Internet Access and Partnership Formation in the United States

Maria Sironi
University College London
m.sironi@ucl.ac.uk

Ridhi Kashyap
University of Oxford and Nuffield College
ridhi.kashyap@nuffield.ox.ac.uk

Abstract

Unlike older communication technologies, the internet has broadened the scope for social interaction and enabled people to meet and interact with people outside their existing social network. This feature of the technology is perhaps most salient for its role in helping people search for mates. While the internet may enlarge the pool of prospective partners, access to a larger pool may also delay the transition to partnership as the option for alternatives may induce individuals to search longer. We empirically examine this effect of the internet on both heterosexual and homosexual partnership formation using nationally-representative data from the National Longitudinal Survey of Youth and the Current Population Survey from the US. We find that while the effect of the internet on the transition to partnership is negative at younger ages, the effect of the internet on increasing the propensity to partner becomes positive as individuals become older, for both homosexual and heterosexual partnerships.

Introduction

The rapid diffusion of the internet has been one of the most significant social phenomena of the new millennium. In the US, internet usage rates grew from 5 to 74 percent between 2000 and 2009 (Dettling 2007). Most recent estimates indicate that 86 percent of Americans now use the internet (Greenwood, Perrin, and Duggan 2016). The expansion of the technology has generated a seismic shift in how people access information, express their ideas, and connect with each other. In contrast to other communication technologies such as the telephone, which improved communication within existing networks, the internet has broadened the scope for social interaction by enabling new possibilities for finding and meeting people outside of one's existing social network.

Although it is still too early to fully assess the varied and wide-ranging social implications of the internet, a domain in which the potential of the internet to communicate more freely and reach outside of one's network has been particularly significant is in finding new romantic partners. The internet has been described as "the new social intermediary in the search for mates" (Rosenfeld and Thomas 2012). Drawing on a nationally representative survey of 4,000 adults in the US who were already in relationships, Rosenfeld and Thomas (2012) found that among those that met in 1994-98, 3.9% reported having met online for the first time. By 2004-06, this number had increased to 20%. While it is not a priori clear whether the proliferation of choice afforded by the internet is likely to result in more dating options only or the formation of more durable partnerships such as marriage, some have argued that the greater amount of information on prospective partners that the internet affords may facilitate better quality and more stable matches (Cacioppo et al. 2013, Hitsch et al. 2010). Internet dating platforms allow users to access a wider pool of partners, but also sort and search for those that meet user-defined criteria, which may arguably result in more efficient matching.

Despite the theoretical speculation about the impact of internet availability on increasing propensities to form a partnership or to marry, the empirical literature on the question has been

relatively limited. An exception here is Bellou (2015) who exploits variation in the timing of broadband diffusion at the county-level in the US to examine its impact on aggregate-level marriage rates. Her study found the effect of broadband diffusion on marriage to be positive. While Rosenfeld and Thomas found that an increasing fraction of couples were likely to meet online, Bellou's findings when combined with those of Rosenfeld and Thomas suggest that the internet may not just be displacing offline modes of meeting partners, but also generating new matches altogether that might not have otherwise occurred. In enabling new types of matches, the role of the internet is likely to be especially salient for couples who might have, in its absence, had limited opportunity to meet and interact. Rosenfeld and Thomas (2012) describe this in terms of those seeking matches in "thin markets", such as gays, lesbians and middle-aged heterosexuals.

Our study examines the effects of the internet on the transition to partnership, examining its effect both on heterosexual and homosexual partnerships. We seek to analyze this relationship across multiple, nationally-representative data sources, given (as outlined later) challenges involved in finding data sources that contain both information on partnership histories and internet access.[1] While the motivation of our paper is closely aligned with Bellou (2015), we use different data sources and analyze partnership formation rather than marriage only among different couple types. In contrast to Bellou who uses aggregate-level data to study the relationship between internet diffusion and marriage rates, we use individual-level data both on internet access and partnership histories, where possible, to analyze how gaining access to the internet changes the propensity to partner. An important contribution of our work compared with previous studies is that we examine how internet access affects the propensities to enter partnerships for both heterosexual and homosexual couples. Furthermore, by drawing on both cohort (longitudinal) and period (cross-sectional) datasets, we assess

[1] This extended abstract of the paper currently uses data from the National Longitudinal Study of Youth (NLSY97) and the Current Population Survey (CPS).

whether access to the internet affects partnership propensities for individuals in a particular cohort as they age, or whether it affects different cohorts ages over the period the internet diffuses.

This paper proceeds as follows: in the next section, we first derive hypotheses related to how we would expect the internet to affect both the propensity to partner and reflect on the age- or timing patterns of this effect. We then present the data and methods, highlighting some of the data limitations and challenges involved in conducting this type of study. We conclude with some results, which at present are preliminary as we aim to expand the analyses to other data sources.

Hypotheses

There are different ways in which the internet could affect partnership formation and the direction of this relationship is a priori not clear. Perhaps most directly, the internet through platforms such as internet dating websites is likely to provide people the chance to meet new prospective partners and draw on a wider network of individuals than those encountered in daily routines and interactions. While internet dating opportunities may reduce the time and search costs associated with finding prospective partners, the wider pool of partners provided by the internet may make the partner search process longer and imply a postponement in partnership formation. Indeed, individuals may find it hard to 'settle down' in the face of potentially unlimited possibilities to meet other, new romantic partners. On the other hand, the opportunity afforded by the internet to collect a lot of information and conduct a targeted search for prospective partners relatively quickly, could hasten the process of partnership formation and increase propensities to partner.

The effect of the internet on partnership formation is not only restricted to internet dating sites, however. Different digital platforms can facilitate the dating and partner search process. Online communication tools such as chat, email and social media enable frequent and fast interactions with others in a way that older communication technologies such as the landline telephone did not.

Communication is also more direct and personalized through the internet, without the need to encounter any intermediaries. While calling a prospective romantic interest in an era of landlines might have meant calling and having to first talk with their parents or family members, communication in a digital era means unrestricted, unmediated and immediate access to a person of interest. For minority communities, such as lesbian, gay, bisexual or transgender (LGBT) individuals who might face greater stigma or resistance towards their romantic interests, this effect is likely to be especially pronounced. For these communities, furthermore, access to online forums and communities may also act as a medium to both recognize and validate their desires. In this way, the effect of the internet on partnership formation for LGBT individuals, may be even stronger through both the information and search mechanisms.

It is plausible that the effects of the internet on facilitating the partner search may vary by age of the individual in their life course. Access to the internet at younger ages may enable individuals to tap into a wider pool of romantic partners, have ready availability and easy communication with prospective partners, and potentially expand their dating opportunities. This ready availability of dates may delay the transition to a more durable partnership. At younger ages, thus, this implies that internet access is likely to be negatively associated with the propensity to partner and a postponement of the partnership transition. As individuals grow older and enter a phase of life when they feel readier to settle down, access to the internet is likely to facilitate the likelihood of entering a partnership.

Empirical Analysis

Longitudinal analysis using NSLY97

Data and Methods

Data sources that include detailed information on internet access, partnership/marital history, and information on the sex of the partners or sexual orientation, within a longitudinal data structure that

enables us to track changes in internet access and partnership simultaneously are fairly limited. The first set of analyses of this work is conducted using the National Longitudinal Survey of Youth 1997 (NLSY97). This survey includes a representative sample of young adults in the United States, who are born between 1980 and 1984. They were interviewed for the first time in 1997 (when they were between 12 and 16), and then every year after that until 2011.[2] The NLSY97 collects data on socio-demographic characteristics, school and employment history, and partnership history. Very important for our analysis is that from 2003 the survey includes questions about internet access. In particular, from 2003 to 2011, the respondents are asked if they had access to the internet, and from 2003 to 2008 they were asked from where they could access it (e.g. home, school, work, library, etc.). While these measures are not perfect in being able to capture exactly how the internet is used by the survey respondents, it is able to approximate if individuals had any form of access. Thus, our key independent variable is based on the question "Do you currently have access to the internet?", and it is coded as 1 if the answer is "yes", and 0 otherwise.

