The Spread of Physical Activity Through Social Networks

David Stück
Evidation Health, Inc.
510 State St, Suite 200
Santa Barbara, CA
dstuck@evidation.com

Haraldur Tómas Hallgrímsson*
University of California
Santa Barbara, CA, USA
hth@cs.ucsb.edu

Greg Ver Steeg
University of Southern California
Los Angeles, CA
gregv@isi.edu

Alessandro Epasto†
Google
111 Eighth Avenue
New York, NY
aepasto@google.com

Luca Foschini
Evidation Health, Inc.
510 State St, Suite 200
Santa Barbara, CA
luca@evidation.com

ABSTRACT

Many behaviors that lead to worsened health outcomes are modifiable, social, and visible. Social influence has thus the potential to foster adoption of habits that promote health and improve disease management. In this study, we consider the evolution of the physical activity of 44.5 thousand Fitbit users as they interact on the Fitbit social network, in relation to their health status. The users collectively recorded 9.3 million days of steps over the period of a year through a Fitbit device. 7,515 of the users also self-reported whether they were diagnosed with a major chronic condition. A time-aggregated analysis shows that ego net size, average alter physical activity, gender, and body mass index (BMI) are significantly predictive of ego physical activity. For users who self-reported chronic conditions, the direction and effect size of associations varied depending on the condition, with diabetic users specifically showing almost a 6-fold increase in additional daily steps for each additional social tie. Subsequently, we consider the co-evolution of activity and friendship longitudinally on a month by month basis. We show that the fluctuations in average alter activity significantly predict fluctuations in ego activity. By leveraging a class of novel non-parametric statistical tests we investigate the causal factors in these fluctuations. We find that under certain stationarity assumptions, non-null causal dependence exists between ego and alter's activity, even in the presence of unobserved stationary individual traits. We believe that our findings provide evidence that the study of online social networks have the potential to improve our understanding of factors affecting adoption of positive habits, especially in the context of chronic condition management.

*Work done while interning at Evidation Health
†Work partially done while at Brown University

ACM 978-1-4503-4913-0/17/04.
http://dx.doi.org/10.1145/3038912.3052688

Keywords

Digital Health; Wearables; Network Science; Dynamic Networks

1. INTRODUCTION

It is estimated that up to 40% of premature deaths in the United States can be attributed to behavior and lifestyle [32], such as level of physical activity and eating habits. Body weight abnormalities alone are estimated to cost \$190.2 billion, 21% of annual medical spending [7, 14].

Nevertheless, negative behaviors conducive of poor health outcomes have been on the rise. Average daily occupation-related energy expenditure has decreased by over 100 calories from 1960 to 2010 [10] and, even if new policies and campaigns [1] have been successful in slowing down the prevalence of obesity in recent years [28], the overall picture is still daunting. A recent study reported that the adult obesity rate in the U.S. increased by more than two percentage points from 25.5% to 27.7% in just six years [25].

Some of the behaviors leading to poor health outcomes are modifiable and can be both social and visible. A large body of research has shown that observing behavior in others has a profound effect on one's own behavior, highlighting that the normative effects of one's social circle can vary in polarity depending on the outcomes under examination and how the social interaction is designed [33]. For example, a randomized control trial on African-American veterans showed that pairing a diabetic individual with a mentor significantly outperformed usual care and even financial incentive to achieve glucose control [20]. On the contrary, another randomized field experiment in a large corporation showed that treadmill usage declined among employees when participants were given information on coworkers' usage levels, due to a tendency to converge to their least-active coworkers [16].

The advent of online social networks has greatly advanced our understanding of how to characterize and shape people's behavior in a social context [8, 17]; however, effects on "offline" behavior, such as those on lifestyle and health outcomes, have been hard to study due to lack of data. This has changed very recently with the advent and rapid adoption of *mobile health* (mHealth) enabled by wearable technologies, which has made continuous monitoring of environment and lifestyle ("life logging" [18]) a concrete possibility. The ability to continuously collect data from very large cohorts in addition to online interactions can finally facilitate our understanding of how social network interactions affect the user's behavior in the real world.

Contribution

In this work we study the relationship between a social network and physical activity in a population of 44.5 thousand Fitbit users, 7,515 of which self-reported whether they had been diagnosed with major chronic conditions. The users interacted through the Fitbit social network over the course of one year.

- We shed light on the health conditions that burden the trackers and quantify how those factors affect the interaction between the social network and physical activity.
- We perform a between-subject analysis and show that people who have more friends on the Fitbit social network are more likely to have higher physical activity, and provide evidence that other features of the social network, such as average BMI, are also associated with increased physical activity, sometimes in a counter-intuitive way.
- To study how the social network and physical activity evolve over time, we perform a within-subject analysis confirming the direct association between fluctuations in physical activity of a user and the average physical activity of the user's friends over time.
- We perform a causality test on the network over time and we reject the hypothesis that the activity can be explained exclusively by homophily (intrinsic or extrinsic) under certain stationarity assumptions. We also provide an in-depth discussion of the limitations of our analysis in the case when non-stationary unobserved external causes of correlation are present.

2. RELATED WORK

Several studies have looked at the influence of social networks on physical activity both in retrospective and prospective settings. Before the advent of widespread mHealth technologies, either the physical activity or the social structure had to be self-reported, which significantly limited the studies in the size of the cohort studied, the longitude of the study, or both. In a study [24] conducted on 310 primary school students it was shown that the perceived activity levels of self-reported friends were associated with outside-of-school physical activity and sedentary time, as measured by accelerometers. The friendship structure of the network was self-reported and did not change over time. In a similar setting, an insufficient level of physical activity was found to be tied with lower social support, with a strong effect due to gender, in a sample of 2,729 male and female college students [19].

The availability of datasets including repeated measurements for the same user over time has enabled the study of the effect of network dynamics on health-related outcomes longitudinally. A large body of work has studied the contribution of social networks on physical health in the context of the obesity epidemic [9, 35]. Christakis and Fowler [9] studied the increased clustering of obese people over time in a social network, considering homophily, confounding, and inductive, or causative, effects over both geographical as well as social distance, and found a significant social contagion effect that was found to be stronger in friends of the same sex. However, a later analysis by Russell Lyons [21] disputed the causal implication derived from Christakis and Fowler's results.

