The Long-term Effects of Expanding Social Capital via Social Media

Luis Armona
Harvard Kennedy School

Michael Bailey
Meta

Preliminary and incomplete - do not cite or circulate without authors' permission

February 28, 2025

Abstract

Even though social media use is prevalent worldwide, there is little empirical evidence on the impact of social media access on people's long-term networks and social capital. In this paper, we use Facebook's sequential rollout to 1,200 U.S. colleges from 2004-2006 to estimate the returns to social capital accumulation during college. We find exposure to Facebook changed the trajectory of people's social networks: 10-15 years after graduation they have larger online social networks, are more closely connected to peers from the same college, but otherwise engage in lower group homophily, and are more likely to have delayed marriage after college. Access to social networks during college is consistent with students building more social capital: Access is associated with increased connections with high socioeconomic status individuals and more diverse networks. Using social media access as a proxy to the returns to social capital accumulation, we find access is associated with moving further from hometowns to neighborhoods with more economic opportunity, and sorting into higher income occupations.

JEL Codes: D85, I2, J6

Keywords: Social Media, Social Capital, Network Formation, Neighborhood Choice, Mobility

1 Introduction

Facebook launched in 2004 at Harvard University and began as a campus-centered online social network for colleges in the United States. Now, over half the world's population uses some form of social media.[1] Facebook made it easier to connect and communicate with college classmates and remain connected after graduation and forever altered social life at college campuses across the US. The impact of social media on people and the social connections they formed is an active area of debate and research.

In this paper, we use the introduction of Facebook at U.S. colleges as a natural experiment to estimate the impact of access to social media on social connectivity, mobility, and occupational choice. Facebook started at Harvard in 2004 and was sequentially rolled out to most colleges in the U.S. by 2006. Prior to being released to the public in September 2006, users could only sign up for the social media platform using `.edu` email addresses associated with colleges which Facebook had granted early access. Students enrolling in college during this period had differing levels of access to Facebook depending on when Facebook entered their school. By using variation between college cohorts before and after Facebook's launch at their college, and between colleges who had access to Facebook earlier or later, we measure the differential impact on both the networks student form, and the choices students make, after college, resulting from access to Facebook during college.

Our sample consists of 1.38 million Facebook users and is drawn from those reporting graduating from one of the colleges where Facebook launched between 2004-2006 and who enrolled between 1995 and 2005. We observe the Facebook friendship graph over time as well as individual characteristics such as age, gender, location, and hometown for much of our sample until 2023, allowing us to measure the impact access to Facebook had on social networks and long-term outcomes many years after college, including location and occupational choice.

We find that access to Facebook during college had a large and pervasive impact on online social networks. An additional year of access to Facebook is associated with a 9.3% increase in friends across all measurement years with the largest difference in the years directly succeeding college (and the difference is still large in the final measurement year, 2019). Access to Facebook in college also produces systemic differences in network composition after graduation: students have networks with lower gender homophily (more likely to be friends with those of the opposite gender), less geographic homophily (smaller proportion of

[1] Source: `https://www.statista.com/statistics/278414/number-of-worldwide-social-network-users/`

local friends), more friends who live in high SES zipcodes, and more connections with college classmates. Of these changes, the largest occurs with respect to how connected students are to classmates after college. Because of this, we interpret the impact of Facebook access during college primarily through the lens of increasing *social capital* students accumulate from college, increasing the strength of ties between college alumni.

We estimate that Facebook exposure in college also leads to changes in students' social interactions on the marriage market, inducing students choosing to delay marriage after college. An additional year of access is associated with a reduced 3.7% chance of being married over our sample, though the difference is insignificant in later years. This reduction in marriage is almost entirely accounted for by substitution to non-marital relationships.

We find that social capital expansion via access to Facebook impacts offline post-college outcomes. Students exposed to Facebook choose to live in more dense, higher educated, more racially diverse, and higher income neighborhoods after graduating from college. These choices are consistent with students choosing to live in more urban areas. They also experience more geographic mobility: they are less likely to move back to their hometown after college, and, conditional on moving away, they move further away from their hometown. This suggests that earlier access to Facebook's social networking technology lessens student ties to local geographies, which may drive the differential choice of the types of neighborhoods students live in. Occupational choice is also impacted by exposure to Facebook, for those in the sample for whom we observe occupation on their Facebook profile, we find an average effect of an 1.4% increase in occupation quality for each year of access to Facebook.

Overall, the findings of this paper suggest that social media access during college leads to significant changes to students'post-college network structure. These changes to social capital cause students to move to neighborhoods with characteristics associated with greater economic opportunity, and facilitate geographic mobility. By directly measuring changes to network structure over time, and associating these changes with a natural experiment, our study is one of the first to quantify and link changes in social capital to outcomes many years later in life, and directly demonstrate social capital's importance for long-term welfare.

This paper adds to the literature on estimating the returns to college. Several studies have documented large pecuniary and non-pecuniary returns to college across contexts, even after accounting for selection Weiss [1995], Black and Smith [2004], Hoekstra [2009], Arcidiacono et al. [2010], Oreopoulos and Petronijevic [2013], Hastings et al. [2013], Kirkeboen et al. [2016], Cellini and Turner [2019]. A focus of existing research is estimating the returns to human capital accumulation and differentiating it from the signaling benefit of a degree after college. Some studies have theorized that some of these returns are due to social capital accumulation during college [Lleras-Muney et al., 2020, Calvo-Armengol and Jackson, 2004]. We contribute to this literature by showing that the channel of social capital expansion during college, has significant consequences both in terms of the structure of alumni social networks after college, and the choices these alumni make in where to reside and what jobs to pursue.

There have been some studies that examine peer effects and social networks in educational and college contexts [Sacerdote, 2001, Zimmerman, 2003, Lyle, 2007, Epple and Romano, 2011, Carrell et al., 2013, Michelman et al., 2022]. The most common measurement approach of this literature are small cohort studies that link group or dorm assignments to outcomes like grades, entrepreneurship, or habit formation. These studies often do not follow students after graduation and are unable to estimate the long-terms impacts of social networks on employment, mobility, or other outcomes [Calvó-Armengol et al., 2009]. We complement these literature in several ways. First, we are able to link variation in social capital and social networks from college to later life outcomes including employment and mobility, allowing us to estimate the link between social capital and long-term welfare. Second, our dataset consists of most colleges in the U.S. and covers a large fraction of total college students, allowing us to analyze heterogeneous effects in the relationship between social networks and social capital formation.

This study is also one of the first to use large-scale social network data over many years and observe long-term changes in social connections and mobility. Recent work analyzes large-scale online social network data and the structure and evolution of college social networks [Bailey et al., 2018, Overgoor et al., 2020a,b, Chetty et al., 2022]. In particular, Overgoor et al. [2020b] study the evolution of social networks at college and analyze the differences between cohorts and public and private colleges. By tying these measurements back to a natural experiment (the launch of Facebook in 2004), and tracking the long-term evolution of these networks after college, we are able to add evidence to the growing research body on the impacts of social media. That we find such a large and pervasive effect of social media access on social networks many years after college indicates the pivotal role social media plays in the post-college trajectory of people's social networks and the choices influenced by their peers.

Finally, this paper contributes to the large literature on the importance of social networks on a broad swath of economic, health, educational, and social outcomes Bourdieu [1986], Coleman [1988], Glaeser et al. [1996], Hoxby [2000], Jackson et al. [2008], Carrell et al.

[2009], Bifulco et al. [2011], Sacerdote [2011], Aral and Nicolaides [2017], Merlino [2019]. Usage and access to social networking sites (e.g. Facebook) in particular has been linked to online and offline outcomes including time spent outside, factual news knowledge, subjective well-being, and online activity Allcott et al. [2020], Braghieri et al. [2022]. Chetty et al. [2022] use large-scale social media data to estimate neighborhood level social capital in the US and find a link between cross socioeconomic links and upward income mobility. Most similar to our paper are Braghieri et al. [2022], and particularly Armona [2023], which leverage the same natural experiment, the release dates of Facebook to colleges, to estimate the effects of Facebook access on mental health during college and labour market outcomes, respectively. Using micro-data on online networks and the timing of the introduction of Facebook, we find that this differential timing leads to differences in friendship composition with access to Facebook making it more likely to be connected online with their college peers later on. We find that this channel explains later geographic mobility of students after college.

2 Data

We now describe the data elements we combine to conduct our analysis on the impact of Facebook access on social capital and later outcomes.

User Self-Reported Data: We collect data on user self-reported demographics from Facebook profiles. This includes the self-reported gender of each user, date of birth, the first postsecondary college attended by each user, their enrollment and graduation year at this college, the hometown of each user (the city they grew up in), the high school of each user, the self-reported jobs history of each user, and the relationship status of each user. For high school, we augment the self-reported data with a high school prediction that matches users to high schools based on the commuting zone of their hometown, as some high schools names are duplicated across the United States, so users accidentally misreport their high school as one implausibly far away from where they grew up.

Location Data: Each Facebook user has an assigned primary location which Meta determines based on factors such as user activity on Meta products, the location the user provides on their profile, as well as device and connection information. We obtain data on individual residential locations on June 30th of the given year, from 2012 to 2019. For those predicted to live in the United States, we map the predicted city of residence to the commuting zone and county of residence. We use this data to describe the geographic mobility of users after college.

Occupation Data: Facebook users can report their job titles and positions on their public user profile page. We use a fuzzy string matching procedure[2] to match these job titles/positions to O*NET job titles from the U.S. Department of Labor[3], which are then matched to SOC occupation codes. We then use the BLS Occupational Employment Statistics (OES) Survey to calculate the average wage of each SOC occupation each year, which is assigned to each individual with the appropriate job title during that year. We deflate these wages to 2019 dollars. If a O*NET job title is matched to multiple job codes is matched to multiple SOC codes, we weight these SOC codes uniformly to construct the average wage of that position[4] Note that this is not a stand-in for each individuals' wage; there is scope for large differences in wages within a precise occupation.[5] Instead, the average occupation wage serves as an "occupation index" to vertically arrange jobs as typically being better or worse paying. We use this to assess the impact of access to Facebook during college on later employment opportunities.

College Characteristics: We obtain data on college characteristics from the Integrated Postsecondary Education Data System (IPEDS) from the Department of Education. It contains characteristics for colleges where Facebook was released including whether it is a public university, the location type of the school (urban, suburban, rural), and the admissions rate of the school in 2003. We use these to evaluate whether Facebook access has a differential impact at different types of colleges.

High School Characteristics: We match self-reported high school on users' profile to a NCES ID and then use the NCES' Common Core Data (CCD)[6] for public schools, and the Private School Survey (PSS)[7] to characterize these high schools. Namely, we construct an indicator of whether the individuals attended a public or private school, and then the

[2] We use `thefuzz` (`https://github.com/seatgeek/thefuzz`) Python library to perform the string matching, and retain those jobs with a >75% match score to O*NET titles.

[3] See `https://www.dol.gov/agencies/eta/onet` for more information

[4] For example, "Cameraman" is matched to both SOC codes 27-4021 "Photographers" and 27-4031 " Camera Operators, Television, Video, and Motion Picture" in O*NET.

[5] On average, the interquartile range of wages within 6-digit SOC code is about 2/3 the variation across all jobs, meaning these occupations only explain about 1/3 of the wage variation across the labor force.

[6] Downloaded from `https://nces.ed.gov/ccd/files.asp`

[7] Downloaded from `https://nces.ed.gov/surveys/pss/pssdata.asp`

average % of students at the high school from 1998-2002 with free or reduced lunch (FRL) status, to approximate the socioeconomic status of the individual.

Hometown Characteristics: We characterize the socioeconomic status of each individual's hometown using the 1998 zip-code level statistics on income from the Internal Revenue Service (IRS).[8]. 1998 is chosen because it is the oldest year of IRS data available, and coincides with when the average person in our sample would have graduated high school, prior to entering college. Using this data, we map zip codes to hometowns and calculate the average income in 1998 of each hometown, along with the population density, measured as the number of personal exemptions, divided by the area covered by the zip codes contained in the hometown. We define an individual as coming from a rural hometown if the population density is less than 1000 people per square mile. These characteristics measure both the affluence of the hometown students come from, and how urban or rural the hometown of students is.

Zipcode Characteristics: We obtain annual data on the characteristics of zipcodes from 5-year ACS estimates, from 2011-2019.[9] The variables we use include the mean per-capita income, the employment rate of working age (18-64) adults, the poverty rate of both children and working age adults, the % of individuals aged 25-64 with at least an associate's degree, bachelor's degree, or graduate degree, and the fraction of racial minorities (the % Hispanic, African American, and Native American). We also use the distribution of race in the population to construct a metric for racial diversity in the zipcode, based on the Hirschman Herfandahl Index (HHI) to measure concentration in market:

$$D_{z,t} = 1 - \sum_r p_{z,t,r}^2$$

where z is a zipcode, t is a year, and $p_{z,t,r}$ denotes the fraction of individuals in the zipcode belonging to ethnicity group r.[10] Thus, a larger number indicates that the area is more racially diverse, or less racially segregated. We use this data to characterize the neighborhood choices of Facebook users after college.

[8] Available for download at `https://www.irs.gov/statistics/soi-tax-stats-individual-income-tax-statistics-`

[9] For example, the 2011 estimate comes from the 2007-2011 5-year ACS. Data is downloaded from `https://www.census.gov/data/developers/data-sets/acs-5year.html`

[10] These groups are: Black non-hispanic, White non-hispanic, Asian non-hispanic, Native American non-hispanic, Pacific Islander non-hispanic, Multi-racial non-hispanic, Other Race non-hispanic, and Hispanic.

Network Structure: We record snapshots of the Facebook friendship graph on June 30th of each year, from 2010 to 2019, to characterize the structure of online social networks for Facebook users. These includes the degree (# of friends), second degree (friends of friends), clustering coefficient (Fraction of one's friends who are friends with each other), as well as various measures of social homophily. Our homophily measures are characterized as the fraction of friends that share a characteristic with the individual. These include same gender, same self-reported employer, having attended the same college, belonging to the same cohort (defined as attending the same college and entering college in the same year), and residing in the same geographic location. We also borrow from Chetty et al. [2022] and measure an individual's "Economic Connectedness" as the % of friends living in high SES zipcodes (defined as zipcodes where the median household income in 2016 is greater than the U.S.-wide median of $54,483).[11]

Close Friendships: We complement our data on social network structure from users' friend lists with data on the closest friends of each user, as of June 8th, 2023. Specifically, we use an algorithm that predicts each user's closest friends based on certain user interactions. We use this data to evaluate whether early access to Facebook causes changes to social network structure beyond binary friendship decisions including the friends individuals communicate with the most.

Facebook College Release Dates: We impute the release dates of Facebook to 1,200 4-year colleges that received early access to Facebook between February 2004, when it was founded, to September 26th, 2006, when it was released to the public. Because college students received early access to Facebook using their college `.edu` emails, we use the registration emails of all Facebook accounts created before September 2006 to identify release dates at each college. We identify which emails are associated with which college using the domains (e.g. `harvard.edu,stanford.edu`) corresponding to the URLs that each college lists as its official website in the 2003-2006 IPEDS Institutional Characteristics Surveys. We then impute the release date to be the date where at least 34 users with email addresses from each college have signed up for the platform. This minimum number of 34 is chosen

[11] In Chetty et al. [2022], the authors use data on census block-level income to determine an individual's SES, then train a machine learning model using a subset of users' mobile phone GPS coordinates, to predict an individual's census block income for their entire sample. We rely on a more coarse measure of zipcode-level income data. However, we note that a regression of census block income with zip code fixed effects for the 2015-2019 5-year ACS captures 60% of the variation in income, suggesting zipcode-level income is a meaningful proxy for the more granular measure of economic connectedness used in that paper.

because it is found to minimize the error rate between the imputed dates and those released dates inferred in Armona [2023], which uses snapshots of the Facebook website from the Internet Archive, where a subset early access college were listed.[12] Panel (a) of Figure 1 plots the distribution of the rollout dates by Facebook to selective four-year colleges in the United States over time. In the first year, the rollout is fairly gradual, then there is a wave of colleges that gain access around May and September 2005. By the time Facebook was released to the public, over 90% of four-year colleges had early Facebook access. In Panel (b), we plot the release date against the admissions rate of each college, as a measure of selectivity of each college. We see clearly that which colleges received earlier access is not random: more selective colleges received access to the Facebook network earlier. We discuss how we overcome this selection problem to estimate the causal effect of access to Facebook in Section 3.

Sample And Summary Statistics: Our universe of potential schools are the 1,342 four-year colleges in the United States with an average cohort size of at least 100 students. We begin our sample construction by first identifying Facebook users that self-report attending one of these schools as their first postsecondary higher education experience that self-report their years of enrollment. We limit our sample to those who first enrolled between 1995 and 2005; because Facebook gave access to some high schools in September of 2005, we remove later cohorts as they may have had access during high school. We further restrict our sample to those who are matched to high schools, with a valid NCES ID, take 3 to 6 years to graduate, and those who report enrolling to college between the ages of 17 and 21, to prune out graduate students. This yields our final sample of 1.38 million Facebook users.[13]

Table 1 describes our sample of Facebook users. In the right three columns, we break up the summary statistics by the academic year that students received access to Facebook. On average, students who received access earlier register for a Facebook account earlier in time. Compared to the average income in the United States in 1998 ($41,000), our sample comes from richer areas of the United States. For schools receiving Facebook access later, student SES decreases, as measured by the average % receiving FRL programs in their high school,

[12] All colleges with access prior to August 2005 were displayed on Facebook's landing page. In September 2005, Facebook changed the design of its website so this information was no longer accessible. Because Facebook snapshots are unavailable for every day, the method of Armona [2023] for inferring release dates only had bounds on when Facebook was released to each college. Our chosen minimum of 34 users places the release date within these bounds 77% of the time, and only has an average gap of 6.5 days when the imputed release date from registration emails falls outside the bounds.