Our main outcome of interest is partnership formation, and we are also interested in who the individuals partnered with to distinguish between heterosexual and homosexual relationships. In order to study whether internet access is associated with partnership formation, we use both the residential household roster, and the nonresidential roster. During the interview, the respondent is asked to answer questions about household members (co-residential household roster), and also questions about nonresident relatives (nonresident roster). The respondents identify the relationship with the household members, and we categorize them as "in a partnership" if they name a wife/husband or a lover/partner as a residential or nonresidential member of the household. Additionally, the respondents are asked the sex of these household members/relatives. Using this information and the

[2] There is a more recent Round, collected in 2013. However, since there is a gap of a year between the last two rounds, we decided to stop the analysis in 2011. Moreover, in 2013, the question about access to internet is replaced by the question on frequency of internet use.

sex of the respondent we are able to identify whether a partnership is heterosexual or homosexual. Our dependent variable is a categorical variable equal to 0 if the respondent is not in a partnership, 1 if he/she is in a heterosexual relationship, and equal to 2 if in a homosexual relationship. Since these questions are included in the survey every year, the partnership status (i.e. in a partnership or not) and the type of partnership (i.e. heterosexual or homosexual) can vary over time.

Our sample includes individuals who are interviewed every year from 2003 to 2011, given that we need data on internet access (only available from 2003 onwards), and we also need information of partnership formation and dissolution over time. Hence, we have a sample of 5,729 individuals (from 8,984 in 1997). Finally, we include several control variables in our analysis in order to take into account sociodemographic characteristics that can influence both the 'risk' of partnership formation and the probability of having access to internet: Other than age (and age^2) and gender, we include *race* (White, Black, Hispanic, and Other), *region of residence* (Northeast, North Central, South, and West), if living in a *rural* or *urban* area, the *years of education* and if *enrolled in school*, *parental level of education* (less than high school, high school, or more than high school), and *family income in 1997*. The final sample size – excluding those who do not have information for the control variables – is of 5,513 respondents, of which 52.6% are women and 47.4% are men.

After presenting some descriptive statistics on the sample used in the analysis, we implement discrete-time event history analysis regression models with competing risks (Allison, 1982) to study the association between Internet access and being either in an opposite-sex or same-sex relationship (i.e. the competing events). Our models are multilevel models, in which partnership episodes are nested within individuals (Barber et al., 2000). Multilevel event history models allow us to introduce random effects, which represent individual-specific unobservables. We follow individuals in the sample over time, and current events give a two-level hierarchical structure: episodes – i.e. partnership formation – are clustered into individuals.

Descriptive Findings

Respondents in our sample were born between 1980 and 1984. Therefore from 2003 to 2011 (our time period of interest) they were between 19-23 and 27-31. This is the age span in which most people enter significant relationships, and possibly get married. As we can see in Table 1, the proportion of respondents in a relationship increases substantially over time, from 28% in 2003 to 71.6% in 2011. This increase can be observed for both heterosexual and homosexual partnerships, which showed a 61% and 57.1% positive change over 9 years respectively. The number of respondents in homosexual relationships is considerably lower (ranging between 36 and 84 people over time) than those in heterosexual relationships.

Table 1. Partnership Status (NLSY97)

	In a heterosexual partnership		In a homosexual partnership		Total in a partnership	
Year	N	%	N	%	N	%
2003	1,508	27.4	36	0.7	1,544	28.0
2004	1,957	35.5	39	0.7	1,996	36.2
2005	2,342	42.5	48	0.9	2,390	43.4
2006	2,705	49.1	60	1.1	2,765	50.2
2007	3,037	55.1	68	1.2	3,105	56.3
2008	3,311	60.1	76	1.4	3,387	61.4
2009	3,519	63.8	74	1.3	3,593	65.2
2010	3,732	67.7	77	1.4	3,809	69.1
2011	3,866	70.1	84	1.5	3,950	71.6

N=5,513

In the same time span, there has been an increase in the percentage of young adults in the sample having access to internet. Table 2 shows that in 2003 80.7% of the sample had access to internet, and in 2011 this figure increased to 87.7%. The same growth can be seen in the proportion of those having internet access at home: from 61.5% in 2003 to 71.7% in 2011.

In Table 3 we report descriptive statistics for the other variables used in the analysis. Half of the sample is white, while 27% is Black, and 20% is Hispanic. Respondents' parents have on average 13

years of education, and the family income was $41,079 in 1997 when the survey started and the respondents were between 13 and 17 years old. Looking at time varying covariates, the mean age increases from 21 to 29. The geographical distribution of the respondents and their urban/rural location remains quite stable over time. As expected the proportion of people in school decreases, from 38.1% in 2003 to 12.2% in 2011, and the average number of years of education goes up from 12.6 to 14.1.

Table 2. Internet Access (NLSY97)

	Access to Internet		Access to Internet at Home	
Year	N	%	N	%
2003	4,450	80.7	3,388	61.5
2004	4,409	80.0	3,352	60.8
2005	4,507	81.8	3,475	63.0
2006	4,640	84.2	3,625	65.8
2007	4,745	86.1	3,806	69.0
2008	4,547	82.5	3,952	71.7
2009	4,593	83.3		
2010	4,651	84.4		
2011	4,804	87.1		

N=5,513

Given that our main interest is the relationship between having access to the internet and partnership formation, we look at the correlation between these two variables over time (Figure 1). The proportion of respondents in partnerships is lower among those who do have access to internet until 2006 (age 22-26). In 2007 and 2008 the two groups almost overlap, and from 2009 onwards having access to internet is associated with a higher proportion of people being in a partnership. Hence, internet access seems to become more important as respondents get older. The figure also shows that internet access might work in the same direction as higher education and socioeconomic status in terms of postponement of partnership formation, followed by recuperation.

Table 3. Control Variables (NLSY97)

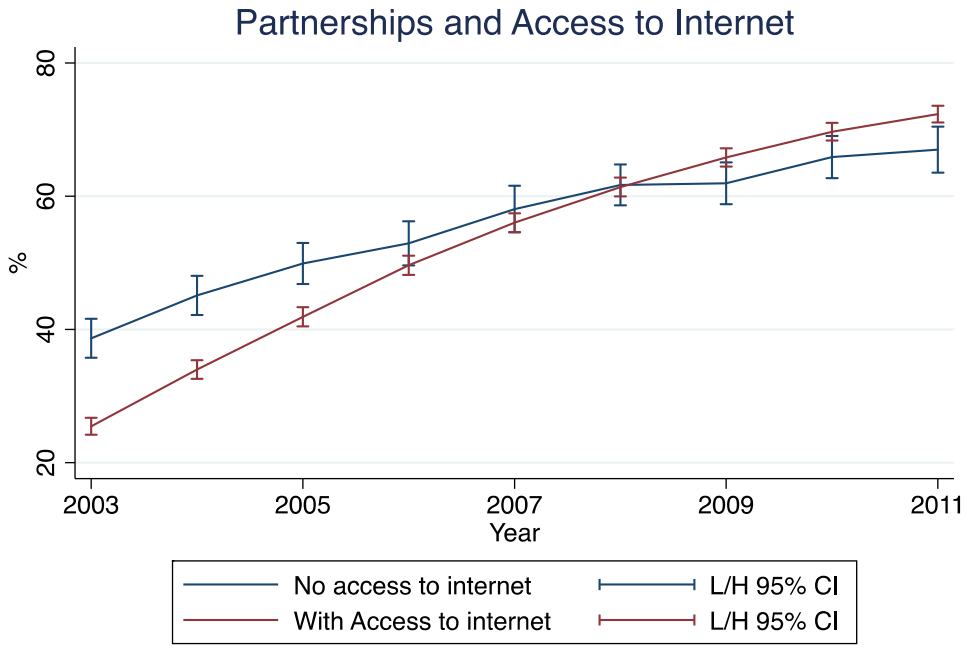
Time Constant									
% Female	52.6								
Race (%)									
White	49.5								
Black	27.0								
Hispanic	20.0								
Other	3.56								
Parents' Education, Avg. (years)	13.2								
Family Income 1997, Avg. USD	41,079								
Year	**2003**	**2004**	**2005**	**2006**	**2007**	**2008**	**2009**	**2010**	**2011**
Mean Age	20.9	21.9	22.9	23.9	24.8	25.8	26.7	27.8	28.7
Region (%)									
North East	16.2	16.0	15.6	15.5	15.3	15.3	15.2	15.5	15.4
North Central	23.4	23.2	23.0	22.7	22.3	22.1	21.9	21.8	21.8
South	38.5	38.9	39.3	39.2	39.7	39.7	40.1	40.1	39.9
West	21.9	22.0	22.1	22.6	22.6	22.8	22.7	22.7	22.9
% Urban	78.8	80.6	81.3	80.9	80.9	81.3	80.9	79.0	78.9
Years of Education, Avg.	12.6	13.0	13.3	13.5	13.7	13.8	13.9	14.0	14.1
% in School	38.1	31.1	25.4	20.3	16.5	15.0	14.5	13.7	12.2

N=5,513

Multilevel Multinomial Logistic Regression Analysis

The plots in Figure 1 do not take into account any possible confounders, and so the fact that the association between internet access and partnership formation can be driven by other individual characteristics. With the greater diffusion of the technology, we may expect users with different socioeconomic characteristics (e.g. less educated or lower socioeconomic status) to adopt its use, and their different characteristics with regards to partnering behavior may result in the patterns we observe in Figure 1.