More recently, the increased popularity of the Internet of Things has made available datasets for large populations where social interactions and health-related outcomes are tracked electronically over time.

In the context of weight tracking, Ma et al. [22] analyzed an online social network of 107 thousand users including five months of self-reported weigh-ins. Users' weight change was found to correlate positively with number of friends and friends' weight change performance, with effects greater than what was measured in a real-world social network. Physical activity has been considered recently in [12] which studied participants using weekly pedometer logs as well as a social network associated with the pedometer device that provided friendship and posting capabilities. They used sentiment analysis to determine whether posts visible from friends had a positive (e.g., walk, run) or negative (e.g., sick) connotation. They examined the group-level influence these posts had on other users who could read them and found that a person's positive post has a significantly higher probability of propagating physical activity to other people in the network than does a negative post. Despite the online nature of social interaction and physical activity measurements, the cohort considered only contained 254 users.

During the preparation of this article, a preprint of the concurrent and independent work of Althoff et al. [3] was made available to us by the authors. Althoff et al. study how social networking features influences user behavior in a physical activity tracking application. They consider the evolution of a social network of 211,383 users over 3 years for which friend requests and acceptances are observed. By comparing activity increase over all friend requests accepted either immediately (N=34,324) or after a week (N=3,146) they estimate that edge formation is associated with an increase of 328 daily steps in the sender of the request. The authors perform a difference-in-difference analysis to conclude that 55% of the total increase in activity post-tie formation can be attributed to social influence (sender becoming aware of the receiver's activity) and 45% is due to the user's elevated intrinsic motivation (which increases both the likelihood of sending friend requests and activity). In addition, for 6,076 out of the 211,383 social network users matching inclusion/exclusion criteria the authors show an increase of 7% in average daily steps in the first week after joining the social network, as compared to a matched control group. The activity spike diminishes slowly over the course of 20 weeks.

Our work departs from previous research in that i) it is the first to consider the influence of social networks and physical activities in relation to self-reported chronic health conditions on a large cohort; ii) studies how average activity of alters influences ego's average activity, which takes into account both the addition and removal of friends and change of activity in current friends; and iii) it considers the causality implication of social network effects on physical activity under the lens of graphical models.

3. DATA

3.1 Data Collection

The dataset considered consists of 44,468 members of a commercial reward platform for aggregating healthy activities (Achievemint, powered by Evidation Health, Menlo Park, CA) observed in the period between 9/30/2015 and 9/30/2016. On the Achievemint platform, members link their activity trackers (e.g., Fitbit pedometers and Wi-Fi scales, Jawbone trackers, etc.) and mobile applications (e.g., MyFitnessPal, RunKeeper) by authorizing their data to be relayed to their reward platform account. Members can connect multiple applications and trackers to the platform. For example, some members might have connected both a Wi-Fi scale and a food journaling application, while others may have connected a pedometer and a workout-tracking application. For every new activity reported through their third-party applications and devices, members earn points. Points are redeemable for cash rewards: after a member has achieved 1,000 points, they will earn $1.00. Members receive a check for every $25.00 earned.

Consent for participation in this study was obtained electronically by accepting a terms of service contract for the reward platform. The study was approved by Solutions Institutional Review Board and determined to be exempt from the OHRP's Regulations for the Protection of Human Subjects (45 CFR 46).[1] Personally Identifiable Information (PII) has been collected, processed, and stored by Evidation Health. It has not been shared with any institutions the authors are affiliated with, except Evidation Health.

Inclusion/Exclusion Criteria

Subjects included in the study are those who are Achievemint members on 9/30/2016,[2] have connected a Fitbit device (any model) to the platform, and authorized access to their Fitbit data, which includes their profile information. Their profile contains information such as the user's age, gender, height, weight, Fitbit join date, and links to de-identified friends' profiles.

For each Achievemint member we consider activity data logged by their Fitbit devices that was synced with the Achievemint platform. We include members with at least 7 days of recorded activity and daily average step counts on the recorded days between 500 and 40,000.

Friendship on the Fitbit network is a symmetric relationship. Two Achievemint members listing each other as friends on Fitbit would appear with consistent identifiers in both members' friends list and can be matched to each other; however, many Fitbit users listed as friends in some Achievemint member's Fitbit profile may not be part of Achievemint. For Fitbit users who were not Achievemint members we can only observe age and gender (as publicly shared by 70.1% of the users) but not activity or friendship relations, except to Achievemint members.

[1]Under the following categories: Category 4–Research involving the collection or study of existing data, documents, records, pathological specimens, or diagnostic specimens, if these sources are publicly available or if the information is recorded by the investigator in such a manner that subjects cannot be identified, directly or through identifiers linked to the subjects.

[2]Members that unsubscribed had their data deleted and are not included in the analysis.

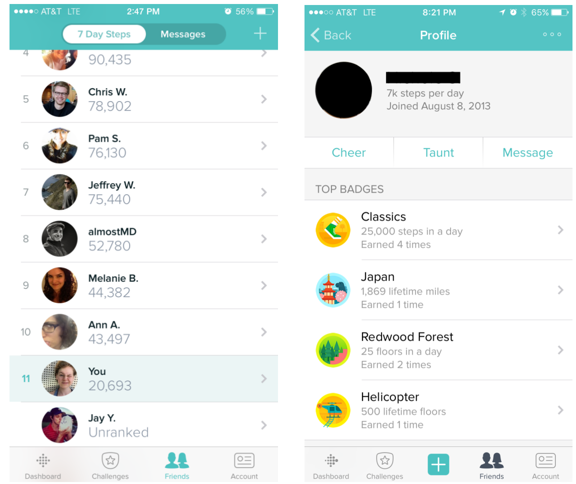

Figure 1: Social Engagement component of the Fitbit mobile application, from [13].

3.2 Social Interactions

It is important to analyze the details of how interaction with friends is presented to users in the Fitbit application, as it is well known that subtle changes in presentation and user interaction may be the true drivers of the effects under examination [26]. At the time of writing, in the Fitbit mobile application positive social influence is fostered by displaying the user's total step count over the last seven days ranked with that of their friends, as seen in Fig. 1. Users are also encouraged by being awarded trophies for achieving certain physical milestones as measured by their Fitbit devices.

