we estimated a discrete-time, proportional hazards regression on the simulated dataset. The regression equation is described in the text (Eq. (3)); it mimics the analyses reported in studies like those cited in the top panel of table 1. We then converted estimated log odds differences (δ_2 in the regression equation) to percentage-point risk differences for the period of peak divorce risk. This exercise generated 1000 risk differences (one for each execution of the simulation tool); this histogram shows the distribution of the 1000 values. The 10^{th} percentile, 25^{th} percentile, mean, 75^{th} percentile, and 90^{th} percentile are marked and labeled on the horizontal-axis, and the vertical red line is at 0. Dotted vertical lines indicate reported effect sizes from two of the studies cited in the top panel of table 1.

Figure 3: Simulation results: Cross-sectional analyses, base parameter values



Notes for figure 3: We executed the simulation tool 1000 times, using the base parameter vector outlined in table 3 (and in the series traced by the solid red lines in figure 1) each time. At the end of each execution, we randomly drew one time period for each couple, and dropped all childless couples. This mimics the process of drawing data on parents with children from a cross-sectional dataset like the Census. We then estimated the difference in risk of having divorced as a function of sex of the first birth, using a linear probability regression. The regression equation is described in the text (Eq. (4)), and mimics the analyses reported in studies like those reported in the bottom panel of table 1. That exercise generated 1000 risk differences (δ_2 in the regression equation)– one for each execution of the simulation tool. This histogram shows the distribution of the 1000 values. The 10^{th} percentile, 25^{th} percentile, mean, 75^{th} percentile, and 90^{th} percentile are marked and labeled on the horizontal-axis, and the vertical red line is at 0. Dotted vertical lines indicate reported effect sizes from two of the studies cited in the bottom panel of table 1.

Table 4: Simulation results: Alternative scenarios

Scenario	Son "effect" (pctg pts)	
	Prospective (δ_2)	Cross-sectional ($\tilde{\delta}_2$)
Base case	-0.10 [0.04]	-0.44 [0.15]
170 male pregnancies per 100 female	-0.16 [0.04]	-0.75 [0.14]
110 male pregnancies per 100 female	-0.01 [0.04]	-0.05 [0.14]
30% of pregnancies end in live birth	-0.06 [0.04]	-0.24 [0.15]
70% of pregnancies end in live birth	-0.13 [0.04]	-0.72 [0.15]
"High Association" cases in figure 1c and 1d	-0.32 [0.04]	-1.92 [0.13]
"Low Association" case in figure 1d	-0.02 [0.04]	0.08 [0.15]
"Threshold Value" case in figure 1c	-0.09 [0.05]	-0.42 [0.14]
Positive correlation between fertility and marriage "fitness"	-0.10 [0.04]	-0.55 [0.15]
Negative correlation between fertility and marriage "fitness"	-0.07 [0.04]	-0.31 [0.13]

Notes for table 4: For each scenario indicated, we executed 250 simulations on 250,000 hypothetical couples each time, using base parameters given in column 3 of table 3, except as indicated in that scenario's label. At the end of the simulation, we estimated a discrete time, proportional-hazards survival regression on the simulated dataset and converted the estimated log odds difference to a percentage-point risk difference, as described in the text and in the notes to figure 2. The average risk difference across the 250 simulations is what we report in column 2. We also simulated the process of drawing data on parents with children from a cross-sectional dataset like the Census, as described in the text and in the notes to figure 3, and estimated the difference in risk of having divorced as a function of sex of the first birth, using a linear probability regression. That average risk difference across the 250 simulations is what we report in column 3. In square brackets beneath each risk difference, we report the bootstrap standard error (i.e., the mean squared deviation across the 250 simulations from the reported son "effect").

Supplementary Materials

A Mathematical Details

A.1 Selective Live Birth

Here, we carry through the formal discussion on potential empirical implications of selective live birth, as a complement to the discussion in section 2.

Eq. (2) explicates the fact that differences in marriage fitness between parents of firstborn sons and parents of firstborn daughters is only observed *conditional* on live birth. Recall that equation:

$$\mathrm{E}\left[\Delta m^{\star}\left|\mathbf{x}, \sum_{p=1}^{P}\omega_p < P\right.\right] = \left(\delta_2 - \delta_1\right) + \underbrace{\left\{\mathrm{E}\left[\eta\left|\mathbf{x}, b=1, \sum_{p=1}^{P}\omega_p < P\right.\right] - \mathrm{E}\left[\eta\left|\mathbf{x}, g=1, \sum_{p=1}^{P}\omega_p < P\right.\right]\right\}}_{\boldsymbol{\Delta}_s}$$

And recall that interpreting $\left(\hat{\delta}_2 - \hat{\delta}_1\right)$ in Eq. (2) as reflecting *nothing but* the parental divorce risk that daughters *would have faced* if they had been male requires assuming that $\boldsymbol{\Delta}_s = 0$.

To examine this critical assumption, we will clean up notational clutter by suppressing the $\mathbf{x}$ in conditional expressions, and introduce a few new terms. Denote the probability that male pregnancy survives to live birth with π^b, and for female pregnancies, π^g; the weighted average of these will be Π. We will denote the probability density function of η with $f(\eta)$, and note that the assumption that sex of pregnancy is assigned as if by coin flip also implies that $f(\eta\,|b=1\,) = f(\eta\,|g=1\,)$. With all that notation, and applying Bayes' rule and the definition of the expected value function, we arrive at this expression for $\boldsymbol{\Delta}_s$:

$$\begin{aligned}\boldsymbol{\Delta}_s &= \int \eta \tfrac{f(\eta|b=1\,)}{(1-\Pi)^{P-1}\pi^b}\,d\eta - \int \eta \tfrac{f(\eta|g=1\,)}{(1-\Pi)^{P-1}\pi^g}\,d\eta \\ &= \int \eta \tfrac{f(\eta)}{(1-\Pi)^{P-1}}\left[\tfrac{1}{\pi^b} - \tfrac{1}{\pi^g}\right]\,d\eta\end{aligned}$$