Figure 1. Partnerships and Access to Internet (NLSY97)



The multinomial logistic regression models that control for potential confounders are presented in Table 4. Our reference group is 'not in a partnership', and the two outcome groups are 'in a heterosexual relationship' and 'in a homosexual relationship'. Model (1) includes internet access, age and age^2, gender, race, region of residence, and urban vs rural location. In Model (2) we include some socioeconomic characteristics, i.e. years of education, parental level of education, and the 1997 family income quintile. Finally, in Model (3) we include the interaction between internet access and age, given that we hypothesize the role of internet access to become more relevant as the respondents get older.

The results show that there is a negative association between internet access and entering a partnership, both for heterosexual (OR=0.632, $p<0.01$) and homosexual (OR=0.698, $p<0.05$) partnerships. The probability of being in a relationship increases with age, and it is much higher for women than for men. Once we include the socioeconomic characteristics in Model (2), we see how the association between access to internet and being in a partnership becomes weaker among those in

a heterosexual partnership (OR=0.768, p<0.01), and non-significant among those transitioning to a homosexual partnership. Being enrolled in school and years of education are also negatively associated with being in a relationship, showing a 'postponement effect' due to the level of education. Finally, in Model (3), including the interaction term between age and access to internet, we see how the main effect of internet access is still negative and that of age still positive. Most importantly, the interaction term is greater than 1 and significant for both heterosexual and homosexual partnerships, 1.302 (p<0.01) and 1.232 (p<0.01) respectively. This shows how the association between internet access and partnership formation is negative at younger ages, but it becomes positive as age increases (individuals in our sample are possibly too young to observe this change in the association).

Table 4. Multinomial multilevel regression models (NLSY97)

Y = being in a partnership (Ref: No)	Heterosexual Partnership			Homosexual Partnership		
	(1)	(2)	(3)	(1)	(2)	(3)
Internet Access	**0.632*** **	**0.768*** **	**0.149*** **	**0.698** **	**0.831**	**0.240*** **
	(0.058)	**(0.065)**	**(0.029)**	**(0.102)**	**(0.123)**	**(0.089)**
Age	**7.855*** **	**7.980*** **	**6.767*** **	**7.262*** **	**7.593*** **	**6.673*** **
	(0.421)	**(0.457)**	**(0.383)**	**(0.660)**	**(0.713)**	**(0.653)**
Age^2	**0.946*** **	**0.946*** **	**0.942*** **	**0.947*** **	**0.946*** **	**0.944*** **
	(0.003)	**(0.003)**	**(0.003)**	**(0.006)**	**(0.006)**	**(0.006)**
Female	3.679***	10.913***	10.598***	3.039***	9.018***	8.738***
	(0.581)	(1.450)	(1.260)	(0.543)	(1.421)	(1.276)
Race (Ref: White)						
Black	0.032***	0.026***	0.029***	0.039***	0.032***	0.036***
	(0.007)	(0.004)	(0.006)	(0.009)	(0.006)	(0.008)
Hispanic	0.654**	0.469***	0.572***	0.692*	0.450***	0.550**
	(0.123)	(0.087)	(0.121)	(0.151)	(0.100)	(0.135)
Other	0.020***	0.037***	0.039***	0.013***	0.022***	0.024***
	(0.011)	(0.011)	(0.011)	(0.009)	(0.010)	(0.010)
Region (Ref: Northeast)						
North Central	6.136***	3.066***	2.764***	2.473***	1.218	1.098
	(1.497)	(0.517)	(0.506)	(0.681)	(0.258)	(0.246)
South	4.495***	3.922***	3.493***	2.143***	1.845***	1.642***
	(0.795)	(0.605)	(0.535)	(0.444)	(0.348)	(0.308)
West	3.693***	3.489***	3.066***	1.656**	1.489*	1.304
	(0.702)	(0.608)	(0.554)	(0.378)	(0.321)	(0.288)

Urban Area	1.244**	1.378***	1.373***	2.127***	2.288***	2.282***
	(0.114)	(0.116)	(0.118)	(0.330)	(0.347)	(0.348)
Enrolled In School		0.283***	0.287***		0.410***	0.411***
		(0.022)	(0.023)		(0.057)	(0.058)
Years of Education		0.769***	0.766***		0.722***	0.721***
		(0.020)	(0.018)		(0.024)	(0.023)
Parents Education (Ref: < High School)						
High School Diploma		0.828	0.888		0.448***	0.479***
		-0.141	(0.174)		(0.096)	(0.113)
More than High School		0.365***	0.377***		0.362***	0.373***
		-0.064	(0.075)		(0.078)	(0.088)
Family Income in 1997 (Ref: 1st Quintile)						
2nd Quintile		0.869	0.855		1.891***	1.864***
		(0.166)	(0.157)		(0.468)	(0.449)
3rd Quintile		0.492***	0.398***		0.535**	0.432***
		(0.092)	(0.090)		(0.143)	(0.127)
4th Quintile		1.291	1.287		1.362	1.361
		(0.273)	(0.260)		(0.389)	(0.379)
5th Quintile		0.067***	0.070***		0.134***	0.138***
		(0.014)	(0.014)		(0.036)	(0.037)
Income missing		0.312***	0.229***		0.587**	0.432***
		(0.058)	(0.040)		(0.141)	(0.099)
Internet Access*Age			**1.302*** **			**1.232*** **
			(0.037)			**(0.062)**
Constant	0.000***	0.002***	0.006***	0.000***	0.000***	0.000***
	(0.000)	(0.001)	(0.003)	(0.000)	(0.000)	(0.000)
N	49,617 (5,513 individuals over 9 years)					

* p<0.10, ** p<0.05, *** p<0.01

Cross-sectional Analysis Using CPS Data

Data and Methods

The second set of analyses is performed using the Current Population Survey data. The CPS is a household survey that collects monthly data in the US on several different topics, including demographic and socioeconomic information. Since 1997 the CPS started collecting data on computer use and internet access in a 'Computer and Internet Use' supplement, and this information is available for the years 1997, 1998, 2000, 2001, 2003, 2007, 2009, 2010, 2011, 2012, 2013, and 2015. In this

supplement, all the respondents in the household are asked whether they have access to internet at home and if the household has an internet connection. For the purpose of our analysis, we use this information to build our independent variable, i.e. having access to internet at home. This variable is equal to 1 if the household has an internet connection or – when the answer to that question is missing – if a member of the household has access to internet from home.

The partnership status of the main respondent is established using the household roster, and the presence of a spouse or an unmarried partner in the same household. In the same way as for the NLSY data, we can distinguish between heterosexual and homosexual unions based on the sex of the household head and of the partner (if present). Therefore, our dependent variable using the CPS data is the same as for the NLSY analysis, and it is equal to 0 if the respondent is not in a partnership, 1 if he/she is in a heterosexual relationship, and equal to 2 if in a homosexual relationship.

