PREMARITAL COHABITATION AND MARITAL DISSOLUTION:

A REPLY TO MANNING, SMOCK AND KUPERBERG

Michael J. Rosenfeld *

Katharina Roesler †

PUBLISHED IN THE *JOURNAL OF MARRIAGE AND FAMILY,* 2020 EARLY ONLINE

OUR ORIGINAL ARTICLE (RR 2019) IS HERE: https:/doi.org/10.1111/jomf.12530

MANNING, SMOCK AND KUPERBERG'S COMMENT IS HERE: https://doi.org/10.1111/jomf.12724

THE JMF FORMATTED VERSION OF OUR REPLY IS HERE: https://doi.org/10.1111/jomf.12744

* Michael J. Rosenfeld, Department of Sociology, Stanford University, 450 Jane Stanford Way, Stanford, CA 94305. Email: mrosenfe@stanford.edu. Web: www.stanford.edu/~mrosenfe.

† Katharina Roesler, Square Inc., San Francisco CA. Email: katroesler@gmail.com.

Acknowledgements: Roesler, with feedback from Rosenfeld, prepared (with substantial commitment of time) the harmonized NSFG event history dataset used in the paper. Rosenfeld, with feedback from Roesler, performed the analyses and wrote the paper. We thank Stanford's Graduate Family Workshop for comments on previous drafts.

Abstract:

Objective: Our goal is to show how premarital cohabitation's association with marital dissolution can be measured consistently over time.

Background: Rosenfeld and Roesler (2019) showed that premarital cohabitation led to lower rates of marital dissolution early in marital duration, but higher rates of marital dissolution after a few years of marital duration. Analysis based on marriage cohorts can erroneously appear to show that the association between premarital cohabitation and marital dissolution has disappeared in the most recent cohort.

Method: We use discrete time event history models to study the association between premarital cohabitation and marital dissolution across calendar time and across marriage cohorts, with data from the NSFG. We examine a variety of modeling strategies and data filters including those in Manning, Smock, and Kuperberg's (MSK) comment to show how different data and modeling choices can bias the results.

Results: MSK's analysis rests on simple misunderstandings of our models. MSK discarded more than 70% of the valid couple years available in the NSFG data, and then used the resulting lower statistical power to declare a key interaction insignificant. MSK treat children as a time-invariant variable despite the fact that the presence of children changes over the course of relationships. MSK discard one wave of the data without explanation. All of these choices affect the outcome.

Conclusion: Our prior result stands: premarital cohabitation consistently predicts higher rates of marital dissolution in the U.S. Research into marital dissolution should be made more robust, transparent, and replicable.

Keywords: Cohabitation, Divorce, Marriage, Relationship Dissolution, Event history analysis

Running head: Premarital Cohabitation and Divorce: Reply to MSK

Premarital Cohabitation and Marital Dissolution: A reply to Manning, smock and Kuperberg

Introduction:

We are happy to receive the Manning, Smock, and Kuperberg (hereafter MSK) comment on our paper (Rosenfeld and Roesler 2019; hereafter RR 2019). We stand by our results. Premarital cohabitation has a stable association over time with higher rates of marital dissolution in the U.S.

Ideally, we would show readers a version of this comment and debate wherein all sides worked from the same data extract and defined the variables the same way. That would have been a more productive exchange. However, MSK never reached out to us with questions. We would have been happy to send them a replication package, but they never requested one. Instead of basing their critique on our actual data and models, their analysis of our work relied on speculation resulting in a variety of errors that should have been easy to avoid. We asked MSK for a replication package for their data analysis, but they did not provide one. We have now made replication packages for our work publicly available (Rosenfeld and Roesler 2020a, Rosenfeld and Roesler 2020b).

MSK's results are driven by several choices they never properly explain. First, MSK drop the 1988 wave from their analysis without explanation. Second, they treat fertility and children in each household as time invariant, when in fact people marry in part to have and raise children. Third, by filtering out couples who married more than 10 years before the surveys they eliminate more than 70% of the couple-years of available data. By discarding most of the available data, MSK have reduced statistical power that leads them to one key erroneous null finding. MSK overlook findings in their own tables that do not fit their main argument, and they write inaccurately about our results as well.

.

Background, and the Recent Cohort Fallacy:

In RR 2019 we showed that in the very early stages of their marriages, couples who had previously cohabited had an advantage in marital stability. We referred to this advantage as the Practical Experience of Cohabitation, by which we mean the advantage that cohabiters have in the practical daily experience of living together compared to couples who have never lived together before. This advantage dissipated after a year or two of marital duration, and by year 5 of marital duration, the couples who had not cohabited before marriage had a significantly better cumulative marital stability. Because the marital stability advantage of non-cohabiters takes a few years to manifest, the most recent marital cohort will always seem to be free of the association (between premarital cohabitation and marital dissolution), even if the association is constant over time. Let's call this the Recent Cohort Fallacy. As time goes by and marital duration extends beyond 5 years, we find that the association between premarital cohabitation and marital dissolution is quite consistent across marriage cohorts in the U.S.

The Recent Cohort Fallacy has analogues in other areas of family demography. Changes in the life course timing of demographic transitions like birth, marriage and divorce can mean that models based on data from the past might not apply well to the most recent cohorts, and can lead to prediction errors. The delay of marriage caused demographers to predict that women in then-recent birth cohorts were destined to never marry in large numbers, but most did eventually marry (Cherlin 1992). Similarly, Goldstein el al (2009) showed that a delay in women's age at childbirth can lead to under-estimates of the Total Fertility Rate, and therefore an under-estimate of what the lifetime fertility of the more recent birth cohorts will be.