[13] Due to privacy concerns, we cannot disclose the number of users droppped from the sample in each intermediate step.

and the income of their hometown. Later access students are also more likely to come from more urban areas of the United States. Students in our sample are more likely to be female, often attend out-of-state colleges, and enroll in college around age 18.

Table 2 displays summary statistics of users' Facebook networks structure and personal outcomes over time. The number of friends increases over time, and begins with a high (20%) fraction of users being friends with those from the same college in our first year of measurement, 2010. There is significant geographic homophily among our sample: about 40% of friends live in the same commuting zone across years. On average, our sample is more likely than not to be friends with individuals who live in high SES zipcodes. These networks are fairly dispersed in terms of their clustering: the mean clustering coefficient is .07, meaning only around 7% of a user's friends are friends with each other, on average. Moreover, clustering of the networks decreases over time, suggesting new friends are not simply friends of friends, but rather often represent brand new connections. In terms of relationship status, the likelihood of being single is relatively stable, while the probability of being married (in a non-marital relationship) increases (decreases), as our sample gets older. Overall, these individuals are likely to work in high paying occupations: the average wage in 2019 in the BLS data is $53,000 across the U.S. labor force. They also tend to live fairly far (700km, or $\approx$ 430 miles) away from their hometown.

In Table 3, we report the characteristics of the zipcodes users live in after college, and for comparison the zipcode characteristics of the U.S. adult population. Over time, users move in richer, more educated neighborhoods, that also decrease in population density, possibly representing a move towards suburbs as individuals get older. Compared to the U.S. adult population, our sample lives in richer, more educated, and denser neighborhoods, across years.

3 Empirical Strategy

Using the inferred release dates at each college, we construct a measure of social capital expansion during college. Following Armona [2023], our measure is the years of access a student had access to Facebook during college, denoted τ:

$$\tau_i = \max(0, \text{Date}(6/30/g(i)) - \text{FB Release Date}_{j(i)})/365 \tag{1}$$

Where i indexes a user, $j(i)$ is the college j user i attended, $g(i)$ denotes the year of graduation, and FB Release Date$_{j(i)}$ is the inferred release date of Facebook to college j. For users graduating before Facebook was released, they have no exposure to Facebook during college. For all other users, this cohort-level variable varies depending on when they graduate relative to the introduction date. Because we do not observe the date of graduation across schools, we assume the end of the academic year, June 30th, marks the graduation date. In all of our analysis, we include school fixed effects; insofar as different colleges have differing, time-invariant graduation dates, any error in this assumption will be absorbed by the college fixed effects.

Figure 2 plots the average access time, along with the minimum and maximum, during college by entry year into college for four-year completers in our sample. We limit this figure to four-year completers so that variation is only attributable to the changes in the release dates within each class rather than how long it takes students to complete a degree. In our analysis, we control for differences in graduation time. For those entering in 1999 and before (graduating in 2003), there is no exposure during college. This is our control group for each college. Then, starting with those entering in the fall of 2000, we see that there is substantial variation across users attending differing colleges in our sample in how much exposure they had to Facebook, up to 2.4 years of difference. This large range of treatment time by cohort will generate substantial variation in treatment exposure we will use to identify the effect of Facebook access during college.

We use the following linear panel regression specification to estimate the causal effect of Facebook access in college on future labor market outcomes:

$$y_{i,t} = \alpha_{j(i)} + \kappa_{e(i),g(i),t} + \beta\tau_i + \gamma\mathbf{X_i} + \epsilon_{i,t} \tag{2}$$

$y_{i,t}$ denotes a user outcome measured on June 30th of year t. We restrict our panel of outcomes to those that occur at least one year after a user reports graduating college,[14] so that effects appropriately captures how exposure during college impacts post-college trajectories. α denote college fixed effects, to capture unobservable differences of outcomes of students graduating from differing colleges, and not attribute differences in outcomes stemming from college selectivity / quality to differences in Facebook exposure. κ denote enrollment-year $\times$ graduation year $\times$ outcome year fixed effects, These fixed effect controls for changes in the return to education over time, changes in the evolution of outcomes over the post-college lifecycle, across measurement years, but also ensures that comparisons in access time are only made within students who take the same amount of time to complete a degree (since

[14] For example, if a user enrolls in college in 2005 and graduates in 2010, we do not include their outcomes in our sample until 2011.

those taking, for example, six years in college, mechanically have longer access time). X_i denotes pre-college controls for individual i, and $\epsilon_{i,t}$ is a mean zero i.i.d. error. We include a comprehensive set of fixed effects to limit the role of differences attributable to student background. These include fixed effects for gender, age (years), the hometown of the user, and the high school the user attended. These latter two fixed effects effectively control for the neighborhood or socioeconomic environment that each individual in the sample grew up in. The coefficient of interest is β, on the access time treatment variable, and can be interpreted as the effect of Facebook being launched a year earlier during a student's college tenure. We cluster standard errors at the university level to capture any unobserved correlations between cohorts at the same college that may have overlapped and influenced each others earnings in a manner unrelated to Facebook access. This is also the level at which the treatment variable, Facebook's entry date into a campus, was assigned.

The initial recipients of the Facebook network were composed of the United States' most elite universities, as can be seen by Panel (b) of Figure 1. Facebook's release dates across schools are then unlikely to be random with respect to student outcomes. Instead, we assume that within a university exposed to the treatment (Facebook access), any particular *cohort* within that university is randomly exposed to Facebook. It is fairly plausible that Facebook did not coordinate to introduce their network to particular cohorts at these elite universities. For example, it is unlikely that Facebook decided to withhold access to UT Austin until the 2004-05 academic year because they determined the class of 2004 that had just graduated that June would not have signed up for their social network website. Instead, it is likely UT Austin received Facebook later than other universities because these students were less likely to sign up for Facebook than historically elite universities, or Facebook's founders, who were students at Harvard, had less ties to UT Austin, so it was a less natural place to expand to. The independence of the timing of which cohort is exposed to Facebook at a given university is the primary identifying assumption, and are be taken care of with university-level fixed effects in my regression analysis. Because τ is mechanically correlated with being in a later class, within school, we also worry about secular trends over time for college graduates that may be spuriously correlated with access time. Graduation year fixed effects should absorb this variation, so that residual differences in τ within cohort years come from differences in Facebook's release date to these colleges. Finally, through the inclusion of hometown and high school fixed effects, we control for differences in socioeconomic background by students within university.

One can frame our identification strategy for β in a differences-in-differences framework.

The thought experiment is as follows. With hometown and high school fixed effects, we can think of our sample for identification purposes as a group of users who went to the same high school and come from the same city, but attend different colleges. College fixed effects act as controls for the "preperiod" students outcomes at the same university that never received Facebook while in college, because they graduated too early. This is the first difference. We can then define treatment and control groups as cohorts that entered and graduated at the same time , but had differential exposure (in terms of τ) to Facebook during college due to variation in release dates. Subtracting the treatment group earnings from the control group earnings yields the second difference. ψ is then identified by two sources of variation: exposure differences between students at the same university, and differences in the release date of Facebook between cohorts of the same year. By assuming that conditional on university, cohort exposure is random, we can interpret the coefficient β as the causal effect of Facebook access on later life outcomes.

4 Network Structures After College

We begin our analysis using historical snapshots of the Facebook friendship graph to investigate whether there are persistent effects on the structure of social networks (as measured by Facebook) due to early access to the Facebook platform during college.

Before examining how networks of students change after college, we verify that earlier access results in earlier takeup of the Facebook platform. To do so, we regress the date of registration (in terms of days) of each user on Facebook access time, τ. Table 4 displays the results. In columns (1), when we only include college fixed effects, we find for each year of earlier access (365 days) during college results in a takeup of the platform 200 days earlier. Adding graduation×enrollment year fixed effects increases the coefficient magnitude to 250 days, and then finally inclusion of demographic controls X_i suggests an earlier sign-up of Facebook of 240 days. The estimates imply for each day of earlier access, users sign up for the platform 0.65 days earlier. This provides direct evidence that access to Facebook during college meaningfully changed student adoption of the social networking technology.

4.1 Network Size

We now turn to how student networks on Facebook change as a result of access during college. In Table 5, We regress the logged number of friends on access time. This regression pools across all years, to estimate an average effect from 2010-2019. We iteratively add controls from left to right in the tables, but our estimates remain relatively consistent in magnitude. The estimate of column (3) with a full set of controls implies that an additional year of access to Facebook implies a 9.3% increase in friends on average across all measurement years. In Figure 3, we estimate the impact on individual years of measurement. We plot the effect of access during college on both number of friends (degree), and friends of friends (second degree). For both measures of network size, the effect is largest in earlier years, and then decreases over time, but is still large in magnitude (8.1%) in the final measurement year, 2019. The larger effect size on second degree suggests not only that earlier access students make more friends, but the friends they make also have larger networks.

Because our data collection procedure begins in 2010, by which the vast majority of individuals in our sample have graduated, it may be that these differences in network size stem purely from a "head start" of larger networks in college, followed by little increased growth. To test this, we also examine the effect on the annual change in friendships each year, or the first difference $\Delta y_{i,t} = y_{i,t} - y_{i,t-1}$. Figure 4 show the estimated results. from 2010-2015, early access students make 2-3 more friends per year on average, after which the effect size is not statistically distinguishable from zero. This shows that even after college, the evolution of student network size changes as a result of college access and earlier takeup of the platform.

One may be concerned that these changes in network size are driven by differences in activity on the platform; that is, earlier recipients of FB are more active on the platform later in life, leading them to friend more individuals. This would be consistent with their "online networks", as measured on Facebook, increasing in size, while their "offline" or true social networks are unchanged. Appendix Figure A1 tests the extent that earlier recipients are more active, by regressing the # of days users have logged in the platform in June, when we measure networks, on college exposure. Estimates are reported in terms of standard deviations of this outcome variable.[15] The estimates suggest, if anything, the opposite is true. Until 2016, the estimates are indistinguishable from zero, and extremely small in magnitude; from 2017-2019, there is a slight negative effect of about -.02 standard deviations. Confidence intervals rule out positive effects in any year greater than .016 standard deviations, suggesting an effect on network size via the channel of increased activity on the platform, if present, is economically small.[16]

[15] Due to privacy concerns, we cannot disclose estimates in levels of the outcome variable

[16] For example, taking the upper confidence interval limit of .004 standard deviations and comparing it to the pooled estimate in Column (3) of Table 5 would imply that, if these effects are driven entirely by increased FB usage, then a 1 SD increase in FB usage would imply a 1600% increase in network size, which

4.2 Network Composition

How do the larger networks of students with access during college differ in structure? Our primary measure of network composition is the degree of *homophily* [Jackson et al., 2008] in students' post-college networks , or how likely they are to be friends with others that share certain characteristics. We will measure homophily as the fraction of friends in their network at a given point in time that share a characteristic with the student.[17] A natural expectation is that by having access during college, students might be more closely connected to their college classmates after they graduate. However, the shock of Facebook to how students formed friends in college may affect other types of social homophily.

In Figure 5, we plot estimates by year (and pooled estimates across years on the right hand side) of the changes in homophily along five dimensions: friends from the same college, friends from the same cohort (e.g. same college and the same college entry year), friends with the same gender, friends living in the same geographic area (measured by the U.S. commuting zone), and friends living in high SES zip codes (e.g. economic connectedness). On average, one additional year of access during college leads to a 2.5 percentage point increase in the % of friends from the same college, and a 0.24 percentage point increase in those from the same cohort. Relative to the average values in the estimation sample, these represent a 15% and 19% increase, respectively, in school homophily. While the effect on cohort homophily is relatively stable over time, school homophily decreases, consistent with students forming new friendships outside their college networks after graduating.

We find that Facebook access during college also affects other types of social homophily later in life. We estimate that students are more likely to be friends with those of the opposite gender from earlier access to Facebook. This suggests the social networking technology of Facebook may partly assist in breaking down social norms that have historically led to gender assortativity in social interactions in a broad set of context [Gallen and Wasserman, 2023].

There is a decrease in geographic homophily: students with more access during college are less likely to be friends with individuals living near them, representing a 3.1% decrease relative to the baseline degree of geographic homophily (40.5%). This is consistent with the social networking technology of Facebook, easing the ability to stay in touch with those far away, decreases the dependence of students on local geographic networks for friendship formation after college. In Appendix Figure A3, we plot the evolution of geographic homophily

seems implausible.

[17]For some measures, such as geography, certain characteristics are missing for friends. We will measure homophily for these characteristics as the fraction of friends with non-missing values of the characteristic.

along other geographic measures, from region (equivalent of U.S. state) to zipcode. For all except city, we see that access during college decreases geographic homophily along these measures as well.

Finally, we estimate that access during college increases students' economic connectedness for a significant period after college. On average students, are likely to have .28% more friends who live in high SES zipcodes, though the effect is statistically insignificant in the later years of the panel. Economic connectedness has been shown to be correlated with a broad array of outcomes important for economic opportunity [Chetty et al., 2022], which suggests that social capital expansion from Facebook access during college may be arming students with new ways to improve their own economic status.

Beyond the observable traits friends may share on the Facebook platform, there may be other ways in which network composition changes as a result of social capital expansion in college via this technology. We consider the degree of network clustering, which measures how often a student's friends are friends with each other, as an additional outcome of interest. In Figure 6, we present estimates by year of the effect of access during college on how clustered students' networks are after college. There is a decrease in all years, with the effect diminishing over time. This suggests Facebook access during college makes friendship networks more diffuse and diverse.

We summarise the effects on network composition in Appendix Table A1, which lists effect sizes in terms of network structure variables in our sample, in order to make the estimated coefficients more comparable. Of the measures of network composition, the largest effects are on those relating to college homophilly: one year of access increases the effect on those from the same school and cohort by .2 and .16 standard deviations, respectively. the next largest coefficient is more than three times smaller in terms of standard deviations. This suggests the primary shift from exposure to Facebook in college is to make students more embedded in their alumni social networks. Because of this, we interpret our later results as stemming primarily from an increase in social capital from college; students are more connected to college classmates, which may enable new opportunities for economic mobility.

Across these measures, the network composition changes from access to Facebook during college appear substantial. Increasing social capital in college via access to Facebook appears to lead to more diverse social networks that are less tied to social norms of gender assortativity, less constrained by local geography, friends living in richer neighborhoods, and social interactions with those less likely to be existing friends of friends. Given the prior evidence on these social metrics impacting later life outcomes, this suggests significant scope for this intervention to also affect the livelihood of students after college.

4.3 Do Networks Actually Change?

We document that networks, as measured by users' friends lists, are significantly altered from social media access during college. However, friends on Facebook may not be perfectly representative of the relationships that are important to users in real life, as some friendships on the platform may be "stale" and not reflect true changes to social interactions. To this end, we supplement our analysis of the effect on Facebook's friendship graph with two pieces of evidence.

First, we use the close friend predictions from June 2023, which serve as an index of overall online communication between users, to see if students change communication patterns due to early access in college. We vary the cutoff of what constitutes "close friends" from the top 5 to 25 predicted closest friends, and measure what fraction of close friends share a particular characteristic for each cutoff. Figure 7 shows the estimates. For the top 10 to top 25 friends, we find that early access college alumni are more likely to have close friends that come from the same school. The estimate implies that a top 10 friend is 3.4% more likely (in percentage points) to be an alumni from each student's alma mater. There is a smaller, but statistically significant effect, on being from both the same college and class for the top 20/25 friends, and a decrease in gender homophily among close friends as well from top 15 to top 25.

Second, we examine how the relationship choices of students after college change due to Facebook, using the self-reported relationship status of individuals. In Figure 8, we plot annual estimates of the effect on access to Facebook during college on relationship outcomes over time. In 2012, when our sample is younger (on average 30 years old), there is a large decline of 7.8 percentage points in the likelihood of marriage, and it appears students are primary substituting to non-marital relationships (6.6% increase). Over time, this substitution declines, as students move from non-marital to marital relationships as they grow older. The results on marriage are consistent with those found in Armona [2023], using the same natural experiment, finds a decline in marriage rates in 2014 from Facebook exposure. What we contribute is what the change in marriage rates leads to: rather than representing students failing to find romantic partners, instead it appears students remain embedded in significant relationships, and are instead delaying matrimony, which has been shown to provide benefits on the labor market for career development [Wang and Wang, 2017] and marriage stability Becker et al. [1977], Rotz [2016].

These results support our findings that access to Facebook during college meaningfully changed both how users interact socially after college, and also the choices they make in their relationships.

4.4 Heterogeneous Effects

We now consider whether the effects on network structure differ by user demographics. We consider 4 sources of heterogeneity: SES of the high school (measured by the % of students in federal free or reduced lunch (FRL) programs, whose eligibility is linked to poverty status), the gender of the student, the density of the hometown (measured by the population density in 1998), and whether the college is located in a city, according to IPEDS. For high school SES, we split by whether the % FRL of a user's high school is greater than 25% (about 1/4 of the sample). We split hometowns into rural or urban based on whether the density is > 1000 people per square mile, consistent with prior U.S. Census definitions[18]. We explore the heterogeneous effects by college locale in order to determine whether students attending colleges in cities, which already may have access to large local economic networks dur to their location, experience differential returns to social networking technology. For each demographic split, we re-estimate Equation 2 on the corresponding sub-sample of users.

Figure 9 presents the average effects across measurement years on the total number of friends on Facebook. Students with higher SES expand their networks further, as do females, and those from rural/suburban hometowns. Broadly speaking however, the effect sizes are large across the population. In Figure 10, we explore the effects on network composition. For school homophily (Panel (a)), we also see higher SES background students and females are more connected to their college cohorts. At the same time, for economic connectedness (Panel (b)), low SES students are the largest beneficiaries. This may be related to the fact that those who come from low SES backgrounds and attend college, by being more closely connected to their college alumni through social media technology, have a larger pool of high SES friends to form connections to, compared to what counterfactuals friendships would look like, outside of college networks. We also see that low SES students experience larger decreases in geographic homophily (Panel (c)). Finally, access to Facebook in college is more impactful for female students in decreasing gender homophily (Panel (d)), suggesting they are likely to make male friends and not discriminate on gender when forming friendships.

