

Marriage, Choice, and Couplehood in the Age of the Internet

© Michael J. Rosenfeld*, 2017

Draft Paper

Forthcoming in *Sociological Science*

July 25 2017

Do Not Cite Without Permission

* Michael J. Rosenfeld, Department of Sociology, Stanford University, 450 Serra Mall, Stanford, CA 94305. Email: mrosenfe@stanford.edu. Web: www.stanford.edu/~mrosenfe. This project was generously supported by the National Science Foundation, grants SES-0751977 and SES-1153867, M. Rosenfeld P.I., with additional funding from Stanford's Institute for Research in the Social Sciences and Stanford's UPS endowment. Thanks to Reuben J. Thomas, Amanda Mireles, Kate Weisshaar, Jasmine Hill, Ariela Schachter, Taylor Orth, and Stanford's Graduate Family Workshop for comments on earlier drafts.

Abstract:

How do the Internet and social media technology affect our romantic lives? Critics of the Internet's effect on social life identify the overabundance of choice of potential partners online as a likely source of relationship instability. This study examines longitudinal data showing that meeting online does not predict couple breakup. Meeting online (and particularly meeting through online dating websites) predicts faster transitions to marriage for heterosexual couples. I do not claim to measure any causal effect of Internet technology on relationship longevity or marriage formation. Rather, I suggest that the data are more consistent with a positive or neutral association between Internet technology and relationships, than with a negative association between the Internet and romantic relationships.

Marriage, Choice, and Couplehood in the Age of the Internet

Introduction:

Internet-based social media and cellular phones play an increasing role in our personal and social lives. Scholars disagree about whether the new technologies improve or degrade the commitment and longevity of our primary relationships. Internet skeptics argue that the suddenly wider choice set of potential partners available online could undermine marriages and other primary relationships (Slater 2013; Turkle 2015). Turkle (2011; 2015), one of the most prominent Internet skeptics, has argued that the new technologies have robbed us of the skills to be effective listeners in face-to-face interactions. If the Internet undermines our relationships, then the social effects of the Internet are to be feared.

Other scholars view the Internet as having a more positive role to play in personal and romantic relationships. McKenna, Green and Gleason (2002) argue that the asynchronicity of computer mediated communication yields potential relationship advantages. Wellman (2001) and McKenna and Bargh (2000) and Glassner (2010) all argue that fears of the Internet's supposed negative effects on relationships are unfounded. The debates over the Internet's effect on primary relationships have been hampered by a lack of nationally representative data.

In this paper I examine data from a nationally representative longitudinal study of American couples followed for 6 years, from 2009 to 2015. I test whether meeting online is associated with higher or lower rates of breakup, and higher or lower rates of transition to marriage. I show that couples who met online and offline have similar rates of breakup, consistent with prior literature (Rosenfeld and Thomas 2012; Cacioppo et al. 2013 found couples who met online had a slightly lower chance of breakup). I demonstrate in this paper that heterosexual couples who met through online dating transition to marriage more quickly than other heterosexual couples, which is a new finding.

Theoretical Background: Choice Overload versus the Advantage of Choice

Choice Overload theory is predicated on what psychologist Barry Schwartz (2005) refers to as the "paradox of choice." According to Choice Overload theory, the larger the choice set available to people in advanced economies, the lower quality decisions individuals will make,

and the worse they will feel. For individuals who are maximizers (i.e. people who always try to find the *best* solution to every problem), using Herbert Simon's (1990) division of people into maximizers and satisficers, the problem of choice overload is that the surfeit of choices frustrates maximizers' desires to find the best solution. How can one know if they have found the best choice, when there are so many possible choices? Without the possibility of knowing they have found the optimal solution, maximizers become unhappy, and (out of frustration) make hasty choices that make them even more unhappy.

Choice overload is one of the key theoretical ideas invoked to explain why Internet dating might yield lower quality partnership matches (Yang and Chiou 2010), and also why Internet dating would undermine existing romantic relationships. In an example of Choice Overload theory applied to Internet Dating, Sherry Turkle (2015 p.183) described a man named Danny who broke up with his girlfriend because "technology made it harder to commit." Danny suspected (because of the large choice set of potential partners available online) that he might be able to go online and find a woman more attractive than his girlfriend, so he broke up with her. After Danny broke up with his girlfriend, he went on a few dates with people he met online, but none of them turned out to be better or more attractive than his ex-girlfriend. Danny started to doubt his own judgement. He spent more time online, and was less satisfied. Danny eventually described the infinite-seeming choice of online dating sites as "toxic." Turkle argues that the abundance of potential new romantic partners available online has weakened individuals' commitments to their existing partners and has perhaps made Americans lonelier. Slater (2013) tells similar stories (consistent with Choice Overload theory) of people whose online dating experience makes them view all romantic relationships as less permanent, and more disposable.

One of the canonical Choice Overload experiments in the field of consumer behavior is Iyengar and Lepper's (2000) supermarket jam experiment. Iyengar and Lepper showed that supermarket customers presented with samples of 6 jams tried an average of 1.4 of the samples, and 30% used an offered coupon to buy jam, whereas customers presented with a larger choice of 24 jams also tried an average of 1.5 samples, but only 3% of the customers who saw the larger selection used the coupon. Iyengar and Lepper concluded that the larger choice set was demotivating to consumers. The reason that larger choice sets are demotivating, according to Choice Overload theory, is that too much choice makes it more difficult to determine which is the *best* choice (Schwartz 2005); furthermore the consideration of the multitude of choices might

make individuals feel less sure of their options, and therefore less inclined to make any choice (Schwartz 2005; Simon 1990).

Although Choice Overload theory (a theory initially based on experiments about decisions over non-essential consumer goods, such as the Iyengar and Lepper jam experiment) has been extended by several scholars (Ansari and Klinenberg 2015; Slater 2013; Turkle 2015; Yang and Chiou 2010) to Choice Overload critiques of Internet dating, the extension of Choice Overload to Internet dating is potentially problematic, for three reasons. First, even if Choice Overload theory is applicable to lower order needs such as jam, the theory might not be applicable to higher order needs such as romance, sex, and companionship, as the higher order needs are more difficult to deter. Second, in order to make the jam experiment work, Iyengar and Lepper had to exclude the flavors of jam that people prefer most (strawberry and raspberry). One of the great advantages of choice is that if a customer has a particular preference, they are more likely to satisfy their preference with a large choice set than a small choice set. A dating website without attractive candidates (i.e, without the analogs of strawberry and raspberry jam) would not be successful for long. A third critique of Choice Overload theory questions whether Choice Overload theory accurately explains consumer behavior. Scheibehenne et al's (2010) meta-analysis found no evidence of Choice Overload across the entire range of published studies that have tested Choice Overload in consumer retail situations.

Whereas Choice Overload theory would predict negative impacts of broad choice sets, other literature identifies relationship advantages of larger choice sets. Rosenfeld and Thomas (2012) showed that meeting online was especially common among gays, lesbians, and middle aged heterosexuals, groups that have difficulty identifying potential partners in the offline world. The technology to search across large choice sets to find the particular kind of partner each subject is looking for (whether by demographics or by acquired characteristics) is what makes the potential large choice set of Internet dating theoretically efficient. Whereas Choice Overload theory predicts that larger choice sets would result in weaker romantic matches and more relationship instability, the Advantages of Choice theory argues that larger choice sets should lead to better matches and therefore to more stable romantic unions.

Along with using the technology of search to take advantage of larger choice sets, Internet dating has the potential to increase the efficiency of data gathering about the smaller set of potential partners chosen for first dates. Among the Internet dating sites that cater to people

looking for relationships rather than for hookups exclusively, long detailed questionnaires are the norm (Finkel et al. 2012). Many kinds of personal attributes of potential partners that might have taken weeks or months to discern in the course of a relationship are, via Internet dating profiles, discernable before the first date. To the extent that mate selection is an information gathering process (Oppenheimer 1988), Internet dating with its search technology and its presumably rich and extensive database to search from, could theoretically speed up the transition from dating to commitment to marriage.¹

Prior Empirical Findings about the Internet and Social Life;

The Community Lost Paradigm

Kraut et al (1998) describe a classic experiment which began in 1995 and 1996 with 93 families in Pittsburgh, Pennsylvania who had Internet access in their homes for the first time. The Pittsburgh families were given computers that recorded the amount of time spent online. In the subsequent 1-2 years, greater Internet use in the Pittsburgh families was associated with modest but statistically significant negative changes: less family communication, more loneliness, and more depression. The Kraut et al initial Pittsburgh experiment has been widely cited, and echoed other studies from the early days of the graphical Internet, which suggested that time spent online was negatively correlated with time spent in face-to-face social interaction (Nie and Hillygus 2002). When Kraut and his colleagues followed the same Pittsburgh cohort for an additional year, they found that most of the negative outcomes associated with Internet use they had previously reported had disappeared (Kraut et al. 2002).