The social component of the Fitbit application is not particularly prominent, as it does not appear on the initial landing page, or dashboard, of the application and it only displays the ego ranking seen in Fig. 1. The application allows users to "cheer", "taunt", and send text messages to their friends seen in their ego ranking.

3.3 Network Analysis Statistics

General statistics of the network and attributes of the Achievemint members are reported in Table 1. The union of all (partially observed) ego networks of the 44.5 thousand fully observed Achievemint members consists of a total of 891 thousand edges of which 23.8 thousand are between Achievemint members in the last month of observation.

Observation Period	9/30/2015-9/30/2016
# Achievemint Members	44,468
# Edges total	891,578
# Edges between Achievemint members	23,829
Average daily steps	8,498
Median age (IQR)	34 (28, 41) years
Female	82.8%
Underweight (BMI < 18.5)	1.1%
Overweight (25 <= BMI < 30)	29%
Obese (BMI >= 30)	42%

Table 1: Statistics of the dataset under analysis.

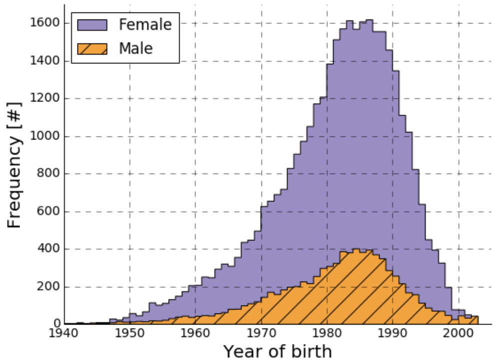

Figure 2: Distribution of the date of birth of the Achievemint members in the population considered.

More than 80% of Achievemint members are female; see Fig. 2 for a breakdown of year of birth by gender. We observe a power-law distribution of friendship ties, but with a marked difference by gender where female members tend to have a larger number of friends, as seen in Fig. 3.

Gender Homophily

We observe in our dataset signs of gender homophily, by which members tend to form ties with members of the same gender more frequently than would be expected in a random network with the same participants and characteristics. The effect is well-known in social network literature [37]. As seen in Fig. 4, we observe that the proportion of female ties for female members is higher than that of male members for most degree buckets.

To verify whether our observation of gender homophily is statistically significant we perform a null-model test. Our dataset presents various nuances that makes it difficult to apply an off-the-shelf test of homophily in social networks. In particular, we only observe edges from Achievemint members to other Fitbit users (whether Achievemint members or not). Achievemint membership status, gender and node degree can interact in various ways. To take into account these aspects we introduce a novel null-model and apply Monte Carlo simulations to establish an empirical p-value for our observation of same-gender preference in forming social ties. We first introduce the null-model, then we describe how to perform the simulations efficiently and finally we report our results.

Let $G = (V, E)$ be the graph representing our social network. For a member $u \in V$, let $a(u)$, $s(u)$, $d_a(u)$ and $d_n(u)$ be the Achievemint membership status of the member, the gender of the member, and the number of connections with Achievemint members and non-Achievemint members of member u in G. Notice that in G there are no edges between pairs of non-Achievemint members. Let $H = (V, E')$ be a multigraph (henceforth a graph possibly with parallel edges and allowing self-loops) over the same set of nodes V and let $d_a^H(u)$ and $d_n^H(u)$ be the number of connections of u in H to Achievemint members and non-Achievemint members, respectively—notice that $a(u)$ and $s(u)$ are not properties of the edge set E', hence they are the same in H.

We assume a null-model distribution χ that is a uniform distribution over all multigraphs $H = (V, E')$ respecting the properties aforementioned and such that: $\forall u \in V$, $d^H(u) = d(u)$, $d_n^H(u) = d_n(u)$, and such that in E' there are no edges between pairs of non-Achievemint members. First notice that this distribution is well-defined (it includes at least the graph G and many others in its support) and that it implies $|E'| = |E|$. For a (multi)graph $H = (V, E')$ so defined, let $S(H)$ be the number of edges between same-gender edges in E' (i.e. the edges between pairs of users of the same-gender counting parallel edges). Let χ' be the distribution of $S(H)$ for $H \sim \chi$. Now that we have defined this null-model, it is possible to compute a p-value for the probability of observing $s' \sim \chi'$ s.t. $s' \geq S(G)$ (i.e. more same-gender edges than observed in our dataset).

The distribution of χ' is quite involved, however it is possible to sample a multigraph $H \sim \chi$ from exactly the distribution χ (and hence compute $S(H)$) efficiently in $O(|V| + |E|)$ time using a variation of the techiniques in Brach et al. [6] for sampling multigraphs from the configuration model. This algorithm is complicated by the presence of the Achievemint membership status; a full discussion of the details is omitted due to lack of space. Notice however that defining this null-model to allow only simple graphs (instead of multigraphs) would be computationally expensive, as efficiently sampling simple graphs uniformly with a given arbitrary degree sequence is still an open research problem [15].

Using the null-model defined we observe that our dataset has a statistically significant gender homophily. We observe 70.1% same-gender edges vs 66.4% expected by the null-model with empirical p-value $< 10^{-3}$.

3.4 Health Questionnaires

Of the 44.5K members on the social network, 7,515 filled an online questionnaire asking the question "Have you been diagnosed with any of the following health conditions at any point in your life? Please select all that apply." 32 common conditions were included as multiple, non-exclusive choices, in addition to "none of the above" and "other, specify". In our analysis we select the most common conditions that are likely to have an impact on activity and social networks: hypertension (N=786), type 2 diabetes (N=257), dyslipidemia (N=229), and depression (N=1,665). General statistics of the sub-population that reported conditions are reported in Table 2.

4. STATIC ANALYSIS

4.1 Between-subject Analysis

In this section, we characterize the physical activity of members as a function of several features of their ego network. Both ego's activity and alter's features are computed as averages over the period of 9/30/15 to 9/30/16. Members have reported on average 209 days of activity during this 366-day period. This kind of analysis examines differences among users without considering their variation over time.

We regress average ego activity on age, BMI, and binarized gender (female=1) of the ego and those same attributes averaged across all alters in the ego network. We also include the size of the ego network (number of alters) and average activity for alters. Alter information is only available for alters that are Achievemint members, with the exception of age and gender for non-Achievemint members that shared them publicly on their Fitbit profile (70.1% of the total).

