Online Social Network Effects in Labor Markets: Evidence From Facebook's Entry to College Campuses

Luis Armona*

August 16, 2021

Abstract

Using quasi-random variation from Facebook's entry to college campuses, I exploit a natural experiment to estimate the effect of online social network access on future earnings. My estimates imply that access to Facebook for an additional year in college causes a .61 percentile increase in a cohort's average earnings, translating to an average wage increase of around $970 in 2014. My results also suggest that Facebook access decreases income inequality within a cohort. I provide evidence that wage increases comes through the channel of increased social ties to former classmates, which leads to strengthened employment networks between college alumni.

Keywords: social networks, labor economics, education, Facebook, technology, online platforms

JEL Codes: J2,, I2, J3.

---

*Department of Economics, Stanford University. Email: larmona@stanford.edu I am thankful to Mike Bailey, Nick Bloom, Sarah Eichmeyer, Matthew Gentzkow, Matt Jackson, and Rose Tan for their helpful comments and feedback on this paper. I am also particularly thankful to Luca Braghieri for aiding me in scraping Facebook release data used in this paper, in addition to his feedback on this paper.

1 Introduction

Few Internet websites that exist today are as ubiquitously known as Facebook.com. Facebook is the largest social networking site (SNS) in the world today, with over 2 billion users worldwide. As of January 2018, 68% of all adults in the United States are active Facebook members, and 50% of all Americans use this website every day.[1] It provides users with a low-cost way to instantly keep in touch with friends and acquaintances, even if they are thousands of miles away.

It is well documented that "weak ties", relationships with those who are informal acquaintances, are key in providing individuals with employment opportunities Granovetter [1973]. Facebook as a platform offers an easy-to-use, widely adopted electronic social network for people to remain in contact with others they have previously met. When asked about the success of Facebook in 2004, founder Mark Zuckerberg supported this notion, saying: "We think we have been particularly successful in strengthening those relationships that exist between people who are only 'fringe friends.'"[2] Facebook should decrease the costs of maintaining weak ties, since users can communicate without having to track down an address that may no longer be valid.

This paper examines the effects of student access to Facebook in the 760 selective colleges that received Facebook before May 2005. Facebook's entry to colleges in my sample occurred over a period of 15 months, which created sharp differences in exposure to Facebook across college students in the United States, both by the university they attended, and the year students entered college. I use the differential timing of access to Facebook between cohorts at the same school to identify the effect of Facebook access on long term labor outcomes, earnings measured in 2014 (5-12 years after exiting college). My empirical approach is to control for endogeneity concerns via two-way fixed effects. I argue this allows me to identify the causal effect of Facebook access on future labor outcomes.

At the time of Facebook's introduction to colleges, the few other large SNSs available, Friendster and Myspace, experienced significant declines shortly after. Friendster began a precipitous decline in 2009 when a technical issue on the website caused Friendster to lose an estimated 80% of its users, leading to its closure in 2011 [Garcia et al., 2013]. By the fall of 2010, 52% of Americans interacted with Facebook daily, while this number was just 8% for Myspace.[3] Other SNSs that are now a part of everyday life were still gaining traction. LinkedIn, a professional networking site launched in 2003, was used daily by only 6% of Americans in 2010.[23]. This period of Facebook dominance in social networking suggests that those who were able to sign up for Facebook during its introduction to colleges would have been fortunate, in the sense that friendships added to Facebook early on would likely still be accessible years later. Variation in Facebook access during this time period would be akin to variation in long-term access to any electronic social network.[4] Moreover, Facebook's introduction constituted a network-level shock, to each social network within a college campus, which allows me to more plausibly construct a treatment and control group for estimating the effect of social network strength on future outcomes.

The diffusion rates of Facebook during its entry were substantial. Jacobs et al. [2015] look at adoption rates among active students at the first 100 schools that received access 1-2 years after the introduction of Facebook, and find adoption rates from 80-100% of the student population. Using the same data as in Jacobs et al. [2015], controlling for entry year and university fixed effects, I estimate that an additional year of access to Facebook at college at the first 100 schools (known as the "Facebook 100") is associated with a 4.3% increase in students creating a Facebook account, and a 23% increase in the number of Facebook friends.

I restrict my attention to the intent-to-treat effect of access to Facebook on labor outcomes, with the understanding that Facebook access is a meaningful proxy for actual adoption. I estimate that receiving an additional year of access to Facebook while in college causes a .4% increase in a cohort's employment rate, and an increase in average wages of .6 percentiles, which translates to a $970 increase in annual earnings in 2014. I supplement this with evidence that students with more Facebook access were more likely to co-sort into employers, as measured on LinkedIn in 2018. In Appendix A.4, I include one instrumental variable regression that establishes a relationship between Facebook network strength in 2005 and future employment outcomes in 2014, though these results should be interpreted with caution since the exclusion restriction may not hold. In Appendix B, I provide a stylized model of job acquisitions via social networks that highlights the economic intuition for why Facebook access can increase earnings among college graduates. Altogether, these results suggest that the returns to online social networking technology are economically meaningful on the labor market. My finding that access time explains Facebook adoption means that this natural experiment could be used to estimate the causal effects of online social networking technology on other outcomes. For this reason, I share the full list of Facebook release dates at college campuses that I have collected for this paper.[5]

This paper relates to both the literature on the structure of the Facebook social network, and the literature on social network effects on individual outcomes. Traud et al. [2011, 2012] examine the structure of the Facebook social networks at the first 100 schools Facebook was released. They document high clustering on observables such as college dorm and major, and short diameters in these networks, consistent with the general attributes of social networks described in Jackson [2010]. Ugander et al. [2011] perform a large-scale study of the entire Facebook social graph in 2011, and confirm that these attributes of local density / low distance persists in the structure of the Facebook network once it was a widely adopted platform. Jacobs et al. [2015] looks at the Facebook network structure at Facebook 100 schools, in order to identify global trends in network formation across these schools. They exploit the same differences used in this paper in Facebook's release date by college to determine whether the networks evolve over time in a consistent fashion. Overgoor et al. [2020a,b] use Facebook data to examine the structure and dynamics of college networks over time. Mayer and Puller [2008] estimates a structural model of friendship formation in college using Facebook data to demonstrate the inefficacy of university intervention to limit segregation in student social networks . Enikolopov et al. [2020] uses an identification strategy similar to mine that shows access to a SNS in Russia increased participation in collective action.

Regarding more general social network effects on labor outcomes, Simon and Warner [1992] is one of the first papers to examine the implications for labor outcomes when they are based on social network structures and job referrals. They find that workers acquiring jobs through social ties are given higher initial salaries and are matched to jobs for longer. Laschever [2013] finds that post-war employment outcomes are positively correlated among those who served together in the same military groups. In a series of papers, Calvó-Armengol [2004], Calvó-Armengol and Jackson [2004, 2007] postulate a theoretical model that finds employment rates are increasing in one's degree, but decreasing in the degree of one's neighbors, while wages are positively correlated between neighbors, suggesting both complementary and substitutability effects of network size on employment prospects.

The only papers I am aware of using information from Facebook social networks to draw inference on labor market outcomes are Gee et al. [2017a] and Mayer [2012]. Gee et al. [2017a] provides descriptive evidence of a positive relationship between connection strength and working at the same employer using Facebook data. Mayer [2012] examines the structure of a Facebook network at one college and combines it with a survey taken at graduation to relate social network structure with employment outcomes. They find a positive association between employment and network size, but no relation between clustering and employment. I build on these descriptive findings using conditionally exogenous variation from Facebook release dates to estimate a causal effect of the Facebook network on future labor market outcomes.

2 Data

This paper primarily draws on five datasets. My primary sample is the set of (undergraduate) graduating classes of 2002 to 2009 at the first 760 selective 4-year universities to which

Facebook was released.

2.1 Mobility Report Card

I use earnings data from the Equality of Opportunity Project's Mobility Report Cards (MRC) [Chetty et al., 2017a][6]. This dataset contains annual IRS data on 2014 earnings of students who attended postsecondary schools in the United States, aggregated to the cohort level, along with cohort-level demographic characteristics. I use the term college cohort to reference a group of students belonging to a university and graduating class (e.g. Harvard class of 2004). This is the unit of observation for most of my analysis.

For my measure of labor market outcomes, I focus on the average *earnings rank* of each college cohort. This statistic measures the average percentile of students from a given college cohort relative to the national distribution of earnings for all individuals in the United States who were born in the same year. The MRC also provides me with measures of unemployment (the fraction of students who earned $0 in 2014), marital status, and a rich set of controls for parent income. My focus in this paper is on "selective colleges", those colleges that require a prospective student to submit an application to enroll in the college and has the ability to deny applications.[7] The Equal Opportunity dataset classifies the 760 selective 4-year schools in my sample into one of 6 selectivity "tiers". These tiers come from Barron's 2009 selectivity index [Barrons Educational Series, 2009], and capture differences in vertical quality across institutions. They include the following categories: Ivy Plus (Ivy league and comparable elite colleges, 12 schools in my sample), other elite schools (64 schools), highly competitive schools (107 schools), very competitive schools (195 schools), competitive schools (322 schools),and less competitive schools (60 schools). I use these classifications to better control for unobservable differences across schools.

While the data is generally quite detailed, there are a couple of disadvantages. The MRC income dataset is reported by birth cohort, rather than college entry cohort (i.e. freshmen class), and data is reported only for the 2014 fiscal year, instead of a fixed number of years after college graduation. Chetty et al. [2017a] documents that at selective colleges,"the vast majority of students enter at age 18 and graduate in four years", and provides some statistical evidence that this is the case. For my measure of earnings of an undergraduate college cohort entering college in year t and exiting school in year $t+4$, I use the 2014 earnings of college students who were 18 in year t. For the rest of this paper, I reference college cohorts by their graduating classes, and assume a 1:1 correspondence between birth cohorts, freshmen cohorts, and graduating classes.[8]

To account for the fact that earnings are measured in a single year across cohorts of different ages, I limit most of my analysis to the earnings rank variable discussed above. By assigning each student to a percentile of income relative to other individuals born the same year, the earnings rank normalizes earnings to a measure that is comparable across classes, since it only ranks students against those at the same stage in the earnings lifecycle. It is also a more robust measure of earnings which limits the influence of outliers. For example, a millionaire would not have excessive influence on the cohort's average earnings rank since they represent earnings in the 100th percentile, only twice the value of someone who has median earnings. This alleviates concern results may be driven by a few "superstar" earners within each cohort who could inflate the sample mean. For these reasons, along with earnings rank being the most detailed measure of earnings available in the MRC, I use average earnings rank as my primary measure of cohort-level earnings in this paper. In a robustness check, I measure the effect of Facebook access on nominal 2014 earnings (mean wages within cohort).

My sample of students, those who attended the first 760 4-year campuses Facebook is released to, comprises approximately 27% of all students born in each year. Because this is a non-trivial fraction of students born in each calendar year, a positive shift in earnings induced by Facebook will not only cause exposed students' earnings rank to increase, but also other students' ranks to decrease mechanically. Due to this limitation of my earnings data, I remain agnostic throughout this paper on what sort of changes to earnings Facebook technology induces. One could interpret earnings improvements generated from Facebook in two ways. The first is that Facebook, via information diffusion, creates high-quality matches between workers and employers that in a world without Facebook would never be realized. The other interpretation is that those better connected via Facebook technology are able to hear about better jobs earlier than their peers without Facebook access, which allows them to “cut in line” and get high-paying jobs before unexposed students here about it. At the extreme, this would yield zero net welfare gain but simply a reallocation of good jobs to those using social networking technology. This second interpretation may still yield insight on the redistributive implications of SNS technology.

2.2 IPEDS Dataset

The second dataset I use is supplementary data from Department of Education on the demographic composition of each college graduating class, along with educational outcomes such as bachelor’s degree completions and graduation rates at each of the universities in my sample. This data is collected from the “Fall Enrollment”, “Graduation Rates”, and “Completions” portions of the IPEDS survey, which are available here. I use this data to better control for differences in class composition (both in demographic background, prior ability (SAT math scores), and areas of study) within a school that may influence later earnings.

2.3 Facebook Release Dates

The third dataset used is the list of the date of introduction of Facebook into each of the first 842 postsecondary institutions Facebook rolled out to. Initially, Facebook was an exclusive network that only allowed individuals with `.edu` email addresses at invited universities to register with the network, and was primarily designed for users to interact with other students at their own universities. Data from Jacobs et al. [2015] lists the date students at Facebook 100 schools were allowed to register on the website. I use the same strategy implemented in Jacobs et al. [2015] to expand this list to the first 842 colleges to which Facebook was released. I accomplish this by either referencing articles from student newspapers indicating the release date when available, or inferring the release date from backlogged scrapes of `thefacebook.com` on the Wayback Machine throughout 2004 and 2005.[9] Following Jacobs et al. [2015], when a new college appears on Facebook's website between Wayback Machine archives on dates d_1 and d_2, I use the most recent date d_2 as the date Facebook is released to this college for my analysis. two-thirds of the colleges have their release date identified to within 3 days, and the maximum range of possible release dates is 2 weeks. Considering the variation in access time I use to explain labor market outcomes is one year, I assume any errors associated with release dates inferred by this procedure are negligible.

Figure 1 plots the penetration rate of Facebook to *all* selective schools in the United States over time, by schools' selectivity tier. The entry of Facebook to campuses is heterogeneous by tiers, with higher tier schools receiving access earlier. The figure shows that the distribution of release dates is fairly continuous over the 15 month period colleges in my sample received access. By May 2005, 55% of all selective 4-year college in the United States had access to Facebook.

Combining the campus-level start date of Facebook access with the assumption that each class spends 4 years at each of these universities (overwhelmingly the case for students at selective colleges [Chetty et al., 2017a]), I construct my primary treatment variable, the length of time each university cohort (indexed by i, t) was exposed to Facebook while forming friendships in college. The equation for this variable, denoted τ, is given below:

$$\tau_{i,t} = \begin{cases} 0 & \text{if } 0 > \frac{g_t - d_i}{365} \\ \frac{g_t - d_i}{365} & \text{if } 0 \leq \frac{g_t - d_i}{365} \leq 4 \\ 4 & \text{if } 4 < \frac{g_t - d_i}{365} \end{cases} \tag{1}$$

g_t denotes the final date of the academic year an individual in class t graduates in (coded as June 30th of that year, the end of the academic calendar year), and d_i is the first date

Facebook becomes available at university i. While some schools end their year earlier that June 30th, because I only use within-school variation to identify the effect of Facebook, this will be aborbed in university fixed effects and does not affect my identification strategy.

2.4 2005 Facebook Network Graph

My fourth dataset is the Facebook social network graph structure in September 2005 at each of the first 100 universities Facebook was released, from Traud et al. [2012]. The dataset provides a snapshot of the full graph of Facebook users who had at least one connection (friend) on the network as of September 2005. The data consists of 100 separate subgraphs of the Facebook network, one within each university. Each user, or node on the graph, has self-reported demographic information including type of user (undergraduate, graduate student, alumni, faculty, or staff), major, and graduating class. For each campus' Facebook graph, I use the subgraph of users that are undergraduate students or alumni[10], and self-report themselves to be in the class of 2002 to 2009 (87% of Facebook users in the data self-report their class). Using IPEDS data on class sizes at each school, I construct the count of students in each class that attended these schools but are not on Facebook, and add these to the September 2005 Facebook graphs as degenerate nodes with no connections/edges.

The two main measures I use from the Facebook graphs are the fraction of students in each class who are on Facebook, and the number of friends students in each class have on Facebook (their degree on the Facebook graph). My measure of the number of friends on Facebook is the number of *peer* friends. I define peers as students within 3 years of one's own class, so that the students would have overlapped with each other during the 4-year college experience. These account for 90% of a student's total friends on Facebook in 2005. I focus on these friendships to more directly examine the effect of Facebook on friendships corresponding to students who went to college at the same time, which are the relevant set of friends that would be affected by access to Facebook during college, my main treatment variable.[11] I use this data to establish that Facebook access is an important determinant of

Facebook takeup.

2.5 LinkedIn Career Insights

Finally, I use scraped data from November 2018 of LinkedIn's Careers Insights portal[12] for all schools in my sample to provide evidence of job sorting as a result of Facebook's entry into college campuses. I collect data on the graduating classes of 1997 to 2013. This allows me to measure the number of Linkedin Users in 2018 from each cohort. Linkedin also provides statistics of the top 15 firms students from each cohort work at, in terms of numbers of users listing a firm as their employer. I use this data to examine whether Facebook access had an effect on students' sorting patterns into firms on the labor market after graduating college.

2.6 Sample Summary

My sample is composed of students in cohorts from the class of 2002, when the Chetty et al. [2017a] dataset begins, up to the class of 2009, who entered college in the fall of 2005. I limit the sample to those entering college before 2006 because Facebook released a high school version for its platform in September 2005.[13], so students entering in 2006 could have been previously exposed to Facebook. Chetty et al. [2017a] documents that earnings ranks among younger cohorts are noisy signals of future income, so inclusion of younger cohorts may misrepresent the effects of Facebook on long-term employment outcomes. The correlation coefficient between earnings at 27 and 36 is 0.9 [Chetty et al., 2017a], suggesting age 27 earnings (the youngest cohorts in my sample when earnings are measured) are still good indicators of long-term labor market outcomes.

