Online Exogamy Reconsidered: Estimating the Internet's Effects on Racial, Educational, Religious, Political and Age Assortative Mating

Reuben J. Thomas

The University of New Mexico

reubenjthomas@unm.edu

February 16, 2018

DRAFT: Please ask the author for the most recent version before citing or sharing.

Abstract

Increases in the rates of interracial and interreligious couples within the U.S. have occurred seemingly in tandem with the rise of the Internet and online dating, but the evidence connecting online sources of romance and couple heterogeneity have been limited and mixed. Using a unique dataset on how U.S. couples met, collected in 2009 and again in 2017, I find that couples who met online are more likely to be interracial, interreligious, and of different college degree status, but also more similar in age. Parental education endogamy, a proxy for social class endogamy, is not related to whether or not couples met online. Different types of online meetings affect couple characteristics differently: dating websites and apps are not associated with greater racial diversity than offline meetings, but other Internet sources are, which includes online communities, games and chat rooms. These non-dating online sources are also related to greater political diversity in couples, but dating sites are not. On the other hand, only dating websites/apps are predictive of greater educational diversity in couples, though they also produce more age-similar couples than offline sources of romance. Couples who meet through a combination of on- and offline settings are not significantly different from those who meet purely offline across these social dimensions. Population-level estimates suggests that only a small part of the recent changes in couple diversity can be directly attributed to couples meeting online, but there is the potential for larger Internet-induced change if it becomes the primary source of romantic introductions.

The potential for the Internet to drastically change the structure of social relationships seems profound, yet the social impact of technological transformations can often be underwhelming, with stratification and social boundaries largely reproduced across technological epochs. A seemingly very straightforward and direct way that the Internet is altering social relationships is through online dating, the creation of new couples from strangers who otherwise may never have had the opportunity to meet one another. Now that Internet sources account for over a fifth of new U.S. couples, and most new same sex couples (Rosenfeld and Thomas 2012), the structure of online romantic interactions is of growing importance in determining the kinds of couples and families that are formed within the U.S. If the Internet is bringing together people from different social groups and class positions, resulting in more families that blend together multiple groups, then those social boundaries could be substantially weakened in a world of increasing online dating. However, it's also quite possible that the ways people meet online actually reinforce some group boundaries, creating more homogenous couples and families than would result otherwise.

What we know about the diversity of couples that result from meeting online has so far been very limited. National probability samples that both include data on how couples met and are recent enough to capture the rise of online dating have been very rare: studies using one survey of the U.S. (Potarca 2017; Rosenfeld and Thomas 2012) and one of Germany (Potarca 2017) found mixed results on the impact of the Internet on couple assortativity. Both found that couples who met online had greater odds of being interreligious, but otherwise found that, controlling for other factors, they differed only

from specific offline settings that are exceptionally segregated on a social dimension: Potarca (2017) found that online dating is related to less educational endogamy than are school-formed romances in both countries, and related to less racial endogamy than meeting through family members in Germany. Using non-probability samples from opt-in online survey panels, Dutton et al. (2009) found that couples in Spain and the UK who met online were more dissimilar in age and education than those who found each other offline. None of these previous studies controlled for the local geographic diversity the respondents were embedded within, however, which is a potentially major confounding factor in determining couple composition.

Evidence from profiles and messaging within online dating websites has been more plentiful (Hitsch et al. 2010; Huber and Malhrota 2017; Lewis 2016; Lin and Lundquist 2013; Robnett and Feliciano 2011; Skopek et al. 2011), but this kind of evidence only offers a limited window into one early phase of couple formation, within specific websites. While these have consistently shown that online interactions are segregated by most social dimensions, it is not clear that the extent is greater or less than what occurs in offline romantic markets. This evidence also doesn't connect online interaction patterns to the couples that result (one exception is Lee's [2015] study of an online dating website in South Korea, which found increased age and educational endogamy in the resulting couples, but less occupational endogamy). Assortativity can change significantly between different stages in online relationship progression (Bruch et al. 2016), and it is reasonable to assume it likewise can change in the transition to offline stages of romance. While studies that only look at online interactions do shed light on important processes behind

couple formation, they can't really tell us about the Internet's effects on the endogamy of couples in a population.

This is the first nationally representative study of couples to examine differences between couples who meet online and offline controlling for the diversity of their local geography, as well as the first to demonstrate a general and robust effect of online dating on the racial, educational and age composition of the couples that result, compared to all other couples. I also offer here estimates of the population changes in endogamy from meeting online, which has largely been a matter of speculation.

The Internet's Potential to Break Down Group Boundaries

The initial, and perhaps greatest barrier to the creation of diverse couples is the lack of opportunities for contact between people from different groups. If the spaces, groups and networks through which people find their romantic partners are already highly segregated on a social dimension, then the resulting couples will also tend to be highly segregated on that dimension, regardless of other factors (Blau, Blum and Schwarz 1982). Early-to-mid 20th century U.S. marriages typically occurred between those who lived within the same or nearby neighborhoods (Bossard 1932), and couple endogamy mirrored residential segregation. In the mid-to-late 20th century U.S., introductions through friends was the most common source of marriages (Rosenfeld and Thomas 2012), creating couples that represented the homogeneity of friendship networks. Interactions online, on the other hand, can potentially occur between any two people who share a common language.

While some dating websites are organized around ethnic and cultural niches, the vast majority of online dating occurs in general audience and openly-accessible apps and websites (Smith and Duggan 2013). Even the chat rooms, online organizations and games that friendships and romances begin within have far fewer barriers to entry than offline groups and voluntary organizations. Internet spaces for meeting new people are radically less exclusive than traditional sources of romance. This expands both the scope and the diversity of the local market for mates that people interact with, and reduces the involvement of family and friends in screening potential partners (i.e. third party interference) (Kalmijn 1998).

The Potential for the Internet to Reinforce Group Boundaries

There are also reasons to suspect that the Internet may actually increase romantic assortativity, beyond just reproducing the patterns in offline couplings. Communication is easier and more likely to occur when there is shared cultural content (Carley 1991). Cultural differences are correlated with a number of other social distinctions, such that sorting on cultural similarity also tends to produce sorting on race, ethnicity, social class, religion, etc., to varying degrees (Blau 1977; Stark and Flache 2012). While this process can occur in any communication medium, online interfaces and structures may heighten cultural homophily.

By design, online dating is an exercise in cultural homophily. The profile options of many dating websites and apps encourage the listing of activities and consumptions that the daters like and dislike, to explicitly signal their specific position within the larger

culture space. Even when apps and websites don't explicitly prompt for that, the presentation of self through the construction of profiles and the framing and selection of images conveys a wealth of cultural information about the online dater. The focus on this information may prompt online daters to sort and select on cultural factors: some have argued that online dating fosters a shopping mentality in sizing up potential mates (Heino et al. 2010), encouraging a check-list approach to sorting prospects, augmenting discriminatory preferences that would otherwise be weaker and more easily overcome by other personal characteristics in face to face communication.

Couples who find each other online, but not through online dating, tend to do so through online communities, chat rooms, games and social networking websites like Myspace and Facebook (Rosenfeld and Thomas 2012). Online games and communities are voluntary organizations much like their offline counterparts, selected on shared cultural interests, and may exhibit similar segregation by gender, social class and race (McPherson and Smith-Lovin 1987). Chat rooms can likewise function as interest-focused communities, though they can also be very open public spaces that facilitate general social interaction between strangers. Social network apps/websites, by contrast, are loose reflections of the user's offline social ties, though sometimes more expansive. Online extensions of offline networks may help broker new ties, but these connections may be expected to still reflect the homogeneity of those offline social circles. Communication in online communities and through social network websites has been shown to be highly segregated by political ideology (Adamic and Glance 2005) and other cultural factors (Lewis et al. 2008), as well as by race (Wimmer and Lewis 2010).

Self-selection into online spaces can determine the degree of segregated interaction between online venues, but the interface(s) through which users interact and perceive each other can affect segregation within venues. In some online spaces, demographic characteristics of the users are predominantly featured, but in others they are obscured and underemphasized. Users of an online game may spend many hours together without guessing each other's ethnicity or education levels, but online dating profiles typically display this information front and center. The assumed intentions behind most user interactions may also influence the assortativity of the ties that result. When romance is the explicit goal of an online space, interactions are perceived through that lens from the onset, which can heighten a shopping mentality for desired mate attributes (Heino et al. 2010). When other activities are the focus, friendship can come first, and people can later be surprised by an emerging romance.

