“Measuring the Value of Connections Across Platforms”

Sinan Aral, Seth Benzell, Avi Collis, Christos Nicolaides

Abstract

We measure heterogeneity in the value of social media connections by user and connection characteristics across platforms through detailed surveys of US users of Instagram, Twitter, Facebook, and LinkedIn. Our study’s main outcomes are users' rankings of their contacts by importance to platform value and users' overall value from platform usage. Across types of connections, our strongest and most consistent result is that connections between individuals that are “closer” tend to create more platform value. Connections that survey takers only know online are valued less across all platforms, but especially so on Facebook. The connection value premium for intimate connections was strongest for Facebook and Instagram, and less strong for Twitter and LinkedIn, consistent with the latter platforms being more public facing. The low observed value for connections that are only known online, especially relative to college friends, is consistent with the hypothesis that Facebook surpassed Myspace by cultivating high-value connections.

WORK IN PROGRESS PLEASE DO NOT CITE OR DISTRIBUTE

1 - Introduction

A 2021 Pew Poll found that 69% of American adults used Facebook, 40% used Instagram, 28% used LinkedIn, and 23% used Twitter. For each of these platforms, at least 30% of users visited multiple times per day (Pew Research Center, 2021).[1] Willingness-to-Accept experiments and surveys have found that users place great monetary value on these services, with median values of $48 a month (Brynjolfsson et al, 2019) or more. While social media platforms have drawn condemnation for anti-competitive and anti-social practices, stated and revealed preferences agree -- individual Americans' demand for social media is high and macroeconomically important.[2]

In this paper, we seek to better understand how and why people get value from social media, with an emphasis on measuring heterogeneity in the value from connections to people of different types. We measure the relative strength of network effects to and from users of different types on four of the most common social media platforms.

To do so, we conducted detailed, focus-group style surveys of US users of Instagram, Twitter, Facebook, and LinkedIn. Our main outcomes of interest are user's overall value from platform usage and focal users' relative rankings of connections by importance. We document how these valuations vary with the characteristics of respondents, the characteristics of their connections, the social media platform, and the interaction of these features.

Understanding how and which connections create value across social media platforms is important for several reasons. Most dramatically, this heterogeneity may determine why some platforms succeed while others are flops. Consider the following proposition: If a new platform launches with a tight network of individuals who value following each other highly, it will be more likely to develop a devoted user base and attract further users. This effect may explain how Facebook was able to overcome Myspace's head start as social media king. Myspace had 88 million more users than Facebook in September 2006, but MySpace was built around interest groups, primarily related to music. On the other hand, Facebook was rolled out and built around college relationships, which are presumably more intimate. It is theorized this might have contributed to their divergent outcomes (Aral 2020).

More broadly, understanding heterogeneity in network effects and platform value is critical for platform strategy. It is a key result in the two-sided platform literature that sides of the platform that provide more positive marginal network effects should face lower marginal costs of platform use, or even subsidy, and vice versa (Jullien et al 2021). Sides of the network which have more elastic demand (i.e. there is a larger mass of users who are close to the border of use vs. non-use) should also face lower monetization at the margin. Further, because of network effects'

[1] Excepting LinkedIn, for which this question was not studied.
[2] E.g. Imputing Facebook's consumer surplus for users to be an average of $76 a month, the Windserized average in our sample, places the total American adult consumer surplus in 2021 from Facebook at over $200 Billion.

importance to both firm strategy and consumer value, better measuring its heterogeneity will also help improve platform regulation and antitrust policy.

Across platforms, Facebook had the highest average total valuation, while LinkedIn had the lowest. However, all platforms had a similar share of users with total valuations of less than $5 or $10 a month, suggesting they all have similar marginal elasticities of demand.

We find significant associations between connection counts, connection characteristics and platform value. For Facebook we find significant relationships between total platform value and friend count respectively -- another Facebook friend is related to .5 cents more value from Facebook per month.

Across types of connections, our strongest and most consistent result is that connections between individuals that are "closer" tend to create more platform value. Across platforms, users had a statistically significant preference for connections to individuals who they met often or lived with (vs. those seen less often) and with individuals with whom they had a more intimate relationship (e.g. family members and friends were preferred to college friends, and both were preferred to connections only known online). The connection value premium for intimate connections was strongest for Facebook and Instagram, and less strong for Twitter and LinkedIn, consistent with the former platforms being more private. The low observed value for connections that are only known online, especially relative to college friends, is consistent with the hypothesis in Aral (2020) that Facebook surpassed Myspace by cultivating high-value connections.

Finally, in addition to documenting this and other interesting dimensions of heterogeneity in connection value within and across platforms, we hope this paper is valuable for introducing a new methodology for estimating network effects. This approach is replicable, and could be paired with other measures to further understand the graph structure of connection value and demand dependence. We provide some closing thoughts on the strengths and weaknesses of our approach, and how future research could improve upon it, in the discussion.

2 - Background

The concept of equilibrium in a communication network with network effects was first proposed by Jeffrey Rohlfs in 1974 (Rohlfs 1974). In the context of multi-sided platforms, this theory was formalized and popularized in seminal papers by Rochet and Tirole (2003) and Parker and Van Alstyne (2005). Since then, there has been an explosion of empirical work on measuring network effects on platforms using natural and field experiments. For example, Tucker (2008) uses an exogenous shock to video messaging platform adoption in an organization to quantify network effects. Boudreau (2021) conducts a field experiment on a new platform where he exogenously varies the size of future expected installed user base that is displayed to a potential adopter, finding people prefer to use platforms that are anticipated to be popular.

Some of this literature has looked at quantifying the monetary value of network effects and connections and exploring heterogeneity across different user groups. Powdthavee (2008) looks at panel survey data on people's real-world interactions to quantify the value of in-person meetings with friends, relatives and neighbors, based on estimated opportunity costs.

Previous work has also explored the importance of network connections in real life on various non-monetary outcomes. Outcomes include grades in the context of a university (Jackson et al. 2022) and conflict in the context of European monarchs (Benzell and Cooke 2021). Rajkumar et al. (2022) use a large-scale randomized experiment on LinkedIn to explore the importance of weak ties on job mobility.

Our study contributes to this literature by empirically measuring the monetary value of network effects across major social media platforms and digging deeper into heterogeneity in valuations across different user groups. While in some cases it may be straightforward to estimate a network effect, especially when collaborating with a platform, often platforms are uninterested in external studies of their most fraught aspects. We provide a method to perform these measurements through online surveys without directly collaborating with the platforms, hence offering a scalable tool for market researchers, policymakers and regulators.

Why is understanding the nature of these network effects so important? Previous case studies (Halaburda and Oberholzer-Gee 2014, Aral 2020) and analytical models (Sundararajan 2008) show that platforms fail to take off when they do not take into consideration heterogeneity in network effects across different user groups. Mathematically, the reason for this is that participation by users on a social media platform can be modeled as a recursive function, where the quality of a platform, and therefore a user's desire to participate on it, is a function of the participation by other users of different types.

Let P be a vector of participation rates for different sides of a social media platform, and let ϕ be a vector of monetization levels (e.g. fees, advertising levels, subsidies etc) for each group. Suppose participation rates are a function of lagged participation rates and this level of monetization. Then we can write:

$$P_{t+1} = D(P_t, \phi)$$

Let's define the marginal elasticity of each group's participation to the participation of others as:

$$\frac{\partial P_{t+1}}{\partial P_t} = \mathbf{B}$$

This equation just says that the change in each group's participation as a function of each other group's participation can be written as a matrix $\mathbf{B}$. Following Benzell and Collis (2022) each element of $\mathbf{B}$, that is the elasticity of participation of each group i to each other group j, will be the product of two terms:

$$\underbrace{B_{i,j} = \frac{\partial D_i}{\partial P_{j,t}}}_{\text{Elasticity of Group i's participation to Group j}} = \underbrace{\frac{\partial \mu_i}{\partial P_j}}_{\text{Network Effect of } j \text{ on } i} \times \underbrace{\frac{\partial P_i}{\partial \mu_i}}_{\text{Demand Elasticity of Group } i}$$

The first is the partial derivative of platform value μ for group *i*, with respect to a person *j*'s participation. The second is the elasticity of participation to platform quality, or demand elasticity.

Platform participation rates are fully determined, in this model, by the directed graph of network effects and the vector of elasticity of demand. By innovating in the measurement of these parameters, we can enhance all stakeholders' models of platforms.

3 - Study Details

We conduct online surveys on US users of four major social media platforms: Facebook, Twitter, Instagram and LinkedIn. Subjects are recruited through Lucid, a market research firm. Lucid provides high-quality responses for online surveys (Coppock and McClellan 2019) and is widely used in academic studies (e.g. Pennycook et. al. 2021, Benzell et al. 2020 and Solís Arce et al. 2021). Surveys were conducted between June 2020 and April 2022. Our study was approved by Massachusetts Institute of Technology's (MIT's) Institutional Review Board.

We first recruit users of one of the four social media platforms from Lucid to take part in our study. They have been verified by Lucid to be monthly active users of the platform under consideration. We refer to the subjects of these surveys as "ego"s, and connections that they are asked to characterize and evaluate as "alter"s. Each ego is only surveyed about their use of a single platform.

Egos are first asked about their demographics (gender, age, ethnicity, income and political ideology) and usage of social media platforms. Egos are then asked to identify eight individuals (alters) -- their four closest friends (who they may or may not be connected to on the platform) and four additional alters they are connected to on the platform. We then ask the egos to characterize their relationship with these alters.

In the next section of the survey, we solicit egos' valuation of a social media platform by asking them for their minimum willingness to accept (WTA) to give up access to the platform for one month (following Brynjolfsson et al. 2019 and Allcott et al. 2020). Finally, we have egos rank the eight alters solicited earlier by their value as a connection on the platform. Survey instruments are reported in Appendix C. We obtained responses from 4,149 users of Facebook, 2,899 users of Twitter, 1,981 users of Instagram and 1,499 users of LinkedIn.

A key aspect of our survey is that egos are required to follow/ friend our research account on the platform to qualify for the study. Those who choose not to are screened out. This step is necessary to ensure that surveys are accurate, and verify that egos actually use the platform. When egos follow/friend our research account, we can look at the list of their alters' as well as their public profile information, subject to the different constraints of each platform. We manually check every response to the best of our ability to ensure survey accuracy. For each alter named in the survey to whom the ego is connected on the platform, we check whether that alter appears in the list of connections of the ego.

In addition, connecting with our ego on social media platforms allows us to collect additional information about their profiles (Table 1). To the extent possible for different platforms, we take advantage of this possibility.

Platform	Additional Information Collected
Facebook	Ego's public friend list
LinkedIn	Ego's profile (e.g. career history)
Twitter	Ego's tweet history, list of followers and followees
Instagram	N/A

Table 1. Additional information collected by survey platform

For LinkedIn, we download egos' profile which contains their education and career history. For Facebook, we download the entire Friends list of our egos and match these names with census data from the US Social Security Administration to estimate the race and gender of alters.[3] For Twitter, we use the Twitter API to download rich information about egos' tweeting activity, as well as the number of followers and followees. Instagram was the only platform we were unable to gather complementary information from due to restrictions imposed by the platform.

We analyze three different subsets of the data. In the main text of this article and the additional tables in appendix A, excepting tables 5 and tables A1 through A3, we use the full universe of survey responses. Tables 5, A1, A2, and A3 draw on information gathered from the social media accounts of survey respondents which we were able to access, and regard a smaller subset of survey takers. Finally, we re-perform the analyses of the main text and report summary statistics for a subset of responses that were deemed to be of particularly high quality. Appendix B gives the details of this filtering criteria for each platform, and reports the results summary statistics and results restricting attention to these surveys. Filtering out responses that did not reach a high quality standard, leaves responses from 1,516 users of Facebook, 656 users of Twitter, 1,260 users of Instagram and 381 users of LinkedIn.

Table 2 reports summary statistics on the demographics of egos in the main analysis sample and their monthly willingness to pay by platform. Our survey respondents skew more female and older compared to a survey of social media users conducted by Pew.[4]

[3] https://www.ssa.gov/oact/babynames/limits.html
[4] https://www.pewresearch.org/internet/2021/04/07/social-media-use-in-2021/

	FACEBOOK (N = 4,149)	INSTAGRAM (N= 1,981)	LINKEDIN (N = 1,499)	TWITTER (N = 2,899)
Gender				
Female	2181 (53%)	1,424 (72%)	853 (57%)	1,236 (43%)
Male	1951 (47%)	543 (27%)	637 (42%)	1,641 (57%)
Other	17 (0.4%)	14 (0.7%)	8 (0.5%)	22 (0.8%)
Age				
18-24 years old	646 (16%)	587 (30%)	239 (16%)	520 (18%)
25-34 years old	1,396 (34%)	713 (36%)	489 (33%)	1,048 (36%)
35-44 years old	1,310 (32%)	388 (20%)	436 (29%)	1,050 (36%)
45-54 years old	434 (10%)	157 (8.0%)	169 (11%)	223 (7.7%)
55-64 years old	214 (5.2%)	72 (3.7%)	91 (6.1%)	42 (1.5%)
65 years or older	142 (3.4%)	49 (2.5%)	73 (4.9%)	12 (0.4%)
Ethnicity				
White	2,952 (71%)	1,189 (60%)	1,014 (68%)	1,873 (65%)
Black	573 (14%)	354 (18%)	227 (15%)	534 (18%)
Hispanic or Latino	387 (9.3%)	278 (14%)	150 (10%)	317 (11%)
Asian/Pacific Islander	159 (3.8%)	100 (5.0%)	66 (4.4%)	106 (3.7%)
Native American.	36 (0.9%)	19 (1.0%)	14 (0.9%)	33 (1.1%)
Other	42 (1.0%)	41 (2.1%)	27 (1.8%)	36 (1.2%)
Income				
Less than $25,000	809 (19%)	448 (23%)	226 (15%)	384 (13%)
$25,000 - $49,999	1,003 (24%)	627 (32%)	386 (26%)	560 (19%)
$50,000 - $99,999	1,047 (25%)	598 (30%)	507 (34%)	711 (25%)
$100,000 - $149,999	751 (18%)	210 (11%)	230 (15%)	742 (26%)
$150,000 or more	539 (13%)	98 (4.9%)	149 (9.9%)	502 (17%)
Political orientation				
Extremely conservative	341 (8.2%)	101 (5.1%)	70 (4.7%)	241 (8.3%)
Conservative	405 (9.8%)	169 (8.5%)	147 (9.8%)	228 (7.9%)
Slightly conservative	274 (6.6%)	165 (8.3%)	143 (9.5%)	165 (5.7%)
Moderate	1,431 (34%)	798 (40%)	566 (38%)	908 (31%)
Slightly liberal	295 (7.1%)	207 (10%)	137 (9.1%)	217 (7.5%)
Liberal	746 (18%)	312 (16%)	255 (17%)	580 (20%)
Extremely liberal	657 (16%)	229 (12%)	180 (12%)	560 (19%)
Platform Valuation				
Median	96	99	75	82
Mean	74	70.5	65.8	69.9
SD	34.1	35.6	35.7	33.6

Table 2: Summary Statistics

4 - Results

We report two sets of results: results on which connection characteristics predict the relative rank of specific friends, and results on what ego and network characteristics predict an ego's valuation for the platform.

4.1 - Relative Friend Value by Ego, Alter, and Connection Characteristics

Our first set of results are on what characteristics of alters are correlated with them being highly ranked as friends. We define this outcome, for each ego i and alter j as

$$\text{Friend Value Percentile}_{i,j} = 1 - \frac{\text{Friend Rank}_{i,j} - 1}{\text{Total Friends Ranked}_i}$$

Higher values of 'Friend Value Percentile' (FVP) indicate that an ego-alter connection type is more highly valued by egos, and vice versa. The median and modal amount of contacts ranked by each ego is 8, so typically an FVP of one indicates the contact was ranked first of eight by the ego, and each decrease in friend percentile of 12.5 percentage points indicates moving one friend down the list of eight. In Figures 1 through 3, we report sample averages and 95% confidence intervals for the average FVP by subgroup.

