Digital Safety Nets: How Social Networks Shape Online Medical Crowdfunding Performance*

Xu Han Yiqing Xing Junjian Yi Haochen Zhang

Abstract

This paper investigates how social network structure influences online medical crowdfunding performance. Using administrative data from a leading medical crowdfunding platform in China, we exploit detailed records of campaign information diffusion through social media and construct novel measures that distinguish between "wide but near" versus "narrow but far" sharing network structures. We find that campaigns achieving broader network penetration—characterized by lower concentration of views from the fundraiser's immediate connections—raise significantly more funds. This relationship may operate through "reaching beyond homophily", a mechanism we formalize in a theoretical model: deeper network layers contain viewers with higher socioeconomic status and greater financial capacity, despite reduced altruistic motivation due to social distance. Furthermore, using exposure to China's Compulsory Schooling Law as an instrument, we show that network structure largely accounts for the causal effect of fundraisers' education on crowdfunding success. Our findings demonstrate how social capital structure shapes inequality in online crowdfunding, with broader implications for understanding the distributional consequences of social networks in the digital economy.

Keywords: Online crowdfunding; social capital; network structure; homophily; inequality
JEL Codes: D64; D85; D91; O12; O15

*We are grateful to XXX for their helpful comments and suggestions. All views expressed in this paper are those of the authors and do not represent the opinions of the Waterdrop Medical Crowdfunding platform. Contact authors: Xu Han: National School of Development, Peking University, Email: hanxu@stu.pku.edu.cn; Yiqing Xing: National School of Development, Peking University, Email: xingyq@nsd.pku.edu.cn; Junjian Yi: National School of Development, Peking University, Email: junjian.yi@gmail.com; Haochen Zhang: School of Economics, Zhejiang University, Email: zhanghaochen@zju.edu.cn

0 Contribution statement for writing sample

This paper serves as Xu Han's writing sample and is a modified version of a collaborative project with his co-authors Yiqing Xing, Junjian Yi, and Haochen Zhang. While this is a joint effort, Xu played a major role in the development of the project, making it suitable as his writing sample. Below is a summary of his specific contributions:

- *Research ideas.* Xu proposed the initial research idea when exploring the dataset, and developed the key measures for the study. Junjian played a crucial role in refining the idea and structuring the paper.

- *Model and Monte Carlo experiment.* Xu constructed the theoretical model and conducted the Monte Carlo experiment to validate the model. Yiqing provided invaluable guidance in refining the model and ensuring its robustness.

- *Empirical analysis and visualization.* Xu implemented all aspects of the empirical analysis, wrote the codes, and created figures and tables. Haochen provided essential guidance in refining the technical details of the empirical analysis and the presentation of the results.

- *Writing.* Xu wrote the initial draft of the paper and played a major role in revising and finalizing the manuscript.

Note: This manuscript is a draft version of the writing sample and requires further refinement and simplification.

1 Introduction

TBA.

2 Background

Medical crowdfunding has emerged as an important form of informal insurance, enabling individuals facing urgent medical expenses to seek public donations through online platforms (Sisco and Weber, 2019; Saleh et al., 2020). These platforms have gained global popularity in recent years, serving as bridges between those in medical need and potential donors while leveraging social media networks to amplify campaign reach and fundraising effectiveness.

The Waterdrop Medical Crowdfunding platform Our study utilizes administrative data from the Waterdrop Medical Crowdfunding platform, one of China's leading online medical crowdfunding platforms (Huang et al., 2021; Chan et al., 2024).[1] Established in 2016, the platform helps address critical gaps in China's healthcare coverage by enabling individuals facing high, unforeseen medical expenses—often arising from serious illnesses or surgeries—to seek public donations. By the end of 2023, approximately 450 million donors had collectively contributed 62.6 billion yuan (about 8.6 billion USD) to over 3.1 million beneficiaries. On December 25, 2024, China's Ministry of Civil Affairs officially designated Waterdrop Medical Crowdfunding as an approved online platform for personal fundraising assistance, underscoring its credibility and nationwide reach.

Platform integration with WeChat The platform's integration with WeChat—a multifunctional application developed by Tencent with over one billion active users—provides a unique setting for studying social network dynamics in medical crowdfunding. WeChat combines messaging, social media, and mobile payment services, making it one of the most dominant applications in China and globally. This integration creates a natural setting where

[1] See Tsai and Wang (2019) for an overview of charitable crowdfunding in China.

fundraising campaigns spread through social networks.

On the platform, fundraisers (beneficiaries) create campaigns by providing personal and medical details, including descriptions of their health conditions, treatment costs, and fundraising goals, supplemented by images and supporting documentation. After platform verification, fundraisers can distribute their campaigns through WeChat's social features: direct sharing with contacts and, more critically, dissemination via WeChat Moments (analogous to Facebook's News Feed) and Group Chats. The platform's low barriers to campaign creation (requiring only basic verification) and flexible fund withdrawal options (allowing multiple withdrawals aligned with treatment needs) enable help-seekers to quickly reach out to their networks for financial help during medical emergencies.[2]

The sharing network formation process When a campaign link is shared via Moments or Group Chats, a preview appears featuring the beneficiary's thumbnail image, fundraising title, and call-to-action prompt.[3] WeChat users who click the link become "viewers" tracked by the platform.