Column (1) of Table 1 displays summary statistics from the cohort-level dataset across the colleges in my sample. Overall, the sample earns more than individuals in the U.S. of the same age (the average earnings rank is around the 70th income percentile). In Columns (2) and (3), I compare means between the first 100 schools that received Facebook access, for which I have network data, and those receiving Facebook later, respectively. Summary statistics for all selective colleges in IPEDS and the Mobility Report Card are displayed in Column (4). This includes schools that did not receive Facebook by May 2005, the end of my data on Facebook release dates. Column (5) displays a p-value from a test of equality for each variable, comparing my sample with cohorts at selective colleges not in my sample. On average, students in my sample schools earn more, come from more affluent backgrounds, marry less often, and have higher math SAT scores, suggesting the rollout was not random across the universe of selective colleges in the United States. These qualitative differences are the same when comparing schools in my sample that received Facebook earlier versus later. For cohorts at Facebook 100 schools, 41% of individuals had signed up for Facebook by September 2005, and on average, individuals with Facebook accounts had 57 Facebook friends, suggesting takeup of Facebook was quite substantial at schools where it was released, even at Facebook's infancy. Though my sample is not representative of students at selective colleges, it covers a significant fraction of this population, about two-thirds of students.

3 Access Time and Online Social Network Structure

I now document that Facebook access directly influences online social network connections using the Facebook 100 network graph data. I consider the effect of access time, the number of years a student would have been exposed to Facebook while on campus, as of September 30th, 2005, when the Facebook graph data is measured, on both the probability of a student being on Facebook, and the number of Facebook friends. This measure of access time is defined as follows:

$$\tilde{\tau}_{i,t} = \begin{cases} 0 & \text{if } 0 > g_t - d_i \\ \frac{09/30/2005 - \max(s_{i,t}, d_i)}{365}) & \text{if } 0 \leq g_t - d_i \end{cases} \tag{2}$$

where d_i indicates the beginning date of Facebook access at university i, g_t indicates the final date of cohort t in college, and $s_{i,t}$ is the start date of cohort t at university i. I use variation in the start date of college across campuses because the class of 2009 enters college in the fall of 2005, only weeks before the Facebook graph snapshot is taken, and minor differences in start dates may greatly impact a class' exposure to Facebook as of September 2005. The start dates for the class of 2009 are acquired from the Jacobs et al. [2015] dataset. For other classes, Facebook access is granted at their university in the middle of their college tenure.

In Table 2, I show regression results relating Facebook access time with Facebook network strength, using individual level data from the 2005 Facebook graph snapshot. All regression estimates include class and university fixed effects, to control for each university's time-invariant propensity to sign up for Facebook, as well differences in usage across classes. An additional year of access time increases the likelihood of signing up for Facebook by 4.3%. I use the inverse hyperbolic sine (IHS) of degree [Burbidge et al., 1988], as the left hand side variable measure of connectedness. This allows me to include zero-degree students (students not on Facebook) in these regressions, while also allowing me to interpret the coefficient on access time in percentage terms. The number of friendships is increasing in time exposed to Facebook. The estimates of column (2), which includes all students in each cohort, suggest an increase in one year of access time leads to a 23% (unconditional) increase in Facebook connections. Conditioning on Facebook usage, and including demographic controls, yields a similar estimate of 20% in Column (4). In Appendix A.1, I complement this regression analysis with an event study that estimates the effect of having Facebook sufficiently early that the class of 2004 would have been exposed while in college. I similarly find large effects on signing up for Facebook and the number of friends. Overall, the effects of access time on the Facebook network structure are large in magnitude, even though I control for unobservable differences across universities and classes. This suggests that variation in access time to Facebook had large implications (at least in 2005) for social network formation on the platform. Given the central role of social networks play in labor market outcomes, we may expect that there are economically large effects on long-term labor market outcomes if these usage patterns induced by earlier access persist later in life.

4 Reduced Form Model

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

$$S(w_{i,t}) = \delta_i + \kappa_{\text{tier},t} + \beta \mathbf{X_{i,t}} + \psi \tau_{i,t} + \epsilon_{i,t} \tag{3}$$

where i indexes university, t indexes class, $\mathbf{X_{i,t}}$ are a vector of demographic controls for a cohort, and $\epsilon_{i,t}$ is a mean zero i.i.d. error. The coefficient of interest is ψ on the access time treatment variable. I report estimates of the effect of Facebook in terms of one year, because a year roughly corresponds to the in-sample variation of entry dates in my sample. ψ can be interpreted as the effect of Facebook being launched a year earlier during a student's college tenure. Time invariant university effects on earnings are captured via δ_i. I supplement this with flexible time fixed effects, $\kappa_{\text{tier},t}$, that vary by the Barron's selectivity tier of college. This fixed effect controls for changes in the return to elite education over time[14], age differences in earnings in 2014, and differential effects on wages from when students entered the labor market (e.g. graduating during a recession). They also provide controls over time for the main stratification variable on which the Facebook rollout was implemented, as documented in Figure 1.

The demographic controls in $\mathbf{X_{i,t}}$ include parent income (flexibly controlled for as the fraction of parents in 5 earnings quintiles, and the fraction of parents in the top 10%,5%,1%, and .1%), the fraction of students married at time of earnings measurement (2014), the gender/racial composition of each college cohort, SAT math score percentiles (25th and 75th percentile)[15], and the admissions rate (% admitted / % applied) of each cohort. $S(w_{i,t})$ is a statistic of the distribution of wages for cohort t at university i. This will typically be the average earnings rank of a cohort, but I will also look at other moments of the distribution of wages. I 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 variation I use to identify ψ, Facebook's entry date into a campus, was implemented.

4.1 Identification

Facebook introduced its social network to the world by iteratively rolling out to selected universities whose students needed to sign up with a `.edu` email address for the relevant school. The first 100 schools are shown in the Appendix in Tables A8 and A9. The initial recipients of the Facebook network were composed of the United States' most elite universities. Facebook's release dates across schools is then unlikely to be random with respect to student outcomes.

A more believable assumption, and the one made in this paper, is that, conditional on the 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 their platform 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 UT Austin students overall were less likely to sign up for Facebook than historically elite universities such as Stanford and Yale, 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 my first identifying assumption, and can be taken care of with university-level fixed effects in my regression analysis to capture unobservables in the earnings of students who graduate from the same university.

Because τ is mechanically correlated with being in a later class, I also worry about secular trends over time in the wages of college graduates that may be spuriously correlated with access time. 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. With year×tier fixed effects, my reduced form model captures a considerable amount of the heterogeneity between cohorts that may challenge my identification of a causal effect of Facebook access on wages, ruling out endogeneity threats done at or above the year×tier level. My second identifying assumption is that within these year×tier cells, any additional time-varying unobservables in labor outcomes, conditional on $\mathbf{X_{i,t}}$, are independent of Facebook's release date decision. For example, within highly selective colleges, Northeastern University in Boston received access to Facebook in April 2004, while University of San Diego, a private university in Southern California with demographically similar students, received Facebook access 7 months later in November. Thus, with university fixed effects, my linear model identifies ψ essentially through the difference in changes in earnings between cohorts at these similar universities.

One can frame my identification strategy for ψ in a difference-in-difference framework. Ignoring demographic covariates $\mathbf{X_{i,t}}$, university fixed effects act as controls for the "preperiod" earnings of students 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 within a entry year×selectivity tier cell that had differential exposure (in terms of τ) to Facebook during college. Subtracting the treatment group earnings from the control group earnings yields the second difference. Because the treatment variable τ is not binary, but rather continuous, we then effectively weight these double differences by the difference in treatment and control group access time to get an estimate of ψ. ψ is then identified by two sources of variation: exposure differences between cohorts at the same university, and differences in the release date of Facebook between cohorts of the same year within the same selectivity tier.

In a potential outcomes causal inference framework [Rosenbaum and Rubin, 1983], I assume unconfoundedness on earnings and timing of Facebook access, conditional on university, time-varying differences at the selectivity tier level, and observable demographics:

$$w_j(s) \perp \tau_{i,t} | \delta_i, \kappa_{\text{tier},t}, \mathbf{X_{i,t}} \quad \forall i, t$$

where $w_j(s)$ is the earnings rank of student j that is exposed to Facebook access for s years, belonging to university i and cohort t. In the Appendix, I provide evidence that Facebook access is exogenous conditional on these controls by running placebo regressions (Table A1) of the effect of Facebook access on predetermined observables such as parent income, gender, and race, and show no significant relationship. I also establish that this is not driven by differential trends at schools receiving Facebook earlier, by estimating a regression that shifts the year of Facebook launched as a platform, but preserves the relative differences in Facebook's release schedule across schools (Table A4).

5 Results

5.1 Main Specification

Table 3 shows the regression results for the baseline specification (Equation 3), iteratively adding controls successively for four dependent variables of interest. Panel A displays the effects on overall earnings rank on average. Note that in all columns, I control for marriage status because of its mechanically positive relationship with earnings in the MRC.[16] Inclusion of year×tier fixed effect (Column 2) reduces the magnitude of the coefficient. In the preferred specifiation, Column (3), I include pre-determined demographic controls, and the estimated effect does not noticeably change. In Column (4), I include controls for intermediate schooling outcomes, including the fraction of each major chosen by each cohort[17], and the graduation rate (measured as % completing a bachelor's degree in 6 years). This does not change the magnitude of the effect significantly, suggesting that improvements in labor market outcomes are not coming through the channel of changing college outcomes. The estimates in Column

(3) suggest that having been exposed to Facebook for an additional year, leads to an increase of 0.61 percentiles in a cohort's average earnings, relative to the national distribution. To translate this effect size into dollars, I estimate a treatment effect on the treated using the following formula:

$$ATT_{FB} = \frac{1}{N_{\tau>0}} \sum_{i,t:\tau>0} \frac{F_t^{-1}(\bar{w}_{i,t}) - F_t^{-1}(\bar{w}_{i,t} - \hat{\beta}\tau_{i,t})}{\tau_{i,t}}$$

where F_t denotes the CDF of earnings from Chetty et al. [2017b] in each age/class t, $N_{\tau>0}$ is the number of cohorts with positive access time in my sample, $\bar{w}_{i,t}$ is the mean earnings rank in a cohort i, t, and $\bar{w}_{i,t} - \hat{\beta}\tau_{i,t}$ is the predicted earnings rank of each cohort exposed to Facebook if they had no access while in college, according to Equation 3. This represents the average effect (per year) of Facebook access predicted by my linear model for cohorts with positive exposure to Facebook in college. I estimate an effect of \$972 on average per year of access, which is sizeable.[18] Because the average earnings are \$49,711 across cohorts, this translates to a 1.95% increase in average earnings.

The effects on average wages could come from multiple paths: those already employed individuals receiving better wage offers from an expanded social network, and those unemployed transitioning to employment and thus increasing the average earnings rank. In order to determine the magnitude of these two complementary effects on earnings, I use the following decomposition from Fairlie and Bahr [2017] of the effect on the earnings rank of wage, w, from exposure to Facebook:

$$\begin{aligned}
\beta_w &= E[w_{i,t}|T] - E[w_{i,t}|C] \\
&= \Big((Pr(E_{i,t}|T) - Pr(E_{i,t}|T))E[w_{i,t}|E,T]\Big) + \Big((E[w_{i,t}|E,T] - E[w_{i,t}|E,T])Pr(E_{i,t} = 1|C)\Big) \\
&= \beta_E E[w_{i,t}|E = 1, T] + \beta_{w|E} Pr(E_{i,t} = 1|C)
\end{aligned}$$

where β_w denotes the effect on earnings rank, β_E denotes the effect on the employment rate, and $\beta_{w|E}$ denotes the effect on earnings rank conditional on already being employed, denoted by E. C denotes control groups, and T denotes treatment groups. The first term captures the contribution from increased employment, while the second term captures the contribution from improved wages for those already employed. I estimate versions of β_E and $\beta_{w|E}$ in Panels B and C of Table 3, which show estimates of the effect of access on employment and earnings rank of individuals in each cohort who are already employed.[19] The coefficient in Panel B, Column (3) suggests an additional year of access to Facebook leads to a 0.39 percentage point increase in a cohort's employment rate. The effect on the wages among those who are employed is 0.36 percentiles in 2014 in Column (3) of Panel C. To implement the decomposition, I use $E[w_{i,t}|E = 1, T] = \hat{E}[w_{i,t}|E = 1, \tau_{i,t} > 0]$, the mean earnings rank conditional on working among cohorts with positive access, and $Pr(E_{i,t} = 1|C) = \hat{E}[E_{i,t}|\tau_{i,t} = 0]$, the mean employment rate among cohorts with no access to Facebook in college. The decomposition implies that 45% of the effect on earnings rank is from increased employment, while 55% of the effect is from wages of those already employed. Thus, it appears that Facebook effects the earnings rank of cohorts equally through the employment likelihood and the wages of those employed in a cohort. In Appendix B, I present a stylized model of an economy where job offers are acquired through friendships and the employment rate is unaffected by the social network structure; this model highlights the economic intuition for how earnings might increase from access to Facebook, even if employment changes little from access. In the model, Facebook decreases the cost of maintaining friends, which leads to more friendships being acquired. With more friends, individuals hear about more job offers, and choose the offer with the highest wage, which leads to more connected individuals (through Facebook) to obtain higher wages.

In terms of effect sizes, the most relevant paper is Mayer [2012], which relates social networks in college to employment outcomes. They find that a standard deviation increase in Facebook friends at one university leads to a 3 percentage point increase in the probability of being employed immediately after college. In Section A.4, I similarly relate Facebook friendships in 2005 attributable to access time to later outcomes via an instrumental variables regression, and find a standard deviation increase in friends leads to a 1.7% increase in the employment rate. Our results differ both because I am examining longer term outcomes, our samples differ, and my empirical strategy uses exogenous variation, whereas that paper simply controls for demographic characteristics of students. However, these results are broadly similar in magnitude to what is found in that paper.

Finally, I also estimate the effect on dispersion in the earnings rank, measured by the standard deviation of the earnings rank distribution within each cohort i, t. This dispersion measure is constructed using an estimate of the earnings rank CDF from the Chetty et al. [2017a] data. Further details of this measure's construction can be found in Appendix A.2. I estimate that access time decreases the dispersion of earnings within cohort. The effect size from a year of access is equal to 15% of the standard deviation of earnings dispersion in my sample, across cohorts. This is consistent with a story where greater information sharing, induced by Facebook technology, makes the set of wage offers received by students more homogeneous, and there is less segmentation within cohort of who acquires high-paying versus low-paying jobs. Gee et al. [2017b] documents that within country networks, income inequality is associated with stronger ties within a network. This suggests that Facebook access decreases social homophily within cohort, since I find a negative effect from access to earnings dispersion.

5.2 Heterogeneous Effects

While Table 3 suggests strong effects on wages on average, it is an open question whether this average effect is equal across the wage distribution within a cohort. To assess the impact of Facebook on the cohort wage distribution, I look at the effects of increased access to Facebook on the fraction of students within quantiles of the national earnings distribution. The quantiles I examine are the fraction of students who are in the $1-60$th, $60-80$th,$80-$

90th, and $90-100$th percentile. I chose these quantiles since they divide the cohorts into approximately equally populated groups for my sample, due to the overall selective nature of the schools in my sample. For the $1-60$th percentile group, I calculate this as the fraction of students in a cohort below the 60th percentile, conditional on being employed (earnings > 0). This excludes those unemployed (examined in Table 3), and instead captures students in low-paying jobs.

Table 4 displays the estimates. The table suggests an overall shift in earnings; an additional year of exposure decreases the probability of being in the lowest quantile of earnings by .57%, but increases the probability of cohort members being in the top decile by 0.78%. There are statistically insignificant intermediate effects on the fraction of students in the middle quantiles (60%-90%). I find that the effect on quantiles below the 80th percentile are statistically significantly different from the effect on the top decile, by running a joint regression with all four quantile variables. This table is consistent with the findings of decreased dispersion in earnings from Table 3 Panel D; the fraction of individuals in the lower tail of the distribution declines, and concentration is increased into the right tail of the earnings rank distribution among students at selective colleges. This is consistent with my finding of decreased earnings rank dispersion within cohort in Table 3. It is not the case, for example, that there is a null effect for earnings below the 80th percentile, a negative effect on the $80-90$th percentile, and a strong positive effect on the $90-100$th percentile, which would cause us to interpret the average coefficient in Table 3 Panel A differently (i.e. only those already likely to make relatively high wages benefit from Facebook).