Now, suppose $\pi^b \approx 1$ and $\pi_g \approx 1$ (that is, almost all pregnancies end in live birth). It follows then that for any couple $P = b + g$, so that all denominators above would be about equal to 1. As long all the covariates *other* than b and g (if any) are properly chosen, then $\int \eta f(\eta) d\eta = 0$, so Δ_s would reduce to zero. In that

sense, belief that $\pi^b \approx 1$ and $\pi^g \approx 1$ could implicitly underlie an inclination to assign a counterfactual interpretation to $\left(\hat{\delta}_2 - \hat{\delta}_1\right)$. However, in the text and in online supplement B.1, we discuss evidence that survival probabilities may be considerably less than 1.

Alternatively, suppose that $\pi^g \approx \pi^b$, but both are less than 1. In that case, the term in square brackets would reduce to zero, and again the assumption that $\Delta_s = 0$ would be reasonable. In that sense, a counterfactual interpretation of $\left(\hat{\delta}_2 - \hat{\delta}_1\right)$ may be justified based on a belief that $\pi^g \approx \pi^b$. However, in the text and in online supplement B.2 we have discussed evidence that there may be a substantial prenatal female survival advantage.

Finally, suppose that $\pi^b < \pi^g$, but both survival probabilities and P are constant with respect to η. In that case, the denominator and the term in square brackets can be factored out of the integral. As long as covariates *other* than b and g (if any) are properly chosen, then a counterfactual interpretation is justified. However, in section 3 we have discussed evidence that at least one factor that is left in η in almost any marital "fitness" regression– namely, stress– may be negatively associated with π^b and π^g.

Most generally, a *purely* counterfactual interpretation of $\left(\hat{\delta}_2 - \hat{\delta}_1\right)$ effectively rests on this **assumption**:

$$\frac{\partial}{\partial \eta}\left(\frac{1}{(1-\Pi)^{P-1}}\left[\frac{1}{\pi^b} - \frac{1}{\pi^g}\right]\right) = 0 \tag{A.1}$$

Our study investigates what could *realistically* be the consequences of a failure of that assumption.

We conclude this discussion by emphasizing that assumption (A.1) is **not** equivalent to assuming that the impact of η on female pregnancies is equal to its impact on male pregnancies. In fact, assumption (A.1) probably cannot hold unless exactly the *opposite* is true. **In order for assumption** (A.1) **to hold in the presence of underlying female survival advantage, unmeasured characteristics in η (including biological stress) must be *more* pernicious (in an absolute sense) for female pregnancies than male ones.**

The same dynamics could also be observed using the principle of the force of mortality (Keyfitz and Caswell, 2005). Consider a cohort aging from fertilized egg (at time 0) to live birth (at time 1). Represent female survivors as of time $s \leq 1$:

$$l_F(s) = l_F(0) \exp\left\{-\int_0^s \mu(t)\, dt\right\}$$

and male survivors:

$$l_M(s) = l_M(0) \exp\left\{-\int_0^s a(t)\mu(t)\, dt\right\}$$

Where $\mu(\cdot)$ represents the force of mortality and $a \geq 1$ indicates the size of any prenatal female survival advantage. It follows that the sex ratio (males per female) in the cohort at time s will be

$$r(s) = r(0) \exp\left\{ \int_0^s (1 - a(t))\, \mu(t)\, dt \right\}$$

Since $\mu(t) \geq 0$, it follows that this ratio will decline with an upward shift in $\mu(\cdot)$ as long as the shift does not have an offsetting effect on $a(\cdot)$, and as long as $a(t) > 1$ for at least some t. Therefore if there is heterogeneity in the population with respect to $\mu(\cdot)$, then (by Bayes' rule) at time 1, average fecundity will be lower among parents of daughters than among parents of sons.

It remains an empirical question whether assumption (A.1) holds. We find it unsettling.

In our simulation tool, we require that the effect of biological stress on pregnancy survival is independent of the sex of the pregnancy.

A.2 Simulation Tool

Here, we describe the mathematical details underlying the simulation tool, as a complement to the discussion in section 4.2.

At the outset of each simulation, every couple is randomly assigned four characteristics:

- An underlying **fertility behavior/fecundability** (ν). This parameter establishes a distribution across the couples in terms of likelihood of establishing a pregnancy in any given period.
- An underlying **fecundity** (μ). This parameter establishes a distribution across the couples in terms of the likelihood that a pregnancy, once established, will survive to live birth.
- An underlying **marital stability** (λ). This parameter establishes a distribution across the couples in terms of the likelihood that they will divorce in any given period.
- A level of biological **stress** (θ). Specifically, the cohort is divided into 1001 categories, from the "least stressed" 0.1% of the population ($\theta = -500$), to the "most stressed" 0.1% ($\theta = 500$).

Here is how the simulation proceeds.

New Pregnancies

The simulation tool follows these steps to assign pregnancies to still intact couples who are not already pregnant:

1. It computes for each intact couple who is not pregnant, a period-and-parity specific latent fecundability, using the following equation:

$$P^{\star}_{i,t} = \nu_i + K^P_{t,c} - \eta_{1,i,t}$$

   $K^P_{t,c}$ is a period-specific constant chosen to ensure that the fraction of couples with c children already born who get pregnant will match the pre-specified overall parity-specific pregnancy hazard for that period, and η_1 is an independent random draw from a standard normal distribution.