The above arguments represent competing visions of the Internet, one as a more egalitarian space freed from the structural and hierarchical boundaries of the offline social world, and the other as a highly Balkanized place in which users self-select into echo-chambers of similarly minded others. The reality of the Internet is a mix of these characterizations, but it is not entirely clear *a priori* which aspect should have a stronger effect on the kinds of couples that find each other online.

Data and Methods

The How Couples Meet and Stay Together survey (HCMST) (Rosenfeld, Thomas, and Falcon 2015) was gathered using GfK\Knowledge Networks' (GfK\KN) panel of respondents, with an initial sample surveyed in 2009, five follow-up waves with those same respondents over the next several years, and then a second sample of new respondents in 2017. My analyses here focus on the initial wave in 2009 and the new sample in 2017 (the only waves that included information on how couples met and their partners' characteristics). The GfK\KN panel altered its sampling\recruitment method between these sample years: the 2009 subjects were selected by a random digit dial of U.S. phone numbers, with follow-up recruitment by phone and certified mail. Panel members in the 2017 sample were selected by an address-based sample of the 50 U.S. states and Washington D.C. The GfK\KN panel consisted of over 40,000 members in 2009 and over 55,000 in 2017, of which 4,002 and 3,510 were included in the HCMST study, respectively, with response rates of 71% and 50%. When thinking of the response rate as a cumulative process through all of the steps from the initial recruitment into the panel, to attrition within the panel, to screening for and then completion of the HCMST survey, then the cumulative rate would only be somewhere in the teens. However, such a rate isn't directly comparable to traditional surveys, as KN partially controls for potential attrition bias by using demographic information collected at each survey stage (Couper 2000). KN panel surveys have been shown to perform better than traditional random-digit-dial samples (Baker et al. 2010, p. 743; Chang & Krosnick 2009; Fricker et al. 2005), and are used by a number of prominent social science research data collection projects, such as the American National Election Survey (2017).

Both samples only cover the U.S. English-speaking population of the 50 states and Washington D.C. Surveys were conducted through a web interface, and GfK\KN provided the necessary equipment and Internet access to those who did not already have home access, through television set-top devices at the time of the 2009 survey, and through tablet computers by the time of the 2017 survey. Panel members received 3-4 surveys a month, requiring an average of 15 minutes to complete, which they could finish piecemeal over the course of a week.

Sampling Couples

From the initial samples of 4,002 and 3,510 respondents (in 2009 and 2017 respectively), the survey asked if they were currently in a relationship, of which 3,009 and 2,862 responded that they were. The 993 un-partnered respondents in 2009 were not asked any further questions, and are not included in the analyses here. The 648 un-partnered respondents in 2017 were instead asked about their most recent relationship, of which 541 reported that they had such a relationship. To account for differences between current versus retrospective relationship sampling, I control for both whether the couple was still together at the time of the survey, as well as the duration of the relationship in years, which is the time since relationship onset for current couples. I limit the analyses here to only couples who met after 1995, to ensure comparisons between offline- and online-formed couples are only within the era that Internet usage was at least somewhat common. This reduces the sample of couples to 3,130. This sub-sample includes an oversampling of same sex couples, so I include a binary indicator of same sex composition in all analyses, as well as survey weights that account for the oversample.

Analyses without same sex couples produced very similar results. See Table 1 for sample means and proportions of all of the variables described below.

Detecting Internet-Created Couples

I use two approaches towards classifying couples as those who met online versus not. The first is a simple dichotomy of those who indicate any role of the Internet in their first meeting versus those who do not. The second is a more detailed classification that compares those who met strictly online as otherwise-strangers to both those who met purely offline and those who met through a combination of online and offline connections. This second classification scheme also breaks down the strictly online meetings into websites and phone apps specifically designed for finding partners versus all other online spaces.

The HCMST data contains highly detailed information about how couples first met, uniquely so for a national sample. The survey included an open-ended question asking respondents to tell their story in detail:

> 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.

This was followed by a large text box, and the web interface prompted respondents to add more after their initial submission (regardless of how much they wrote), prompting them a second time if their story was less than 100 characters. The median story was 185 characters long, with 75% longer than 100 characters and 25% longer than 353 characters. The first wave of these stories was open-coded by two of the data authors to

determine non-mutually-exclusive categories of how the couples met, and then coded by these categories by the authors and an additional research assistant, with a high degree of inter-coder reliability. The second wave of the data was coded by a research assistant according to the categories from the first wave. I additionally read and coded the stories of all respondents who identified an online meeting in the multiple choice question or by the existing story coding, for my more detailed classification.

After telling their story, the respondents were asked a series of multiple choice questions about their first meeting with their partner. This included the following question about the role of the Internet in their meeting, and the kind of online venue they met it:

Did you use an Internet service to meet [Partner_Name]?

No, I did NOT meet [Partner_Name] through the Internet
Yes, a social networking site (like Facebook or Myspace)
Yes, an Internet dating or matchmaking site (like eHarmony or match.com)
Yes, an Internet classified advertising site (like Craigslist)
Yes, an Internet chat room
Yes, an app on my phone (like Tinder or Grindr) [2017 sample only]
Yes, a different kind of Internet service

If there was any indication of an online meeting in either the respondent's answer to this question or in their open-ended story, I classified the couple as having met online in the first, more inclusive measure of Internet-created couples, following Rosenfeld and Thomas (2012).

To distinguish between couples that met as strangers online and those that met through a combination of online and offline introductions, I first use two questions in the HCMST that identify prior social network/context overlap (Kalmijn and Flap 2001). The first of

these asked whether the respondent's parents and their partner's parents knew each other before they first met, and the second asked whether they had friends in common, or their friends knew each other, before meeting each other. The survey also asked whether they attended the same high school, college, and/or grew up in the same town. While these may often be a good indicator of prior offline social network overlap, this is highly dependent on the size of those units, which is unknown, so I don't use school and town overlap here. If a respondent classified as having met online also identified prior family or friend overlap in these questions, I code that couple as having met through a combination of online and offline sources. I additionally re-read the meeting stories of all respondents who were coded as having met online above, and code respondents as having met their partner both online and offline that identified any offline brokerage in their story. This results in a high bar of classifying couples as Internet-created, meant to identify previously unconnected strangers who found each other only through the Internet, and would likely have never met otherwise.

I further break down this met-strictly-online category into those who identified having met through a dating website or phone app (either in their multiple choice answer or in their story) and all other strictly-online meetings. This results in four categories of couple-formation: purely offline-formed couples, purely online-formed couples who met through an Internet dating site or phone app, purely online-formed couples who met in another (or unidentified) Internet venue, as well as couples who met through a combination of online and offline sources. This last on/offline hybrid category includes couples who knew each other in a previous life stage, lost touch, and then reconnected

through a dating website, social network site, etc. It also includes couples who met friends-of-friends through social networking sites, their initial communication purely online but the shared context that connected them largely offline. The Other Internet category includes a wide variety of online venues such as chat rooms, online games, online organizations and interest-oriented online communities, and unspecified online meetings.

Measuring Couple Characteristics

The HCMST survey asked about respondents' partners' race and Hispanic ethnicity separately, which I combined into a single scale that treats Hispanic as a master category, and also includes White, Black, Asian\Pacific-Islander, Native American, and Other. Respondents had already answered a similar set of questions about themselves for the GfK\KN panel demographic survey, which I recoded into the same categories. Respondents could identify themselves as multiple races: I coded anyone who identified as Hispanic into that category, and anyone who identified as Black but not Hispanic into the Black category, and all other multiracial respondents into the Other category. As a robustness check, I replicated the findings below with alternate coding schemes, coding all multiracial as others, or treating Asian as a third master category, and found very similar results to those presented here.

Education: Respondent's, Partner's and Mothers' Education was measured on a 14 point scale for both the respondent (in the Gfk\KN panel survey) and the respondent's partner, as well as for both the respondent's and partner's mothers, as reported by the respondent.

This scale asked for the highest grade or degree completed, including no formal education, 1st-4th grade, 5th-6th grade, 7th-8th grade, 9th grade, 10th grade, 11th grade, 12th grade without a diploma, high school diploma or equivalent, associate's degree, bachelor's degree, master's degree, and professional or doctorate degree. I use a dichotomized indicator of college graduate status in the analyses below. Note that there are no significant differences between online and offline formed couples in the absolute value differences of these full scales, whether comparing respondents' and partners' or their mothers' education levels.