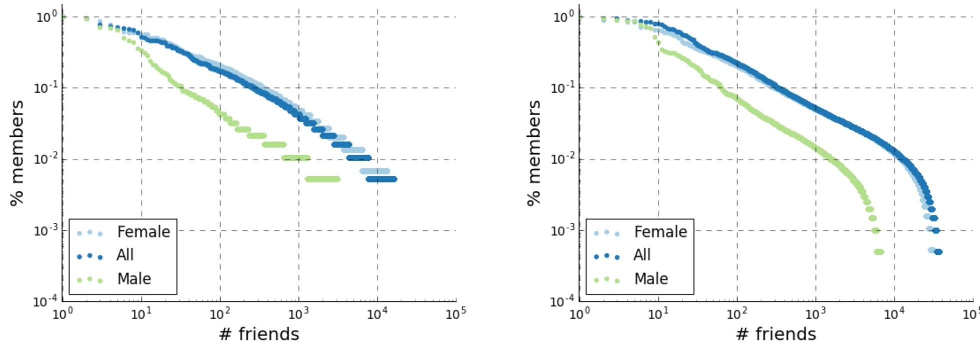

Figure 3: Complementary Cumulative Distribution Function (CCDF) of social network degree, measuring how many members have more than a given number of alters in the network, broken down by gender. On the left, only Achievemint members in each member's ego net are considered. On the right, non-Achievemint members are counted in the degree as well.

Group	Average daily steps	Median age (IQR)	% female	% underweight	% overweight	% obese
Depression (N=1,665)	8,423	35 (30, 41)	91	0.5	27	48
Dyslipidemia (N=229)	8,720	43 (36, 52)	80	0.4	26	56
Hypertension (N=786)	8,609	42 (36, 50)	77	0.1	23	66
Diabetes (N=257)	7,934	43 (37, 51)	80	0.0	16	77
Completed survey (N=7,515)	9,231	34 (29, 41)	84	1.1	30	39

Table 2: Statistics of most four common self-reported diagnosed health conditions, as answered by 7,515 members out of 44.5 thousand surveyed.

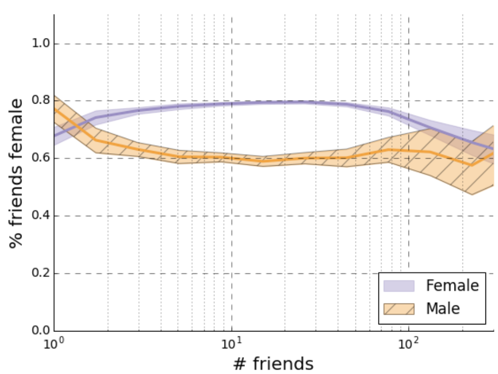

Figure 4: Mean proportion of female friends of a member given his/her total social ties. 95% confidence intervals estimated by bootstrap sampling 40% of the population 10,000 times.

Table 3 shows the coefficients and associated p-values for a linear regression. We see that all regressors are highly significant except for alter's average age. Some results are expected, such as that lower BMI correlates to higher activity and that men take more steps than women on average when controlling for everything else, an effect which has been reported previously [4]. Additionally, we see that users with highly active contacts tend to have higher activity themselves, consistent with [3]. Specifically, ego's activity increases on average by 26 steps for each additional 100 steps of average alter's activity. We also observe an average increase of 6.5 daily steps for each additional social tie.

Less intuitively, the coefficients for the average alter's BMI and gender each have a positive sign, with an additional alter average BMI point corresponding to an increase in ego's activity by 14.1 steps controlling for all other features. In both cases, the unadjusted coefficients (the effect we'd see without controlling for the other regressors) are the opposite sign which signals that these effects are correlated to the other regressors. Due to gender homophily, having more female ties correlates with the ego being female and thus lower average daily steps, but what these results tell us is that if we control for ego gender and the other features, having a higher rate of female ties is associated with higher activity. We find the same set of significant predictors when we repeat the analysis using logarithms of counts instead of absolute numbers.

4.1.1 *Interactions with Health Status*

We now consider the effects of hypertension, dyslipidemia, diabetes, and depression by including in the regression model interaction terms of each regressor with each medical condition. We do not report the full set of coefficients due space limitations, but we discuss significant coefficients of interaction terms with the medical conditions, as compared to the same coefficient for the no condition case.

Members with hypertension show a significant interaction with their condition, with ego BMI having a coefficient of -82.4 ($p = 0.001$). When compared to the BMI coefficient of -69.1 ($p < 10^{-10}$) for the no-condition case, the change highlights that keeping all other factors constant, users with hypertension see a decrease in activity per BMI point that

Covariate	Coefficient	Unadjusted	p-value
(Intercept)	8,762	8,895	$< 10^{-10}$
Ego's Gender	-1,590	-1,527	$< 10^{-10}$
Ego's Age	7.02	1.73	$< 10^{-6}$
Ego's BMI	-72.06	-70.21	$< 10^{-10}$
Average Alter's Daily Steps	0.26	0.28	$< 10^{-10}$
Average Alter's Age	0.07	1.03	0.9
Average Alter's Gender	356	-201	0.0009
Average Alter's BMI	14.1	-18.9	0.0001
Count of Alters	6.50	7.01	$< 10^{-10}$

Table 3: Standard and unadjusted coefficients and p-values for linear regression on ego average daily steps. Gender is defined as 1 for female members, 0 for male. Alter statistics are computed from an average over all alters of the ego. All covariates are included in the table.

is significantly larger than those without the condition. In the smaller dyslipidemia group, while the base correlation of average alter activity is 0.32 ($p < 10^{-10}$), the interaction term with the condition (-0.23, $p = 0.05$) cancels out 71.8% of the association. This result implies that the member's social network is not predictive of physical activity for users with dyslipidemia. A possible explanation could be that users with high cholesterol are more sophisticated about their activity level as they follow external guidelines, and thus tend to be less engaged with the social network and affected by its normative influence.