4.1.1 - Relative Friend Value by Ego, Alter, and Connection Characteristics: Conditional Averages

Figures 1 through 3 report average FVPs conditional on various connection, ego, and alter characteristics both overall within specific platforms.[5]

Figure 1a shows a monotonic relationship between how frequently the pair of contacts meet and the relative value of the connection. The largest increases in predicted rank are due to increases from "multiple times per week" to ''I live with this person." There is also a significant difference between "Never" meeting in person and the other responses.

[5] These conditional averages are the same as the coefficient from an OLS regression of FVP on a dummy variable encoding whether the connection belongs to the condition added to the intercept term.

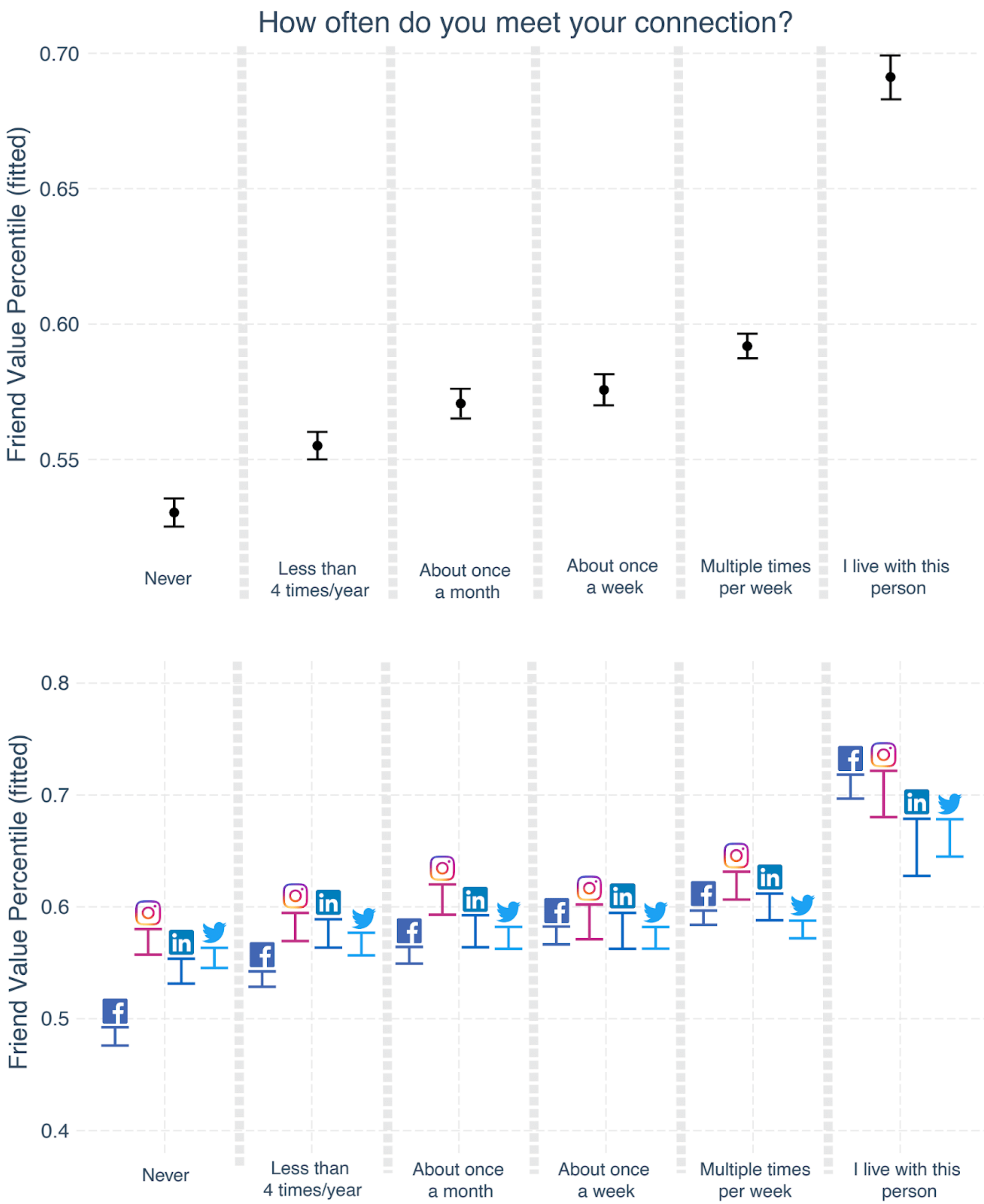
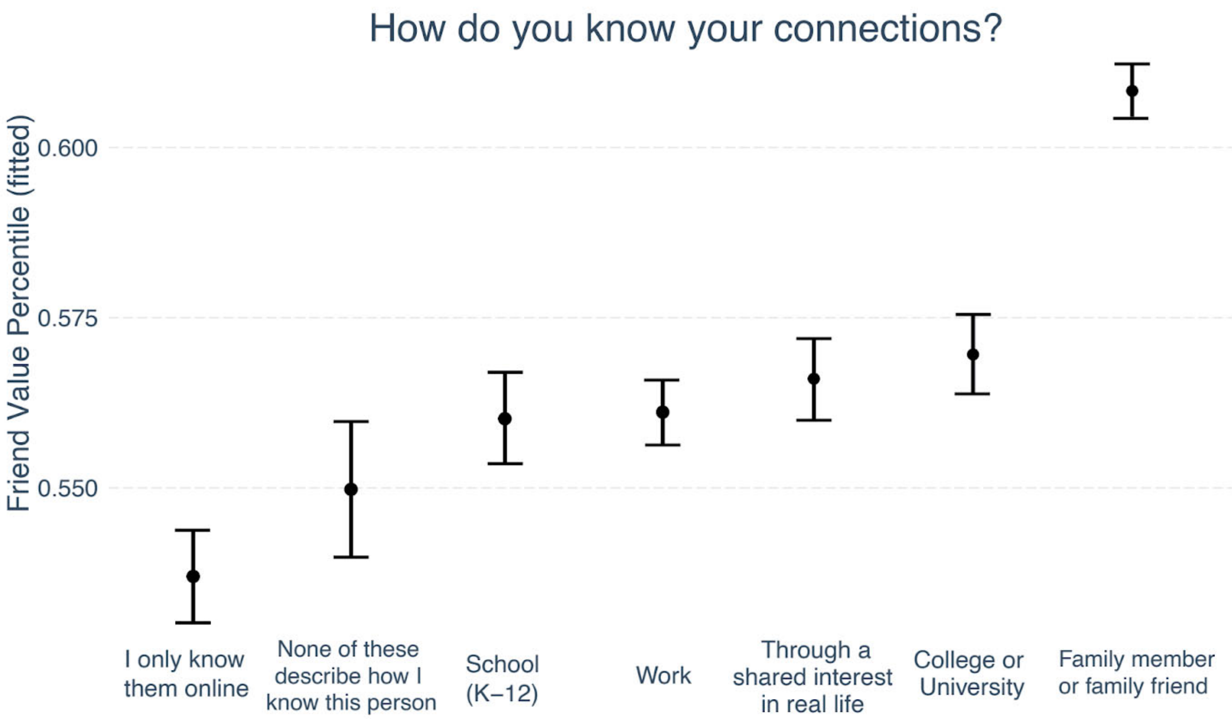


Figure 1a: Conditional sample average FVP, and 95% confidence interval, by how often the ego reports meeting their connection. Figure 1b: Conditional sample average FVP, and 95% confidence interval, by how often the ego reports meeting their connection and platform.

Figure 1b splits these results by platform. While across all platforms there is a trend for more frequent connections to be higher valued, this trend is most pronounced for Facebook, weaker for Instagram, and least pronounced for LinkedIn and Twitter. The relatively high value from individuals the ego has “Never” met for Instagram and Twitter are consistent with people valuing these platforms as places to follow celebrities and influencers. LinkedIn’s relatively low value from people whom one lives with is consistent with this being a platform where one connects with work colleagues, as co-workers may be relatively unlikely cohabitants.

It is illuminating to juxtapose these results on LinkedIn with the important recent findings of Rajkumar et al (2022). That paper found that ties on LinkedIn of medium strength (those with about 10 mutual connections) are more likely to lead to new employment opportunities than stronger or weaker ones. This suggests that LinkedIn’s relatively weaker increases in network effects as a function of meeting frequency may be in part due to the importance of weaker ties in the job search.



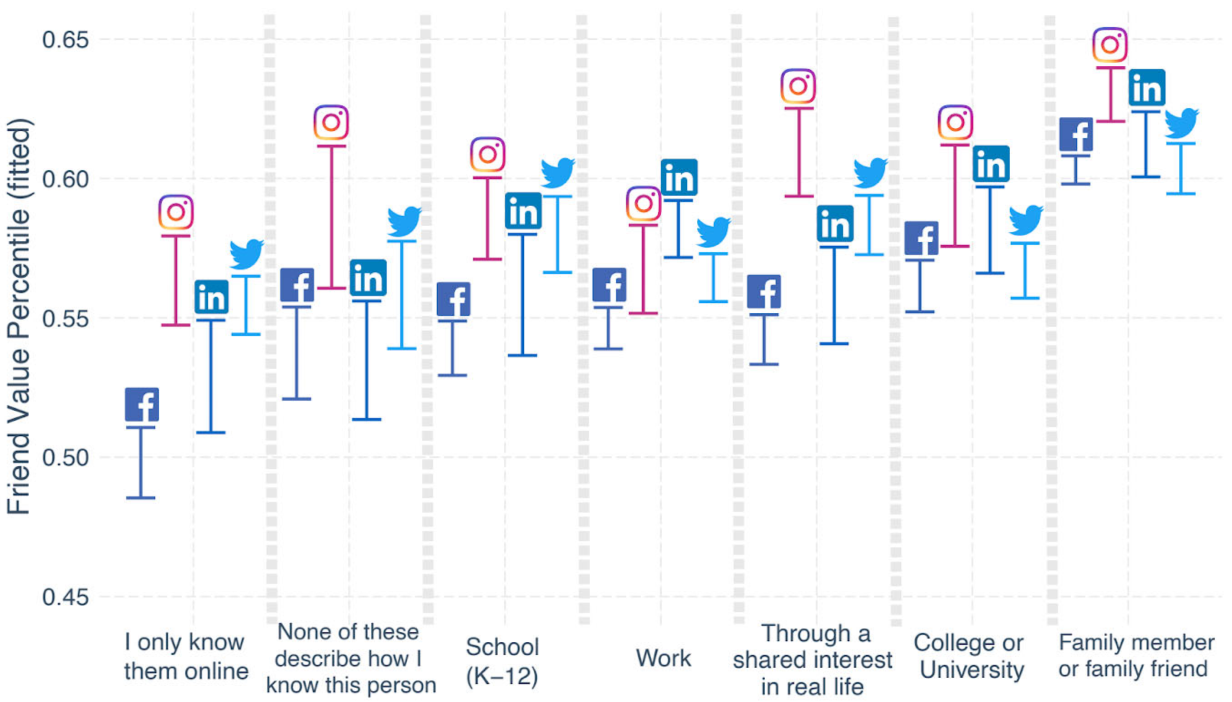


Figure 2a: Conditional sample average FVP, and 95% confidence interval, by how the ego reports knowing their connection. Figure 2b: Conditional sample average FVP, and 95% confidence interval, by how the ego reports knowing their connection and by platform.

Figure 2a is similar to Figure 1a, reporting conditional average FVPs for connections by how that connection was formed and pooling responses across all platforms. Family members and friends are valued more highly than other connections while individuals only known to the ego due to their online relationship are valued less than other connections. College friends also have a high average FVP.

Figure 2b splits the analysis of FVP by how the connection is known to the ego and platform. Similar to Figure 1b, we see the largest difference between the lowest and highest average connection type for Facebook, with less variation in FVP by connection type for the other platforms. This is because of two categories that have very low value on Facebook: “I Only Know Them Online” and “Work” friends. For both of these friend types, average FVP on Facebook is significantly lower than on other platforms.

Across Figures 1a through 2b, one theme which emerges is that closer types of friendships and connections have higher average FVP than looser connections. This is an important result. While, intuitively, communications with closer contacts are likely to be more valuable, on the other hand, close contacts and family members may have a diversity of ways of being in touch. So the higher average value of being able to connect to an individual might be outweighed by the existence of many substitute lines of communication in calculating the value of communication on any *particular* platform. To give a concrete example, it may be more important to communicate a fact to a roommate, but roommates probably have many alternatives to Facebook for communication. What these results suggest is that, at least at the margin, the friend value effect dominates the existence of substitute modes of communication effect.

Another interesting observation is the relatively high value from “work” friends for LinkedIn. LinkedIn was the only platform where work contacts had a higher sample average FVP than “School” or “College” friends. This is consistent with LinkedIn being a source of labor market value more than immediate social use.

For LinkedIn, it particularly makes sense that work contacts should be highly valued, but work friends are valued highly on Instagram and Twitter as well. Despite Facebook historically being strongly associated with school in general and college in particular (it was initially rolled out on college campuses), average FVP on Facebook for “College” friends is not significantly different than on other platforms. That said, “College” friends are the second highest point estimated source of friends on Facebook.

Figure 3a: Conditional sample average FVP, and 95% confidence interval, by ego and alter gender. Other and don't know responses not displayed. Figure 3b: Conditional sample average FVP, and 95% confidence interval, by ego and alter gender and platform. Other and don't know responses not displayed.

Figures 3a and b report average FVP by the gender of the ego and alter, pooling all platforms and across all platforms respectively. Male egos had significantly higher FVPs from female alters than male alters, while there was no significant difference in average FVPs of women for people of different genders.[6] Splitting the results by platform, we see that male preference for female alters is primarily prevalent on Facebook and Instagram. Female egos did not display a significant preference for gender on any platform.

In addition to these results featured in the main text, additional results on conditional sample average FVPs by ego and alter characteristic and platform are reported in Appendix A. We do not find a significant effect of ego and alter relative age except on Instagram, where contacts of the same age as the ego or older are preferred (appendix figure A.1.1). We find that alters that are connected to egos on at least one additional platform (i.e. are connected on at least two platforms) are significantly more valued on that platform. This effect is particularly large for Facebook connections (appendix figure A.1.2). This is consistent with our findings that “closer” connections are generally more valued social media contacts, even though those who are more closely connected have more alternate means of communication. Also consistent with our finding that “closer” connections tend to be more valued, we find that when categorizing individuals into two racial categories, white and non-white, egos give higher FVP for connections of the same race (appendix figure A.1.3).

4.1.2 - Relative Friend Value by Ego, Alter and Connection Characteristics: LASSO Results

In order to more fully explore correlations between ego and alter qualities in our sample, we performed a LASSO regression, where FVP is explained as a function of the survey platform, an ego characteristic, an alter characteristic, and an up to three-way interaction of these features. LASSO regression is a machine learning method appropriate to extracting predictive features when there are a large number of features relative to observations, and is therefore appropriate for an exploratory analysis.

Table 3 reports the results of an OLS regression of FVP on the retained predictors from the LASSO regression. Several interesting factors emerge. First, we see the result that closer connections (in terms of relationship type and frequency visited) tend to have higher FVPs is retained as a robust predictor. We further learn that the effect is particularly strong for female, slightly conservative and middle-income egos. Also, women are more valued as contacts overall, consistent with the male preference for females and female indifference to gender found above.

The only three-way interactions retained by LASSO highlight unique features of Instagram. For that platform only, LASSO retains coefficients for a female preference for young contacts and Millennials'[7] preference for contacts of the same age.