Religion The 2009 HCMST survey included questions about the respondent's partner's religious identity, and both the respondent's and partner's identity at age 16, but the 2017 survey did not. These were modeled on a GfK\KN panel question that was previously asked of all respondents. The response options were Baptist, Other Protestant, Catholic, Mormon, Jewish, Muslim, Hindu, Buddhist, Pentecostal, Eastern Orthodox, Other Christian, Other Non-Christian, and None. I recoded these into a 5 category scheme: Catholic, Protestant/Other Christian, Jewish, and None (following Potarca [2017] and Rosenfeld and Thomas [2012]). I replicated the analyses below using the original 13 categories, which produced only minor differences in the results. Similar results are also found when examining endogamy by religious identity at age 16 instead of current identity.

Politics, Age and Controls Both surveys included information about the respondent's partner's political identification. As the scales were different between the two surveys, I simplify both to three categories here: Republican, Democrat and Other. I also examine

age assortativity below, operationalized as the absolute age difference in years between partner and respondent, from both samples. I include additional covariates in the multivariate models to control for potentially confounding factors, including the duration of the relationship, when the couple first met, whether the couple was married, coresidency, the respondent's and couple's gender, respondent's household income scale, the number of children in the respondent's household, and the region the respondent lived in at the time of the survey.

Local Diversity Measures

A major threat of spuriousness to any claim of the Internet's effect on exogamy lies in the potential differences in the diversity of places where online dating is popular versus the places where people less often meet online. If people in more diverse urban areas are more likely to find long term partners online, for instance, then it could spuriously appear that the Internet is creating more heterogeneous couples, even if couples within the same geographic units are no different if they meet online versus offline. Despite this concern, this is the first study to compare online and offline endogamy that controls for the local diversity the respondents are embedded within.

I use two different geographic identifiers from the HCMST data: where the respondent lived at the time they met their partner, and where the respondent lived at the time of the survey. The first directly and clearly reflects the demographics of the local opportunities for mates that the respondent was embedded in at the time they found their partner, but has two major drawbacks: first, the measurement is based on the respondent naming the

place they lived, and thus can only be matched to metro\micro-politan areas rather than more precise local geographic units. The second drawback is that this question was only asked of the 2017 sample, which both limits the N of the analyses that use it and precludes its use in the religious exogamy analyses, which is only possible with the 2009 data. Gfk\KN collected the zip codes of respondents at the time of both surveys, which allows for more specific measures of racial and educational diversity in additional to MSA level measures, as well as the inclusion of the 2009 cases in the analyses. Segregation tends to be recreated across geographic relocations (Sampson and Sharkey 2008), such that current local area diversity is typically a good proxy for the diversity that people were exposed to in the past as well. I'll test this assumption in the results below, but most tables and models will use the data about the respondents' location when surveyed.

I matched respondents to information about their current zip code and their current and past metro/micropolitan area (MSA). Some rural respondents' zip codes and places they previously lived were not within an MSA, so I include information about their county instead. Some respondent's lived in zip codes that could be matched to multiple MSAs or counties; for these I calculated a weighted average of the MSA\county measures by the distribution of their zip code between them, as apportioned by the MABLE/Geocorr12 Geographic Correspondence Engine (Missouri Census Data Center 2012).

Diversity can be measured in a number of different ways. For race, religion and college degree diversity, I simply control for the proportion of the geographic unit of the same group as the respondent. This ego-centric way of coding diversity is the most directly

determinant of the dependent variables in the analyses: whether the respondents matched with a partner different from themselves. I don't have data on proportions of political party identification in these geographic areas, so I use a partisan voting scale instead (see below).

I use data from the Decennial Census in 2000 and the American Community Survey's (ACS) 5-year estimates in 2010 and 2015 (United States Census Bureau 2015), to measure the racial\ethnic, educational and income composition of the respondent's zip codes and their larger metro-\micro-politan area (MSA). The year of the survey or the year the couple met determines which of these three sources is used, such that the greatest temporal mismatch is no more than five years (e.g. a 2005 meeting matched to the 2010 ACS data). The models below control for the proportion of the respondent's zip code and MSA/county that are the same racial category and same college graduate status as the respondent. The race\ethnicity measures use the ACS's reported categories, which treat Hispanic as a separate non-racial category, and only separate out non-Hispanic respondents for one racial category, White. This yields 7 categories: White Non-Hispanic, Hispanic, Black, Asian, Pacific Islander, Native American, Other, and Multiracial. Despite the double counting of many non-White Hispanics, these are nonetheless good indicators of the proportion of people in respondent's local areas that identify as the same category as themselves.

I measure the respondents' local political context using Cook's Partisan Voting Index (Wasserman and Flinn 2017) for their congressional district (or Washington D.C.), from 2010 for the 2009 sample and from 2017 for that newest sample. These scores represent

the Democrat or Republican advantage in the district based on the most recent presidential election, in terms of vote percentage (e.g. "D+5"). I recoded these into absolute values to create a scale of partisanship, ignoring which party had the advantage in the district, such that D+10 and R+10 both become scores of 10, with zeroes representing evenly split districts.

I measure local religious diversity at the MSA\county level using the 2010 U.S. Religion Census Religious Congregations and Membership Study (Grammich et al. 2012), a survey of the 236 major religious organizations in the U.S. as identified by the Association of Statisticians of American Religious Bodies. Note that this is information about religious groups' official membership counts in each geographic area, not a sample of people. As such, it cannot tell us about rates of self-identification or private religious identity (e.g. what proportion of the MSA identifies as Baptist, or Agnostic), only about the rates of the formal identity of congregation membership. I calculated the proportion of the MSA\county population that are the same religion as the respondent based on the estimated membership of 3 broad types of congregations (Catholic, other Christian, other religion), as well as a 4^{th} category that is the inferred proportion of the population that are not members of any congregation.

Models

To test whether the Internet is leading people to be in more or less diverse relationships than they otherwise would have been in requires controlling for a variety of other potentially confounding factors such as education, income, local geographic density and

diversity, etc. To control for these potential sources of spuriousness, I use multivariate logistic regressions predicting couple diversity, and OLS regression models to predict years of age difference between respondents and their partners.

There are a small proportion of cases with missing data, but their exclusion or inclusion in the analyses does not substantively alter the results. Over 94% of the respondents who gave any information at all about their current or most recent relationship gave full information on all of the covariates used here. Two percent of the respondents lacked geographic identifiers. The remaining cases with missing data did not answer one or more questions in the survey, most often information about their partner, such as partner's education, partner's mother's level of education, partner's politics, and partner's race. I replicated the models below using multiple imputation of the missing independent variable data, finding very similar results. I present here the models using listwise deletion of missing data.

Results

The Expansion of Online Couple Formation

Figure 1 illustrates local-regression-smoothed survey-weighted rates of the three types of online-initiated couples: those that met through dating websites, those that met through other Internet venues, and those that met by a hybrid of offline and online venues/networks. While all have risen in prevalence since the 1990's, online dating websites and apps have risen more steadily and strongly. Other Internet venues were once more common sources of couples than dating websites, but their rise has tapered off since

the mid-2000's. This may reflect the fading stigma of online dating and its growth in popularity around this time, largely replacing chat rooms and other online romantic options. Hybrid online/offline first meetings have grown at a steady but slower rate, and may track the rise in social networking websites/apps that make reconnections and introductions to mutual friends easier. Of the three online categories, only online dating sites and apps has a clearly upward trajectory that is suggestive of continued rapid growth.

Respondents who found their partners purely online do exhibit some notable differences from those who found a partner offline (Table 2), which reinforces the need for multivariate controls in estimating the Internet's effects on endogamy. U.S. residents who met their partner online are more often college educated ($p<.01$), were over four years older on average when they met their partner ($p<.001$), and those relationships are on average almost three years newer ($p<.001$). Those who found their partner online also tend to live in metros with higher median incomes ($p<.05$).

Meeting Online and Exogamy

Couples who met online since 1996 are 6 percent more often interracial than couples who met purely offline (Table 3), but this ignores potentially confounding factors. From the multivariate model in Table 4, I estimate that couples who met online have over one and a half times greater odds of crossing the major racial\ethnic boundaries of the U.S., or an average 6% greater probability across the observed values of the covariates ($p<.01$), similar to the observed bivariate difference. While there's not a difference between online

and offline formed couples in their closeness on the educational scale (not shown), couples who met online are more likely to include both a college graduate and a non-graduate, both bivariately (Table 3) and with controls (Table 4). Those who met purely online have over 1.5 times greater odds of this kind of educational exogamy, or a 7.5% increased probability at observed covariates ($p<.01$). This is also similar to the observed bivariate difference of 7.5% in couples formed since 1996. As reported in previous studies (Potarca 2017; Rosenfeld and Thomas 2012), couples who meet online are more likely to be interreligious. In simple bivariate terms (Table 3), couples who met purely online are 13% more often interreligious. Controlling for other factors, I estimate online formed couples to have over 1.8 greater odds (Table 4), or a 12.5% increased probability across observed values of the covariates ($p<.01$).