For users with diabetes and with depression, we see a large effect in the interaction term of the size of the ego network and the medical condition. The base effect of the size of the ego network on average daily steps is 6.2 ($p < 10^{-10}$) in the case of no condition, but the interaction term with diabetes and depression increases this effect by 31.1 steps ($p < 10^{-4}$) and 9.3 steps ($p < 10^{-5}$) per new tie added for users with these conditions, respectively. Given that users with diabetes have a similar number of social contacts as those without, (mean=36.8, SD=51.5) (mean=39.3, SD=86.9) respectively, this suggests that there is a stronger correlation between activity and social network size in users with diabetes (and to a lesser extent, depression). One possible explanation for the increased association observed may be due to an overlap between the offline disease management communities interacting with the member and the member's ego net.

5. NETWORK DYNAMICS

5.1 Within-subject Analysis

In this section, we model the association between a member's activity and several features of the ego network as they vary over time. Changes in the member's ego network are driven by the addition or removal of social ties (structural changes) and by the change in the observable characteristics of the alters.

To capture changes in the ego nets over time, the Fitbit profile information of each member was downloaded every three days between 2/3/2016 and 9/30/2016, as authorized by the member. Fig. 5 shows the growth of the network considered (when considering the graph induced on the Achievemint members only) over the 8-month period under analysis. The size of the network has significantly increased due to the growth in number of members of the Achievemint platform. When considering both Achievemint and non-Achievemint members in ego nets, we observe a total of 479 thousand new edges being added to the network in the eight-month period considered.

In the following analysis we use fixed-effects panel regression for controlling inter-user variability [11]. Panel regression removes all static confounders into a per-user fixed-effect term. Additionally we run a two-ways effects model by including a month-of-year indicator variable to account for seasonality effects that are very prominent in activity data [31].

We regress on ego BMI and average alter's activity, BMI, and membership (an indicator variable set to 1 if the alter is an Achievemint member, 0 otherwise) as well as the difference between ego and average alter age and gender. We use the difference between ego and average alter in lieu of demographic features of the ego, since time-invariant ego features are incorporated into the per-user fixed effect.

5.2 Results

Table 4 shows the coefficients and associated p-values from the panel regression. An immediate observation consistent with previous research [31] is that when considering only the ego's BMI (i.e., keeping social network features constant), fluctuations in BMI are significantly inversely correlated with ego activity over time, with an increase in activity corresponding to a decrease in BMI.

Next, we consider the effects of the social network itself by looking at changes in the size of the ego net. We see a significant impact of tie formation on ego activity. Consistent with what is reported in recent research [3], this association can be thought of as arising from a combination of a selection effect where users who increase their activity are more likely to add new friends (homophily), and causal effects where expanding one's social network increases activity due to social interactions (influence). The effect sizes of the associations surfaced are not directly comparable with those reported in [3] as in our case the effect of a tie formation is spread across several independent variables related to alters' features.

We also note that ego activity significantly increases when average alter membership to Achievemint decreases, which may suggest that adding "fresher" users (i.e., Fitbit users who are not yet Achievemint members) has a more positive association with ego activity. Finally, we consider the association between ego and alter physical activity over time. We see a significant, positive association between average friend activity and ego activity; members whose friends increase

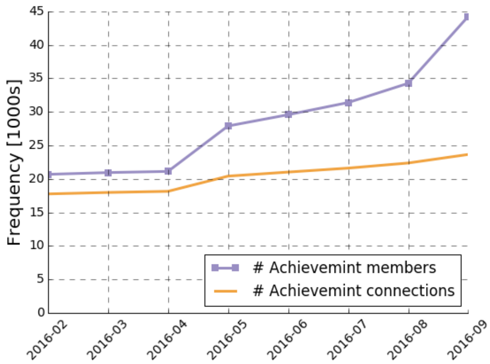

Figure 5: Number of Achievemint members and edges between them over the eight 30-day periods between 2/3/2016 and 9/30/2016.

their steps or who add more active friends see an increase in their own steps as well.

The fact that fluctuations in monthly steps are correlated among friends may be due to social influence and/or time varying confounders. Since seasonality has been removed at the population level by incorporating month in the regression, time-varying (non-stationary) causes of confounding may be ascribed to external phenomena selectively affecting activity of a subset of the network, such as local weather affecting friends in geographic proximity, or groups of friends joining group "challenges" offered by Fitbit. An in-depth discussion of causality observed on the social network is carried out in the following section.

Covariate	**Coefficient**	**p-value**
Ego's BMI	-39.5	$< 10^{-10}$
Count of Alters	0.7535	0.02
Average Alter's Daily Steps	0.065	$< 10^{-10}$
Average Alter's BMI	6.01	0.3
Average Alter's Membership	-1.66	$< 10^{-5}$
Difference in Age	-2.20	.7
Difference in Gender	-52.3	.7

Table 4: Coefficients and p-values for fixed effects panel regression on ego average daily steps. Alter statistics are averaged across all alters for each ego. Membership indicates if an alter is also a member of Achievemint(1=Achievemint member). Differences in age and gender are defined as the difference between the average alter's age/gender and the ego's age/gender respectively. Fixed effects are included on a per user and per month basis. All covariates are included in the table.

6. CAUSAL INFLUENCE

6.1 Method

In this section, we continue the analysis of the network over time to characterize the role of social influence (contagion) vs. homophily in driving correlated fluctuations in members' behavior over time. The ability to make inferences about causality relies heavily on what can be assumed at the

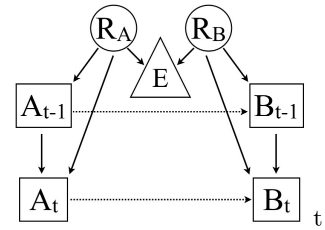

Figure 6: Solid lines reflect a graphical model of latent homophily where Alice's state A_t depends on her previous state and unobserved latent traits, R_A. The formation of an edge, E, between Alice and Bob depends on both of their latent traits. The dotted line alters the model to include contagion.

outset. The modern approach to causal inference relies on assumptions about conditional independence encoded in an intuitive way using graphical models [30]. These assumptions can be used to construct powerful causal tests using the d-separation criterion. D-separation essentially helps us to determine the conditional independence relationships implied by a set of graphical causal assumptions. Human behavior remains a challenging domain for causal inference, due to the difficulty of specifying all the dependencies affecting such a complex phenomenon. Especially challenging is the fact that in any study we measure only a tiny fraction of the variables that might be affecting human behavior.