[6] We did not solicit the sexual orientation of the egos or alters, although this is an obvious mediating factor. In our survey we also allowed for egos and alters to have an “other” or unknown gender, but the sample size for these responses were too small to produce significant differences from other populations.
[7] More precisely, 24-34 year olds in 2021, which roughly corresponds to the Mil

	Friend Value Percentile
Survey Platform (Instagram)	0.018***
	(0.006)
Survey Platform (Twitter)	0.017***
	(0.003)
Alter Female	0.009***
	(0.003)
Ego knows Alter only online	−0.008*
	(0.005)
Alter a Family member or family friend	0.030***
	(0.003)
Ego meet Alter about once a week	0.018***
	(0.004)
Ego meet Alter multiple times per week	0.027***
	(0.004)
Ego lives with Alter	0.093***
	(0.007)
Survey Platform (Instagram) x Alter Female	0.003
	(0.006)
Survey Platform (LinkedIn) x Alter Female	0.016***
	(0.006)
Survey Platform (Twitter) x Alter Female	0.003
	(0.005)
Survey Platform (Instagram) x Alter 25-34 years old	0.010
	(0.009)
Survey Platform (LinkedIn) x Alter 25-34 years old	0.020***
	(0.006)
Survey Platform (LinkedIn) x Ego knows Alter through Work	0.025***
	(0.006)
Survey Platform (Instagram) x Ego knows Alter Through a shared interest in real life	0.028***
	(0.009)
Survey Platform (Twitter) x Ego knows Alter Through a shared interest in real life	0.019***
	(0.006)
Survey Platform (Instagram) x Ego meets Alter About once a month	0.025***
	(0.008)
Ego Female x Ego meets Alter Less than 4 times/year	−0.013***
	(0.004)
Ego Female x Ego meets Alter About once a week	0.007
	(0.006)
Ego Female x Ego meets Alter Multiple times per week	0.012***
	(0.005)
Ego Female x Ego lives with Alter	0.038***
	(0.008)
Ego 35-44 years old x Ego knows Alter online	−0.016**
	(0.007)
Ego 45-54 years old x Alter is Family member or family friend	0.016**
	(0.006)
Ego 45-54 years old x Ego lives with Alter	0.040**
	(0.016)
Ego Income \$50K - \$100K x Alter is Family member or family friend	0.015***
	(0.004)
Ego Income \$25K - \$50K x Ego meets Alter Multiple times per week	0.016***
	(0.006)
Ego Income \$50K - \$100K x Ego lives with Alter	0.025***
	(0.009)
Ego Slightly conservative x Ego lives with Alter	0.046***
	(0.016)
Survey Platform (Instagram) x Ego Female x Alter 18-24 years old	0.024***
	(0.008)
Survey Platform (Instagram) x Ego 25-34 years old x Alter 25-34 years old	0.015
	(0.011)
Constant	0.528***
	(0.003)
N	70,767
R^2	0.025
Adjusted R^2	0.024
Residual Std. Error	0.282
F Statistic	59.474***

*p<0.1; **p<0.05; ***p<0.01

Table 3. Results from an OLS Regression of LASSO retained predictors of FVP (selected at lambda 1SE). The original model used is Friend Value Percentile ~ Survey Platform x Ego Characteristics x Alter Characteristics. The following ego and alter characteristics were used in the original model, with the baseline-response in parentheses: Survey Platform (Facebook), Ego Gender (Male), Ego Age Bin (65 and older), Ego Income (Less than \$25,000), Ego Ethnic group (White), Ego Political

orientation (Moderate), Alter Age (I Don't know), Alter Gender (Male), Alter Ethnic group (White), How do you know Alter (None of these describe how I know this person), How frequently you meet Alter (Never).

4.2 Platform Value By Ego and Network Characteristics

Our second set of results are on the predictors of total platform value (TPV) as reported by egos in response to the question “What is the minimum amount of money (in US $) you would require to deactivate [Platform] for 1 month?” and presented with a slider of $0 to $100.[8]

The bottom section of Table 2, above, reports the mean, median, and standard deviation of responses to this question. As can be seen, Facebook has the highest average value to users, while Instagram has the highest median value.

Appendix figures A.2.1 and A.2.2 report the empirical CDF of the share of users valuing a platform by different amounts. The first figure top-codes monthly valuations at $100 a month, while the second, with a logged x-axis, does not. The steepness of these curves at different points give insight into the amount of users prepared to leave the platform after a certain degree of degradation. The four platforms all have a comparable share of users with less than $5 and $10 a month in surplus, indicating that the managers of each platform face similar average marginal elasticities of demand.

[8] Survey takers who selected the maximum value of the slider ($100 a month) were then prompted to report their full willingness to pay in a free response field.

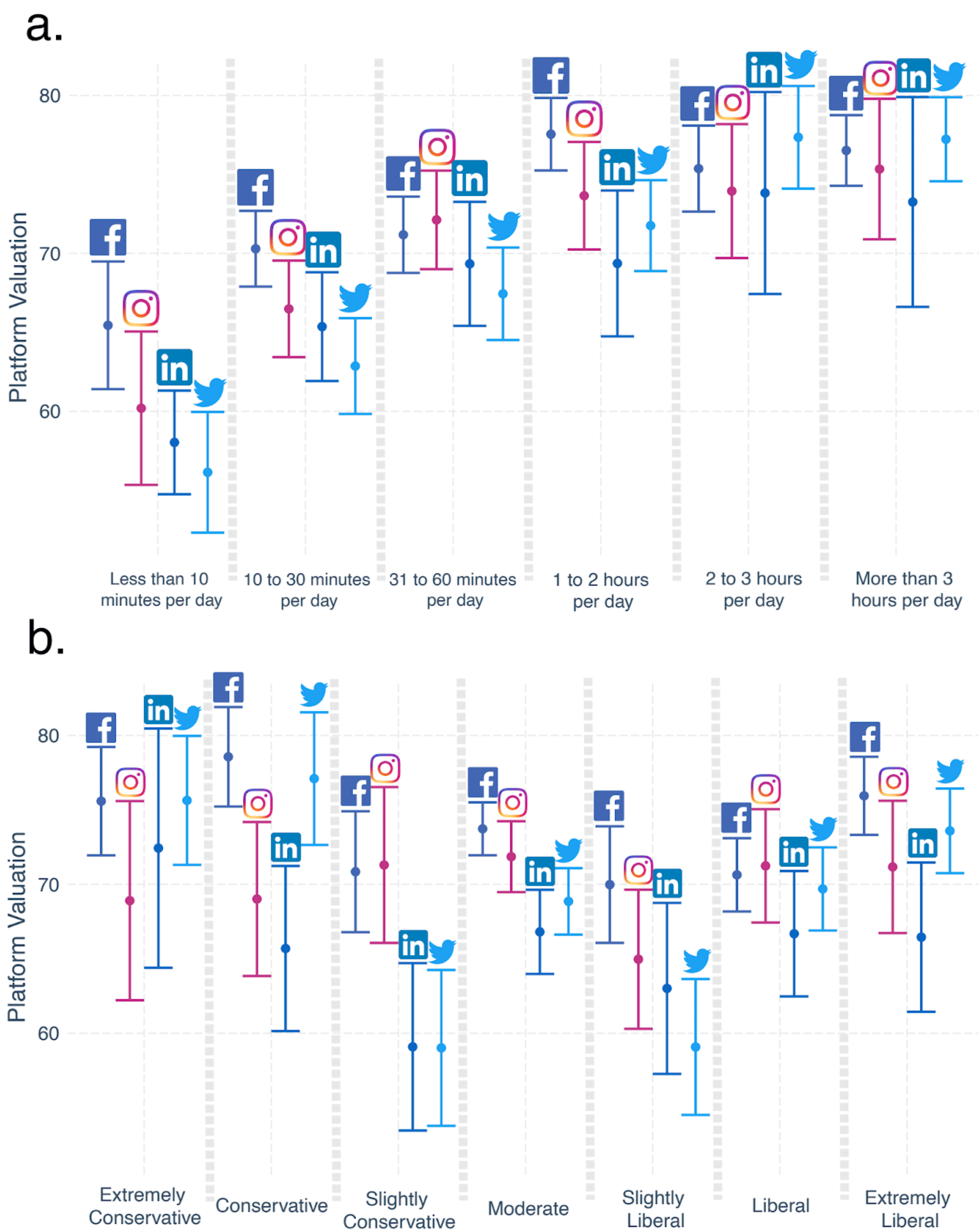


Figure 4a: *Conditional sample average TPV, and 95% confidence interval, by ego's time spent on the platform and platform.* Figure 4b: *Conditional sample average TPV, and 95% confidence interval, by ego's political orientation and platform.*

Figures 4a and 4b plot the conditional sample average monthly total platform value (TPV) by ego's time spent on the platform, ego's political orientation, and platform. Two results are apparent. First, the more time one spends on any platform, the more the platform is valued. However, this asymptotes somewhat around the 3 hours a day level, perhaps because those with that quantity of free time have lower incomes, and therefore lower Willingness-to-Accept. It also may be because of our censoring of very high platform values. The trend is least pronounced for Facebook, indicating these considerations might be more important on that platform. There is little effect of political orientation on TPV, except that there is a pronounced decrease in value to slightly liberal and conservative Twitter users, relative to more extreme users. This may be because it is those with more radical political views that prefer the rough and tumble of Twitter's political tussling. At the population level, our survey found that users of Twitter were less moderate on average as well (Table 2).

4.2.1 Ego Characteristic and Platform Value Across Platforms

Mirroring our analysis of FVPs, we now turn to LASSO to search for robust predictors of TPV. Table 4 reports the results of an OLS regression of LASSO retained predictors on TPV. The results show that our previous finding, that slightly conservative and liberal egos have less of a taste for social media in general and specifically to Twitter, is robust to LASSO. Middle aged people value social media more than very young or old individuals, and 35-44 year olds have a particular taste for Twitter. Low to middle income people value Twitter less, as do hispanics.

	Total Platform Valuation
Survey Platform (LinkedIn)	-6.331^{***}
	(1.023)
Ego 18-24 years old	-1.727^{*}
	(1.045)
Ego 35-44 years old	1.822^{**}
	(0.919)
Ego 45-54 years old	4.408^{***}
	(1.211)
Ego Slightly conservative	-2.430
	(1.691)
Ego Slightly liberal	-4.685^{***}
	(1.416)
Survey Platform (Twitter) x Ego 18-24 years old	-1.402
	(1.924)
Survey Platform (Twitter) x Ego 35-44 years old	3.565^{***}
	(1.348)
Survey Platform (Twitter) x Ego Income $25K - $50K	-5.118^{***}
	(1.617)
Survey Platform (Twitter) x Ego Income $50K - $100K	-6.043^{***}
	(1.436)
Survey Platform (LinkedIn) x Ego Slightly conservative	-5.738^{*}
	(3.450)
Survey Platform (Twitter) x Ego Slightly conservative	-9.181^{***}
	(3.193)
Survey Platform (Twitter) x Ego Slightly liberal	-5.383^{*}
	(2.796)
Survey Platform (Twitter) x Ego Hispanic or Latino	-4.515^{**}
	(2.047)
Constant	72.596^{***}
	(0.584)
N	10,527
R^2	0.019
Adjusted R^2	0.018
Residual Std. Error	34.217
F Statistic	14.467^{***}

8

$^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

Table 4. OLS regression of monthly total platform valuation on LASSO retained predictors (at lambda 1SE). The original model used is Total Platform Valuation ~ Survey Platform x Ego Characteristics.The following ego characteristics were used in the original model, as well as two-way interactions of these, with the baseline-response in parentheses: Survey Platform (Facebook), Ego Gender (Male), Ego Age Bin (65 and older), Ego Income (Less than $25,000), Ego Ethnic group (White), Ego Political orientation (Moderate).

4.2.2 Platform-Specific Ego and Network Characteristics and Platform Value

Our final analyses exploit the special information we can extract on egos and their connections on each platform. These findings are summarized in Table 5. Note that this analysis is only possible for a smaller set of survey takers from whom we were able to gather data from their social media accounts.

	Platform Monthly Valuation		
	Facebook	Twitter	LinkedIn
Number of Friends	0.005**		
	(0.002)		
Count of Followers		-0.0004	
		(0.003)	
Count of Followees		0.004*	
		(0.002)	
Connections Count			0.001
			(0.001)
N	829	692	344
R^2	0.005	0.008	0.006
Adjusted R^2	0.006	0.003	0.003
Residual Std. Error	30.606	34.154	35.392
F Statistic	4.502**	2.922*	1.982

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

Table 5. Results of OLS Regression of TPV on the number of connections across three platforms.

As can be seen, a larger count of connections is significantly associated with ego's platform value for Facebook, but not for LinkedIn. For Twitter, the number of followees (i.e. the number of people the ego follows) positively and significantly predicts platform value, while the number of followers does not. On Facebook, an additional friend is correlated with another .5 cents per month of TPV at the margin, at the 5% level, while an additional followee on Twitter is associated, at the 10% level, with another .4 cents. The result on Twitter suggests that for the typical ego surveyed, hearing what others are saying on Twitter is valuable, but being heard is not.[9] On LinkedIn there is no significant effect of the number of contacts on platform value,

[9] This stands in contrast with the findings of Filipas and Horton (2021) who find that "econtwitter" users getting value from their posts being viewed by peers is important for explaining observed patterns in their networks. However, the discrepancy may be due to the relevant populations --

suggesting that non-friend sources of platform value may be particularly important on that social network.

Appendix tables A1 through A3 dive deeper into the platform-specific data's correlation with TPV. For Facebook, we are able to estimate the number of egos' friends by ethnicity and sex based on the last and first names respectively. While the point estimates suggest that female friends are worth slightly more than male friends on average, this difference is not significant. We do find significant differences in the value of friends by race. Platform users with lots of contacts with native and bi-racial names tend to value the platform more, and users with friends with Asian names value the platform less. We speculate this may be because of the prevalence of WeChat as a common alternative to Facebook for Asian Americans.

For LinkedIn, we also have data on the egos' work history. Because LinkedIn is often harnessed as a job search network, it makes sense that those with secure employment might value the platform less. Indeed, we find that individuals who are employed and worked at that position for more than one-year value access to LinkedIn by about $9 less a month.

For Twitter, we have data on the total engagement the accounts have. We find at the 1% level, a significant positive relationship between the number of retweets an ego receives and their platform value, and a negative relationship between the number of replies an ego receives and their platform value (significant at the 10% level). Retweets are more likely to be uncritical endorsements, while replies are likely to be critiques of data, so this is consistent with people enjoying positive attention while disliking negative attention.[10]

4.3 Robustness to Survey Quality Filtering

One possible concern is that low-effort survey responses may degrade the quality of our data. To address this, appendix section B re-estimates our above analysis, but restricting attention to the highest quality surveys. We find that our qualitative findings are all unchanged by these restrictions, although confidence intervals are larger due to the decreased amount of data.

5.0 Discussion

Digital platforms, and social media platforms in particular, have been some of the most successful and popular businesses of the 21st Century. These businesses create much of their value from network effects -- that is, the value of the platform is dependent on how many people of different types use it. Individuals' decision to use these platforms can therefore be modeled by combining measures of their elasticity of demand with measures of the matrix of network effects that each group provides to each other group. These measures are therefore also of critical value in conducting platform strategy or regulation.

---

economics academics on Twitter may be more narcissistic and self-promoting than the average ego in our sample.

[10] Receiving more replies than retweets of a tweet is called being "ratioed", and is considered a sign of general disapproval for one's ideas (Birney 2022).

In this paper, we propose and deploy a new approach to measuring heterogeneity in connection and total platform value within and across platforms. Across types of connections, our strongest and most consistent result is that connections between individuals that are “closer” tend to create more platform value. The exceptions are rich survey-takers on Twitter and male survey-takers on LinkedIn, for whom more distant contacts are preferred. This result is consistent with the hypothesis that Myspace was eventually surpassed by Facebook because the latter focused on creating higher-value college friend connections, while the former fostered looser shared-interest connections. Still, it is important to recall that our relative friend value results are per-connection and measured at the margin. They do not rule out other friend types having more importance in aggregate or infra-marginally.

Our study has several limitations. A key limitation of our study is that our surveys are hypothetical choices without real incentives (i.e. we did not enforce someone’s choice to disconnect with a connection in exchange for actual payment). While we had to resort to a hypothetical survey design given the nature and scale of our study, future research could build upon our work by conducting similar surveys but with monetary incentives in controlled environments where respondents have to actually disconnect from certain connections.

Readers of this study should also be cautious in interpreting significance findings too strongly. While we do not formally conduct hypothesis tests for differences in means between sub-groups, the very large amount of possible comparisons possible lead to multiple hypothesis testing concerns. In part, we deal with these concerns by using LASSO regressions to identify characteristics that robustly predict total platform or relative friend value. Still, readers should focus on differences between populations with very large differences in sample means which would survive such multiple hypothesis testing adjustments.

We do not view this study as the last word on measuring the sources of platform value. We hope future research will continue to improve quantitative measurement of how individuals get value from social media, with the goal of improving both platform management and regulation.