Next, I examine the heterogeneous effects of Facebook exposure on wages by student demographics, specifically gender and parent income. I show estimates of the effect on earnings rank by parent income quintile, the most granular measure of earnings rank (conditional on parent income) available. For gender, I split by male and female students. I drop the demographic control variables for these regressions because they are unavailable for these subgroups, so I only include controls for marital status, year$\times$tier fixed effects, and university fixed effects. For comparison, I also report the baseline regression for all students in a cohort on overall earnings ranks with the same set of controls. Table 5 displays the results. Quantitatively, the most benefited group in terms of earnings rank whose parents are those coming from the 20th-40th percentile, which roughly corresponds to parents making between \$22,000 and \$40,000 in 2014. The effect for those with parents making below this amount is statically insignificant; this is likely due to the low count of students coming from this income quintile (6.6%) at 4-year selective schools. Facebook access also increases the earnings rank for those coming from the top quintile of parent income, which is unsurprising given they comprise 42% of students. With respect to gender, I find that the effect on female students' earnings rank is 80% larger than their male counterparts, however the difference in these coefficients is statistically insignificant. In a joint regression of all outcome variables in the table, I find that none of the estimates differ in a statistically significant manner from the baseline coefficient in Column (1). This suggests that the benefits of Facebook were largely shared by these subpopulations of college students.

In Table 6, I examine heterogeneity in the response to Facebook access by school characteristics. We focus on three types of characteristics: whether a school is located in a city (the central city of an MSA), whether the school is large (average undergraduate enrollment > 5,000) and whether a school is private. Each of these characteristics split my sample roughly in half. In Columns (1)-(3), I estimate the marginal effect of each of these characteristics on the coefficient of Facebook access on earnings rank. None of the interaction coefficients are statistically significant at the 5% level. In Column (4), I include all characteristics, and find significant heterogeneity across school types in responses. This comes from the fact that these characteristics are correlated: private schools are less likely to be large ($\rho = -0,7$), so the effects of each of these characteristics may cancel each other out in the marginal interactions. Schools located in urban cities are less responsive to Facebook access, which is consistent with the story that Facebook access is less valuable if there are more job opportunities immediately near the campus. Facebook access increases earnings rank more at schools where the student body is larger, and thus networks made in college can be larger. Finally, Facebook access has a larger effect at private schools. This is consistent with evidence from Overgoor et al. [2020a] that shows students at private schools have larger and more clustered networks, controlling for the size of the school, and are more likely to remain close to college friends after college [Overgoor et al., 2020b].

5.3 Nonlinear Effects

One might expect that the greatest return to Facebook access occurs when Facebook is first released to a campus, when offline friendships can be instantly converted onto the platform. At the same time, it may be that economies of scale are required for Facebook to have an effect on later earnings. With economies of scale in Facebook access, earnings may be affected only after the cohort is exposed to Facebook long enough that the campus' Facebook network is able to reach a critical mass. I test the extent to which these two forces play a role in the effect of access time τ estimated in equation 3, which required the effect to be linear across years. To do so, I estimate a version of Equation 3 that allows for flexible, nonlinear effects of Facebook exposure, that depends on how many academic years of college a student had access to Facebook while in college:

$$S(w_{i,t}) = \delta_i + \kappa_{g,\text{tier},t} + \beta \mathbf{X_{i,t}} + f(\tau_{i,t}, \vec{\psi}) + \epsilon_{i,t} \tag{4}$$

where $f(.)$ specifies a natural cubic spline with knots at $[0, 1, 2, 3, 4]$ years of exposure, and $\vec{\psi}$ are the parameters governing the non-linear relationship. This allows us to estimate a nonlinear relationship between τ and earnings while controlling for the same fixed effects and demographic controls as in Equation 3.

Figure 2 plots the marginal effects of increased Facebook exposure ($\partial f / \partial \tau_{i,t}$), from zero to four years of Facebook access. The estimated function $\hat{f}(\tau_{i,t})$ suggests that the marginal return to earnings rank from increased Facebook access is lower initially, at 0.33 percentiles per year of exposure, and the returns to further access increase until around 2 years of access in college to 0.85 percentiles, after which the returns decrease to 0.44 percentiles. This figure is broadly consistent with the economies of scale story, in which Facebook is not useful for labor outcomes if access time is so short that not enough offline friendships are converted onto online friendships.

5.4 Robustness Checks

I now consider variations to my baseline specification to see if the results are robust to spurious causes.

Even though I control for time invariant university effects on earnings, there may be concern, because elite universities were the first to be exposed to Facebook, that my results are being driven by differential trends in the return to education at an elite university during my sample period [Brewer et al., 1999]. Although I address this by including year $\times$tier fixed effects in my baseline specification, this may incompletely control for these trends if the classification tiers are too broad. To account for this, I perform a version of my baseline specification (Equation 3) that excludes the "Ivy Plus" and "Elite" tier universities, the 76 most selective schools in the country, according to the Barron's selectivity index.

Another concern is that the earnings data from Chetty et al. [2017a] is collected at an aggregated level (called a super OPEID in their paper), which combines earnings from multiple campuses due to the way institutions report taxes. Because Facebook was released to specific campuses, earnings data from super OPEID schools contain information about the earnings at other campuses, and one may be worried my positive results are attributable to earnings information from unexposed campuses, or that pooling campuses with different exposure dates is mis-measuring the effect of exposure on earnings. Therefore, I also estimate a version of Equation 3 dropping the 7 super OPEID schools in my sample.

Even though I include year$\times$tier fixed effects, there may be concern that my results may be driven by a differential response, within these cells, to weathering the Great Recession by students from more selective colleges, possibly through substituting to graduate education [Kahn, 2010]. I also consider a version of my model that excludes the last 2 years of my sample, so that I estimate the effect of access on earnings rank for students who graduated before 2008.

One might argue that my measure of earnings (individual pretax earnings) does not appropriately capture an individual's true income. To address this, I use supplementary data from Chetty et al. [2017a] that calculates earnings rank by household wages, which includes spousal earnings. I also consider wages measured as adjusted gross income (AGI), which includes spousal earnings *and* capital gains experienced by the household, for earnings rank calculations.

It may seem unreasonable to equally weight the earnings rank average of a Caltech cohort containing 150 students and the earnings rank average of a University of Florida cohort containing 7,000 students. Therefore, I also estimate my baseline specification where I weight cohort-level observations by the number of students in each campus' cohort.

Finally, I also use a nominal measure of earnings as my outcome variable, the logged mean average earnings of a cohort. This measurement is subject to concerns discussed in Section 2.1, but it is still useful to see if my effects translate when using a measure of nominal earnings within a cohort, since the coefficient is more easily interpretable.

Table 7 displays the results from these robustness checks. Across specifications, there is largely no noticeable difference between these estimates and those from Table 3 column (4). In columns (4) & (5) of Table 7, the coefficient is approximately halved. This is a consequence of the earning rank in these columns including the earnings of an individual's spouse, which are less likely to be affected by Facebook access[20]. Using nominal earnings measures, the estimate of Facebook access is statistically significant at only the 10% level, suggesting cohort earnings increase by 1.2% from an additional year of Facebook access. Because this effect is smaller and noisier than results for earnings rank, which are relative wage measures, this suggests that Facebook is delivering better jobs at least in part through exposed students acquiring good jobs at the expense of non-exposed students. Overall though, this table verifies that my baseline results are robust to the alternatives considered.

I also consider robustness checks that alter the main empirical specification. In Appendix Table A2, I consider alternative year fixed effects that account for regional labor market shocks, by considering inclusion of BEA Region and State in my time fixed effects specification. The results are robust to various levels of time fixed effects, though the magnitudes are larger when selectivity tier is not included, suggesting there are differential labor market shocks by school selectivity (which is related to student ability) correlated with Facebook access. Appendix Table A7 shows that the estimated effects on the main outcome variables are robust to using two-way clustering of standard errors, at the school and time (year of entry) level. Appendix Table A5 shows the effect of the IHS of Facebook access, so that effects can be understood in terms of logged increases in access time. The table suggests the earnings rank increases by 0.73 percentiles from a 100% increase in access time during college.

In Table A3, I estimate the effect of Facebook access on both schooling outcomes and the likelihood of marriage. I find that there is no significant effect of Facebook access on the likelihood of graduation, though access does appear to be correlated with students sorting into differential majors: students appear to shift away from the Arts/Humanities and into the Social Sciences when they have access to Facebook longer during college. I also estimate that receiving Facebook a year earlier leads to a 1% decrease in a cohort's marriage probability. While labor outcomes are the focus of this paper, it's effect on other outcomes, both during and after university, documented in this table suggest access to Facebook in college may be a useful research tool in future work to study the effect of social networking on non-labor outcomes.

6 Mechanism Evidence: Job Sorting on Linkedin

While the previous section established a clear link between Facebook access time and earnings, I still lack empirical evidence for what channel these earnings increases occurred through. One possible mechanism is that Facebook access translates to a student better connecting with their peers on the labor market after they graduate. Since Facebook reduces the cost of maintaining social ties, it should increase the pool of friends through which students could learn about job opportunities after college. If friend a tells friend b about a job, the vacancy is likely to be at a's own firm. If b then applies and is hired, then a and b would now be working for the same company. Therefore, one consequence of Facebook increasing job opportunities learned from social networks would be greater co-location of college alumni at firms in the labor market.

To test this mechanism, I use scraped LinkedIn data from 2018 on the cohorts in my sample to determine if students are more likely to positively sort across firms when Facebook access is longer during college. To do this, I construct a measure of employment concentration using data from the top 15 firms each cohort works at, according to LinkedIn. The measure is based on the Herfindahl-Hirschman index (HHI) used to measure industry concentration in merger analysis. It is defined as:

$$\text{Employment HHI}_{i,t} = \sum_{k=1}^{15} s_{i,t,k}^2$$

where $s_{i,t,k}$ indicates the share of students at top 15 firms for cohort t from university i that work at the firm k. Shares are calculated conditional on working at a top 15 firm so that $\sum_{k=1}^{15} s_{i,t,k} = 1$. A high employment HHI indicates that there is a relatively high concentration of students in the same class in their employment choices, and thus alumni choose to co-locate at firms more frequently. I use logged employment HHI as my main dependent variable in this analysis, so that changes are interpreted as in percentage terms. If Facebook increases job opportunities acquired through one's friends from college, I would expect employment concentration to increase with more access time.

This measure is not without its limitations. Ideally, I would be able to construct a cohort-level measure of concentration using all firms students from cohort i, t work at, but I am only able to use data from the top 15 firms in each cohort among LinkedIn users, because this is the maximum LinkedIn reports per cohort. On average, the top 15 firms on LinkedIn within a cohort on average employ 5.3% of *all* students in a graduating class, and 8.8% of those on on LinkedIn. This is a non-trivial fraction, but clearly this measure will not capture most of the employment patterns among college graduates in a cohort. Note also that LinkedIn classes pool graduate and undergraduate classes, so the measure may not accurately reflect job sorting occurring at the undergraduate level.

Table 8 displays the effect of access time in college on the employment HHI measured on Linkedin. The results in Column (2) suggest an effect of Facebook access on job concentration for an additional year in college of approximately 5.4%, a large magnitude, consistent with the large effect on earnings I estimate and discuss in Section 5. I also measure the effect on the raw number of coworkers who are peers: the number of students from the same college (within 3 years of their class) who also work at that firm. Because I do not have data outside top 15 firms in a cohort, this measure is also fairly noisy; if a firm drops out of the top 15 for a cohort, I do not measure how many students are employed there, even if it is within 3 years of the cohort for which I am constructing the measure. I find that receiving Facebook one year earlier increases the expected number of workers with which you were peers in college by around 7.3% (Column 4).

Note that this measure does not deal with endogeneity concerns between LinkedIn takeup and Facebook access. It may be that Facebook access increases a student's propensity to sign up for other social networks, such as LinkedIn, so that changes to the employment HHI occur solely via changes in LinkedIn's user composition. To address these endogeneity concerns, I consider a regression that estimates the effect of access time on the fraction of alumni on LinkedIn in each cohort. Linkedin adoption is quite high in my sample: on average, 60%

of each graduating class is on LinkedIn in 2018. This suggests that a non-trivial fraction of students' job location decisions are being captured in the LinkedIn employment HHI measure considered in this section, assuming firms below the top 15 have similar employment patterns. The regression estimates are shown in columns (5) and (6) of Table 8. For LinkedIn takeup, inclusion of only university effects shows a positive relationship with Facebook access time. This is due to a secular increase in Linkedin adoption in younger cohorts; inclusion of year×tier fixed effects eliminates this relationship, which is now insignificantly different from zero in Column (6) and low in magnitude (-.2%). Overall, Table 8 gives credence to the story that job opportunities shared in one's social network are driving the results on cohort earnings.

7 Discussion and Conclusion

The empirical results above suggests a relatively large positive effect on earnings from increased exposure to Facebook access. It is likely that the explosive growth the Facebook platform experienced during the sample, documented in Table 2, contributes to the large estimated effects from exposure on income. It is useful to remind readers that Facebook was *the* social network for the group of college students comprising my sample, made evident by Facebook's extraordinary penetration rates and network density as early as 2005. This bears out in Facebook's usage patterns at the time: in 2004, 50-90% of *all* Facebook users logged in to the site daily.[21]. In the words of one student from Southeast Missouri State University in 2005, "It is so addictive. It is now part of my daily routine, no matter what else is going on that day".[22] The campus-focused nature of the Facebook platform made it central to student life during this period.

The effect of Facebook's online social network platform on 4-year selective college campuses is unlikely to translate equally to other settings, even holding usage rates fixed. 4-year selective colleges in the United States are often composed of more geographically dispersed individuals than other colleges. For example, compared to other schools tracked by the Department of Education, which are mostly 2-year commuter colleges and vocational schools, students in the class of 2009 at the schools in my sample have 4 times as much diversity in their state of origin, and 3.3 times as many international students per cohort.[23] Individuals with less geographically dispersed social networks may benefit less from Facebook's social network technology, since they could stay in touch with local ties without an electronic communication platform. The difference in Facebook usage during my period of interest, along with the difference in the composition of selective 4-year colleges compared to other postsecondary schools, gives context to the estimated effect size and its potential external validity.

Overall, this paper provides evidence that Facebook access during college has a positive effect on student earnings later in life. This positive effect matters on both the intensive and extensive margin of wages. I also provide direct evidence of a positive relationship between Facebook access time and job referrals, in the form of a LinkedIn job sorting measure, consistent with improved labor outcomes stemming from changes to employment networks.

While a more complex estimation procedure is necessary to fully understand the underlying mechanisms relating access time, Facebook usage/network structure, and earnings, this paper provides a first pass at estimating the causal effects of online social network structure on future earnings. Further work in this setting is needed to better understand the effects of social networks on future labor market outcomes. In particular, whether these earnings increases for students with Facebook access come from better employment matches, or simply a reallocation of desirable jobs within the economy to those with Facebook, is a first-order question for understanding the welfare consequences of SNS technology with respect to labor market outcomes. This paper also documents a natural experiment, the entry of Facebook to college campuses, that could be used to investigate other relationships between social network structures and economic outcomes.

References

J. D. Angrist and J.-S. Pischke. *Mostly harmless econometrics: An empiricist's companion*. Princeton university press, 2008.

S. Aral. The future of weak ties. *American Journal of Sociology*, 121(6):1931–1939, 2016.

C. D. E. . Barrons Educational Series. Barrons profiles of american colleges, 2009.

D. J. Brewer, E. R. Eide, and R. G. Ehrenberg. Does it pay to attend an elite private college? cross-cohort evidence on the effects of college type on earnings. *Journal of Human resources*, pages 104–123, 1999.

J. B. Burbidge, L. Magee, and A. L. Robb. Alternative transformations to handle extreme values of the dependent variable. *Journal of the American Statistical Association*, 83(401): 123–127, 1988.

A. Calvó-Armengol. Job contact networks. *Journal of economic Theory*, 115(1):191–206, 2004.

A. Calvó-Armengol and M. O. Jackson. The effects of social networks on employment and inequality. *American Economic Review*, 94(3):426–454, June 2004. doi: 10.1257/0002828041464542. URL `http://www.aeaweb.org/articles?id=10.1257/0002828041464542`.

A. Calvó-Armengol and M. O. Jackson. Networks in labor markets: Wage and employment dynamics and inequality. *Journal of economic theory*, 132(1):27–46, 2007.

R. Chetty, J. N. Friedman, E. Saez, N. Turner, and D. Yagan. Mobility report cards: The role of colleges in intergenerational mobility. Technical report, National Bureau of Economic Research, 2017a.

R. Chetty, D. Grusky, M. Hell, N. Hendren, R. Manduca, and J. Narang. The fading american dream: Trends in absolute income mobility since 1940. *Science*, 356(6336):398–406, 2017b.

R. Enikolopov, A. Makarin, and M. Petrova. Social media and protest participation: Evidence from russia. *Econometrica*, 88(4):1479–1514, 2020.

R. W. Fairlie and P. R. Bahr. The labor market returns to computer skills: Evidence from a field experiment and california ui earnings records. 2017.