Upon arriving at a campaign page, viewers see the fundraiser's story, medical documentation,[4] fundraising goal, and current amount raised. Viewers can support campaigns in two ways: making donations through the integrated WeChat Pay system or sharing the campaign to their personal networks.[5] Sharing is particularly important as it amplifies campaign visibility by extending reach beyond the beneficiary's immediate contacts to secondary connections and even strangers. Notably, approximately one-third of those who share do so without making a donation, highlighting sharing as a form of costless yet consequential behavior. This process would repeat as new viewers in turn share the campaign to their own

[2]To sustain its operations, the platform charges a service fee of six percent on the total amount raised for each campaign, along with a payment processing fee of 0.6 percent.

[3]Common fundraising titles include phrases such as "Please help my 6-year-old daughter fight leukemia" or "Urgent: 45-year-old father requires heart surgery."

[4]The platform provides updates on beneficiaries' medical conditions, which enhances trust and transparency.

[5]The interface prominently displays "donating" and "sharing" buttons, making both forms of support easily accessible.

networks, creating multiple layers of diffusion.

This iterative sharing process creates what we term the "sharing network"—a tree-like diffusion structure that builds on the underlying social networks. Views generated directly from the fundraiser's sharing represent depth 1, while views accessed through one degree of separation (via sharing by depth-1 viewers) represent depth 2, and so forth. The platform records these depths, allowing us to observe precisely how far campaign information travels through online social networks.

3 Data

In this section, we describe the data source, sample, and variables for our empirical analysis.

3.1 Data source and sample

Our study utilizes unique administrative data from the Waterdrop Medical Crowdfunding platform.[6] The dataset contains a complete sample of all crowdfunding campaigns posted on the platform during a five-day period from November 20 to 24, 2023. This five-day window was selected to capture typical platform activity while avoiding potential confounding effects from holidays or special events. Our observation period covers six months following campaign launch. For each campaign, the data include complete viewing records for all viewers, regardless of whether they donated or shared the campaign. For each view, we observe the viewer's location, viewing timestamps, donation amount (if any), and sharing behavior. Crucially, the dataset precisely records the sharing network depth for each view—the number of degrees of separation between the original fundraiser and the viewer. Over 95 percent of views and donations occur within one month after campaign launch, indicating that our sample effectively captures the overall fundraising performance of these campaigns.

[6]We designed our data analysis plan and submitted it to the platform. All analyses were conducted in a secure lab within the platform. Data remain within the platform's premises and are not extracted. We only obtained the results of the analyses as requested.

Our primary analysis is conducted at the campaign level. From an initial universe of 4,971 campaigns posted during the sample period, we exclude those with no money raised,[7] and those with missing information on beneficiaries' baseline characteristics, such as age, gender, and the city of residence, which together account for 8.1% of the total. The final sample consists of 4,569 campaigns with beneficiaries located in 308 of China's 338 prefecture-level geographical units. Collectively, these campaigns raised approximately 52.65 million yuan (about 7.34 million USD).

Table 1 presents summary statistics for campaign characteristics. Approximately 36 percent of the campaigns feature female beneficiaries. Beneficiaries average 46.8 years of age, ranging from infancy to 92 years. Among the beneficiaries, 31.9 percent are diagnosed with cancer, 26.3 percent are covered by commercial insurance, and 89.3 percent have medical insurance coverage (either public or commercial). Fundraising targets range from 10,000 to 600,000 yuan. Regarding crowdfunding performance, campaigns receive an average of 4,142 views and 209 times of sharing (50 to Group Chats and 159 to WeChat Moments), raising an average of 11,523 yuan.[8] The completion rate—defined as the proportion of fundraising target achieved—ranges from 0 to 98 percent, with a mean of 5.8 percent. Figure 1 plots the distribution of total amounts raised across the campaigns in our sample, which shows substantial variations in fundraising success. Over 60 percent of campaigns raised less than 10,000 yuan.

3.2 Measures of sharing network structure

A unique feature of our dataset is the ability to observe the complete viewing records for each campaign, including the precise network depth of every view. This allows us to characterize the structure of the information sharing network for the fundraising process of each campaign.

[7] These likely represent campaigns that were created for technical tests but not actively utilized for fundraising purposes.

[8] To mitigate the impact of outliers, we apply winsorization to individual donations at the 99th percentile at the view level prior to aggregating the data to the campaign level. We further winsorize total fundraising amounts at the 99th percentile at the campaign level.

Figure 2 illustrates the distribution of sharing network depths for campaigns with different fundraising performance. We divide all campaigns into two groups based on their total fundraising amount: those above the median (successful campaigns) and those below the median (less successful campaigns). For each campaign, we calculate the share of total views occurring at each network depth—that is, the proportion of views from direct sharing by the fundraiser (depth 1), second-degree connections (depth 2), and so forth. We then average these proportions within each group and present them in Figure 2.