References

Aral, S. 2020. The Hype Machine: How Social Media Disrupts Our Elections, Our Economy, and Our Health–and How We Must Adapt. Penguin Random House, New York, NY.

Allcott, H., Braghieri, L., Eichmeyer, S., & Gentzkow, M. (2020). The welfare effects of social media. *American Economic Review*, *110*(3), 629-676.

Benzell, S. G., & Collis, A. (2022). Regulating Digital Platform Monopolies: The Case of Facebook

Benzell, S. G., Collis, A., & Nicolaides, C. (2020). Rationing social contact during the COVID-19 pandemic: Transmission risk and social benefits of US locations. *Proceedings of the National Academy of Sciences*, *117*(26), 14642-14644.

Benzell, S. G., & Cooke, K. (2021). A network of thrones: Kinship and conflict in europe, 1495–1918. *American Economic Journal: Applied Economics*, *13*(3), 102-33.

Birney, A. (2023). *What does "ratioed" mean on Twitter and how can you spot it?* Android Authority. Published March 26, 2023. Retrieved May 7, 2023, from https://www.androidauthority.com/what-is-ratio-twitter-3141551/

Boudreau, K. J. (2021). Promoting Platform Takeoff and Self-Fulfilling Expectations: Field Experimental Evidence. *Management Science*, *67*(9), 5953-5967.

Brynjolfsson, E., Collis, A., & Eggers, F. (2019). Using massive online choice experiments to measure changes in well-being. *Proceedings of the National Academy of Sciences*, *116*(15), 7250-7255.

Coppock, A., & McClellan, O. A. (2019). Validating the demographic, political, psychological, and experimental results obtained from a new source of online survey respondents. *Research & Politics*, *6*(1), 2053168018822174.

Filippas A, Horton JJ (2021) The production and consumption of social media. National Bureau of Economic Research, Working Paper No. 28666.

Halaburda, H., & Oberholzer-Gee, F. (2014). The limits of scale. *Harvard business review*, *92*(4), 94-9.

Jackson, M. O., Nei, S. M., Snowberg, E., & Yariv, L. (2023). *The Dynamics of Networks and Homophily* (No. w30815). National Bureau of Economic Research.

Parker, G. G., & Van Alstyne, M. W. (2005). Two-sided network effects: A theory of information product design. *Management science*, *51*(10), 1494-1504.

Pennycook, G., Epstein, Z., Mosleh, M., Arechar, A. A., Eckles, D., & Rand, D. G. (2021). Shifting attention to accuracy can reduce misinformation online. *Nature*, *592*(7855), 590-595.

Pew Research Center. (2021). Social Media Use in 2021. Published April 7, 2021. Retrieved May 6, 2023, from https://www.pewresearch.org/internet/wp-content/uploads/sites/9/2021/04/PI_2021.04.07_Social-Media-Use_FINAL.pdf

Powdthavee, N. (2008). Putting a price tag on friends, relatives, and neighbours: Using surveys of life satisfaction to value social relationships. *The Journal of Socio-Economics*, *37*(4), 1459-1480.

Rajkumar, K., Saint-Jacques, G., Bojinov, I., Brynjolfsson, E., & Aral, S. (2022). A causal test of the strength of weak ties. *Science*, *377*(6612), 1304-1310.

Rochet, J. C., & Tirole, J. (2003). Platform competition in two-sided markets. *Journal of the european economic association*, *1*(4), 990-1029.

Rohlfs, J. (1974). A theory of interdependent demand for a communications service. *The Bell journal of economics and management science*, 16-37.

Solís Arce, J. S., Warren, S. S., Meriggi, N. F., Scacco, A., McMurry, N., Voors, M., ... & Omer, S. B. (2021). COVID-19 vaccine acceptance and hesitancy in low-and middle-income countries. *Nature medicine*, *27*(8), 1385-1394.

Sundararajan, A. (2008). Local network effects and complex network structure. *The BE Journal of Theoretical Economics*, *7*(1).

Tucker, C. (2008). Identifying formal and informal influence in technology adoption with network externalities. *Management science*, *54*(12), 2024-2038.

Jullien, B., Pavan, A., & Rysman, M. (2021). Two-sided markets, pricing, and network effects. In *Handbook of Industrial Organization* (Vol. 4, No. 1, pp. 485-592). Elsevier.

Appendix A - Additional Analyses Excluded for Conciseness

A.1 - Additional FVP Analyses

Figure A.1.1 Panel A: Conditional sample average and 95% confidence interval on FVP by relative ego and alter age. Figure A1 Panel B: Conditional sample average and 95% confidence interval on FVP by relative ego and alter age and platform.

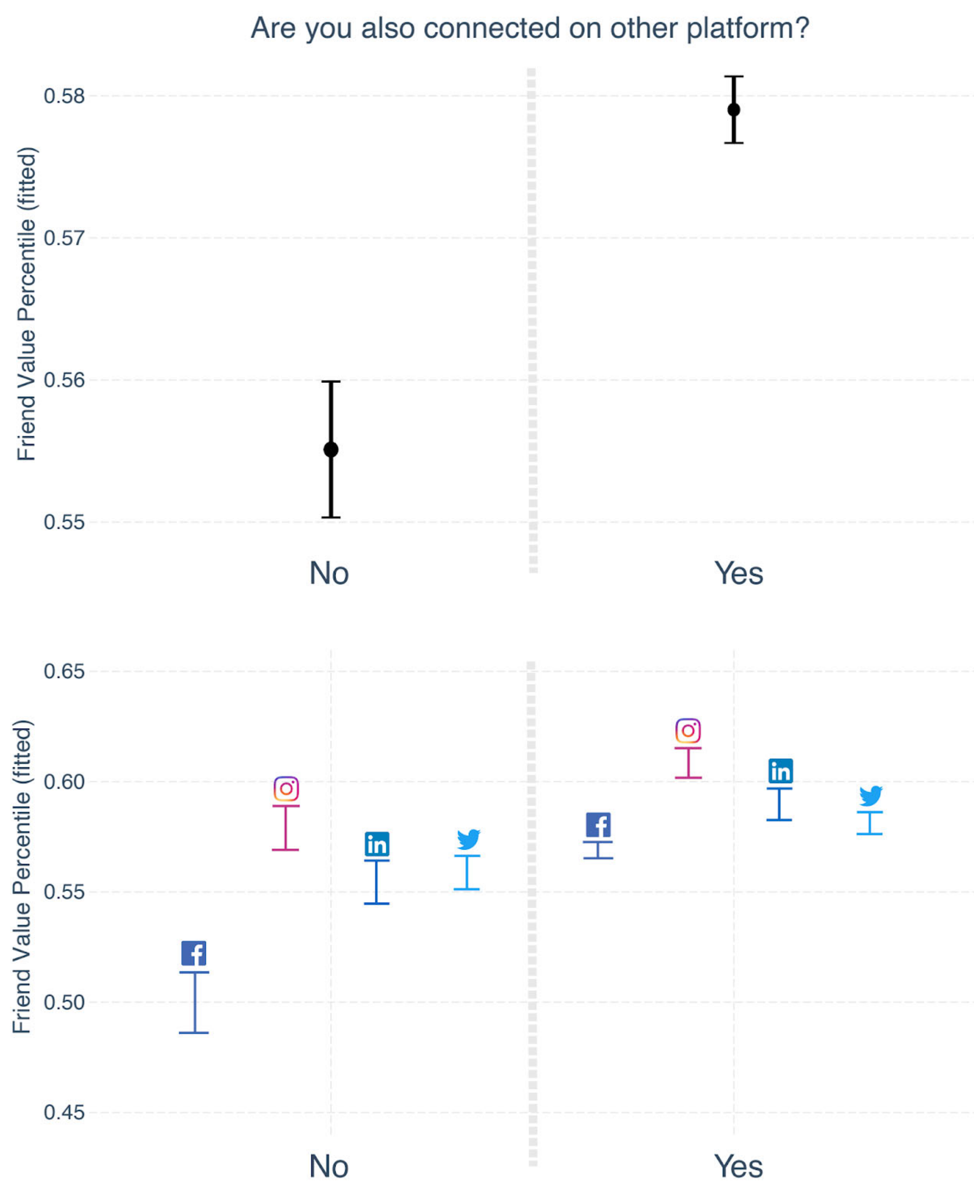


Figure A.1.2 Panel A: Conditional sample average and 95% confidence interval on FVP by whether the ego and alter are connected on at least one other platform. Figure A1 Panel B: Conditional sample average and 95% confidence interval on FVP by ego survey platform and whether the ego and alter are connected on at least one other platform.

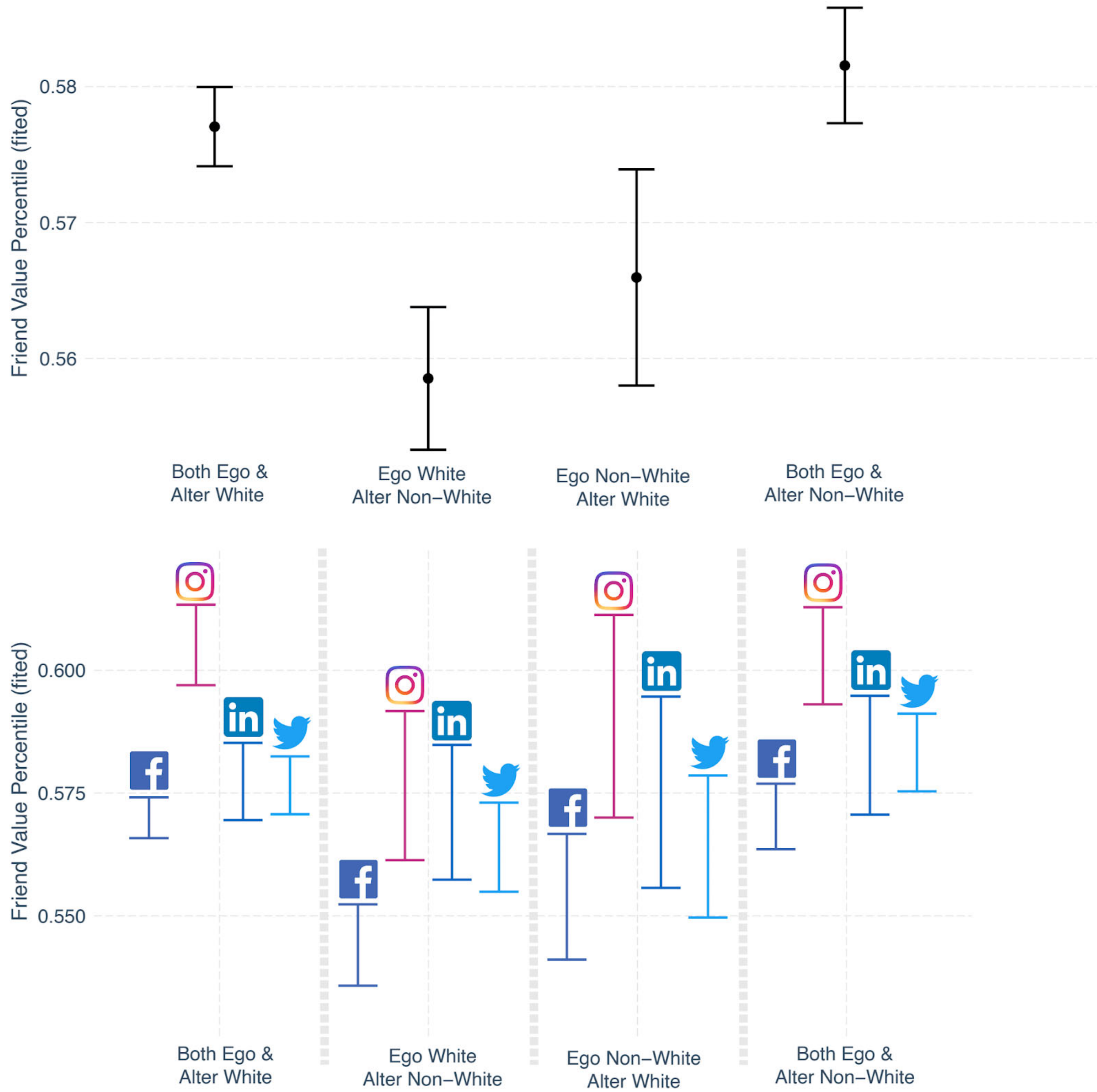


Figure A.1.3 Panel A: Conditional sample average and 95% confidence interval on FVP by whether the ego and alter are identified by the ego as white. Figure A1 Panel B: Conditional sample average and 95% confidence interval on FVP by ego survey platform and whether the ego and alter are identified by the ego as white.

A.2 Additional Total Platform Value Analysis

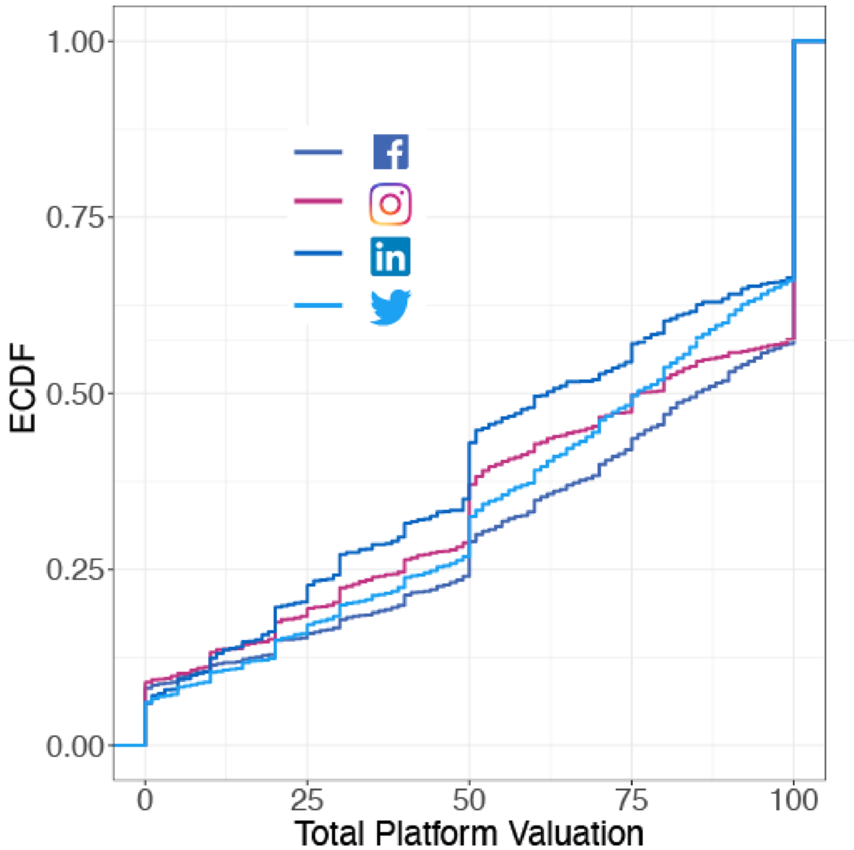


Figure A.2.1: CDF of Platform valuations for each platform.