The magnitude of the challenge posed by latent confounders in observational social network studies should not be underestimated. If there are any latent traits that affect the variate and affect the formation of ties in the social network, then this mechanism, which we will refer to as "latent homophily", can lead to effects mimicking contagion (influence) in the network [34]. The latent traits act as confounders for identifying contagion and cannot be removed without measuring all potentially relevant traits. Most studies attempt to measure and control for confounders, but it is difficult to be exhaustive when it comes to the intricacies of human behavior and this leaves many openings for latent homophily to intrude.

In this work, we consider a recently introduced approach to causal testing in the presence of unobserved confounders [36]. We allow for an arbitrary set of latent traits for each ego, and we allow that the effect of these traits on activity or on tie formation can be arbitrary. The only assumption we impose is that activity for each ego at time t is binary, $A_t \in \{0, 1\}$, and obeys a stationary Markov assumption: the probability of an action at time t depends only on the latent traits of that ego, R_A and actions at time $t-1$, A_{t-1}, and this dependence does not change over time. This model is shown in Fig. 6. While Shalizi & Thomas showed that this model is unidentifiable, i.e., the relative strength of contagion or latent homophily cannot be determined [34] (potential contagion is represented with a dotted line in the figure), the model is nevertheless *partially identifiable*. Partial identifiability means that we cannot exactly determine how much correlation is from homophily or influence, but in some cases we can rule out the possibility of correlations arising solely from latent homophily (with no contagion at all). Despite the generality of the latent homophily model and the existence of possibly infinite latent traits affecting activity, the distribution of observations compatible with this model is constrained. Using algebraic geometry, we can devise tests that tell us whether an observed probability distribution violates these constraints. If the constraints are violated, we can rule out the possibility that all observed correlations are due to latent homophily alone. We now summarize the considerations behind constructing these tests before applying them to our physical activity data.

For a pair of users, A, B, whose activity over time $t = 1, \ldots, T$, is captured by the latent homophily model described above, we write the probability distribution over their observed activity sequences, conditioned on the presence of a tie between them, as follows.

$$P(A_1, \ldots, A_T, B_1, \ldots, B_T|E) = \sum_{R_A, R_B} P(R_A, R_B|E) \prod_{t=1}^{T} P(A_t|A_{t-1}, R_A) P(B_t|B_{t-1}, R_B)$$

We sum over traits R_A and R_B for Alice and Bob respectively since these traits may be hidden. The stationary Markov assumption means that the transition probabilities, such as $P(A_t|A_{t-1}, R_A)$, do not change over time. We can represent the set of probability distributions that are compatible with this expression as vectors in a 2^{2T} dimensional space. The structure of this space is like a convex mixture (from the term $P(R_A, R_B|E)$) of points that are specified by polynomials in the parameters that characterize the probability of a transition, like $P(A_t = 1|A_{t-1} = 0, R_A)$. A convex set can be represented as an intersection of half-spaces, which are described by linear inequalities of the form,

$$\langle c(A, B) \rangle_{LH} \equiv \sum_{A,B} c(A, B) P(A, B|E) \leq 0, \forall P \in \mathcal{P}_{LH}, \quad (1)$$

where $\mathcal{P}_{LH}$ represents any distribution compatible with our latent homophily model. We are now ready to devise statistical tests that indicate if the observed probability distributions are compatible with the distributional constraints developed. It has been shown that we can find tests that describe a sequence of convex relaxations to the space $\mathcal{P}_{LH}$ by solving linear programming (LP) problems [36]. The tightness of these bounds is determined by the maximum degree of the polynomial representations used in the LP construction, d. The LPs give us guarantees of the form in Eq. 1. If we can find a test, $\langle c(A, B) \rangle_{LH} \leq 0$, while $\langle c(A, B) \rangle_{\hat{P}} > 0$ for an observed distribution, $\hat{P}$, we can rule out any latent homophily model as the sole mechanism generating this distribution. After putting our data into the appropriate form, we consider various tests of this form below and their interpretation.

6.2 Data Preparation

We use the framework described to model social influence between Achievemint members connected on the Fitbit network for activity level and BMI separately. To comply with the limitations of the framework described, we convert our dataset as follows. When looking at the possible influence of activity among friends, we binarize monthly changes in activity based on whether the change is greater or less than a threshold. This threshold is chosen as the median change for that month to remove global seasonality effects. We then randomly impute missing values as 0 or 1 to allow the data to be incorporated without bias.

For the BMI set, we take a similar approach by modeling changes in BMI as an increase or decrease relative to the median for that month. For this analysis we only consider BMI values that have been frequently updated, as we are interested in fluctuations over time and many BMI values are statically set in a profile with self-reported weight rather than by weight tracking over time.

In both the BMI and activity dataset we consider only the last 5 months of the 8 one-month periods available due to limitation in the computational feasibility of larger causality tests. We use BMI and ACT to refer to the BMI datasets and the activity datasets respectively.

6.3 Result

As a warm-up we consider the simplest tests that result from setting $d = 0$ in the convex relaxation of Eq. 1, and restrict our attention further to a single test, $c_I(A, B)$ which is optimized so it would be maximally violated by a simple synthetic model of contagion while $\langle c_I(A, B) \rangle_{LH} = 0$ for any latent homophily model. Code implementing this test is provided [2]. When we input the empirical distributions for BMI and ACT, we get $\langle c_I(A, B) \rangle_{\mathsf{BMI}} = 0.008$ and $\langle c_I(A, B) \rangle_{\mathsf{ACT}} = 0.017$. Although these both violate the equality, a question remains as to whether this is due to sampling error in estimating the empirical distribution. To determine this, we estimate the probability to see a violation this large under the null hypothesis that the true distribution satisfies the equality. This leads to p-values of $p_{\mathsf{ACT}} = 0.0014, p_{\mathsf{BMI}} = 0.147$. The first test gives evidence to rule out the null model that all correlations come from latent homophily. The second test should be regarded as inconclusive. Even without any violation, contagion may or may not be present. Also, it could be that this test was simply not powerful enough, and a different test will rule out the null model.