One interesting aspect of the two Kraut et al studies (1998; 2002) is that the initial study which reported negative impacts of the Internet has been cited more (more than 4,600 times) than the updated and revised second study (cited more than 1,900 times) which reported neutral social impacts of the Internet (Google Scholar 2017). Even in the most recent completed years, 2015 and 2016, Kraut et al's initial negative findings received almost twice as many citations as Kraut et al's updated neutral findings about the social impacts of the Internet. The greater popularity (in citation counts) of Kraut et al's initial negative findings as compared to their later neutral findings for social impacts of the Internet is consistent with McKenna and Bargh's (2000) and

¹ Of course, the theoretical advantage of information gathering from Internet profiles depends on the veracity of the information within those profiles. Research by Hancock et al (2007) suggests that online dating profiles are reasonably accurate along the dimensions that can be readily measured (age, height, and weight).

Wang and Wellman's (2010) views that a negative bias is present in popular and scholarly writing about the Internet's social impact.

Critiques of the Internet's supposedly negative effects on social life (Kraut et al. 1998; Nie and Hillygus 2002; Sales 2015a; Turkle 2015) contribute to a long tradition of arguments that modernity (in its various manifestations) undermines family and communal social bonds; Wellman (1979) refers in a different context to critiques of modernity as "community lost." In Wellman's (1979) analysis, urbanization and industrialization were the technological changes whose impact on community was in question. Wellman (1979) found that urban dwellers were not as isolated as the "community lost" paradigm led scholars to expect. Wang and Wellman (2010) argued that contemporary critiques of Internet technology's supposedly isolating effect on individuals echo a prior generation's critiques of urbanization. Parigi and Henson (2014) explicitly made the connection between criticism of technology's social impact and the "community lost" paradigm. I therefore use the term "community lost," to describe a broad set of critiques of Internet technology's supposed isolating effect on individuals.

The "community lost" paradigm and the Choice Overload hypothesis both predict that Internet technologies would undermine the stability of primary romantic relationships. The Choice Overload hypothesis is specific to one pathway (exposure to an over-abundance of potential partners) through which the Internet would undermine relationships.

Prior Empirical Findings about the Internet and the Stability of Romantic Relationships

Manning (2006 p.141) reported on an informal 2002 survey of divorce lawyers, finding that "68% of the divorce cases involved one party meeting a new love interest over the Internet." The potential ability of the Internet (which expanded the supply of new people any individual can meet) to undermine existing relationships is consistent with Choice Overload theory, but note that divorce lawyers see a population of couples that is selected on the dependent variable (divorce).

Young's (1998) study of "Internet addiction" was similarly based on a non-representative sample selected on the dependent variable, in Young's case selection on people who self-reported that Internet use was interfering with other aspects of their lives. Young (1998 p.134) wrote:

“I heard of many cases of seemingly perfect terminal love among both married and single cyberlovers that instantly failed as real-life relationships. Stripped of their fantasy masks, cyberlovers seldom embrace the other person when they discover how he or she really looks, acts, feels and talks, and they catch on to the bigger lies easily concealed from the safety of the computer”

Neither Manning’s (2006) study of divorce lawyers nor Young’s study of “Internet addicts” were nationally representative. Non-representative data (especially data selected on outcomes) can yield biased estimates of population behavior.

In contrast to the negative findings of the Internet’s impact based on non-representative data, recent studies based on nationally representative data show more benign or even positive associations between the Internet and relationship quality and stability. Cacioppo et al (2013) used a nationally representative² retrospective survey of 19,000 subjects and found that subjects who had met their spouses online were slightly less likely to report marital breakup than respondents who had met their spouse offline (controlling for year of marriage, age, ethnicity and other factors). Rosenfeld and Thomas (2012) used nationally representative longitudinal data from the How Couples Meet and Stay Together (HCMST) surveys and found no differences in breakup rate or relationship satisfaction by whether the couple had met online or offline.

Paul (2014) used HCMST data and found that couples who met online were more likely to break up, but Paul’s results were not replicable (see Appendix 1). I provide a fresh event history analysis of HCMST data, and I show below that the HCMST data yield results that differ substantively from Paul’s results.

In this paper, I expand on Rosenfeld and Thomas (2012) by using five years of follow-up HCMST data, along with retrospective HCMST data, rather than the one year of follow-up data used by Rosenfeld and Thomas. I add transitions to marriage along with breakup as relationship outcomes, and I use meeting online as the primary measure of the influence of the Internet. Cacioppo et al (2013) found that a higher percentage of married couples in the US had met

² Cacioppo et al used data from Harris Interactive, an opt-in Internet survey sample. Whereas the Knowledge Networks/ GfK data I use below includes US adults in the sample universe regardless of whether they had Internet access or not, the Harris Interactive sample had only Internet users in their sample universe. Whether opt-in sampling such as Harris Interactive yields nationally representative data is subject to a lively debate among scholars of survey methods (Baker et al. 2010).

online than Rosenfeld and Thomas reported for all couples meeting online.³ One way to reconcile the Cacioppo et al with Rosenfeld and Thomas would be to hypothesize that couples who met online progressed to marriage more quickly. I show below that couples who met online did transition to marriage more quickly.

The association between meeting online and faster transitions to marriage (as I document in the HCMST data below) does not necessarily imply that meeting online *caused* couples to marry more quickly. An alternative explanation is that marriage-ready individuals selectively chose the Internet as a venue to meet partners. Selection bias is a plausible alternative hypothesis to most hypotheses about romantic relationships that are based on observational data, where neither the relationship outcome (e.g. marriage) nor the inputs (e.g. how couples met) can be experimentally manipulated by researchers.

Hypotheses:

Hypothesis 1: Couples who meet online will be more likely to break up.

Consistent with the Choice Overload hypothesis, Hypothesis 1 implies that couples who met through the large choice set of potential partners available online, will have higher breakup rates and lower rates of transition to marriage, controlling for other relevant factors.

Hypothesis 2: Meeting online is associated with greater relationship stability.

Consistent with the Advantages of Choice, Hypothesis 2 implies that Internet meeting will be associated with lower breakup rates, and higher rates of transition to marriage, controlling for other relevant factors. Hypothesis 2 is consistent with some recent empirical research on the Internet's effect on romantic relationships (Cacioppo et al. 2013).

Hypothesis 3: Of all ways of meeting online, meeting through an Internet dating website will be especially associated with transitions to marriage.

Hypothesis 3 is an extension of Hypothesis 2. If the Internet provides advantages due to partner choice (following Hypothesis 2), then Internet dating websites should provide the

³ Cacioppo et al (2013) reported that 35% of couples married in the US in the 2005-2012 period had met online. Rosenfeld and Thomas (2012) found that the percentage of heterosexual couples who met online rose from about 8% in 2000 to about 22% in 2009.

greatest relationship advantage, because Internet dating websites are dedicated to efficient search for particular partner criteria (Finkel et al. 2012) in a way that provides, in theory, better matches and more rapid information gathering on matched partners. Many other ways of meeting online, such as through gaming, or through chat are serendipitous and would lack the maximized choice set or potential informational advantage of online dating. Additionally, individuals who are more interested in finding a partner for a committed relationship might self-select into the Internet dating market.