2. It assigns a pregnancy to all couples for whom the latent fecundability score is greater than or equal to zero, or equivalently:

$$P_{i,t} = \begin{cases} 1 & \text{if} \quad \nu_i + K^P_{t,c} \geq \eta_{1,i,t} \\ 0 & \text{if} \quad \nu_i + K^P_{t,c} < \eta_{1,i,t} \end{cases}$$

3. It assigns a sex to each new pregnancy, as follows:

$$M_{i,t} = \begin{cases} 1 & \text{if} \quad \Phi^{-1}\left(\frac{r_0}{1+r_0}\right) \geq \eta_{2,i,t} \\ 0 & \text{if} \quad \Phi^{-1}\left(\frac{r_0}{1+r_0}\right) < \eta_{2,i,t} \end{cases}$$

   $\Phi^{-1}(\cdot)$ represents the inverse cumulative normal distribution, and η_2 is an independent random draw from a standard normal distribution.

New Births

The simulation tool follows these steps to determine the outcome of pregnancies established in the previous period:

1. It computes for each pregnancy a latent "fitness" score, using the following equation:

$$L^{\star}_{i,t} = \mu_i + \alpha_1(\theta) + \gamma M_{i,t-1} + K^L_t - \eta_{3,i,t}$$

K_t^L is a constant chosen to ensure that the fraction of overall pregnancies matches the pre-specified overall mortality risk, and γ reflects the female survival advantage which is necessary to drive a decline in the sex ratio from r_0 at fertilization to r_1 at live birth. $\alpha_1(\theta)$ is a function that reflects the relationship between biological stress and pregnancy "fitness." In most of our simulation runs, this function is linear in form– $\alpha_1 \times \theta$, but we also experiment with more flexible forms. η_3 is an independent random draw from a standard normal distribution.

2. It assigns a first birth to all couples who have a pregnancy with nonnegative "fitness" score. Or, equivalently:

$$L_{i,t} = \begin{cases} 1 & \text{if} \quad \mu_i + \alpha_1(\theta) + \gamma M_{i,t-1} + K_t^L \geq \eta_{3,i,t} \\ 0 & \text{if} \quad \mu_i + \alpha_1(\theta) + \gamma M_{i,t-1} + K_t^L < \eta_{3,i,t} \end{cases}$$

3. If a couple has three live births, its family size is topcoded and it is removed from eligibility for future pregnancies (since birth hazards beyond the third are generally low enough that they cannot be reliably identified in data).

New Divorces

The simulation follows these steps to assign divorces to still intact couples:

1. It computes for each intact couple a period specific latent divorce "vulnerability" score, using the following equation:

$$d_{i,t}^{\star} = \lambda_i + \alpha_2(\theta) + K_t^d - \eta_{4,i,t}$$

K_t^d is a period-specific constant chosen to ensure that the fraction of couples who divorce in each period will match the pre-specified period-specific divorce hazard. η_4 is an independent random draw from a standard normal distribution.

2. It assigns a new divorce to all couples with nonnegative latent divorce "vulnerability." Equivalently:

$$d_{i,t} = \begin{cases} 1 & \text{if} \quad \lambda_i + \alpha_2(\theta) + K_t^d \geq \eta_{4,i,t} \\ 0 & \text{if} \quad \lambda_i + \alpha_2(\theta) + K_t^d < \eta_{4,i,t} \end{cases}$$

3. For every couple experiencing a divorce in time t, the simulation tool assigns to them a permanent indicator $D_{i,\tau} = 1$ for every $\tau \geq t$, and excludes them from any further followup.

B Survival Bias and Offspring Sex

Within a cohort of pregnancies, those reaching live birth may comprise a minority of the original cohort. Furthermore, the cohort of pregnancies may become increasingly female as it ages– reflecting the fact that male pregnancies more likely to drop out of the cohort.

B.1 The Extent of Selection into Live Birth

Human reproduction is strikingly inefficient when compared against that of other species. Even after an egg cell is fertilized, the probability that it will develop to live birth is smaller than may be widely appreciated. Epidemiologists have estimated this probability from as high as 70% to as low as 30% (Benagiano et al., 2010; Grudzinskas and Nysenbaum, 1985; Macklon et al., 2002). This range is wide, because measurement is difficult; even using state of the art techniques, a fertilized egg– or "zygote"– cannot be reliably detected until it is about 10 days old.

In four important longitudinal studies conducted in the United States (Wilcox et al., 1988), China (Wang et al., 2003), Bolivia (Nepomnaschy et al., 2004), and Guatemala (Vitzthum et al., 2006), epidemiologists recruited women volunteers who were in sexual unions, were not contracepting, and had no known fertility problems. Biological samples from these volunteers were then tested regularly for the signature hormone of early pregnancy, in an attempt to detect surviving zygotes and thereby to map out a survival curve for normal early pregnancies.

In these studies, 30-40% of observed pregnancies terminated without a live birth. Although these studies aimed to represent the broader population by specifically excluding couples who had demonstrated difficulty establishing a pregnancy, it is important to note that they did not involve probability-based, population representative sampling designs. To our knowledge, no study has combined this measurement approach with standard population survey sampling techniques in order to identify the prevalence of early pregnancy loss among non-contracepting humans.

These studies can only speak to the fraction of zygotes that make it to live birth, among those that survive the first 10 days. Even less is known about so-called "pre-implantation loss"– that is, the risk that a zygote will pass from a woman without ever implanting in the placenta. Distinguishing that event from a normal menstrual period is prohibitively expensive, even in small samples. Given these technological limits to measurement, physicians and epidemiologists have used clever approaches to assess the likely scale of this type

of loss in humans, including demographically comparing the expected and observed number of clinical pregnancies in populations (Roberts and Lowe, 1975).

Taken together, epidemiological evidence collected over the past 50 years indicates that live births may represent something on the order of a 50%-30% subsample of zygotes (Benagiano et al., 2010). When one accounts for the fact that about half of the zygotes lost contain major, fatal defects resulting from random errors in cell reproduction (Macklon et al., 2002), this still leaves only about 40-50% of normal, viable zygotes surviving to live birth.