D. Garcia, P. Mavrodiev, and F. Schweitzer. Social resilience in online communities: The autopsy of friendster. In *Proceedings of the first ACM conference on Online social networks*, pages 39–50. ACM, 2013.

L. K. Gee, J. Jones, and M. Burke. Social networks and labor markets: How strong ties relate to job finding on facebooks social network. *Journal of Labor Economics*, 35(2): 485–518, 2017a.

L. K. Gee, J. J. Jones, C. J. Fariss, M. Burke, and J. H. Fowler. The paradox of weak ties in 55 countries. *Journal of Economic Behavior & Organization*, 133:362–372, 2017b.

M. S. Granovetter. The strength of weak ties. *American journal of sociology*, 78(6):1360–1380, 1973.

M. O. Jackson. *Social and economic networks*. Princeton university press, 2010.

M. O. Jackson and B. W. Rogers. Meeting strangers and friends of friends: How random are social networks? *American Economic Review*, 97(3):890–915, 2007.

A. Z. Jacobs, S. F. Way, J. Ugander, and A. Clauset. Assembling thefacebook: Using heterogeneity to understand online social network assembly. In *Proceedings of the ACM Web Science Conference*, page 18. ACM, 2015.

L. B. Kahn. The long-term labor market consequences of graduating from college in a bad economy. *Labour economics*, 17(2):303–316, 2010.

R. Laschever. The doughboys network: Social interactions and the employment of world war i veterans. 2013.

S. K. Lewis and V. K. Oppenheimer. Educational assortative mating across marriage markets: Nonhispanic whites in the united states. *Demography*, 37(1):29–40, 2000.

A. Mayer. The structure of social networks and labour market success. *Applied Economics Letters*, 19(13):1271–1274, 2012.

A. Mayer and S. L. Puller. The old boy (and girl) network: Social network formation on university campuses. *Journal of public economics*, 92(1):329–347, 2008.

J. Overgoor, L. 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.

J. Overgoor, B. State, and L. Adamic. The dynamics of us college network formation on facebook. *arXiv preprint arXiv:2011.01967*, 2020b.

P. R. Rosenbaum and D. B. Rubin. The central role of the propensity score in observational studies for causal effects. *Biometrika*, pages 41–55, 1983.

C. J. Simon and J. T. Warner. Matchmaker, matchmaker: The effect of old boy networks on job match quality, earnings, and tenure. *Journal of labor economics*, 10(3):306–330, 1992.

A. L. Traud, E. D. Kelsic, P. J. Mucha, and M. A. Porter. Comparing community structure to characteristics in online collegiate social networks. *SIAM review*, 53(3):526–543, 2011.

A. L. Traud, P. J. Mucha, and M. A. Porter. Social structure of facebook networks. *Physica A: Statistical Mechanics and its Applications*, 391(16):4165–4180, 2012.

J. Ugander, B. Karrer, L. Backstrom, and C. Marlow. The anatomy of the facebook social graph. *arXiv preprint arXiv:1111.4503*, 2011.

1. Source: http://www.pewinternet.org/fact-sheet/social-media/
2. Source: https://web.archive.org/web/20050820005604/http://www.msnbc.msn.com:80/id/6596533/site/newsweek/
3. Source: http://www.pewinternet.org/wp-content/uploads/sites/9/media/Files/Reports/2011/PIP-Social-networking-sites-and-our-lives.pdf/
4. Even in 2014, when I measure labor outcomes, among American college graduates less than age 30 (so entering college from 2003-2010), 73% used Facebook daily, while this number was 18% for Twitter, 4.4% for Google Plus, and 9% for LinkedIn, the other social networks inquired about in a national Pew survey. Source: https://www.journalism.org/dataset/american-trends-panel-wave-1/
5. The release dates at each school can be found here: www.web.stanford.edu/~larmona/fbearn/FB_introduction_dates.csv
6. The dataset is available for download here: http://www.equality-of-opportunity.org/data/
7. I exclude the 4 U.S. military academies (West Point, Navy, Coast Guard, Air Force) because they lack income data in the MRC.
8. One exception is Northeastern University, which has 5 year undergraduate degree programs; for this school, I code graduating classes as 5 instead of 4 years after the freshmen entry year. I still assume freshmen entry year corresponds perfectly to birth cohort for this college.
9. Facebook used to list campuses with access on its homepage. Example homepage I infer these dates from: https://web.archive.org/web/20050822175050/http://www.thefacebook.com:80/index.php?showall=1, In August 2005, Facebook changed the format of its homepage, making it impossible to infer dates from Wayback Machine archives. For this reason, I only infer release dates until the end of the 2004-2005

academic year.

10. I exclude graduate students from the Facebook networks since this paper only examines the effects of Facebook on undergraduate labor market outcomes, so the relevant network structure is among undergraduates.

11. At the same time, there is ample evidence that weak ties, which may be formed through non-physical interactions during college, may be important for later outcomes [Aral, 2016]. I am unable to estimate the effect of access to Facebook during college on the formation of these weak ties. Additionally, advertising was allowed on Facebook as early as 2004 `https://www.adtaxi.com/blog-roll/2018/03/08/ad-evolution-history-facebook`, and this may have allowed employers to recruit candidates from specific campuses. This may also be a relevant input into the effect of Facebook access on later labor market outcomes.

12. An example page is displayed in Figure A3

13. Source: Facebook S1 Document, p.43

14. Changes over time to the returns of an elite college degree is documented by Brewer et al. [1999] for the pre-period to my sample

15. These are the only moments of the score distribution reported to IPEDS. For schools with a higher fraction of students reporting ACT scores, I convert ACT scores to SAT scores using `https://www.ets.org/Media/Research/pdf/RR-99-02-Dorans.pdf`, and include a dummy for test scores derived from the ACT.

16. Married individuals are given half of non-wage income on tax form 1040 as part of their individual earnings. In Table A6, I show that excluding marriage from controls does not qualitatively change results with year fixed effects. Year fixed effects are necessary since the propensity to marry changes as students get older, and income changes are correlated with age & access time mechanically through this channel.

17. Aggregated to 8 major types using the College Board's major type designation `https://bigfuture.collegeboard.org/majors-careers`.

18. Note CDF earnings data is only available for the cohorts entering college between 2000 and 2002, but effects do not differ quantitatively across years.

19. The earnings rank conditional on employment is computed by using the fact that unemployed individuals in a cohort are assigned the mean rank nationally of individuals with zero income (typically around the 9th percentile since around 20% of students are unemployed on average in each cohort) Chetty et al. [2017a], so the following identity $E[w_{i,t}|\text{Employed}] = (E[w_{i,t}] - Pr_{i,t}(\text{Not Employed}) * E[w_{i,t}|\text{Not Employed}]) / Pr_{i,t}(\text{Employed})$ allows me to recover the average earnings percentile among the employed in a cohort

20. This assumes no sorting on university cohort in the marriage market, which may not be a valid assumption (for example, Lewis and Oppenheimer [2000] documents marital sorting on education level). Without additional data on the cohort of spouses, it is not clear how I would adjust my estimates for marital sorting.

21. Source: `https://web.archive.org/web/20060521084814/http://www.post-gazette.com:80/pg/04333/417839.stm`

22. Source: `https://web.archive.org/web/20051123232204/http://semissourian.rustcom.net:80/story/1112288.html`

23. Diversity is measured as the Shannon entropy for the probability of a student originating from each of the 50 states or D.C. Data for these statistics come from the IPEDS "Fall Enrollment" survey in 2005.

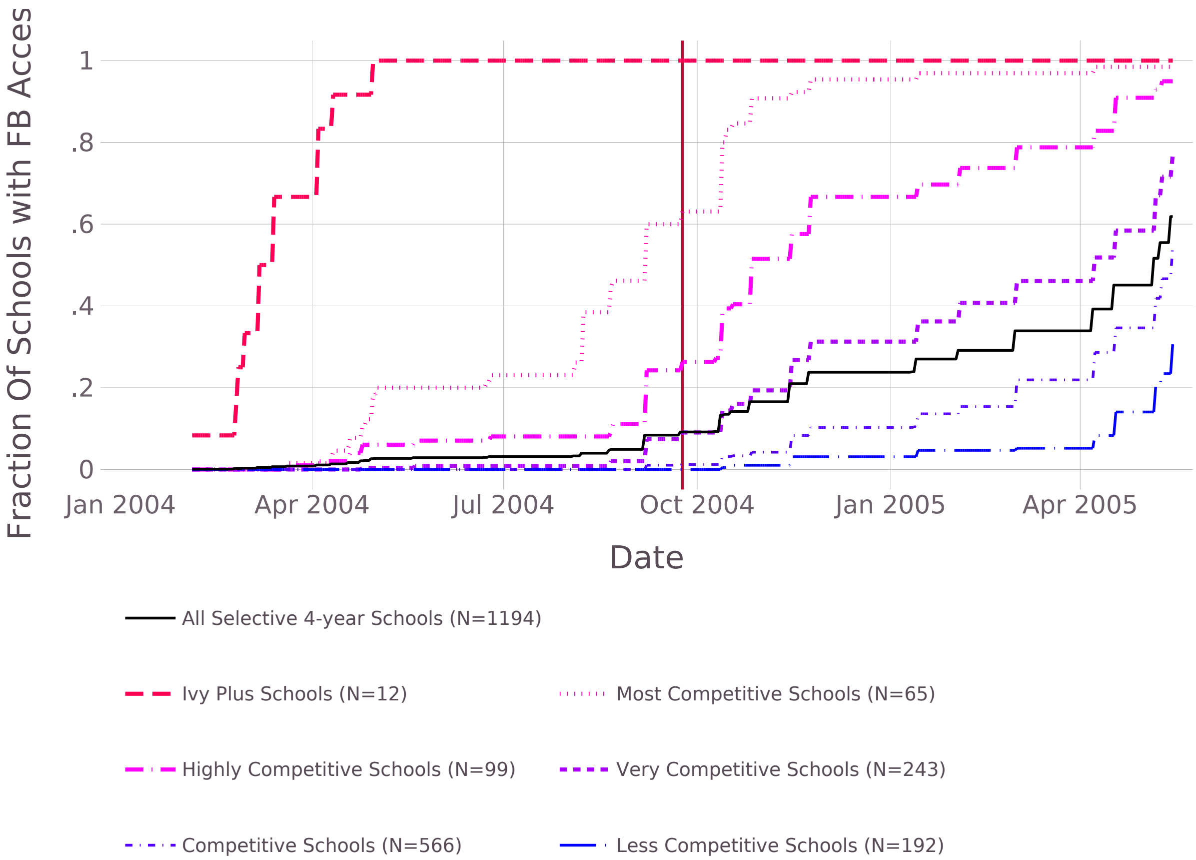


Figure 1: Facebook Entry to Colleges over Time, By School Tier

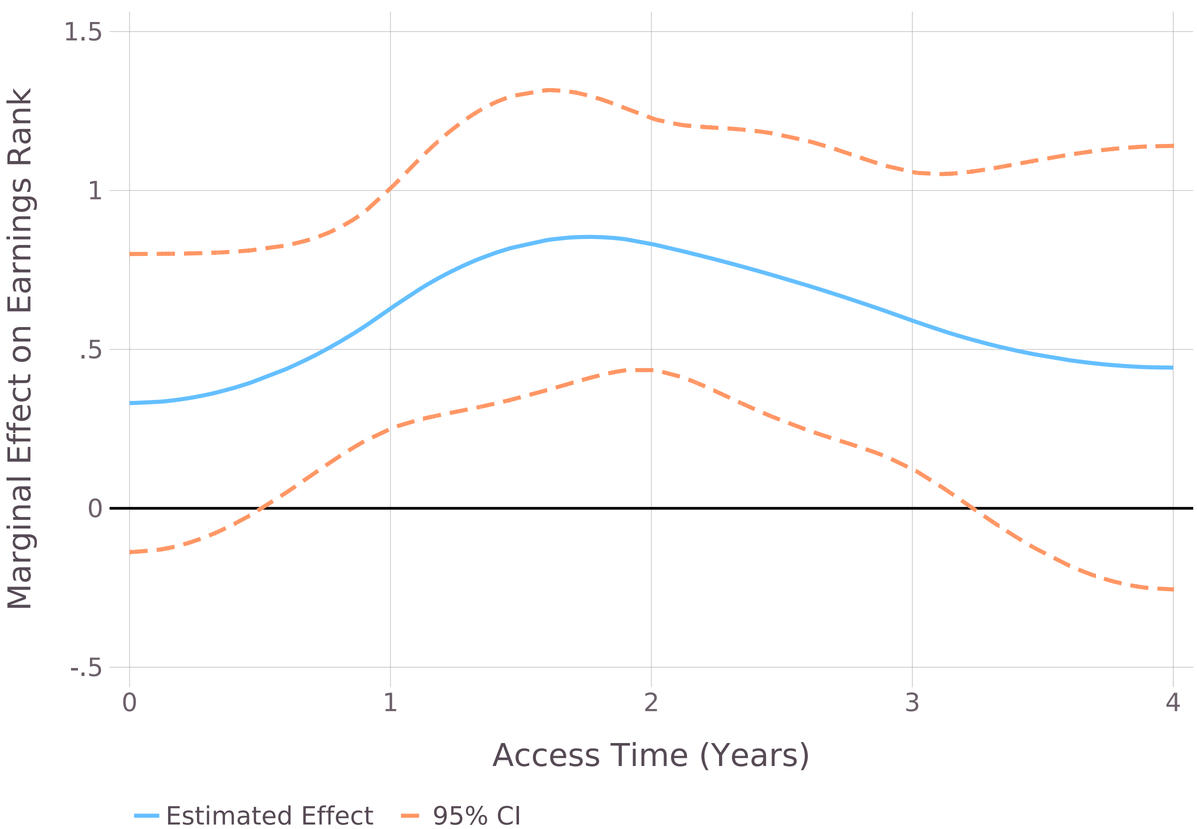


Figure 2: Nonlinear Effects from Facebook Exposure on Earnings Rank

Schools:	(1) Sample	(2) FB 100	(3) Late FB	(4) All Colleges	(5) P-value
Access Time (Years)	1.47	1.73	1.43		
	(1.54)	(1.61)	(1.53)		
Earnings Rank	65.41	70.61	64.59	63.25	0.000
	(6.26)	(5.21)	(6.02)	(6.74)	
Earnings Mean (USD)	49,711	70,292	46,456	45,433	0.000
	(17,348)	(27,357)	(12,359)	(15,680)	
Employment Rate	92.13	92.11	92.14	91.81	0.000
	(2.58)	(1.75)	(2.69)	(2.87)	
% Married	46.64	43.01	47.21	47.79	0.000
	(14.81)	(15.56)	(14.60)	(15.04)	
Parent Earnings Rank	67.86	75.67	66.62	64.09	0.000
	(8.84)	(6.57)	(8.51)	(9.92)	
% Female	53.01	53.24	52.98	51.67	0.000
	(16.76)	(12.25)	(17.35)	(19.08)	
% Hispanic	5.51	6.34	5.39	5.51	0.362
	(7.48)	(4.06)	(7.86)	(7.46)	
% Black	9.83	6.43	10.35	9.80	0.011
	(16.44)	(7.22)	(17.36)	(16.41)	
SAT Math Score (200-800)	553.09	646.27	538.07	531.47	0.000
	(67.48)	(59.48)	(55.61)	(66.33)	
Cohort Size	2,018	3,059	1,853	2,946	0.000
	(3,736)	(3,603)	(3,730)	(5,964)	
% on FB (2005)		40.86			
		(28.17)			
# of FB Friends (2005)		57.05			
		(34.95)			
Observations	6,071	799	5,272	11,863	
Number of Schools	728	100	634	1,177	
Number of Students (1000s)	6,542	1,552	4,990	9,313	
# of Ivy Plus Schools	12	12	0	12	
# of Elite Schools	64	39	25	65	
# of Highly Selective Schools	94	25	72	99	
# of Selective Schools	558	24	537	1001	

Standard deviations in parentheses. Column (1) displays summary statistics for the set of cohorts in my sample. Column (2) displays summary statistics for the set of cohorts at the first 100 schools to receive Facebook. Column (3) displays summary statistics for the set of cohorts at the remaining 660 schools that receive Facebook by August 2005. Column (4) displays summary statistics for the set of cohorts at all selective colleges that are covered by IPEDS and the Mobility Report Card in the continental United States. P-value is from ANOVA test of equality of means between cohorts in my sample and those at other selective colleges. Number of schools/students are reported at the super-OPEID level, the level of aggregation for the Mobility Report Card Data. # of FB friends denotes the conditional average number of friends within a cohort, conditional on having a Facebook account.