The sample includes individuals that are interviewed in the 'Computer and Internet use' supplement and that have answered the questions on internet access. We start with a sample of 952,892 individuals, 15 years old and older, over the 12 years, and we are left with 619,158 of them after we exclude those who do not report data on internet access. We included several control variables, trying to include the same variables used in the previous set of analyses: *age* at interview (and age^2), *gender*, *ethnicity* (White, Black, Hispanic, Asian, American Indian, and Other/Mixed), *state*, if *living in a metro area*, the *level of education* (less than high school, high school, some college, college degree or more), and *family income* (< $25,000, $25,000-49,999, $50,000-74,999, $75,000 and over).

We present some descriptive statistics on the partnership status of the respondent, access to internet and control variables over the years in our selected sample, and then we run multinomial logistic regressions to investigate the relationship between internet access and partnership status, distinguishing between heterosexual and homosexual unions.

Descriptive Findings

Table 5 reports the proportion of individuals in the CPS data that are in a coresidential union and distinguishes between heterosexual and homosexual partnerships. This proportion declines slightly over time for heterosexual unions, from 56.4% in 1997 to 53.2% in 2015, but increases for homosexual unions, from 45 couples in 1997 (0.09% of the total sample) to 313 couples in 2015 (0.63%).

In Table 6, we can see the proportion of households with an internet connection. The pattern shows an increase in the prevalence of internet access over time, since the percentage increases from 18.1% to 73.4% between 1997 and 2015, with a very sharp increase between 1997 and 2009 and a slower growth between 2010 and 2015.

Table 5. Partnership Status

	In a heterosexual partnership		In a homosexual partnership		Total in a partnership	
Year	N	Weighted %	N	Weighted %	N	Weighted %
1997	27,251	56.4	45	0.09	27,296	56.5
1998	27,229	56.2	50	0.10	27,279	56.3
2000	27,093	56.5	53	0.10	27,146	56.6
2001	31,908	55.8	96	0.18	32,004	55.9
2003	31,574	55.9	109	0.19	31,683	56.1
2007	29,786	54.4	148	0.26	29,934	54.6
2009	29,789	54.0	166	0.30	29,955	54.3
2010	29,635	54.0	206	0.34	29,841	54.3
2011	29,264	54.3	245	0.44	29,509	54.7
2012	29,156	53.8	239	0.41	29,395	54.2
2013	21,632	53.6	194	0.45	21,826	54.0
2015	28,106	53.2	313	0.63	28,419	53.8
Total	342,423	54.8	1,864	0.30	344,287	55.1

N=619,158

Table 6. Internet Access

	Access to Internet at Home	
Year	N	Weighted %
1997	8,620	18.1
1998	12,709	26.4
2000	19,924	41.6
2001	28,864	50.5
2003	30,911	54.8
2007	33,930	61.9
2009	37,539	68.7
2010	38,804	71.1
2011	37,898	70.5
2012	40,462	74.8
2013	29,933	74.2
2015	38,462	73.4
Total	358,056	57.9

N=619,158

Table 7 includes descriptive statistics for the control variables in our analysis. The racial composition of the sample becomes more heterogeneous over time, with 75.3% the respondents being white, 12.1% Blacks and 9.0% Hispanics in 1997, and 66.9%, 12.7% and 13.6% in 2015, respectively. The geographical distribution of the respondents across regions remains rather stable over time, while the proportion not living in a metro area decreases and the average level of education increases over the years.

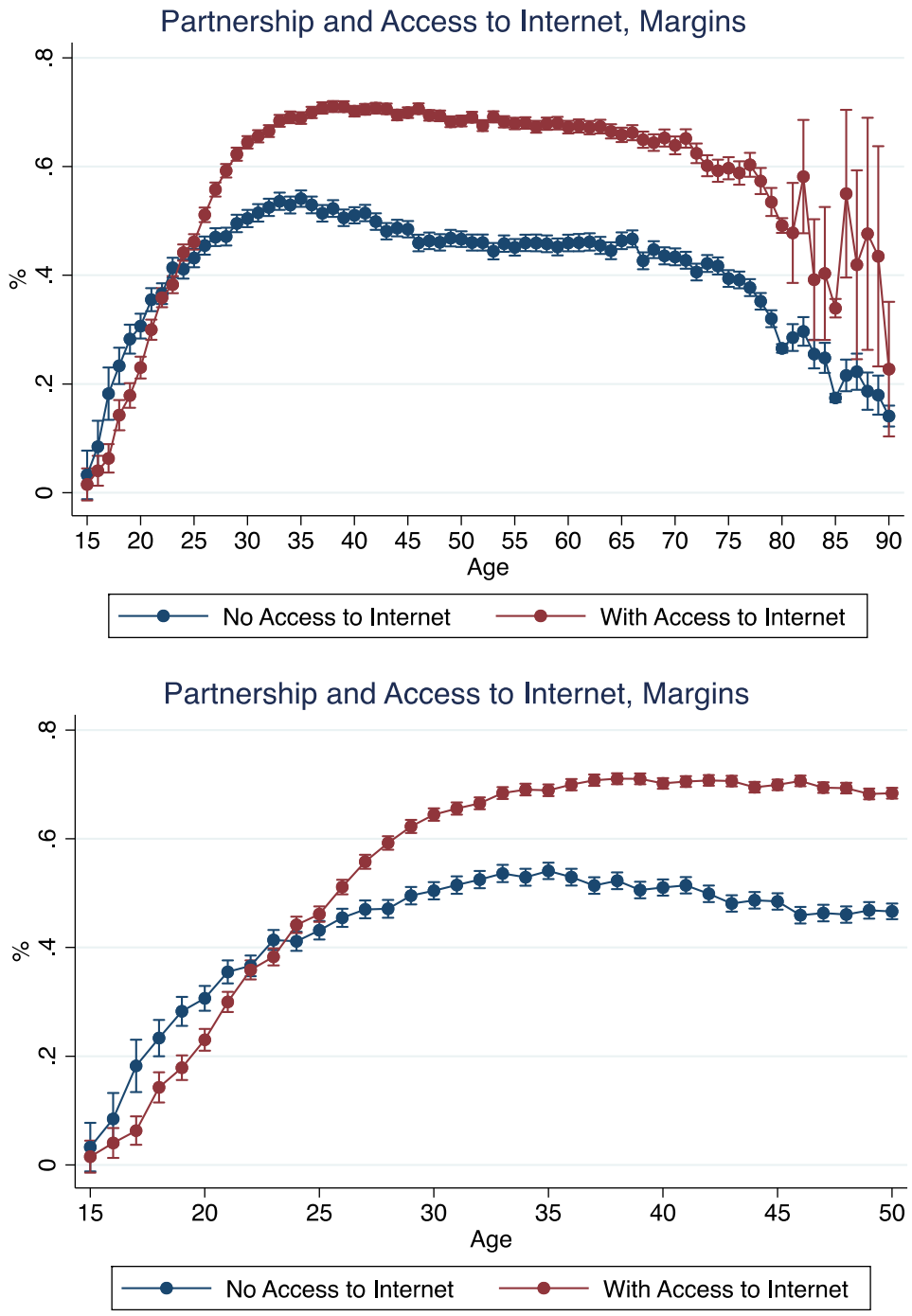
Table 7. Control Variables (Weighted)