Couples who meet online are more endogamous in one regard: they are more similar in age than those who met offline. While the bivariate difference is very small and not statistically significant (Table 3), once other factors are controlled for I estimate that those who met online are more than six tenths of a year closer in age than those who met purely offline (Table 4). Couple who met online are not more or less likely to identify with a different political party, however, nor are they more or less likely to have educationally dissimilar mothers.

Why Haven't We Seen This Before? Changes between the 2009 and 2017 Samples

While previous research found limited evidence of the Internet's effect on exogamy when controlling for other factors (Potarca 2017; Rosenfeld and Thomas 2012), I find robust

effects of meeting online on exogamy by multiple social dimensions. What changed? The most important factor is that there are now more couples who met online, as well as more couples in general who met during the era of online dating, increasing the statistical power to detect differences between the offline and online couples in that time window. Table 5 compares results from the 2009 HCMST sample used by previous studies to results that only use the new 2017 HCMST sample. Even though the total sample size was smaller in 2017, the subsample of couples who met in the Internet era is several hundred cases greater, allowing for larger N models here. There are likewise many more couples who met online in the newer sample: 454 compared to 274 in the 2009 sample. As a results, the standard errors for the 2017 met-online coefficients are substantially smaller. This isn't the only factor at play in producing these results: the effect sizes predicting racial and age assortativity are also larger in the 2017 data's results. This could be due to more accurate estimation of effects that have been constant, or it could be due to social change during this period. However, interaction effects between the year the couple met and meeting online were not significant for any of the exogamies examined here, so there isn't evidence here of a change over time in the Internet's effects on exogamy.

Local Diversity: When Met versus When Surveyed

I argued above that one could use the diversity of respondents' local areas when interviewed as a good proxy for the diversity they were embedded in when they met their partners, and I do so in most of the models I present here. As I also know the MSA or county that the 2017 respondents lived in when they met their partner, I can test here

whether this assumption holds for these analyses, and it does. The diversity measures of where the respondents currently live account for over 85% of the variance in the racial diversity of where they lived when they met their partner ($r=0.92$, $p<.001$), and over 90% of the variance in college degree diversity ($r=0.95$, $p<.001$), but only 15% of the political partisan voting index ($r=0.39$, $p<.001$). Table 6 compares models predicting interracial, inter-educational, and age diversity in couples, with some using information about the respondents' current metropolitan area, and others using information about the metro area they lived in when they met their partner. As the later data is only available in the 2017 sample, this requires reducing the N and statistical power of the models, so I compare these "MSA When Met" results (models 3) to the both full sample models that use current MSA (models 1) and the same current-MSA models with the smaller 2017-only sample (models 2). I exclude the zip code level data used in the models in Table 4, as zip code level information was not deducible from respondents' answers about where they used to live.

The findings are quite consistent regardless of which geographic diversity measures are used. The differences between models 2 and models 3, which only differ in whether the current or former MSA information is being used, are minimal. The fit statistics are nearly the same, as is the strong effect of MSA level racial diversity on racial exogamy, and the other MSA diversity measures have mostly similar effects across models. The coefficient for meeting online becomes a little stronger in predicting racial exogamy in models 3, crossing the boundary between marginal and statistical significance, while changing little in predicting educational and age exogamy. Information about the MSA

where the couple met should generally be preferred over the MSA when the respondent was surveyed, but here that decision doesn't affect the results, and the benefits of a more accurate estimate of local geographic diversity at the time of couple formation are outweighed by the benefits of a larger sample size, as well as the benefits of the more locally-scaled zip code level diversity data.

Breaking Down Internet Effects on Exogamy

Comparisons with Specific Offline Sources. Table 7 compares couples who met online in any way with those who met through a variety of offline sources without any detectable role of the Internet (similarly to Potarca [2017]). This division of the offline meetings are based on the data authors' coding of the respondent's stories of how they met their partner. Those codes are not mutually exclusive, so I apply the following rules to categories cases that identify multiple of these categories: any case that identifies family brokerage are always classified as Family, and any remaining cases that identify a religious setting is classified as Religious, and then any remaining cases that identify a school setting are classified as a School meeting. Then I continue this process for Work, Voluntary Organizations, Friend Introductions, and lastly Neighbors, in that order. Remaining cases that identify none of these fall into the Other category, which includes bars and restaurants, pre-Internet singles services, vacations, and meetings in public. Thus an introduction through friends at a bar would be classified as a Friend introduction, while introduction through friends at church would be a Religious introduction. Work

meetings include business trips, customer service encounters, and military service. The Friend category includes private parties and significant other introductions.

The Internet's effects on exogamy depend upon the offline source it is being compared to (or replacing): online-formed couples are more likely to be interracial than those met in educational settings, through family, or in the other offline settings category. They aren't more interracial than neighbor-brokered romances, and the differences between online and friend introductions, work or voluntary organizations are statistically insignificant. Unsurprisingly, couples who meet in schooling have half the odds of being different in college degree status than those who meet online, but there are also strong educational exogamy effects of meeting online versus family and friendship brokered romances. Family introductions are the only romantic source that differs from online sources in terms of parental education exogamy: online-formed couples have one and a half greater odds of having one mother with a college degree and one without. Neighbors, workplaces, and settings in the Other category (largely public places) produce more age disparate couples than meeting online, but school romances are more similar in age than Internet formed couples. Religious settings are the only offline romantic source that predict less political exogamy than Internet sources, less than two thirds the odds of producing couples of different political affiliations than meeting online. Couples who meet through religious settings are also far less likely to be of different religions than online couples (eight times lower odds), as are, to a lesser extent, those who meet through schooling, family, and voluntary organizations.

Specific Online Sources and Exogamy. In Table 8, the models from Table 4 are replicated with a more detailed categorization of online sources of romance, with the comparison group of couples who met offline without any hint of Internet involvement. Also included is a hybrid category of those who met through both online and offline introductions, so that the two other online categories are meant to capture only those meeting as otherwise strangers online. The effect of meeting online on racial exogamy is strongest when comparing non-dating online sources of romance to strictly offline couples, and only marginally significant when comparing online dating to offline couples. Online sources other than dating websites/apps (including chat rooms, games, online organizations and unspecified online meetings) predict almost double the odds of an interracial union than do offline romantic sources, or on average a 10.2% greater probability, across observed covariate values ($p<.05$). Internet effects on educational exogamy, on the other hand, are only clearly detectable between the couples that meet through dating websites compared to offline couples, not through other online sources. Couples who meet through online dating have one and a half times greater odds of including one college graduate and one non-graduate, compared to those meeting purely offline, or on average a 7.4% greater probability across the observed covariates ($p<.05$).

Dating websites and apps also decrease age exogamy, meaning they produce more age-similar couples than those who meet offline, by almost 1 year less age difference. This is the only characteristic (studied here) for which meeting online creates more similar couples, and it seems to be confined to the romance-focused spaces of the Internet. Both online dating and other online spaces creates couples who are more likely to be of

different religions than those who meet offline, both to similar extents: roughly double the odds, and on average a 14% greater probability across cases. While there was no general effect of meeting online on political exogamy, there is when comparing only the non-dating Internet sources to strictly offline couples. First meetings in the non-dating parts of the Internet predict 1.6 times greater odds of politically heterogenous couples, compared to couples who meet strictly offline, or on average a 10.9% higher probability across the observed covariates. By comparison, the effects of online dating and combined off/online sources on political exogamy versus strictly offline meetings are close to zero, meaning that there are very similar political combination outcomes from online dating and couples who meet partially or fully offline.

Across all of these types of exogamy, there aren't detectable differences between couples who meet through a combination of offline and online sources versus those who meet purely offline. This isn't entirely due to a smaller sample size than online dating, as the Other Online category is actually the smallest one, and still shows multiple effects. The hybrid off/online category includes people who knew each other offline and later reconnect online, as well as those who are introduced through offline friends and settings, but initiate their communication in an online venue. Unsurprisingly, these semi-online meetings are the least different from offline sources of romance.