B.2 Prenatal Female Survival Advantage

What determines which half of normal, viable zygotes become live births? Another body of evidence indicates that sex may be one factor; specifically, female zygotes appear overrepresented among those that survive to live birth.

Because sex does not change over the course of pregnancy, a prenatal female survival advantage will lead to a decrease in the number of males per female as a cohort of pregnancies progresses. Since the sex ratio at *live birth* is readily observed wherever vital registration is complete, pinning down the extent of prenatal female survival advantage "only" requires identifying the sex composition of normal zygotes as early as possible in gestation. If the sex ratio declines over the course of pregnancy– just as it does throughout the life course after birth– then this implies that the female survival advantage begins before birth.

Owing to high costs of measurement, studies have not been conducted that directly quantify sex ratios in very early pregnancy. Instead, studies have tried to shed light on this question by examining the products of induced abortions. A typical approach involves collecting biological material from embryos and fetuses that were medically aborted as early as 5 weeks into pregnancy. Many such studies were conducted decades ago, before technology existed that would allow parents or their physicians to know the sex when the pregnancy was aborted. In that case, it may be plausible that the sex ratio among the *aborted* embryos and fetuses represents the sex ratio among *all* embryos and fetuses at a similar stage of development. One of the earliest studies to do this (Kellokumpu-Lehtinen and Pelliniemi, 1984) found sex ratios of 164 males per 100 females among Finnish women in the first 8 weeks of pregnancy, and 111 males per 100 females for the 8-24 week stage. Comparable studies in different populations and at different times, have found similar patterns– the earlier a fetus or embryo is recovered in the course of a pregnancy, the greater is the probability that it is male (McMillen, 1979; Pergament et al., 2002).

Other research draws inferentially from in vitro fertilization (IVF) records. Depending on the strategy favored by parents and their physicians, an embryo produced by IVF can be introduced into a woman very early in its development, or it can be allowed to mature through several rounds of cell division first. The earlier an embryo is introduced into the woman, the less viable it is (Gardner et al., 1998). In populations around the globe, live births resulting from pregnancies involving further developed (and, thus, more stable) embryos are more likely to be male than those resulting from pregnancies involving the less stable embryos (Chang et al., 2009; Dean et al., 2010).

Based largely on evidence from studies like these, many epidemiologists have speculated that the sex ratio of newly formed viable zygotes may be substantially higher than the sex ratio of live births (Vatten and Skjaerven, 2004); this ratio may vary from as low as 110 male zygotes per 100 female to as many as 170 male per 100 female (Chahnazarian, 1988; Kellokumpu-Lehtinen and Pelliniemi, 1984; McMillen, 1979; Pergament et al., 2002).

B.3 Is Sex of Offspring Like a Coin Flip?

We have discussed evidence that survival bias may generate associations between fecundity and sex of offspring– even if the sex of *pregnancies* is randomly assigned. This would imply that the sex of live births would *not* be indpendent and identically distributed across parents– specifically, within a population, less fecund parents would be more likely to have daughters.

The sex composition of birth cohorts is remarkably stable over time and across place. In the absence of conscious manipulation, ratios typically fall within the range of 102-107 boys per 100 girls (Wilcox and Baird, 2011). The mechanisms generating variation within this range are still poorly understood, although the nature and extent variation of systematic in these ratios has been a topic of interest for over a quarter of a millenium (Laplace, 1781; Trivers and Willard, 1973; Éric Brian and Jaisson, 2007). Detecting patterns rests largely on observational studies, which risk discovering false positive associations driven by measurement error or failure to account for the multiplicity of tests of the same hypothesis (Gelman and Weakliem, 2008; Wilcox and Baird, 2011; Simpson, 2012; Maconochie and Roman, 1997).

On the other hand, *within* a population in the same time and place, the sex of live births has been observed to covary with some demographic characteristics. Perhaps the most salient example is the well-documented racial patterning in sex ratios at birth. Black women giving birth in the United States are more likely to have a girl than white women. This pattern is robust, stable over years, and has been widely noted in the epidemiology literature (Pergament et al., 2002; Chahnazarian, 1988; Davis et al., 2007; Marcus et al., 1998).

In this sense, the sex of live births is *not* like a coin flip– flipped many times, a coin will come up heads about the same fraction of times, whether a black woman or white woman is doing the flipping.

Other characteristics and behaviors that have been observed to statistically predict offspring sex relate directly with fecundity– they include nutritional status (Cagnacci et al., 2004; Song, 2012; Rosenfeld and Roberts, 2010), parental age (Nicolich et al., 2000; Jacobsen et al., 1999), when in the ovulation cycle the relevant insemination occurs (James, 2012), and coital frequency (Wadley and Martin, 1997). Sex ratios at birth have also been observed to covary with the occurrence of disruptive events (Hansen et al., 1999; Fukuda et al., 1998; Bruckner et al., 2010; Torche and Kleinhaus, 2012; Song, 2012).

All of these characteristics and behaviors remain controversial as candidate predictors of offspring sex. Many hypotheses have been proposed and tested to account for both the observed variability *within* a place and time of sex ratios at birth, and their consistency *across* place and time; consensus remains elusive. Factors affecting the viability of pregnancies may play a role (James, 2012; Chahnazarian, 1988; Pergament et al., 2002).

In light of this controversy, our approach does not directly specify any characteristics as determinants of sex ratio at birth. Instead, we draw two lessons from this literature– first, a fecundity-related mechanism is plausible, by which the sex of a couple's offspring could provide information about that couple's circumstances. Such a mechanism amounts simply to survival bias; variation in fecundity interacts with prenatal female survival advantage to generate small but systematic differences in the probability that a couple's offspring will be female. The second lesson we draw from the ongoing empirical controversy regarding determinants of sex ratio at birth is that the assumption that offspring sex is independent and identically distributed across all parents, in the way that the outcome of a coin flip would be, is not self-evident (Wilcox and Baird, 2011).