Year	1997	1998	2000	2001	2003	2007	2009	2010	2011	2012	2013	2015	Total
% Female	40.0	41.8	44.5	46.2	46.7	48.7	48.6	49.4	49.2	49.6	49.6	49.3	47.1
Mean Age	48.2	48.4	48.7	48.7	48.4	49.0	49.6	49.8	50.0	50.3	50.5	50.8	49.4
Race (%)													
White	75.3	75.1	74.4	73.8	72.1	70.3	69.7	69.6	69.5	68.2	68.0	66.9	71.0
Black	12.1	12.1	11.9	12.0	11.8	12.3	12.5	12.4	12.4	12.4	12.4	12.7	12.3
Hispanic	9.0	9.0	9.3	9.7	10.8	11.9	12.0	12.2	12.3	13.1	13.0	13.6	11.4
Asian	2.8	-	-	-	3.5	3.8	3.9	3.9	4.0	4.3	4.5	4.6	3.0
American Indian	0.7	0.8	0.8	0.8	0.5	0.5	0.7	0.6	0.6	0.7	0.7	0.6	0.7
Other/Mixed	-	3.1	3.6	3.7	1.2	1.3	1.2	1.3	1.3	1.4	1.5	1.6	1.7
Region (%)													
North East	19.4	19.0	19.3	19.4	19.1	18.3	18.3	18.2	18.0	17.8	17.8	17.6	18.5
North Central	23.4	23.6	23.1	23.5	23.1	22.7	22.3	22.3	22.3	22.3	22.0	21.8	22.7
South	35.7	36.0	36.2	35.9	36.0	36.7	37.1	37.2	37.5	37.7	37.8	38.0	36.8
West	21.5	21.5	21.4	21.3	21.8	22.4	22.4	22.3	22.2	22.3	22.5	22.6	22.0
% Not in metro area	19.5	19.4	19.4	19.1	18.8	15.9	15.8	16.0	15.9	15.5	15.6	14.1	17.0
Level of Education (%)													
Less than High School	17.6	17.4	16.6	16.0	15.0	12.8	12.5	12.0	11.8	11.4	11.1	10.4	13.6
High School	31.7	31.3	30.7	30.7	30.4	29.9	29.3	29.2	29.3	28.5	28.4	27.4	29.7
Some College	26.0	25.9	26.4	26.9	26.8	28.1	28.2	28.5	28.5	28.8	28.6	29.2	27.7
College Degree or more	24.7	25.4	26.4	26.4	27.8	29.2	30.0	30.3	30.4	31.4	31.9	33.0	29.0
Family Income (%)													
< $25,000	34.9	31.2	27.1	26.1	24.4	20.9	22.2	28.9	28.4	27.9	27.2	24.7	26.9
$25,000-49,999	27.4	27.0	25.3	24.5	24.0	21.6	21.8	27.5	27.1	26.7	26.0	25.8	25.4
$50,000-74,999	14.2	14.6	14.7	14.9	14.6	14.6	14.7	17.8	18.1	17.9	18.0	18.3	16.1
$75,000 and over	11.7	13.5	16.4	17.5	17.6	20.4	21.0	25.9	26.4	27.5	28.8	31.3	21.7
Missing	11.9	13.6	16.5	17.0	19.4	22.5	20.3	-	-	-	-	-	10.0

N=619,158

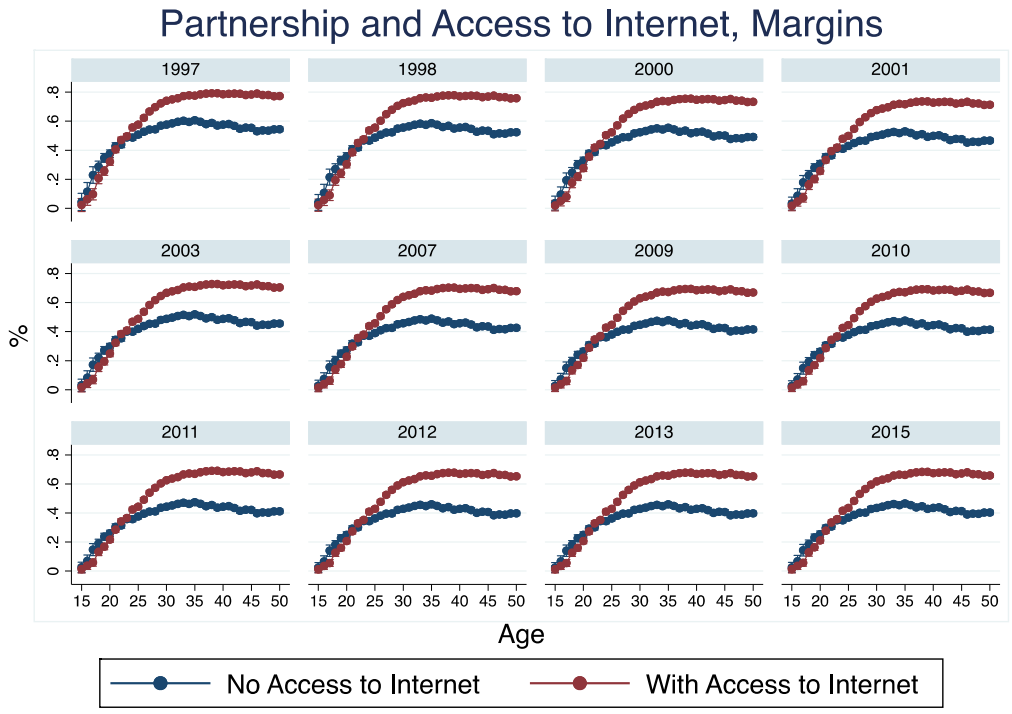
Figure 2 replicates Figure 1 using the CPS data, and shows the relationship between access to internet and partnership status over age on the whole sample and on those who are between 15 and 50 years old.

Figure 2. Partnership and Access to Internet (CPS)



We observe how the predicted probability of being in a coresidential union is lower for those who have internet access at home at younger ages (until age 22-24), but it becomes higher after age 25 and it remains higher. Figures 3 reports the same association across different years, and it shows how the association between internet access and partnership formation is consistent over time.

Figure 3. Partnership and Access to Internet, by year (CPS)



Multinomial Logistic Regression Analysis

As for the analysis using the NLSY97, we run some multinomial logistic regression models in order to consider confounders that could influence the association between internet access and partnership status, and to distinguish between heterosexual and homosexual unions. Table 8 reports the results of three different specifications: Model (1) includes internet access, age and age^2, gender, ethnicity, if

living in a metro area, state and year dummies; Model (2) adds level of education and family income; Model (3) includes the interaction term between age and internet access.

Table 8. Multinomial multilevel regression models

Y = being in a partnership (Ref: No)	Heterosexual Partnership			Homosexual Partnership		
	(1)	(2)	(3)	(1)	(2)	(3)
Internet Access	**2.488*** (0.016)**	**1.826*** (0.013)**	**1.234*** (0.024)**	**3.661*** (0.262)**	**2.383*** (0.180)**	**0.500*** (0.110)**
Age	**1.106*** (0.001)**	**1.066*** (0.001)**	**1.057*** (0.001)**	**1.172*** (0.012)**	**1.129*** (0.012)**	**1.093*** (0.012)**
Age^2	**0.999*** (0.000)**	**0.999*** (0.000)**	**0.999*** (0.000)**	**0.998*** (0.000)**	**0.999*** (0.000)**	**0.999*** (0.000)**
Female	0.375*** (0.002)	0.406*** (0.002)	0.408*** (0.002)	0.783*** (0.037)	0.817*** (0.039)	0.821*** (0.039)
Race (Ref: White)						
Black	0.512*** (0.005)	0.608*** (0.007)	0.606*** (0.006)	0.236*** (0.027)	0.317*** (0.036)	0.313*** (0.035)
Hispanic	1.337*** (0.014)	1.600*** (0.018)	1.578*** (0.018)	0.627*** (0.058)	0.896 (0.086)	0.863 (0.083)
Asian	1.293*** (0.024)	1.405*** (0.027)	1.418*** (0.027)	0.385*** (0.065)	0.396*** (0.067)	0.402*** (0.068)
American Indian	0.729*** (0.021)	0.874*** (0.026)	0.867*** (0.026)	0.611 (0.188)	0.785 (0.242)	0.774 (0.238)
Other/Mixed	0.929*** (0.020)	1.067*** (0.024)	1.071*** (0.024)	0.692** (0.125)	0.805 (0.146)	0.808 (0.146)
Metro Area (Ref: Not in Metro Area)						
Central city	0.566*** (0.005)	0.512*** (0.005)	0.513*** (0.005)	1.866*** (0.164)	1.490*** (0.133)	1.501*** (0.133)
Outside central city	0.952*** (0.008)	0.783*** (0.007)	0.783*** (0.007)	1.274*** (0.112)	0.986 (0.088)	0.987 (0.088)
Missing/Unknown	0.868*** (0.008)	0.805*** (0.008)	0.805*** (0.008)	1.243** (0.109)	1.102 (0.097)	1.104 (0.097)
Level of Education (Ref: < High School)						
High School Diploma		0.891*** (0.009)	0.880*** (0.009)		0.954 (0.121)	0.929 (0.118)
Some College		0.678*** (0.007)	0.671*** (0.007)		1.143 (0.142)	1.115 (0.139)
College Degree or more		0.531*** (0.006)	0.525*** (0.006)		1.383*** (0.173)	1.353** (0.169)
Family Income in 1997 (Ref: < $25,000)						
$25,000-49,999		2.369*** (0.019)	2.362*** (0.019)		1.464*** (0.121)	1.456*** (0.120)
$50,000-74,999		4.433*** (0.044)	4.430*** (0.044)		2.239*** (0.195)	2.231*** (0.194)