Data and Methods:

I use the How Couples Meet and Stay Together surveys (HCMST; Rosenfeld, Thomas and Falcon 2015) which started with a nationally representative survey of 3,009 adults who had romantic partners in 2009, and included longitudinal follow-up with the same individuals in 2010, 2011, 2013, and 2015. HCMST surveys were implemented by survey company Knowledge Networks/GfK (hereafter KN/GfK). KN/GfK panel participants were initially recruited into the panel through a nationally representative random digit dialing (RDD) telephone survey. Subjects who did not have Internet access at home were given Internet access. The HCMST wave 1 survey was an Internet survey, and waves 2-5 were Internet and phone surveys.⁴

Among subjects eligible for follow-up in HCMST, the response rate was 85% at wave 2, 73% at wave 3, 60% at wave 4, and 46% at wave 5. In waves 2-5 combined, the HCMST response rate was 93% for subjects who remained in the KN/GfK panel, compared to a 29% response rate among subjects who had retired or withdrawn from the panel. The key determinant of response to the HCMST follow-up surveys was whether the respondent was still in the KN/GfK panel at the time of the follow-up survey, rather than any factor that predicts couple

⁴ Seventy one percent of KN/GfK panelists contacted for the wave 1 HCMST survey consented to participate. If one includes the initial RDD phone contact to join the KN panel (participation rate 32.6%) which took place months or years before HCMST wave 1, and each subject's completion of the KN background survey (56.8% completion rate), and multiplies those rates together to derive a composite response rate (a composite rate which considers individuals asked to join the KN/GfK panel as having been eligible to respond later to HCMST wave 1), the composite response rate for the wave 1 HCMST survey is $.71 * .326 * .568 = 13\%$ (Callegaro and DiSogra 2008). Despite the low composite response rate of KN/GfK surveys compared to single-stage RDD surveys, the quality of data derived from the KN/GfK panel has been shown to equal or exceed the quality of data derived from industry standard RDD surveys (Fricker et al. 2005; Chang and Krosnick 2009), in part because KN/GfK gathers information from subjects at each survey stage.

longevity (such as relationship duration or marriage), which is why loss-to-follow-up does not bias estimates of relationship transitions in HCMST. In Appendix 2, I compare a key model from Table 2, weighted by the ordinary sampling weights, and then re-weighted with attrition adjusted weights (McGuigan et al. 1997). The results of the models with and without attrition weights are substantively the same, suggesting that the bias from attrition is negligible.

Methodologically, I rely on discrete time event history logistic regression (Yamaguchi 1991). My logistic regressions are weighted using the weight variable “weight2,” with robust standard errors (White 1980) and clustering to account for the non-independence of repeated observations of the same couple over time (Rogers 1993). Regressions without weights, including the HCMST variable “recsource” that identifies the oversampled groups and therefore predicts the weights (Winship and Radbill 1994) yield similar substantive results (see Table 3 below).

[Table 1 here]

Using the HCMST data, I create and analyze two separate event history datasets described in Table 1. The first dataset is a prospective dataset, starting with HCMST wave 1 in early 2009, following the 2,669 subjects who responded to at least one follow-up survey. Of these 2,669 partnered subjects, 1,341 met their partners in the Internet era or after, which I operationalize as 1995 or later because the graphical web browsers were first introduced in 1994 and 1995, and couples first started to meet online around 1995 (Rosenfeld and Thomas 2012). I analyze transitions to marriage only for heterosexual (i.e. different-sex) couples, because same-sex couples did not have access to legal marriage in most of the states and periods under study, and because even informal marriage among same-sex couples was historically constrained by the lack of legal marriage as an option (Rosenfeld 2014). The prospective dataset is a couple-month dataset because survey dates are specific to month. For breakups reported in the 12 months between HCMST waves 1 and 2, month of breakup was not asked and was therefore randomly allocated.⁵ For marriages reported in waves 2 and 3, the year of marriage was known and the

⁵ For married couples who broke up between wave 1 and wave 2, whose rate of breakup was less than 2% per year, breakups were randomly distributed to months between wave 1 and wave 2. For unmarried couples, breakup rate is much higher in the early stages of the relationship; the rate of breakup was more than 60% for unmarried couples who had been together for less than a year (Rosenfeld 2014), meaning the breakups would have been distributed

month of marriage was randomly allocated. For marriages reported in waves 4 and 5, and for breakups reported in waves 3, 4, and 5, both the month and year of transition were reported.

The second event history dataset is a retrospective dataset based on the history of relationships from wave 1 of HCMST. The time unit for the retrospective dataset is years. The retrospective dataset lacks a measure of subject education. The retrospective data include more transitions to marriage than the prospective dataset (814 compared to 109), but the retrospective data includes no breakups, because all HCMST couples were intact in 2009. The retrospective dataset has the advantage of covering events as early as 1998, when a modest number of HCMST couples who met online were first exposed to the risk of marriage. Together, the prospective and retrospective event history HCMST datasets provide for a more robust analysis than either dataset alone.

The prospective and retrospective datasets both rely on a measure of whether the survey subject and their partner met online, derived primarily from the open-ended wave 1 question, q24, “Please write the story of how you and [Partner_Name] first met and got to know one another and be sure to describe ‘how’ and ‘where’ you first met.” These open-ended responses were coded by the study investigators (Rosenfeld and Thomas 2012). In addition to the 270 subjects who were identified from the open-ended question as having met their partner online, an additional 19 subjects were identified from closed-ended question q32 from HCMST wave 1, “Did you use an Internet service to meet [partner_name]?” Of the 289 subjects who met their partner online, a subset of 134 met their partners through internet dating, identified either through their answers to the open-ended q24, or else by selecting the reply “Yes, an internet dating or matchmaking site (like eHarmony or match.com)” to q32.

Additional controls that are explored in the models or in the appendices include subject’s household income (time varying), subject’s race, whether subject described themselves as “born again or evangelical Christian” in 2009, whether subject and partner were a same-sex couple, whether subject lived in a metropolitan area (time varying), subject’s region of the U.S. (time

more in the beginning of the year than in the end of the year between wave 1 and wave 2. To accommodate the front-loading of breakups of nonmarital unions in the period between wave 1 and wave 2, I used the following

function: $M_b = (M_e) r^{\frac{2+rd}{1+rd}}$ Where M_b is the allocated month of breakup after wave 1, M_e is the number of months elapsed between wave 1 and wave 2, r is a random uniform number between zero and 1, and rd is relationship duration in years. For short relationship duration, the random factor is nearly squared, reducing the allocated months before breakup.

varying), number of minor children in the subject’s household (time varying), and for predicting breakups, whether subject and partner were coresident (time varying) and whether subject and partner were married or had a domestic partnership (time varying).

[Table 2 here]

Results: The Internet’s influence on couples analyzed with longitudinal data

Table 2 shows the influence of meeting online on relationship outcomes using the prospective event history HCMST dataset. In column 1, the couples who met online had odds of breakup 2.46 times higher than the couples who met offline. In the weighted HCMST prospective dataset, including married couples and unmarried couples, the breakup rate was 3.9% per year, compared to a breakup rate of 9.5% per year for couples who met online. The ratio of the odds of breakup rates at the monthly level yield the highly significant odds ratio of 2.46, but this initial comparison is misleading.⁶ The raw odds ratio for breakup without controls is highly significant in column 1 because the couples who met offline had been together much longer than the couples who met online, and were (as a consequence of longer relationship duration up to that point) less likely to break up. At wave 1, the average relationship duration for couples who met online was 4.4 years, and the average relationship duration for couples who met offline was 19.1 years. Most married couples in HCMST were married before 1995, before the graphical Internet. Controlling for relationship duration, as the adjusted odds ratio does, reduces the odds ratio of breakup (for couples who met online compared to couples who met offline) to an insignificant 1.10.

The full set of adjusted odds ratio coefficients for Columns 1 and 2, are provided in Appendices 2 and 3. Appendix 3 contains the same models with an expanded set of controls, including controls that predict Internet use and Internet dating, such as household income, urbanicity, and subject race (Sautter, Tippett and Morgan 2010) along with couple coresidence,

⁶ The prospective HCMST dataset used in Table 2 is a couple-month dataset, so odds ratios are calculated with monthly rates. If AP_1 and AP_2 are annual breakup rates, and monthly breakup rate $MP = AP / 12$, then

$$OR = \frac{\left(\frac{MP_2}{1 - MP_2} \right)}{\left(\frac{MP_1}{1 - MP_1} \right)}$$

the number of children in the household, and same-sex couple status. Appendix 3 shows that the inclusion of additional predictors of breakup does not alter the (null) effect of meeting online on couple breakup.