The figure reveals striking differences in sharing network structure between the two groups. Less successful campaigns (Panel A) show a highly concentrated depth distribution, with the vast majority of views occurring at depths 1 and 2—primarily from the fundraiser's direct connections and their immediate further connections. In contrast, campaigns with higher donation amounts raised (Panel B) exhibit a more dispersed depth distribution, with substantially lower proportions of views at depths 1 and 2, and greater reach into deeper network layers.

These patterns reveal two distinct types of sharing network structure: "wide but near" networks versus "narrow but far" networks. As illustrated in Figure 3, holding the total number of views constant, "narrow but far" networks are characterized by a higher proportion of views from deeper network layers (beyond the first degree of separation), while "wide but near" networks concentrate most views within the fundraiser's immediate social circle. The pattern observed in Figure 2 suggests that "narrow but far" networks tend to achieve better fundraising outcomes.

To quantitatively characterize these network structures, we develop three key measures that differentiate between "narrow and far" versus "wide but near" sharing networks:

(i) The proportion of views at depth 1 (λ^1) We begin with the most intuitive measurement: the proportion of views from the fundraiser's direct connections (at depth 1). This measurement captures the essential distinction between direct and indirect connections. A "wide but near" sharing network is characterized by views concentrated

at lower depths, indicating a larger proportion of views at depth 1. In contrast, a "narrow but far" network typically contains fewer views from direct connections and more from indirect ones. We denote the the proportion of views from the fundraiser's direct connections (at depth 1) as λ^1, defined as $\lambda^1 = \frac{\#\text{ views at depth 1}}{\#\text{ views in total}}$. A "wide but near" sharing network would exhibit a higher λ^1 as compared to a "narrow but far" one.

(ii) Herfindahl-Hirschman Index (HHI) To more precisely quantify the level of dispersion in the depth distribution, we employ the Herfindahl-Hirschman Index, commonly used to measure the degree of concentration in a distribution. The definition is $HHI = \sum_{i \in \mathbb{Z}^+} (\lambda^i)^2$. Since views with a depth greater than 10 are rare, we use λ^1 to λ^{10} to compute the HHI in practice. A smaller HHI indicates a more dispersed distribution. Hence, a "narrow but far" sharing network would exhibit a lower HHI, while a "wide but near" network would show a higher HHI, indicating that views are concentrated in just a few low depths. By the properties of the HHI, $0 < HHI \leq 1$.

(iii) Entropy Entropy is a measure commonly used to quantify the dispersion of a distribution. In our context, entropy captures how evenly views are distributed across different depths. Entropy of depths in the sharing network is given by $Entropy = -\sum_{i \in \mathbb{Z}^+} \lambda^i \log(\lambda^i)$.[9] Similar to the HHI, we use λ^1 to λ^{10} to compute entropy. Higher entropy indicates greater dispersion, and thus a "narrow but far" sharing network would exhibit higher entropy. By the properties of entropy, $Entropy \geq 0$.

The summary statistics and distributions of the three variables of network structure are presented in Panel B of Table 1 and Figure 4. As shown in the histograms, there are substantial variations across campaigns in all three measures. The proportion of views at depth 1 (λ^1) averages 24.3% but exhibits considerable heterogeneity, with most campaigns receiving

[9] We use the binary logarithm ($\log_2$) here to express entropy in bits. Both the HHI and entropy can be interpreted as a weighted average of some function of the proportion of views at each depth. The difference is that the HHI uses a linear function, while entropy employs a negative logarithmic function, which is more sensitive to the level of dispersion.

between 5% and 80% of their views from the fundraiser's direct connections. Notably, there is a significant peak at 100% in the distribution, indicating that a non-negligible portion of campaigns rely almost entirely on views from depth 1. The HHI distribution shows a similar pattern, with an average value of 0.253 and substantial variations ranging from highly concentrated networks (HHI close to 1) to more dispersed ones (HHI around 0.1). The entropy measure, averaging 2.258, also demonstrates significant cross-campaign variations, with most values falling between 0.5 and 3. Furthermore, Appendix Table A1 presents the pairwise correlations between the three measures of sharing network structure. As expected, all three measures are strongly correlated with each other, with correlation coefficients ranging from 0.808 to 0.899 in absolute value (all statistically significant at 1% level), suggesting that they capture the same underlying aspect of network structure dispersion from different perspectives.

4 Network structure and crowdfunding outcomes

We now formally examine the relationship between network structure measurements and crowdfunding outcomes. To this end, we estimate the following equation:

$$Y_j = \alpha + \beta Network_j + \mathbf{X}_j\gamma + \delta_{c(j)} + \varepsilon_j, \tag{1}$$

where j indexes individual campaigns, and $c(j)$ denotes the fundraiser's prefecture of residence. Y_j captures fundraising performance at the campaign level; our primary dependent variable is the logarithm of total donation amount raised. $Network_j$ represents our measurements of sharing network structure—specifically, the three variables constructed in Section 3.2: the proportion of views at depth 1 (λ^1), HHI, and entropy. $\mathbf{X}_j$ is a vector of campaign-level controls including beneficiary's age and gender, insurance coverage status, fundraising target, an indicator for cancer diagnosis, and the number of images and title/description characters. $\delta_{c(j)}$ represents fixed effects of fundraisers' prefectures based on the location of hukou registration, which absorb unobserved heterogeneity related to local social norms and

socioeconomic conditions. ε_j is the error term. Standard errors are clustered at the prefecture level. Our parameter of interest is β, which captures the association between sharing network structure and fundraising outcomes.