[18]Source: `https://www2.census.gov/geo/pdfs/reference/GARM/Ch12GARM.pdf`

5 Student Outcomes

We established in the prior section that online social networks experience large changes from earlier access to Facebook in college. In particular, we found that students become more connected to their college classmates, individuals in high SES neighborhoods, and less connected to local geographic networks. In this section, we examine whether these changes to network structure result in concrete changes in student outcomes after college. We focus on the implications for geographic mobility of students, due to the available data on residential location from Facebook, but also analyze the occupation choices of students due to a shock to social capital from college.

5.1 Neighborhood Choice

We begin by analyzing the characteristics of neighborhoods (zipcodes) students choose to reside in after college. Table 6 provides estimates of Equation 2 for characteristics relating to neighborhood income, racial diversity, education, and density, averaged across all years.[19] The largest effect across outcomes is on the population density of zipcodes: an additional year of access to Facebook during college leads to a 12% increase in population density. This effect is consistent across measurement years (Figure A4). These zipcodes are also more likely to be populated by college-educated individuals (.5% increase in percent with bachelor's degree per additional year of access in college). Examining across years, and alternative definitions of education, the effect is largest in the early portion of the sample, from 2012-2015, and then dwindles in later years of the sample (Figure A5). This pattern is consistent across alternative definitions of education, such as % with associates plus degrees, or % with graduate degrees. We find that students choose to live in more racially diverse neighborhoods: for each year of access, diversity along racial lines increases by 1.4%. As seen in Figure A6, this effect is consistent across years of the sample, and are coupled with increases in the overall fraction of racial minorities in neighborhoods. With regards to income measures, neighborhoods chosen by students are both richer on average, by 1.2% per year of access in college, and have higher poverty rates. This points to neighborhood choices that are more diverse in income. This effect is also consistent across years (Figure A7).

Overall, students with more exposure to Facebook during college choose to live in neighborhoods that are more diverse, both along racial and income lines, denser, and more educated. These choices are all consistent with Facebook access leading students to live in

[19]For these regressions, we cluster standard errors by both college and zipcode.

more urban areas, where income and racial segregation are low. These factors have been shown to be strongly correlated with upward mobility Chetty et al. [2014], suggesting that the changes to social capital fromcollege, via access to Facebook, induces students to reside in areas with higher economic opportunity.

In Figure 11, we plot the heterogeneous average effects on neighborhood characteristics by user demographics from Facebook access in college. Across all the metrics we consider, female students are more impacted than males in terms of moving to denser, more affluent and less racially segregated neighborhoods. This is consistent with the evidence from Figures 9 and 10 that females also had the largest changes to their network structure. We also find across the board that those from rural/suburban hometowns are more impacted in neighborhood choice in an analogous manner. No significant differences are found on neighborhood choice when measuring SES by high school.

While the evidence above clearly demonstrates the types of neighborhoods students live in changes, this may be induced by individuals simply living in nearby, more affluent neighborhoods to where they may have lived absent of Facebook. We next explore whether geographic mobility after college is altered by access to Facebook during college. To do so, we use the predicted city of residence for each user after college, and calculate the geographic distance from the user's reported hometown, from 2012-2019. We then calculate the effect of access on both (a) whether the student lives in their hometown (e.g. moves back to where they grew up after college) and (b) the distance students live from their home. In Figure 12, we plot these estimates by year. Our primary metric for geographic mobility is the inverse hyperbolic sine of geographic distance from one's hometown, to calculate a % effect on distance but include those living in their hometown after college. On average, we estimate that geographic mobility increases by 10% due to an additional year of access to Facebook. We estimate students are less likely to live in their hometown after college by 1.0 percentage points for each year of access. Our effect of distance is not driven solely by those choosing to not move back home. We estimate that, conditional on not living in their hometown, students still live 5.1% further away from their hometown on average for each year of access. Together, these results suggest that access to social networking technology in college increases geographic mobility. Intuitively, through using social media to stay connected to friends, students are less tied to local geography for maintaining social connections, and are able to move further away from their existing social networks, possibly to pursue economic opportunity.

Though we find that students are engaging in more geographic mobility, an open question is whether the changes to neighborhood characteristics come from moves to entirely new areas of the country, or differential choices of neighborhoods within a metropolitan area. To determine this, we rerun the regressions in Table 6, but include various levels of *contemporaneous* geographical area fixed effects. Specifically, we re-run Equation 2 with zipcode characteristics as the dependent variable, but include state-of-zipcode and commmuting zone-of-zipcode fixed effects. This specification means that the residual variation for identifying β must come solely from within state or CZ, respectively. Figure 13 displays the estimates. We note first that, controlling for CZ, all effect sizes are still statistically significant at conventional levels, except for neighborhood income, which suggests the benefits estimated of more affluent neighborhoods come primarily from living in richer metropolitan areas. For the other outcomes, We can use the estimate with CZ fixed effects to decompose the fraction of the changes in neighborhood characteristics attributable to social media access in college from moving to better neighborhoods within a commuting zone:

$$\% \text{ Explained By Intra-Metro Choice} = \frac{\beta_{\text{CZ FE}}}{\beta_{\text{No Area FE}}} \times 100 \tag{3}$$

For both density and racial diversity, 1/2 of the variation is explained by neighborhood choice within a commuting zone. For education, about 1/3 of the variation is due to within commuting zone. These results suggests that both choices of which metrpolitan areas to live in, and what neighborhood to live in within a metropolitan area, both play important roles in the changes to neighborhood characteristics of early access Facebook users.

5.2 Occupation Choice

We explore the question of whether these higher opportunity areas induced by changes to social network structure result in better job opportunities for early Facebook users. To do so, we use the self-reported employment history of Facebook individuals recorded on user profiles.

This data is not without it's limitations. We do not observe the income of individuals, so we cannot disentangle whether someone working as, for example, an orthodontist, is at the top of their profession and making substantially more income than their peers in the same job. Our primary outcome will instead be the average wage within an occupation in the United States, which will serve as an "occupation index" of occupation quality in order to investigate whether those with early access to social media during college sort into better or worse jobs. There is still substantial variation in quality within these occupation designations. Additionally, unlike the predicted residential choices of students after college, coverage for this variable is relatively low- about 20-25% of users in our sample report employment in each year sufficiently rich that it can be matched with reasonable accuracy to an SOC-designated occupation. With this in mind, we view the results as suggestive in how employment opportunities are impacted by Facebook.

In Figure 14, we plot annual estimates of the effect of early access on average occupation wages of users.[20] Until 2012, there is a statistically indistinguishable effect from zero, but a clear trend of increasing wages, that becomes statistically significant in 2013 and remains so until 2019. Across years, the average effect from one year of access is a 1.4% increase in occupation quality for each year of access to Facebook. The impact of Facebook on occupation quality peaks in 2015, when the effect size is 2.3%, and dwindles slightly thereafter, suggesting eventually the social capital advantage from Facebook access during college will fade to zero. In terms of heterogeneous effects, we find that female alumni benefit more from Facebook access during college (Figure 15), but the increases are broadly shared across demographics.

This does not appear to be driven by differences in the overall sectors students work in after college as a result of Facebook; In Appendix Figure A8, we plot the effect of access during college on the probability of occupation at the 2-digit SOC code level. There are small increases (<1%) in STEM-related occupation sectors, such as engineering, computer science, and the natural sciences, but for the vast majority of occupations, the effect is zero. This suggests that instead, the change in wages we estimate is coming from primary within-sector changes, such as a real estate agent becoming a real estate broker.

6 Conclusion

In this paper, we examine the effect of an expansion of social capital from college, via Facebook's early entry to college campuses, on how social networks and outcomes evolve. We estimate online networks are larger, more diverse, and more tied to college classmates through this intervention, as well as individuals in high SES neighborhoods. Beyond connecting students more to college alumni, this intervention reduces student ties to local geography, by decreasing geographic homophily in social networks, and increasing geographic mobility. We estimate that this induces changes in what neighborhoods students live in after college: on average, college graduates live in urban neighborhoods that are denser, richer, and more edu-

[20]For this section, we two-way cluster regressions by both college and SOC-code of matched occupation.

cated, with less racial and income segregation. We provide additional evidence that personal outcomes change as well: students sort into higher wage occupations, and choose to delay marriage (but not romantic partnerships). In total, the changes to social network structure suggest that by fostering stronger ties to high human capital college alumni, individuals are able to move to areas with higher economic opportunity, and pursue better jobs as a result.

References

Hunt Allcott, Luca Braghieri, Sarah Eichmeyer, and Matthew Gentzkow. The welfare effects of social media. *American Economic Review*, 110(3):629–676, 2020.

Sinan Aral and Christos Nicolaides. Exercise contagion in a global social network. *Nature communications*, 8(1):14753, 2017.

Peter Arcidiacono, Patrick Bayer, and Aurel Hizmo. Beyond signaling and human capital: Education and the revelation of ability. *American Economic Journal: Applied Economics*, 2(4):76–104, 2010.

Luis Armona. Online social network effects in labor markets: Evidence from facebook's entry to college campuses. *The Review of Economics and Statistics*, pages 1–47, 2023.

Michael Bailey, Rachel Cao, Theresa Kuchler, Johannes Stroebel, and Arlene Wong. Social connectedness: Measurement, determinants, and effects. *Journal of Economic Perspectives*, 32(3):259–80, 2018.

Gary S Becker, Elisabeth M Landes, and Robert T Michael. An economic analysis of marital instability. *Journal of political Economy*, 85(6):1141–1187, 1977.

Robert Bifulco, Jason M Fletcher, and Stephen L Ross. The effect of classmate characteristics on post-secondary outcomes: Evidence from the add health. *American Economic Journal: Economic Policy*, 3(1):25–53, 2011.

Dan A Black and Jeffrey A Smith. How robust is the evidence on the effects of college quality? evidence from matching. *Journal of econometrics*, 121(1-2):99–124, 2004.

Pierre Bourdieu. The forms of capital.(1986). *Cultural theory: An anthology*, 1:81–93, 1986.

Luca Braghieri, Ro'ee Levy, and Alexey Makarin. Social media and mental health. *American Economic Review*, 112(11):3660–3693, 2022.

Antoni Calvo-Armengol and Matthew O Jackson. The effects of social networks on employment and inequality. *American economic review*, 94(3):426–454, 2004.

Antoni Calvó-Armengol, Eleonora Patacchini, and Yves Zenou. Peer effects and social networks in education. *The review of economic studies*, 76(4):1239–1267, 2009.

Scott E Carrell, Richard L Fullerton, and James E West. Does your cohort matter? measuring peer effects in college achievement. *Journal of Labor Economics*, 27(3):439–464, 2009.

Scott E Carrell, Bruce I Sacerdote, and James E West. From natural variation to optimal policy? the importance of endogenous peer group formation. *Econometrica*, 81(3):855–882, 2013.

Stephanie Riegg Cellini and Nicholas Turner. Gainfully employed? assessing the employment and earnings of for-profit college students using administrative data. *Journal of Human Resources*, 54(2):342–370, 2019.

Raj Chetty, Nathaniel Hendren, Patrick Kline, and Emmanuel Saez. Where is the land of opportunity? the geography of intergenerational mobility in the united states. *The Quarterly Journal of Economics*, 129(4):1553–1623, 2014.

Raj Chetty, Matthew O Jackson, Theresa Kuchler, Johannes Stroebel, Nathaniel Hendren, Robert B Fluegge, Sara Gong, Federico Gonzalez, Armelle Grondin, Matthew Jacob, et al. Social capital i: measurement and associations with economic mobility. *Nature*, 608(7921): 108–121, 2022.

James S Coleman. Social capital in the creation of human capital. *American journal of sociology*, 94:S95–S120, 1988.

Dennis Epple and Richard E Romano. Peer effects in education: A survey of the theory and evidence. In *Handbook of social economics*, volume 1, pages 1053–1163. Elsevier, 2011.

Yana Gallen and Melanie Wasserman. Does information affect homophily? *Journal of Public Economics*, 222:104876, 2023.

Edward L Glaeser, Bruce Sacerdote, and Jose A Scheinkman. Crime and social interactions. *The Quarterly journal of economics*, 111(2):507–548, 1996.

Justine S Hastings, Christopher A Neilson, and Seth D Zimmerman. Are some degrees worth more than others? evidence from college admission cutoffs in chile. Technical report, National Bureau of Economic Research, 2013.

Mark Hoekstra. The effect of attending the flagship state university on earnings: A discontinuity-based approach. *The review of economics and statistics*, 91(4):717–724, 2009.

Caroline M Hoxby. Peer effects in the classroom: Learning from gender and race variation, 2000.

Matthew O Jackson et al. *Social and economic networks*, volume 3. Princeton university press Princeton, 2008.

Lars J Kirkeboen, Edwin Leuven, and Magne Mogstad. Field of study, earnings, and self-selection. *The Quarterly Journal of Economics*, 131(3):1057–1111, 2016.

Adriana Lleras-Muney, Matthew Miller, Shuyang Sheng, and Veronica T Sovero. Party on: The labor market returns to social networks and socializing. Technical report, National Bureau of Economic Research, 2020.

David S Lyle. Estimating and interpreting peer and role model effects from randomly assigned social groups at west point. *The Review of Economics and Statistics*, 89(2):289–299, 2007.

Luca Paolo Merlino. Informal job search through social networks and vacancy creation. *Economics Letters*, 178:82–85, 2019.

Valerie Michelman, Joseph Price, and Seth D Zimmerman. Old boys' clubs and upward mobility among the educational elite. *The Quarterly Journal of Economics*, 137(2):845–909, 2022.

Philip Oreopoulos and Uros Petronijevic. Making college worth it: A review of research on the returns to higher education. 2013.

Jan Overgoor, Lada A Adamic, et al. The structure of us college networks on facebook. In *Proceedings of the International AAAI Conference on Web and Social Media*, volume 14, pages 499–510, 2020a.

Jan Overgoor, Bogdan State, and Lada Adamic. The dynamics of us college network formation on facebook. *arXiv preprint arXiv:2011.01967*, 2020b.

Dana Rotz. Why have divorce rates fallen?: The role of women's age at marriage. *Journal of Human Resources*, 51(4):961–1002, 2016.

Bruce Sacerdote. Peer effects with random assignment: Results for dartmouth roommates. *The Quarterly journal of economics*, 116(2):681–704, 2001.

Bruce Sacerdote. Peer effects in education: How might they work, how big are they and how much do we know thus far? In *Handbook of the Economics of Education*, volume 3, pages 249–277. Elsevier, 2011.

Chunbei Wang and Le Wang. Knot yet: minimum marriage age law, marriage delay, and earnings. *Journal of Population Economics*, 30:771–804, 2017.

Andrew Weiss. Human capital vs. signalling explanations of wages. *Journal of Economic perspectives*, 9(4):133–154, 1995.

David J Zimmerman. Peer effects in academic outcomes: Evidence from a natural experiment. *Review of Economics and statistics*, 85(1):9–23, 2003.

FB Released at College:	All Users	Before Sept 2004	Sept 2004 - Aug 2005	Sept 2005- Sept 2006
Enrollment Year	2000.761	2000.743	2000.764	2000.758
	(3.083)	(3.075)	(3.081)	(3.099)
Years In College	4.167	4.103	4.181	4.144
	(0.721)	(0.592)	(0.726)	(0.788)
Public College	0.673	0.520	0.736	0.455
	(0.469)	(0.500)	(0.441)	(0.498)
Registration Date (In Years)	2007.611	2007.113	2007.579	2008.242
	(2.728)	(2.949)	(2.683)	(2.655)
Exposure to FB During College	1.505	1.715	1.528	1.186
	(1.695)	(1.749)	(1.706)	(1.536)
Male	0.392	0.448	0.388	0.366
	(0.496)	(0.500)	(0.493)	(0.504)
Year of Birth	1982.280	1982.413	1982.281	1982.151
	(3.186)	(3.141)	(3.186)	(3.223)
Pct FRL in HS	0.165	0.126	0.159	0.235
	(0.174)	(0.154)	(0.166)	(0.207)
Public HS	0.867	0.786	0.872	0.911
	(0.345)	(0.413)	(0.338)	(0.298)
Avg. Hometown Income (1998)	45831.338	54736.604	45665.574	39078.115
	(22021.520)	(31737.201)	(20829.360)	(14215.959)
Rural Hometown	0.574	0.389	0.580	0.700
	(0.501)	(0.491)	(0.498)	(0.475)
Out of State Student	0.326	0.482	0.299	0.342
	(0.475)	(0.502)	(0.463)	(0.494)
Number of Users	1,384,014	163,938	1,036,355	183,721
Number of Colleges	1202	55	648	499

Table 1: User-Level Summary Statistics

Year of Measurement:	2010	2012	2015	2019
Number of Friends	356.510	463.668	576.924	747.156
	(294.769)	(396.782)	(513.980)	(696.234)
Pct Friends From Same School	21.346	18.534	15.849	13.609
	(14.194)	(12.847)	(11.608)	(10.538)
Pct Friends From Same Class	1.770	1.445	1.194	0.986
	(1.777)	(1.529)	(1.384)	(1.277)
Pct Friends with Same Gender	56.719	57.049	57.949	58.878
	(12.010)	(11.709)	(12.312)	(12.961)
Pct Friends in Same Region		53.055	53.225	52.475
		(27.169)	(27.278)	(26.856)
Pct Friends in Same County		26.726	27.509	28.451
		(20.484)	(20.740)	(20.689)
Pct Friends In Same City		17.349	17.601	18.443
		(17.604)	(17.725)	(18.329)
Pct Friends in Same Zipcode		3.957	6.004	6.669
		(7.274)	(9.850)	(9.559)
Pct Friends in High SES Zipcodes		55.948	58.652	52.922
		(15.810)	(19.324)	(18.806)
Pct Friends with Same Employer	0.826	0.770	0.772	0.732
	(1.995)	(1.865)	(1.803)	(1.660)
Clustering Coef	0.085	0.074	0.064	0.054
	(0.057)	(0.053)	(0.053)	(0.050)
Days Logged on to FB in June*		-0.015	0.074	-0.161
		(0.931)	(0.921)	(1.185)
Pct Married		52.517	62.972	71.160
		(50.841)	(49.156)	(46.156)
Pct In Relationship		26.305	17.241	11.131
		(44.845)	(38.463)	(32.076)
Pct Single		21.178	19.788	17.709
		(41.717)	(40.638)	(38.938)
Average Wage of User's Occupation	77496.556	78294.675	80206.223	74756.879
	(39206.731)	(39316.464)	(41210.135)	(35450.103)
Distance From Hometown (km)		703.493	705.894	730.939
		(1721.137)	(1685.316)	(1714.820)
Distance From College (km)		622.885	548.847	581.423
		(982.104)	(947.064)	(970.925)
Users	1,189,853	1,289,244	1,353,261	1,385,322
Users with Known Location (Zipcode)	0	1,253,085	1,329,269	1,356,196
Users with Known Occupation	243,674	322,083	361,199	283,100

Table 2: Network Structure and Outcomes Summary Statistics

Table displays summary statistics of outcome measures in June 30th of each year. Cells report mean and standard deviations in parentheses. For Days logged on to FB in June, due to privacy reasons, we cannot disclose the mean during the sample. Instead, we report the outcome in terms of standard deviations, with 0 denoting the mean over the entire sample.