In column 2, the raw odds ratio of breakup (for couples who met online compared to couples who met offline) is a non-significant 1.31, because column 2 includes only couples who met during the Internet era, thus excluding couples who met before 1995 (nearly all of whom met offline) and who as a result of their long relationship durations had the lowest rate of breakup. Consistent with Rosenfeld and Thomas (2012), column two shows that meeting online had no significant effect on couple longevity, either with or without controls. The adjusted odds ratio of column 2 shows that controlling for other factors that predict breakup for couples who met during the Internet era, the odds ratio for breakup was not significantly different (odds ratio of 0.96) for couples who met online compared to couples who met offline. The difference between the raw odds ratio of 1.31 in column 2, and the adjusted odds ratio of 0.96 in column 2 is explained by controls for relationship duration, because even among the couples who met on or after 1995, the couples who met online met more recently (as meeting online became more popular over time). Across the 6 years of the HCMST prospective dataset, the couples (including married couples, cohabiting couples, and informal non-cohabiting couples) who met after 1995 had a breakup rate of 7.3% per year between 2009 and 2015. According to column 2, if the couples who met online after 1995 had the same demographic profile as the couples who met offline after 1995, their breakup rate would have been 7.0% per year (not statistically distinguishable from 7.3% per year).

Given the null findings of the effect of meeting online on breakup, I report the power of the data to reject null hypotheses. If the odds of breakup were truly 1.5 times higher for couples who met online than for couples who met offline, the HCMST data would have a 0.802 power to reject the null hypothesis of no difference, assuming a two-tailed alpha of 0.05. And if the real world odds of breakup were 2.0 times higher for couples who met online compared to couples who met offline, the power to reject null hypotheses of no difference with HCMST data would be 0.999, near certainty, with a standard two-tailed alpha of 0.05. HCMST data have sufficient power to reject null hypotheses for differences in the breakup rate between couples who met online and couples who met offline, if the differences between the true breakup rates corresponded to an odds ratio of 1.5 or higher.

Column 3 of Table 2 shows that heterosexual couples who met online had odds of transition to marriage almost twice as high (odds ratio 1.98) compared to heterosexual couples who met offline. Most transitions to marriage occur in the first 10 years of the relationship. The significantly faster transition to marriage among couples who met online was not quite robust to other predictors of transitions to marriage (odds ratio 1.86 after controls are applied, P value falling just above 0.05 at 0.057), providing some support for Hypothesis 2.⁷

[Figure 1 here]

Figure 1 shows the Kaplan-Meier (1958) cumulative survival (as intact couples) function for couples who met online compared to couples who met offline, using the prospective HCMST event history data. Consistent with Table 2, Figure 1 shows no significant difference in couple survival rates for couples who met online compared to couples who met offline (at the same relationship duration).

[Table 3 here]

Table 3 revisits the question of whether meeting online was associated with faster transitions to marriage. Model 1 of Table 3 replicates Column 3 of Table 2, showing that heterosexual couples who met online had odds of transition to marriage 1.98 times as high as the odds of transition to marriage for heterosexual couples who met offline. Models 2-5 of Table 3 break meeting online into two subcategories: those who met through internet dating or matchmaking websites, and those who met online in other ways (generally less structured and with smaller choice sets).

In Model 2, using the weighted version of the prospective data, heterosexual couples who met through online dating had an odds ratio (equivalent to a hazard ratio) of transitions to marriage of 3.19 times greater than couples who met offline, and the ratio was highly significant.

⁷ Not shown in Table 2 are associations between meeting online and transitions to cohabitation, which were not statistically significant. Also not shown in Table 2 are associations between having Internet access at home and various relationship transitions, which were not statistically significant, with one exception: having Internet access at home was associated with faster transition to marriage, until respondent's age was controlled for. Once respondent's age was controlled for, the association between having Internet access at home and transitions to marriage was no longer significant.

In Model 2, heterosexual couples who met online but not through online dating had an odds ratio of transitions to marriage of 1.52 times greater than couples who met offline, and this odds ratio was not significantly different from 1. In each of Models 2 through 5, meeting through online dating was significantly associated with a higher odds ratio for transitions to marriage, and in models 2-5 (prospective or retrospective data, weighted or unweighted), couples who met online but not through online dating had odds of transition to marriage that were not significantly different from couples who met offline. In Model 3 of Table 3, people with college degrees transitioned to marriage more quickly, as did evangelicals and fundamentalist Christians.

In Model 5, the odds ratio of transition to marriage was $(1.74/0.93)=1.87$ times higher for couples who met through online dating compared to couples who met online but not through online dating. The results of Table 3 support Hypothesis 3, because Internet dating is the most choice-intensive type of online meeting, and is also the type of online meeting most associated in the HCMST data with the transition to marriage.

[Figure 2 here]

Figure 2 shows the cumulative transition to marriage over the 2009-2015 period, for heterosexual couples who were unmarried in HCMST wave 1. Unmarried couples were at risk for marriage until breakup, or until censoring at the time of their last HCMST survey response. Consistent with the analyses in Table 3 above, couples who met through online dating had the fastest transitions to marriage. Of the couples who met through online dating and who remained together, it took between 3 and 4 years of the relationship for 50% of the couples who remained intact to transition to marriage. For couples who met offline, more than 10 years of couple duration were required before half of the relationship cohort had transitioned to marriage according to Figure 2's unweighted Kaplan-Meier (1958) estimates.

It is important to note that the confidence interval in Figure 2 for heterosexual couples who met through online dating is wide because there were only 13 transitions to marriage in the prospective HCMST data for heterosexual couples who met their partners in online dating.⁸ Even though the number of transitions to marriage among the couples who met through online dating

⁸ Because 9 of the 13 marriages in Figure 2 of couples who met through online dating had their month of marriage randomly allocated within twelve month windows, Figure 2 shows transitions to marriage in yearly, rather than monthly increments (so that Figure 2 is unaffected by random allocations of the month of transition to marriage).

was small in the prospective HCMST, the marriages occurred so early in the relationships that the cumulative marriage rate was significantly higher than for couples who met offline. The retrospective HCMST data yields the same result (couples who met through online dating transitioned to marriage fastest) with much larger sample size of marriage transitions (see Table 1). The main competing risk for transition to marriage is transition to breakup. Table 2 and Figure 1 above show that couples who met online and couples who met offline have indistinguishable breakup rates (controlling for relationship duration), so the competing risk from breakup should not bias Figure 2.

Discussion:

Analysis of individual level data from the HCMST project shows that meeting online is not associated with couple breakup (though the number of breakups in HCMST is modest, so the power to reject null hypotheses of small differences in breakup rates is modest as well). Meeting online was significantly associated with transitions to marriage. Heterosexual couples who met online transitioned to marriage faster than heterosexual couples who met offline. My results overturn earlier results (Paul 2014) that claimed that HCMST data showed that couples who met online had higher rates of breakup.

The HCMST result showing that couples who met online transition to marriage more quickly helps explain why Cacciopo et al (2013) found that a high percentage of married couples in the U.S. (35%) had met online. My results are consistent with Bellou (2015), who found that marriage rates were *highest* in US states that had faster adoption of broadband Internet, though Bellou's marriage data were not individual level data.

The association between meeting online and faster transitions to marriage in the HCMST data was entirely driven by couples who met through Internet dating websites, the sites on the web where choice of potential partners is maximized. Meeting through the choice-intensive process of Internet dating was significantly associated with transitions to marriage for heterosexual couples in both the prospective and the retrospective versions of the HCMST event history data.

One of the key advantages to Internet dating is information gathering from the potential partner's profile. Many of the kinds of questions that the Internet dating websites gather are the

kinds of questions that are difficult to ask a potential partner at the first date or early in the relationship (e.g., what is your philosophy about money, what particular kind of travel do you like, do you have herpes, do you own a gun). One 28-year-old HCMST respondent who later married his partner of 3 years, wrote: “We met through [an Internet dating website]. We started getting to know one another through the website, then through phone/email, then after about a month, started dating in person. Our first phone conversation lasted into the middle of the night... We are soul mates... We complement each other well. We are planning to be engaged in the next month...”

Religious individuals tend to be particularly goal-oriented about marriage, and they often find themselves in thin dating markets where similarly religious potential partners are difficult to identify. A 41 year old white HCMST respondent wrote: “Met on a Christian website. We first met in person 3 days after making initial online contact. We got engaged 2 months after meeting.” A 59 year old black Christian HCMST respondent found his mate because the woman saw a picture of him leaving church with his mother on his profile, and she contacted him because she liked the picture. They had to wait 4 months for first meeting because they lived in different cities, but the day after they first met in person, they decided to get married.