Main results We begin by examining the simple correlations between our network structure measures and medical crowdfunding performance. Figure 5 presents binscatter plots showing the relationships between log donation amounts and our three key network structure measurements. The patterns are consistent across all three measures and strongly support our conjectures on these relationship. These binscatter plots provide compelling visual evidence that "narrow but far" sharing networks—characterized by lower λ^1 (Panel A), lower HHI (Panel B), and higher entropy (Panel C)—are systematically associated with superior crowdfunding outcomes. Notably, the data points in all three panels are closely distributed around the fitted lines, indicating strong goodness of fit and suggesting that these relationships are not driven by outliers but represent systematic patterns across the entire sample.

Table 2 presents our main regression results examining the relationship between network structure and crowdfunding performance. In columns (1), (4), and (7), we include no control variables, and the regression results are consistent with the binscatter patterns, indicating significant correlations between network structure and fundraising performance. Notably, the R-squared values are remarkably high, suggesting that these network structure variables individually possess strong explanatory power for our dependent variable—λ^1, HHI, and entropy can each explain 38%, 49%, and 64% of the variation in total donation amounts, respectively. In columns (2), (5), and (8), we further include prefecture fixed effects and baseline campaign characteristics as controls. The estimated coefficients remain stable. To better distinguish the role of network structure from network scale effects, we additionally include total number of views, and the number of shares to Group Chats and WeChat Moments in columns (3), (6), and (9). After controlling for these factors, the coefficients on network structure variables decrease slightly but remain statistically significant at the 1% level. Given that total views and shares may be affected by sharing network structure, we

take the regressions in columns (2), (5), and (8) as our preferred specification. The estimated coefficients indicate that a one percentage point decrease in the proportion of views from depth 1 (λ^1) is associated with a 4.3% increase in total donation amounts. Similarly, a 0.1 standard deviation decrease in the HHI (a 0.1 standard deviation increase in entropy) corresponds to a 10.2% (10.3%) increase in fundraising amounts. These estimated coefficients demonstrate high economic significance.

Robustness Our main results remain robust across a battery of sensitivity analyses, as shown in Appendix Table A2. First, we use alternative dependent variables, replacing the log donation amount with total number of views in logarithm and fundraising completion progress in percentage points. Second, we construct the network structure variables based on the distribution of donation amounts rather than views across depths in the sharing network. Third, we re-estimate equation (1) by imposing further sample restrictions, to show that our main findings are not driven by some special subgroups of campaigns: (i) we exclude campaigns with highest and lowest 5% donation amounts, and those with highest and lowest 5% values of the network structure variables, in order to mitigate the impact of extreme values on our results; (ii) we exclude campaigns with beneficiary age below 18 or above 70, as for beneficiaries in these age groups, the fundraisers are typically not themselves but rather their immediate family members; (iii) we exclude campaigns with beneficiaries in coastal prefectures, in prefectures that are provincial capitals or municipalities, and beneficiaries with commercial insurance coverage, considering that these beneficiaries have on average better socioeconomic conditions and may have a lower reliance on the online crowdfunding network for medical expenses; (iv) we exclude campaigns with a fundraising target lower than 100,000 yuan, as these beneficiaries likely have less severe medical conditions and lower financial needs, resulting in reduced reliance on crowdfunding networks and potentially weaker effects of network structure.

5 Mechanisms

In our empirical analysis, we have demonstrated that "narrow but far" sharing networks are systematically associated with superior medical crowdfunding performance. However, the mechanisms underlying this relationship remain unclear. Why do sharing networks that reach deeper into online social connections yield better fundraising outcomes? What forces determine whether extending the reach of a campaign enhances or diminishes its success?

This section develops a theoretical framework to understand the mechanisms underlying the association between network structure and crowdfunding performance. While the scale effect of deeper network penetration constitutes an important driver of fundraising success, our empirical evidence indicates that the relationship persists after controlling for network scale, underscoring the distinct role of network structure.

We propose that such a relationship results from the tension between two countervailing forces that may operate simultaneously. The "pros" channel, which we term "reaching beyond homophily," suggests that deeper network penetration enables fundraisers to access viewers (potential donors) with higher socioeconomic status (SES) who are more capable of making substantial contributions. The "cons" channel, which we call "decaying altruism," posits that individuals become less willing to donate as their social distance from the fundraiser increases.

In the following, we first present our conceptual framework and show stylized facts related to each of the two forces. We then build a theoretical model to illustrate the pros channel as well as its tradeoff with the cons channel. While our theoretical framework shows the potential operation of both channels, our empirical findings reveal the dominance of the pros channel in determining the success of online medical crowdfunding, suggesting that the benefits of reaching potentially higher-SES viewers outweigh the costs of reduced altruism across online social distances.