Table 1: Summary Statistics

	Pr(On Facebook)	IHS(Degree)		
	(1)	(2)	(3)	(4)
Access Time (Years)	0.0428***	0.234***	0.213***	0.198***
	(0.0116)	(0.0638)	(0.0571)	(0.0540)
Observations	2532412	2532412	988412	988411
R-squared	0.286	0.295	0.179	0.189
Outcome Mean	0.390	1.760	4.510	4.510
University Fixed Effects	Yes	Yes	Yes	Yes
Cohort Fixed Effects	Yes	Yes	Yes	Yes
Demographic Controls	No	No	No	Yes
Sample	All	All	FB Users	FB Users

OLS estimates shown. Standard errors clustered at the university level. IHS(Degree) denotes the inverse hyperbolic sine of the number of Facebook friends who were peers in college (within 3 classes of the student). Access time is the length of a student's access to Facebook while in college as of September 30th, 2005. All regressions include university and cohort fixed effects. Demographic controls include fixed effects for major and gender, as self-reported on Facebook. Sample "All" denotes all students attending the first 100 universities to receive Facebook access, while "FB Users" denotes students with Facebook accounts.
* $p < .1$,** $p < .05$, *** $p < .01$

Table 2: Relationship between access time and Facebook network growth

Panel A: Earnings Rank				
	(1)	(2)	(3)	(4)
Access Time (Years)	0.74***	0.62***	0.62***	0.63***
	(0.07)	(0.17)	(0.17)	(0.17)
Observations	5748	5748	5744	5696
R-Squared	0.93	0.95	0.95	0.95
University Fixed Effects	Yes	Yes	Yes	Yes
Tier-Year Fixed Effects	No	Yes	Yes	Yes
Demographic Controls	No	No	Yes	Yes
Degree Composition Controls	No	No	No	Yes
Panel B: Employment Rate				
	(1)	(2)	(3)	(4)
Access Time (Years)	0.62***	0.34***	0.39***	0.39***
	(0.04)	(0.13)	(0.13)	(0.13)
Observations	5748	5748	5744	5696
R-Squared	0.77	0.81	0.81	0.81
University Fixed Effects	Yes	Yes	Yes	Yes
Tier-Year Fixed Effects	No	Yes	Yes	Yes
Demographic Controls	No	No	Yes	Yes
Degree Composition Controls	No	No	No	Yes
Panel C: Earnings Rank| Working				
	(1)	(2)	(3)	(4)
Access Time (Years)	0.43***	0.38**	0.36**	0.37**
	(0.06)	(0.15)	(0.15)	(0.15)
Observations	5748	5748	5744	5696
R-Squared	0.95	0.96	0.96	0.96
University Fixed Effects	Yes	Yes	Yes	Yes
Tier-Year Fixed Effects	No	Yes	Yes	Yes
Demographic Controls	No	No	Yes	Yes
Degree Composition Controls	No	No	No	Yes
Panel D: Std. Dev. of Earnings Rank				
	(1)	(2)	(3)	(4)
Access Time (Years)	-0.28***	-0.23**	-0.28***	-0.29***
	(0.03)	(0.10)	(0.10)	(0.10)
Observations	5748	5748	5744	5696
R-Squared	0.81	0.85	0.85	0.85
University Fixed Effects	Yes	Yes	Yes	Yes
Tier-Year Fixed Effects	No	Yes	Yes	Yes
Demographic Controls	No	No	Yes	Yes
Degree Composition Controls	No	No	No	Yes

OLS estimates shown. Standard errors are clustered at university level. Outcome variable in each column denoted by panel title. All columns include marital status as a control. Demographic Controls denotes controls for marital status in 2014, race, gender, parent income, the 25th and 75th percentile for test scores, a dummy for missing test scores, a dummy for using the ACT to report test scores, the admissions rate, and a dummy for missing the admissions rate. Degree Composition Controls denotes controls for the fraction of students graduating with degrees in each of the 8 major categories defined by the College Board.
* $p < .1$, ** $p < .05$, *** $p < .01$

Table 3: Relationship between Access and Earnings

Outcome: Fraction of Students in Quantiles of National Cohort Earnings Distribution				
	(1)	(2)	(3)	(4)
	$< 1\%$-60%	60%-80%	80%-90%	90%-100%
Access Time (Years)	-0.568**	-0.193	0.371	0.779***
	(0.253)	(0.218)	(0.230)	(0.264)
Observations	5744	5744	5744	5744
R-Squared	0.93	0.91	0.78	0.97
Fraction of Students (%)	28.12	24.89	17.23	21.90
University Fixed Effects	Yes	Yes	Yes	Yes
Tier-Year Fixed Effects	Yes	Yes	Yes	Yes
Demographic Controls	Yes	Yes	Yes	Yes
Pr(Coefficient Same as Column 4)	0.0023	0.0155	0.2953	-

OLS estimates shown. Standard errors are clustered at university level. Demographic Controls denotes controls for marital status in 2014, race, gender, parent income, the 25th and 75th percentile for test scores, a dummy for missing test scores, a dummy for using the ACT to report test scores, the admissions rate, and a dummy for missing the admissions rate. * $p < .1$,** $p < .05$, *** $p < .01$

Table 4: Relationship between Access and Earning Quantiles

Outcome: Earnings Rank of Students by Subgroup								
	All	Parent Income Quintiles					Gender	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	All	0%-20%	20%-40%	40%-60%	60%-80%	80%-100%	Male	Female
Access Time (Years)	0.621***	-0.346	0.898**	0.138	0.508*	0.805***	0.435**	0.806***
	(0.169)	(0.673)	(0.405)	(0.346)	(0.271)	(0.213)	(0.195)	(0.230)
Observations	5748	5748	5748	5748	5748	5748	5597	5700
R-Squared	0.95	0.61	0.73	0.82	0.86	0.90	0.93	0.93
Fraction of Students	100%	6.6%	10.7%	16.3%	24.3%	42.1%	56.2%	43.8%
Outcome Mean	65.414	60.121	62.333	63.980	66.000	67.305	69.093	62.423
Outcome SD	6.263	8.199	6.944	6.329	5.996	6.185	6.339	6.741
University Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Tier-Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Marriage Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Degree Composition Controls	No	No	No	No	No	No	No	No
Pr(Coefficient Same as Column 1)	-	0.1351	0.4606	0.1039	0.6038	0.2119	0.1781	0.2581
Pr(Coefficient Same as Column 6)	-	0.0961	0.8272	0.0726	0.3327	-	-	-
Pr(Coefficient Same as Column 7)	-	-	-	-	-	-	0.1735	-

OLS estimates shown. Standard errors are clustered at university level. Marriage Controls denotes controls for marital status in 2014. * $p < .1$,** $p < .05$, *** $p < .01$

Table 5: Relationship between Access and Earnings, by Parent's Income and Gender

Outcome: Earnings Rank				
	(1)	(2)	(3)	(4)
Access Time (Years)	0.71***	0.59***	0.63***	0.51***
	(0.17)	(0.17)	(0.17)	(0.18)
Access Time * School In City	-0.09*			-0.11**
	(0.05)			(0.05)
Access Time * UG Enrollment > 5000		0.02		0.16**
		(0.05)		(0.07)
Access Time * Private School			0.08	0.20**
			(0.06)	(0.08)
Observations	5683	5683	5683	5683
R-Squared	0.95	0.95	0.95	0.95
University Fixed Effects	Yes	Yes	Yes	Yes
Tier-Year Fixed Effects	Yes	Yes	Yes	Yes
Demographic Controls	Yes	Yes	Yes	Yes

OLS estimates shown. Standard errors are clustered at university level. Demographic Controls denotes controls for marital status in 2014, race, gender, parent income, the 25th and 75th percentile for test scores, a dummy for missing test scores, a dummy for using the ACT to report test scores, the admissions rate, and a dummy for missing the admissions rate. School In City is a dummy variable for a school being located in the central city of an MSA or CMSA, according to IPEDS. UG Enrollment > 5000 is a dummy for a school having on average > 5,000 undergraduates enrolled during my sample. Private school is a dummy for a school being a private (not-for-profit) university. No for-profit private universities are in my sample. * $p < .1$,** $p < .05$, *** $p < .01$

Table 6: Heterogenous Effects of Facebook Access by School Characteristics

	(1) No Elite Schools	(2) No Super OPEIDS	(3) No Recession Grads	(4) HH Wages	(5) AGI	(6) Weighted	(7) Log(Mean $)
Access Time (Years)	0.574***	0.635***	0.559**	0.340**	0.276**	0.713***	0.012*
	(0.176)	(0.190)	(0.231)	(0.142)	(0.135)	(0.138)	(0.007)
Observations	5154	5291	4235	5744	5744	5744	5744
R-Squared	0.94	0.95	0.97	0.97	0.97	0.96	0.97
University Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Tier-Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Demographic Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Degree Controls	No	No	No	No	No	No	No

OLS estimates shown. Standard errors are clustered at university level. Dependent variable is earnings rank in each column. In columns (1), (2), and (3), I exclude certain groups from estimation, while in columns (4) and (5), I use alternative definitions of earnings rank than the baseline version used in the rest of the paper. Column (3) excludes students who entered in 2004 and 2005, so would have graduated in 2008 and 2009 during the Great Recession. Column (4) uses household instead of individual earnings in 2014 to calculate earnings rank, while column (5) uses Adjusted Gross Income (AGI), which also includes capital gains. Column (6) weights each observation by the number of students in the cohort. Column (7) uses logged mean earnings within a cohort as the outcome variable. Demographic Controls denotes controls for marital status in 2014, race, gender, parent income, the 25th and 75th percentile for test scores, a dummy for missing test scores, a dummy for using the ACT to report test scores, the admissions rate, and a dummy for missing the admissions rate. * $p < .1$, ** $p < .05$, *** $p < .01$

Table 7: Robustness Checks

	Log(Employment HHI)		Log(# of Peer Coworkers)		Fraction on LinkedIn	
	(1)	(2)	(3)	(4)	(5)	(6)
Access Time (Years)	0.047***	0.054***	0.199***	0.073***	0.030***	-0.002
	(0.002)	(0.015)	(0.003)	(0.024)	(0.002)	(0.009)
Obs.	6063	5567	6063	5567	6063	5567
R-squared	0.827	0.853	0.981	0.986	0.927	0.941
Outcome Mean	-2.279	-2.287	4.299	4.277	0.604	0.608
Outcome SD	0.272	0.268	1.336	1.335	0.369	0.383
University Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
State-Tier-Year Fixed Effects	No	Yes	No	Yes	No	Yes

OLS estimates shown. Standard errors are clustered at the university level. Employment HHI is the sum of squared employment shares among LinkedIn alumni at the 15 most popular firms per cohort. # of Peer Coworkers the expected number of coworkers at your firm that graduated within 3 years of one's university class, conditional on working at one of the 15 most popular firms in their cohort. Fraction on linkedin is the fraction of all students graduating from a university in a given year that are on linkedin.

* $p < .1$, ** $p < .05$, *** $p < .01$

Table 8: Relationship between LinkedIn Job Sorting and Facebook access

A Appendix

A.1 Event Study Analysis of Early Network Formation

In this section, I provide visual evidence on the effect of Facebook access on early adoption of the Facebook platform. Let $\mathbb{1}\{\text{Early}\}_i$ be an indicator for whether Facebook was released to this university i before May 30th, 2004 (an "early access" school). This indicates whether the class of 2004 would have been exposed to Facebook prior to graduation. I estimate an event study [Angrist and Pischke, 2008] in Facebook usage $D_{i,t}$ by graduating class t according to a cohort's early and late access (after May 30th, 2004) status, by running the following panel regression:

$$D_{i,t} = \alpha_i + \delta_t + \sum_{y=2002}^{2009} \Delta_t \times \mathbb{1}\{\text{Early}\}_i + \epsilon_{i,t}$$

$\mathbb{1}\{\text{Early}\}_i$ is the treatment variable, whose effect I allow to vary by year. Δ_t captures the differential effect of early Facebook access by graduating class. If Facebook access increases (decreases) the outcome $D_{i,t}$, I would expect there to be increase (decrease) in Δ_t in 2004, when only the early access class of 2004 had access to Facebook while in college, followed by a decrease (increase) in Δ_t in 2005, when both schools had Facebook.

Figure A1 plots the level of Facebook usage by class in the 2005 Facebook data, separately drawn for those for who the class of 2004 had access , and those who do not. It confirms adoption was rhigh for those who were in school while Facebook was introduced, and adoption was greater at early access schools. Figure A2 is a plot of the differences in Facebook usage Δ_t by early / late access status among Facebook 100 Schools, with 95% confidence intervals. I document the difference by class / access status on both the intensive and extensive margin of Facebook usage. In Panel (a), I plot the difference in the fraction of students in each cohort signed up for Facebook. Their is a sharp increase in 2004, signifying that on average 13% more of the class of 2004 at early access schools had Facebook accounts than their late access counterparts, relative to the difference in 2002, the base period. After 2005, this difference decreases to 5%, as the late access schools receive access to Facebook, and eventually approaches levels near the difference in 2003, for which no graduating classes would have had access to Facebook while on campus. Statistically, there is no difference between Δ_{2003} and Δ_t for $t \geq 2005$.

In Panel (b) I plot the difference in the intensive margin of Facebook usage, the mean log-degree among Facebook users. Here, $D_{i,t} = \frac{1}{S_{i,t}} \sum_s \log(d_s)$, where d_s signifies the degree / number of Facebook friends of a student s in cohort i, t, *conditional* on having a Facebook account, and $S_{i,t}$ is the number of Facebook users in cohort i, t. Log degree is used because degree obeys a power law distribution, as is common in social network data. The plot documents a similar shape to Panel (a) for how many friends Facebook users have by class, consistent with the timing of Facebook's entry. This suggests that Facebook access mattered for early of Facebook usage. In both plots, before 2005, the difference Δ_t is increasing. This may be due to indirect network effects coming from the Facebook platform. A member of the class of 2003 at an early access school, for example, is likely to have more potential friends on the platform because the cohort below them (the class of 2004) had access to Facebook while on campus.

A.2 Earnings Rank Variance Calculation

Using earnings rank data from Chetty et al. [2017a], I have a measure of the CDF of earnings rank R within each cohort at finite discrete points $r_j \in [0., .2, .4, .6, .8, .9, .95, .99, 1.]$ Therefore, I can construct an estimate of the second moment of earnings ranks as follows, treating these discrete points as an estimate of the empirical CDF of earnings:

$$\begin{aligned} E[R^2] = \int R^2 dF &\approx \int R^2 d\hat{F} \\ &= \sum_j Pr(R \in [r_j, r_{j+1}]) E[R^2 | R \in [r_j, r_{j+1}]] \\ &= \sum_j Pr(R \in [r_j, r_{j+1}]) \frac{1}{3} (r_{j+1}^3 - r_j^3)/(r_{j+1} - r_j) \end{aligned}$$

This estimate assumes the empirical CDF (which is flat between the percentiles r_j) is the true CDF. Thus,conditional on a quantile, the empirical CDF implies that earnings ranks are uniformly distributed between cutoffs r_j, which allows us to calculate the conditional expectation of R^2 between percentiles. An analogous calculation can be done for the first moment. The earnings rank variance is then calculated as $\hat{V}(R) = \hat{E}[R^2] - \hat{E}[R]^2$. Intuitively, this measure will assign a higher variance to cohorts for which I observe a higher fraction of students near $R = 0$ and $R = 1$, and should capture whether earnings within a cohort are highly dispersed.

A.3 Placebo Regression

I provide supplementary evidence in Table A1 that my identifying assumption holds by validating a null effect of Facebook access on predetermined observable characteristics of the cohort, once I control for potential confounders. This includes parent earnings rank, fraction female, fraction of students that are racial minorities (Black, Hispanic, or Native American), and average SAT math score in each cohort. I control for year×tier fixed effects because increased exposure corresponds to later years of the sample, and if their were trends of, for example, increased female participation in postsecondary education, it would be mechanically correlated with access time. I also keep in my demographic controls. The results are shown in Table A1. There is no statistically significant relationship between Facebook exposure and predetermined variables, as I would expect.

Despite being uncorrelated with predetermined observables, once I include appropriate controls, there is still a worry that these cohorts are on differential trends in ways that are unobservable. To test this, I perform a placebo test similar to that in Enikolopov et al. [2020]. I construct a placebo access variable $\tau_{i,t}$ as if Facebook was introduced four years earlier four years after its actual release date. For example, I assume Harvard received Facebook in February 4th, 2000, followed by Columbia on February 25th, 2000, and so on. This preserves the relative differences between schools, but uses variation in these relative differences for cohorts who graduated before Facebook was actually released. This will test whether the cohorts from schools receiving Facebook earlier were on differential wage trends, which could explain the earnings findings in my main results. I shift access by four years so that there is no overlap in exposure across cohorts: all students in the class of 2004, the first cohort to actually receive Facebook, would have full access time of 4 years according to this placebo measure.[24] Table A4 displays the results from regressing the placebo access time on the 4 outcome variables of interest in Table 3. None of the effects are statistically different from zero. We can reject the hypothesis that the effect is equal to the main effects in Table 3 at 10% confidence for all except earnings rank, conditional on working, and at 5% confidence level for earnings rank, the main outcome variable of the paper.