In the National Survey of Family Growth (hereafter NSFG) data for all couples in first marriages for 2005-2016 (using waves through 2015-2017), the odds of actual marital dissolutions (separation or divorce) were 1.32 times higher (95% Confidence Interval 1.20-1.46) for couples who had cohabited before marriage. From 1970-2004, the same odds ratio was 1.31 (95% Confidence Interval 1.25-1.37). Premarital cohabitation consistently predicts higher rates of marital dissolution. Is MSK's claim (that the association between premarital cohabitation and marital dissolution has disappeared) consistent with the actual stability of the association across calendar time? We do not think so. The odds of divorce in NSFG were 1.22 (95% CI 1.11-1.32) times higher in 2005-2016 for married couples who had cohabited before marriage, compared to married couples who had never cohabited. The same odds ratio was 1.18 (95% CI 1.12-1.24) for the 1970-2004 period. These are raw odds ratios, weighted but uncorrected for other factors that predict marital dissolution. As we showed in Table 1 and Figure 2 of RR 2019, correcting for all the predictors of marital dissolution (including marital duration, age at marriage, education, mother's education, parental marital stability, children in the household, and race) in NSFG does not alter the consistency over time in this key association. These odds ratios make a comparison across calendar time (examining the marital dissolution rate in year X for all couples whose marriages were intact in year X regardless of marital duration) rather than across marriage cohorts, so they are not subject to the Recent Cohort Fallacy.

The Recent Cohort Fallacy is one reason why interactions between premarital cohabitation and marriage cohort can appear to show that premarital cohabitation is no longer associated with marital dissolution, despite the consistency of the association over calendar time. All of the changes in the marriage system of the U.S. since 1970 also contribute to the difficulty of modeling marriage and divorce. Between 1970 and the present, every aspect of marriage and divorce in the U.S. has changed. Premarital cohabitation has gone from being rare to being the norm. Age at marriage has risen sharply. The divorce rate rose, then plateaued, and recently has declined sharply. There is no simple way to control away all the underlying changes so that premarital cohabitation can be studied net of everything else. In short, all models of complex reality are flawed. In RR 2019 we showed results from a variety of models not because we love variety or were trying to confuse readers, but to try to explain which kinds of findings were robust, and which findings were not robust. Even in models with interactions between premarital cohabitation and marriage cohort, which tend to erroneously show declines in the association between premarital cohabitation and marital dissolution, the significance of the apparent decline in cohabitation's effect on marital dissolution depends on other (seemingly unrelated) aspects of model design. We now turn briefly to the different aspects of model design, and how they affect the association of interest here.

Data and Models:

We use a harmonized event history dataset combining data on first marriages for women subjects (age 15-44 at the time of the various surveys) from NSFG waves 1988 through 2013-15, consistent with RR 2019. Premarital cohabitation means cohabitation with the woman's first husband before they were married. All the marriages in our dataset are heterosexual marriages, i.e women married to men. The outcome variable is formal divorce or marital separation, whichever comes first. The full dataset has 216,455 couple-years and 8,488 marital dissolutions among 24,888 women in first marriages. MSK invoke a variety of filters that reduce the dataset substantially, which we replicate one at a time and discuss below.

MSK briefly describe a variety of different aspects of model design, but they neglect to explain how the different modeling choices contribute to the result (the supposed disappearance of the association between premarital cohabitation and marital dissolution) they are keen to defend. We take up the discussion where they left off: we endeavor to explain how both our own and MSK's modeling choices influence the result.

MSK's modeling choices and their implications

a) Interacting premarital cohabitation with either calendar year or marriage cohort

MSK point out that previous literature on premarital cohabitation and marital dissolution relied mostly on interactions between premarital cohabitation and marriage cohort. In RR 2019 we interacted premarital cohabitation with *both* calendar time and marriage cohort (in separate models, see RR 2019 Table 1), and we explained why the marriage cohort interaction was vulnerable to what we call here a Recent Cohort Fallacy. MSK prefer to interact premarital cohabitation with marriage cohort, and their rationale seems to be that this is the traditional way to analyze the data. If scholars were limited only to analyzing data using the exact recipe in previously published work, new ideas and results would be hard to come by. We do not think that fealty to prior data analysis recipes is always a virtue.

MSK's Table 1, Model 3 interacts cohabitation with calendar time (odds ratio 0.99, P value 0.334 with the appropriate calendar time interaction), and finds no significant change over time. MSK's Table 1 offered them an opportunity to explain *why* the calendar time interaction in their Table 1 is not consistent with the marriage cohort interaction (odds ratio of 0.98 reported in their Table 1, Model 2 as significantly different from 1) but they did not do so. Our answer is straightforward: the Recent Cohort Fallacy explains why interactions between premarital cohabitation and marriage cohort yield misleading results. MSK's failure to explain the insignificant calendar year interaction with premarital cohabitation in their Table 1, Model 3 is a lost opportunity, especially since the choice (to interact premarital cohabitation with either marriage cohort or calendar year) substantively determines their results (in combination with other data analysis choices).

In discussion of their Table 1, MSK claim not to be sure whether our calendar year variable is fixed or varies over time. We return to this point below (along with other examples where they misconstrue our data and results), but here we simply note: our calendar year variable grows by 1 every year; how could it be otherwise? MSK assume, falsely, that our calendar year variable was fixed in time. MSK's errors arising from this false assumption could have been avoided if they had asked us about our coding or requested a replication package from us.

b) Too many measures of time?

We show in RR 2019 that the relationship between marital duration and marital dissolution is non-linear, especially for couples who cohabited before marriage. The non-linearity is an important new finding, which we fit with a dummy variable for the calendar year in which the marriage took place. In MSK's Table 2 (Model 1, all cohorts, and Model 2 for both the early and the recent cohort), they attempt to fit the nonlinear relationship between premarital cohabitation and marital duration with a single linear term, and find the linear term to be insignificant. They conclude from this exercise that "the interaction term is not statistically significant, indicating that cohabitation operates in a similar manner across marital durations (up to 10 years)." The problem here is that MSK were fitting a line to a curvilinear reality (the curvilinear reality we illustrated in RR 2019 Figure 3 and below in Figure 1), so they miss-fit the data. MSK argued that RR 2019 had too many interaction terms, however our view is that their models needed at least one more interaction term to fit the curvilinear reality shown in Figure 1. Some of MSK's confusion about our models owes to their failure to inquire into how our models were constructed.

Cox proportional hazard models, which have a built-in baseline hazard and therefore control for the association between marital duration and marital dissolution differently than in our discrete time event history models, are shown in Appendix Table 2. The Cox models have substantively the same results as our discrete time event history models in Table 1.

Instead of using a linear term, MSK's Appendix A shows results for a dummy variable interaction between the first year of marriage and premarital cohabitation, more consistent with our approach. MSK find the dummy variable interaction in Appendix A to not be significant, but their tests are all run on approximately one quarter of the full dataset, so their tests are under-powered.