References

Amato, P. (2010). Research on divorce: Continuing trends and new developments. *Journal of Marriage and Family 72*(2), 650–666.

Ananat, E. O. and G. Michaels (2008). The effect of marital breakup on the income distribution of women with children. *Journal Of Human Resources 18*(3), 612–629.

Arck, P. C., M. Rücke, M. Rose, J. Szekeres-Bartho, A. J. Douglas, M. Pritsch, S. M. Blois, M. K. Pincus, N. Bärenstraus, J. W. Dudenhausen, K. Nakamura, S. Sheps, and B. F. Klapp (2008). Early risk factors for miscarriage: A prospective cohort study in pregnant women. *Reproductive Biomedicine 17*(1), 101–113.

Beck, L. A., P. R. Pietromonaco, C. J. DeBuse, S. I. Powers, and A. G. Sayer (2013). Spouses' attachment pairings predict neuroendocrine, behavioral, and psychological responses to marital conflict. *Journal of Personality and Social Psychology 2013*(Jun17), 1–37.

Bedard, K. and O. Deschênes (2005). Sex preferences, marital dissolution, and the economic status of women. *Journal Of Human Resources 15*(2), 411–434.

Belkin, L. (2010). Motherlode: Do daughters cause divorce? `http://parenting.blogs.nytimes.com/2010/09/07/do-daughters-cause-divorce/`. Accessed: July 2013.

Benagiano, G., M. Farris, and G. Grudzinskas (2010). The fate of fertilized human oocytes. *Reproductive Biomedicine Online 21*, 732–741.

Boklage, C. E. (2005). The epigenetic environment: Secondary sex ratio depends on differential survival in embryogenesis. *Human Reproduction 20*(3), 583–587.

Bruckner, T. A., R. Catalano, and J. Ahern (2010). Male fetal loss in the U.S. following the terrorist attacks of September 11, 2001. *Bmc Public Health 10*, 273–279.

Bzostek, S. H., S. McLanahan, and M. Carlson (2012). Mothers' repartnering after a nonmarital birth. *Social Forces 90*(3), 817–841.

Cagnacci, A., A. Renzi, S. Arangino, C. Alessandrini, and A. Volpe (2004). Influences of maternal weight on the secondary sex ratio of human offspring. *Human Reproduction 19*(2), 442–444.

Chahnazarian, A. (1988). Determinants of the sex ratio at birth: Review of recent literature. *Social Biology 35*(3-4), 214–235.

Chang, H. J., J. R. Lee, B. C. Jee, and C. S. Suh (2009). Impact of blastocyst transfer on offspring sex ratio and the monozygotic twinning rate: A systematic review and meta-analysis. *Fertility and Sterility 91*(6), 2381–2390.

Clark-Flory, T. (2010). More daughters, more divorce. `http://www.salon.com/2010/09/07/daughters_divorce/`. Accessed: July 2013.

Conley, D., K. M. Pfeiffer, and M. Velez (2007). Explaining sibling differences in achievement and behavioral outcomes: The importance of within- and between-family factors. *Social Science Research 36*(2007), 1087–1104.

Cox, M. J., B. Paley, M. Burchinal, and C. C. Payne (1999). Marital perceptions and interactions across the transition to parenthood. *Journal Of Marriage And The Family 61*(3), 611–625.

Crimmins, E. M., J. K. Kim, and T. E. Seeman (2008). Poverty and biological risk: The earlier "aging" of the poor. *Journals Of Geronotology 64*(A2), 286–292.

Crouter, A. C., S. M. Mchale, and W. T. Bartko (1993). Gender as an organizing feature in parent-child relationships. *Journal Of Social Issues 49*, 161–174.

Dahl, E., R. S. Gupta, M. Beutel, Y. Stoebel-Richter, B. Brosig, H.-R. Tinneberg, and T. Jain (2006). Preconception sex selection demand and preferences in the United States. *Fertility and Sterility 85*(2), 468–473.

Dahl, G. B. and E. Moretti (2008). The demand for sons. *Review of Economic Studies 75*, 1085–1120.

Das Gupta, M. (2005). Explaning Asia's missing women: A new look at the data. *Population and Development Review 31*(3), 529–533.

Davis, D. L., P. Webster, H. Stainthorpe, J. Chilton, L. Jones, and R. Doi (2007). Declines in sex ratio at birth and fetal deaths in Japan, and in U.S. Whites but not African Americans. *Environmental Health Perspectives 115*(6), 941–946.

Dean, J. H., M. G. Chapman, and E. A. Sullivan (2010). The effect on human sex ratio at birth by assisted reproductive technology (ART) procedures. *BJOG: An International Journal of Obstetrics & Gynecology 117*(3), 1628–1634.

Emery, R. E., M. Waldron, K. M. Kitzmann, and J. Aaron (1999). Delinquent behavior, future divorce or nonmarital childbearing, and externalizing behavior among offspring: A 14-year prospective study. *Journal of Family Psychology 13*(4), 568–579.

Éric Brian and M. Jaisson (2007). *The Descent of the Human Sex Ratio at Brith*, Volume 4 of *Methods*. Springer.

Ferin, M. (1999). Stress and the reproductive cycle. *The Journal Of Clinical Endocrinology and Metabolism 84*(6), 1768–1774.

Fremont, A. M. and C. E. Bird (2000). Social and psychological factors, physiological processes, and physical health. In C. E. Bird, P. Contrad, and A. M. Fremont (Eds.), *Handbook Of Medical Sociology*, Chapter 23, pp. 334–364. Prentice Hall.

Friedman, E. M., A. S. Karlamangla, D. M. Almeida, and T. E. Seeman (2012). Social strain and cortisol regulation in midlife in the U.S. *Social Science and Medicine 74*(2012), 607–615.