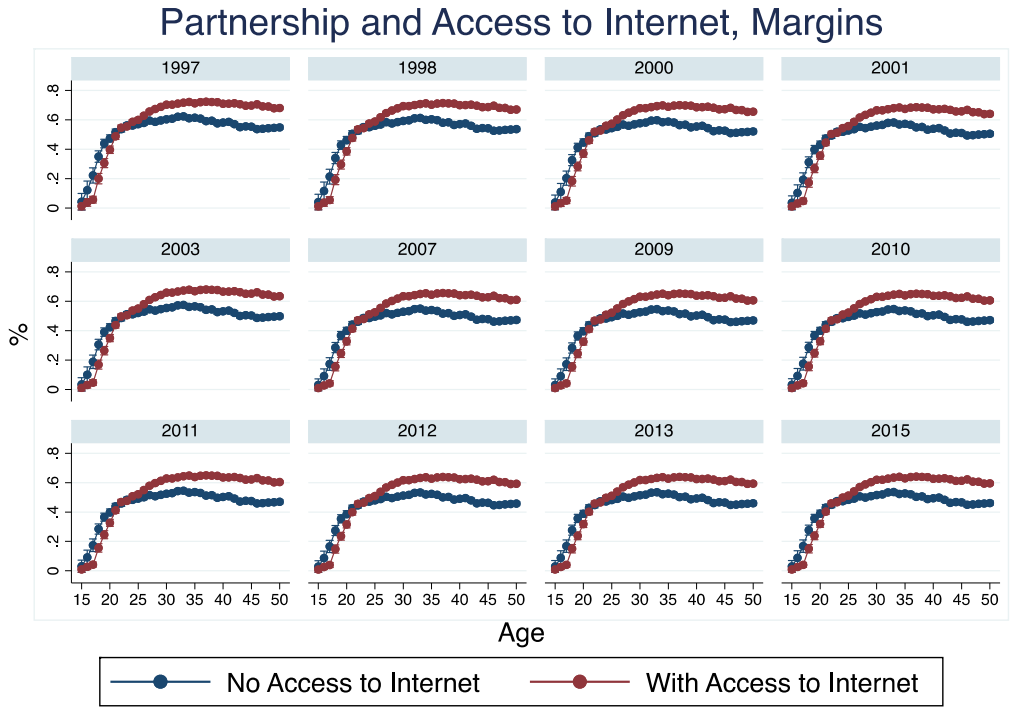
$75,000 and over		9.240***	9.253***		4.878***	4.844***
		(0.099)	(0.099)		(0.399)	(0.397)
Missing		2.339***	2.345***		0.978	0.985
		(0.025)	(0.026)		(0.141)	(0.142)
Internet Access*Age			**1.008*** **			**1.035*** **
			(0.000)			**(0.005)**
Constant	0.252***	0.416***	0.566***	0.000***	0.000***	0.001***
	(0.009)	(0.016)	(0.023)	(0.000)	(0.000)	(0.000)
N	619,158					

* p<0.10, ** p<0.05, *** p<0.01. All the specifications include dummies for State and Year.

The results are very similar to what we find using the NLSY97, except that the odds ratios for internet access are greater than one for both heterosexual and homosexual unions, given the higher proportion of people who are 25 years old and older. The results remain similar also when we include education and family income as control variables, with slightly smaller odds ratios. When we include the interaction between age and internet access we observe that the odds ratio becomes smaller for heterosexual unions (OR=1.234, p<0.01) but still greater than one, while it becomes lower than one for homosexual unions (OR=0.500, p<0.01). The interaction term is greater than one and significant for both heterosexual and homosexual partnerships (OR= 1.008, p<0.01 and OR=1.035, p<0.01).

The multinomial logistic regression models confirm that there is an interaction between age and internet access, and that the predicted probability of being in a union is higher for those who do have access to the internet after a specific age (after 25 on average, but it changes slightly by year). To have a better visual picture of this interaction, and make sure that it holds for every year, in Figure 4 we replicate Figure 3 including the same covariates included in the multinomial logistic regression models (and capping the age at 50).

Figure 4. Partnership and Access to Internet, by year – with confounders (CPS)



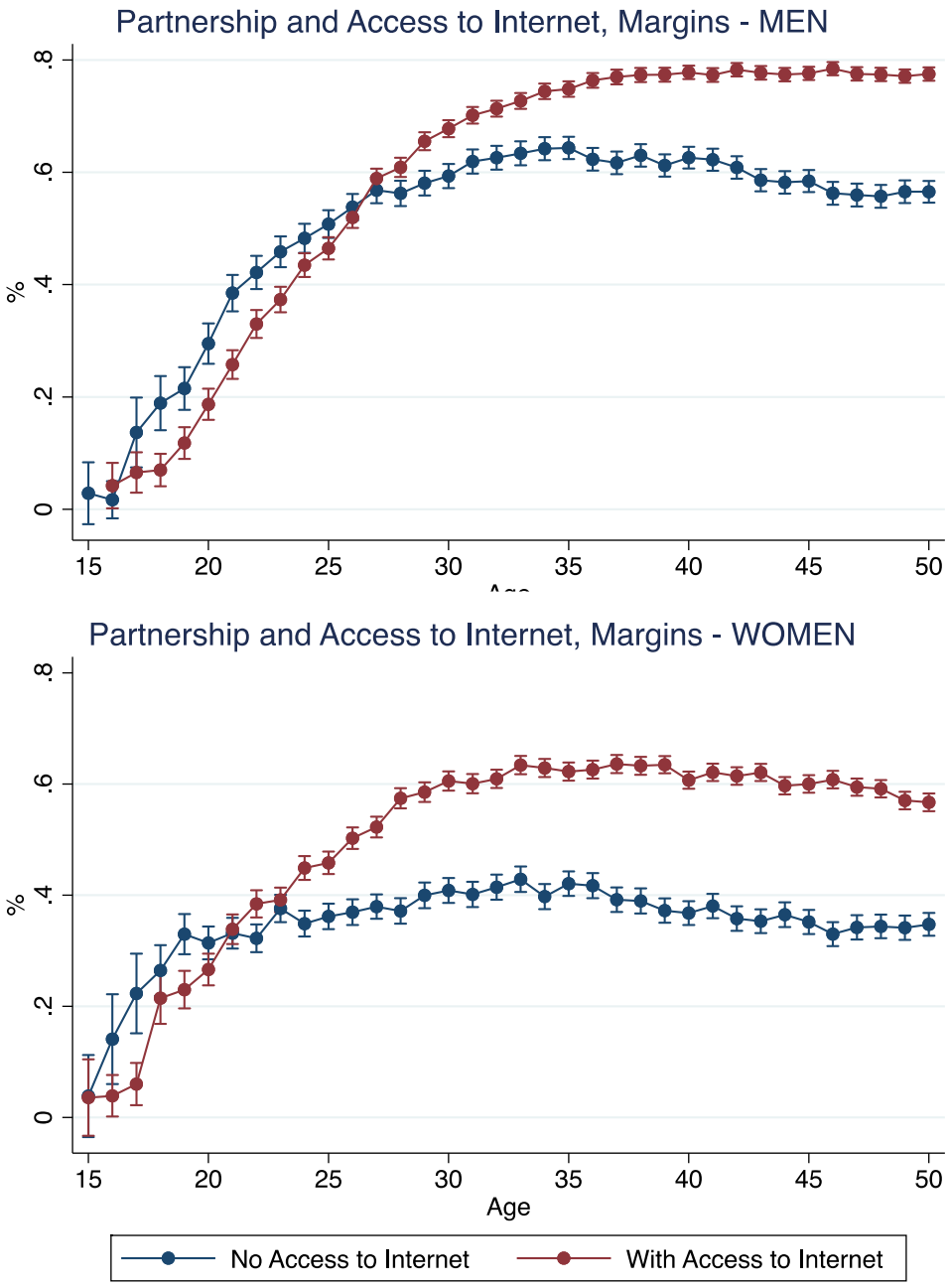
Robustness Checks

We perform three additional analyses, as robustness checks, to make sure that the results picked up so far are not driven by specific subgroups of the population in our sample or by the type of variables we selected and included in the analysis.