For a stronger test, we can look for violations using a tighter convex relaxation built from polynomials of maximum degree $d = 10$ in our LPs. For each dataset, we search for the inequality that is maximally violated. We find linear tests that require $\langle c(A, B) \rangle_{LH} \leq 0$ for all latent homophily models, while for our data $\langle c_{\mathsf{BMI}}(A, B) \rangle_{\mathsf{BMI}} = 0.179$ and $\langle c_{\mathsf{ACT}}(A, B) \rangle_{\mathsf{ACT}} = 0.126$. The p-values to achieve violations this large due to sampling error if the null hypothesis is true are calculated via Hoeffding's inequality giving $p < 10^{-10}$ in both cases (code available [2]). We conclude from this test that we can rule out the hypothesis that the dependence in our data was solely due to latent homophily, even with an arbitrary number of unobserved latent traits (under the assumptions of the model).

6.4 Limitations of the Analysis

Our result is the rejection of a null model. While the model rejected allows arbitrarily many unobserved latent traits, it still relies on assumptions, such as that of stationary Markovian dynamics. While contagion would provide one mechanism for violating the null model, we must consider the alternatives, discussed below.

Data processing artifacts

Our method required activity levels to take binary values over a fixed number of discrete time windows and does not allow for missing values. We made processing choices that are consistent with the Markovian assumptions of the null model, such as binarizing continuous data with respect to the median change in the population, to reduce seasonality effects. However, this may not be sufficient to account for all forms of seasonality. Also, effects may reflect mechanisms in the collection of data, rather than the underlying quantity measured. For instance, BMI is updated infrequently by users. We cannot exclude that it is our observation of the updates that is affected by influence instead of the actual BMI (e.g. a friend influences you to update your profile but your weight is independent of your friend's weight).

Markovian dynamics

The main assumption we relied on to derive these tests is stationary Markovian dynamics. Essentially, the present has to resemble the past for us to draw any meaningful conclusions from observations of time series. There are several plausible mechanisms for breaking this assumption including intra-city seasonality correlating with friendship structure and news that increases social structure mediated by the social network (i.e., Fitbit challenges). Using the same methods described above, we devised a second set of tests with the strengthened null hypothesis that data was generated according to the Markov model in Fig. 6 but this time including latent homophily *and* contagion (the dotted lines). Even in this case, we still see several significant violations, implying that latent homophily and contagion together are still insufficient to explain the data. This fact suggests that the violation of the Markov assumption may be partly responsible for rejecting the null model that allows only for latent homophily previously considered. In summary, our findings suggest that non-Markovian dynamic effects may constitute a barrier to parse out contagion from (latent) homophily in the network under analysis. The ability to break the dependency of physical activity from time-varying external confounders correlating with the social network structure, either through direct manipulation or by using natural experiments (e.g. [3]) is necessary to obtain a conclusive assessment of the importance of contagion driving changes in physical activity in the social network under analysis.

7. CONCLUSION

In this work we study the relationship between a social network and physical activity on a population of 44.5 thousand Fitbit users interacting through the Fitbit social network. We find that each additional social tie corresponds to an increase of 6.5 steps on average, and that users walk 26 additional steps per day for each additional 100 steps of average alter's daily activity. We confirm the directions and significance of the associations between ego and alter's activity over time by modeling the fluctuations of ego's activity via a panel regression with per-user fixed effects. We then qualify the strength of causal factors involved in the network dynamics by testing for violations of null models of increasing power. The causality analysis identified that, albeit non-causal models do not explain the data, the possibility that there exist external causes that drive correlated changes in ego's and alter's activity cannot be ruled out, even after correcting for global seasonality trends. Finally, we provide the first analysis of how the impact of a social network on physical activity interacts with the presence of a (self-reported) chronic condition on the user, revealing an enhanced positive association of alter's and ego's activity for users with diabetes (5.8-fold increase) and depression (2.4-fold increase).

Future work should focus on better understanding the causal factors driving changes in physical activity, and on quantifying how sustainable are such changes over time. Enabling long-term engagement and habit formation remains a challenge in the wearables space [27] and the sustained formation of new positive habits has proven hard to attain, even for widely adopted interventions [29]. From a data perspective, including finer-grained time-aggregate features (e.g., minute level step counts) could also be helpful to explain intra-day influence effects (e.g., by comparing counts of bouts of sedentary times [5]).

The causal analysis highlighted difficulties and potential pitfalls. Even when drivers of influence are identified, caution should be put in deriving general claims on how physical activity can be affected by manipulating the social network without validating it with randomized controlled experiments. A meta-study of 2040 studies found very modest evidence for effectiveness of interventions in online social networks [23], and an increasing amount of recent research warns that effects observed should not be immediately generalized outside the specific context in which they are observed [26].

We hope that the current study can inform comparative effectiveness research with the goal to understand which features and mechanisms of an online social network are the most successful at promoting healthy habits among their members, especially in the context of chronic disease management.

8. REFERENCES

[1] Let's move. http://www.letsmove.gov/, 2015. Accessed: 2015-07-20.

[2] Contagion Test Code. http://drop.isi.edu/sites/default/files/users/gregv/evidation.zip, 2016. Accessed: 2016-10-24.

[3] Tim Althoff, Pranav Jindal, and Jure Leskovec. Online actions with offline impact: How online social networks influence online and offline user behavior. In *Proceedings of the Tenth ACM International Conference on Web Search and Data Mining (to appear)*. ACM, 2017. Online appendix of this paper to be released with publication:http://bit.ly/2asuIeD Accessed: 2016-10-24.

[4] Mario Renato Azevedo, Cora Luiza Pavin Araújo, Felipe Fossati Reichert, Fernando Vinholes Siqueira, Marcelo Cozzensa da Silva, and Pedro Curi Hallal. Gender differences in leisure-time physical activity. *International journal of public health*, 52(1):8–15, 2007.

[5] Sangwon Bae, Anind K Dey, and Carissa A Low. Using passively collected sedentary behavior to predict hospital readmission. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, pages 616–621. ACM, 2016.

[6] Paweł Brach, Alessandro Epasto, Alessandro Panconesi, and Piotr Sankowski. Spreading rumours without the network. In *ACM Conference on Social Network (COSN'14)*. ACM, 2014.

[7] John Cawley and Chad Meyerhoefer. The medical care costs of obesity: an instrumental variables approach. *Journal of health economics*, 31(1):219–230, 2012.