Fukuda, M., K. Fukuda, T. Shimizu, and H. Møller (1998). Decline in sex ratio at birth after the Kobe earthquake. *Human Reproduction 13*(8), 2321–2322.

Gardner, D. K., P. Vella, M. Lane, L. Wagley, T. Schlenker, and W. B. Schoolcraft (1998). Culture and transfer of human blastocysts increases implantation rates and reduces the need for multiple embryo transfers. *Fertility and Sterility 69*(1), 84–88.

Gelman, A. and D. Weakliem (2008, October 27). Of beauty, sex, and power: Statistical challenges in estimating small effects. Accessed: July 2013.

Gourley, S. L., A. M. Swanson, and A. J. Koleske (2013). Corticosteroid-induced neural remodeling predicts behavioral vulnerability and resilience. *The Journal of Neuroscience 33*(7), 3107–3112.

Grudzinskas, J. G. and A. M. Nysenbaum (1985). Failure of human pregnancy after implantation. *Annals Of The New York Academy Of Sciences 442*, 38–44.

Gunlicks-Stoessel, M. L. and S. I. Powers (2009). Romantic partners' coping strategies and patterns of cortisol reactivity and recovery in response to relationship conflict. *Journal of Social and Clinical Psychology 28*(5), 630–649.

Hansen, D., H. Møller, and J. Olsen (1999). Severe periconceptional life events and the sex ratio in offspring: Follow up study based on five national registers. *British Medical Journal 319*, 548–549.

Harris, K. M. and S. P. Morgan (1991). Fathers, sons, and daughters: Differential paternal involvement in parenting. *Journal Of Marriage And The Family 1991*, 531–544.

Hennessy, M. B., S. Kaiser, and N. Sachser (2009). Social buffering of the stress response: Diversity, mechanisms, and functions. *Frontiers in Neuroendocrinology 30*(2009), 470–482.

Het, S., D. Schoofs, N. Rohleder, and O. T. Wolf (2012). Stress-induced cortisol level elevations are associated with reduced negative affect after stress: Indications for a mood-buffering cortisol effect. *Psychosomatic Medicine 74*(1), 23–32.

Hiedemann, B., J. M. Joesch, and E. Rose (2004). More daughters in child care? Child gender and the use of nonrelative child care arrangements. *Social Science Quarterly 85*(1), 154–168.

Hofferth, S. and K. G. Anderson (2003). Are all dads equal? Biology versus marriage as a basis for paternal investment. *Journal Of Marriage And The Family 65*(1), 213–232.

Hutchison, C. (2010). Couples with dauthers more likely to divorce. `http://abcnews.go.com/Health/Wellness/daughters-breed-divorce-boys-blessing-daughters-curse/story?id=11804444`. Accessed: July 2013.

Jacobsen, R., H. Möller, and A. Mouritsen (1999). Natural variation in the human sex ratio. *Human REproduction 14*(12), 3120–3125.

James, W. H. (2012). Hypotheses on the stability and variation of human sex ratios at birth. *Journal Of Theoretical Biology 310*(2012), 183–186.

Katzev, A. R., R. L. Warner, and A. C. Acock (1994). Girls or boys? Relationship of child gender to marital instability. *Journal Of Marriage And The Family 56*(1), 89–100.

Kellokumpu-Lehtinen, P. and L. J. Pelliniemi (1984). Sex ratio of human conceptuses. *Obstetrics and Gynecology 64*(2), 220–222.

Kelly, A. E. (2010). Psychology today insight: Why parents of girls divorce more. `http://www.psychologytoday.com/blog/insight/201008/why-parents-girls-divorce-more`. Accessed: July 2013.

Kennedy, S. and L. L. Bumpass (2007). Cohabitation and children's living arrangements: New estimates from the United States. *Demographic Research 19*(47), 1663–1692.

Keyfitz, N. and H. Caswell (2005). *Applied Mathematical Demography*. Statistics for Biology and Health. Springer.

Landsburg, S. E. (2003a). Maybe parents don't like boys better: A follow-up to the recent column about whether daughters cause divorce. `http://www.slate.com/articles/arts/everyday_economics/2003/10/maybe_parents_dont_like_boys_better.html`. Accessed: July 2013.

Landsburg, S. E. (2003b). Oh no: It's a girl! Do daughters cause divorce? `http://www.slate.com/articles/arts/everyday_economics/2003/10/oh_no_its_a_girl.html`. Accessed: July 2013.

Laplace, P. S. (1781). Mémoire sur les probabilités. *Mémoires de l'Académie Royale des Science de Paris VI*, 621–656.

Lauderdale, D. S. (2006). Birth outcomes for Arabic-named women in California before and after September 11. *Demography 43*(1), 185–201.

Laurent, H. K., S. I. Powers, and D. A. Granger (2013). Refining the multisystem view of the stress response: Coordination among cortisol, alpha-amylase, and subjective stress in response to relationship conflict. *Physiology and Behavior 119*(2013), 52–60.

Lawn, J. E., H. Blencowe, G. L. Darmstadt, and Z. A. Bhutta (2013). Beyond newborn survival: The world you are born into determines your risk of disability-free survival. *Pediatric Research*. Epub ahead of print.

Lhila, A. and K. I. Simon (2008). Prenatal health investment decisions: Does the child's sex matter? *Demography 45*(4), 885–905.

Lundberg, S. (2005). Sons, daughters, and parental behaviour. *Oxford Review of Economic Policy 21*(3), 340–356.

Lundberg, S., S. McLanahan, and E. Rose (2007). Child gender and father involvement in Fragile Families. *Demography 44*(1), 79–92.

Lundberg, S. and E. Rose (2003). Child gender and the transition to marriage. *Demography 40*(2), 333–349.