A.4 2005 Network Effects on Earnings

In the main portion of the paper, I only discussed the indirect effects of Facebook on future outcomes by looking at the effect of access time on outcomes. I am able to do this by arguing that conditional on university and state-tier-year fixed effects, remaining variation in access dates is effectively random across cohorts. While I cannot observe the social network for all students at the time of their graduation, which would be my preferred measure of network strength for future outcomes, I do observe one snapshot of the Facebook social network in September 2005 from the Traud et al. [2012] dataset. Under more demanding assumptions than those made in the main portion of the paper, I can use this data to estimate the long-term effect of the Facebook network structure in 2005 on future earnings in 2014 using instrumental variables (IV) regression. For this section, I focus my analysis on the Facebook 100 schools for which I have 2005 network data. For the first stage of the endogenous variable (Facebook network strength), I use the modified access variable $\tilde{\tau}_{i,t}$ defined in Equation 2 as my instrument to predict network size, denoted $D_{i,t}$. $\tilde{\tau}_{i,t}$ is the same access time measure used in Section 3 for the network strength regressions in Table 2. In order to improve the fit of the first-stage, I use a cubic polynomial of $\tilde{\tau}_{i,t}$ to allow for changes in returns from access time to network formation. I then estimate the following two-stage equation to recover the causal effect of network strength on future earnings:

$$D_{i,t} = \delta_i + \kappa_{t,\text{tier}} + \beta \mathbf{X_{i,t}} + \psi_1 \tilde{\tau}_{i,t} + \psi_2 \tilde{\tau}_{i,t}^2 + \psi_3 \tilde{\tau}_{i,t}^3 + \epsilon_{i,t} \tag{5}$$

$$S(w_{i,t}) = \alpha_i + \phi_{t,\text{tier}} + \gamma \mathbf{X_{i,t}} + \theta D_{i,t} + \nu_{i,t} \tag{6}$$

where $D_{i,t}$ is a measure of a cohort's Facebook network strength. Using the conditionally exogenous variation in network structure coming from access time $\tilde{\tau}_{i,t}$ in Equation 5, I may be able to interpret the parameter θ in Equation 6 as the causal effect of a cohort's Facebook participation levels on a statistic of earnings rank, $S(w_{i,t})$. In contrast to previous results which look at the indirect effect of access time, the IV regression specification requires the assumption of an exclusion restriction that Facebook access effects earnings *only* through the online Facebook network strength measure in 2005. For example, this would rule out a mechanism where Facebook access increases a cohort's more general online activity on other social platforms, or improves offline connectivity within a cohort. More generally, one might imagine my measures of network strength (Facebook participation rate and average number of online friendships) are imperfect measures of the ways Facebook access affects later labor market outcomes. This exclusion restriction is a fairly strong assumption, and for this reason, results of this IV regression should be interpreted with caution (and are limited to treatment in the Appendix). Instead, I view the analysis in this section as a validation exercise that conditionally exogenous variation in Facebook network structure in 2005 can explain earnings increases. I consider 3 endogenous variables as my measures of Facebook network strength, $D_{i,t}$. The first is the fraction of students in a cohort on Facebook, measured according to the number of students who self-report belonging to a graduating class in the network data from Traud et al. [2012]. The second is the average number of Facebook friends of students in cohort i, t among those who have Facebook accounts. The third is a combination of the previous two measures; the unconditional number of Facebook friends. This variable is constructed by imputing zero Facebook friends for students who are not on the Facebook graph in the Traud et al. [2012] data and including these zero-degree nodes in the average taken over students in a cohort.

Table A10 displays the IV estimates of the parameter θ from Equation 6. I consider 2 measures of future labor outcomes: The earnings rank, as I have focused on throughout this paper, and the employment rate, both in 2014 using the Chetty et al. [2017a] income data. Across the board, the F-statistic for the first stage of the excluded instruments (access time) ranges from 7.67-8.6, confirming my instrument is not weak and captures substantial variation in the Facebook network buildout across campuses/cohorts. However, these F-statistics are below standard thresholds for an IV (10). The estimated first-stage cubic relationship between access times and all 3 network measures is documented in Figure A4. The first stage plots suggest most of the effect on the 2005 network is captured in the first 6-9 months of access.

The relationship between earnings rank and Facebook friendship measures/Facebook takeup is statistically insignificant, but qualitatively demonstrates a positive relationship between Facebook friends in 2005 and earnings 9 years later.

For the employment rate, I find a positive, statistically significant relationship for both measures of network friendships. My preferred measure of network strength, unconditional average number of Facebook friends, suggests an additional online friend in 2005 for each student in a cohort increases the cohort's employment rate in 2014 by 0.059 percentage points. This effect size is probably best understood in terms of standard deviations, since the size of Facebook networks in 2005 is not comparable to the monolithic size of Facebook's online social network in 2014.[25] I provide the effect of a standard deviation increase in each endogenous network variable according to the estimated IV coefficients in the bottom row of the table. The estimates suggest a standard deviation increase in Facebook friends in 2005 increases the employment rate in a cohort by 1.7 percentage points in 2014.

B A Simple Model of Labor Markets on Local Networks

B.1 Wages and Friendships

I now present an extremely stylized model of social networks on the labor market to motivate the potential mechanisms through which SNS network effects operate on wages. It is motivated in part by the model in Calvó-Armengol and Jackson [2007], but also adapted to this paper's setting of student formed networks and future labor outcomes. Assume that all workers in the economy are homogeneous in skill and equally attractive to employers. Each worker i is embedded in a static social network graph $\mathcal{G}$ with N neighbors/friends, n of which are employed. Assume employment was not an input into social network decisions, i.e. there are no friendship preferences over individuals' employment status. In this setting, the static social network is formed during college, when no one has a job yet, so this assumption is reasonably plausible. Each period, worker i hears about job offerings from each of their n employed neighbors. Assume that these job openings are announced each period and wages W for those job openings are drawn i.i.d. from a distribution $W \sim F$. In particular, assume F is a Fréchet distribution with CDF $F(w) = Pr(W \leq w) = \exp(-(w/s)^{-\alpha})$ and mean $E[w] = s\Gamma(1 - 1/\alpha)$, where $\alpha > 0$ is a shape parameter, $s > 0$ is a scale parameter, and Γ is the gamma function.

If a worker is currently employed, they always prefer to stay with their current job (possibly due to various transition costs/frictions associated with switching jobs). However, with some exogenous probability p_U, in each period employed workers are separated from their current job. If they are separated from their firm, they first engage in job search, and encounter a job offer $w_{n+1} \sim F$ found via direct job search, in addition to those openings they hear about from their employed friends.

In each period, an unemployed individual can apply to only one of the job offers they hear about. Each job offer has an equiprobable probability of hiring the unemployed worker,

p_H. Critically, I assume this employment probability p_H does not depend on the social network structure of i's neighbors. This means that the probability of being employed next period is unrelated to network size n. I assume this to isolate the effect of social network structure on wages, since that is the primary focus of my empirical analysis. By assuming p_H is exogenous, I assume social network graphs are sufficiently *local* that, between any given pairs of neighbors, there are not enough individuals applying to the same job opening from a common pool of friends that p_H is endogenous to this structure. This assumption could just mean that there are many other applicants to these jobs besides i's neighbors, so that p_H does not change very much if , say, i's neighbors also get unemployed, and apply to the same job offer as i does. Alternatively, if there are many individuals who hear about the job offer via non-social job search, the exogeneity assumption for p_H will approximately hold.

Among the job offers i is aware of, they apply to the highest wage offer they hear about, $\tilde{w}_i = \max(w_1, w_2, \ldots, w_{n+1})$. Using wage as a sole criteria could be because workers only care about wages, or only know about wages from each job offer. This would also be consistent with wages increasing monotonoically with other unobservable characteristics of jobs that make it more appealing, such as workplace culture or benefits. If workers do not get hired, they remain unemployed into the next period.

Conditional on currently being employed with wage w_t, and the number of employed friends n the expected wage of individual i in the next period $t+1$ is:

$$E[w_{i,t+1}|w_{i,t} > 0, n] = (1 - p_U)w_{i,t} + p_U p_H E[\tilde{w}_i] \tag{7}$$

Now, assume that this system, after many periods, converges to a steady state, where the distribution of wages and employment is invariant . This would imply a steady state employment rate $\tilde{p}_E = p_H/(p_H + p_U)$, since being hired to a job is independent of network size n, and that $E[w_{i,t}|n] = E[w_{i,t+1}|n]$. This implies the following steady state mean wage among employed individuals[26]:

$$E[w_{i,t+1}|w_{i,t} > 0, n] = p_H E[\tilde{w}_i] = p_H (n+1)^{1/\alpha} s\Gamma(1 - 1/\alpha)$$

where I use the fact that the maximum of m Frechet random variables with shape parameter α and scale parameter s is itself Frechet with shape parameter α and scale parameter $m^{1/\alpha}s$. Taking logs, the equation is:

$$\log(E[w_{i,t+1}|w_{i,t} > 0, n]) = \log(p_H s\Gamma(1 - 1/\alpha)) + \frac{1}{\alpha}\log(n+1) \equiv \gamma + \frac{1}{\alpha}\log(n+1) \qquad (8)$$

where γ is a constant depending on the primitives α, s, p_H. Thus, log mean wages for an individual i in a social network are positively associated with the number of employed friends in this stylized model in steady state. Since employment is orthogonal to friendship formation, wages are also positively associated with the overall number of friends N in the social network. This motivates some of the later reduced form analysis.

B.2 Friendship Formation

Now focus more explicitly on our Facebook setting to see how the above relationship relates to improvements in social networking technology. I model the network formation process of the above static social network. Suppose that meetings for friendship arrive during college according to a continuous time poisson process with rate λ for a fixed amount of time T (e.g. 4 years of college). Each meeting, the value of the potential friendship is drawn i.i.d. uniform $v \sim U[0,1]$[27]. While friendships all have some value, there is also a cost c_0 associated with maintaining a friendship, so meetings convert into friendships only if $v > c_0$. The expected number of friendships N after t periods is given as:

$$E[N|t] = \lambda t(1 - c_0)$$

Now, suppose that at time $T-\tau$, new online social networking technology is introduced such that the costs of maintaining a friendship are now reduced. Specifically the cost maintaining a friendship during $[T-\tau, T]$ is $c_1 < c_0$). In this case, the expected number of friends after T periods is:

$$E[N|\tau] = \lambda(T-\tau)(1-c_0) + \lambda\tau(1-c_1) = \lambda T(1-c_0) + \lambda\tau(c_0 - c_1)$$

Using the property that if $X \sim Binomial(p, N)$ and $N \sim Poisson(\lambda)$, then X is itself distributed $Poisson(\lambda p)$, the total number of employed friends n of any given individual is :

$$n \sim Poisson(\tilde{p}_E \lambda\big(T(1-c_0) + \tau(c_0 - c_1)\big))$$

Ideally, I would input the random variable n directly into equation 7 and get the unconditional expected value of the maximum wage offer, leading to the unconditional expected wages for the invariant distribution. However this will be intractable and fail to yield a closed form solution due to the additional variation from the maximum wage offer conditional on n. Instead, I take a mean field approximation, as is common in the network formation literature (see for example Jackson and Rogers [2007], Jackson [2010]). Instead of accounting for full random variation of the number of employed friends n at any given moment, I assume n is a deterministic process evolving as a function of τ, and is approximately $n(\tau) \approx E[n|\tau] = \tilde{p}_E \lambda\big(T(1-c_0) + \tau(c_0 - c_1)\big)$. Using this approximation in equation 8, I have an expression for mean wages as a function of τ and the other model primitives:

$$\log(E[w_{i,t+1}|w_{i,t} > 0, \tau]) = \frac{1}{\alpha}\log(\tilde{p}_E \lambda\big(T(1-c_0) + \tau(c_0 - c_1)\big) + 1) + \gamma \tag{9}$$

For exposition, assume that $\log(n(t) + 1) \approx \log(n(t))$, which is reasonable if $E[n|\tau]$ is even moderately large. Then the simplified expression is:

$$\log(E[w_{i,t+1}|w_{i,t} > 0, \tau]) \approx \gamma + \frac{1}{\alpha}\log(\tilde{p}_E \lambda) + \frac{1}{\alpha}\log((1-c_0)T + \tau(c_0 - c_1)) \tag{10}$$

Equation 10 shows that access time to the friendship cost reduction technology of an SNS can be a direct input into mean wages of the employed individuals in this simplified economy.

B.3 Simulation Evidence

In this section, I offer some simulation evidence on what introducing social networking technology earlier to a college campus network might do to wages / employment. For this, I consider a slightly more realistic model of the world, where social network size effects both wages and employment status: wages for job offers are uniformly i.i.d. $w \sim U[0,1]$, the number of friends N is fixed/identical across individuals in the economy (as before, $N = \lambda T(1-c_0) + \lambda\tau(c_0 - c_1)$), and there is no outside job offer w_{n+1} acquired outside the social network.

While each employed individual in the network hears about a job, they now pass it on to exactly *one* of their unemployed friends. If I hear about J job offers from friends, I take the one with highest wages w. Since in this model, multiple individuals do not hear about the same job offer, there is no competitive aspect to applying for a job once an individual has learned about it.

Given an employment rate of p_E in the economy, this implies a probability of hearing about a job from friend j as

$$p_O = Pr(\text{hear of job offer from friend } j) = \frac{p_E}{(1-p_E)N}$$

. Given an individual starting a period unemployed, the probability of remaining unemployed next period is

$$Pr(w_{i,t+1} = 0|w_{i,t} = 0) = \Pi_{j=1}^{n}(1 - p_O) = (1 - \frac{p_E}{(1 - p_E)N})^N$$

In steady state, the following equation governs the employment rate:

$$\begin{aligned}(1 - p_E) &= p_E(p_U) + (1 - p_E)Pr(w_{i,t+1} = 0|w_{i,t} = 0)\\ &= p_E(p_U) + (1 - p_E)(1 - \frac{p_E}{(1 - p_E)N})^N\end{aligned}$$

where p_U is the exogeneous separation rate. Given p_U, N, this equation implicitly defines the steady state employment rate, which cannot be written analytically but can be found via solving the root of the equation. Given an employment rate, steady-state wages are characterized as follows:

$$\begin{aligned}E[w_{i,t+1}|w_{i,t} > 0, N] &= (1 - p_U)E[w_{i,t+1}|w_{i,t} > 0, N] + p_U E[w_{i,t+1}|w_{i,t} = 0]\\ \Rightarrow E[w_{i,t+1}|w_{i,t} > 0, N] &= E[w_{i,t+1}|w_{i,t} = 0]\\ E[w_{i,t+1}|w_{i,t} = 0, N] &= \sum_{j=1}^{N} Pr(j \text{ job offers})E[\max(w_1, \ldots, w_j)|j \text{ job offers}]\\ E[w_{i,t+1}|w_{i,t} = 0, N] &= \sum_{j=1}^{N} Pr(j \text{ job offers})\frac{j}{j+1}\\ \Rightarrow E[w_{i,t}|N] &= \sum_{j=1}^{N} Pr(j \text{ job offers})\frac{j}{j+1}\end{aligned}$$

where $Pr(j$ job offers) obeys a binomial distribution with probability p_0 on N friends. While I cannot express this equation analytically, I can simulate it, to determine what the effect of an additional year of access to Facebook does to individuals in college. To do so, I pick some calibrated parameters for this simulation exercise:

- $p_U = 0.05$: each period, there is a 5% chance one is separated from their job

- $\lambda = 100$: each year, 100 potential friendship meetings occur on campus.

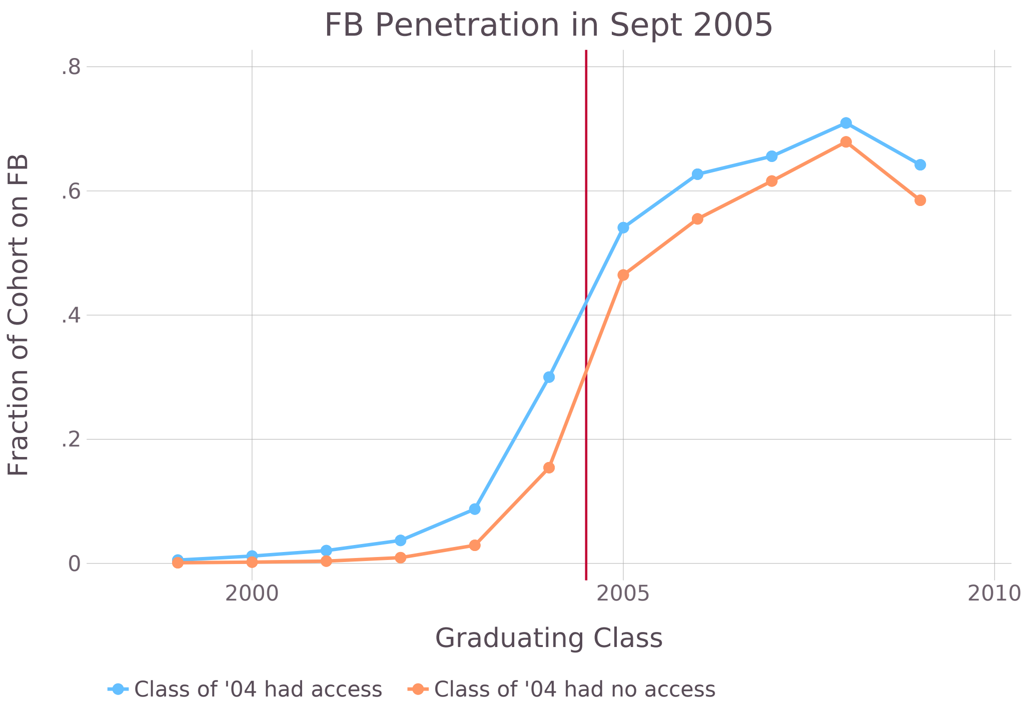
- $T = 4$: Friendship formation occurs over 4 years of college.