[Figure 1 here]

c) Evidence in MSK's results of the interaction between premarital cohabitation, marital duration, and marital dissolution

Figure 1 (reprinted from RR 2019) shows the relationship between marital duration, marital dissolution, and premarital cohabitation in the NSFG data. Note how in the first 12 months after marriage, couples who cohabited before marriage have lower rates of dissolution than couples who did not cohabit. Although MSK claim that "cohabitation operates in a similar manner across marital durations" the data tell a different story. We would have preferred to generate Figure 1 from MSK's version of the NSFG data, but since MSK did not send us a replication package we rely on our own NSFG data extract throughout.

In MSK's Appendix Table A they attempt to replicate our findings (despite having a model design quite different from ours, see below) and they report an odds ratio of 0.77 for the lower rate of marital dissolution for cohabiters compared to non-cohabiters in the first year of marriage. Since MSK reported the coefficient as not significant, they dismiss this interaction.

MSK's models include less than a quarter of the valid NSFG data (through a series of data exclusions whose impact on sample size they fail to describe). Discarding data reduces power and can result in erroneous null findings. We calculated the power to exclude erroneously null findings for MSK's reported odds ratio of 0.77 using the full NSFG dataset and the known distribution of premarital cohabiters and the known breakup rate of marriages in the first year. With the full dataset and an odds ratio of 0.77 the power to reject a null finding would be 0.87 (using MSK's cutoff of alpha of 0.05, two sided test) or 0.70 (with a stricter cutoff of 0.01, two sided test). In the reduced dataset that MSK rely on exclusively, the power is only 0.46 (alpha of 0.05, two sided test) to reject null hypotheses when false.

MSK overlooked the important interaction between premarital cohabitation and the first year of marriage in their Appendix A because their tests were under-powered. Although the harmonized NSFG is a large dataset to begin with, the first year of marriage is a small subset of all marital exposure to the risk of marital dissolution in the data. MSK's choice to discard more than three quarters of the data left them with insufficient power for this specific first-year-of-marriage test.

Building on their erroneous null finding for the interaction between the first year of marriage and premarital cohabitation, MSK's Table 1, Model 2 included no interaction between marital duration and premarital cohabitation, so this model assumed (incorrectly) that premarital cohabitation had a similar effect on marital dissolution across different marital durations. MSK then used the predicted values from this incorrectly specified model to generate their Figure 1.

d) Reducing the dataset part 1: MSK's 10 year filter

MSK claim that because of recall bias, the NSFG data should be limited to couples whose first marriage occurred within 10 years of each survey. MSK cite Hayford and Morgan (2008) to justify their exclusion of couples married more than 10 years before each survey. Hayford and Morgan (2008) found that women in NSFG under-reported cohabitations that occurred further in the past, however Hayford and Morgan were studying *all* cohabitations, as opposed to only cohabitations with the future first husband, which are the cohabitations under study here. We imagine (and Hayford and Morgan 2008 p. 140-141 also suggest) that premarital cohabitation with the future first husband would be easier to recall years later than cohabitation with a different partner. Whereas the stigma against sex outside of marriage (especially for women and especially with multiple partners) might lead to underreporting of past cohabitations with partners other than the future husband, those type of nonmarital cohabitations are not examined here. Furthermore Hayford and Morgan noted that problems in the 2002 NSFG could explain some of their results. Additionally, potential bias in recalling premarital cohabitation from the past would only be problematic for the analysis here if the recall bias was conditional on whether the couple's later marriage had ended. Hayford and Morgan did not examine this last question.

There is a potential rationale for filtering out the older marriages to limit recall bias (as MSK do), but there is also a rationale for the reverse: not throwing valid data away. MSK's 10 year marriage filter eliminates 73% of the otherwise eligible couple-years from consideration; their age-at marriage restriction (including only marriages that occurred when the wife was age 15 to 35) reduces the sample size by another 0.5%, for a total 74% reduction in eligible couple-years.

If one is going to discard the majority of the valid data in any analysis, one should first explain exactly how much of the data is being discarded; this is not only academic standard practice but it is the official policy of this journal (National Council on Family Relations 2020). Second, one should offer supplementary analyses with the full dataset. MSK offer neither the rate of discarded data nor the analysis with a full sample; we provide both.

Filtering out the couples who have been married longer (as MSK do) enhances the Recent Cohort Fallacy because in the very early stages of marriages, premarital cohabitation reduces the risk of marital breakups. The NSFG is already age-limited at the time of survey to subjects younger than 45 years old; throwing away the data from subjects with longer marital duration exacerbates the age filter that is built in to the design of the NSFG.

Marriages can have trajectories that are many decades long. The average married woman with a husband in the U.S. has been in that marriage for more than 20 years (Rosenfeld 2014; Table 1). MSK's 10 year filter means that the average married couple in their analysis is only observed for 5 years after marriage. See Figure 1 for an illustration of why the first 5 years of marriage might not accurately represent the relationship between premarital cohabitation and marital dissolution over a longer marital duration. The short time window unnecessarily limits MSK's ability to find recent effects of premarital cohabitation on marital dissolution.

It is not uncommon for literature on marital dissolution using NSFG to filter the data in some way. Teachman (2002) used 5 waves of NSFG data and did not exclude married couples due to how long they were married before the surveys (as MSK do), but he did truncate marital duration at 10 years, and found that this truncation made no substantive difference. We estimate that Teachman's truncation of the data at 10 years of marital duration combined with his exclusion of couples married before 1950 eliminated approximately 19% of dissolution events and 27% of otherwise valid couple-years, a far cry from MSK's omission of 74% of valid couple-years.

e) Reducing the dataset part 2: dropping the 1988 NSFG wave

MSK's models do not include the 1988 wave, the first NSFG wave with data on premarital cohabitation, and they do not mention why. Dropping the 1988 wave increases MSK's reduction of couple years in the analysis to a 79% reduction. Dropping the 1988 wave has a substantial biasing effect on the key outcome: sharply increasing the apparent (but misleading) decline over time in the association between premarital cohabitation and marital dissolution that MSK report.

f) Using a stealth proxy for premarital cohabitation: children at the time of marriage

One reason couples marry is to have and to raise children. NSFG surveys were designed to study fertility. Datasets derived from NSFG, including the datasets used in RR 2019 and here, control for whether minor children are present in the home for every year that a couple is together. The presence of children is time-varying over the course of each marriage in real life, and should be time-varying in the data as well. Instead of allowing the presence of children to vary year to year, MSK code only whether the couple had children at the time of marriage or not. Fertility during marriage and children growing up and leaving the household are both ignored by MSK. MSK also code subject's education as time invariant, fixed at the time of marriage, which is also problematic (Martin 2006) but this choice matters less to the results so we do not dwell on it. We assume that age of BA achievement is 22, as age at educational milestones is not recorded in NSFG. MSK do not mention or explain their choice to treat children as time-invariant in their main body text, but it is a highly biasing choice because couples who cohabit before marriage are much more likely to enter marriage with children, and the tendency of the premarital cohabiters to enter marriage with children has been increasing over time.