Lupien, S. J., B. S. McEwen, M. R. Gunnar, and C. Heim (2009). Effects of stress throughout the lifespan on the brain, behavior and cognition. *Nature Reviews: Neuroscience 10*(6), 434–445.

Lytton, H. and D. M. Romney (1991). Parents' differential socialization of boys and girls: A meta-analysis. *Psychological bulletin 109*(2), 267–296.

Macklon, S., J. P. M. Geraedts, and B. C. J. M. Fauser (2002). Conception to ongoing pregnancy: The "black box" of early pregnancy loss. *Human Reproduction Update 8*(4), 333–343.

Maconochie, N., P. Doyle, S. Prior, and R. Simmons (2006). Risk factors for first trimester miscarriage: Results from a UK population-based case-control study. *BJOG: An International Journal of Obstetrics and Gynecology 114*(1), 170–186.

Maconochie, N. and E. Roman (1997). Sex ratios: Are there natural variations within the human population? *BJOG: An International Journal of Obstetrics and Gynecology 104*(9), 1050–1053.

Mammen, K. (2008). The effect of children's gender on living arrangements and child support. *American Economic Review: Papers and Proceedings 98*(2), 408–412.

Mansour, H. and D. I. Rees (2012). Armed conflict and birthweight: Evidence from the al-Aqsa Intifada. *Journal of Development Economics 99*(1), 190–199.

Manton, K. G., M. A. Woodbury, and E. Stallard (1995). Sex differences in human mortality and aging at late ages: The effect of mortality selection and state dynamics. *The Gerontologist 35*(5), 597–608.

Marcus, M., J. Kiely, F. Xu, M. Mcgeehin, R. Jackson, and T. Sinks (1998). Changing sex ratio in the United States 1969-1995. *Fertility and Sterility 70*(2), 270–273.

McEwen, B. S. (2003). Mood disorders and allostatic load. *Biological Psychology 54*(2003), 200–2007.

McEwen, B. S. and J. C. Wingfield (2003). The concept of allostasis in biology and medicine. *Hormones and Behavior 43*(2003), 2–15.

McMillen, M. M. (1979). Differential mortality by sex in fetal and neonatal deaths. *Science 204*(4388), 89–91.

Meuwly, N., G. Bodenmann, J. Germann, T. N. Bradbury, B. Ditzen, and M. Heinrichs (2012). Dyadic coping, insecure attachment, and cortisol stress recovery following experimentally induced stress. *Journal of Family Psychology 12*(6), 937–947.

Miller, G. E., E. Chen, and E. S. Zhou (2007). If it goes up, must it come down? Chronic stress and the hypothalamic-pituitary-adrenocortical axis in humans. *Psychological Bulletin 133*(1), 25–45.

Morgan, S. P., D. N. Lye, and G. A. Condran (1988). Sons, daughters, and the risk of marital disruption. *American Journal Of Sociology 94*(1), 110–129.

Morgan, S. P. and M. S. Pollard (2002). Do parents of girls really have a higher risk of divorce? `www.soc.duke.edu/~pmorgan/PAA02.MorganPollard.pdf`. Accessed: July 2013.

Mott, F. L. (1994). Sons, daughters and fathers' absence: Differentials in father-leaving probabilities and in home environments. *Journal Of Family Issues 15*, 97–128.

Nakamura, K., S. Sheps, and P. C. Arck (2008). Stress and reproductive failure: Past notions, present insights and future directions. *Journal of Assisted Reproduction and Genetics 25*(1), 47–62.

Nepomnaschy, P. A., K. Welch, D. Mcconnell, B. I. Strassmann, and B. G. England (2004). Stress and female reproductive function: A study of daily variations in cortisol, gonadotrophins, and gonadal steroids in a rural Mayan population. *American Journal Of Human Biology 16*, 523–532.

Nepomnaschy, P. A., K. B. Welch, D. S. McConnell, B. S. Low, B. I. Strassmann, and B. G. England (2008). Cortisol levels and very early pregnancy loss in humans. *Proceedings of the National Academy of Sciences 103*(10), 3938–3942.

Nicolich, M. J., W. W. Huebner, and A. R. Schnatter (2000). Influence of parental and biological factors on the male birth fraction in the United States: An analysis of birth certificate data from 1964 through 1988. *Fertility and Sterility 73*(3), 487–492.

Parker, V. J. and A. J. Douglas (2010). Stress in early pregnancy: Maternal neuro-endocrine-immune responses and effects. *Journal of Reproductive Immunology 85*(2010), 86–92.

Pebley, A. R. and C. F. Westoff (1982). Women's sex preferences in the United States: 1970 to 1975. *Demography 19*(2), 177–189.

Pergament, E., P. B. Todydemir, and M. Fiddler (2002). Sex ratio: A biological perspective on "Sex and the City". *Reproductive Biomedicine Online 5*(1), 43–46.

Powers, S. I., P. R. Pietromonaco, M. Gunlicks, and A. Sayer (2006). Dating couples' attachment styles and patterns of cortisol reactivity and recovery in response to a relationship conflict. *Journal of Personality and Social Psychology 90*(4), 613–628.

Radloff, L. S. (1977). The CES-D scale: A self-report depression scale for research in the general population. *Applied Psychological Measurement 1*(3), 385–401.

Raley, S. and S. Bianchi (2006). Sons, daughters, and family processes: Does gender of children matter? *Annual Review of Sociology 32*(2006), 401–421.

Rawdin, B., S. Mellon, F. Dhabhar, E. Epel, E. Puterman, Y. Sue, H. Burke, V. Reus, R. Rosser, S. Hamilton, J. Nelson, and O. Wolkowitz (2013). Dysregulated relationship of inflammation and oxidative stress in major depression. *Brain, Behavior, and Immunity 31*(2013), 143–152.

Roberts, C. J. and C. R. Lowe (1975). Where have all the conceptions gone? *The Lancet 305*(7905), 2184–2189.