Although meeting through an Internet dating website is associated with faster transitions to marriage for heterosexual couples, there is likely a strong selection effect in drawing the most marriage-ready individuals into the Internet dating websites. I cannot rule out the possibility that *all* of the association between meeting through online dating and transitions to marriage is due to the self-selection of the most marriage-ready individuals into online dating. If all of the association between online dating and transitions to marriage were due to self-selection, this would still imply that marriage-ready individuals *believed* that online dating improved their chances of finding a mate.

Newer technologies for meeting include GPS-based phone applications Tinder and Grindr, which have reputations for promoting hookups and short term relationships rather than committed relationships (Sales 2015b). Nationally representative data do not yet exist to explore Tinder and Grindr’s effects on social interactions. The HCMST data only cover couple meetings up to March of 2009. Smartphone apps such as Tinder and Grindr post-date 2009. There were pre-2009 Internet services that prioritized hookup culture, such as Craigslist casual encounters, and the website hotornot.com. It is very unlikely that negative social impacts of the Internet will

have emerged since 2009, if (as I document in this paper) there were no discernible negative impacts of the Internet on couples who met online before 2009.

In waves 4 and 5 of HCMST, subjects who reported a breakup were asked to explain, in an open-ended text box, why their relationship broke up. Of these 136 (mostly brief) breakup stories, 28 (21% of all breakups) mentioned infidelity as a reason for the breakup, and of those 28, six individuals (4% of all breakups) mentioned the Internet as a cause of the infidelity. One subject wrote: “She and I had constant disagreements and were unhappy. She then revealed to me that she was in love with one of my friends she met on the Internet (not in real life). We divorced afterwards.” The 4% of HCMST breakups that report Internet infidelity as a cause of the breakup is far less than the 68% that Manning (2006) reported of divorces resulting from Internet infidelity. Infidelities, like other relationships, often have an online component. We lack the data that would be required to know whether the internet era has made infidelity more or less common.

Despite anecdotal evidence for the negative effect of Choice Overload in Internet dating (Slater 2013; Turkle 2015; Heino, Ellison and Gibbs 2010; Vitzthum 2007), representative survey-based studies of the internet’s effect on couples find either null effects or positive effects of meeting online (Cacioppo et al. 2013; Rosenfeld and Thomas 2012). Glassner (2010) argues that cyberspace, being relatively new, is an attractive locus for our unfounded fears.

References:

- Ansari, Aziz, and Eric Klinenberg. 2015. *Modern Romance*. New York: Penguin Press.
- Baker, Reg, Stephen J. Blumberg, J. Michael Brick, Mick P. Couper, Melanie Courtright, J. Michael Dennis, Don Dillman, Martin R. Frankel, Philip Garland, Robert M. Groves, Courtney Kennedy, Jon Krosnick, Paul J. Lavrakas, Sunghee Lee, Michael Link, Linda Piekarski, Kumar Rao, Randall K. Thomas, and Dan Zahs. 2010. "Research Synthesis: AAPOR Report on Online Panels." *Public Opinion Quarterly* 74 (4):711-781.
- Bellou, Andriana. 2015. "The Impact of Internet Diffusion on Marriage Rates: Evidence from the Broadband Market." *Journal of Population Economics* 28:265-297.
- Cacioppo, John T., Stephanie Cacioppo, Gian Gonzaga, Elizabeth L. Ogburn, and Tyler J. VanderWeele. 2013. "Marital Satisfaction and break-ups differ across on-line and off-line meeting venues." *Proceedings of the National Academy of Sciences* 110 (25):10135-10140.
- Callegaro, Mario, and Charles DiSogra. 2008. "Computing Response Metrics for Online Panels." *Public Opinion Quarterly* 72 (5):1008-1032.
- Chang, Linchiat, and Jon A. Krosnick. 2009. "National Surveys via RDD Telephone Interviewing versus the Internet: Comparing Sample Representativeness and Response Quality." *Public Opinion Quarterly* 73 (4):661-674.
- Finkel, Eli J., Paul W. Eastwick, Benjamin R. Karney, Harry T. Reis, and Susan Sprecher. 2012. "Online Dating: A Critical Analysis from the Perspective of Psychological Science." *Psychological Science in the Public Interest* 13 (1):3-66.
- Fricker, Scott, Mirta Galesic, Roger Tourangeau, and Ting Yan. 2005. "An Experimental Comparison of Web and Telephone Surveys." *Public Opinion Quarterly* 69 (3):370-392.
- Glassner, Barry. 2010. *The Culture of Fear: Why Americans are Afraid of the Wrong Things*. Tenth Anniversary Edition. New York: Basic Books.
- Google Scholar. 2017. "Robert E. Kraut Google Scholar profile." <https://scholar.google.com/citations?user=HKGYvu4AAAAJ&hl=en&oi=ao>. Accessed January 7, 2017.
- Hancock, Jeffrey T., Catalina Toma, and Nicole Ellison. 2007. "The Truth about Lying in Online Dating Profiles." in *ACM Conference on Human Factors in Computing Systems*. San Jose, CA: CHI.

- Heino, Rebecca D., Nicole B. Ellison, and Jennifer L. Gibbs. 2010. "Relationshopping: Investigating the Market Metaphor in Online Dating." *Journal of Social and Personal Relationships* 27 (4):427-447.
- Iyengar, Sheena S., and Mark R. Lepper. 2000. "When Choice is Demotivating: Can One Desire Too Much of a Good Thing?" *Journal of Personality and Social Psychology* 79 (6):995-1006.
- Kaplan, E.L., and Paul Meier. 1958. "Nonparametric Estimation from Incomplete Observations." *Journal of the American Statistical Association* 53 (282):457-481.
- Kraut, Robert, Sara Kiesler, Bonka Boneva, Jonathon Cummings, Vicki Helgeson, and Anne Crawford. 2002. "Internet Paradox Revisited." *Journal of Social Issues* 58 (1):49-74.
- Kraut, Robert, Michael Patterson, Vicki Lundmark, Sara Kiesler, Tridas Mukophadhyay, and William Scherlis. 1998. "Internet Paradox: A Social Technology that Reduces Social Involvement and Psychological Well-Being?" *American Psychologist* 53 (9):1017-1031.
- Manning, Jill C. 2006. "The Impact of Internet Pornography on Marriage and the Family: A Review of the Research." *Sexual Addiction & Compulsivity* 13 (2/3):131-165.
- McGuigan, K. A., P. L. Ellickson, R. D. Hays, and R. M. Bell. 1997. "Adjusting for Attrition in School-Based Samples: Bias, Precision, and Cost Trade-Offs of Three Methods." *Evaluation Review* 21 (5):554-567.
- McKenna, Katelyn Y.A., and John A. Bargh. 2000. "Plan 9 From Cyberspace: The Implications of the Internet for Personality and Social Psychology." *Personality and Social Psychology Review* 4 (1):57-75.
- McKenna, Katelyn Y.A., Amie S. Green, and Marci E. J. Gleason. 2002. "Relationship Formation on the Internet: What's the Big Attraction." *Journal of Social Issues* 58 (1):9-31.
- Nie, Norman, and D. Sunshine Hillygus. 2002. "The Impact of Internet Use on Sociability: Time-Diary Findings." *IT & Society* 1 (1):1-20.
- Oppenheimer, Valerie Kinckaid. 1988. "A Theory of Marriage Timing." *American Journal of Sociology* 94:563-591.
- Parigi, Paolo, and Warner Henson, II. 2014. "Social Isolation in America." *Annual Review of Sociology* 40:153-171.