Forty one percent of NSFG's premarital cohabiters were living with minor children in the household in the calendar year of their first marriage (whether the children were born before or immediately after the marriage), compared to 19% of non-cohabiters (and the difference was highly significant). Twenty five percent of the NSFG's premarital cohabiters had their first biological child born before the month of their first marriage, compared to 7% of the non-cohabiters, and again the difference was highly significant.

The difference in premarital fertility between cohabiters and non-cohabiters has been *widening* over time. It is well known that the percent of U.S. born children who are born outside of formal marriage has been rising steadily over time (Acs et al. 2013). For women who married in the 1980s in the NSFG data, 7% of the non-cohabiters had a biological child before first marriage, compared to 19% for the premarital cohabiters. For women who married in the 2010s, 8% of the non-cohabiters entered their first marriage having already given birth, compared to 34% of women who had cohabited with their husband before marriage.

In the footnotes below their tables, MSK write "education and fertility were measured at marriage to avoid confounding associations between covariates and divorce," but this explanation reflects a very basic misunderstanding on their part. Couple-specific fertility trajectories do not confound with any other variables. The creation of the couple-period event history dataset to hold the couple-specific time-varying children-in-the-household variable is a labor intensive enterprise (Mills 2011) which is important and relevant to the analysis, which is why we undertook the work and MSK should have done so as well.

In general, married couples with minor children in the home are moderately less likely to divorce (Brines and Joyner 1999 Table 2; see also RR 2019 Appendix Table 1) because married individuals with children have more to lose in a divorce. In MSK's models, children at marriage *raise* the odds ratio of marital dissolution by an odds ratio of between 1.47 and 1.57, a very substantial increase, because MSK's measure of children at the time of marriage is a proxy for premarital cohabitation. The increasing tendency of premarital cohabiters to enter their first marriage with a child, combined with MSK's finding of a positive association between having children at marriage and later marital dissolution masks what appears (erroneously) in MSK's models to be a decline over marriage cohorts in the association between premarital cohabitation and marital dissolution.

g) The Age, Period, and Cohort issue

MSK raise an issue about age, period, and cohort (APC) terms, or in this case marital duration, calendar year, and marriage cohort. Since marriage cohort+ marital duration=calendar year, the three terms cannot all be in the model in an untransformed state. The untransformed state is the key caveat. There is a small industry in the study of data analysis devoted to APC models where all three types of terms (under some transformations) are in the model. Leaving one of the three APC dimensions out of the models entirely is one of the worst options, according to the survey by Fosse and Winship (2019). We had calendar decade dummy variables in our marriage cohort models, which was perfectly appropriate. These calendar decade dummy variables are not collinear with yearly marriage cohort or with marital duration, they improved the goodness of fit, but had no substantive effect on the association of interest. Dropping the decade dummy variables and the survey wave dummy variables as we do in Appendix Tables 1 and 2 leads to the same substantive conclusions. MSK's complaint about APC is therefore a red herring.

h) The weights

RR 2019 used unweighted multivariable regressions, because the models included the predictors of the weights (Winship and Radbill 1994). MSK insist that weights should be applied in the models, which is fine with us. Because the different NSFG waves had different survey structures with respect to sampling strata and primary sampling units, there is no single way to implement the weights across survey waves. In Tables 1 and Appendix Table 1 below we use the weights as analytic weights, rescaled to 1 within each NSFG wave. In Appendix Table 2 we use the weights as probability weights and we take advantage of the complex survey structure by accounting for strata and primary sampling units in waves that have them. The sample clustering and strata do not affect the substantive outcome. As the weights and strata and primary sampling units do not make a substantive difference in the outcome of interest, the weights are another red herring.

i) Problems arising from MSK's failure to request a replication package from us

It is not always possible to accurately translate the English version of a model description from someone else's paper into an actual model, which is why replication packages promote scientific progress. Several statements in the MSK comment about our models and data were wrong in ways that would have been simple to correct or verify had they cared to do so. We are perplexed by MSK's claim that "Calendar year is a period indicator measuring the year marital dissolution was observed or year of interview if dissolution did not occur. This measure should be a time-varying indicator and not a fixed indicator." This statement of MSK is unfortunately inaccurate. Calendar year is a period indicator of course; how could it be otherwise? And how can the variable for calendar year be fixed in the dataset if the dataset contains a range of calendar years (as the NSFG does)? Calendar year in our models is not a function of dissolution or survey year. When MSK wrote "the year of year of dissolution in 1995 should not be used to predict whether a dissolution occurred in 1987," they were expressing a misunderstanding either about our models or about how event history analysis works. MSK wrote declaratively about our data structure, but what they wrote was both incorrect and illogical.

RESULTS:

[Table 1 here]

Table 1 traverses the model design choices suggested by MSK, as well as the model design choices they use but fail to explain or justify. The design choices that MSK failed to explain or justify are the more relevant ones for the outcome of the analysis. Model 1 is an exact replica of RR 2019 Table 2, Model 3. In scanning across the models, we highlight not only the P value of the key coefficient (premarital cohabitation × time), but also the BIC statistic for the premarital cohabitation × time interaction, since with large sample sizes in the NSFG, the typical 0.05 significance threshold is too easily satisfied (Raftery 1995). BIC values smaller than -10 would indicate significance (by the BIC standard) of the key coefficient.