5.1 Conceptual framework and stylized facts

Medical crowdfunding through online social networks creates an inherent tension between reach and social proximity. Fundraisers facing medical emergencies typically find themselves in severe financial distress, often representing lower socioeconomic strata of the society. Given the well-documented principle of homophily in social networks—that individuals tend to form connections with others who share similar characteristics (McPherson et al., 2001; Currarini et al., 2009; Jackson et al., 2017)—these fundraisers are more likely to be directly connected to individuals of similar SES. This creates a fundamental challenge for successful fundraising: those most accessible through direct social ties—who tend to exhibit highest altruism toward the beneficiaries—may be least capable of providing substantial financial assistance, while those more capable of helping may be socially distant and harder to reach.

The pros channel: Reaching beyond homophily This mechanism suggests that "narrow but far" sharing networks enable fundraisers to overcome the constraints of socioeconomic homophily. As information about a medical crowdfunding campaign diffuses through multiple layers of online social connections, it would progressively reach individuals who are more socioeconomically distant from the fundraiser. Due to the structure of homophilous networks, these more distant connections are increasingly likely to represent higher socioeconomic strata, thus possessing greater financial capacity to make donations. Theoretically speaking, this mechanism is closely related to the concept of *economic connectedness* developed by Chetty et al. (2022), which captures "the degree to which people with low and high SES are friends with each other." In our context, we extend this concept to encompass both direct and *indirect* connections through sharing networks. Fundraisers with higher indirect economic connectedness can leverage their social networks to reach beyond their immediate socioeconomic circle, accessing potential donors with greater financial resources.

Several patterns in our data support this mechanism. First, as shown in Appendix Figure A1, fundraisers in our sample are disproportionately located in regions with low levels of

GDP per capita, consistent with the fact that medical crowdfunding predominantly serves economically disadvantaged populations.[10] Second, Appendix Figure A2 reveals that viewers at greater depths (compared to depth 1) exhibit systematically different characteristics that suggest higher SES. Panel A shows that viewers at deeper network layers have higher historical donation amounts as compared to depth 1, and Panel B indicates they have made more donations in the past, suggesting their greater financial capacity, with possible additional effects from stronger prosocial inclinations.[11] Furthermore, deeper network layers are increasingly likely to include viewers from different provinces (Panel C) and prefectures (Panel D) from the fundraiser. This geographic diversification is particularly relevant in the Chinese context, where regional economic disparities are substantial (Zhang, 2021). Given that the fundraisers are predominantly located in economically disadvantaged regions (Appendix Figure A1), geographic distance often correlates with economic advantage, making geographically distant viewers more likely to be more financially capable to make substantial contributions.

The cons channel: Decaying altruism The opposing mechanism operates through the decay of altruistic motivation as social distance increases. Extensive research in behavioral economics and social psychology demonstrates that individuals exhibit stronger prosocial preferences toward those who are socially closer to them (Bernhard et al., 2006; Chen and Li, 2009; Enke, 2024). This tendency is particularly pronounced in societies with strong cultural emphasis on in-group favoritism, which is deeply embedded in China's kinship-based social value system (Enke, 2019). In the context of medical crowdfunding, this mechanism predicts that while deeper network layers may contain wealthier potential donors, these individuals

[10] Appendix Figure A1 presents the distribution of GDP per capita of the beneficiary's hometown prefecture across campaigns. We use the year 2019 in order to better capture regional levels of economic development in normal times not affected by the COVID-19 shock.

[11] A substantial literature documents that charitable giving exhibits strong income elasticity, with estimates typically ranging from 0.7 to 1.3, indicating that donation behavior is highly responsive to financial capacity (Andreoni, 2006; Vesterlund, 2006; Bakija and Heim, 2011). Given that medical crowdfunding typically requires substantial contributions to generate meaningful impact, financial capacity represents a more plausible explanation for the large differences in historical donation patterns across network layers than variation in altruistic preferences alone.

are less motivated to donate due to their greater social distance from the fundraiser.

Appendix Figure A3 provides evidence consistent with this channel. To isolate the change in potential donors' altruistic tendency across depths of viewing from the difference in viewer compositions across depths, we employ a within-viewer design, focusing on viewing records of potential donors who viewed more than one campaign in our sample period and controlling for both campaign and viewer fixed effects.[12] This approach effectively addresses confounding factors related to endogenous selection of which campaign to view based on social distances with fundraisers (i.e., depths). Compared to depth 1, viewers on average contribute significantly lower amounts at greater depths, suggesting a decrease in altruism with greater social distances.

5.2 Model and Monte Carlo experiment

In this subsection, we present a model of information diffusion to formally illustrate the "pros" channel—i.e., reaching beyond homophily—as outlined above. Consider a large enough community with n individuals, $N = \{1, 2, \cdots, n\}$, who are of different levels of SES. Following Chetty et al. (2022), we divide individuals into two groups: individuals with high SES (denoted by subset H), and those with relatively low SES (subset $L = N \setminus H$).

The underlying social network of the community is represented by a random network (N, p, q) with a homophilous structure, where $p \in (0, 1)$ denotes the probability that two individuals of the same group are connected, and $q \in (0, 1)$ for the probability of two individuals of different groups being connected. The assumption of homophily by SES indicates that $p > q$. Let $p \equiv a + e$ and $q \equiv a - e$, then $a = \frac{p+q}{2}$ captures the average density of the network, and $e = \frac{p-q}{2} > 0$ captures the (absolute) level of homophily.