- Paul, Aditi. 2014. "Is Online Better than Offline for Meeting Partners? Depends: Are you Looking to Marry or to Date?" *Cyberpsychology, Behavior, and Social Networking* 17 (10):664-667.
- Rogers, William. 1993. "Regression Standard Errors in Clustered Samples." *Stata Technical Bulletin* 13:19-23.
- Rosenfeld, Michael J. 2014. "Couple Longevity in the Era of Same-Sex Marriage in the US." *Journal of Marriage and Family* 76:905-918.
- Rosenfeld, Michael J., and Reuben J. Thomas. 2012. "Searching for a Mate: The Rise of the Internet as a Social Intermediary." *American Sociological Review* 77 (4):523-547.
- Rosenfeld, Michael J., Reuben J. Thomas, and Maja Falcon. 2015. "How Couples Meet and Stay Together." Stanford, CA: Stanford University Libraries. waves 1, 2, and 3 version 3.04; wave 4 supplement version 1.02; wave 5 supplement version 1.0.
<http://data.stanford.edu/hcmst>.
- Sales, Nancy Jo. 2015a. *American Girls: Social Media and the Secret Lives of Teenagers*. New York: Alfred A. Knopf.
- . 2015b. "Tinder and the Dawn of the 'Dating Apocalypse'." *Vanity Fair*, September 2015b.
<http://www.vanityfair.com/culture/2015/08/tinder-hook-up-culture-end-of-dating>
- Sautter, Jessica M., Rebecca M. Tippett, and S. Philip Morgan. 2010. "The Social Demography of Internet Dating in the United States." *Social Science Quarterly* 91 (2):554-575.
- Scheibehenne, Benjamin, Rainer Greifeneder, and Peter M. Todd. 2010. "Can There Ever be Too Many Options? A Meta-Analytic Review of Choice Overload." *Journal of Consumer Research* 37:409-425.
- Schwartz, Barry. 2005. *The Paradox of Choice: Why More is Less*. New York: Harper Perennial.
- Simon, Herbert A. 1990. "Invariants of Human Behavior." *Annual Review of Psychology* 41:1-19.
- Slater, Dan. 2013. *Love in the Time of Algorithms: What Technology Does to Meeting and Mating*. New York: Penguin.
- Turkle, Sherry. 2011. *Alone Together: Why We Expect More from our Technology, and Less from Each Other*. New York: Basic Books.
- . 2015. *Reclaiming Conversation: The Power of Talk in a Digital Age*. New York: Penguin Press.

- Vitzthum, Virginia. 2007. *I Love You, Let's Meet: Adventures in Online Dating*. New York: Little, Brown and Company.
- Wang, Hua, and Barry Wellman. 2010. "Social Connectivity in America: Changes in Adult Friendship Network Size from 2002 to 2007." *American Behavioral Scientist* 53 (8):1148-1169.
- Wellman, Barry. 1979. "The Community Question: The Intimate Networks of East Yorkers." *American Journal of Sociology* 84 (5):1201-1231.
- . 2001. "Physical Place and Cyberplace: The Rise of Personalized Networking." *International Journal of Urban and Regional Research* 25 (2):227-252.
- White, Halbert. 1980. "A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity." *Econometrica* 48:817-830.
- Winship, Christopher, and Larry Radbill. 1994. "Sampling Weights and Regression Analysis." *Sociological Methods and Research* 23 (2):230-257.
- Yamaguchi, Kazuo. 1991. *Event History Analysis*. Newbury Park, California: Sage Publications.
- Yang, Mu-Li, and Wen-Bin Chiou. 2010. "Looking Online for the Best Romantic Partner Reduces Decision Quality: the Moderating Role of Choice-Making Strategies." *Cyberpsychology, Behavior, and Social Networking* 13 (2):207-210.
- Young, Kimberly S. 1998. *Caught in the Net: How to Recognize the Signs of Internet Addiction- And a Winning Strategy for Recovery*. New York: John Wiley & Sons.

Table 1: Descriptive Statistics for two event history datasets derived from the How Couples Meet and Stay Together surveys

Data Version	Prospective	Retrospective
Period Covered	2009-2015; couple met in 1995 or later	1998-2014 [¥]
time unit	couple-months	couple-years
Outcome 1: Breakups	407	N/A
Maximum number of couple-periods exposed to risk of breakup	57,081 months (equivalent of 4,757 couple years)	N/A
N of couples exposed to the risk of breakup	1,341	N/A
Weighted Percent of couples who met online	13.7%	N/A
Outcome 2: Transitions to marriage among heterosexual couples	109	814
Maximum number of couple-periods exposed to the risk of marriage	17,153 couple months (equivalent of 1,429 couple years)	7,646 couple years
N of unmarried heterosexual couples exposed to the risk of marriage during the exposure period	619	1,430
Weighted Percent of subjects who met online	15.4%	13.0%

¥ Time limits in the retrospective data set by availability of unmarried heterosexual couples who met online. The difference in N of couples between the retrospective and prospective versions of the data is that the retrospective dataset includes an additional 705 couples who married between 1998 and 2009, plus an additional 106 couples who were unmarried at HCMST wave 1 in 2009 and who were lost to follow-up thereafter. If we relax the filter for prospective data that the couple needed to have met in 1995 or later, there would be 503 breakups, and 121,989 couple-months of exposure to the risk of breakup in 2009-2015.

Table 2: Comparison of event history relationship outcomes by Internet influence, using prospective data for 2009-2015, odds ratios from weighted event history logistic regressions

	1	2	3
Outcome	Broke up	Broke up	Got Married
Filters	none	Met in 1995 or later	Met in 1995 or later; Heterosexual couples
Internet Variable	Met Online	Met Online	Met Online
Raw Odds ratio of rates without controls for Met Online compared to Met Offline [with 95% CI]	2.46¥ [1.72, 3.52]	1.31 [0.91, 1.89]	1.98* [1.06, 3.72]
Adjusted Odds Ratios for Met Online compared to Met Offline (from logistic regressions with controls) [with 95% CI]	1.10 [0.77, 1.59]	0.96 [0.66, 1.39]	1.86 [0.98, 3.52]
Controls	age, relationship duration, relationship duration ^{-0.5} , formal union, same-sex couple, college degree	age, relationship duration, relationship duration ^{-0.5} , formal union, same-sex couple, college degree	age, age ² , relationship duration, relationship duration ² , race, college degree, evangelical

Source: How Couples Meet and Stay Together, all outcomes took place between wave 1 of HCMST and wave 5 of HCMST, 2009-2015. Rates and comparisons of rates are weighted by weight variable “weight2.” Confidence intervals determined by event history logistic regressions with robust standard errors, with standard errors clustered on couples. Controls are all time varying except for the following: same-sex couple status, race, evangelical/fundamentalist identity. Race excluded from the models predicting breakup because the race terms were insignificant, but see Appendix 3.

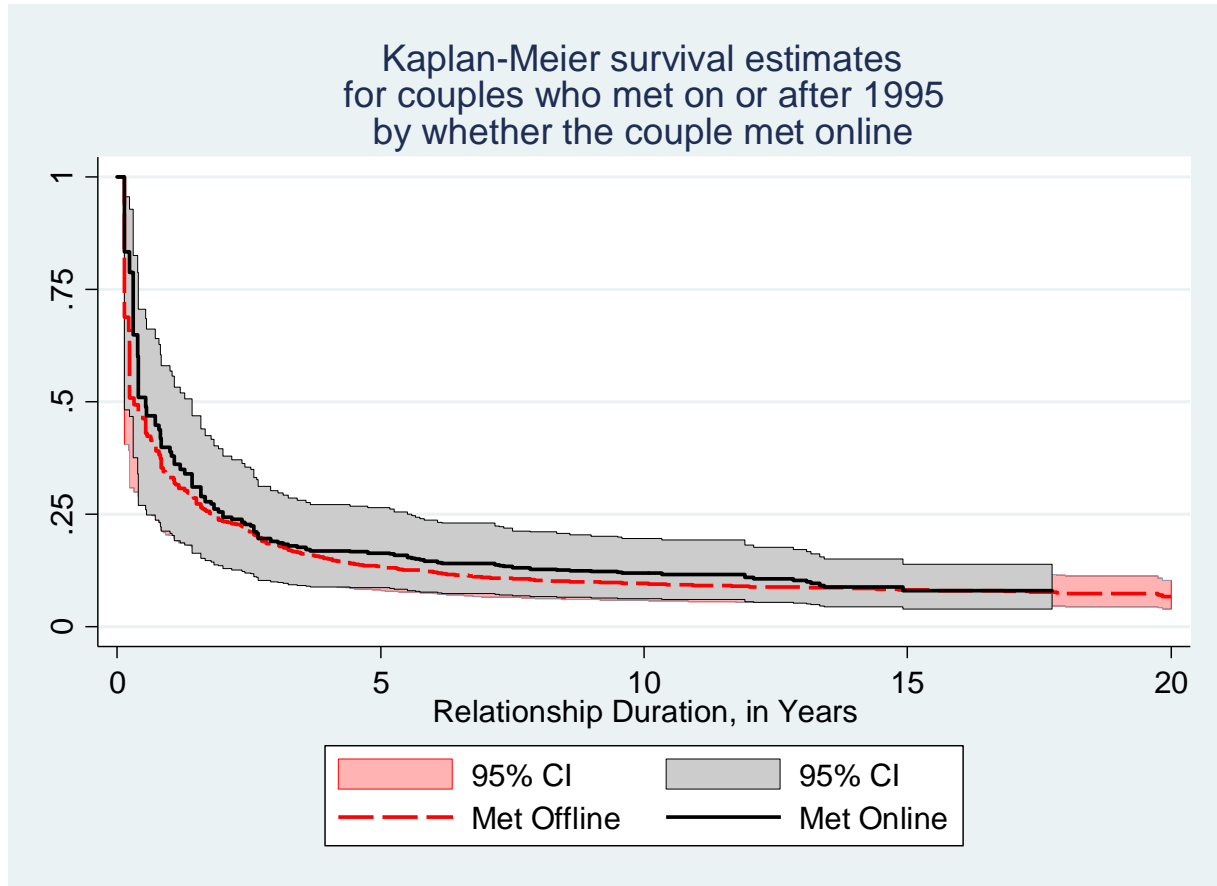
* P<0.05 ¥ P<0.001, two tailed tests.