- $c_0 = 0.9$; before Facebook's introduction to a campus, meetings have a 10% chance of turning into friendships due to maintenance costs

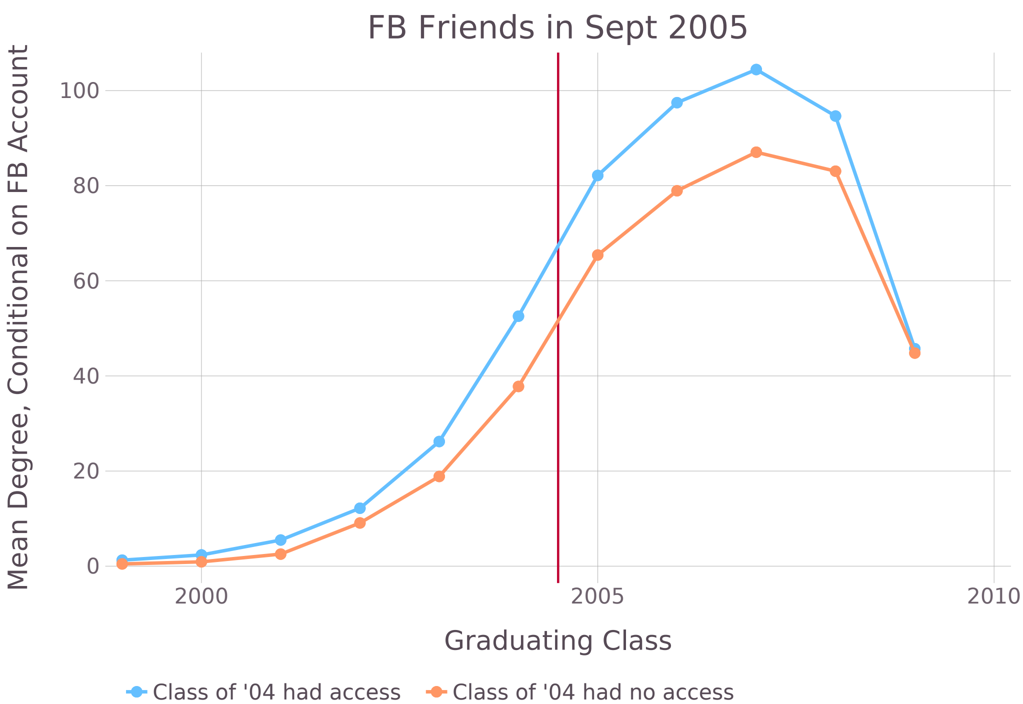
- $c_1 = 0.1$; after Facebook's introduction to a campus, meetings have a 90% chance of turning into friendships due to reduced maintenance costs.

I then simulate the friendships/steady state labor outcomes of students in the classes of 2004-2009 in a hypothetical campus governed by this model of wages/friendship formation, under two scenarios: one where Facebook is introduced in 2004, and one where Facebook is introduced in 2005. This should roughly correspond to the effect of an additional year of access time on wages, my main regression estimate.

Figures A5 and A6 present the results from the simulated economy. The difference peaks at the time of "late access" introduction (2005) then decays slowly after. This is broadly consistent with my finding in Figure 2 that access time is less effective after two years.



(a) Fraction of Users, by Cohort



(b) Number of Peer Friends, by Cohort

Figure A1: Facebook Usage levels based on receiving Facebook in 2004 or 2005 AY

(a) Fraction of students on Facebook

(b) Log # of Facebook Friends

Figure A2: Differences in Facebook Penetration in Sept 2005 Based on Receiving Facebook in 2004 or 2005 AY

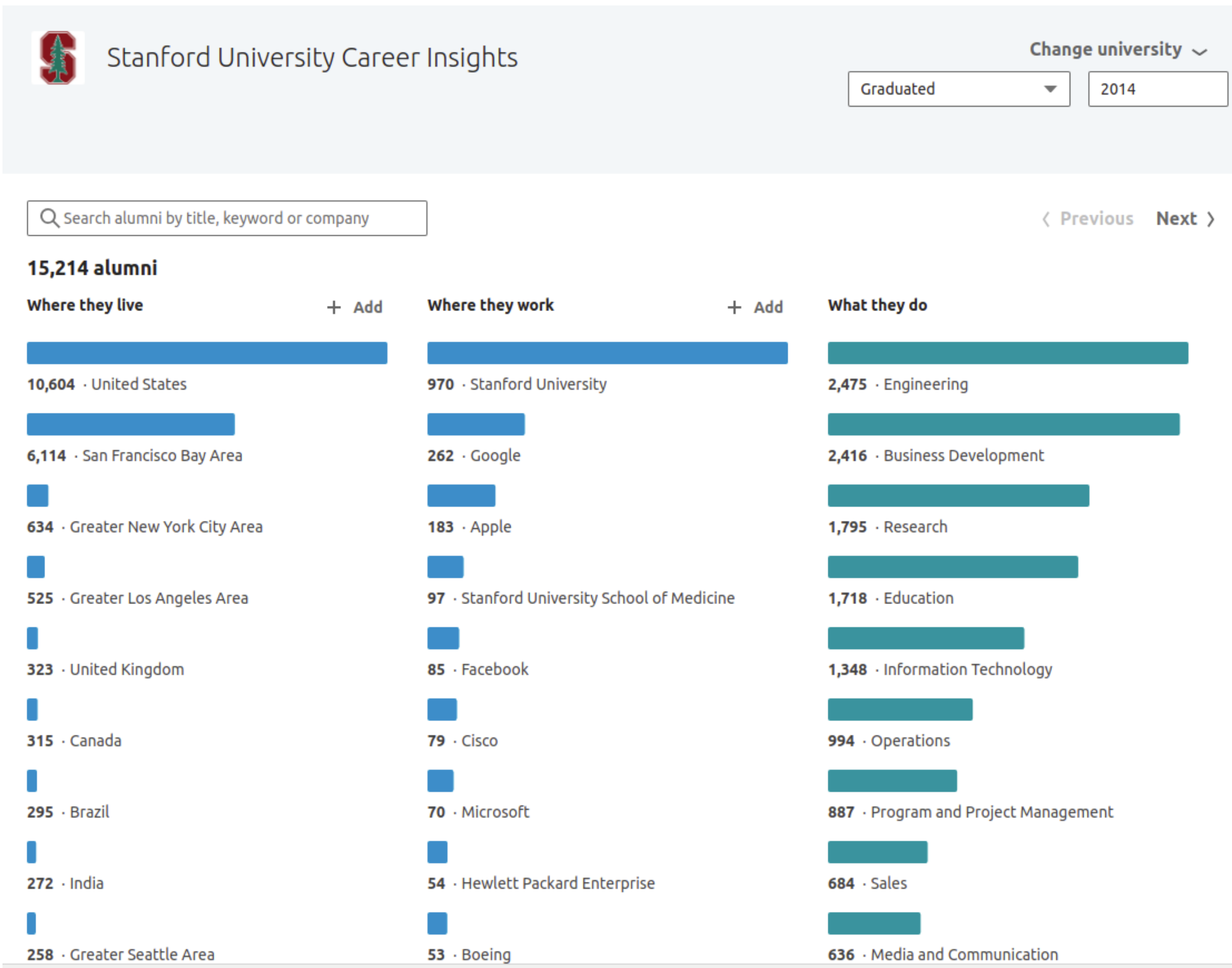


Figure A3: Example of Linkedin Career Insights Portal

	(1)	(2)	(3)	(4)
	Parent Earnings Rank	Fraction Female	Fraction Racial Minority	SAT Math score
Access Time (Years)	0.002	0.001	-0.008	-2.629
	(0.002)	(0.003)	(0.008)	(2.067)
Observations	5744	5744	5744	3492
R-Squared	0.98	0.97	0.98	0.98
University Fixed Effects	Yes	Yes	Yes	Yes
Tier-Year Fixed Effects	Yes	Yes	Yes	Yes
Demographic Controls	Yes	Yes	Yes	Yes

OLS estimates shown. Standard errors are clustered at university-level. SAT Math Score is the average of the 25th and 75th percentile of scores reported by colleges for each cohort to IPEDS. Fraction Racial Minority is the fraction of students who are hispanic, black or native american in a cohort. For Demographic Controls, I exclude controls colinear with each dependent variable. In column (1), I exclude parental income controls, but retain race, gender, SAT score, and marital status; in column (2), I exclude gender controls, but include race, SAT score, marital status, and parent income; in column (3), I include gender, parent income,SAT score, and marital status; finally, in column (4), I include all demographic controls besides SAT score and admissions rate.

* $p < .1$,** $p < .05$, *** $p < .01$

Table A1: Placebo Regression on Predetermined Observables

Outcome: Earnings Rank						
	(1)	(2)	(3)	(4)	(5)	(6)
Access Time (Years)	1.39***	1.37***	1.44***	0.61***	0.66***	0.80***
	(0.15)	(0.16)	(0.16)	(0.17)	(0.17)	(0.20)
Observations	5744	5744	5719	5744	5709	5258
R-Squared	0.95	0.95	0.95	0.95	0.95	0.96
University Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Other Fixed Effects	Year	Region*Year	State*Year	Tier*Year	Region*Tier*Year	State*Tier*Year
Demographic Controls	Yes	Yes	Yes	Yes	Yes	Yes

Table A2: Effect on Earnings Rank with Varying Time Fixed Effects

OLS estimates shown. Standard errors are clustered at university level. Dependent variable is the mean earnings rank in a cohort. Demographic Controls denotes controls for marital status in 2014, race, gender, parent income, the 25th and 75th percentile for test scores, a dummy for missing test scores, a dummy for using the ACT to report test scores, the admissions rate, and a dummy for missing the admissions rate. Region in Other Fixed Effects denotes the BEA Region a school is located in.

* $p < .1$, ** $p < .05$, *** $p < .01$

	Major Choice					Graduation	Marriage
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	STEM	Social Sciences	Arts/Humanities	Business	Health	Graduation Rate (%)	% Married
Access Time (Years)	0.299	1.051***	-0.741***	-0.459	-0.288	0.146	-1.072***
	(0.310)	(0.397)	(0.284)	(0.368)	(0.199)	(0.385)	(0.362)
Observations	5728	5728	5728	5728	5728	5699	5744
R-Squared	0.98	0.95	0.96	0.96	0.93	0.97	0.97
Outcome Mean	17.87	32.00	13.57	18.86	5.52	61.00	46.64
Outcome SD	14.718	11.873	9.013	12.243	6.748	17.123	14.801
University Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Tier-Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Demographic Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes

OLS estimates shown. Standard errors are clustered at university level. STEM denotes the % of students graduating in a science, math, or technology major. Social Sciences denotes the % of students graduating in a social science major. Arts/Humanities denotes the % of students graduating in an arts or humanities major. Business denotes the % of students graduating in a business major. Health denotes the % of students graduating in a health and medicine major. These aggregations of majors come from the College Board's major category definitions `https://bigfuture.collegeboard.org/majors-careers`. Graduation Rate denotes the % of students in a cohort who graduate with a 4-year degree in 6 years. % Married denotes the % of students who are married in 2014. Demographic Controls denotes controls for marital status in 2014, race, gender, parent income, the 25th and 75th percentile for test scores, a dummy for missing test scores, a dummy for using the ACT to report test scores, the admissions rate, and a dummy for missing the admissions rate.

* $p < .1$, ** $p < .05$, *** $p < .01$

Table A3: Effect on Major Choice and Earnings Rank

	(1)	(2)	(3)	(4)
	Earnings Rank	Employment Rate	Earnings Rank | Working	Std. Dev. of Earnings Rank
Access Time Shifted 4 Years Prior	-0.32	-0.12	-0.18	0.08
	(0.30)	(0.18)	(0.26)	(0.13)
Observations	5744	5744	5744	5744
R-Squared	0.95	0.81	0.96	0.85
University Fixed Effects	Yes	Yes	Yes	Yes
Tier-Year Fixed Effects	Yes	Yes	Yes	Yes
Demographic Controls	Yes	Yes	Yes	Yes
Pr(Same as Actual Access Time)	0.0319	0.0817	0.1600	0.0818

OLS estimates shown. Standard errors are clustered at university level. Access Time Shifted 4 Years Prior denotes a constructed measure of access time as if Facebook was released to Harvard on February 4th, 2000, and the subsequent release dates are the same relative to the Harvard. Demographic Controls denotes controls for marital status in 2014, race, gender, parent income, the 25th and 75th percentile for test scores, a dummy for missing test scores, a dummy for using the ACT to report test scores, the admissions rate, and a dummy for missing the admissions rate. Pr(Same as Actual Access Time) denotes the p-value from a test of the coefficient reported in the table being equal to the coefficient of each corresponding outcome variable in Table 3 Column (3). * $p < .1$,** $p < .05$, *** $p < .01$

Table A4: Placebo Test of Access Time Shifted 4 Years Before and After Facebook's Entry

	(1) Earn Rank	(2) Emp. Rate	(3) Earn Rank | Working	(4) Earn Rank SD
IHS(Access Time)	0.730***	0.585***	0.372*	-0.385***
	(0.230)	(0.209)	(0.202)	(0.144)
Observations	5744	5744	5744	5744
R-Squared	0.95	0.81	0.96	0.85
Outcome Mean	Yes	Yes	Yes	Yes
Outcome SD	Yes	Yes	Yes	Yes
University Fixed Effects	Yes	Yes	Yes	Yes

OLS estimates shown. Standard errors are clustered at university level. Dependent variable is the mean earnings rank in a cohort. IHS(Access Time) denotes the inverse hyperbolic sine of Facebook access time while in college (years). Demographic Controls denotes controls for marital status in 2014, race, gender, parent income, the 25th and 75th percentile for test scores, a dummy for missing test scores, a dummy for using the ACT to report test scores, the admissions rate, and a dummy for missing the admissions rate.
* $p < .1$,** $p < .05$, *** $p < .01$

Table A5: Effect IHS Access Time on Labor Market Outcomes

	(1) Earn Rank	(2) Emp. Rate	(3) Earn Rank | Working	(4) Earn Rank SD
Access Time (Years)	0.556***	0.360***	0.306**	-0.284***
	(0.171)	(0.135)	(0.151)	(0.097)
Observations	5744	5744	5744	5744
R-Squared	0.95	0.81	0.96	0.85
University Fixed Effects	Yes	Yes	Yes	Yes
Tier-Year Fixed Effects	Yes	Yes	Yes	Yes
Demographic Controls	Yes	Yes	Yes	Yes

OLS estimates shown. Standard errors are clustered at university level. Dependent variable is the mean earnings rank in a cohort. Demographic Controls denotes controls for marital status in 2014, race, gender, parent income, the 25th and 75th percentile for test scores, a dummy for missing test scores, a dummy for using the ACT to report test scores, the admissions rate, and a dummy for missing the admissions rate.
* $p < .1$,** $p < .05$, *** $p < .01$

Table A6: Effect of Access Time on Labor Market Outcomes without Marriage Controls

	(1)	(2)	(3)	(4)
	Earn Rank	Emp. Rate	Earn Rank | Working	Earn Rank SD
Access Time (Years)	0.614***	0.390***	0.354***	-0.284**
	(0.101)	(0.108)	(0.098)	(0.089)
Observations	5744	5744	5744	5744
R-Squared	0.95	0.81	0.96	0.85
Outcome Mean	Yes	Yes	Yes	Yes
Outcome SD	Yes	Yes	Yes	Yes
University Fixed Effects	Yes	Yes	Yes	Yes

OLS estimates shown. Standard errors are clustered at university level and entry year level. Dependent variable is the mean earnings rank in a cohort. Demographic Controls denotes controls for marital status in 2014, race, gender, parent income, the 25th and 75th percentile for test scores, a dummy for missing test scores, a dummy for using the ACT to report test scores, the admissions rate, and a dummy for missing the admissions rate.
* $p < .1$,** $p < .05$, *** $p < .01$

Table A7: Effect on Earnings Rank with Two-way Clustering

Table A8: Dates of Access to Facebook for 100 Universities in Sample from Jacobs et al. (2015), 1-50