Estimating the Internet's Effects on Population-Level Changes in Exogamy

Is finding romantic partners online directly causing population level changes in rates of exogamy? This is a question about a counterfactual, comparing the observed rates to a hypothetical similar population that didn't develop and widely adopt online dating. A simple and straightforward way to estimate this is to ask: how different would the population look like if we replaced the partners of those who met their mates online with the kinds of partners we would predict them to have met offline? I estimate this by comparing the predicted values from the models with significant Internet effects in Table 4 to the predicted values with the online-formed couples switched to offline-met couples. This is equivalent to calculating average marginal effects, but I graph the predictions by the year couples met to show how the rates over time would change in the counterfactual. Figure 2 displays these predicted rates of interracial couples, interreligious couples, and cross-BA couples (where only one has a college degree), averaged with survey weights by the year the couples met. Note that the interreligious predictions are limited to 2009 and earlier by the available data.

I also include a second set of counterfactual predictions to Figure 2: a hypothetical U.S. in which online dating has increased at a faster rate, closer to how it has increased in the same-sex couple population. Instead of a U.S. that saw online sources increase from less than 1% of new couples in 1992 to over one fourth by 2016, this counterfactual U.S. supposes that online-formed couples rose to account for over half of new couples by the end of that time period. I simulate this by first calculating the predicted probability of meeting one's partner online for all respondents, using the same covariates from the models in Table 4. I then set a lower bar for coding each respondent as having met their

partner online (.17); as year-met is a strong predictor of meeting online, this produces a steeper but otherwise similar increase in the online coupling rate compared to the observed increase.

As Figure 2 illustrates, this evidence suggests that the Internet's impact on exogamy by these four dimensions is not large, but not unimportant. The rise in interracial and interreligious couples are predicted to still have occurred even if the online couples had met offline instead, but with a slight less steep increase, amounting to on average 1.6% fewer couples interracial each year in the 2010s, and 2.2% fewer couples of different religions each year in the 2000s. On the other hand, the hypothetical U.S. with a majority of the couples meeting online would see somewhat steeper increases in these kinds of couple diversity, roughly 3.2% more couples crossing racial boundaries each year over the past several years, and 2.7% more couples interreligious each year in the 2000s. College-education heterogeneity in couples has been nearly flat, but these models predict a modest increase in this kind of educational exogamy in a U.S. with more online dating, 4.4% more couples crossing that educational boundary each year in the 2010s, or a decrease in a U.S. without online dating of about 2.2%. Age exogamy seems to have decreased slightly in the 1990s and then experienced a recent uptick. This model estimates that couples would be slightly less age similar in a world without Internet romances, by about one sixth of a year in the 2010s, or a third of a year more similar on average in a world in which more couples met online.

Why do some of the substantial and significant logistic regression coefficients translate into such small population effects? Effects can be both strong multiplicatively and small

in impact when only a minority of the population is affected, and the outcome is relatively unlikely. In the case of interracial couples, the first model in Table 4 predicts that couples who meet online have over 1.5 greater odds of being interracial ($p<.01$), which translate into a 6% greater probability, averaged across cases. Limiting the temporal window to couples who met since 2009, the model predicts that if the 28% who met online had instead met offline, they would have had (on average) a 7% lower probability of being interracial. So the interracial 32% of that 28% of couples would have been reduced to 25% of 28%, resulting in (.28*.07) a little less than 2 percent fewer couples in the overall population who were interracial. This simple illustrative calculation makes some assumptions that the full predictions in Figure 2 do not make, but the results are roughly the same regardless.

These estimates come with some caveats, as they rest on a number of assumptions, including that there are not substantial confounding factors causing both interracial dating and online dating that have been omitted from the models. They also assume zero-to-minimal reverse causation in the relationship between online dating and endogamy. For instance, a cultural shift towards preferring interracial relationships might also cause people to increasingly use online dating, because of the difficulty of finding such partners offline. If that has happened, then the above calculations would be overestimating the role of online dating in creating interracial couples.

Note that these estimates also do not account for potential indirect effects of the Internet on couple diversity. For instance, interracial friendships and relationships begun online could in turn introduce new interracial couples offline whom wouldn't otherwise meet

(Ortega and Hergovich 2017). Such introductions are acts of transitive closure, which typically compounds homophily in social networks (Goodreau et al. 2009), but can also magnify the bridging of group boundaries (Mark and Harris 2012). Second-order and nth-order network effects can't be estimated with this data, but cannot be ruled out.

Discussion and Conclusion

Online sources of romance are related to greater couple diversity than offline formed romances, but not in all social characteristics. Couples who meet online are more likely to be different in race\ethnicity, religion and college degree status than are couples who find each other offline, but more similar in age. The online venues through which couples meet also seems to matter for most of these types of endogamy, with online dating creating more age similar and more educationally diverse couples than offline, while other online sources are related to more interracial and politically diverse couples. Couples who meet through a hybrid of offline and online sources are not significantly different in any of these dimensions from purely offline-formed couples, which suggests that the connection of socially-distant strangers lies at the heart of the Internet's effect on endogamy. The predicted rates of population level endogamy without Internet-formed couples, illustrate that while meeting online significantly increases the odds of couple heterogeneity for any given respondent, the Internet has had a more limited impact on endogamy in the population as a whole. However, as the predicted rates for a hypothetical U.S. with more Internet-formed couples illustrate, the Internet's potential to

impact endogamy is substantial, though perhaps not the sea change that some have envisioned.

Online interaction in generally characterized by contradictory tendencies towards barrier-less communication and self-selection into culturally narrow venues. How people find romantic partners online is a reflection of this tension, encountering both vaster arrays of potential partners as well as conscious and meticulous sorting on a variety of characteristics. The net result for couple characteristics is more diversity than found in offline-formed couples, but far less than the potential of a boundary-less mating market would suggest.

References

Adamic, Lada A. and Natalie Glance. 2005. "The Political Blogosphere and the 2004 U.S. Election: Divided They Blog." *Proceedings of the Third International Workshop on Link Discovery* pp. 36-43.

The American National Election Studies (www.electionstudies.org). 2017. *The ANES Guide to Public Opinion and Electoral Behavior.* Ann Arbor, MI: University of Michigan, Center for Political Studies.

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: 711-781.

Blau, Peter M. 1977. *Inequality and Heterogeneity*. New York: The Free Press.

Blau, Peter M., Terry Blum, and Joseph E. Schwartz. 1982. "Heterogeneity and Intermarriage." *American Sociological Review* 47: 45-62.

Bossard, James H. S. 1932. "Residential Propinquity as a Factor in Marriage Selection." *American Journal of Sociology* 38,2: 219-24.

Bruch, Elizabeth, Fred Feinberg, and Kee Yeun Lee. 2016. “Extracting Multistage Screening Rules from Online Dating Activity Data.“ *PNAS* 113,38: 10530-10535.

Carley, Katherine M. 1991. "A Theory of Group Stability." *American Sociological Review* 56: 331-354.

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: 661-674.

Couper, Mick P. 2000. “Web Surveys: A Review of Issues and Approaches.” *Public Opinion Quarterly* 64: 464-494.

Dutton, William H., Ellen J. Helsper, Monica T. Whitty, J. Galen Buckwalter, and Erina Lee. 2009. “The Role of the Internet in Reconfiguring Marriages: A Cross-National Study.” *Interpersona* 3, suppl. 2: 3-18.

Fricker, Scott, Mirta Galesic, Roger Tourangeau and Ting Yan. 2005. “An Experimental Comparison of Web and Telephone Surveys.” *Public Opinion Quarterly* 69: 370-92.

Goodreau, Steven M., James A. Kitts and Martina Morris. 2009. “Birds of a Feather, or Friend of a Friend? Using Exponential Random Graph Models to Investigate Adolescent Social Networks.” *Demography* 46(1): 103-25.

Grammich, Clifford, Kirk Hadaway, Richard Houseal, Dale E. Jones, Alexei Krindatch, Richie Stanley, and Richard H. Taylor. 2012. *2010 U.S. Religion Census: Religious Congregations & Membership Study.* Lenexa, KS: Association of Statisticians of American Religious Bodies.

Hitsch, Günter J., Ali Hortaçsu, and Dan Ariely. 2010. “Matching and Sorting in Online Dating.” *American Economic Review* 100: 130-163.

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.

Huber, Gregory A. and Neil Malhotra. 2017. “Political Homophily in Social Relationships: Evidence from Online Dating Behavior.” *The Journal of Politics* 79,1: 269-283.