Now we introduce the information diffusion process, i.e., how a fundraising campaign diffuses through online social connections and results in a sharing network. Without loss of generality, we set individual $1 \in L$ as the fundraiser, who wants to diffuse her campaign link

[12] This subsample contains 19,473,216 views from 9,028,190 viewers.

through the existing network. In addition, we assume that every individual viewing the link (including the fundraiser) shares it with each of their neighbors with probability β. In the diffusion process, let $\phi(k)$ denote the (conditional) probability that a viewer at depth k has high SES (i.e., belonging to H). Then we have the following proposition.

Proposition 1 *For $k \geq 1$, $\phi(k)$ is strictly increasing in k. Specifically, when $\#H = \#L$, $\phi(k) = \frac{1}{2}[1 - (\frac{e}{a})^k]$.*[13]

Proof. See Appendix B.1.1. ■

The proposition indicates that the likelihood of reaching a viewer of high SES increases with the depth. Notably, the case of $\#H = \#L$ suggests that the likelihood $\phi(k)$ is decreasing in both the absolute level of homophily, e, and the relative level of homophily, $r := \frac{e}{a}$, indicating that homophily hinders the low-SES fundraiser from reaching high-SES potential donors at any given layer. In sum, Proposition 1 lends support to our pros channel—reaching beyond homophily: since views at greater depths tend to come from high-SES individuals and these viewers donate larger amounts on average, a "narrow but far" sharing network would benefit from containing more views at greater depths.

Based on the setup of our model, we conduct a Monte Carlo experiment to further support the pros channel. In the experiment, we hold the total number of views constant and manipulate the depth distribution. We examine whether a more dispersed depth distribution is associated with more views from high-SES individuals. We corroborate the consistency of our result using different sets of parameters.

As in Section 3.2, we define the depth distribution of a sharing network as a vector $\mathbf{\Lambda} = (\lambda^1, \lambda^2, \cdots, \lambda^k, \cdots) \in \Delta(\mathbb{Z}^+)$, where $\lambda^k = \frac{\text{\# views at depth k}}{\text{\# views in total}}$. Consistent with our empirical measurements (Section 3.2), we focus on $\mathbf{\Lambda} = (\lambda^1, \lambda^2, \cdots, \lambda^{10})$ in this experiment.

We set the population size $\#N$ at 1,000, and specify four parameters, $\#H$, p, q, and the

[13] $\#H = \#L$ holds when the median of SES is used as the cutoff between high and low SES groups, as in Chetty et al. (2022).

total number of views, denoted by ν. Under each set of parameters, we examine five different depth distributions, $\mathbf{\Lambda}_1$ through $\mathbf{\Lambda}_5$, where

$$\begin{aligned}
\mathbf{\Lambda}_1 =& \frac{1}{100}(90, 10, 0, 0, 0, 0, 0, 0, 0, 0);\\
\mathbf{\Lambda}_2 =& \frac{1}{100}(65, 20, 10, 5, 0, 0, 0, 0, 0, 0);\\
\mathbf{\Lambda}_3 =& \frac{1}{100}(50, 30, 10, 5, 5, 0, 0, 0, 0, 0);\\
\mathbf{\Lambda}_4 =& \frac{1}{100}(35, 30, 10, 10, 5, 5, 5, 0, 0, 0);\\
\mathbf{\Lambda}_5 =& \frac{1}{100}(20, 30, 10, 10, 5, 5, 5, 5, 5, 5).
\end{aligned}$$

All five depth distributions are right-skewed, which is consistent with the regular skewness pattern in Figure 2. The HHI of the five distributions are 0.820, 0.475, 0.355, 0.240, and 0.165, respectively, with a higher HHI indicating a more concentrated depth distribution.

Under a given set of parameters $(\#H, p, q, \nu)$, we randomly generate 100 underlying social networks. For each social network, we then simulate 100 sharing networks following each of the five depth distributions, respectively. Finally, for each depth distribution, we compute the average proportion of views from high-SES individuals across 10,000 simulated sharing networks, denoted by θ_1 through θ_5. Appendix B.2 presents more details about the experiment. Table 3 shows the results with 18 different sets of parameters.

In Table 3, we find that, under all parameter sets exhibiting homophily ($p > q$), the average proportion of views from high-SES individuals is strictly increasing in the dispersion level of the depth distribution, i.e. $\theta_1 < \theta_2 < \ldots < \theta_5$, meaning that a "narrow but far" sharing network is likely to reach more high-SES viewers. In contrast, this does not hold without the homophilous structure (in rows 4 and 13 of Table 3, where $p = q$). In sum, this experiment further supports the pros channel by underscoring the crucial role of network homophily by SES and demonstrates the robustness of the channel across different sets of parameters.

5.3 Discussion on the role of generalized economic connectedness

Theoretically, the pros channel of reaching beyond homophily helps derive deeper economic implications of sharing network structure, associating our key measurements of network structure with an important dimension of social capital: economic connectedness (EC). By definition, EC is "the degree to which people with low and high SES are friends with each other" (Chetty et al., 2022). For an individual with low SES, their individual EC captures their ability to connect with high-SES individuals.