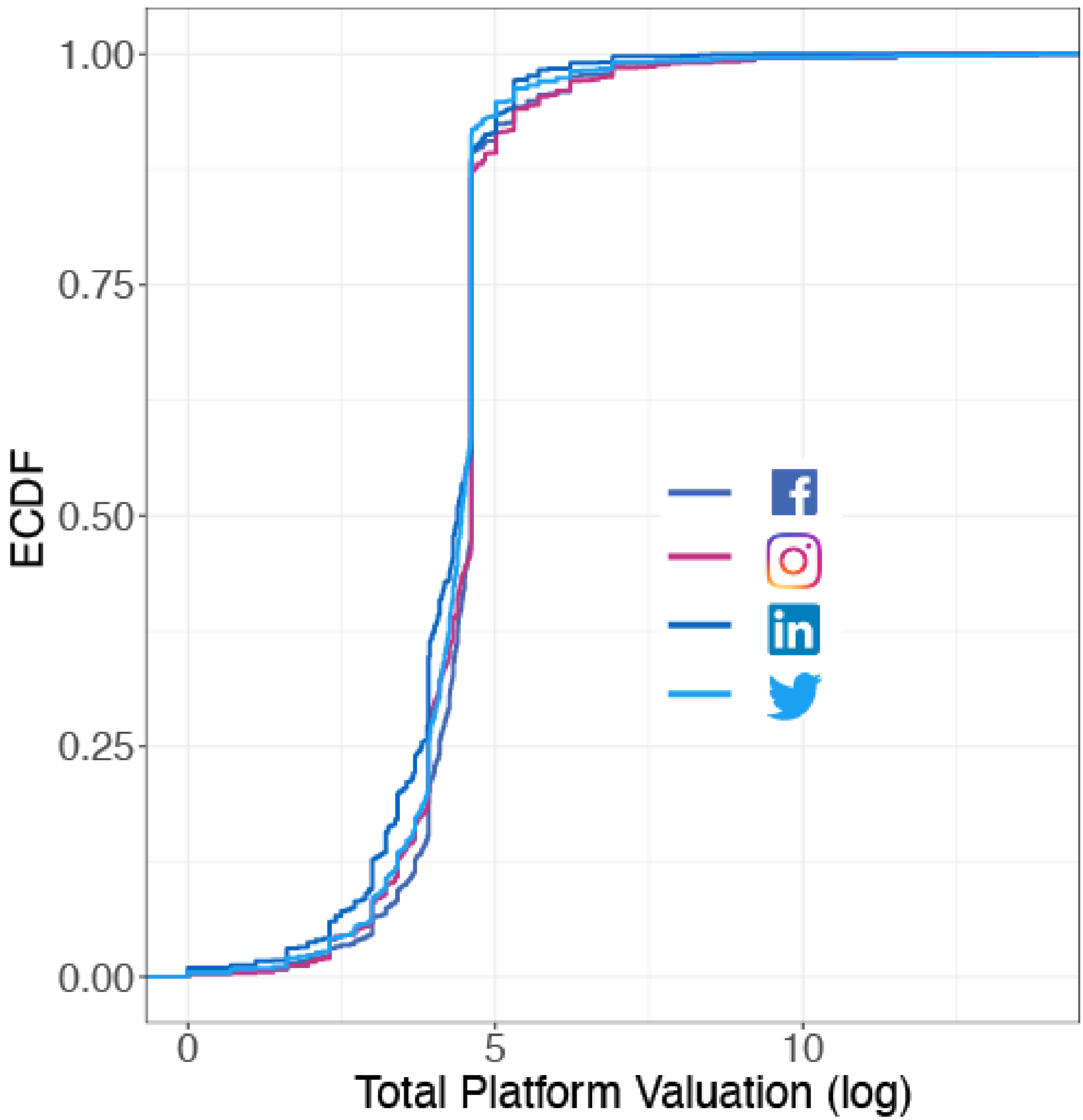


Figure A.2.2: CDF of platform valuations, with log WTA on the x axis. Free responses to platform WTA >$100 included.

	Facebook Monthly Valuation		
	(1)	(2)	(3)
Total number of friends	0.005**		
	(0.002)		
Number of Female Friends		0.007	
		(0.006)	
Number of Male Friends		0.003	
		(0.005)	
Number of Black Friends			−0.080
			(0.070)
Number of White Friends			−0.030
			(0.020)
Number of Hispanic Friends			0.005
			(0.011)
Number of Asian Friends			−0.135**
			(0.060)
Number of Nat. American Friends			0.846
			(0.870)
Number of bi-racial Friends			1.771
			(1.273)
Constant	78.874***	78.778***	79.334***
	(1.452)	(1.473)	(1.491)
N	829	829	829
R^2	0.005	0.006	0.018
Adjusted R^2	0.004	0.003	0.011
Residual Std. Error	30.606	30.622	30.505
F Statistic	4.502**	2.323*	2.506**

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

Table A1 - Facebook TPV by friend counts and their estimated race and sex

	LinkedIn Monthly Valuation		
	(1)	(2)	(3)
Connections Count	0.001		0.002
	(0.001)		(0.001)
Currently working >1 year		-8.423^{**}	-9.093^{**}
		(4.028)	(4.051)
Constant	68.497^{***}	73.038^{***}	72.671^{***}
	(1.947)	(2.600)	(2.623)
N	344	312	309
R^2	0.006	0.014	0.022
Adjusted R^2	0.003	0.011	0.016
Residual Std. Error	35.392	35.079	34.965
F Statistic	1.982	4.372^{**}	3.447^{**}

$^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$

Table A2 - LinkedIn TPV by connection count (1) and if currently ego is working for over 1 year in the current job (2,3).

	Twitter Monthly Valuation	
	(1)	(2)
count of followers	-0.0004	
	(0.003)	
count of followees	0.004^{*}	
	(0.002)	
total retweets		0.011^{***}
		(0.003)
total quotes		0.012
		(0.015)
total replies		-0.004^{*}
		(0.002)
favorites count		0.001
		(0.001)
Constant	67.847^{***}	67.788^{***}
	(1.465)	(1.726)
N	692	621
R^2	0.008	0.027
Adjusted R^2	0.006	0.021
Residual Std. Error	34.154 (df = 689)	33.957 (df = 616)
F Statistic	2.922^{*} (df = 2; 689)	4.306^{***} (df = 4; 616)

$^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$

Table A3 - Twitter TPV by engagement

Appendix B -- Robustness to Filtering Low Quality Responses

B.1 -- Filtering Criteria

To validate the robustness of the above results to low quality responses, we also conducted the main analysis restricting attention to only the highest quality responses. For each platform, our procedure was the following

Facebook: We first check whether the respondent adds our research account as a friend on their primary active Facebook profile (the account with their real name). We scan the names of their contacts and check whether these names appear realistic or not. If the respondent's Facebook friends list is shared with their friends (which was the case with almost all of our respondents), for each of their contacts whom they are connected to on Facebook we verify if these contacts appear in their friend list.

Twitter: We first check whether the respondent follows our research account on Twitter with their Twitter profile (the handle they listed in the survey). If their Twitter profile is public (which was the case with almost all of our respondents), for each of their contacts whom they are connected to on Twitter we verify if these contacts appear in their follower/ following lists.

Instagram, Linkedin: For these two platforms, we verify whether the respondent follows/ connects with our research account with their profile (the handle they listed in the survey). We also verify if their profile appears legitimate (profile picture, profile public details are not completely blank and look legitimate). Besides this, we cannot really do anything else since contacts list are not shared with connections on these platforms.[11]

B.2 -- Results with After Filtering For Low Quality Responses

[11] On Instagram, the follower and following list is shared with contacts but Instagram blocked our research account when we attempted to scroll through these lists and scrape them.

	FACEBOOK (N = 1,516)	INSTAGRAM (N= 1,260)	LINKEDIN (N = 381)	TWITTER (N = 656)
Gender				
Female	956 (63%)	995 (79%)	258 (68%)	363 (55%)
Male	551 (36%)	255 (20%)	122 (32%)	283 (43%)
Other	9 (0.6%)	10 (0.8%)	0 (0.0%)	10 (1.5%)
Age				
18-24 years old	242 (16%)	389 (31%)	53 (14%)	176 (27%)
25-34 years old	495 (33%)	446 (36%)	127 (33%)	233 (36%)
35-44 years old	374 (25%)	223 (18%)	92 (24%)	167 (25%)
45-54 years old	214 (14%)	103 (8.3%)	45 (12%)	56 (8.5%)
55-64 years old	112 (7.4%)	52 (4.2%)	35 (9.2%)	17 (2.6%)
65 years or older	78 (5.1%)	34 (2.7%)	28 (7.4%)	7 (1.1%)
Ethnicity				
White	1,115 (74%)	803 (64%)	285 (75%)	408 (62%)
Black	180 (12%)	192 (15%)	35 (9.2%)	127 (19%)
Hispanic or Latino	124 (8.2%)	164 (13%)	29 (7.6%)	73 (11%)
Asian/Pacific Islander	55 (3.6%)	58 (4.6%)	17 (4.5%)	24 (3.7%)
Native American.	18 (1.2%)	13 (1.0%)	4 (1.1%)	12 (1.8%)
Other	24 (1.6%)	30 (2.4%)	10 (2.6%)	12 (1.8%)
Income				
Less than $25,000	382 (25%)	305 (24%)	51 (13%)	132 (20%)
$25,000 - $49,999	468 (31%)	417 (33%)	91 (24%)	194 (30%)
$50,000 - $99,999	421 (28%)	369 (29%)	136 (36%)	191 (29%)
$100,000 - $149,999	151 (10.0%)	118 (9.4%)	66 (17%)	96 (15%)
$150,000 or more	94 (6.2%)	51 (4.0%)	36 (9.5%)	43 (6.6%)
Political orientation				
Extremely conservative	86 (5.7%)	47 (5.1%)	70 (4.7%)	241 (8.3%)
Conservative	160 (11%)	110 (8.7%)	147 (9.8%)	228 (7.9%)
Slightly conservative	125 (8.2%)	120 (9.5%)	143 (9.5%)	165 (5.7%)
Moderate	600 (40%)	519 (41%)	566 (38%)	908 (31%)
Slightly liberal	122 (8.0%)	138 (11%)	137 (9.1%)	217 (7.5%)
Liberal	274 (18%)	202 (16%)	255 (17%)	580 (20%)
Extremely liberal	149 (9.8%)	124 (9.8%)	180 (12%)	560 (19%)
Platform Valuation				
Median	100	100	90	80
Mean	76	72	67	66
SD	35	36	37	36

