

Supplement to:

Rosenfeld. Michael J. 2017. "Marriage, Choice, and Couplehood in the Age of the Internet." *Sociological Science* 4: 490-510.

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\yen (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

\yen 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 $((\text{weight}2)/P)$ 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