University	Date	Ranking	Uncertainty
Harvard University	2/4/2004	1	0
Columbia University in the City of New York	2/25/2004	2	0
Stanford University	2/26/2004	3	0
Yale University	2/29/2004	4	0
Cornell University	3/7/2004	5	0
Dartmouth College	3/7/2004	6	0
University of Pennsylvania	3/14/2004	7	0
Massachusetts Institute of Technology	3/14/2004	8	0
New York University	3/21/2004	9	0
Boston University	3/21/2004	10	0
Brown University	4/4/2004	11	0
Princeton University	4/4/2004	12	0
University of California-Berkeley	4/4/2004	13	0
Duke University	4/11/2004	14	0
Georgetown University	4/11/2004	15	0
University of Virginia-Main Campus	4/11/2004	16	0
Boston College	4/19/2004	17	0
Tufts University	4/19/2004	18	0
Northeastern University	4/19/2004	19	0
University of Illinois at Urbana-Champaign	4/19/2004	20	0
University of Florida	4/25/2004	21	6
Wellesley College	4/25/2004	22	6
University of Michigan-Ann Arbor	4/25/2004	23	6
Michigan State University	4/25/2004	24	6
Northwestern University	4/25/2004	25	0
University of California-Los Angeles	4/27/2004	26	3
Emory University	4/30/2004	27	3
University of North Carolina at Chapel Hill	4/30/2004	28	3
Tulane University of Louisiana	4/30/2004	29	3
University of Chicago	4/30/2004	30	0
Rice University	4/30/2004	31	0
Washington University in St Louis	5/2/2004	32	0
University of California-Davis	5/20/2004	33	5
University of California-San Diego	5/20/2004	34	5
University of Southern California	6/23/2004	35	0
California Institute of Technology	6/25/2004	36	0
University of California-Santa Barbara	6/25/2004	37	0
University of Rochester	8/4/2004	38	9
Bucknell University	8/4/2004	39	9
Williams College	8/8/2004	40	4
Amherst College	8/8/2004	41	4
Swarthmore College	8/8/2004	42	4
Wesleyan University	8/8/2004	43	4
Oberlin College	8/8/2004	44	4
Middlebury College	8/8/2004	45	4
Hamilton College	8/8/2004	46	4
Bowdoin College	8/8/2004	47	4
Vanderbilt University	8/21/2004	48	1
Carnegie Mellon University	8/21/2004	49	1
University of Georgia	8/21/2004	50	1

Table A9: Dates of Access to Facebook for 100 Universities in Sample from Jacobs et al. (2015), 51-100

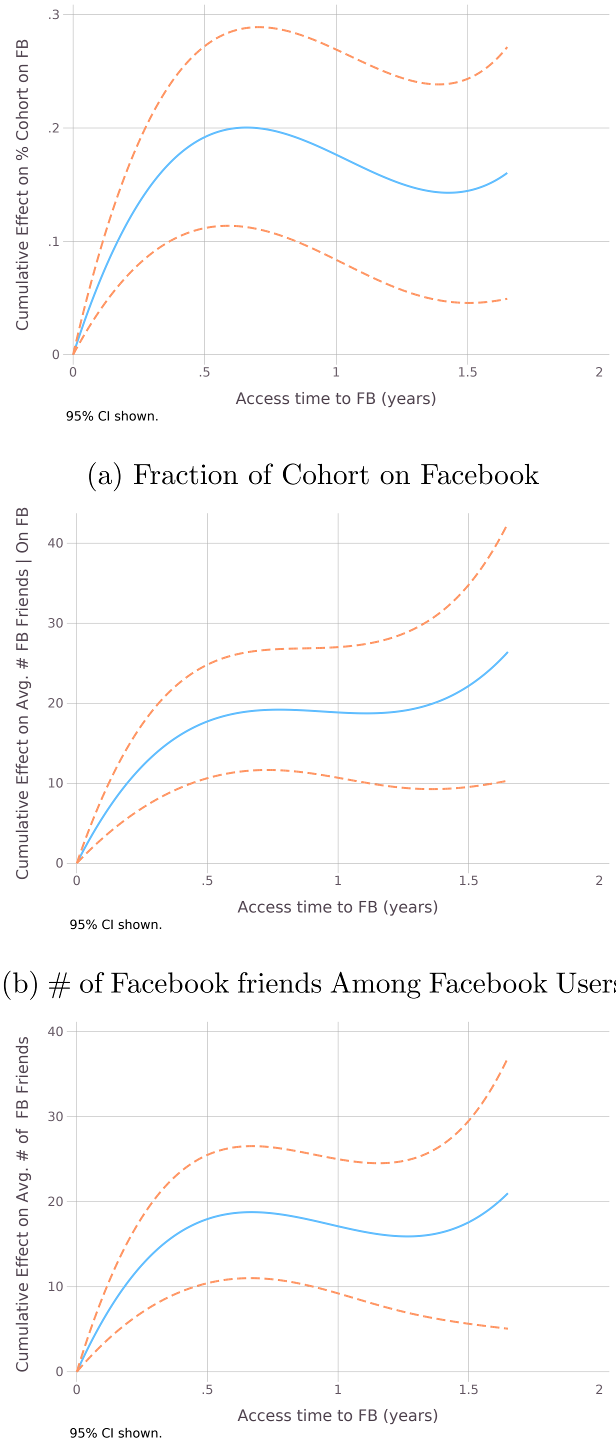
University	Date	Ranking	Uncertainty
University of South Florida-Main Campus	8/21/2004	51	1
University of Central Florida	8/21/2004	52	1
Florida State University	8/21/2004	53	1
George Washington University	8/21/2004	54	0
Johns Hopkins University	8/21/2004	55	0
Syracuse University	8/22/2004	56	1
University of Notre Dame	8/22/2004	57	1
University of Maryland-College Park	8/22/2004	58	1
University of Maine	9/7/2004	59	3
Smith College	9/7/2004	60	3
University of California-Irvine	9/7/2004	61	3
Villanova University	9/7/2004	62	3
Virginia Polytechnic Institute and State University	9/7/2004	63	3
University of California-Riverside	9/7/2004	64	3
California Polytechnic State University-San Luis Obispo	9/7/2004	65	3
University of Mississippi	9/7/2004	66	3
Michigan Technological University	9/7/2004	67	3
University of California-Santa Cruz	9/7/2004	68	3
Indiana University-Bloomington	9/7/2004	69	3
University of Vermont	9/7/2004	70	3
Auburn University	9/7/2004	71	3
University of San Francisco	9/7/2004	72	3
Wake Forest University	9/7/2004	73	3
Santa Clara University	9/7/2004	74	3
American University	9/7/2004	75	3
Haverford College	9/7/2004	76	3
College of William and Mary	9/7/2004	77	3
Miami University-Oxford	9/7/2004	78	3
James Madison University	9/7/2004	79	3
The University of Texas at Austin	9/7/2004	80	3
Simmons College	9/7/2004	81	3
SUNY at Binghamton	9/7/2004	82	3
Temple University	9/7/2004	83	3
Texas A & M University-College Station	9/7/2004	84	3
Vassar College	9/7/2004	85	3
Pepperdine University	9/7/2004	86	3
University of Wisconsin-Madison	9/7/2004	87	3
Colgate University	9/7/2004	88	3
Rutgers University-New Brunswick	9/7/2004	89	3
Howard University	9/7/2004	90	3
University of Connecticut	9/7/2004	91	3
University of Massachusetts-Amherst	9/7/2004	92	3
Baylor University	9/7/2004	93	3
Pennsylvania State University-Main Campus	9/7/2004	94	3
The University of Tennessee-Knoxville	9/7/2004	95	3
Lehigh University	9/7/2004	96	3
University of Oklahoma-Norman Campus	9/7/2004	97	3
Reed College	9/7/2004	98	3
Brandeis University	9/7/2004	99	3
Trinity College	9/24/2004	100	0

	Earnings Rank			Employment Rate		
	(1)	(2)	(3)	(4)	(5)	(6)
% Cohort on FB	3.921			4.055*		
	(3.666)			(2.135)		
Cond. # FB Friends		0.069			0.063*	
		(0.070)			(0.037)	
Uncond. # FB Friends			0.059			0.059**
			(0.057)			(0.029)
Observations	784	783	783	784	783	783
First Stage F-Stat	8.34	8.6	7.67	8.34	8.6	7.67
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Effect of σ Increase	1.1	2.41	1.72	1.14	2.2	1.71

2SLS estimates shown. Standard errors in parentheses clustered at university level . Dependent variable in percentage points. First stage regresses a Facebook social network measure in 2005 on a cubic polynomial of exposure to Facebook as of September 2005. Controls include university and year$\times$tier fixed effects as well as demographic controls.

* $p < .1$,** $p < .05$, *** $p < .01$

Table A10: IV Estimates of Effect of 2005 Facebook Network Strength on Future Earnings

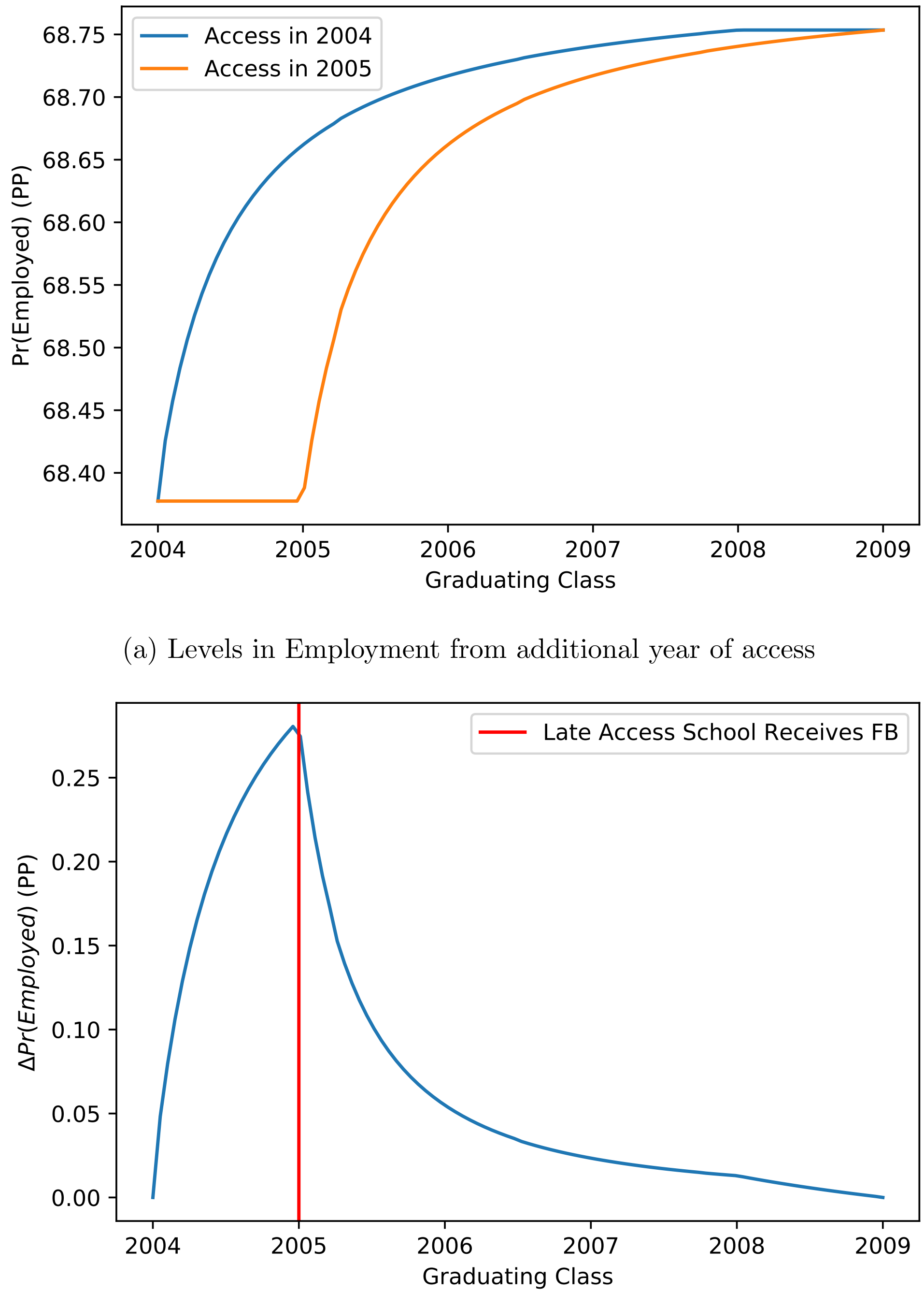


(a) Fraction of Cohort on Facebook

(b) # of Facebook friends Among Facebook Users

(c) Unconditional # of Facebook friends

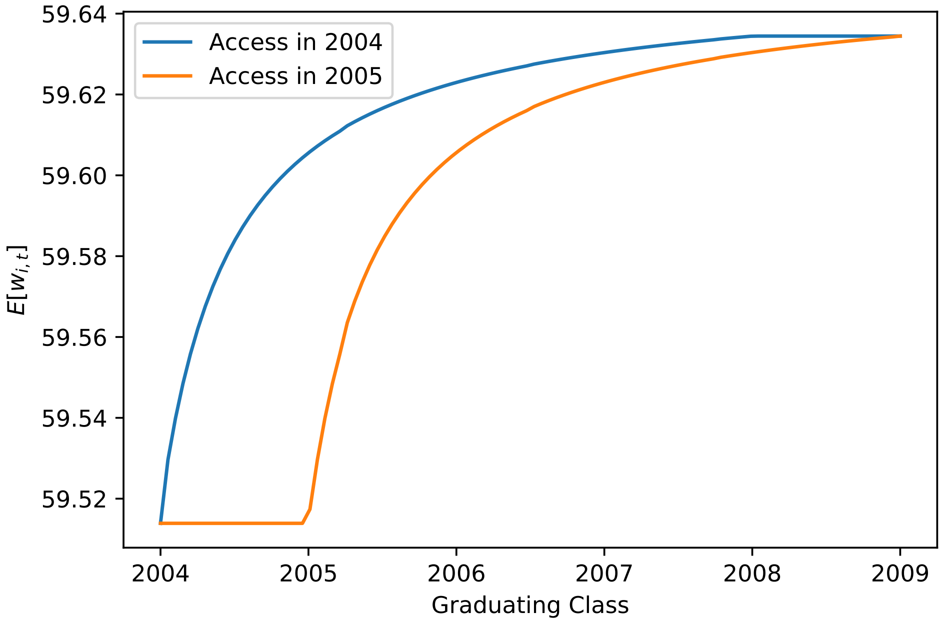
Figure A4: First-Stage Cubic Relationship between Access Time and Facebook Network



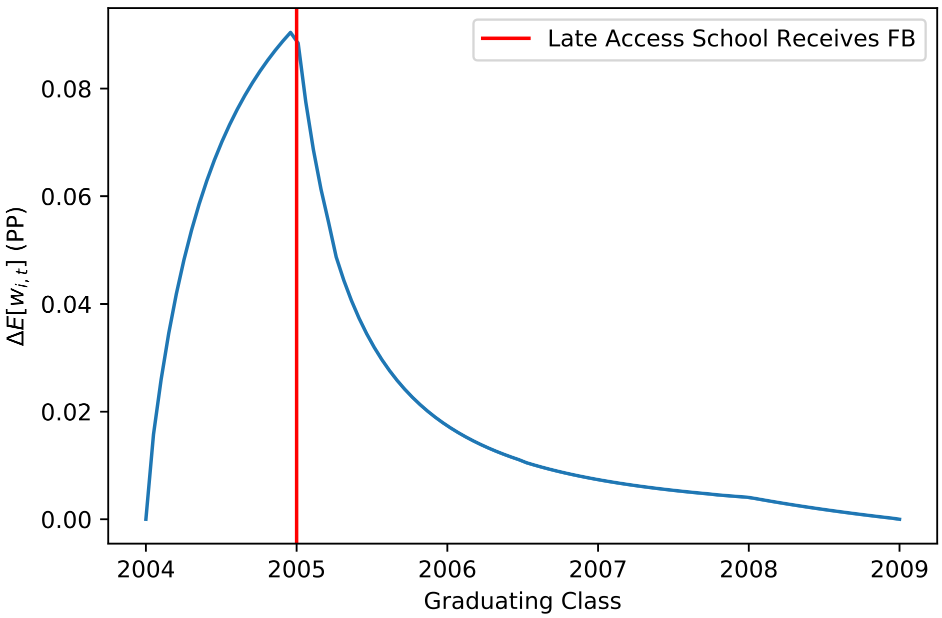
(a) Levels in Employment from additional year of access

(b) Difference in Employment Rate from additional year of access

Figure A5: Simulated effect on employment rate from additional year of reduced friendship maintenance cost



(a) Levels of Wages from additional year of access



(b) Difference in Wages from additional year of access

Figure A6: Simulated effect on wages from additional year of reduced friendship maintenance cost