Kalmijn, Matthijs. 1998. "Intermarriage and Homogamy: Causes, Patterns, Trends." *Annual Review of Sociology* 24: 395-421.

Kalmijn, Matthijs and Henk Flap. 2001. “Assortative Meeting and Mating: Unintended Consequences of Organized Settings for Partner Choices.” *Social Forces* 79,4: 1289-1312.

Lee, Soohyung. 2015. “Effect of Online Dating on Assortative Mating: Evidence from South Korea.” *Journal of Applied Econometrics* 31: 1120-1139.

Lewis, Kevin, Jason Kaufman, Marco Gonzalez, Andreas Wimmer, and Nicholas Christakis. 2008. “Tastes, Ties, and Time: A New Social Network Dataset Using Facebook.com.” *Social Networks* 30: 330-342.

Lewis, Kevin. 2016. “Preferences in the Early Stages of Mate Choice.” *Social Forces* 95,1: 283–320.

Lin, Ken-Hou and Jennifer Lundquist. 2013. “Mate Selection in Cyberspace: The Intersection of Race, Gender, and Education.” *American Journal of Sociology* 119,1: 183-215.

Lundquist, Jennifer and Ken-Hou Lin. 2015. “Is Love (Color) Blind? The Economy of Race among Gay and Straight Daters.” *Social Forces* 93,4: 1423-1449.

Mark, Noah P. and Daniel R. Harris. 2012. “Roommate’s Race and the Racial Composition of White College Students’ Ego Networks.” *Social Science Research* 41: 331-42.

McPherson, J. Miller and Lynn Smith-Lovin. 1987. “Homophily in Voluntary Organizations: Status Distance and the Composition of Face-to-Face Groups.” *American Sociological Review* 52,3: 370-379.

Missouri Census Data Center. 2012. MABLE/Geocorr12, Version 1.2: Geographic Correspondence Engine. Web application accessed October 2017 (http://mcdc.missouri.edu/websas/geocorr12.html)

Ortega, Josué and Philipp Hergovich. 2017. “The Strength of Absent Ties: Social Integration via Online Dating.” arXiv:1709.10478 [physics.soc-ph]. Retrieved November 2017 (https://arxiv.org/abs/1709.10478).

Potarca, Gina. 2017. “Does the Internet Affect Assortative Mating? Evidence from the U.S. and Germany.” *Social Science Research* 61: 278-297.

Robnett, Belinda and Cynthia Feliciano. 2011. “Patterns of Racial-Ethnic Exclusion by Internet Daters.” *Social Forces* 89,3: 807-828.

Rosenfeld, Michael J. 2008. “Racial, Educational and Religious Endogamy In The United States: A Comparative Historical Perspective.” *Social Forces* 87: 1-31.

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, Waves 1, 2, and 3: Public version 3.04, plus Wave 4 Supplement Version 1.02 and Wave 5 Supplement Version 1.0* [Data file]. Stanford, CA: Stanford University Libraries. http://data.stanford.edu/hcmst

Skopek, Jan, Florian Schulz and Hans-Peter Blossfeld. 2011. “Who Contacts Whom? Educational Homophily in Online Mate Selection.” *European Sociological Review* 27,2: 180-195.

Smith, Aaron and Maeve Duggan. 2013. *Online Dating & Relationships.* Washington, D.C.: Pew Research Center. Retrieved October 21, 2013 (http://pewinternet.org/Reports/2013/Online-Dating.aspx)

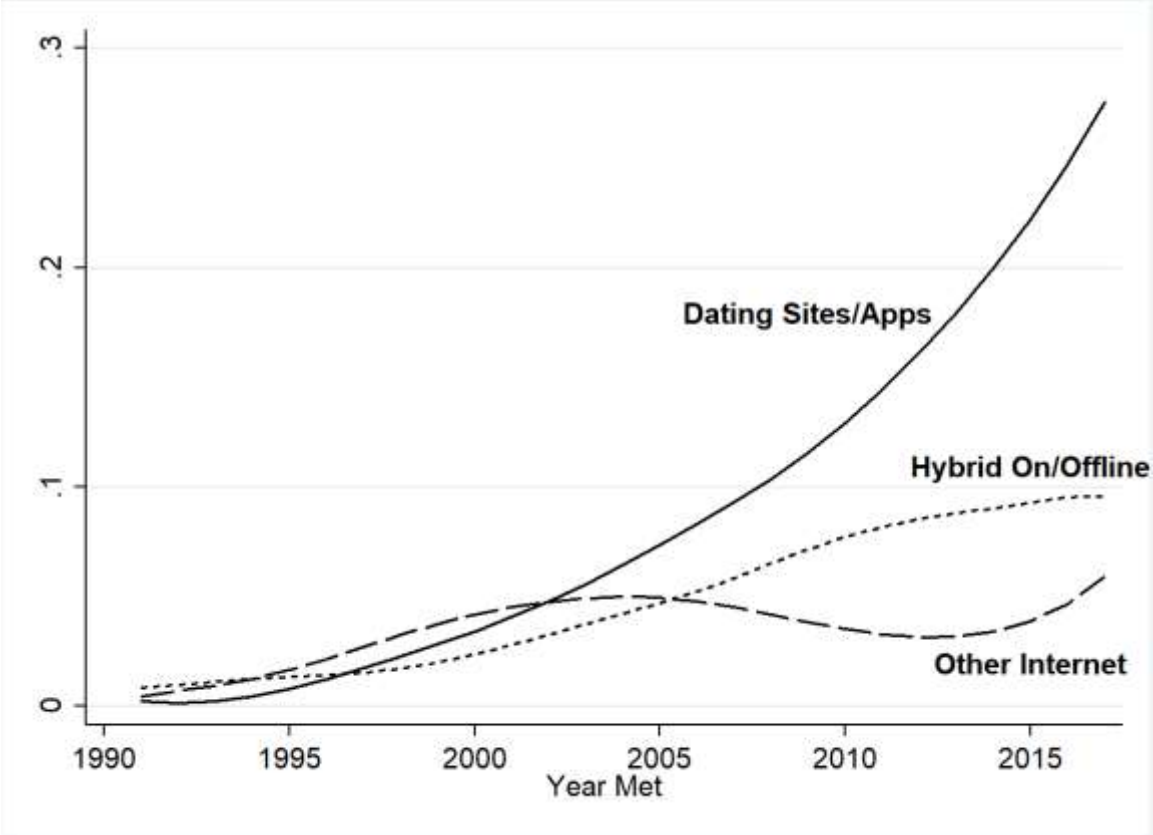
Stark, Tobias H. and Andreas Flache. 2012. “The Double Edge of Common Interest: Ethnic Segregation as an Unintended Byproduct of Opinion Homophily.” *Sociology of Education* 85,2: 179-199.

Wasserman, David and Ally Flinn. 2017. “Introducing the 2017 Cook Political Report Partisan Voter Index.” *The Cook Political Report.* Retrieved April 7, 2017 (http://cookpolitical.com/introducing-2017-cook-political-report-partisan-voter-index)

U.S. Census Bureau. 2015. American Community Survey 5-Year Estimates, Table S0601 generated using American FactFinder. Retrieved October 2017 (http://factfinder.census.gov)