Table B.2.1 - Summary statistics with quality filtered data

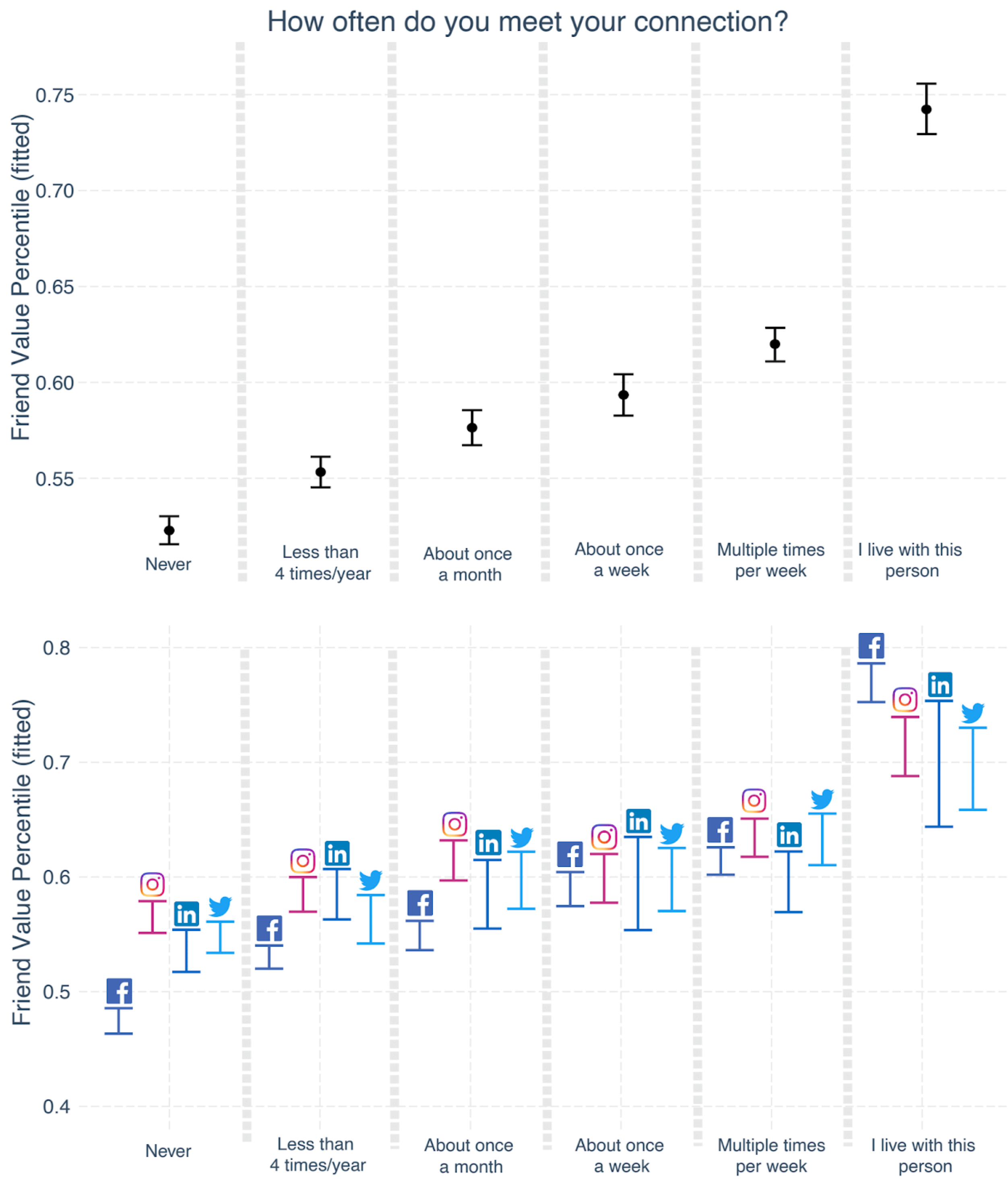


Figure B.2.1 -- Figure 1, restricting attention to quality filtered data

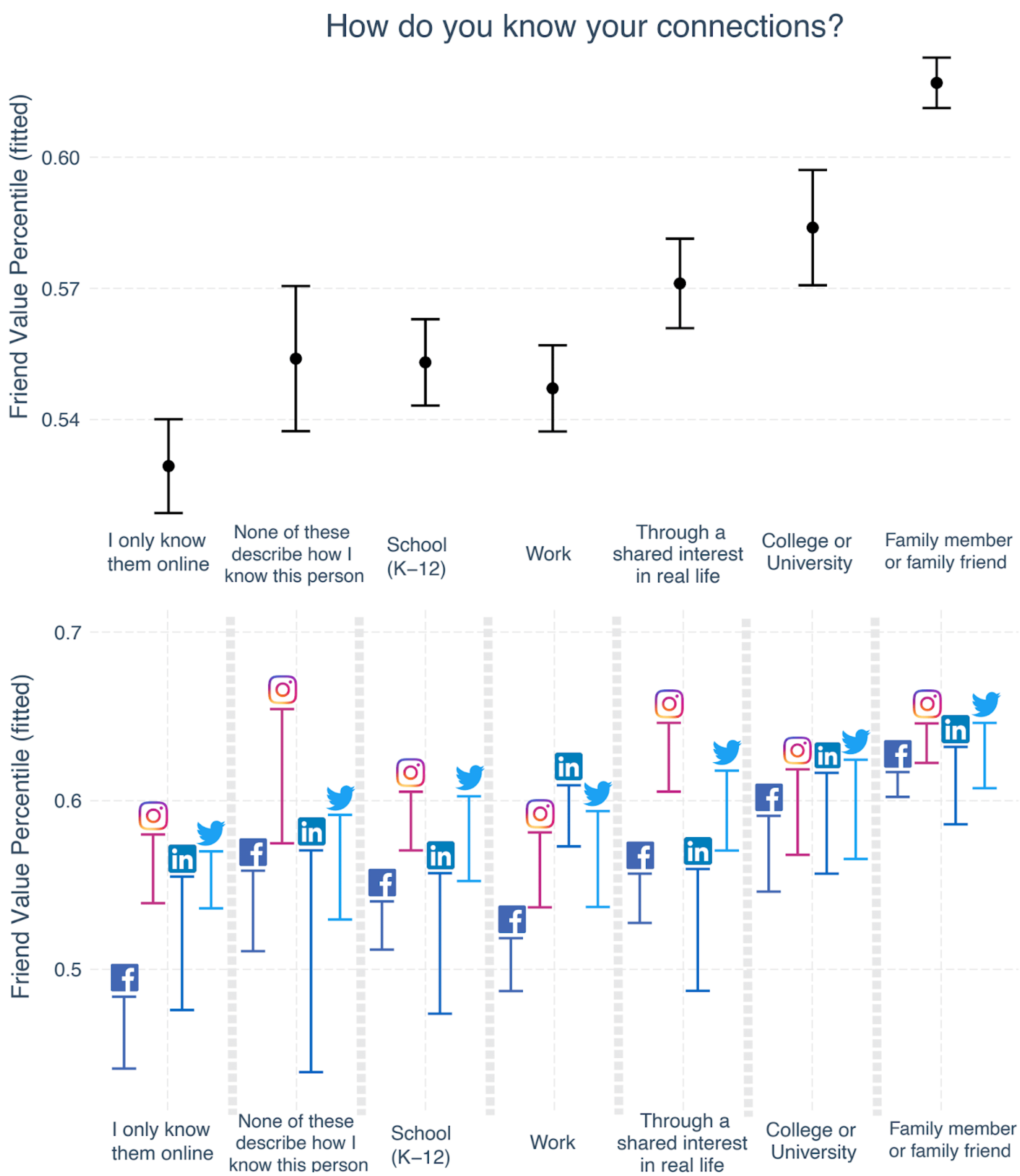


Figure B.2.2 - Figure 2, restricting attention to quality filtered data

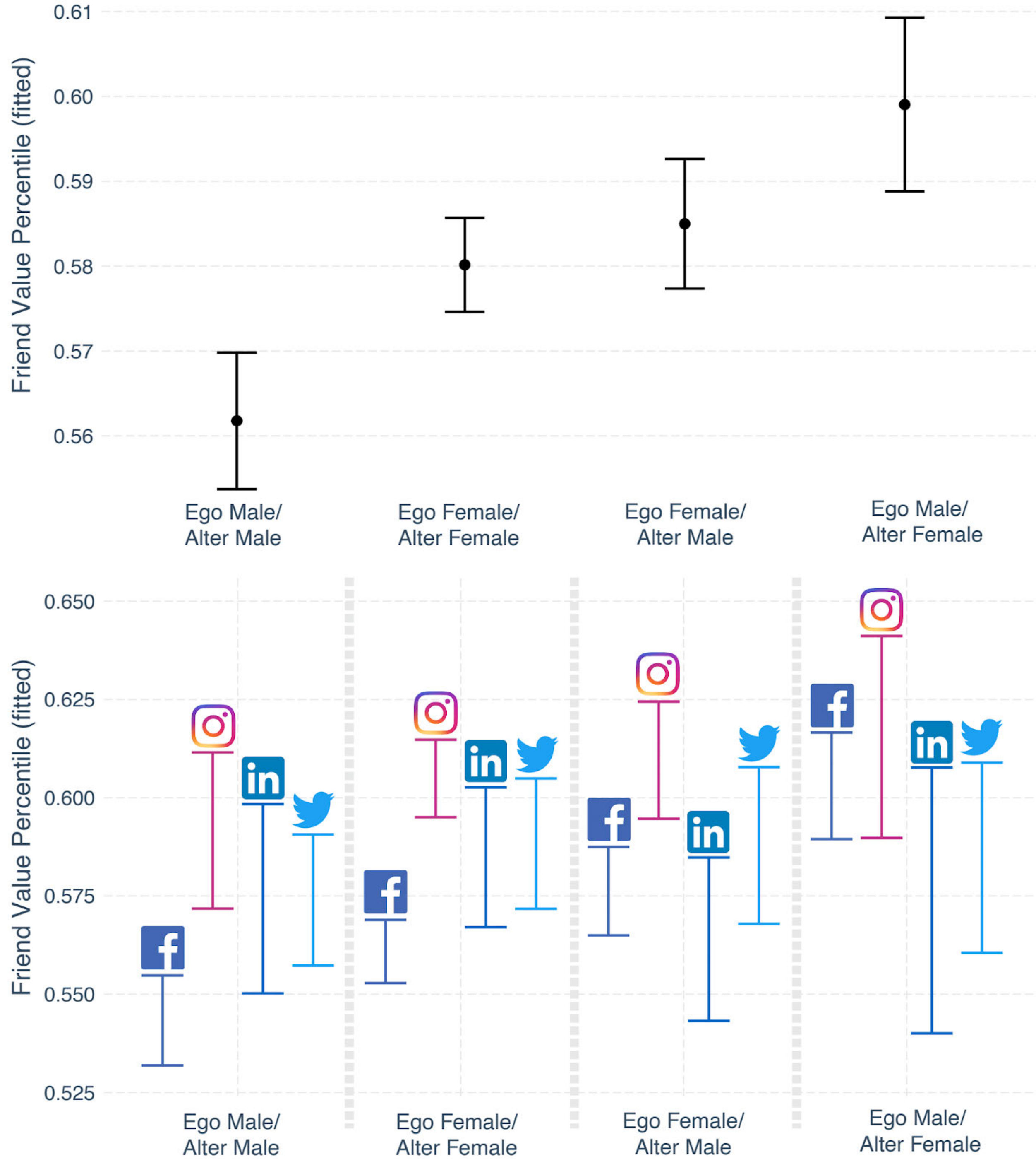


Figure B.2.3 - Figure 3, restricting attention to quality filtered data

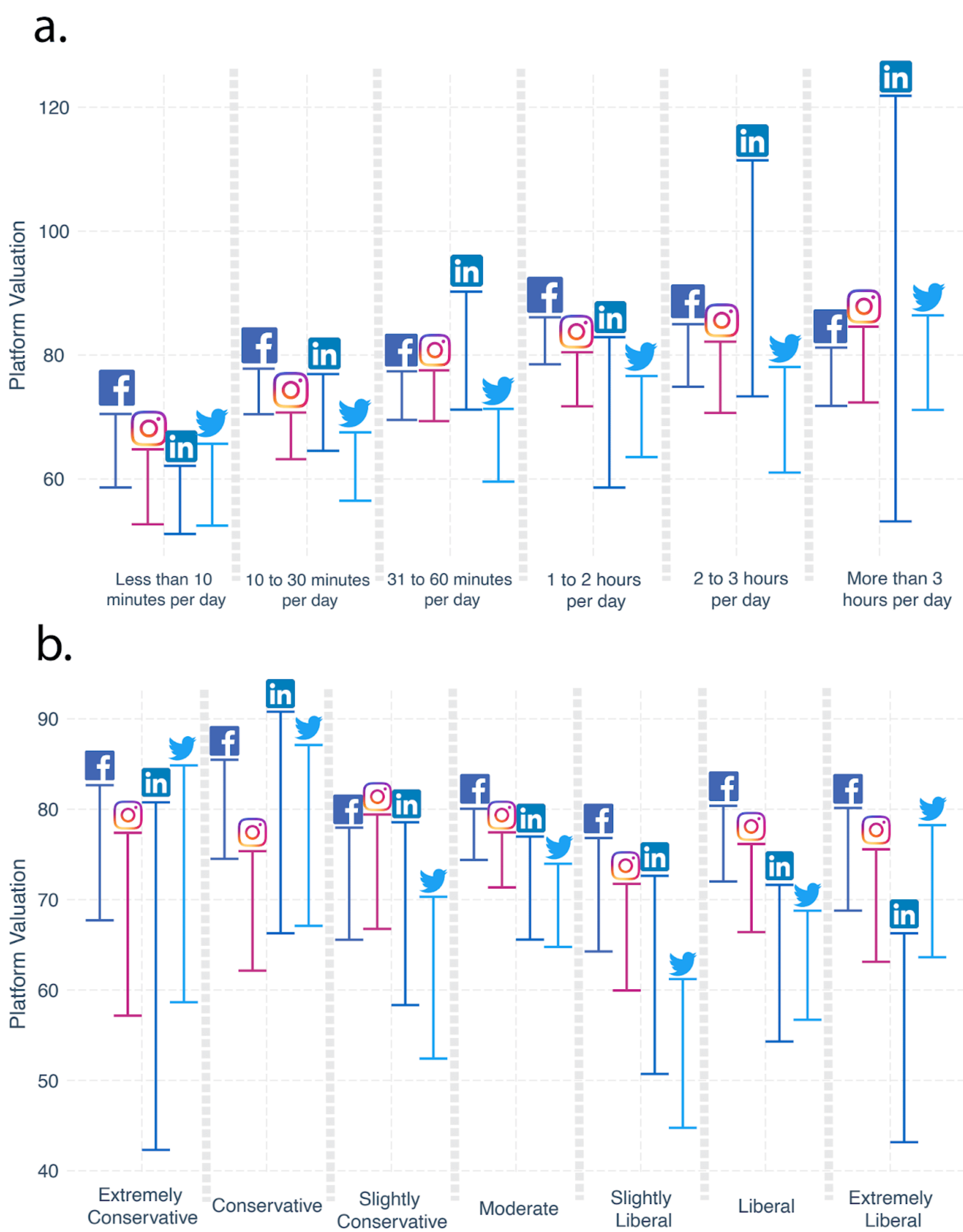


Figure B.2.4 - Figure 4, restricting attention to quality filtered data

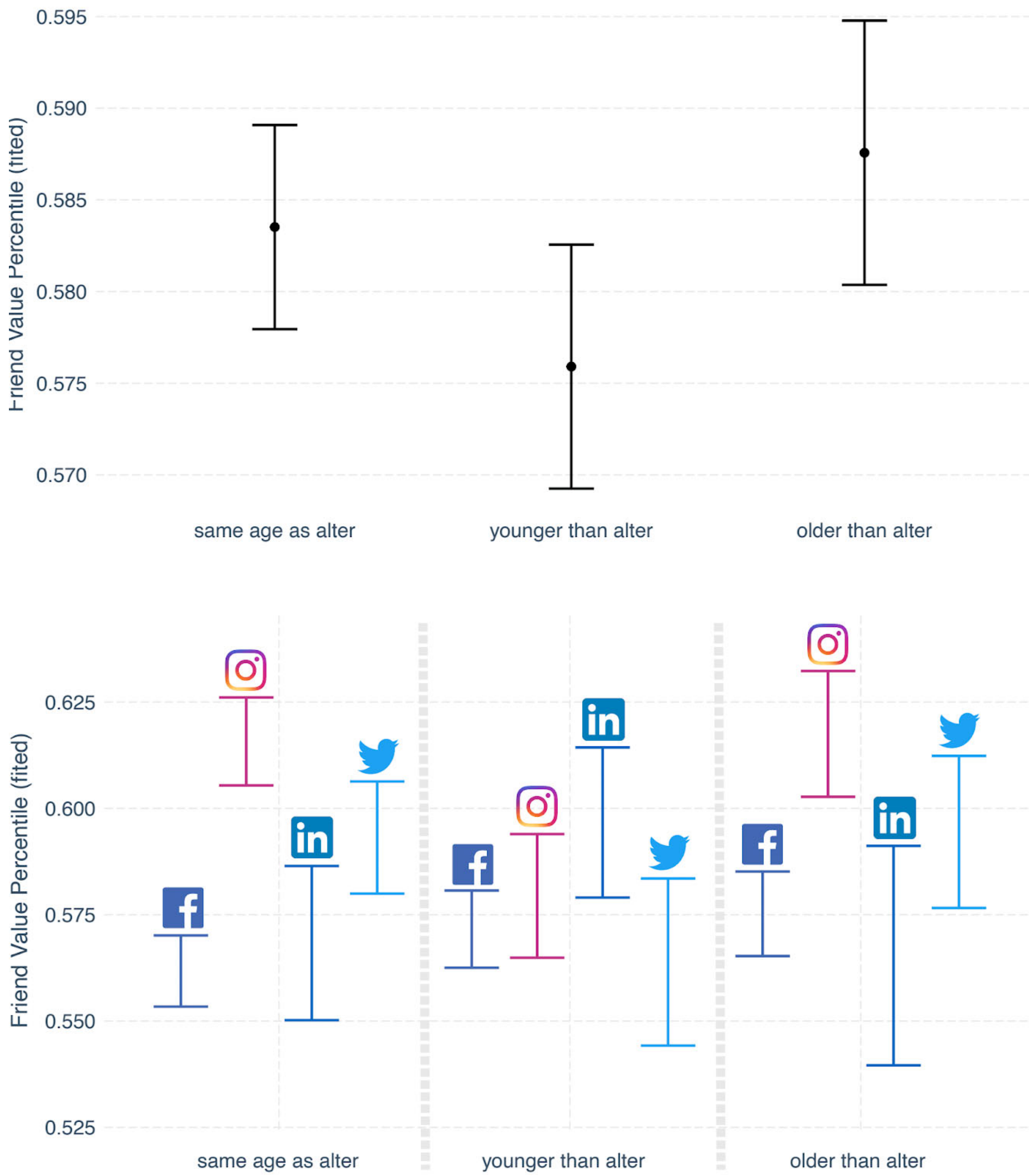


Figure B.2.5 - Figure A.1.1, restricting attention to quality filtered data

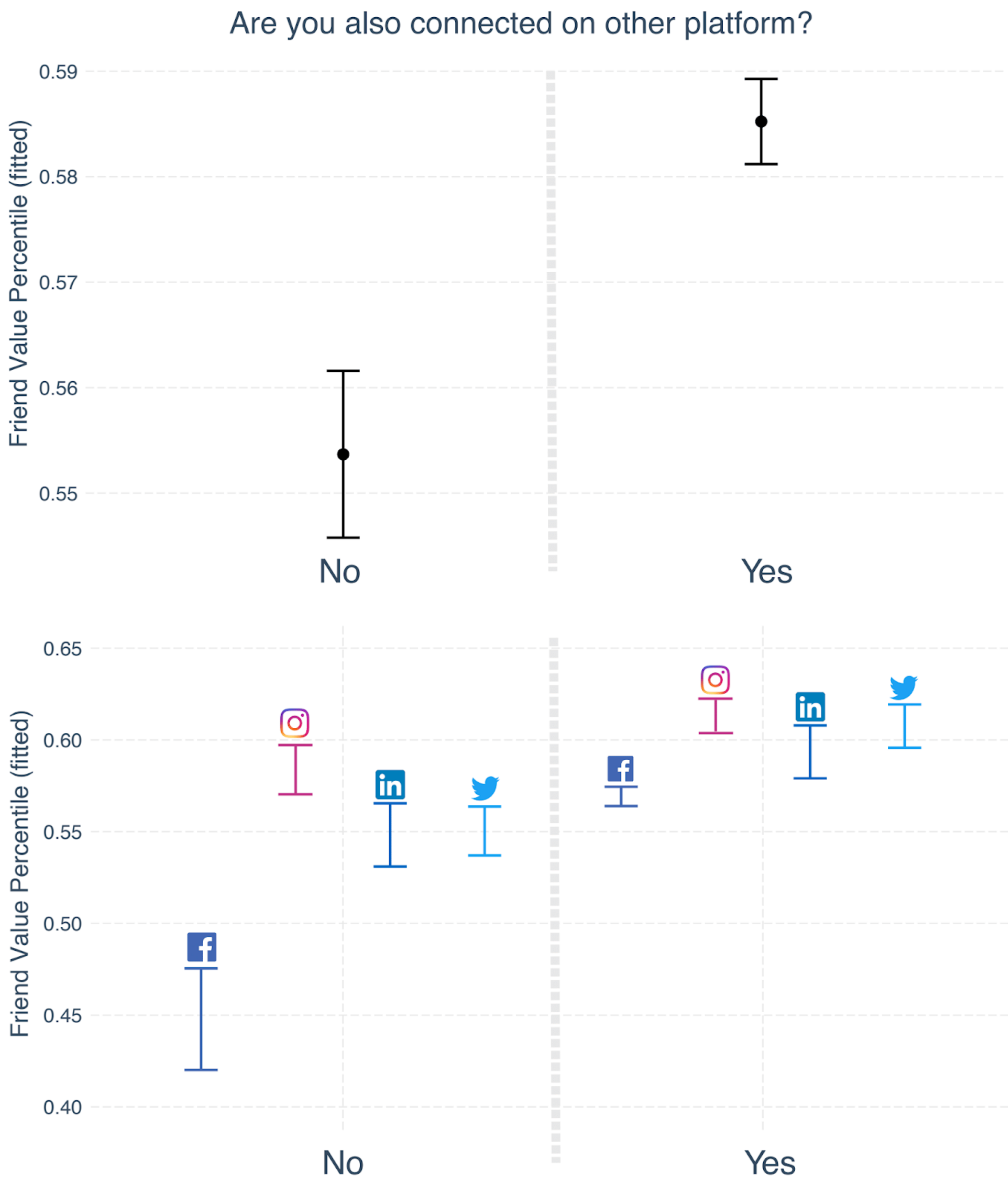


Figure B.2.6 - Figure A.1.2, restricting attention to quality filtered data

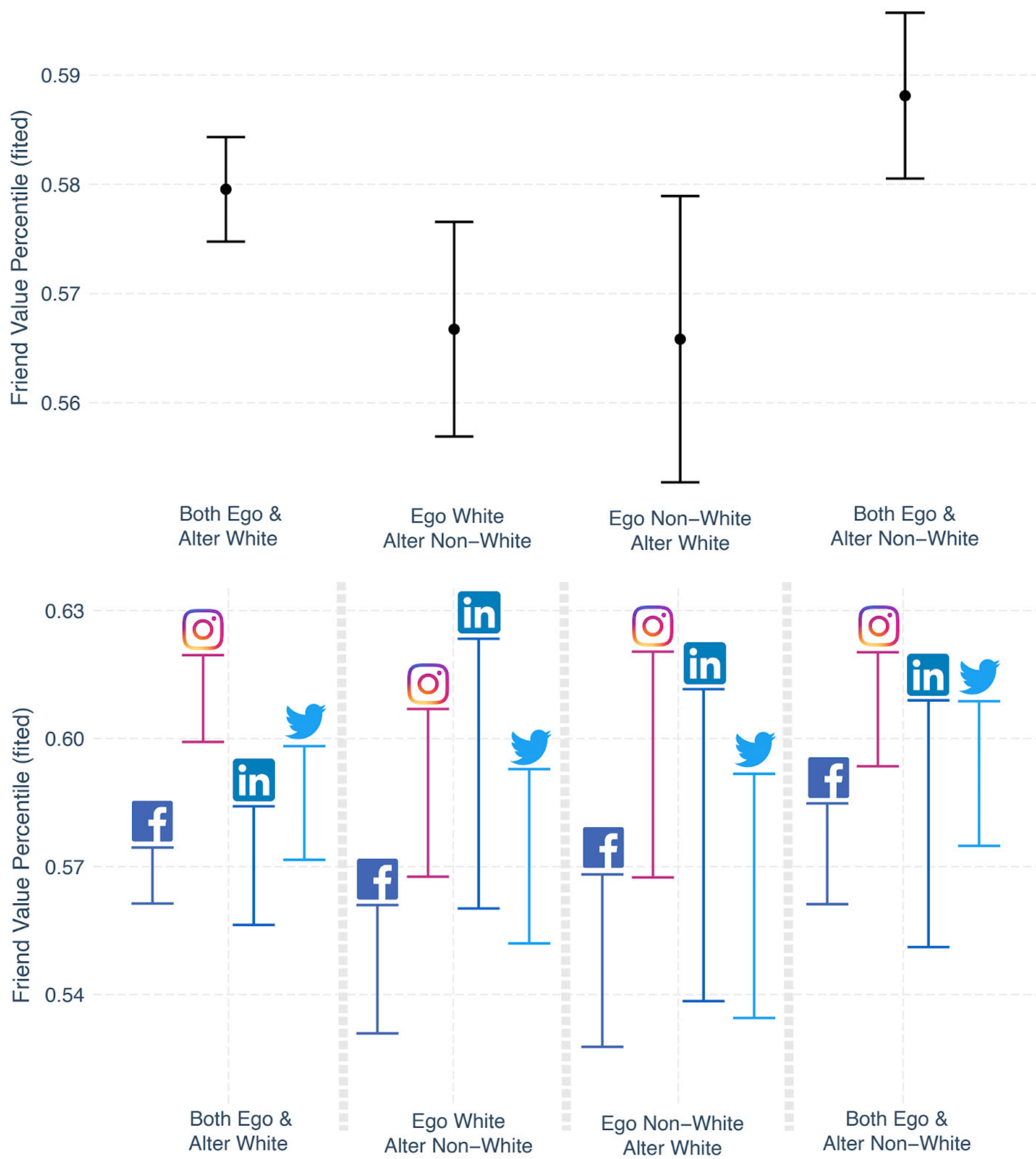


Figure B.2.7 - Figure A.1.3, restricting attention to quality filtered data

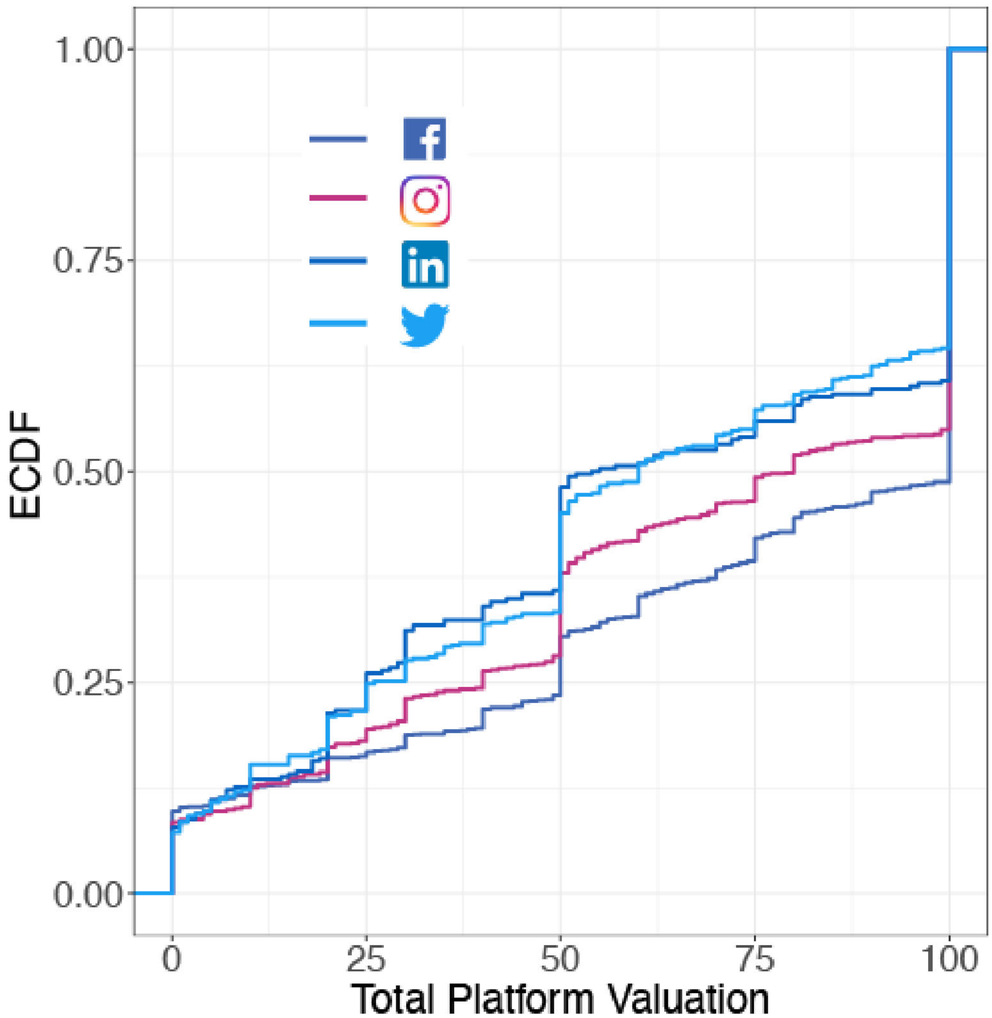


Figure B.2.8 - Figure A.2.1, restricting attention to quality filtered data

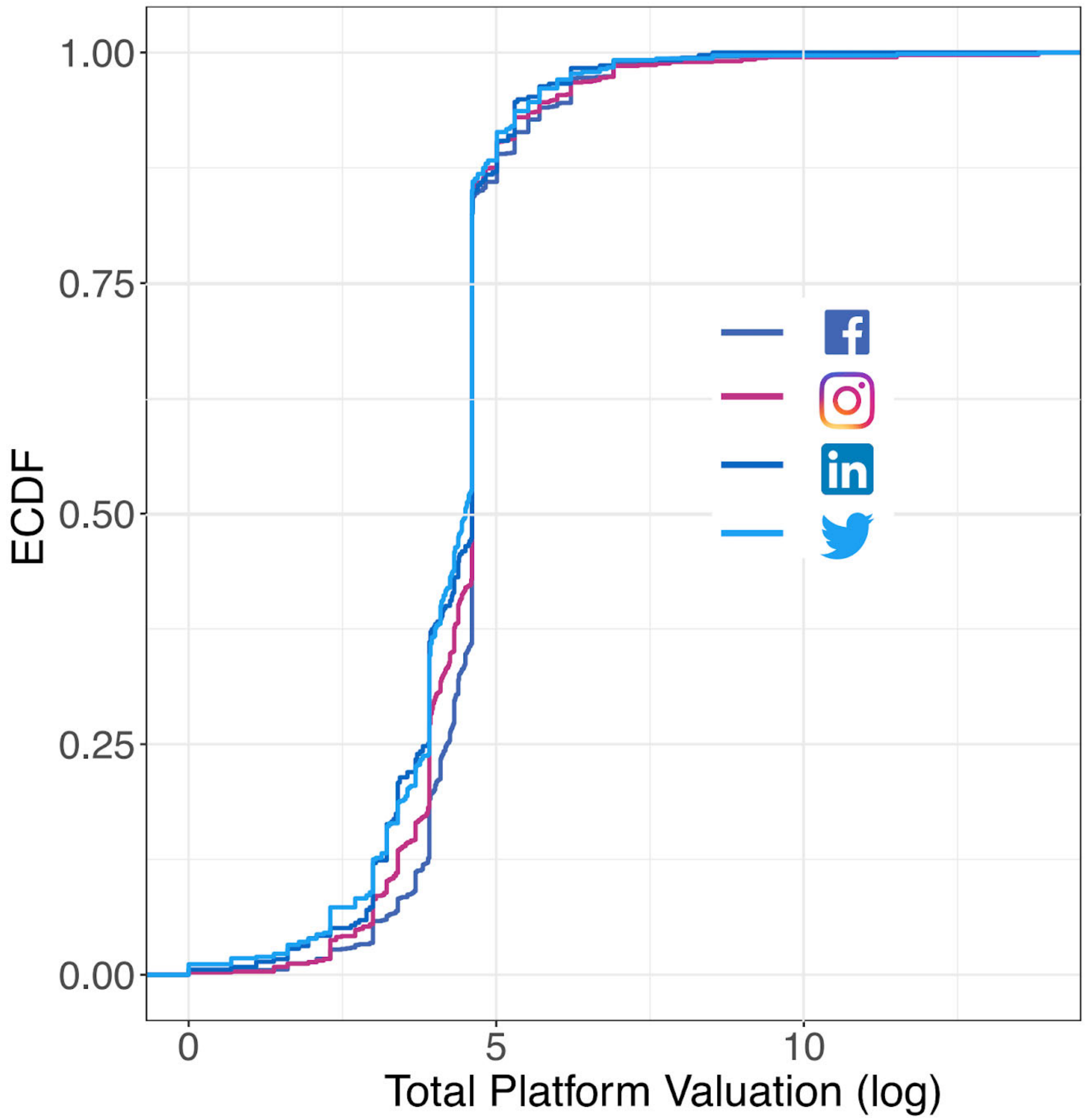


Figure B.2.9 - Figure A.2.2, restricting attention to quality filtered data

	Friend Value Percentile
Survey Platform (Instagram)	0.016^{**}
	(0.007)
Survey Platform (Twitter)	0.031^{***}
	(0.005)
Ego Knows Alter from Work	-0.035^{***}
	(0.007)
Alter a Family member or family friend	0.041^{***}
	(0.004)
Ego meets Alter About once a week	0.050^{***}
	(0.006)
Ego meets Alter Multiple times per week	0.073^{***}
	(0.008)
Ego lives with Alter	0.166^{***}
	(0.009)
Survey Platform (Instagram) x Alter Female	0.011
	(0.007)
Survey Platform (LinkedIn) x Ego Knows Alter from Work	0.062^{***}
	(0.017)
Survey Platform (Instagram) x Ego knows Alter Through a shared interest in real life	0.048^{***}
	(0.011)
Survey Platform (Instagram) x Ego meets Alter About once a month	0.045^{***}
	(0.010)
Survey Platform (Twitter) x Ego meets Alter About once a month	0.044^{***}
	(0.014)
Ego Female x Alter 45-54 years old	-0.025^{***}
	(0.007)
Ego Female x Ego meets Alter Multiple times per week	0.008
	(0.009)
Ego 45-54 years old x Alter a Family member or family friend	0.024^{***}
	(0.009)
Ego 25-34 years old x Ego lives with Alter	0.047^{***}
	(0.015)
Ego 45-54 years old x Ego lives with Alter	0.064^{***}
	(0.024)
Survey Platform (Instagram) x Ego Female x Alter 18-24 years old	0.019^{**}
	(0.009)
Survey Platform (Instagram) x Ego 25-34 years old x Alter 25-34 years old	0.029^{***}
	(0.010)
Survey Platform (LinkedIn) x Ego Female x Ego Knows Alter from Work	0.031
	(0.019)
Survey Platform (LinkedIn) x Ego 25-34 years old x Ego Knows Alter from Work	0.037^{*}
	(0.019)
Constant	0.515^{***}
	(0.003)
N	24,221
R^2	0.053
Adjusted R^2	0.052
Residual Std. Error	0.278
F Statistic	64.496^{***}

$^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

Table B.2.2 -- Table 3, restricting attention to quality filtered data