The first robustness check refers to gender, and we replicated the same analysis reported in Figure 2 separately for men and women. As Figure 5 shows, the general pattern is very similar across gender. Two things need to be noted: First, the predicted probability of being in a union is higher among men than among women at all ages, independently on internet access; Second, the critical age at which internet access is associated with a higher probability of being in a union is higher for men than for women (25-28 among men and 20-23 among women), and the association with internet access after that age is also stronger for women. This result seems to be in line with the fact that women enter

unions earlier than men, and possibly that they are ready to settle down earlier, using the internet to help find their long-term partner.

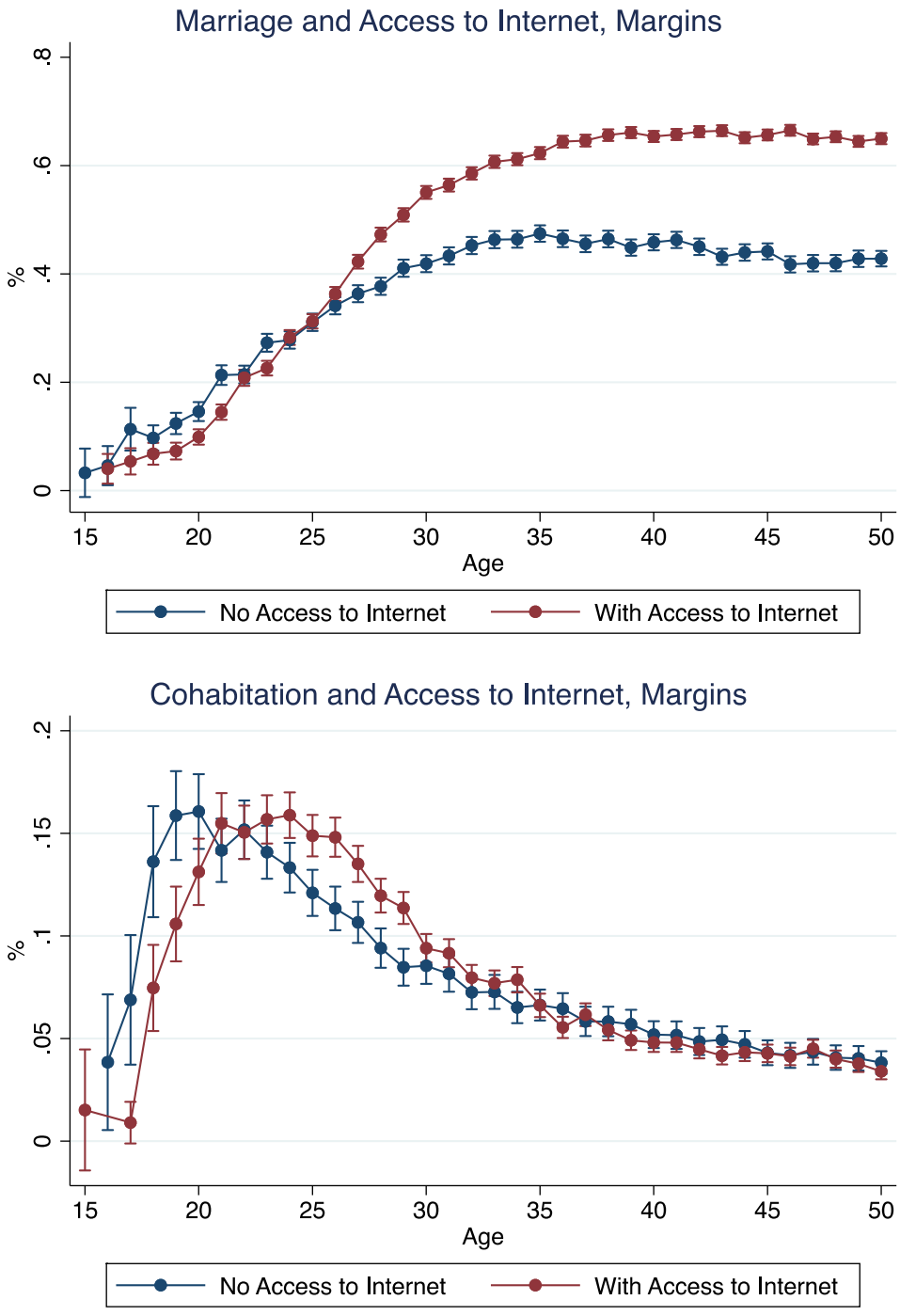
Figure 5. Partnership and Access to Internet, by gender (CPS)



The second robustness check replicates the existing analysis separately for different types of coresidential unions, i.e. cohabitation and marriage. We need to note that the proportion of

respondents who are cohabiting (4.8% of the whole sample) compared to those who are married (50.8% of the whole sample) is much lower, especially for older birth cohorts. The results are reported in Figure 6.

Figure 6. Partnership and Access to Internet, by type of Partnership (CPS)