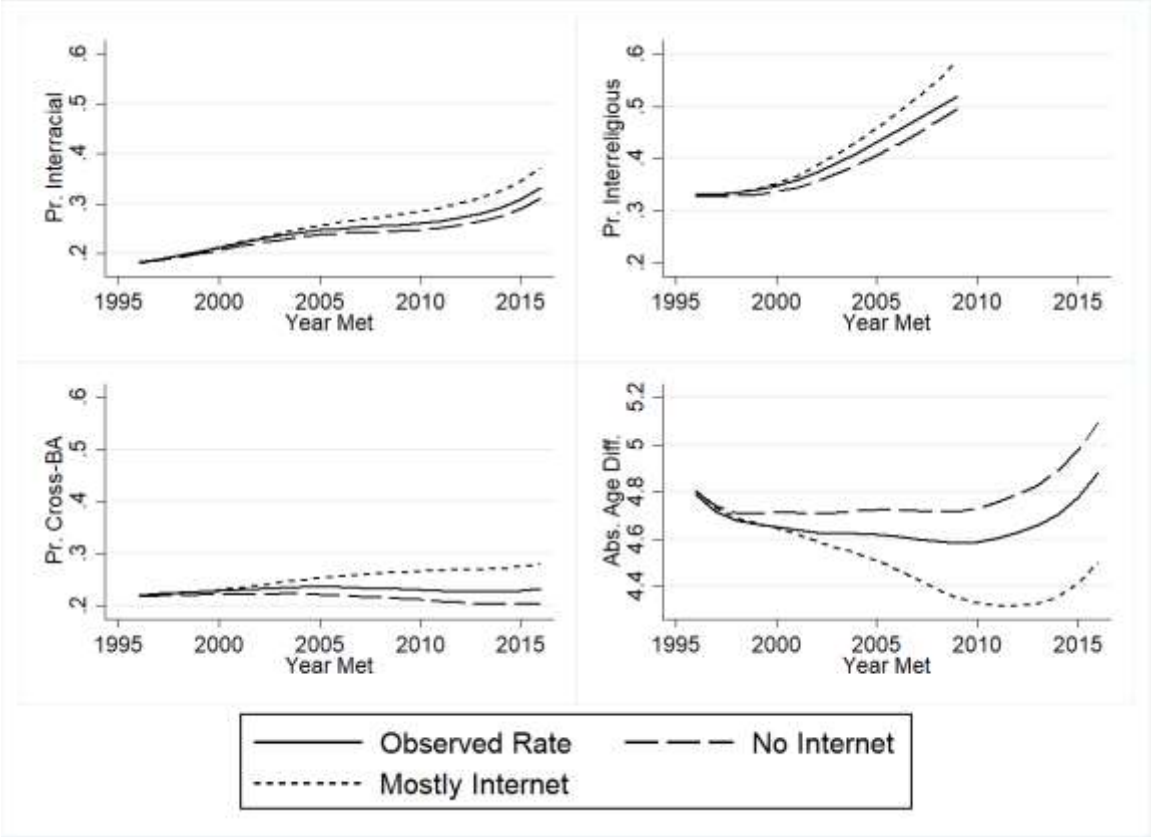
Wimmer, Andreas and Kevin Lewis. 2010. “Beyond and Below Racial Homophily: ERG Models of a Friendship Network Documented on Facebook.” *American Journal of Sociology* 116: 583-642.

Figure 1. Proportions of Couples who Met Online, by Year and Online Source



Survey-weighted lowess regression-smoothed lines (bandwidth=.8), 1991 to 2016

Figure 2. Estimated Population Level Endogamy Given Observed and Hypothetical Rates of Online Romance Formation, by Year Met



Predicted values from models in Table 4.

Table 1. Sample Descriptives

	N	Proportion	Mean	Std. Dev.
Couple Characteristics				
Met Online At All	712	0.23		
Met Through Online Dating Only	368	0.12		
Met Through Other Online Source Only	158	0.05		
Met Both On- and Offline	177	0.06		
Met Purely Offline	2418	0.77		
Different Race/Ethnicity	774	0.25		
Different Religion (2009 sample only)	579	0.43		
Different Political Party	1393	0.45		
Different College Degree Status	808	0.26		
Different Mother's College Degree Status	838	0.27		
Absolute Age Difference			5.03	(5.91)
Same Sex Couple	488	0.16		
Married	1388	0.44		
Coresident	2063	0.66		
Year Met			2005.64	(5.96)
Age Met			30.48	(12.96)
Relationship Length, in years			7.43	(5.62)
Retrospective Relationship Subsample	307	0.10		
Respondent Characteristics				
Age			38.38	(13.40)
College Graduate	1216	0.39		
Mother a College Graduate	755	0.24		
Previously Married	902	0.29		
Home Internet Prior to Survey Panel	2851	0.92		
Race/Ethnicity				
White Non-Hispanic	2147	0.69		
Black Non-Hispanic	349	0.11		
Asian Non-Hispanic	102	0.03		
Hispanic	403	0.13		
Other Race	126	0.04		
Political Party Identification				
Republican	1147	0.37		
Democrat	1899	0.61		
Other	84	0.03		
Religious Identification (2009 sample only)				
Catholic	294	0.22		
Protestant or Other Christian	674	0.50		
Other Religion	122	0.09		
No Religious Identification	255	0.19		

# of Children in Household			0.59	(1.02)
Household Income Scale			12.11 ($50k's)	(4.58)
Local Geography				
% of Zipcode R's Race/Ethnicity			0.60	(0.30)
% of MSA/County R's Race/Ethnicity			0.54	(0.27)
% of Zipcode R's College Degree Status			0.60	(0.24)
% of MSA/County R's College Degree Status			0.56	(0.21)
% of MSA/County R's Religion (2009 only)			0.30	(0.18)
Congressional District Partisan Voting Index			12.66	(9.22)
Zipcode Per Capita Income, in $1000's			15.64	(15.60)
MSA/County Median Income, in $1000's			27.70	(4.81)
Region				
Northeast	579	0.19		
Midwest	701	0.22		
South	1073	0.34		
West	777	0.25		
Sample				
2009	1348	0.43		
2017	1782	0.57		
Total N	3130			

Notes: This subsample only includes respondents who met their partners in 1996 or later.

Table 2. Survey-Weighted Proportions and Means of Covariates, by How Couples Met (Since 1996)

	Met Purely Offline	Met Online or Both	Total
Age Met	28.15	32.31	28.93
Years Known	8.76	5.79	8.20
Household Income	73,603.25	75,522.54	73,944.54
College Graduate R	0.32	0.36	0.33
R's Mother a College Grad.	0.24	0.21	0.24
R's Age at Survey	36.91	38.10	37.14
R Previously Married	0.26	0.37	0.28
R Had Home Internet	0.91	0.96	0.92
R's Race			
White NH	0.60	0.65	0.61
Black	0.14	0.10	0.14
Hispanic	0.17	0.16	0.17
Asian	0.05	0.06	0.06
Other	0.03	0.03	0.03
R's Party			
Republican	0.40	0.41	0.40
Democrat	0.57	0.57	0.57
Other	0.04	0.02	0.03
R's Religion			
Catholic	0.23	0.21	0.23
Other Christian	0.53	0.52	0.53
Other Religion	0.07	0.07	0.07
Non-Religious	0.18	0.20	0.18
Region			
Northeast	0.18	0.18	0.18
Midwest	0.21	0.21	0.21
South	0.37	0.37	0.37
West	0.24	0.24	0.24
MSA Population	3,714,892	3,737,237	3,723,306
Zipcode Pr. R's Race	0.58	0.58	0.58
MSA Pr. R's Race	0.51	0.52	0.51
Zipcode Pr. R's BA	0.63	0.60	0.63
MSA Pr. R's BA	0.60	0.57	0.59
MSA Pr. R's Religion	0.31	0.32	0.31
District Republicanness	-0.72	-0.83	-0.74
District Partisanness	12.91	13.06	12.93
Zipcode Income Per Capita	16,605.99	19,057.50	17,083.41
MSA Median Income	27,428.51	27,794.38	27,496.45
N	2418	712	3130

Survey-weighted estimates from a subsample of couples who met after 1995.

Table 3. Survey-Weighted Proportions and Means of Couple Differentiation, by How Couples Met (Since 1996)

	Different Race	Different Religion	Different College Degree	Different Mother's College	Different Party	Age Difference
Met Online	0.29*	0.51**	0.30***	0.27	0.49	4.63
Met Strictly Offline	0.24	0.38	0.22	0.25	0.45	4.70
Online Sub-Categories						
Dating Site	0.28	0.50+	0.34**	0.32*	0.46	4.66
Other Online	0.31	0.56	0.23	0.16	0.56	5.53
On/Offline	0.32	0.38	0.27	0.25	0.50	4.02
Pure Offline	0.24	0.38	0.22	0.25	0.45	4.70
Total	0.25	0.39	0.23	0.25	0.46	4.69

Survey-weighted estimates from a subsample of couples who met after 1995.
Chi-square tests within columns: + p<.1, * p<.05, ** p<.01, *** p<.001
All t-tests for differences in means of Age Difference were insignificant.