The BIC tests here are not tests between models (which in Table 1 have different sampling frames), but rather are tests within each model description of the model as described compared to the same model without the key interaction of premarital cohabitation and time. The Likelihood Ratio Test of the model (compared to the constant only model) with and without the key interaction is the statistic ΔLRT =LRT(M w/o interaction)-LRT(M with interaction), on one degree of freedom. ΔLRT will be negative or zero, as the model with the interaction cannot fit worse than the model without the interaction. BIC= ΔLRT+ln(events), where events are the number of marital dissolutions in the sample, indicated on the table.

Applying the weights in model 2 yields a log odds coefficient of -0.0044, neither significant at the coefficient level or by the BIC. Model 3 drops all observations from marriages that took place more than a decade before each survey, and drops marriages when the bride's age at marriage was less than 15 or older than 35, dropping 74% of the couple-year observations. The key interaction remains insignificant in Model 3, and BIC criteria continues to reject the interaction between premarital cohabitation and time.

Model 4 drops the 1988 wave, which MSK did not explain or justify, and in Model 4 the key coefficient is significant (P=0.008), but the BIC test prefers the model without the interaction between premarital cohabitation and time (because BIC of 0.57 >-10). After dropping the 1988 wave (as MSK did), the number of couple years has been reduced to 46,331/216,455=21% of the original data (a 79% reduction), and the number of dissolution events has been reduced to 1,923/8,488=23%, a 77% reduction. MSK report 45,586 couple years in the footnote below their Appendix A, slightly lower than the 46,331 we find in our replication of their results.

Model 5 replaces the appropriate time-varying measure of children in the household with MSK's less defensible time-invariant measure of children at the time of marriage. Note how MSK's decision to use a non-time-varying measure of children in Model 5 elevates the statistical significance of the key interaction of premarital cohabitation and time. Model 5 follows MSK's recipe for the data, and here finally the key interaction is significant at the coefficient level and ambiguous by the BIC (BIC=-2.37) Unfortunately Model 5 relies on a very peculiar and biasing set of modeling choices the most relevant of which (dropping the 1988 wave, and treating children as time-invariant) were never justified by MSK.

Model 5 is based as closely as we could on MSK's Appendix A, which they described as replicating RR 2019 Table 2, Model 3. RR 2019's Table 2, Model 3 is reproduced exactly here in Model 1. It should be obvious that Model 5 (closely following MSK's Appendix A) is not a replication of Model 1. MSK applied an entirely different set of design choices in their Appendix A (our Model 5) including a sample size of couple years 21% as large, omission of the 1988 wave, and treating children as non-time-varying. Not only did MSK mis-attribute their model design choices to us, they failed to justify in their text their model design choices that were relevant to their results. MSK's replication of our results from RR 2019 is not a real replication.

We acknowledged in RR 2019 that under certain modeling choices, the key association (the supposed decline of the association of premarital cohabitation and marital dissolution over time), could appear to be significant. MSK's results simply reinforce that aspect of RR 2019's prior results. Our question in RR 2019 and now is simple: is the finding (of declining association between premarital cohabitation and marital dissolution) robust? Table 1 shows that the key association of interest is not robust. Our Appendix Table 1 repeats Table 1, but without the calendar decade dummy or the survey wave dummy variables. Appendix Table 2 repeats Table 1 but with Cox proportional hazard models instead of the discrete time models, and Appendix Table 2 also excludes the calendar year and NSFG wave dummy variables. Appendix Table 2 also accounts for the complex survey design in the NSFG waves that were sampled with clusters and strata. Both appendices yield the same substantive results as Table 1. If one relies on the BIC criteria, which we favor for its strictness, *none* of the models (7 models in Table 1 and 8 more in two appendices with BIC statistics) shows a significant change (BIC<-10) in the association of premarital cohabitation and marital dissolution over time. If one relies on significance of the coefficient at the 0.05 level, the key interaction can be significant but the results are not robust to reasonable changes in research design.

In Model 6 we reinstate the 1988 wave and we return to a calendar year time interaction for premarital cohabitation (eliminating the Recent Cohort Fallacy), and we reinstate a time-varying presence of children. We retain the weights and MSK's problematic 10 year filter, both of which model design choices were at least explained by MSK. Model 6 wipes away the core finding that MSK promote. Lest one imagine that the lack of statistical significance of the key interaction in Model 6 might be due to the reduced sample size, Model 7 removes the 10 year filter and the age at marriage filter imposed by MSK. Model 7 almost quadruples the sample size compared to Model 6, and increases the power of the model accordingly, and the key coefficient (the interaction between premarital cohabitation and calendar year) remains insignificant.

Ordinarily we would treat a coefficient such as the -0.0024 interaction between premarital cohabitation and year in Model 7 as zero, since the interaction's statistical significance falls below the P=0.05 threshold. Alternatively, we might care about the size of the interaction as well as its significance (or lack thereof). At the rate of change reported in Model 7 for the interaction between premarital cohabitation and calendar year (-0.002378 per year more precisely) it would take about 159 years for this change to erase the established interaction between premarital cohabitation and marital dissolution, which has a coefficient of 0.3787 (not shown) in Model 7. At that rate of change, with a baseline year of 1960, premarital cohabitation might cease to be associated with marital dissolution around the year 2119, far enough in the future for us to assume that the association between premarital cohabitation and higher rates of marital dissolution will be with us for the foreseeable future.

CONCLUSION:

We conclude here the same way we did in RR 2019: premarital cohabitation seems to have a consistent association with marital dissolution in the US across calendar time. Premarital cohabitation's association with marital dissolution is reasonably consistent across marriage cohorts in the US if care is taken to account for the Recent Cohort Fallacy. MSK's claim that premarital cohabitation no longer predicts marital dissolution in the U.S. is not a robust finding.

There are theoretical reasons to believe that the association between premarital cohabitation and marital dissolution *should have* declined over time. We explored some of those reasons in RR 2019. The data, however, tell a different story.

We believe in research transparency. Research transparency means sharing data and models where possible to promote scientific progress. MSK failed to share the replication package of their analysis that we requested, and they failed to request a replication package from us. Because MSK's purported reanalysis of our results was based on their speculation rather than on an actual encounter with our data and models, they make simple errors and erroneously attribute errors to us. By pretending to replicate RR 2019 but failing to actually do so, MSK have undermined research transparency.