	Total Platform Valuation
Survey Platform (LinkedIn)	−6.331***
	(1.023)
Ego 18-24 years old	−1.727*
	(1.045)
Ego 35-44 years old	1.822**
	(0.919)
Ego 45-54 years old	4.408***
	(1.211)
Ego Slightly conservative	−2.430
	(1.691)
Ego Slightly liberal	−4.685***
	(1.416)
Survey Platform (Twitter) x Ego 18-24 years old	−1.402
	(1.924)
Survey Platform (Twitter) x Ego 35-44 years old	3.565***
	(1.348)
Survey Platform (Twitter) x Ego Income \$25K - \$50K	−5.118***
	(1.617)
Survey Platform (Twitter) x Ego Income \$50K - \$100K	−6.043***
	(1.436)
Survey Platform (LinkedIn) x Ego Slightly conservative	−5.738*
	(3.450)
Survey Platform (Twitter) x Ego Slightly conservative	−9.181***
	(3.193)
Survey Platform (Twitter) x Ego Slightly liberal	−5.383*
	(2.796)
Survey Platform (Twitter) x Ego Hispanic or Latino	−4.515**
	(2.047)
Constant	72.596***
	(0.584)
N	10,527
R^2	0.019
Adjusted R^2	0.018
Residual Std. Error	34.217
F Statistic	14.467***

8

*p<0.1; **p<0.05; ***p<0.01

Table B.2.3 -- Table 4, restricting attention to quality filtered data

Appendix C -- Survey Instruments

END OF PAPER

****
**
****
STRAY NOTES
****

>Appendix explaining filtering + in text how many were filtered
>details for data on what data was sent back to qualtrics
>Survey instruments
[Christos: RACE FVP INTERACTION HERE?] (I think it’s kind of cool if we don’t find anything)

>conclusion: Seth writes most, get a few sentences from Avi on IC
****
****

Appendix B notes:

Here are the criteria to pass the main filter, and have your data used in the main paper:

Facebook:
>Actually has account which follows us
>Checked whether 5 random friends were actually friends

Twitter:
>Actually has account
>Checked whether 5 random followers were actually followers

LinkedIn:
>Is account real
>Can’t see friends

Instagram:
>For first quarter of obs, same as above,
>But after we got kicked out we had to accept everyone.

****

[Seth: Confirm Details of this With Faiz]

- For Facebook:
  - For people with public friend lists, we confirmed that they list of alters seemed to appear (match is often fuzzy)
  - We also did a survey quality check on the names listed not being silly
- For Twitter
  - For people with public friend lists, we confirmed that they list of alters seemed to appear, minimum number of contacts at least (extremely fuzzy match)
- For Instagram:
  - Friend lists (almost?) always private, so just confirmed they followed us For LinkedIn
  - For people with public friend lists, we confirmed that they list of alters seemed to appear (match is often fuzzy)
  - We also did a survey quality check on the names listed not being silly

Notes towards a stricter filter

1. LINKEDIN:
    a. The verified accounts are the ones where “linkedinprofile” variable is non-missing.
    b. For many of these accounts we also have “connectionscount” ,”jobtitle”, “jobdaterange”,”school”, “major” etc.