In what follows, we propose a generalized definition of EC and discuss how our key measurements of network structure as described in Section 3.2 capture this generalized EC. In a social network, a low-SES individual might access high-SES individuals through both direct and indirect connections. Chetty et al. (2022) primarily focus on the former. They measure EC as follows: a low-SES individual's EC is the share of high-SES *direct* friends among all her *direct* friends, further divided by the share of high-SES individuals in the population to quantify the degree of under-representation of high-SES friends. Using notations in our model, for a low-SES individual $i \in L$, we have

$$EC_i = \frac{1}{\#H/\#N} \times \frac{q\#H}{p\#L + q\#H}.$$

However, indirect connections also play a crucial role in the context of medical crowdfunding, as well as in other important scenarios such as job referrals, marketing, and knowledge diffusion. As emphasized by Jackson (2008), "much of the interest in networked relationships comes from the fact that individual nodes benefit (or suffer) from indirect relationships." In medical crowdfunding, viewers from indirect connections (at depths greater than 1) are more likely to have high SES, who are on average more likely to make substantial contributions compared with directly connected viewers (at depth 1). Similar logic applies to other contexts. For instance, in job referrals, job seekers may not know potential employers directly, and thus often connect with employers through intermediaries; in marketing, entrepreneurs in poor regions often lack direct ties to individuals with high purchasing power, yet may still

reach wealthier potential customers via indirect connections.

We generalize the concept of economic connectedness (EC) as the degree to which people with low and high SES are either directly or indirectly connected to each other through the sharing network. For a low-SES fundraiser, her EC at depth k is the share of high-SES depth-k viewers among all depth-k viewers divided by the share of high-SES people in the community. Formally, for a low-SES fundraiser $i \in L$, her depth-k EC is $EC_i^k = \frac{\phi(k)}{\#H/\#N}$. Then, we define her generalized EC (GEC) as the average EC across all depths, weighted by her ability to reach each depth as represented by the depth distribution $\mathbf{\Lambda}_i$:

$$\begin{aligned} GEC_i =& \sum_{k=1}^{\infty} \lambda_i^k \underbrace{\frac{\phi(k)}{\#H/\#N}}_{\text{Depth-k EC}} \\ =& \lambda_i^1 \underbrace{\frac{1}{\#H/\#N} \times \frac{q\#H}{p\#L + q\#H}}_{\text{Direct EC}} + (1-\lambda_i^1) \underbrace{\frac{1}{\#H/\#N} \sum_{k=2}^{\infty} \frac{\lambda_i^k}{1-\lambda_i^1} \phi(k)}_{\text{Indirect EC}}. \end{aligned} \tag{2}$$

This decomposition expresses GEC as the linear combination of direct EC and indirect EC, weighted by the proportion of views at depth 1.[14] It highlights three components of an individual's generalized EC (GEC): (a) Direct EC, reflecting the extent to which her immediate social circle includes high-SES individuals; (b) Indirect EC, capturing her exposure to high-SES individuals through indirect connections, which is related to the EC of her direct friends; (c) Depth distribution of her sharing network, representing her ability to reach deeper layers, which governs the relative weights between direct and indirect EC.

Hence, our measures of network structure capture the third component of GEC. By Proposition 1, a low-SES fundraiser's indirect EC is always greater than her direct EC. Consequently, a fundraiser with a more dispersed depth distribution exhibits a higher GEC by placing more weight on the indirect component. Intuitively, for individuals in less-developed regions, the capacity to reach beyond their immediate social circle through deep diffusion in the online network enables access to higher-quality social capital.

[14] Notably, the value of direct EC (i.e., depth-1 EC) coincides with EC in Chetty et al. (2022).

In the following, we establish a proposition (Proposition 2) that mathematically shows the relationship between our key measures of network structure and economic connectedness. We first define a *far-reaching transition* of a depth distribution used in Proposition 2.

Definition 1 *For an individual $i \in L$ whose sharing network has depth distribution $\mathbf{\Lambda}$, we define a far-reaching transition (j, k, ε) of $\mathbf{\Lambda}$ as follows, where $j, k \in \mathbb{Z}^+$ with $j < k$; $0 < \varepsilon < \lambda^j$.*

The transition shifts $\mathbf{\Lambda}$ to a more narrow-but-far depth distribution $\widetilde{\mathbf{\Lambda}}$ such that

(a) $\widetilde{\lambda}^j = \lambda^j - \varepsilon$, $\widetilde{\lambda}^k = \lambda^k + \varepsilon$;

(b) For any positive integer $\ell \neq j$ or k, $\widetilde{\lambda}^\ell = \lambda^\ell$.