Research transparency also means presenting results in a way that highlights rather than hides the contributors to the substantive results. MSK were not transparent about the design choices that drove their results. MSK did not mention how much of the data they had discarded, or how discarding three quarters of the valid data may have enabled their dismissal of the association between premarital cohabitation and marital stability in the first year of marriage. Their main text never mentioned dropping the 1988 wave or treating parenthood as time-invariant, yet these decisions were substantively consequential. The supplementary tables they sent to us were silent on these subjects as well. Replicability of social science research depends on research transparency.

The issues that MSK did raise included a variety of data analysis red herrings. They speculated unreasonably about aspects of our data structure that a simple email or a glance at our data would have cleared up. They suggested that the weights or the presence of decade dummy variables might explain the divergence in conclusions about the association between premarital cohabitation and marital dissolution. We show in our tables and appendices that the weights and the decade dummy variables do not substantively alter the results. It is unfortunately all too common in social science for published work to perseverate on technical details that do not affect the results, and to hide from the reader the more material issues on which the results actually depend.

MSK complain that we fail to follow traditions in the literature, citing themselves primarily as the sources of accepted traditions. We are pleased to challenge some of the traditions in this literature about the predictors of marital dissolution. We do not believe that there is sufficient blanket justification for discarding (without strong specific theoretical necessity, sensitivity analyses and robustness checks) large swathes of the valid NSFG data. Even before discarding any data, the NSFG is limited to younger respondents and therefore excludes couples who have been married the longest. Discarding the couples with longer marriages from the data, as MSK do, further biases the sample and lowers the power of statistical tests and can lead to crucial associations being overlooked or ignored.

We also seek to overturn the tradition of using time-invariant measures of children and education when the data are rich enough to allow for time-varying measures of these key predictors. We showed in this rejoinder that a time-invariant measure of children, such as used by MSK, is highly biasing on the association between cohabitation and later marital dissolution.

Social science exists in a world of constantly evolving multivariable statistical tools. Multivariable statistical models as presented in published papers are inherently lacking in transparency because the reader generally cannot know how the results would be different under different design choices. Simple bivariate associations and displays of raw data can be misleading just as multivariable models can, but at least bivariate associations and raw data are transparent in the sense that a reader can identify the shortcomings readily (Rosenfeld 2005). The raw data have a story to tell (Tufte 2001) that is too often ignored in social science because social scientists trust our own multivariable models too much.

We note in this context that MSK failed to show raw data or bivariate associations for the central questions in their comment. They relied entirely on multivariable models and predictions from models that were biased for reasons we explain above. Research transparency requires scholars not to overlook what the raw data and the bivariate associations have to say. We document in this brief reply and in RR 2019 that the association between premarital cohabitation and marital dissolution is positive and generally unchanged across calendar time and unchanged across marital cohorts (when appropriate controls are applied).

REFERENCES:

Acs, G., Braswell, K., Sorensen, E. & Turner, M. A. (2013) The Moynihan Report Revisited. The Urban Institute. http://www.urban.org/UploadedPDF/412839-The-Moynihan-Report-Revisited.pdf

Brines, J. & Joyner, K. (1999). The Ties that Bind: Principles of Cohesion in Cohabitation and Marriage. *American Sociological Review, 64*, 333-355. http://dx.doi.org/10.2307/2657490

Cherlin, A. J. (1992). *Marriage, Divorce, Remarriage,* Cambridge, Mass.: Harvard University Press.

Fosse, E. & Winship, C. (2019). Analyzing Age-Period-Cohort Data: A Review and Critique. *Annual Review of Sociology, 45*, 467-492. https://doi.org/10.1146/annurev-soc-073018-022616

Goldstein, J. R., Sobotka, T. & Jasilioniene, A. (2009). The End of 'Lowest Low' Fertility? *Population and Development Review, 35*, 663-699. http://dx.doi.org/10.1111/j.1728-4457.2009.00304.x

Hayford, S. R. & Morgan, S. P. (2008). The Quality of Retrospective Data on Cohabitation. *Demography, 45*, 129-141. http://dx.doi.org/10.1353/dem.2008.0005

Martin, S. P. (2006). Trends in Marital Dissolution by Women's Education in the United States. *Demographic Research, 15*, 537-560. http://dx.doi.org/10.4054/DemRes.2006.15.20

Mills, M. (2011). *Introducing Survival and Event History Analysis,* Los Angeles: Sage Publications.

National Council on Family Relations (2020) Journal of Marriage and Family (JMF) Style Guide. https://www.ncfr.org/jmf/submit-jmf/jmf-style-guide Accessed on October 7, 2020.

Raftery, A. E. (1995). Bayesian Model Selection in Social Research. *Sociological Methodology, 25*, 111-163. http://dx.doi.org/10.2307/271063

Rosenfeld, M. J. (2005). A Critique of Exchange Theory in Mate Selection. *American Journal of Sociology, 110*, 1284-1325. http://dx.doi.org/10.1086/428441

Rosenfeld, M. J. (2014). Couple Longevity in the Era of Same-Sex Marriage in the US. *Journal of Marriage and Family, 76*, 905-918. http://dx.doi.org/10.1111/jomf.12141

Rosenfeld, M. J. & Roesler, K. (2019). Cohabitation Experience and Cohabitation's Association with Marital Dissolution. *Journal of Marriage and Family, 81*, 42-58. https://doi.org/10.1111/jomf.12530

Rosenfeld, M. J. & Roesler, K. (2020a) Replication package for reply to MSK. https://web.stanford.edu/~mrosenfe/data_share/RR_2020_NSFG_replication_package_v2.zip

Rosenfeld, M. J. & Roesler, K. (2020b) Replication package for Rosenfeld and Roesler JMF 2019. https://web.stanford.edu/~mrosenfe/data_share/Rosenfeld_Roesler_2018_NSFG.zip

Teachman, J. (2002). Stability across Cohorts in Divorce Risk Factors. *Demography, 39*, 331-351. https://doi.org/10.1353/dem.2002.0019

Tufte, E. R. (2001). *The Visual Display of Quantitative Information,* Cheshire, Conn.: Graphics Press.