Sample	Facebook User Sample			U.S. Adult Population		
Year of Measurement:	2012	2015	2019	2012	2015	2019
Income Per Capita	32550.729	34500.739	40045.237	27943.342	28954.500	34144.805
	(16966.036)	(15992.192)	(19668.160)	(293387.349)	(12862.397)	(15131.585)
Poverty Rate (Working Age Adults)	13.676	13.350	12.658	14.152	14.864	12.955
	(9.690)	(8.914)	(8.885)	(9.414)	(9.537)	(8.708)
Pct Racial Minority	28.292	26.368	28.488	29.420	30.230	31.139
	(25.606)	(22.764)	(23.926)	(27.003)	(26.937)	(26.797)
Pct BA+ Over Age 25	37.640	40.708	41.962	30.087	31.140	33.406
	(20.135)	(19.634)	(20.277)	(17.389)	(17.718)	(18.348)
Population/Sq. Mi.	997.279	948.217	874.374	596.101	612.988	615.727
	(2402.282)	(2262.341)	(1880.542)	(1478.487)	(1510.279)	(1490.416)

Table 3: Characteristics of Zipcode Residence After College

Table 4: Effect of Exposure to FB in college on FB Registration Date

	Outcome: Registration Date (Days)		
	(1)	(2)	(3)
Years Exposed to FB in College	-201.40^{***}	-248.05^{***}	-239.62^{***}
	(2.02)	(18.19)	(15.63)
College FE	Yes	Yes	Yes
Grad Year FE	No	Yes	Yes
Demographic Controls	No	No	Yes
Observations	1,384,014	1,384,014	1,384,014
R^2	0.17	0.22	0.29
Adjusted R^2	0.17	0.22	0.26

Note: *p<0.1; **p<0.05; ***p<0.01

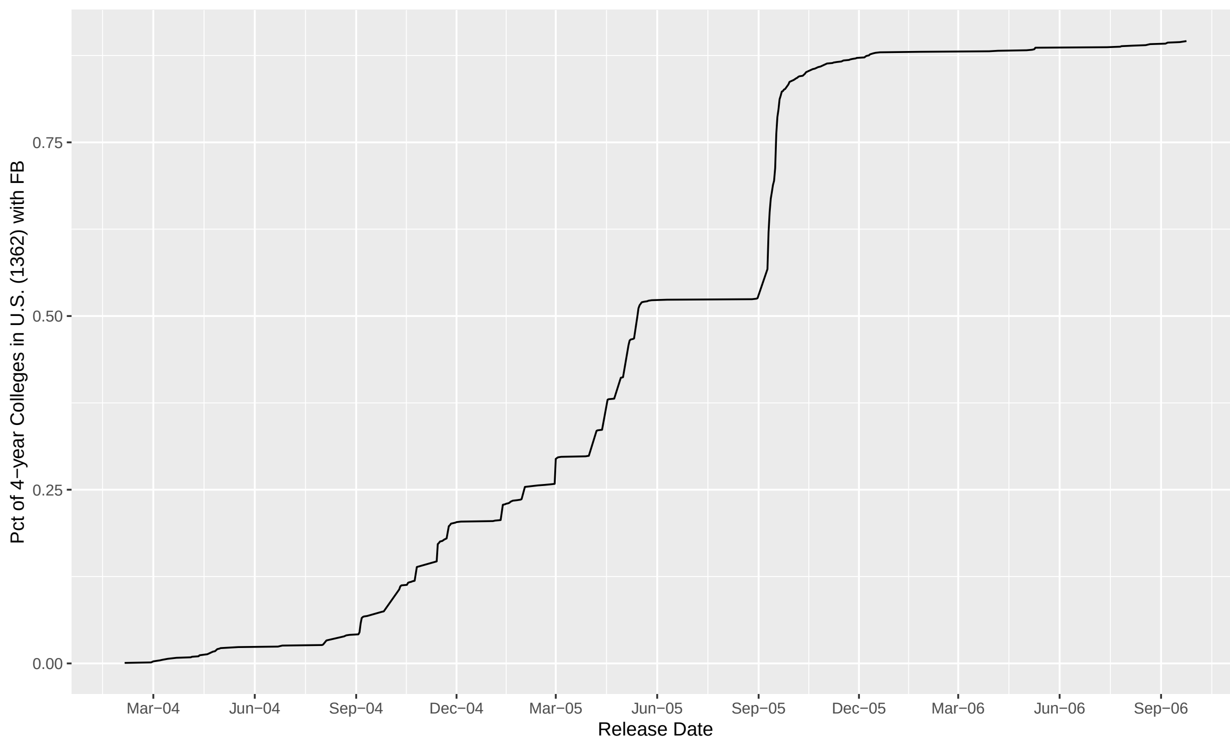
Table 5: Effect of Exposure to FB in College on Friendships

	Outcome: Log(Number of Friendships)		
	(1)	(2)	(3)
Years Exposed to FB in College	0.085^{***}	0.097^{***}	0.064^{***}
	(0.001)	(0.007)	(0.009)
College FE	Yes	Yes	Yes
Enroll * Grad * Outcome Year FE	No	Yes	Yes
Demographic Controls	No	No	Yes
Observations	13,193,959	13,193,959	13,193,959
R^2	0.065	0.121	0.178
Adjusted R^2	0.065	0.121	0.172

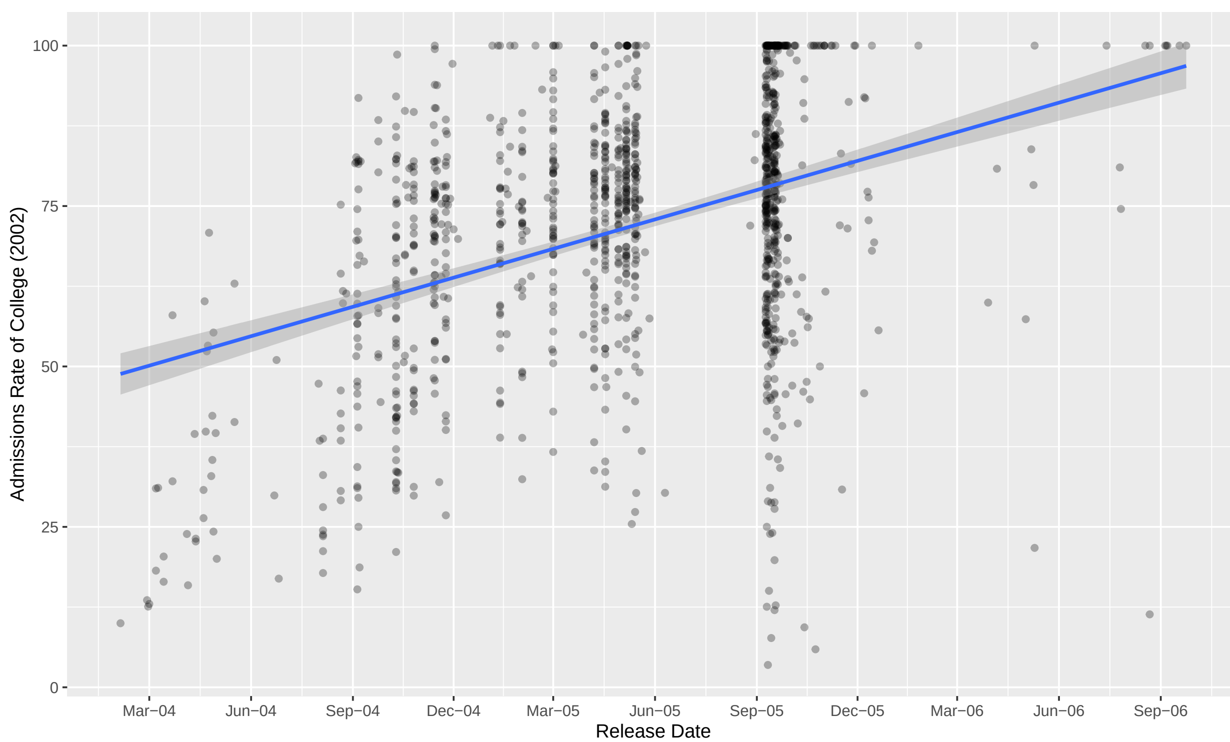
Note: *p<0.1; **p<0.05; ***p<0.01

	(1)	(2)	(3)	(4)	(5)
Outcome:	Log(Density)	Pct BA+	Log(Racial Diversity)	Log(Per Capita Income)	Poverty Rate (Pct)
Years Exposed to FB in College	0.120	0.526	0.014	0.008	0.351
	(0.010)	(0.115)	(0.002)	(0.003)	(0.039)
Observations	9,763,271	9,760,742	9,754,454	9,762,221	9,751,747
Adjusted R^2	0.316	0.189	0.266	0.191	0.096
College FE	Yes	Yes	Yes	Yes	Yes
Grad Year FE	Yes	Yes	Yes	Yes	Yes
Demographic Controls	Yes	Yes	Yes	Yes	Yes

Table 6: Effect on Neighborhood Characteristics of Zipcode Users Reside in (All Years)



(a) Distribution of College Release Dates



(b) Release Data vs Admission Rate of College

Figure 1: Rollout Schedule of Facebook to Colleges

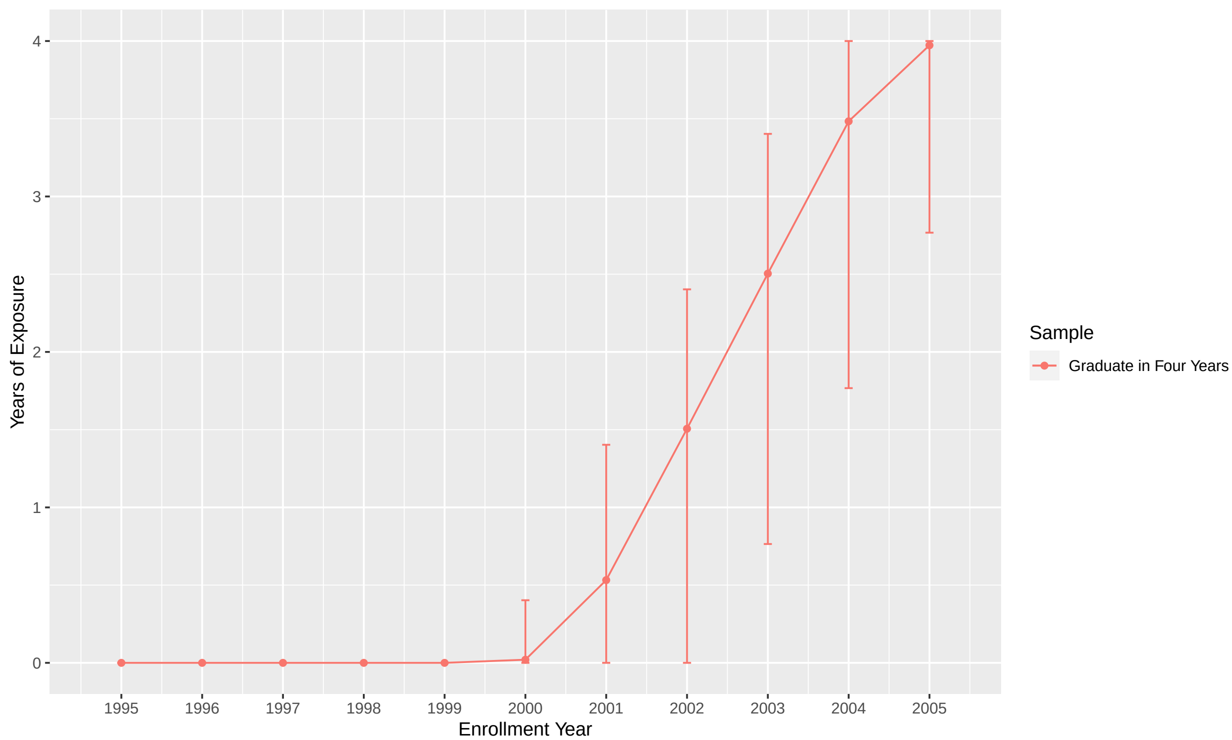


Figure 2: Average Access time During College and Range by Entry Year, Four-year Graduates