Table 3: Predicting Transitions to marriage for heterosexual couples, odds ratios [and 95% CI] from event history logistic regressions

	1	2	3	4	5
Data type	Prospective	Prospective	Prospective	Retrospective	Retrospective
Weighted	Yes	Yes	No	Yes	No
Met Online (ref: met offline)	1.98* [1.06, 3.72]				
A) Met Online subset: not Internet Dating (ref: met offline)		1.52 [0.63, 3.66]	1.37 [0.62, 3.04]	1.21 [0.86, 1.69]	0.93 [0.69, 1.24]
B) Met Online subset: Online Dating (ref: met offline)		3.19¥ [1.48, 6.88]	3.38¥ [1.64, 6.99]	1.83¥ [1.36, 2.45]	1.74¥ [1.35, 2.24]
Relationship Duration in Years		1.47* [1.04, 2.08]	1.42* [1.09, 1.84]	1.06 [0.99, 1.13]	1.05 [0.99, 1.11]
Relationship Duration Squared		0.97 [0.94, 1.00]	0.98* [0.95, 0.99]	0.99* [0.99, 0.999]	0.99* [0.99, 0.999]
Subject Age		1.17 [0.96, 1.42]	1.05 [0.80, 1.37]	1.06 [0.99, 1.13]	1.07* [1.02, 1.13]
Subject Age squared		1.00 [0.99, 1.00]	1.00 [1.00, 1.00]	0.99* [0.99, 0.999]	0.99¥ [0.99, 0.999]
Subject is black		0.82 [0.31, 2.18]	0.89 [0.38, 2.07]	0.63* [0.42, 0.93]	0.64* [0.47, 0.88]
Subject has BA		2.26* [1.34, 3.80]	2.24¥ [1.39, 3.63]		
Subject Evangelical or Fundamentalist		1.29 [0.67, 2.47]	1.81* [1.09, 3.00]		
Calendar year				0.94¥ [0.92, 0.96]	0.94¥ [0.92, 0.96]
Test: B/A		2.10 [0.68, 6.49]	2.46 [.89, 6.81]	1.51 [1.00, 2.30]	1.87¥ [1.30, 2.68]

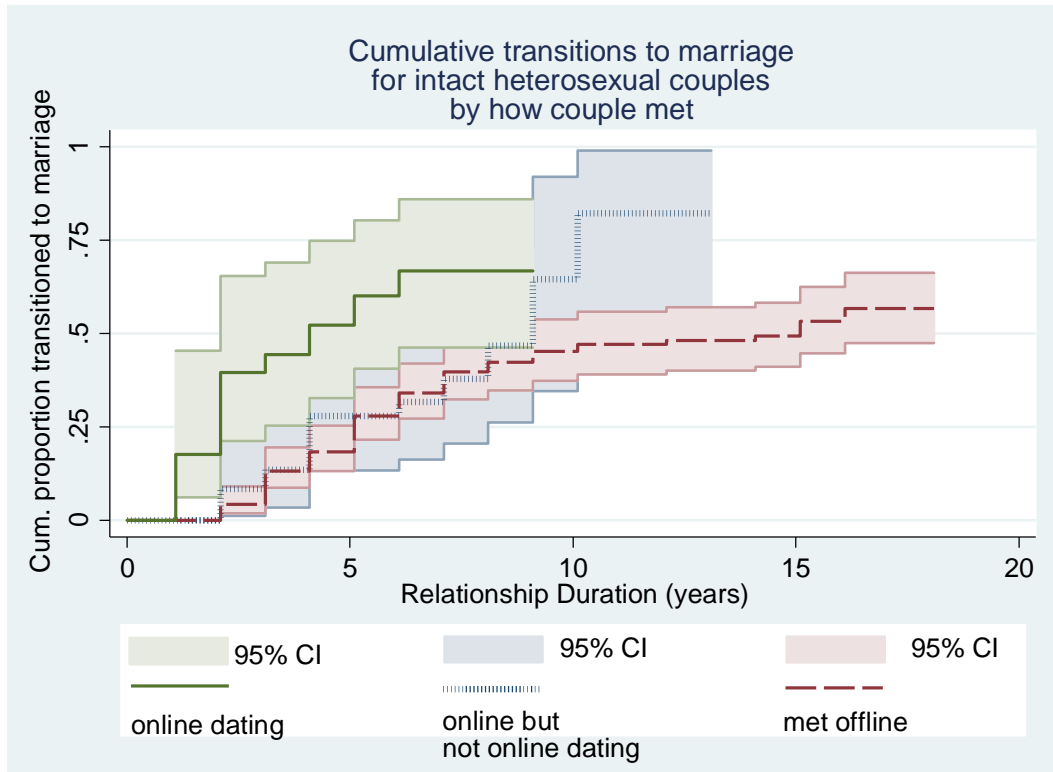
Source: How Couples Meet and Stay Together, waves 1-5. Prospective data cover exposure to the risk of marriage in 2009-2015, for heterosexual couples who met in 1995 or later. Retrospective data cover exposure to the risk of marriage in 1998-2014 for heterosexual couples regardless of when they met. All models are clustered on individual couples, with robust standard errors. Weighted logistic regressions are weighted by weight variable "weight2." Unweighted regressions add variable "recsource" as a predictor because "recsource" predicts the weights. Coefficients for "recsource" and for 3 additional race categories not shown. Calendar year not entered into the prospective models as a predictor because it was not significant. All predictors are time varying except for met online, race, and evangelical/fundamentalist which was asked in 2009. Subject's education is not known in the retrospective data. * P<0.05; ¥ P<0.001, two tailed tests

Figure 1: Meeting online not associated with lower couple longevity



Source: HCMST prospective event history data, covering breakups between 2009 and 2015. Data are unweighted. Cumulative survival in this case means couple survival as an intact couple, i.e. without breakup. Number of breakup events: 94 for couples who met online, and 300 for couples who met offline.

Figure 2: Meeting through Internet dating associated with faster transitions to marriage



Source: HCMST prospective data. Unweighted cumulative transitions to marriage in 2009-2015 (with 95% confidence intervals) for intact heterosexual couples who were unmarried at HCMST wave 1. Cumulative transitions to marriage is $1 - S$ where S =(Kaplan- Meier survival curve for remaining unmarried). Number of transitions to marriage: 13 for heterosexual couples who met through online dating, 12 for heterosexual couples who met online but not through online dating, and 84 transitions to marriage for heterosexual couples who met offline.

