Comments on Kin Effects on Fertility

The authors develop a framework for comparing various accounts of why kin affect fertility. They test these possibilities by compiling and analyzing a database of kin effects on fertility from the literature. Focusing on how parents and in-laws might affect different fertility outcomes, the authors find that parents are more likely than in-laws to have anti-natal effects on a woman's total fertility, and age at first birth, but not on inter-birth intervals.

The authors divide the evolutionary proposals into mechanisms; ones whereby kin *are selected* to promote each others' fitness, and situations where kin's interests are less aligned. In the first group the authors include kin; 1) providing direct help, 2) buffering a woman from extrinsic risks of mortality, and 3) providing cultural information or transmitting information about reproductive norms (e.g., parents imparting pronatalist norms). In the second they include 4) kin as a cue to environmental context. The authors discuss these mechanisms and then consider their interaction with the measure of fertility being studied. The authors then consider and test the contexts that may moderate kin effects.

The general question of kin influences on fertility is of widespread general interest to the field of demography, and the authors objective to expand evolutionary theory makes an important potential theoretical contribution to the broader field. My comments come from the perspective of a family demographer who does not typically use or apply evolutionary theories to questions of kin influences on demographic behavior, so perhaps they will assist the authors to reach a wider audience of family demographers. Below are comments for the authors to consider as they revise their paper.

1. To start, the authors might consider defining some general terms at the outset that may not be family to evolutionary family demographers. The terms "fitness" and "inclusive fitness" for example appear to imply higher fertility, and if this is intended, then a clear definition of terms specific to evolutionary theory might be useful to reach a wider audience in demography.
2. In general, from a stylistic perspective, I found the text somewhat concentrated and dense. The authors may, again, reach a wider audience by defining basic terms from evolutionary theory at the outset (fitness, adaptive, maladaptive) and editing the text to remove jargon that may not be familiar to other family demographers.
3. The mechanisms by which kin influence fertility as described in the text do not line up explicitly with the mechanisms that are summarized in table 1 (e.g., two main mechanisms, with the first comprised of three parts and the second comprised of one part). For clarity, I recommend that the authors more closely match the summary of hypothesized mechanisms in the text with those in Table 1.
4. According to Table 1, almost one third of the outcomes considered in this analysis are not informed by the various evolutionary mechanisms that the authors discuss (indicated by "." In Table 1).
5. In general, I found the myriad of predictions, given the various mechanisms of kin influence and the multiple (four outcomes) a lot to sift through in a single paper. I would like to see

the authors (a) provide a rationale for the large number of outcomes for which they offer predictions, (b) hone the number of predictions by reducing the number of outcomes, or synthesizing the text in some way so that the major patterns of prediction are more apparent from reading the text.

6. In the authors discussion of contextual moderators of kin influences on fertility, the authors do not seem to have considered fully other potentially relevant contexts, for example the extent of a woman' exposure to non-family social network (e.g., in school, at work, in community organizations, etc) that may offer alternative or competing reproductive norms. Is women's participation in these other networks a relevant consideration?

7. Regarding the methodology for selection of articles for review and empirical testing, the authors claim to have done a systematic review of articles up to 2008 and then a different kind of review for articles after 2008 or 2012 (if I understand correctly). First, the details of the systematic review are not sufficient for the reviewer to assess whether the review was, in fact systematic. For example, the search terms and data bases that were used to identify articles are not clearly spelled out, nor are the inclusion / exclusion criteria for the selection of article. The Second, the rationale for the different in search methodology before and after 2008 is unclear, and the inclusion of articles after 2008 from a non-systematic review would to me suggest that the "systematic review" before 2008 is no longer relevant, as the authors analyze the full body of articles collectively.

8. The coding of 140 articles for 500 kin effects was an impressive undertaking; however, more detail is needed on what actually was coded, whether inter-rater reliability of coding was attempted and, if so, what level of reliability in coding was achieved. Would it be useful to provide an appendix of the coding schema? How well does the coding of articles align with the various hypotheses laid out in Table 1?

9. The fertility measure or outcomes in Figure 1 do not match the fertility outcomes that are presented in Table 1 (the summary of theoretical mechanisms and hypotheses). Why is this the case, and is this another rationale for honing the predictions to be more in line with the data at hand?

10. Likewise, the kin categories that are summarized in Figure 1 do not align totally with the theoretical kin influences, as specified in Table 1 (which speak more to quality of kin ties than structure of kin ties).

11. If the authors conducted a systematic review of the literature, was it feasible to conduct a meta-analysis of findings for at least a subset of outcomes rather than the "vote tallying method" that is described? Also, what are the actual implications of the way in which significant and non-significant results are coded for the interpretation of the findings?

12. Table 2 is difficult for this reviewer to interpret, and its alignment with the hypotheses generated from the outset is unclear.

13. It is not clear to this reviewer why the results are presented by discipline in Table, if a similar research question and similar kin and fertility variables are used across research articles. Is this table necessary? If yes, perhaps a stronger justification is needed.

14. Figure 3 seems to be the heart of the purpose of the review and empirical analysis. It would seem worthwhile, if the empirical section is retained, to focus most of the attention in the results section to this set of results and to assess the correspondence of the findings with the predictions in Table 1. Also, if the authors wish to retain the table on results by

discipline, are there systematic differences in the results presented in Table three by discipline. If so, then understanding why might be relevant, if not then it remains unclear to this reviewer why results are presented by discipline.

15. In general, the authors seem to have taken on two very large tasks of developing an expansive new theory (not all of which generates directional hypotheses) and then attempting a vast empirical review to test only a portion of the mechanisms described at the outset. The authors might consider couple of strategies for revision going forward (a) focusing only on refining the theory framework, as described in the first half of the paper, (b) narrowing the theoretically mechanisms that are discussed to those that are testable with the systematic review, as a way to more explicitly link the theoretical and empirical sections of the paper.

Thank you for the opportunity to read this ambitious work. I hope the thoughts above are helpful going forward.