[8] Justin Cheng, Lada Adamic, P Alex Dow, Jon Michael Kleinberg, and Jure Leskovec. Can cascades be predicted? In *Proceedings of the 23rd international conference on World wide web*, pages 925–936. ACM, 2014.

[9] Nicholas A Christakis and James H Fowler. The spread of obesity in a large social network over 32 years. *New England journal of medicine*, 357(4):370–379, 2007.

[10] Timothy S Church, Diana M Thomas, Catrine Tudor-Locke, Peter T Katzmarzyk, Conrad P Earnest, Ruben Q Rodarte, Corby K Martin, Steven N Blair, and Claude Bouchard. Trends over 5 decades in us occupation-related physical activity and their associations with obesity. *PloS one*, 6(5):e19657, 2011.

[11] Yves Croissant, Giovanni Millo, et al. Panel data econometrics in r: The plm package. *Journal of Statistical Software*, 27(2):1–43, 2008.

[12] Javid Ebrahimi, NhatHai Phan, Dejing Dou, Brigitte Piniewski, and David Kil. Characterizing physical activity in a health social network. In *Proceedings of the 6th International Conference on Digital Health Conference*, pages 123–129. ACM, 2016.

[13] Brodwin Erin. I tried fitbit for a month, and taking it off was the best decision i've made. http://finance.yahoo.com/news/tried-fitbit-month-taking-off-205114536.html, 2015. Accessed: 2016-10-24.

[14] Eric A Finkelstein, Justin G Trogdon, Joel W Cohen, and William Dietz. Annual medical spending attributable to obesity: payer-and service-specific estimates. *Health affairs*, 28(5):w822–w831, 2009.

[15] Pu Gao and Nicholas Wormald. Uniform generation of random regular graphs. In *Foundations of Computer Science (FOCS), 2015 IEEE 56th Annual Symposium on*, pages 1218–1230. IEEE, 2015.

[16] Leslie K John and Michael I Norton. Converging to the lowest common denominator in physical health. *Health Psychology*, 32(9):1023, 2013.

[17] David Kempe, Jon Kleinberg, and Éva Tardos. Maximizing the spread of influence through a social network. In *Proceedings of the ninth ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 137–146. ACM, 2003.

[18] Quantified Self Labs. The quantified self. www.quantifiedself.com, 2016. Accessed: 2016-05-20.

[19] Eva Leslie, Neville Owen, Jo Salmon, Adrian Bauman, James F Sallis, and Sing Kai Lo. Insufficiently active australian college students: perceived personal, social, and environmental influences. *Preventive medicine*, 28(1):20–27, 1999.

[20] Judith A Long, Erica C Jahnle, Diane M Richardson, George Loewenstein, and Kevin G Volpp. Peer mentoring and financial incentives to improve glucose control in african american veterans: a randomized trial. *Annals of internal medicine*, 156(6):416–424, 2012.

[21] Russell Lyons. The spread of evidence-poor medicine via flawed social-network analysis. *Statistics, Politics, and Policy*, 2(1), 2011.

[22] Xiaoxiao Ma, Guanling Chen, and Juntao Xiao. Analysis of an online health social network. In *Proceedings of the 1st ACM international health informatics symposium*, pages 297–306. ACM, 2010.

[23] Carol A Maher, Lucy K Lewis, Katia Ferrar, Simon Marshall, Ilse De Bourdeaudhuij, and Corneel Vandelanotte. Are health behavior change interventions that use online social networks effective? a systematic review. *Journal of medical Internet research*, 16(2):e40, 2014.

[24] Jennifer Marks, Kayla de la Haye, Lisa M Barnett, and Steven Allender. Friendship network characteristics are associated with physical activity and sedentary behavior in early adolescence. *PloS one*, 10(12):e0145344, 2015.

[25] Justin McCarthy. In u.s., adult obesity rate now at 27.7%, May 2014. [Online; posted 22-May-2014].

[26] Susan Michie, Rachel N Carey, Marie Johnston, Alexander J Rothman, Marijn de Bruin, Michael P Kelly, and Lauren E Connell. From theory-inspired to theory-based interventions: A protocol for developing and testing a methodology for linking behaviour change techniques to theoretical mechanisms of action. 2016.

[27] Patel MS, Asch DA, and Volpp KG. Wearable devices as facilitators, not drivers, of health behavior change. *JAMA*, 313(5):459–460, 2015.

[28] Cynthia L Ogden, Margaret D Carroll, Brian K Kit, and Katherine M Flegal. Prevalence of childhood and adult obesity in the united states, 2011-2012. *Jama*, 311(8):806–814, 2014.

[29] Leslie Oley. Can augmented reality alter reality? quantifying the pokémon go effect. http://bit.ly/2dF2G5X, 2016. Accessed: 2016-10-24.

[30] Judea Pearl. *Causality*. Cambridge University Press, 2009.

[31] Arya Pourzanjani, Tom Quisel, and Luca Foschini. Adherent use of digital health trackers is associated with weight loss. *PloS one*, 11(4):e0152504, 2016.

[32] Steven A Schroeder. We can do better—improving the health of the american people. *New England Journal of Medicine*, 357(12):1221–1228, 2007.

[33] P Wesley Schultz, Jessica M Nolan, Robert B Cialdini, Noah J Goldstein, and Vladas Griskevicius. The constructive, destructive, and reconstructive power of social norms. *Psychological science*, 18(5):429–434, 2007.

[34] Cosma Rohilla Shalizi and Andrew C Thomas. Homophily and contagion are generically confounded in observational social network studies. *Sociological methods & research*, 40(2):211–239, 2011.

[35] Thomas W Valente, Kayo Fujimoto, Chih-Ping Chou, and Donna Spruijt-Metz. Adolescent affiliations and adiposity: a social network analysis of friendships and obesity. *Journal of Adolescent Health*, 45(2):202–204, 2009.

[36] Greg Ver Steeg and Aram Galstyan. Statistical tests for contagion in observational social network studies. In *AISTATS*, pages 563–571, 2013.

[37] Yana Volkovich, David Laniado, Karolin E Kappler, and Andreas Kaltenbrunner. Gender patterns in a large online social network. In *International Conference on Social Informatics*, pages 139–150. Springer, 2014.