Winship, C. & Radbill, L. (1994). Sampling Weights and Regression Analysis. *Sociological Methods and Research, 23*, 230-257. https://doi.org/10.1177/0049124194023002004

Table 1: Different ways to measure the change over time in the interaction between premarital cohabitation and marital dissolution with data from NSFG, following MSK design choices in Models 2-5

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7
Description	Same as M3 in Table 2 of RR 2019	M1 plus weights	Married within 10 years, at age 15-35	Drop the 1988 wave	Make children and education time-invariant	add 1988 wave back in, and use calendar year, time-varying fertility	Remove the 10 year and age at marriage filters
Waves	1988-2013/15	1988-2013/15	1988-2013/15	1995-2013/15	1995-2013/15	1988-2013/15	1988-2013/15
Weighted?	No	Yes	Yes	Yes	Yes	Yes	Yes
Married within 10 years, and at age 15-35.	No	No	Yes	Yes	Yes	Yes	No
Premarital cohabitation interacted with calendar year or marriage year	marriage year	marriage year	marriage year	marriage year	marriage year	calendar year	calendar year
children and education time varying?	yes	yes	yes	yes	no	yes	yes
The key association: Log odds Interaction coefficient for premarital cohabitation × time (SE)	-0.0035 (0.0022)	-0.0044 (0.0023)	-0.0071 (0.0049)	-0.018** (0.0067)	-0.021*** (0.0067)	-0.0035 (0.0049)	-0.0024 (0.0024)
P value	0.115	0.054	0.15	0.008	1.6×10^{-3}	0.48	0.31
OR version of coefficient	0.997	0.996	0.993	0.982	0.979	0.997	0.998
BIC	6.56	5.35	5.68	0.57	-2.37	7.26	8.03
N of couple years	216,455	216,455	57,172	46,331	46,331	57,172	216,455
N of marital dissolution outcomes	8,488	8,488	2,372	1,923	1,923	2,372	8,488

Notes: Results are from discrete time event history models in logistic form, predicting marital dissolution. Data for female respondents from NSFG. OR version of the coefficient is the odds ratio version, which is the coefficient exponentiated. BIC is a test of the one degree of freedom interaction between premarital cohabitation and time. BIC<-10 is the cutoff for significance recommended by Raftery. BIC>-10 rejects the interaction between premarital cohabitation and time, and prefers the model with no such interaction. Additional controls in all the models: marital duration in years (1df); dummy variable for calendar year of marriage versus other years (1df); premarital cohabitation × calendar year of marriage dummy (1df) to isolate the different relationship of premarital cohabitation and marital dissolution in the first year of marriage, see Figure 1; subject's age at marriage categorical (3 df); presence of children time varying unless otherwise noted (1df); subject's education time varying unless otherwise noted (3df); Mother's education (3 df); subject's race (2 df); survey wave (4 or 5 df); calendar decade (3 or 4 df); stable family of origin (1 df).

* < P.05; ** P<0.01; *** P<0.001, two tailed tests.