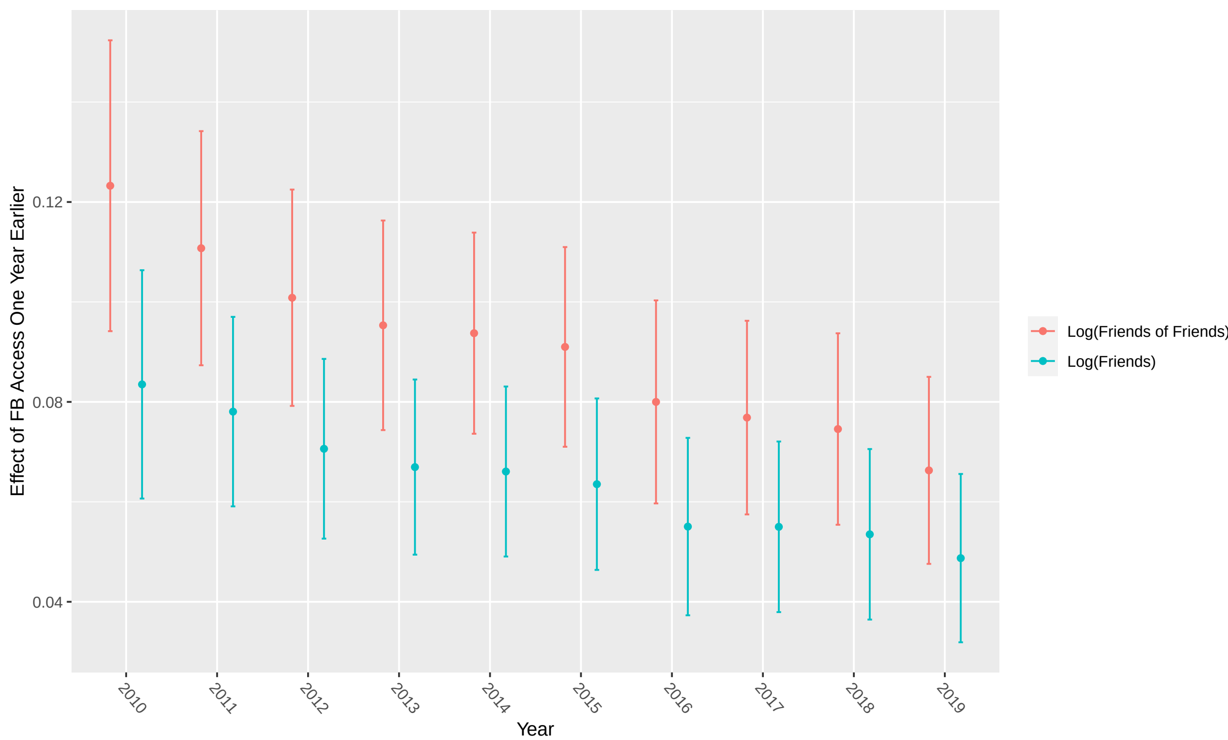


Figure 3: Effect of Exposure to FB in College on Total Friendships, by Year

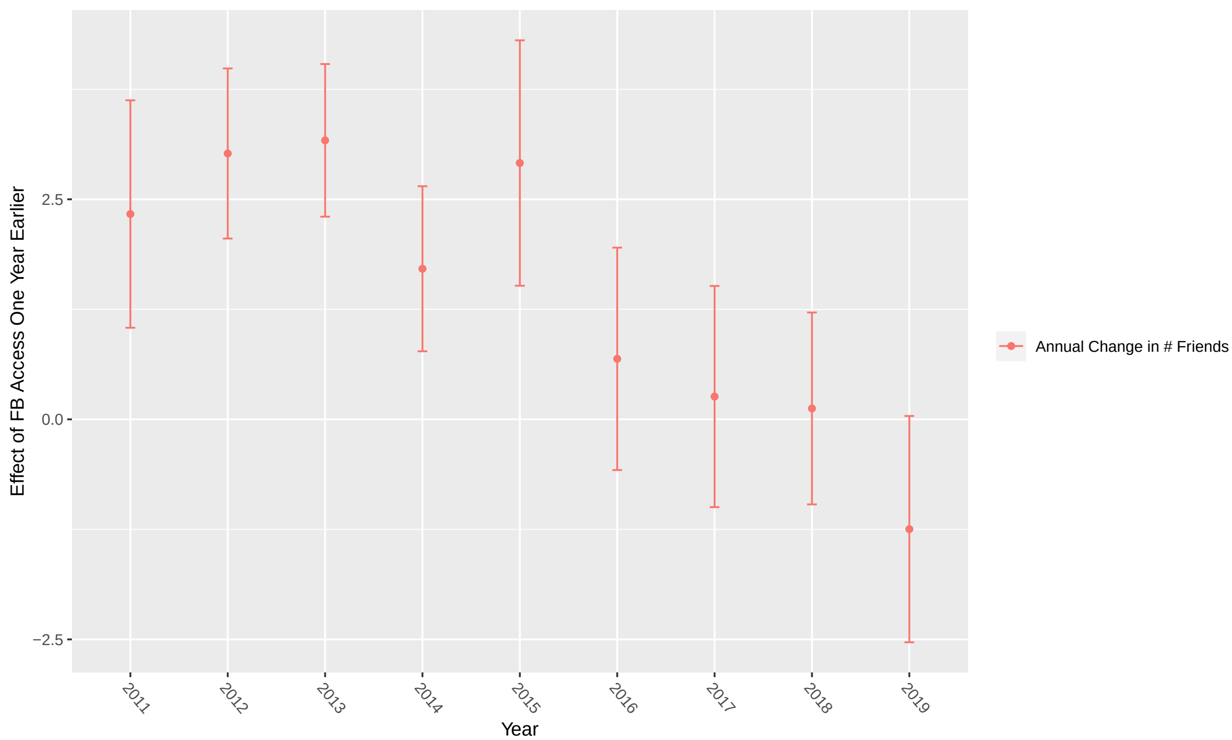


Figure 4: Effect on Annual Change in Friendships, by Year

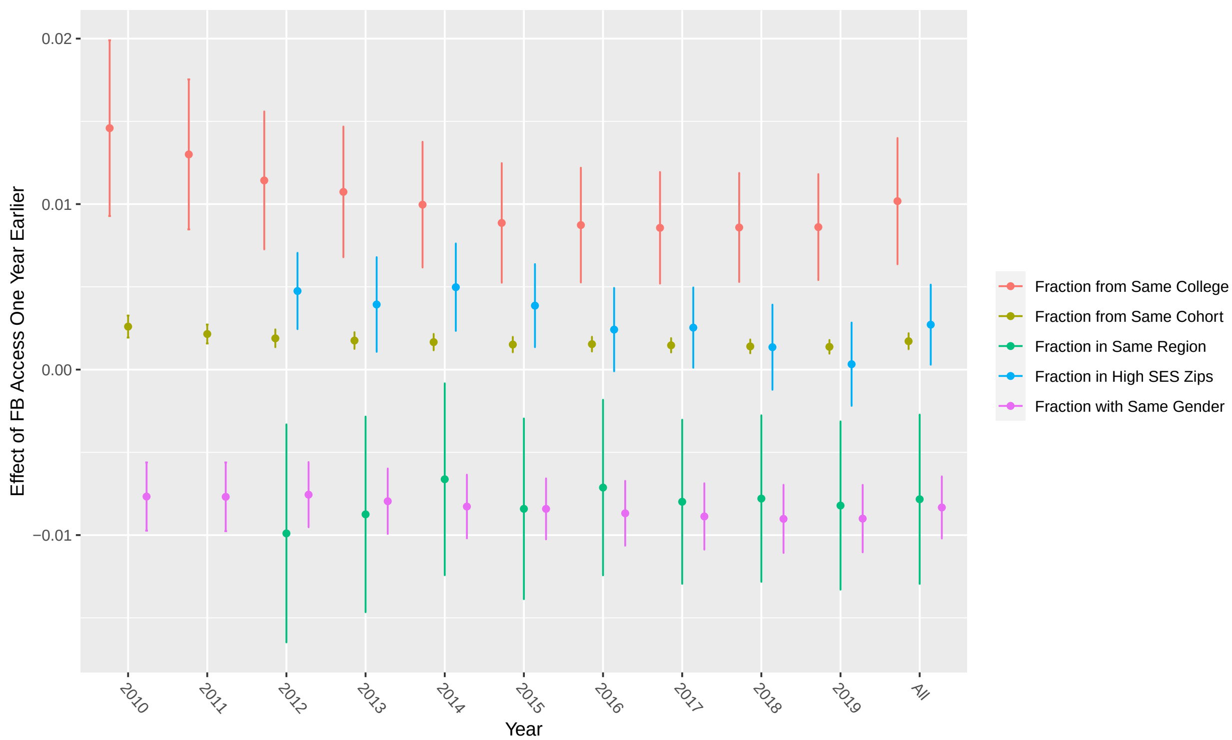


Figure 5: Effect of Exposure to FB in College on Network Homophily, by Year

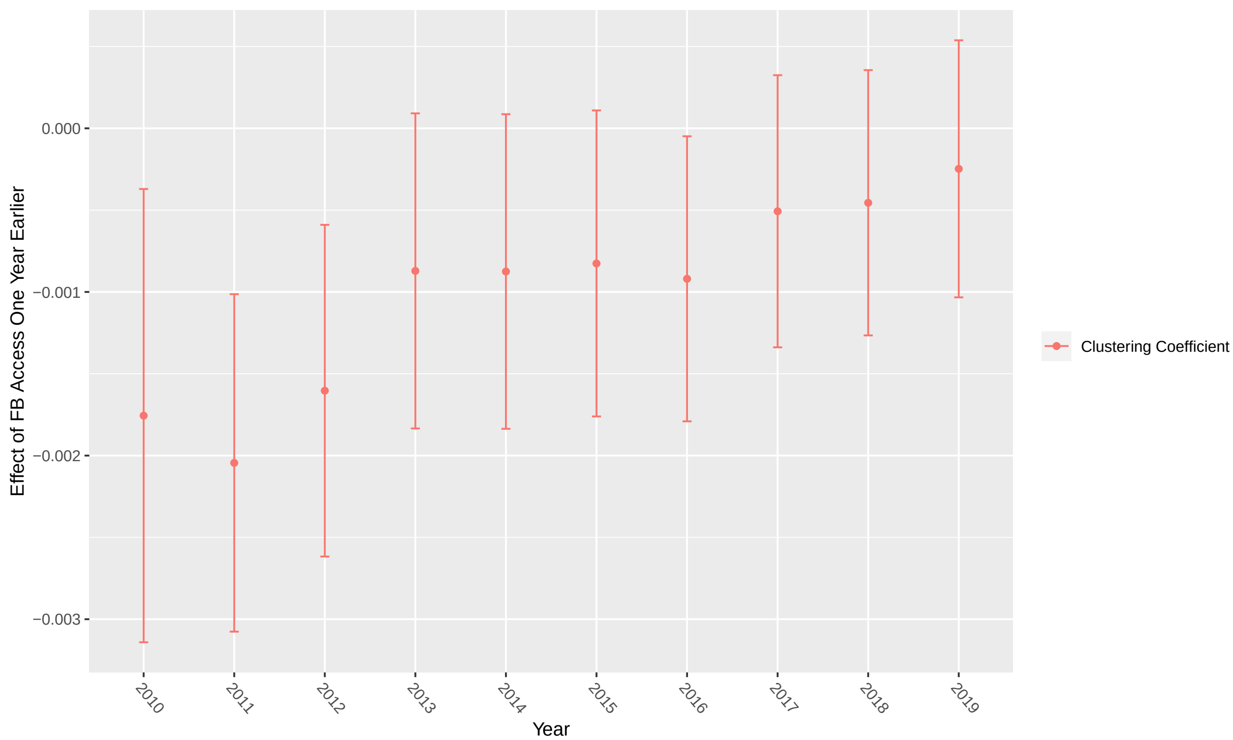


Figure 6: Effect of Exposure to FB in College on Network Clustering, by Year

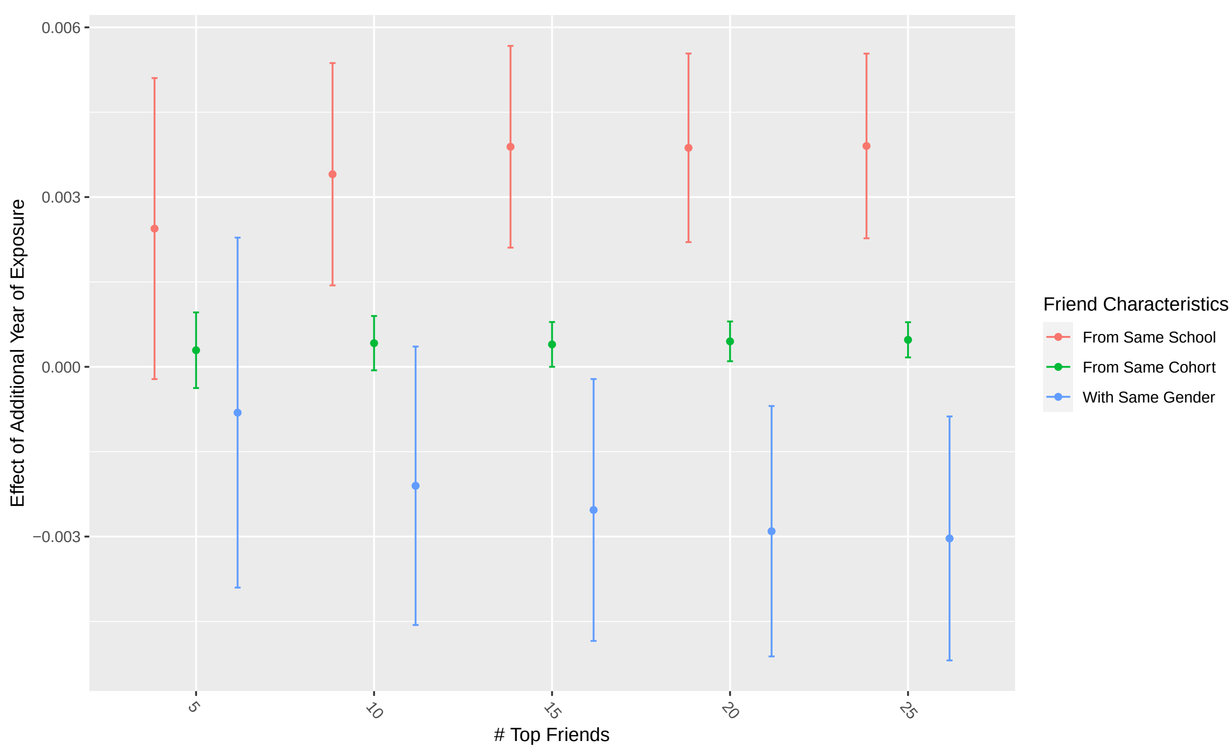


Figure 7: Effect on Composition of Close Friends in 2023, By Year

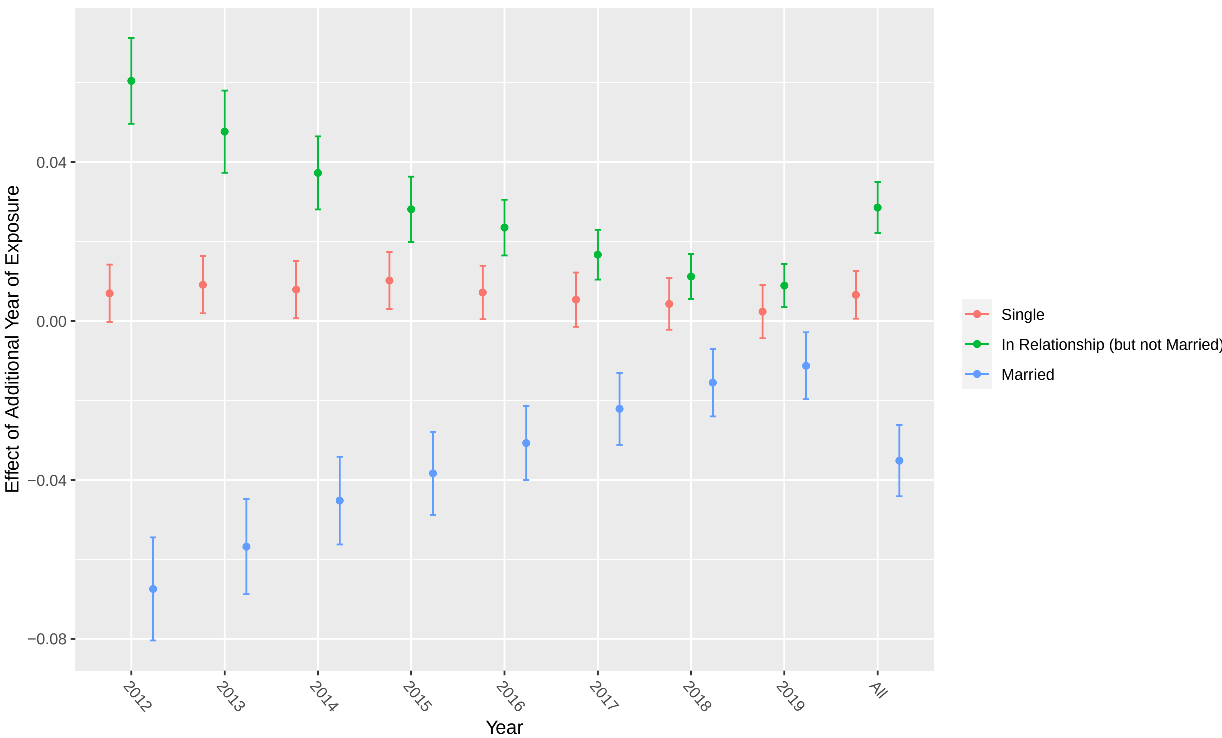


Figure 8: Effect on Relationship Status, By Year

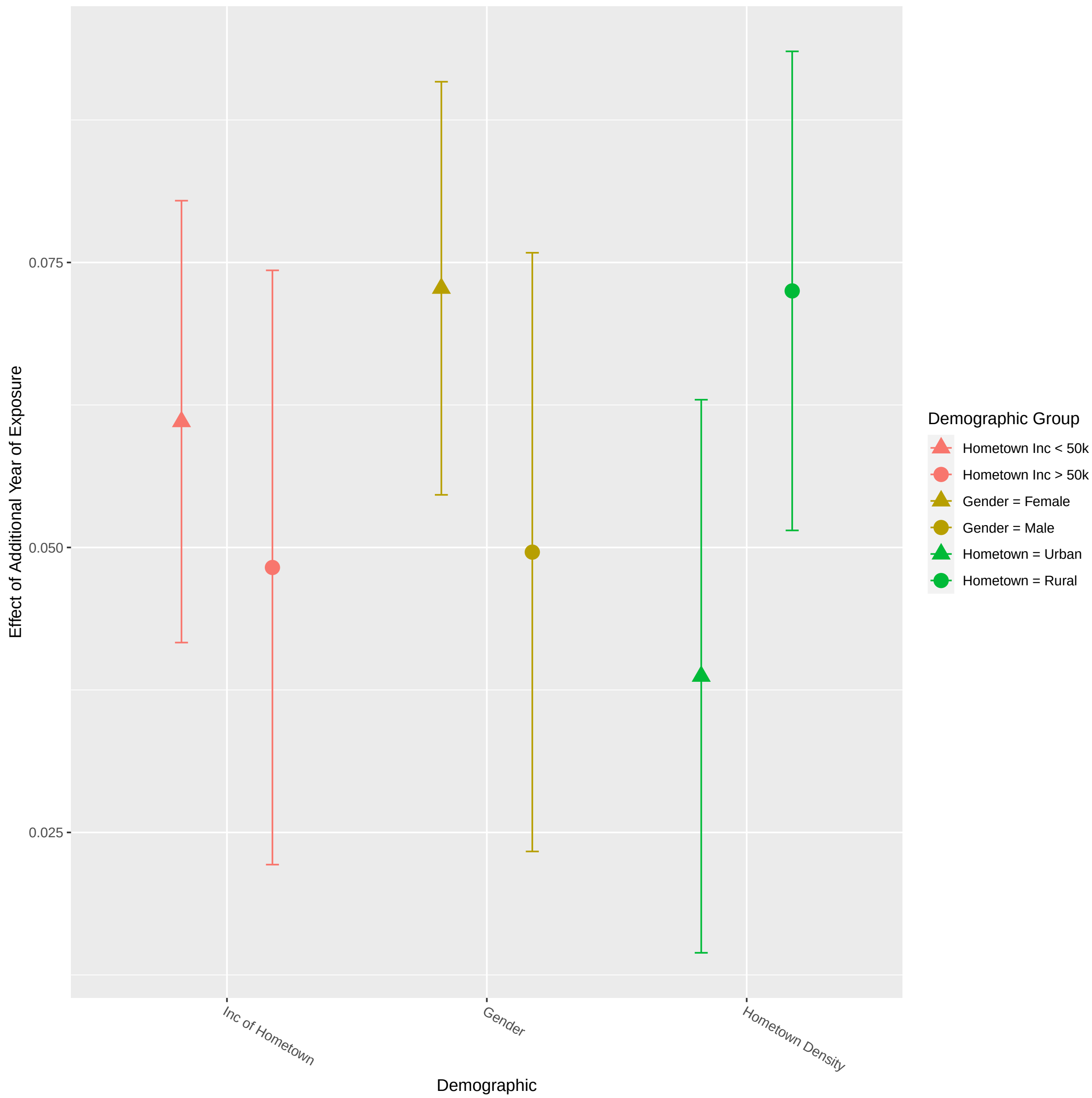
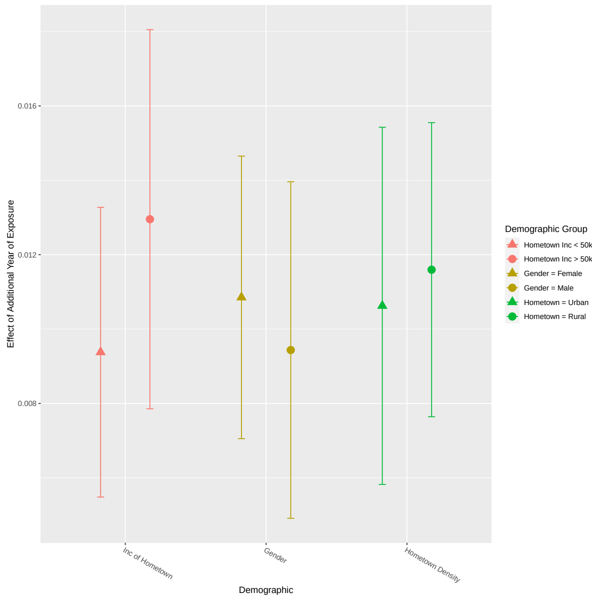
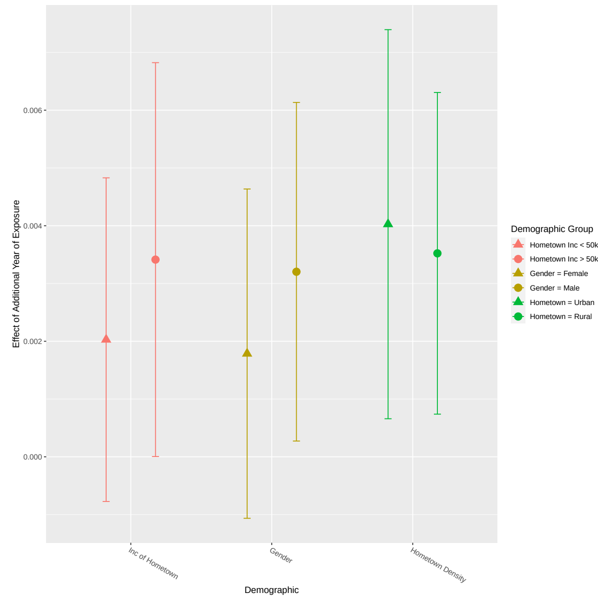