Table 4. Logistic and OLS Regression Coefficients Predicting Exogamy by Meeting Online (Since 1996)

	Different Race/ Ethnicity	Different College Degree Status	Different Mother's College Graduate Status	Absolute Age Difference	Different Religion	Different Political Party
Model	Logit	Logit	Logit	OLS	Logit	Logit
Met Online At All	0.42**	0.44**	0.02	-0.62*	0.61**	0.11
Local Proportion Same as R:						
Zipcode – Race	-1.70***	0.05	0.05	-0.87	-0.00	0.01
MSA – Race	-1.67**	-0.37	-0.77+	-0.56	0.14	0.27
Zip – BA Status	0.12	-1.91***	-0.36	0.02	-0.92	0.28
MSA – BA Status	-0.55	0.59	0.06	1.52	0.39	-0.24
MSA – Religion					-2.23**	
District Partisaness	0.01	-0.00	-0.01	0.01	0.01	-0.01
R^2				0.12		
McFarland's Psuedo-R^2	0.23	0.08	0.13		0.13	0.05
Sample Years	Both	Both	Both	Both	2009	Both
N	3030	3032	2989	3017	1335	3015

+ p<.1, * p<.05, ** p<.01, *** p<.001

Survey-weighted estimates from a subsample of couples who met after 1995. Covariates not shown: Race, Age, Education Scale, Income, Married, Coresident, Years Known, Number of Children in Household, Same Sex, Year Sampled, Relationship Duration, Still Together, Previously Married, Home Internet Access, Region, MSA\County Population and Median Income, and Zipcode Per Capita Income. Religion models include R's religion and political models include R's political party. For full model coefficients, see Appendix Table A1, available at: http://www.unm.edu/~reubenjthomas/OnlineExogamyAppendix.pdf

Table 5: Effects of Meeting Online on Couple Exogamy, by Sample and Controls

	2009 Sample			2017 Sample		
	1	2	3	1	2	3
	Minimal Controls	Full Respondent Controls	Full Respondent Controls + Geo Controls	Minimal Controls	Full Respondent Controls	Full Respondent Controls + Geo Controls
Effect of Meeting Online on:						
Interracial Log Odds	0.12	0.26	0.37	0.34*	0.39*	0.42*
	(0.27)	(0.28)	(0.29)	(0.16)	(0.17)	(0.18)
N	1404	1400	1340	1748	1739	1693
Pseudo R^2	0.32	0.33	0.37	0.11	0.12	0.17
Different College Dgr. Log Odds	0.29	0.38+	0.45+	0.37*	0.38*	0.42*
	(0.23)	(0.23)	(0.24)	(0.16)	(0.16)	(0.17)
N	1405	1402	1342	1749	1739	1693
Pseudo R^2	0.05	0.07	0.10	0.05	0.06	0.07
Age Difference	-0.32	-0.38	-0.48	-0.41	-0.62+	-0.74*
	(0.50)	(0.48)	(0.48)	(0.35)	(0.35)	(0.35)
N	1402	1398	1338	1737	1728	1682
Adjusted R^2	0.09	0.11	0.13	0.12	0.13	0.13

+ p<.1, * p<.05, ** p<.01, *** p<.001

Survey-weighted estimates from a subsample of couples who met after 1995. Minimal controls include the respondent's characteristic in question, respondent age, years since the couple met, and home internet access prior to joining the survey panel. Full respondent controls additionally include race, education scale, income, political party, married, coresident, years known, number of children in household, same sex, year sampled, relationship duration, still together, previously married, home internet access and region. Geo controls include MSA\county population and median income, zip code per capita income, racial\ethnic and college degree diversity for both zip code and MSA, and political partisanness for the respondent's congressional district. For full model coefficients, see Appendix Tables A4a-c, available at: http://www.unm.edu/~reubenjthomas/OnlineExogamyAppendix.pdf

Table 6: Endogamy by Online-vs-Offline: Current MSA vs. MSA When Met

	Different Race/Ethnicity (Logit)			Different College Degree (Logit)			Absolute Age Difference (OLS)		
	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)
Sample Years	Both	2017	2017	Both	2017	2017	Both	2017	2017
Met Purely Online	0.41**	0.34+	0.39*	0.39**	0.44*	0.46**	-0.63*	-0.84*	-0.84*
	(0.15)	(0.19)	(0.19)	(0.14)	(0.17)	(0.17)	(0.28)	(0.35)	(0.34)
Local Proportion Same as R:									
Current MSA – Race	-3.27***	-3.55***		-0.36	0.24		-1.35	-0.23	
MSA When Met – Race			-3.60***			0.27			-0.74
Current MSA – BA Status	-0.38	-0.30		-1.19**	-0.64		1.56+	1.06	
MSA When Met – BA Status			0.06			0.10			1.68+
Current District Partisaness	0.00	0.00		-0.01	-0.01		0.01	0.02	
District When Met Partisanness			0.01			-0.02			0.03
R^2							0.12	0.15	0.15
McFarland's Psuedo-R^2	0.22	0.18	0.18	0.07	0.07	0.07			
N	3030	1552	1552	3032	1552	1552	3017	1544	1544

+ p<.1, * p<.05, ** p<.01, *** p<.001

Survey-weighted estimates from a subsample of couples who met after 1995. Covariates not shown: Race, Age, Education Scale, Income, Political Party, Married, Coresident, Years Known, Number of Children in Household, Same Sex, Year Sampled, Relationship Duration, Still Together, Previously Married, Home Internet Access, Region, and MSA\County Population and Median Income. For full model coefficients, see Appendix Tables A5a-c, available at:
http://www.unm.edu/~reubenjthomas/OnlineExogamyAppendix.pdf

Table 7. Odds Ratios and OLS Regression Coefficients Predicting Exogamy by Romantic Source (Since 1996)

	Different Race/ Ethnicity	Different College Degree Status	Different Mother's College Graduate Status	Absolute Age Difference	Different Political Party	Different Religion
Coefficient Type	Odds Ratios	Odds Ratios	Odds Ratios	OLS Coefficients	Odds Ratios	Odds Ratios
Source of Introduction (vs Internet)						
Friends	0.68^{+}	0.65*	1.00	0.31	1.13	0.69
Family	0.50**	0.58*	0.60*	0.56	0.83	0.48*
Neighbors	1.81	0.58	0.73	5.12***	0.88	1.73
Work	0.79	0.74^{+}	1.18	1.35***	0.92	0.65
School or College	0.60*	0.48***	1.03	-0.70*	0.87	0.52*
Religious Org.	0.65	0.61^{+}	1.03	-0.42	0.56*	0.12***
Vol. Org.	0.96	1.67	1.65	1.82	0.52^{+}	0.37*
Other Source	0.49**	0.76	0.84	1.61**	0.72	0.67
McFarland's Psuedo-R^2	0.23	0.08	0.13		0.06	0.15
R^2				0.15		
Sample Years	Both	Both	Both	Both	Both	2009
N	2896	2898	2860	2885	2884	1309

+ p<.1, * p<.05, ** p<.01, *** p<.001

Survey-weighted estimates from a subsample of couples who met after 1995. Covariates not shown: Race, Age, Education Scale, Income, Political Party, Married, Coresident, Years Known, Number of Children in Household, Same Sex, Year Sampled, Relationship Duration, Still Together, Previously Married, Home Internet Access, Region, MSA\County Population and Median Income, and Zipcode Per Capita Income. Religion models include R's religion. For full model coefficients, see Appendix Table A3, available at: http://www.unm.edu/~reubenjthomas/OnlineExogamyAppendix.pdf

Table 8. Logistic and OLS Regression Coefficients Predicting Exogamy by Online Source (Since 1996)

	Different Race/ Ethnicity	Different College Degree Status	Different Mother's College Graduate Status	Absolute Age Difference	Different Religion	Different Political Party
Model	Logit	Logit	Logit	OLS	Logit	Logit
Source of Introduction (vs Purely Offline)						
Dating Site/App	0.35+	0.43*	0.17	-0.93*	0.71*	-0.02
Other Online	0.69*	0.30	-0.21	-0.05	0.68*	0.47*
Both Off- & Online	0.23	0.44+	-0.18	-0.49	-0.14	0.04
R^2				0.12		
McFarland's Psuedo-R^2	0.23	0.08	0.13		0.14	0.05
Sample Years	Both	Both	Both	Both	2009	Both
N	3025	3027	2985	3012	1334	3010

+ p<.1, * p<.05, ** p<.01, *** p<.001

Survey-weighted estimates from a subsample of couples who met after 1995. Covariates not shown: Race, Age, Education Scale, Income, Political Party, Married, Coresident, Years Known, Number of Children in Household, Same Sex, Year Sampled, Relationship Duration, Still Together, Previously Married, Home Internet Access, Region, MSA\County Population and Median Income, and Zipcode Per Capita Income, Racial\Ethnic and College Degree Diversity for both zip code and MSA, and Political Partisanness for the respondent's congressional district. Religion models include R's religion. For full model coefficients, see Appendix Table A2, available at: http://www.unm.edu/~reubenjthomas/OnlineExogamyAppendix.pdf