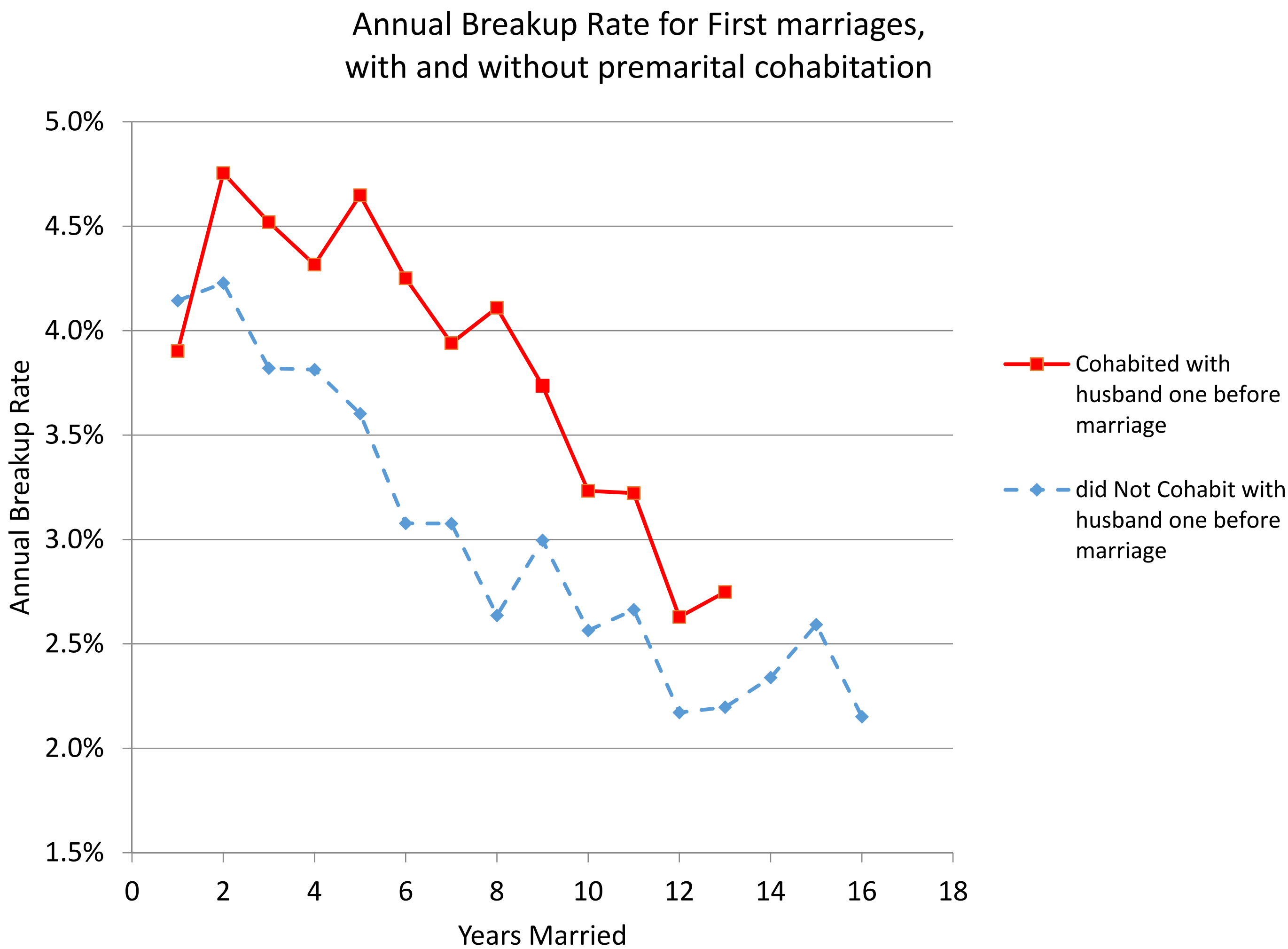

Figure 1. Weighted NSFG data on first marriages, female subjects age 15-44, waves 1988 and later. In this graph, the 12 months between marriage and the first anniversary is year 1, and the next year is year 2, and so on. Breakup rate in year X is conditional on the marriage surviving the previous X-1 years intact. Reprinted from RR 2019 Figure 3. In our Table 1 models we interact premarital cohabitation with a dummy variable for the calendar year of marriage. Given that the average couple marries in the middle of the calendar year, newlyweds are exposed to the risk of marital dissolution for an average of only 6 months in the calendar year of marriage. In those first 6 months, the marital stability advantage of the premarital cohabiters is greater than at 12 months of marital duration, and therefore greater than what is shown in Figure 1.

PREMARITAL COHABITATION AND MARITAL DISSOLUTION:

A REPLY TO MANNING, SMOCK AND KUPERBERG

SUPPLEMENTAL TABLES

Appendix Table 1: Different ways to measure the change over time in the interaction between premarital cohabitation and marital dissolution with data from NSFG. Same as Table 1 but all these models lack calendar decade dummy and survey wave dummy variables that are present in Table 1 models.

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7
Description		M1 plus weights	Married within 10 years, at age 15-35	Drop the 1988 wave	Make children and education time-invariant	add 1988 wave back in, and use calendar year, and time-varying fertility,	Remove the 10 year and age at marriage filters
Waves	1988-2013/15	1988-2013/15	1988-2013/15	1995-2013/15	1995-2013/15	1988-2013/15	1988-2013/15
Weighted?	No	Yes	Yes	Yes	Yes	Yes	Yes
Retrospection Filter	No	No	Yes	Yes	Yes	Yes	No
Premarital cohabitation interacted with calendar year or marriage year	marriage year	marriage year	marriage year	marriage year	marriage year	calendar year	calendar year
children and education time varying?	yes	yes	yes	yes	no	yes	yes
The key association: Log odds Interaction coefficient premarital cohab x linear time (SE)	-0.0028 (0.0021)	-0.0055* (0.0022)	-0.0039 (0.0049)	-0.013 (0.0069)	-0.017* (0.0069)	-0.00026 (0.0049)	-0.0040 (0.0022)
P value	0.18	0.012	0.44	0.057	0.016	0.96	0.068
OR version of coefficient	0.997	0.9945	0.996	0.987	0.984	0.9997	0.996
BIC	7.26	2.81	7.16	3.95	1.81	7.77	5.73
N of couple years	216,455	216,455	57,172	46,331	46,331	57,172	216,455
N of marital dissolution outcomes	8,488	8,488	2,372	1,923	1,923	2,372	8.488

Notes: Results are from discrete time event history models in logistic form, predicting marital dissolution. Data for female respondents from NSFG. BIC<-10 is the cutoff for significance recommended by Raftery. BIC>-10 rejects the interaction between premarital cohabitation and time, and prefers the model with no such interaction. Additional controls in all the models: marital duration in years (1df), dummy variable for calendar year of marriage versus other years (1df), premarital cohabitation × calendar year of marriage dummy (1df), subject's age at marriage categorical (3 df), presence of children time varying unless otherwise noted (1df), subject's education time varying unless otherwise noted (3df), Mother's education (3 df), subject's race (2 df), stable family of origin (1 df).

* < P.05; ** P<0.01; *** P<0.001, two tailed tests.

Appendix Table 2: Different ways to measure the change over time in the interaction between premarital cohabitation and marital dissolution with data from NSFG. Same as Table 1 except these are results from Cox proportional hazard models instead of the discrete time event history models in Table 1. These models account for complex survey design and lack decade and wave dummy variables.

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7
Description		M1 plus weights	Married within 10 years, at age 15-35	Drop the 1988 wave	Make children and education time-invariant	Add 1988 wave back in, and use calendar year, time-varying fertility	Remove the 10 year filter
Waves	1988-2013/15	1988-2013/15	1988-2013/15	1995-2013/15	1995-2013/15	1988-2013/15	1988-2013/15
Weighted?	No	Yes	Yes	Yes	Yes	Yes	Yes
Retrospection Filter	No	No	Yes	Yes	Yes	Yes	No
Premarital cohabitation interacted with calendar year or marriage year	marriage year	marriage year	marriage year	marriage year	marriage year	calendar year	calendar year
children and education time varying?	yes	yes	yes	yes	no	yes	yes
The key association: Interaction coefficient premarital cohab x linear time (SE)	-0.0027 (0.0021)	-0.0047 (0.0028)	-0.0038 (0.0061)	-0.012 (0.0082)	-0.016* (0.0080)	-0.00017 (0.0062)	-0.0035 (0.0029)
P value	0.19	0.089	0.54	0.14	0.05	0.98	0.23
Hazard Ratio version of coefficient	0.997	0.995	0.996	0.988	0.984	0.999	0.997
BIC	7.37	N/A	N/A	N/A	N/A	N/A	N/A
N of couple years	216,455	216,455	57,172	46,331	46,331	57,172	216,455
N of marital dissolution outcomes	8,488	8,488	2,372	1,923	1,923	2,372	8,488

Notes: Data for female respondents from NSFG. BIC<-10 is the cutoff for significance recommended by Raftery. BIC>-10 rejects the interaction between premarital cohabitation and time, and prefers the model with no such interaction. The models here do not include decade dummy variables or wave dummy variables; other controls the same as Table 1. Models 2 and later use weights and account for complex survey implementation, including survey strata starting with the 1995 wave and primary sampling units starting with the 2002 wave. Accounting for complex surveys means models are not maximized through likelihood maximization. As the BIC test is based on the likelihood ratio test, models 2 and later do not report BIC test statistics.

* < P.05; ** P<0.01; *** P<0.001, two tailed tests.