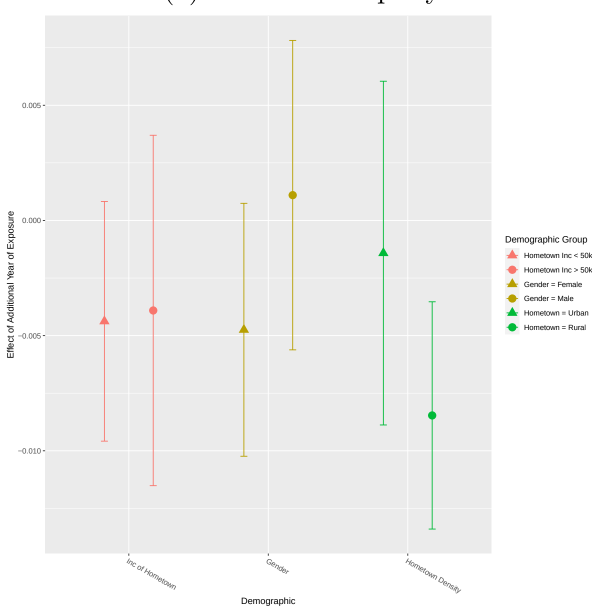
Figure 9: Heterogeneous Effect on Number of Friendships, by Demographics (All Years)



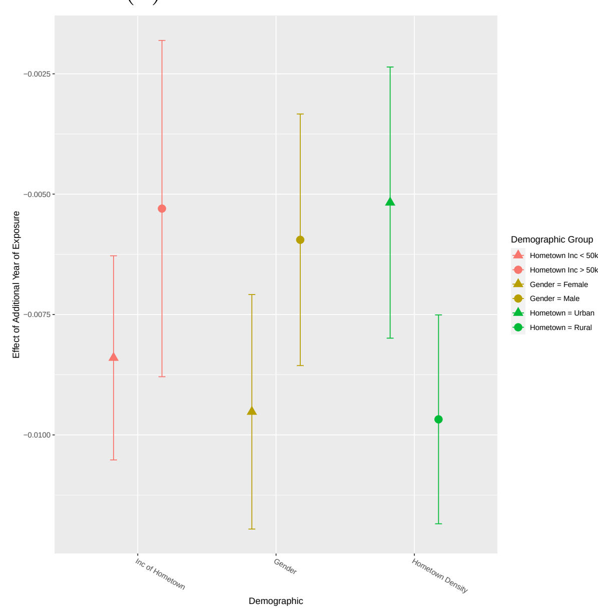
(a) School Homophily



(b) Economic Connectedness

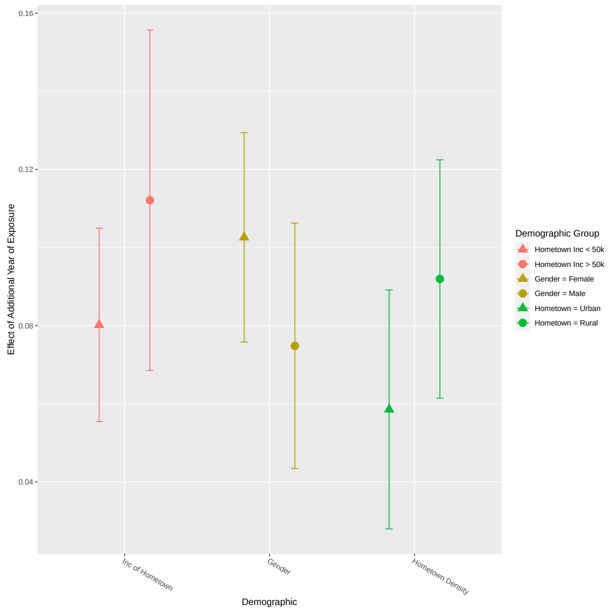


(c) Geographic Homophily

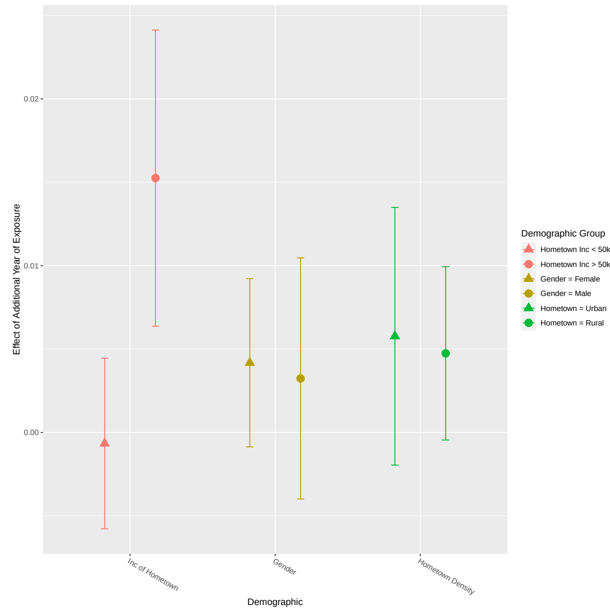


(d) Gender Homophily

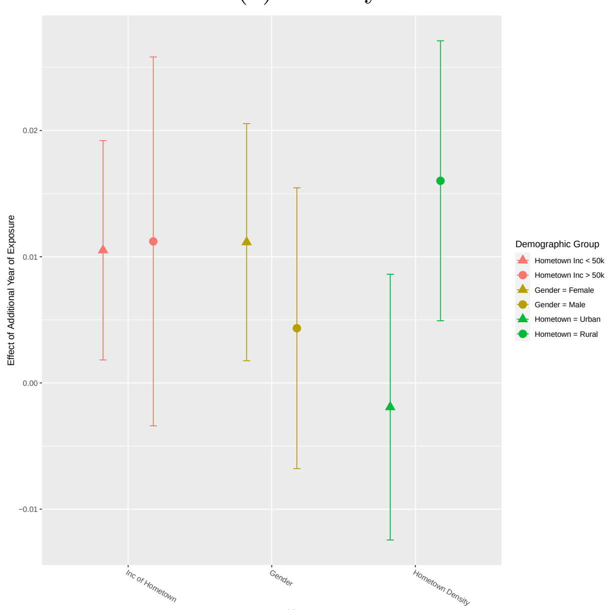
Figure 10: Heterogeneous Effects on Network Composition



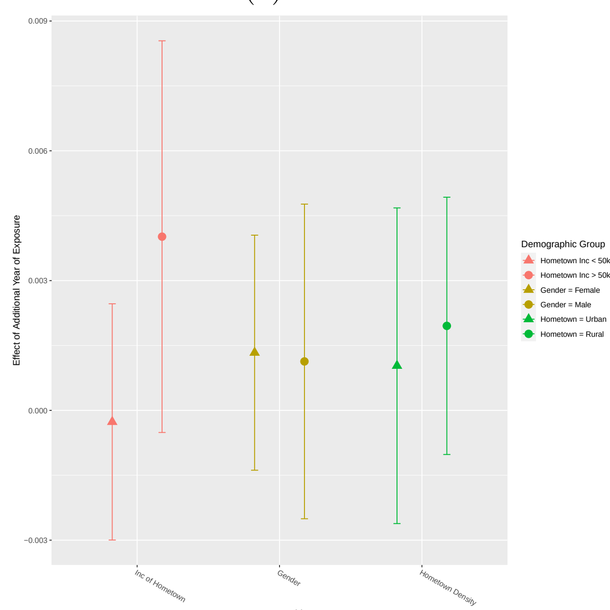
(a) Density



(b) Income



(c) Racial Diversity



(d) Education

Figure 11: Heterogeneous Effect on Characteristics of Zipcode Users Reside in, by Demographic (All Years)