2. FACEBOOK:
    a. The subset you’re working with was already filtered once for friend name quality
    b. To be even stricter, we should see what happens when you drop all the observations where FB_Survey_Quality_Check_Notes are “.” or FB_friend_count_tot>10 or the intersection of these.

3. TWITTER:
    a. As a stricter subset, let’s restrict attention to those with screen_name not missing, which means they followed us on Twitter

4. INSTAGRAM:
    a. For the stricter subset, only keep the Yes and yes from the “Accept” variable

<u>Re-running results using all of WTA platform value instead of just <100</u>

Seth 5-7-23 Notes:

- Bits still needing a pass:
    - Abstract -- write last
    - Intro - first draft done
    - Background -- avi needs to write
    - Data description -- avi + christos need to write
    - Results - first draft done
    - Discussion -- Seth has an outline, needs to finish
    - Appendix A - I’d like FVP by ego and alter race added here, as well as FVP by whether the alter has a connection on another platform
    - Appendix B - I’d like to re-run some FVP results here on the stricter subset, finish describing the inclusion criteria in the main subset (Faiz help?) and see TPV with the un-windsorized valuations
    - Appendix C - Include survey instruments

- Additional results desired (Most to least important):
    - Fixing platform specific data results:
        - We want the LinkedIn appendix table to have additional results on the ego’s job history
    - Appendix A:
        - FVP as a function of connection on other platforms
        - FVP results by ego and alter race
    - Appendix B:

        - Results using stricter ego filtering criteria
        - Re-running WTA results using the full free-text response (>$100) instead of just the slider.

- Some details Seth needs help filling in:
    - Table 5: Christos can you please list the things being thrown into the lasso bucket and the baseline levels into the caption? You can follow the wording in table 3
    - Table A1: Facebook friend value results. Why do standard errors explode in column 2?
    - Let's talk about what results and limitations to focus on in the discussion/conclusion

- Other 5/9/23 meeting notes: Exact timing for Instagram closeout moves; go through Instagram exit questionnaire; Do we want to apply for MISQ author-panel session due today, papers discussed on June 19.

- May 21 is end of Instagram study period
- May 22 -- Have megan send out compliance payments
- May 22 - Followup questionnaire once the payments go out
- May 29 - Deadline for getting in questionnaires
- May 30 - Final questionnaire payouts
- May 31 (deadline? For payments)

Social media platforms, where users both produce and consume content from other users, have sometimes been conceptualized as a one-sided platform [citation], or as a two-sided platform with ordinary users constituting one side and advertisers the other [citation]. For this project, we conceptualize the sides of a social media platform as being constituted of different demographic, social and familial groups. This is a necessary framing for understanding whether, for example, it is best to subsidize usage on colleges or among older users.This is an important question not just for a platform manager, but for any regulator that hopes to maximize the social benefit of a platform, because you need to take into account the platform's optimal response to any policy.

More broadly, measuring the matrix of network effects that users of different communities provide to each other is essential for modeling participation on social media platforms. It is therefore key to both platform managers and aspiring regulators.

Understanding the matrix of network effects is important, because that parameter, in combination with an estimate of elasticity of demand for the platform, is essential for modeling both the dynamics of platform participation and equilibrium platform participation in any detail.

To Do Today:

- Decide how to discuss multiple hypothesis test issues, and in-text sig differences
    - Maybe we just have a discussion of this in a “limitation” section, and can mention if any of the results survive there?
        - Answer: Qualitative discussion in the limitations for now; note that LASSO doesn’t have this problem
- Decide on which figures tables are main text vs. appendix
    - MS Fast Track Restrictions:
        - 6,000 words, including references
        - No explicit figure maximum?
- Decide on what our “Main” contributions we want to highlight in the intro are
- I think we have everything we need (in terms of results) for a first draft (except data description). Is there anything else we absolutely need? Anything else we want?
    - Potential Needs:
        - Data description table
        - Combine the three different connection value tables into one table with one column per platform, put full results in the appendix
    - Potential Wants:
        - Checks to see which significant differences survive a multiple hypothesis adjustment
        - FVP on ego and alter race by platform
        - Re-running all main results with restricted samples
        - Try value of all connection results with log friends explaining log value; or just log friends explaining straight value
        - See if we can get additional LinkedIn results based on other ego page characteristics

Our research also connects to a larger and older literature on the value of social networks. These have been shown to play a decisive role in everything from early modern European warfare [Benzell and Cooke 2020] to modern job search [Matt Jackson 2011 Handbook Chapter]. [Avi: Say something here about how our results speak to older literature]

Our paper’s main contributions and findings are [Seth: Summarize Main Findings]

- First paper applying the same methodology to measuring heterogeneity in network effects across multiple platforms.
- Specifically find that (Make sure all these results hold up with multiple testing adjustment)
    - Across almost all platforms “closer” friends are valued more than less close friends (in terms of freq of meet and how know) with this effect most extreme for Facebook, and less extreme for Twitter. This is consistent with Facebook being built around close personal connecctions, and Twitter around interacting with

strangers and celebrities. It is also consistent with LinkedIn creating value through ties of “medium” strength as consistent with (new IDE paper)

- Across all platforms male egos reported a preference for female alters over male alters, with the effect strongest on Facebook and Instagram. Female egos did not report a preference for alters by gender on any platform.

Most important results ranking tier list:

A. Each friend on FB is worth .5 cents a month
B. People value more being connected to close connections (people seen more often, and with deeper social connection) than loose connections:
    a. I.e. direct effect of person being valuable contact overwhelms the substitution effect of close contacts having alternate modes of communication
    b. For Twitter and Instagram, this effect is lessened 0-- strangers valued relatively more
C. Lots of other heterogeneity across LOTS of dimensions, by age and gender for example
    a.
    b. Age Results:
        i. Political moderates don’t like Twitter and LI as much
            1. Because Twitter is for spicy hot takes
            2. (but this is only about the ego so less interesting?)
        ii. On insta, people don’t like being connected to others younger than them as much
    c. Gender Results:
        i. Men value female connections on the non-work platforms
    d. Middle aged value twitter, young people don’t
    e. Value of platforms is increasing in time used, but this flattens at 3 hours
    f. More Retweets = more value
    g. Asians are bad friends

% In the text, we also report t-tests for significant differences between groups. These t-tests are adjusted for multiple hypothesis testing using [Christos: Choose a multiple hypothesis testing adjustment]

Priorities for Christos (most important first):

1. One table of LASSO results on predicting FVP using ego, alter characteristics, and platform, and their interactions. For section 5.1.1

	Friend Value Percentile
Survey Platform (Instagram)	0.011**
	(0.005)
Alter Female	0.012***
	(0.003)
Alter a family member or family friend	0.030***
	(0.003)
Ego knows Alter only online	−0.007
	(0.005)
Ego meet Alter multiple times per week	0.024***
	(0.004)
Ego lives with Alter	0.120***
	(0.009)
Survey Platform (Instagram) x Alter Female	0.001
	(0.006)
Survey Platform (Instagram) x Ego knows Alter through a shared interest	0.026***
	(0.009)
Survey Platform (Instagram) x Ego meets alter once a month	0.025***
	(0.008)
Ego Female x Ego meets alter less than 4 times/year	−0.013***
	(0.004)
Ego Female x Ego meets alter once a week	0.024***
	(0.005)
Ego Female x Ego meets alter multiple times per week	0.017***
	(0.006)
Ego Female x Ego lives with Alter	0.018*
	(0.010)
Ego 35-44 years old x Ego knows Alter only online	−0.021**
	(0.008)
Ego Income \$25,000 - \$49,999 x Ego meets alter multiple times per week	0.014**
	(0.007)
Ego Income \$25,000 - \$49,999 x Ego lives with Alter	0.018
	(0.011)
Survey Platform (Instagram) x Ego Female x Alter 18-24 years old	0.015**
	(0.007)
Survey Platform (Twitter) x Ego income 150,000 or more x Ego lives with Alter	−0.095***
	(0.022)
Constant	0.544***
	(0.002)
N	50,679
R^2	0.023
Adjusted R^2	0.023
Residual Std. Error	0.282
F Statistic	67.691***

*p<0.1; **p<0.05; ***p<0.01

2. One table of LASSO results on predicting total platform value using ego characteristics, platform and their interaction. Use variables that are available for all platforms. For

section 5.2.1

	Total Platform Valuation
Survey Platform (LinkedIn)	-6.566^{***}
	(1.030)
Political Orientation (Slightly liberal)	-5.058^{***}
	(1.666)
Survey Platform (Twitter) x Ego 18-24 years old	-5.970^{***}
	(1.639)
Survey Platform (Twitter) x Ego 35-44 years old	4.355^{***}
	(1.190)
Ego Income \$50,000 - \$99,999	-5.455^{***}
	(1.418)
Survey Platform (Twitter) x Ego political orientation (Slightly conservative)	-12.334^{***}
	(2.748)
Survey Platform (Twitter) x Ego political orientation (Slightly liberal)	-6.034^{**}
	(2.926)
Constant	72.838^{***}
	(0.527)
N	7,894
R^2	0.018
Adjusted R^2	0.017
Residual Std. Error	34.538
F Statistic	20.505^{***}

$^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

3. Two tables for 5.2.2 Platform-Specific Ego and Network Characteristics and Platform Value. For Linkedin predict platform WTP as a function of additional ego characteristics we scraped -- about 2- 4 specifications. For Facebook, predict WTP as a function of friend counts, friend count by type, and these interacted with an interesting ego characteristic or two. About 4 specifications.

	Facebook Monthly Valuation		
	(1)	(2)	(3)
Total number of friends	0.005**		
	(0.002)		
Number of Female Friends		0.007	
		(0.495)	
Number of Male Friends		0.003	
		(0.503)	
Number of Black Friends			−0.080
			(0.070)
Number of White Friends			−0.030
			(0.020)
Number of Hispanic Friends			0.005
			(0.011)
Number of Asian Friends			−0.135**
			(0.060)
Number of Nat. American Friends			0.846
			(0.870)
Number of bi-racial Friends			1.771
			(1.273)
Constant	78.874***	78.778***	79.334***
	(1.452)	(1.473)	(1.491)
N	829	829	829
R^2	0.005	0.006	0.018
Adjusted R^2	0.004	0.003	0.011
Residual Std. Error	30.606	30.622	30.505
F Statistic	4.502**	2.323*	2.506**

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

	Twitter Monthly Valuation	
	(1)	(2)
count of followers	-0.0004	
	(0.003)	
count of followees	0.004^{*}	
	(0.002)	
total retweets		0.011^{***}
		(0.003)
total quotes		0.012
		(0.015)
total replies		-0.004^{*}
		(0.002)
favorites count		0.001
		(0.001)
Constant	67.847^{***}	67.788^{***}
	(1.465)	(1.726)
N	692	621
R^2	0.008	0.027
Adjusted R^2	0.006	0.021
Residual Std. Error	34.154 (df = 689)	33.957 (df = 616)
F Statistic	2.922^{*} (df = 2; 689)	4.306^{***} (df = 4; 616)

$^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$

	LinkedIn Monthly Valuation
Connections Count	0.001
	(0.001)
Constant	68.497^{***}
	(1.947)
N	344
R^2	0.006
Adjusted R^2	0.003
Residual Std. Error	35.392
F Statistic	1.982

$^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$

	Instagram Monthly Valuation
Pre-account blockage	-4.813^{***}
	(1.621)
Constant	72.530^{***}
	(1.048)
N	1,981
R^2	0.004
Adjusted R^2	0.004
Residual Std. Error	35.578
F Statistic	8.818^{***}

$^{*}\ p < 0.10$, $^{**}\ p < 0.05$, $^{***}\ p < 0.01$

4. Data description section:
    a. A data summary table for the main text
    b. A data summary table for the even more restricted sample
    c. 150 words of discussion in the main text (maybe get Avi's help)
5. Redo current section 5.2 figures to be more like 5.1 (version b) figures
6. Evaluate the multiple hypothesis adjusted t-tests I point you towards in the text (maybe something you could delegate to Avi if you give him the raw data on point estimates and intervals?)
7. More overall platform value results (section 5.2)
    a. WTP by platform:
        i. Can we plot the full CDFs (which are demand curves) rather than only focusing on the average?
    b. Platform x Ego Race -- make figures analogous to section 5.1 figures, with platforms superimposed within a category, unlike current section 5.2 figures
    c. Platform x Gender -- make figures analogous to section 5.1 figures, with platforms superimposed within a category, unlike current section 5.2 figures
8. A table on Twitter for 5.2.2 explaining Twitter WTP as a function of Twitter API ego characteristics (still hoping to collect a bit more info here, so low priority).
9. Alt versions of all figures with restricted sample for appendix

Monthly Total Platform Valuation	<u>Mean</u>	<u>Median</u>	<u>SD</u>
Facebook	\$76.0	\$100	34.6

Instagram	$70.5	$99	35.6
LinkedIn	$65.8	$75	35.7
Twitter	$69.9	$82	33.6

Table 4: Summary statistics for the total monthly valuation of the platform for the different platforms.

OLD MAIN TABLE

	FACEBOOK (N = 1,516)	**INSTAGRAM (N= 1,981)**	**LINKEDIN (N = 1,499)**	**TWITTER (N = 2,899)**
Gender				
Female	956 (63%)	1,424 (72%)	853 (57%)	1,236 (43%)
Male	551 (36%)	543 (27%)	637 (42%)	1,641 (57%)
Other	9 (0.6%)	14 (0.7%)	8 (0.5%)	22 (0.8%)
Age				
18-24 years old	242 (16%)	587 (30%)	239 (16%)	520 (18%)
25-34 years old	495 (33%)	713 (36%)	489 (33%)	1,048 (36%)
35-44 years old	374 (25%)	388 (20%)	436 (29%)	1,050 (36%)
45-54 years old	214 (14%)	157 (8.0%)	169 (11%)	223 (7.7%)
55-64 years old	112 (7.4%)	72 (3.7%)	91 (6.1%)	42 (1.5%)
65 years or older	78 (5.1%)	49 (2.5%)	73 (4.9%)	12 (0.4%)
Ethnicity				
White	1,115 (74%)	1,189 (60%)	1,014 (68%)	1,873 (65%)
Black	180 (12%)	354 (18%)	227 (15%)	534 (18%)
Hispanic or Latino	124 (8.2%)	278 (14%)	150 (10%)	317 (11%)
Asian/Pacific Islander	55 (3.6%)	100 (5.0%)	66 (4.4%)	106 (3.7%)
Native American.	18 (1.2%)	19 (1.0%)	14 (0.9%)	33 (1.1%)
Other	24 (1.6%)	41 (2.1%)	27 (1.8%)	36 (1.2%)
Income				
Less than $25,000	382 (25%)	448 (23%)	226 (15%)	384 (13%)
$25,000 - $49,999	468 (31%)	627 (32%)	386 (26%)	560 (19%)
$50,000 - $99,999	421 (28%)	598 (30%)	507 (34%)	711 (25%)
$100,000 - $149,999	151 (10.0%)	210 (11%)	230 (15%)	742 (26%)

$150,000 or more	94 (6.2%)	98 (4.9%)	149 (9.9%)	502 (17%)
Political orientation				
Extremely conservative	86 (5.7%)	101 (5.1%)	70 (4.7%)	241 (8.3%)
Conservative	160 (11%)	169 (8.5%)	147 (9.8%)	228 (7.9%)
Slightly conservative	125 (8.2%)	165 (8.3%)	143 (9.5%)	165 (5.7%)
Moderate	600 (40%)	798 (40%)	566 (38%)	908 (31%)
Slightly liberal	122 (8.0%)	207 (10%)	137 (9.1%)	217 (7.5%)
Liberal	274 (18%)	312 (16%)	255 (17%)	580 (20%)
Extremely liberal	149 (9.8%)	229 (12%)	180 (12%)	560 (19%)
Platform Valuation				
Median	100	99	75	82
Mean	76	70.5	65.8	69.9
SD	34.6	35.6	35.7	33.6

Table 2. Descriptive Statistics of the survey takers (ego) demographics.