Rosenfeld, C. S. and R. M. Roberts (2010). Maternal diet and other factors affecting offspring sex ratio: A review. *Biology of Reproduction 71*(4), 1063–1070.

Ruttle, P. L., E. A. Shirtcliff, L. A. Serbin, D. B.-D. Fisher, D. M. Stack, and A. E. Schwartzman (2011). Disentangling psychobiological mechanisms underlying internalizing and externalizing behaviors in youth: Longitudinal and concurrent associations with cortisol. *Hormones and Behavior 59*(2011), 123–132.

Sassler, S., A. Cunningham, and D. T. Lichter (2009). Intergenerational patterns of union formation and relationship quality. *Journal of Family Issues 30*(6), 757–786.

Seeman, T., E. Epel, T. Gruenewald, A. Karlamangla, and B. S. McEwen (2010). Socioeconomic differentials in peripheral biology: Cumulative allostatic load. *Annals of the New York Academy of Sciences 1186*(1), 223–239.

Sen, A. (1992). Missing women: Social inequality outweighs women's survival advantage in Asia and North Africa. *British Medical Journal 304*(6827), 587–588.

Sen, A. (2003). Missing women– revisited. *British Medical Journal 327*, 1297–1298.

Shumow, L., D. L. Vandell, and J. K. Posner (1998). Harsh, firm, and permissive parenting in low-income families. *Journal Of Family Issues 19*(5), 483–507.

Simpson, L. (2012). Letter to the editor. *American Journal of Epidemiology 175*(9), 973.

Snyder, J. R. (1998). Marital conflict and child adjustment: What about gender? *Developmental Review 18*(1998), 398–420.

Song, S. (2012). Does famine influence sex ratio at birth? evidence from the 1959-1961 Great Leap Forward famine in China. *Proceedings of the Royal Society: Biological Sciences 279*(1739), 2883–2890.

Spanier, G. B. and P. C. Glick (1981). Marital instability in the United States: Some correlates and recent changes. *Family Relations 30*, 329–338.

Steelman, L. C., B. Powell, R. Werum, and S. Carter (2002). Reconsidering the effects of sibling configuration: Recent advances and challenges. *Annual Review of Sociology 28*(1), 243–269.

Steinsaltz, D. (2013, January). What do we think we know about prenatal sex ratio, and when did we think we knew it? Berkeley Population Center Symposium.

Sugiura-Ogasawara, M., T. A. Furukawa, Y. Nakano, S. Hori, K. Aoki, and T. Kitamura (2002). Depression as a potential causal factor in subsequent miscarriage in recurrent spontaneous aborters. *Human Reproduction 17*(10), 2580–2584.

Teachman, J. D. and P. T. Schollaert (1989). Gender of children and birth timing. *Demography 26*(3), 189–199.

Torche, F. (2011). The effect of maternal stress on birth outcomes: exploiting a natural experiment. *Demography 48*(4), 1473–1491.

Torche, F. and K. Kleinhaus (2012). Prenatal stress, gestational age and secondary sex ratio: The sex-specific effects of exposure to a natural disaster in early pregnancy. *Human Reproduction 27*(2), 558–567.

Trivers, R. L. and D. E. Willard (1973). Natural selection of parental ability to vary the sex ratio of offspring. *Science 179*(4068), 90–92.

Tucker, C. J., S. M. Mchale, and A. C. Crouter (2003). Dimensions of mothers' and fathers' differential treatment of siblings: Links with adolescents' sex-typed personal qualities. *Family Relations 52*(1), 82–89.

van Eck, M. M. M., N. A. Nicholson, H. Berkhof, and J. Sulon (1996). Individual differences in cortisol responses to a laboratory speech task and their relationship to responses to stressful daily events. *Biological Psychology 43*(1996), 69–84.

Vatten, L. J. and R. Skjaerven (2004). Offspring sex and pregnancy outcome by length of gestation. *Early Human Development 76*(1), 47–54.

Vaupel, J. and A. Yashin (1985). Heterogeneity's ruses: Some surprising effects of selection on population dynamics. *The American Statistician 39*(3), 176–185.

Vitzthum, V. J., H. Spielvogel, J. Thornburg, and B. West (2006). A prospective study of early pregnancy loss in humans. *Fertility and Sterility 86*(2), 373–379.

Wade, T. J. and D. J. Pevalin (2004). Marital transitions and mental health. *Journal of Health and Social Behavior 45*(2), 155–170.

Wadley, R. L. and J. F. Martin (1997). On secondary sex ratios and coital frequency with an Iban case. *Current Anthropology 38*(1), 79–81.

Waffarn, F. and E. P. Davis (2012). Effects of antenatal corticosteroids on the hypothalamic-pituitary-adrenocortical axis of the fetus and newborn: Experimental findings and clinical considerations. *American Journal Of Obstetrics and Gynecology 207*(6), 446–454.

Wang, X., C. Chen, L. Wang, D. Chen, W. Guang, and J. French (2003). Conception, early pregnancy loss, and time to clinical pregnancy: A population based prospective study. *Fertility and Sterility 79*(3), 577–584.

Wilcox, A. J. and D. D. Baird (2011). Natural versus unnatural sex ratios–a quandry of modern times. *American Journal of Epidemiology 174*(12), 1332–1334.

Wilcox, A. J., C. R. Weinberg, J. F. O'Connor, D. D. Baird, J. P. Schlatterer, R. E. Canfield, E. G. Armstrong, and B. C. Nisula (1988). Incidence of early loss of pregnancy. *The New England Journal Of Medicine 319*(4), 189–194.

Center for Demography and Ecology
University of Wisconsin
1180 Observatory Drive Rm. 4412
Madison, WI 53706-1393
U.S.A.
608/262-2182
FAX 608/262-8400
comments to: jnobles@wisc.edu
requests to: cdepubs@ssc.wisc.edu