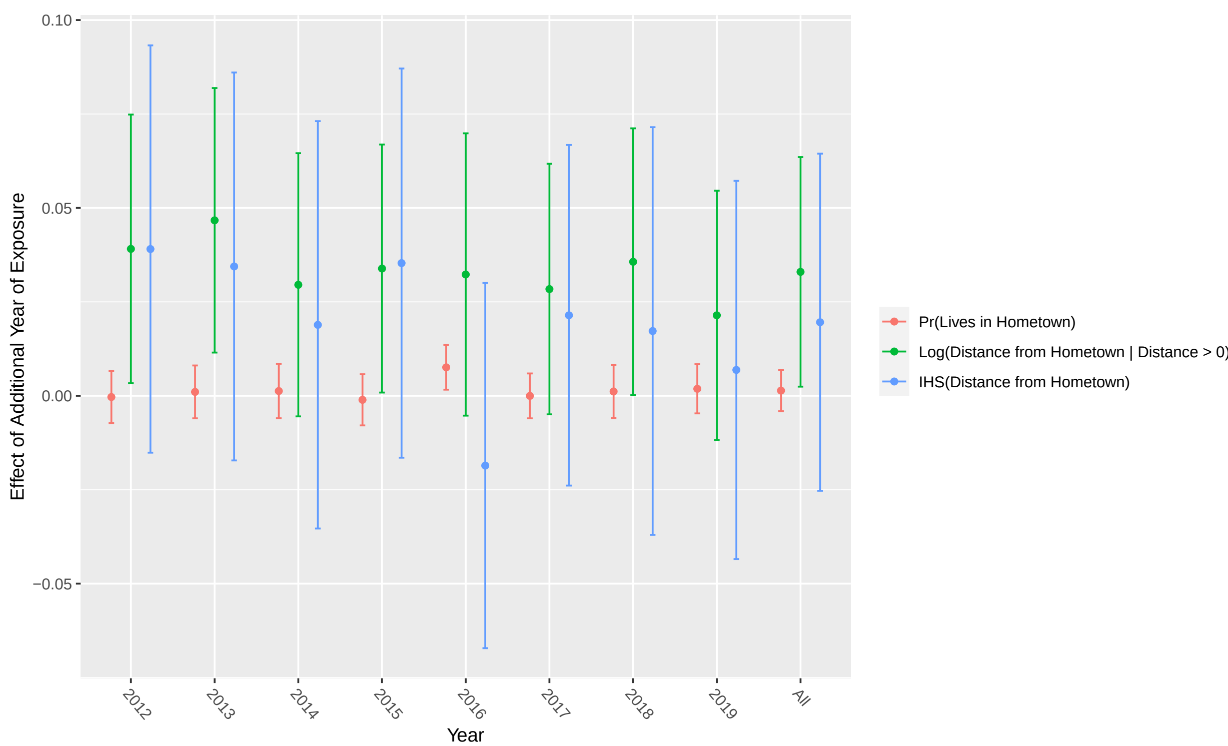


Figure 12: Effect on Distance in Residence from Hometown

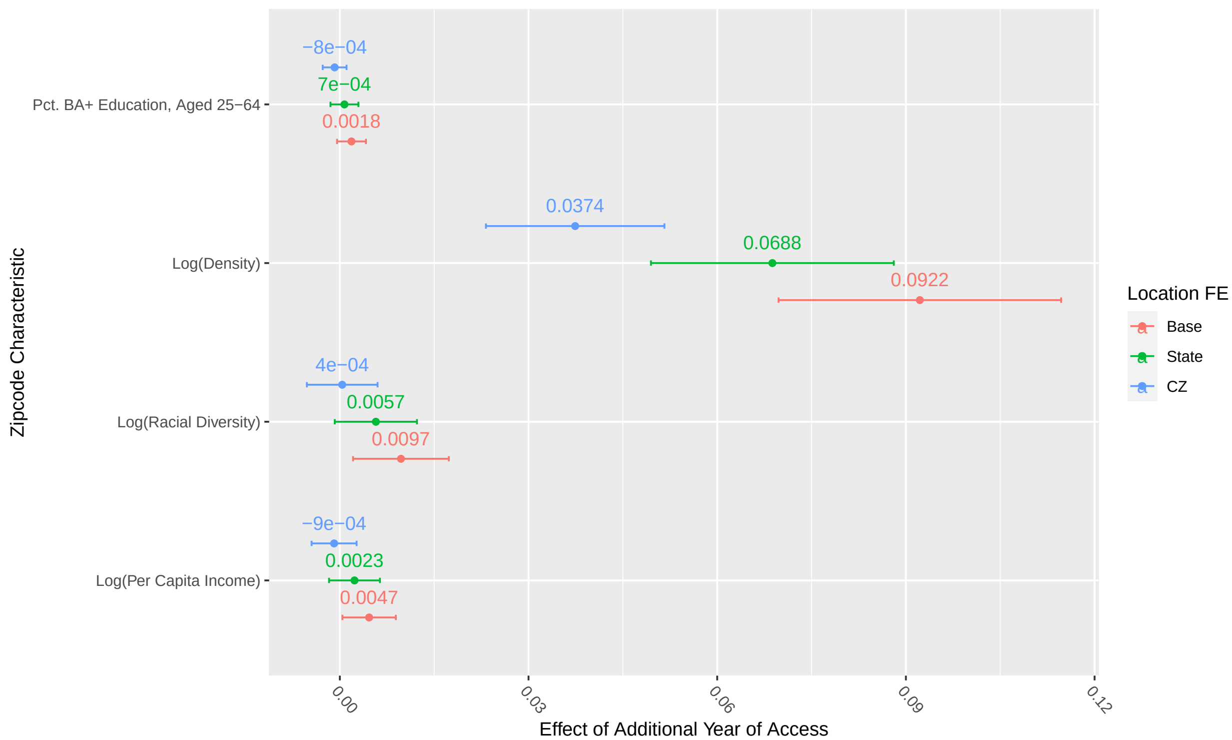


Figure 13: Effect on Zipcode Characteristics, Controlling for Region

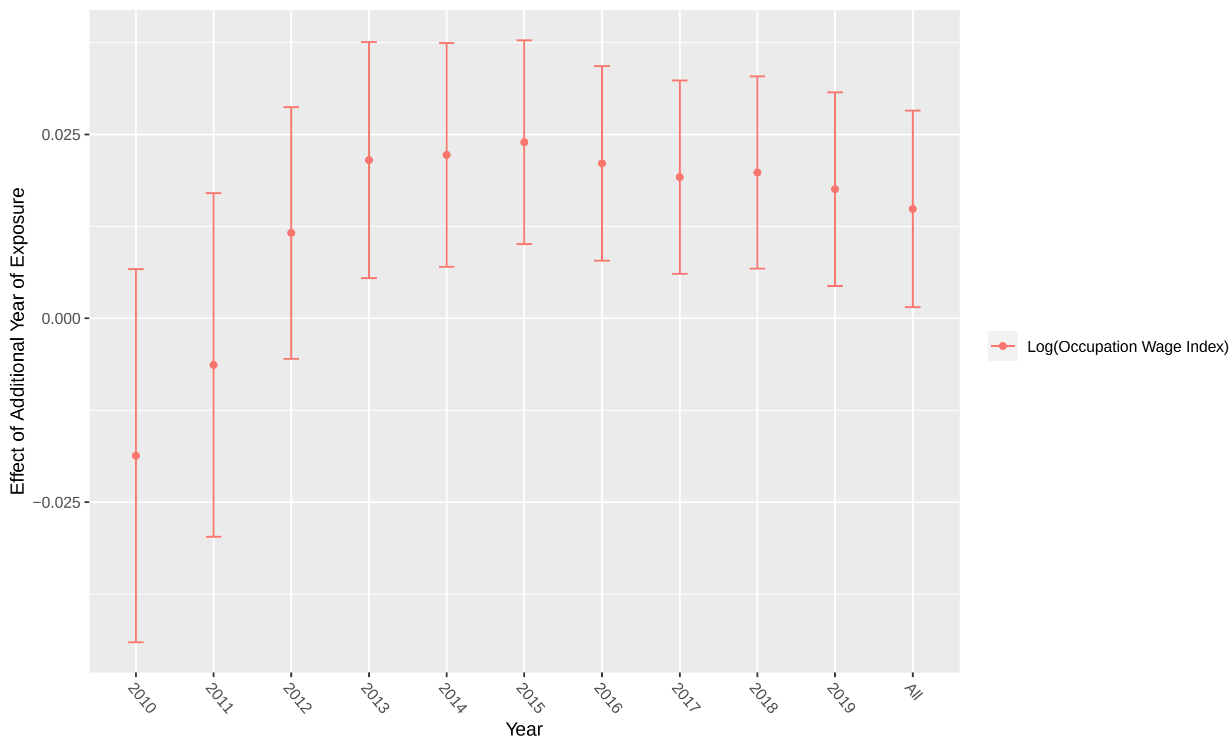


Figure 14: Effect on Occupation Index, by Year

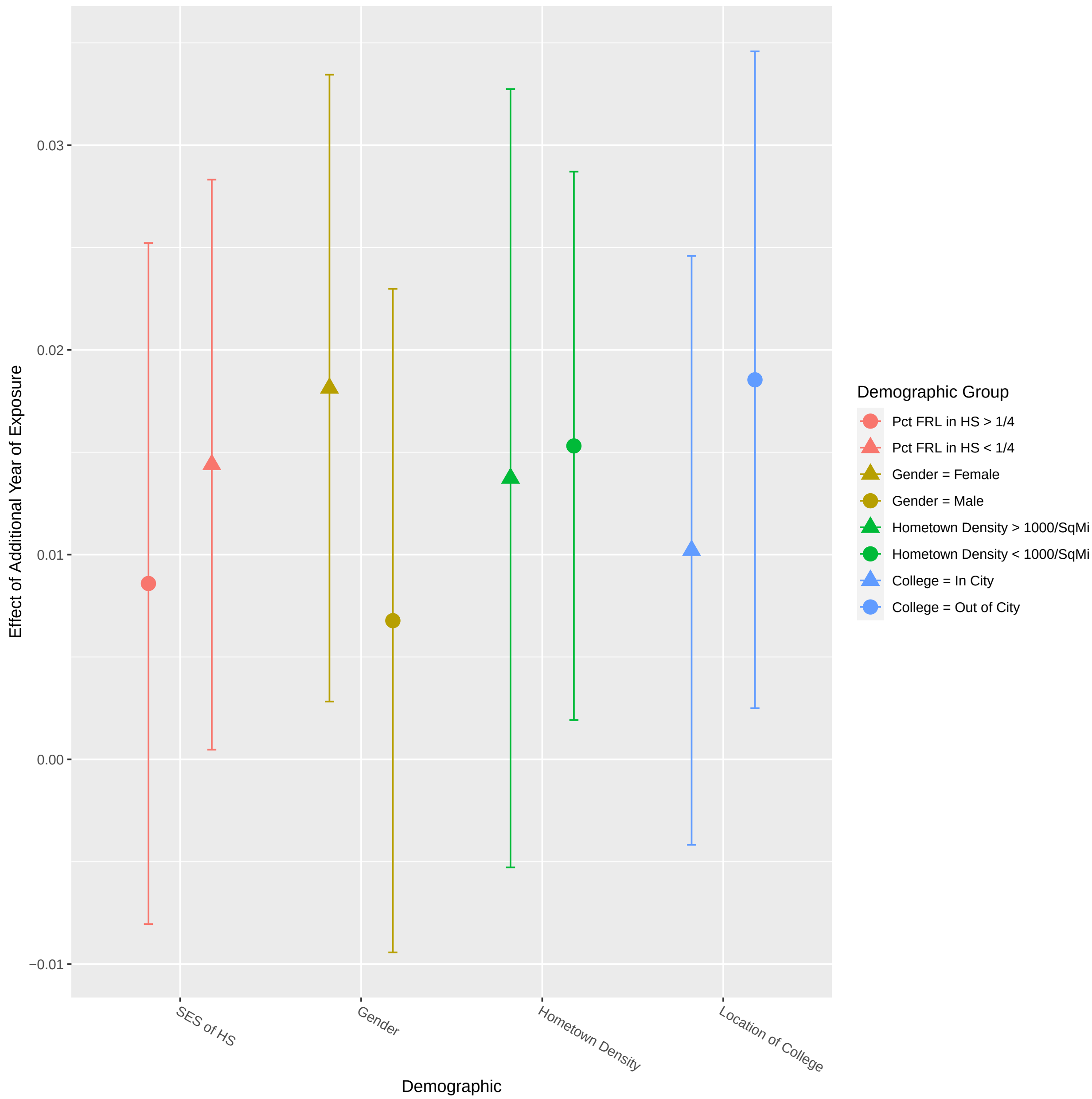


Figure 15: Heterogeneous Effect on Occupation Index (All Years)

Outcome	Estimate	t-value	Estimate in SD Terms
Pct From Same School	2.5218	17.541	0.206
Pct From Same Cohort	0.2424	13.185	0.165
Pct From Same CZ	-1.2577	-7.175	-0.052
Pct With Same Gender	-0.6040	-8.690	-0.049
Pct in High SES Zips	0.2835	3.572	0.015
Clustering Coefficient	-0.0014	-4.355	-0.027

Table A1: Effects on Network Composition in Terms of Standard Deviation

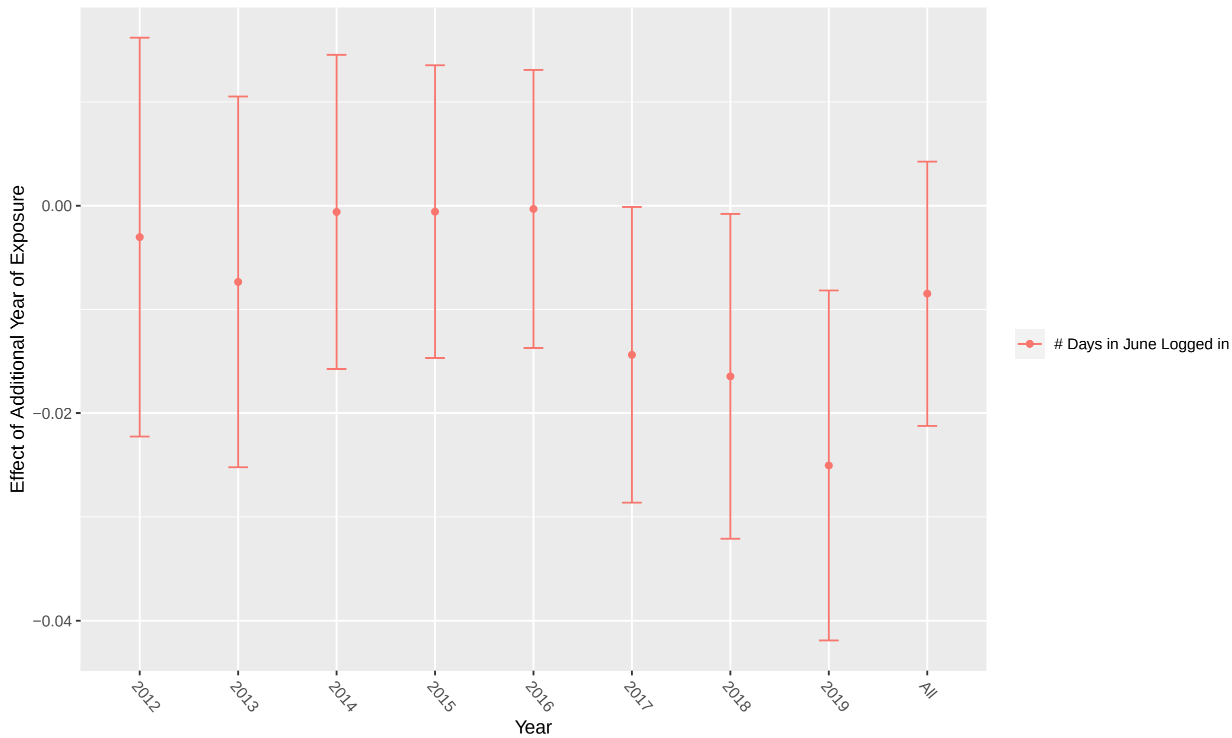


Figure A1: Effect of Exposure to FB in College FB Usage, by year

Figure reports the effect on the number of days logged in to FB in June of a particular year. Due to privacy reasons, we report coefficients in terms of standard deviations of the outcome (# days logged in) over the entire sample.