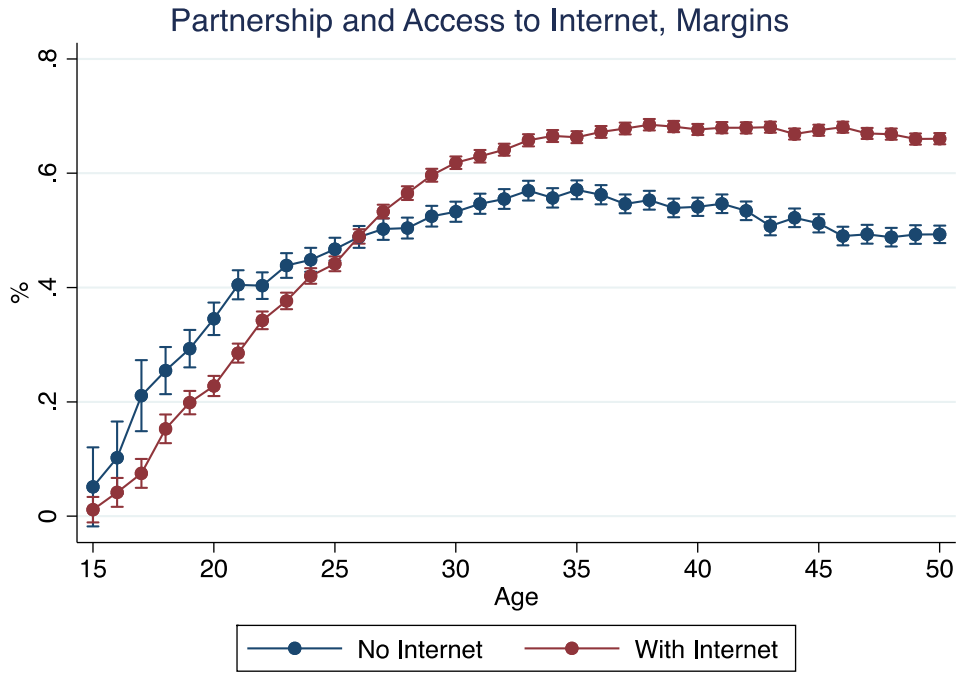


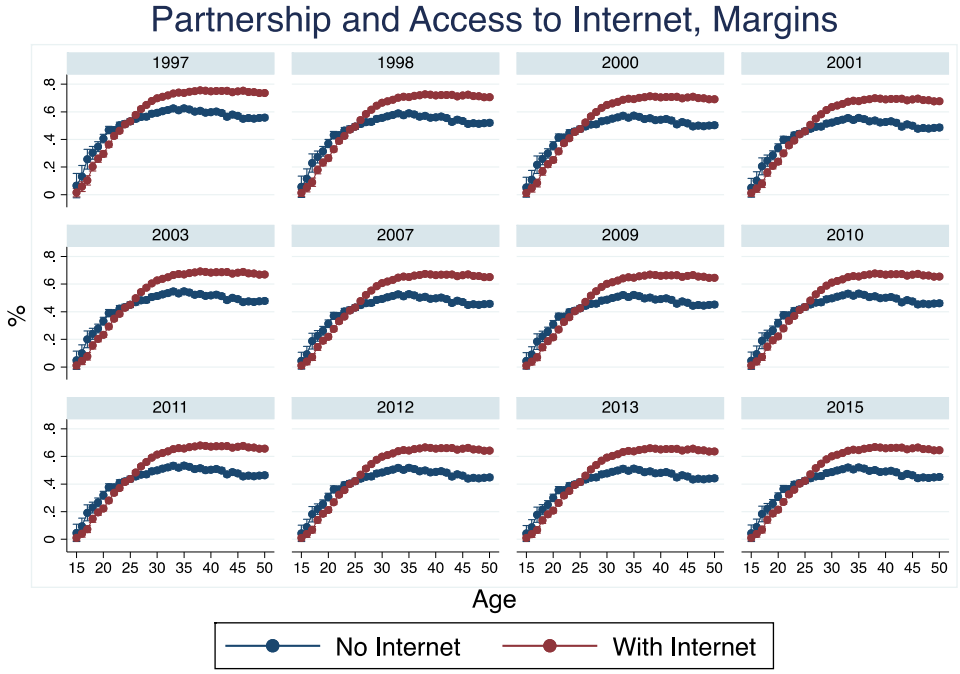
As expected, the picture for marriage resembles very closely the general results reported in Figure 2, given that 91.4% of the coresidential partnerships recorded in our sample are marriages. The graph for cohabitation shows that the predicted probability of being in a cohabitation increases until age 23-24 and then consistently declines at older ages, predominantly because people move into marriages. However, we still observe an interaction between age and internet access: the predicted probability of being in a cohabitation is higher for those with no internet access at younger ages (up until age 19) and then it becomes higher for those who do have access to the internet. This is true until age 29 and after that the difference between the two lines is not statistically significant.

The final robustness check has to do with the way in which respondents access the internet. So far, all the analyses have been carried out using internet access from home (either they answer that they have access to internet at home or they answer that the household has an internet connection). As already mentioned in the theoretical background, people might access the internet outside the household and – especially for more recent years – they can have internet connections through their smartphone. So it is possible that by considering only internet access at home we are not picking up those who still use the internet, but do not have a household connection. To overcome this limitation, we repeat the analysis reported in Figure 2 and Figure 3 using all the available information on internet access. In particular we use the following variables: whether the *household* has internet connection, whether the respondent accesses internet at *home*, whether the respondent accesses the internet at *any location*, whether someone in household accesses the internet *outside of the home*, whether the respondent uses the internet at *someone else's house*, whether the respondent accesses the internet at a *public library*, whether the respondent accesses the internet at *school*. The results, reported in Figure 7, show that independently on how the respondents in our sample access the internet, the association with being in a partnership and the interaction do not change. Also, the results are consistent across different years. The only differences worth noticing are that the predicted probability of being in a union is

higher for those with internet access after age 27, while in Figure 2 it was after age 25, and that the difference between the two groups at older ages is less marked than in the analysis with only internet access at home.

Figure 7. Partnership and Access to Internet, All Locations (CPS)





Discussion

This paper analyzes the relationship between internet access and propensities to partner. We use individual-level data from the NLSY97, which contains data on internet access starting 2003 and partnership histories, as well as other socio-demographic characteristics that enable us to examine the effects of internet availability on partnership propensities. The NLSY allows us to follow a cohort and examine how internet access is associated with partnership transitions as the cohort grows older. We also analyze data from the CPS, which although cross-sectional in design, enables us to analyze the relationship between internet access and partnership status for multiple years. Although the potential role of the internet for partnership formation has theoretical plausibility and has attracted significant both scholarly and public interest, nationally-representative data sources that ask questions on internet access and partnership history are surprisingly limited. Our work explores this interesting and important question by drawing on multiple data sources. Our motivation for using different data sources is to assess if the effects are similar and constant across them, even if their questions and design imperfect to analyze the effect of internet diffusion on partnership formation.

One of these limitations is that our data do not allow us to measure exactly how individuals use the internet. Nevertheless, our findings suggest that internet access is associated with lower probabilities to transition into a partnership for both heterosexual and homosexual partnerships at younger ages relative to those without access to the internet. At older ages, however, the relationship between internet access and the transition to partnerships begins to change to a positive one. This effect emerges for both the transition to homosexual and heterosexual partnerships. These results are found in both the NLSY and CPS analyses.

Access to the internet generates a postponement effect, consistent with the hypothesis that while at younger ages the internet may enable the opportunity to expand one's network and meet new people, it does not encourage young people to enter partnerships. As individuals grow older, the effect

of the internet has a positive effect on their propensity to partner, which is consistent with ideas that emphasize the role of the internet in facilitating an efficient partner search. Although existing work suggests how the effect of the internet in facilitating partnership formation is likely to be stronger for those in 'thin markets', such as gays and lesbians, the effect sizes in both the NLSY and CPS are not larger among homosexual partnerships than those for heterosexual partnerships. At least in the NLSY, this may be driven by the small number of homosexual partnerships that we observe.

The fact that our findings are similar across the NLSY and CPS suggest that the association between internet access and partnership outcomes is not restricted to a specific cohort or set of cohorts. The age-specific effect of internet access and partnership status features across different periods in the CPS, suggesting that internet access is a period change that affects different cohorts.

An important shortcoming of our data is that we are unable to capture the diffusion of the internet as the internet question is only asked from 2003 onwards, when a significant fraction of the users already have access. Although we do have variation in the access to the internet variable over time in the NLSY, we do not fully capture the diffusion process of the new technology from a state of limited availability to saturation. To the best of our knowledge, such a dataset that tracks individual change in access to the internet in a longitudinal perspective over a long period of time, including when the technology had more limited use, and captures partnership transitions does not exist. This, to a certain extent, limits our ability to fully measure and understand how the diffusion of the internet, arguably the most important technological innovation of the twentieth century, has affected if, when, and with whom individuals partner.

Despite these limitations, however, we believe our findings provide support for hypotheses about the importance of the internet for partnership formation. We find that the internet affects partnership propensities when individuals feel ready to settle down, but delays these transitions at

younger ages. It affects this for different types of partnerships, including both homosexual and heterosexual partnerships.

References

Allison, P., 1982. Discrete-time methods for the analysis of event histories. Sociol. Methodol. 13, 61–98.

Barber, J.S., Murphy, S.A., Axinn, W.G., Maples, J., 2000. Discrete-time multilevel hazard analysis. Sociol. Methodol. 30, 201–235. doi:10.1111/0081-1750.00079

Bellou, Andriana, "The impact of Internet diffusion on marriage rates: evidence from the broadband market," Journal of Population Economics, 2015, 28 (2), 265–297.

Cacioppo, John T, Stephanie Cacioppo, Gian C Gonzaga, Elizabeth L Ogburn, and Tyler J VanderWeele, "Marital satisfaction and break-ups differ across on-line and off-line meeting venues," Proceedings of the National Academy of Sciences, 2013, 110 (25), 10135–10140.

Dettling, Lisa J, "Broadband in the labor market: the impact of residential high-speed internet on married womens labor force participation," ILR Review, 2017, 70 (2), 451–482.

Greenwood, S., Perrin A., and Duggan M.. 2016. "Social Media Update 2016." Pew Research Center: Internet, Science & Tech. Retrieved September 29, 2017 (http://www.pewinternet.org/2016/11/11/social-media-update-2016/).

Hitsch, Gunter J, Ali Hortacsu, and Dan Ariely, "What makes you click? Mate preferences in online dating," Quantitative marketing and Economics, 2010, 8 (4), 393–427.

Rosenfeld, Michael J and Reuben J Thomas, "Searching for a mate: The rise of the Internet as a social intermediary," American Sociological Review, 2012, 77 (4), 523–547.