Appendix 1: Replicating Paul (2014). Log Odds coefficients from unweighted logistic regression models predicting couples staying together (over SE)

	Wave 2 alone, Married Couples	Wave 3 alone, Married Couples	Waves 2-5 Combined in Event History Model for Married Couples	Waves 2-5 Combined in Event History Model for Married Couples, met on or after 1995
Logit Coefficient for Met Online reported by Paul	N/A	-1.64_¥ (0.54)		
My replication of Paul's Models: Logit coefficient for Met Online	N/A	-0.86 (0.59)		
Logit coefficient for met online, from event history model			-0.56 (0.29)	-0.46 (0.32)
breakups of couples who met online	0	5	16	14
breakups of couples who met offline	30	18	96	56
Paul's Controls	Yes	Yes	Yes	Yes

¥ P<0.01, two tailed test

Source: HCMST wave 2, wave 3, and waves 1-5, and Table 2 of Paul, Aditi. 2014. "Is Online Better than Offline for Meeting Partners? Depends: Are you Looking to Marry or to Date?" *Cyberpsychology, Behavior, and Social Networking* 17 (10):664-667. Paul's controls are: relationship quality, duration of relationship (continuous), partner age, respondent age, partner race (categorical), and respondent race (categorical). Following Paul, the dichotomous dependent variable here is couple stability compared to breakup; reversing the dichotomous outcome so that breakup is compared to stability would simply reverse the signs of the log odds ratio coefficients.

Notes:

Paul analyzed each wave separately, rather than combining waves into one event history dataset, and Paul analyzed married couples separately from unmarried couples. Paul's only significant association with meeting online was for the couple stability of married couples from Wave 3. Because there were zero breakups of married couples who met online in HCMST Wave 2, and therefore there was no coefficient for "met online" predicting married couple stability using Wave 2, Wave 2's married couples were discarded.

Examining only Wave 3 by itself, Paul claimed that married couples who met online were more likely to break up than married couples who met offline, and this was the only significant association Paul reported between meeting online and couple stability. In my replications of Paul's models I have

been unable to replicate Paul's lone significant association between meeting online and couple stability of married couples in Wave 3. Where Paul reported a significant coefficient of -1.64 for married couple stability for couples who met online (see above) from Wave 3, my replication results found instead an insignificant coefficient of -0.86. Paul did not respond to my email requesting clarification about Paul's models. The fact that there were only 5 breakups of married couples who met online in Wave 3, makes it unlikely in my view that any reasonable model could be constructed that would find the 5 breakups to be significantly more than expected.

Column 3 shows an event history model predicting couple stability, using HCMST data from waves 2-5, and using Paul's set of controls (to the extent that I was able to replicate them). Note that in column 3, the coefficient for met online comes close to significance at the 5% level, with $Z = (.56 / .29) = 1.93$, and two tailed P value of 0.054. Adding additional controls (not shown), such as "resource" which is recommended when weights are not used, diminished the "met online" association with married couple breakup further. Standard errors rose further when clustering and weights were taken into account (not shown).

In column 4, I make one simple adjustment from column 3, excluding couples who met before 1995, because couples who met before 1995 were meeting before meeting online was a generally available option. This simple adjustment, which results in a comparison of couples who met online to couples who met their partners in the same era, is more appropriate, and further reduces the association of meeting online with marital breakups to insignificance.

Appendix 2: Predicting breakup in HCMST, Comparing models with regular survey weights and with attrition-adjusted weights, odds ratios from weighted logistic regressions [with 95% CI]

	1	2
Data type		
Weighted	regular weights	attrition-adjusted weights
Met Online (all)	1.10 [0.77, 1.59]	1.10 [0.76, 1.57]
Subject Age	0.997 [0.99, 1.01]	0.997 [0.99, 1.01]
Relationship Duration in Years	0.96¥ [0.94, 0.98]	0.96¥ [0.94, 0.98]
Relationship Duration ^{-0.5}	2.02¥ [1.67, 2.45]	2.03¥ [1.67, 2.47]
Subject is a formal union	0.14¥ [0.10, 0.20]	0.14¥ [0.10, 0.20]
Same-sex couple	1.34 [0.90, 2.00]	1.47 [0.98, 2.21]
Subject has BA	0.74* [0.57, 0.97]	0.76 [0.58, 1.00]
N of subjects	2593	2593
N of couple-months	119,240	119,240
Wald Chisquare (7df)	539.4	532.6

Source: How Couples Meet and Stay Together, waves 1-5. Prospective data cover exposure to the risk of breakup in 2009-2015. Both models are clustered on individual couples, with robust standard errors. Column 1 is the same as column 1 of Table 2, weighted by survey weight “weight2.” Following McGuigan et al (1997), the attrition adjusted weights are ((weight2)/P) where P is the predicted probability of response to any of HCMST waves 2-5, and where P is determined by a logistic regression with the following wave 1 predictors: gender, marital status, relationship duration, race, recruitment source, log income, number of minor children in household, age, age squared, educational attainment, and having Internet access at home, plus a dummy variable for whether the subject ever retired or withdrew from the KN/GfK panel in waves 2-5. As noted in the text, retirement or withdrawal from the KN/GfK panel was the best single predictor of wave 2-5 response rate. Because of the way the original HCMST weights were designed, the weights create especially large errors for same-sex couples in HCMST, which is why analysis of the same-sex couples is better done without weights, see Rosenfeld (2014).

* P<0.05; ¥ P<0.001, two tailed tests

Appendix 3: Predicting breakup in HCMST, with additional controls, odds ratios from weighted logistic regressions [with 95% CI]. Effect of meeting online on odds of breakup not much affected by presence of other controls.

	1	2	3	4	5	6
Filter	none	none	none	met >=1995	met >=1995	met >=1995
Met Online (all)	1.10 [0.77, 1.59]	1.10 [.77, 1.56]	1.14 [0.82, 1.59]	0.96 [0.66, 1.39]	0.95 [0.66, 1.37]	1.01 [0.72, 1.42]
Subject Age	0.997 [0.99, 1.01]	0.99 [0.98, 1.003]	0.99 [0.98, 1.00]	0.99 [0.98, 1.02]	0.99 [0.99, 1.01]	0.99 [0.98, 1.01]
Relationship Duration in Years	0.96¥ [0.94, 0.98]	0.96¥ [0.95, 0.98]	0.96* [0.94, 0.99]	0.92* [0.87, 0.97]	0.92* [0.88, 0.97]	0.92¥ [0.87, 0.97]
Relationship Duration ^{-0.5}	2.02¥ [1.67, 2.45]	2.03¥ [1.67, 2.48]	1.79¥ [1.48, 2.18]	1.81¥ [1.44, 2.27]	1.80¥ [1.42, 2.28]	1.62¥ [1.29, 2.05]
Subject is a formal union	0.14¥ [0.10, 0.20]	0.15¥ [0.10, 0.21]	0.26¥ [0.17, 0.40]	0.16¥ [0.11, 0.24]	0.17¥ [0.11, 0.24]	0.23¥ [0.15, 0.36]
Same-sex couple	1.34 [0.90, 2.00]	1.34 [0.91, 2.00]	1.27 [0.69, 2.32]	1.51 [0.99, 2.31]	1.46 [0.96, 2.23]	1.29 [0.70, 2.38]
Subject has BA	0.74* [0.57, 0.97]	0.79 [0.60, 1.02]	0.84 [0.64, 1.09]	0.72* [0.53, 0.96]	0.75* [0.56, 0.99]	0.81 [0.61, 1.07]
Additional Controls	None	Set 1	Set 2	None	Set 1	Set 2
N of subjects	2593	2593	2590	1309	1309	1309
N of couple-months	119,240	119,240	119,105	56,140	56,140	56,140
Wald Chisquare (df)	539.4 (7)	592.3 (12)	742.3 (20)	339.9 (7)	386.9 (12)	561.4 (20)

Source: How Couples Meet and Stay Together, waves 1-5. Prospective data cover exposure to the risk of breakup in 2009-2015. All models are clustered on individual couples, with robust standard errors. All controls are time varying except: same-sex couple status, and race. The main predictor, met online, is fixed over time within couples. Model 1 is the same as Table 2, Model 1, row G. Model 4 is the same as Table 2, Model 2, row G. Additional controls, set 1 (5df): region (3 df), metro (1 df), log income in 2009 dollars (1df). Additional controls, set 2 (13 df): region (3 df), metropolitan area vs rural (1 df), log income in 2009 dollars (1 df), coresident(1 df), number of minor children in the HH (1 df), couple is lesbian (1 df), respondent race/ ethnicity (5 df).

* P<0.05; ¥ P<0.001, two tailed tests