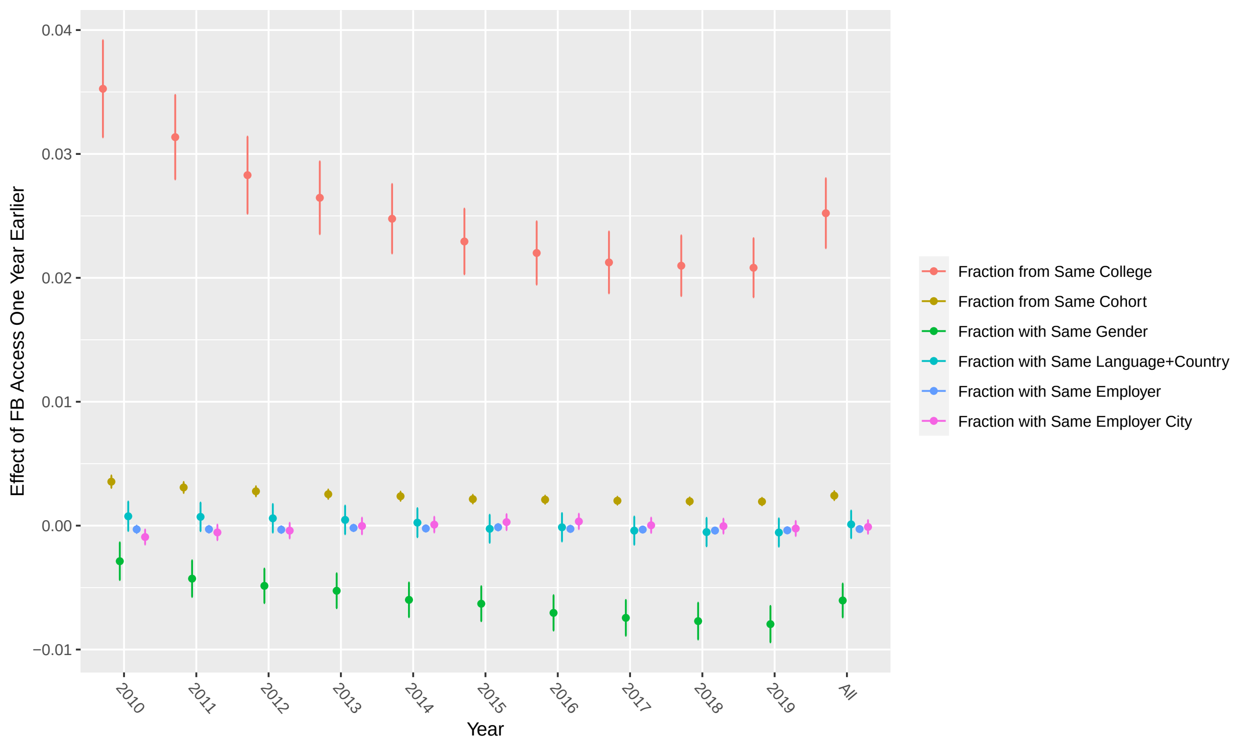


Figure A2: Effect of Exposure to FB in College on Demographic Homophily, by Year

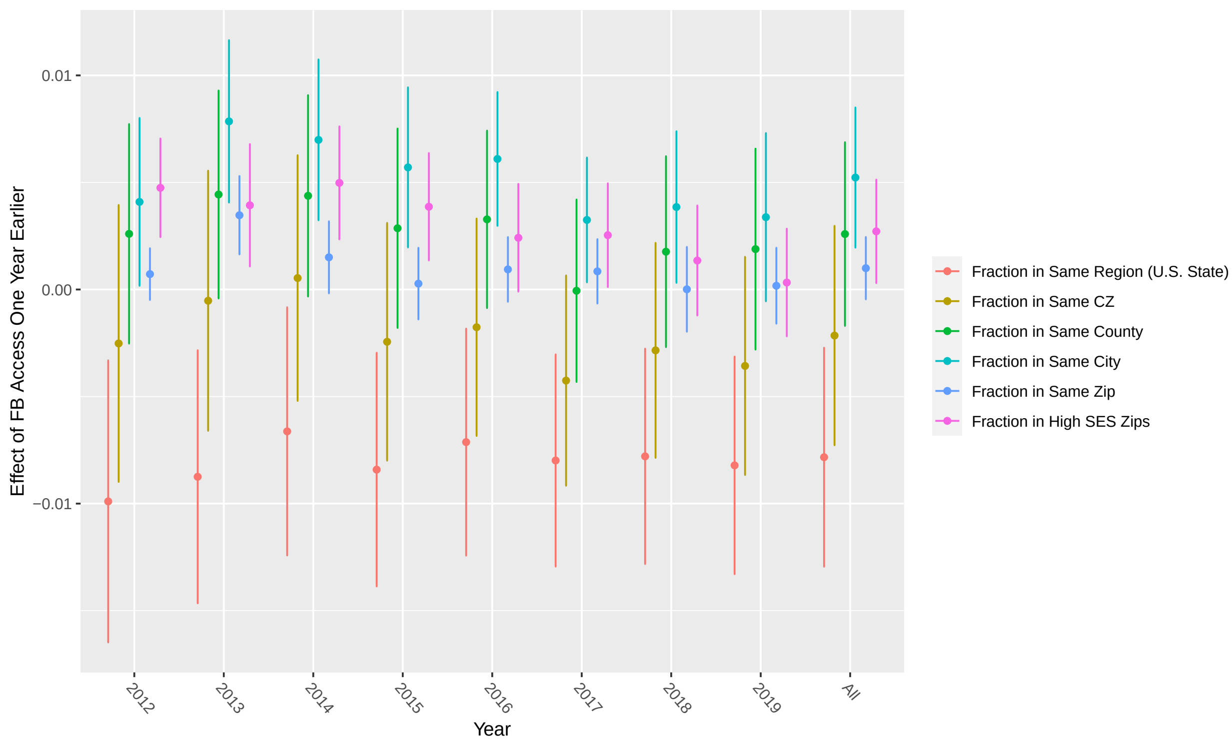


Figure A3: Effect of Exposure to FB in College on Geographic Homophily, by Year

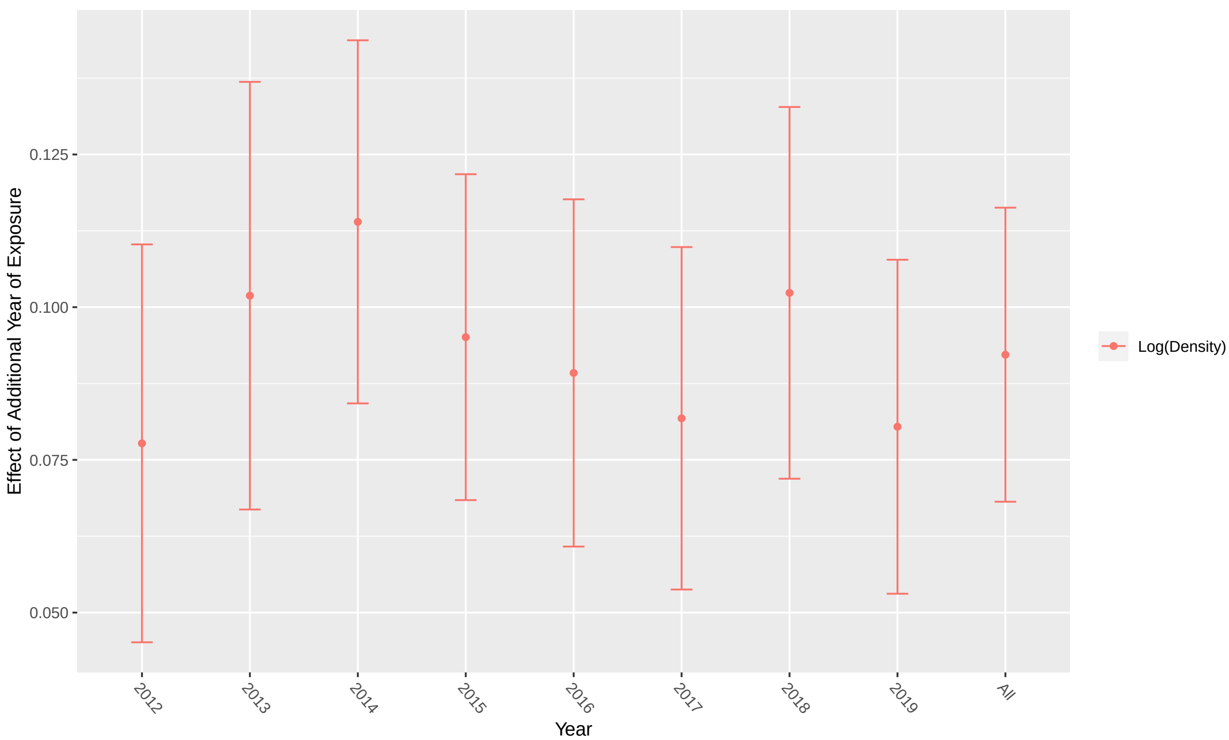


Figure A4: Effect on Zipcode Density, by Year

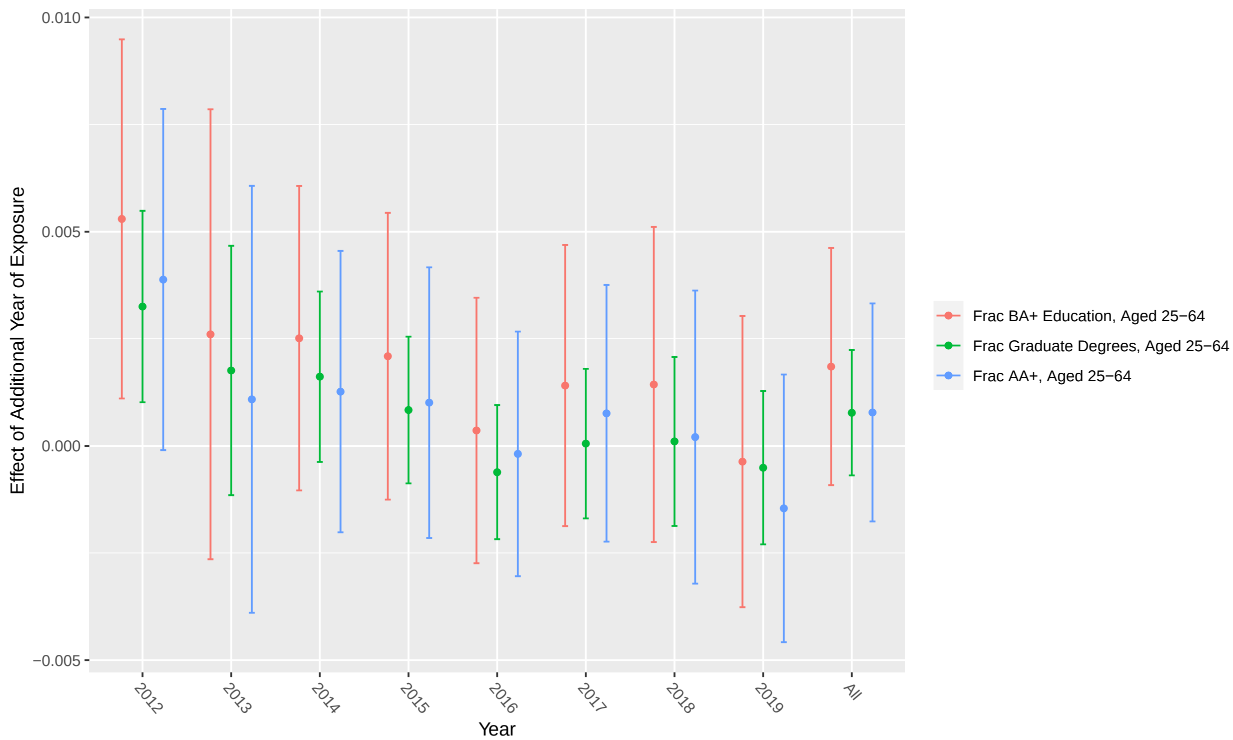


Figure A5: Effect on Zipcode Education Composition, by Year

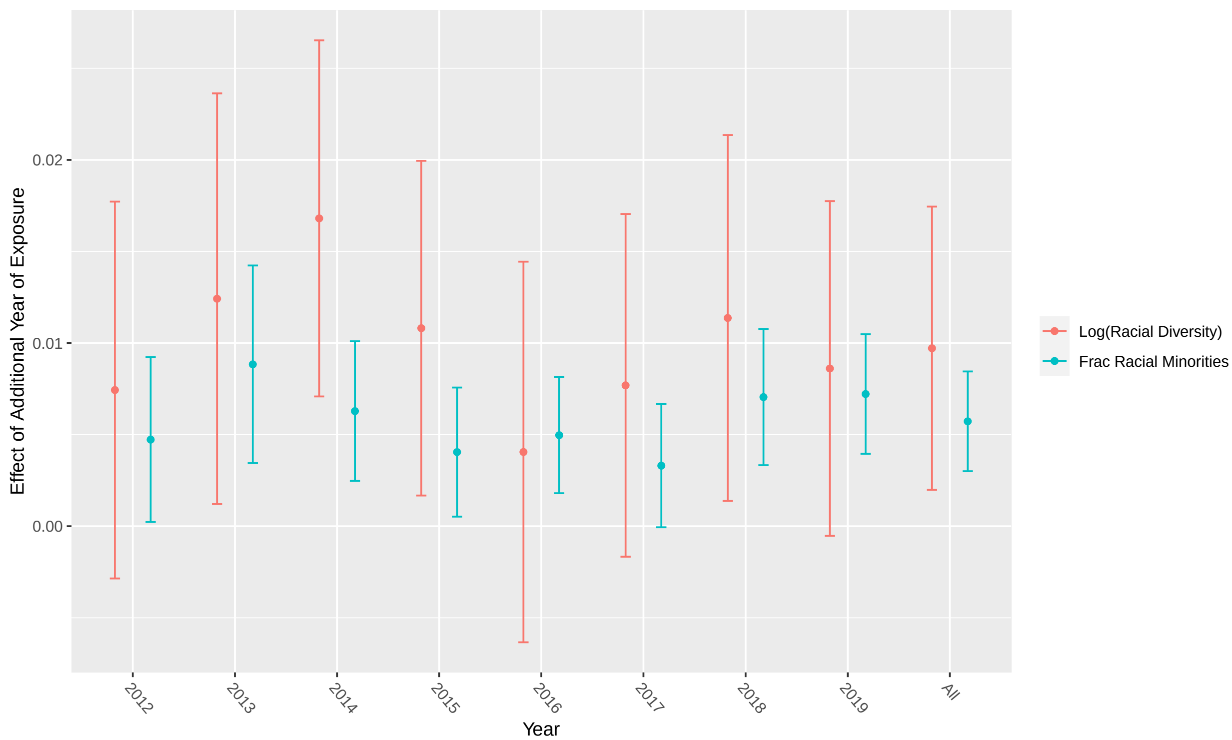


Figure A6: Effect on Zipcode Racial Composition, by Year

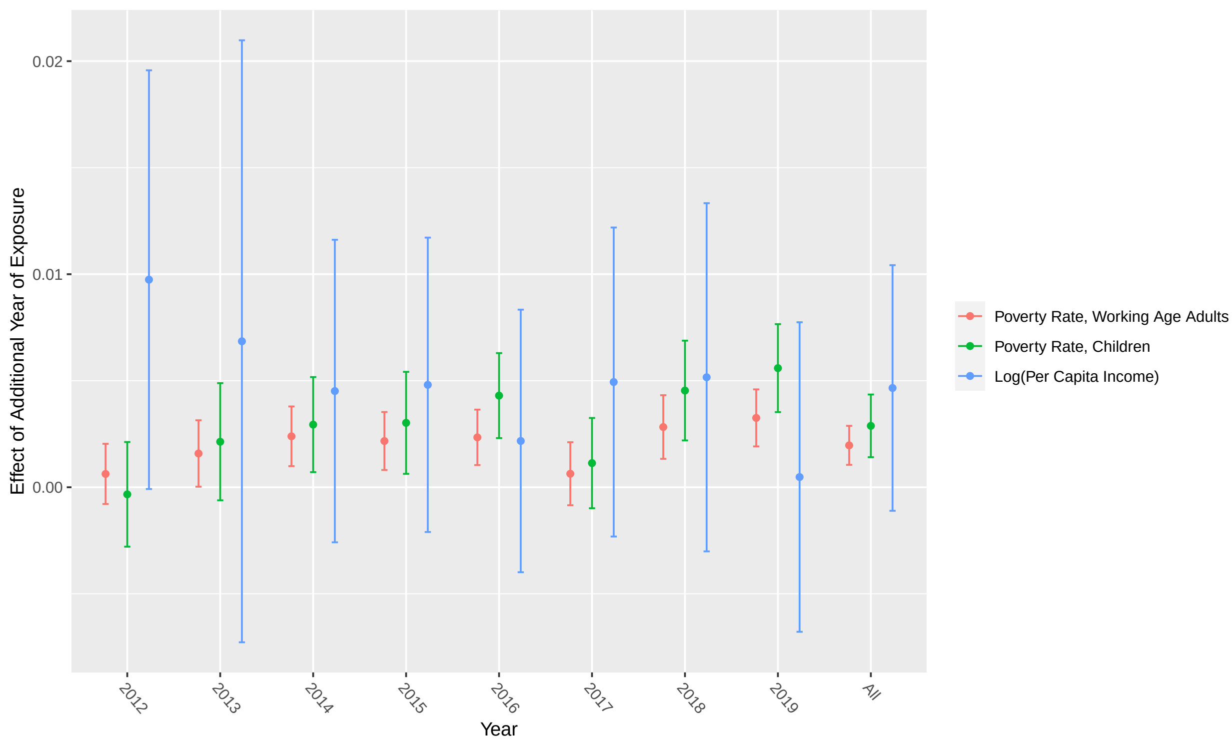


Figure A7: Effect on Zipcode Income Composition, by Year

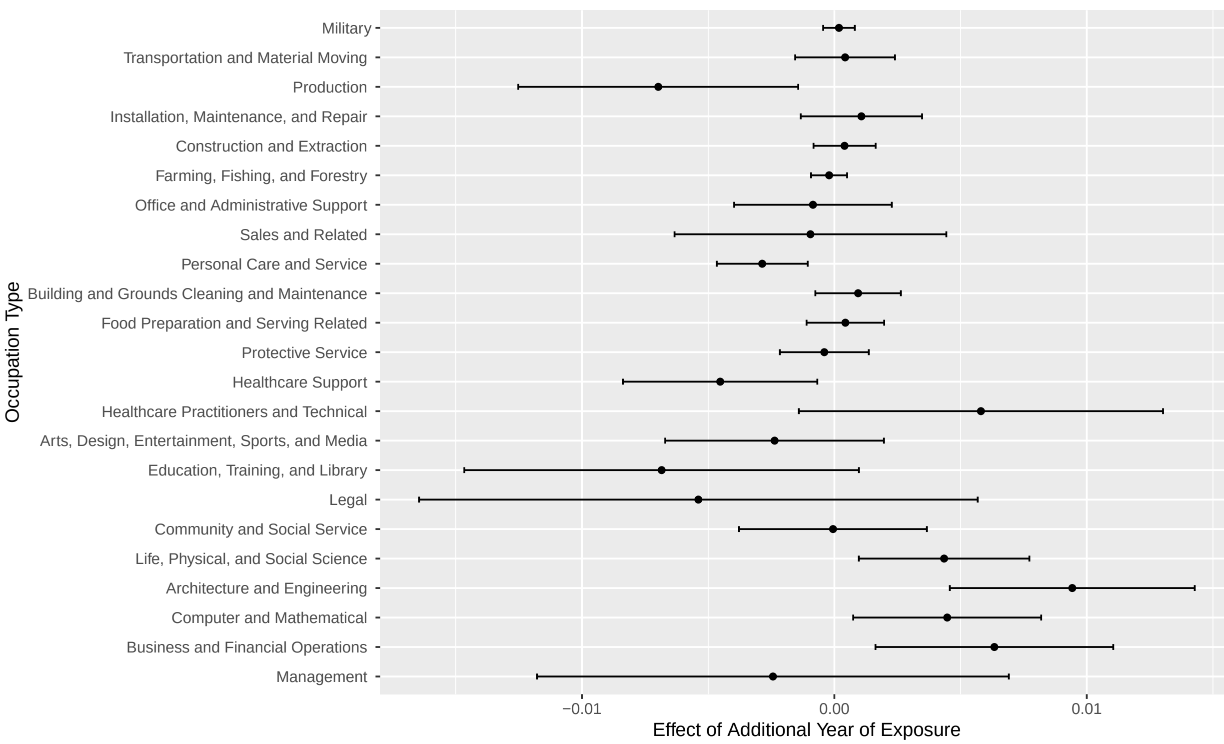


Figure A8: Effect on Occ Sorting

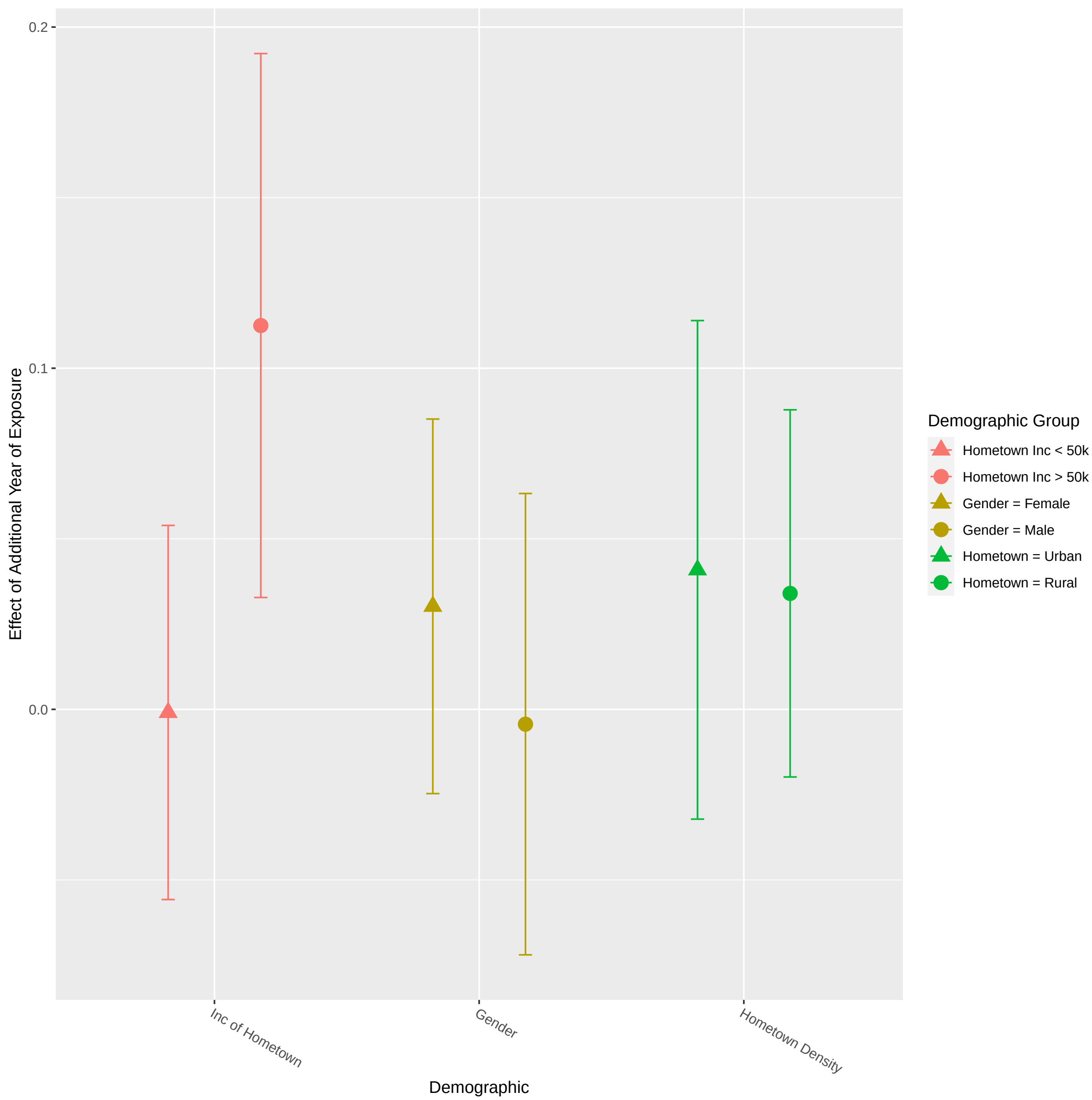


Figure A9: Heterogeneous Effect on Distance From Hometown, All Years

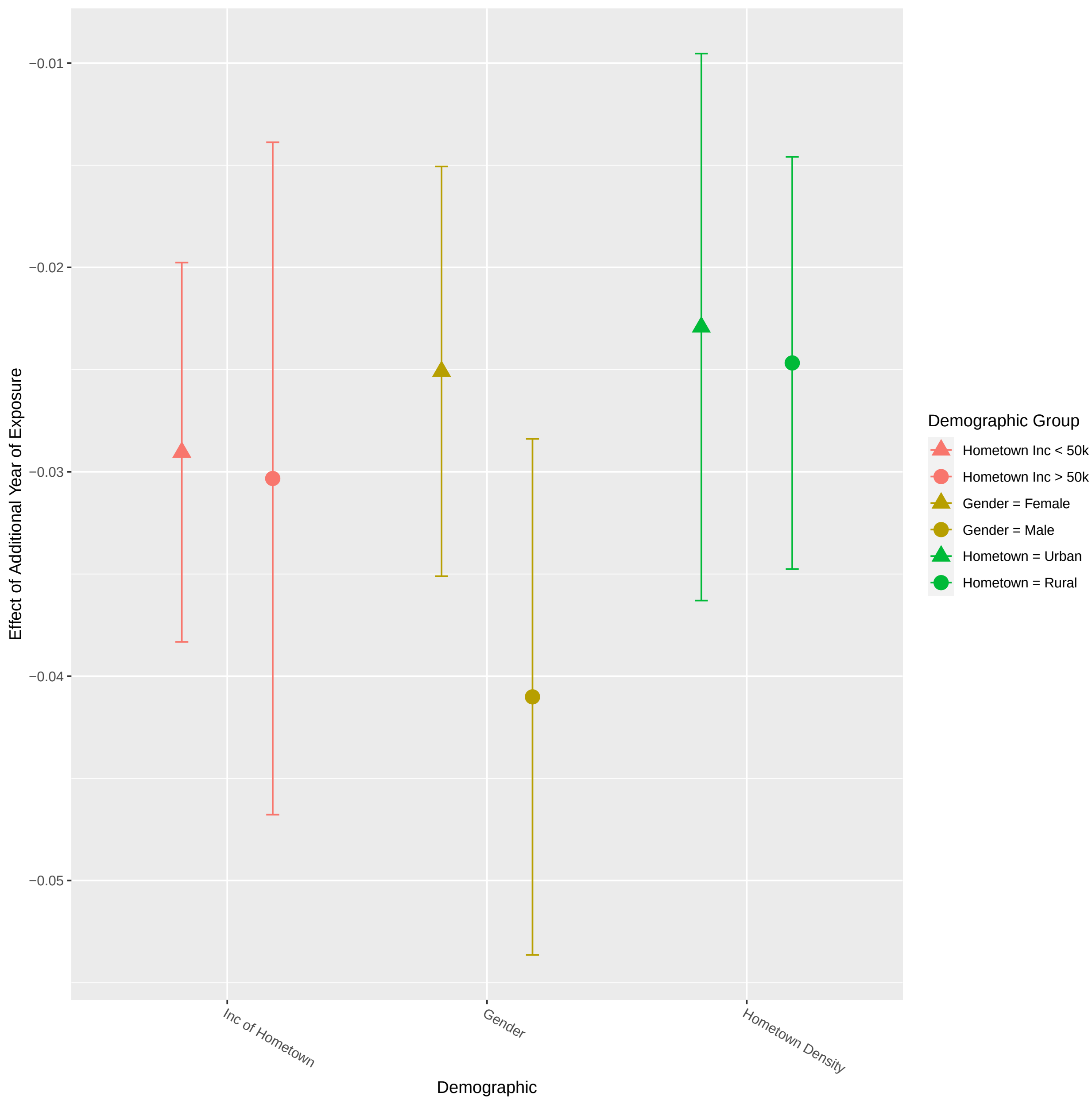


Figure A10: Heterogeneous Effect on Marriage Probability, All Years