Proposition 2 *For an individual $i \in L$ whose sharing network has depth distribution $\mathbf{\Lambda}$, we impose a far-reaching transition (j, k, ε) on $\mathbf{\Lambda}$, which yields a more narrow-but-far depth distribution $\widetilde{\mathbf{\Lambda}}$. Consequently, her* GEC, λ^1, HHI *and* Entropy *shift to* $\widetilde{\text{GEC}}$, $\widetilde{\lambda^1}$, $\widetilde{\text{HHI}}$, *and* $\widetilde{\text{Entropy}}$. *It follows that:*

(i) $\widetilde{\text{GEC}} > \text{GEC}$;

(ii) Suppose $j = 1$. Then $\widetilde{\lambda^1} < \lambda^1$;

(ii) Suppose $\varepsilon < \lambda^j - \lambda^k$. Then $\widetilde{\text{HHI}} < \text{HHI}$, *and* $\widetilde{\text{Entropy}} > \text{Entropy}$.

Proof. See Appendix B.1.2. ■

Proposition 2 suggests that, at the margin, GEC (as defined in equation (2)) moves in the opposite direction with λ^1 and HHI, and in the same direction with entropy. Iterating over multiple far-reaching transitions extends these relationships to more general cases. Thus, the variation across these measures might reflect underlying differences in the fundraiser's (generalized) economic connectedness, which in turn helps explain inequality in fundraising outcomes.

5.4 Extension: Tradeoff between pros and cons channels

The above discussion formally characterizes the "pros" channel—reaching beyond homophily—and delves into its implications. We now introduce the "cons" channel—decaying altruism—into the model to capture the tradeoff between two opposing forces, explore under what conditions the pros channel dominates, and discuss the linkage with our empirical findings.

We begin with a model of individual donation decision to examine how viewers' donation behavior varies with SES and sharing network depths (i.e., social distance to the beneficiary). We building our model on the canonical model of altruistic behavior (Andreoni, 1989, 1990, 2006) and specify a Cobb-Douglas utility function. We further assume that the viewer's altruistic level decreases exponentially with the depth—they care less about beneficiaries who are more socially distant. Given a certain campaign, a viewer j at depth k has utility:

$$U_j(d) = \log(w_j - d) + \alpha^k \log(D + d),$$

where d denotes her donation amount, w_j is her initial wealth, and D represents the wealth level of the fundraiser. The baseline altruistic level $\alpha \in (0, 1)$ is assumed to be constant.

The viewer maximizes $U_j(d)$ subject to the constraint $0 \leq d \leq w_j$. By the first order condition (FOC), the viewer's optimal donation amount is

$$d_j^* = \begin{cases} \frac{\alpha^k w_j - D}{1+\alpha^k} & \text{if } \alpha^k w_j - D \geq 0, \\ 0 & \text{if } \alpha^k w_j - D < 0. \end{cases}$$

Comparative statics analysis shows that the optimal donation d_j^* increases with her initial wealth, w_j, and decreases with her depth in the sharing network, k. The positive wealth effect explains the association between donations and individual SES, which corresponds to the pros channel. In our theoretical framework in Section 5.2, we incorporate this effect by assuming that $w_j = \begin{cases} w_H & \text{if } j \in H \\ w_L & \text{if } j \in L \end{cases}$, where $w_H > w_L > 0$. In addition, the decrease in donations at higher depths corresponds to the cons channel of decaying altruism.

Then, we examine crowdfunding success at the campaign level. For a fundraiser i whose sharing network has a depth distribution $\mathbf{\Lambda}_i$ and total number of views ν, her expected total donation amount is

$$\mathscr{D}_i = \nu \sum_{k=1}^{\infty} \lambda_i^k [\phi(k) \frac{\alpha^k w_H - D}{1+\alpha^k} \mathbf{1}\{\alpha^k w_H - D \geq 0\} + (1-\phi(k)) \frac{\alpha^k w_L - D}{1+\alpha^k} \mathbf{1}\{\alpha^k w_L - D \geq 0\}].$$

In order to compare different depth distributions, we continue to consider a far-reaching transition (j, k, ε) of $\mathbf{\Lambda}_i$. This corresponds to a small shift toward a distribution that is more "narrow but far", which is denoted by $\widetilde{\mathbf{\Lambda}}_i$. Consequently, the expected total donation amount shifts to $\widetilde{\mathscr{D}}_i$. If $\widetilde{\mathscr{D}}_i > \mathscr{D}_i$ holds, then a sharing network that is more "narrow-but-far" tends to generate more donations.

Based on our conceptual framework, this requires the pros channel to dominate the cons channel. Intuitively, the tradeoff between two channels is governed by two key parameters: the relative level of homophily $r = \frac{e}{a}$, and the level of altruism α.[15] Since reaching beyond homophily becomes more valuable in highly homophilous networks, the strength of the pros channel increases with r. And since $-\ln \alpha$ is the decay rate of altruism with depth k, the strength of the cons channel declines as the altruistic level α rises. Hence, when both r and α are large enough, the pros channel tends to outweigh the cons channel, implying that a more "narrow-but-far" sharing network will raise more donations. Appendix B.3 illustrates this mechanism by examining the case when $j = 1$ and $k = 2$, thereby exemplifying the intuition and further clarifying the tradeoff between the two opposing channels.

Accordingly, our empirical findings are consistent with a high level of overall SES-based homophily and a relatively low level of decrease in altruism with depth. Moreover, discussion in Appendix B.3 suggests that higher level of wealth inequality (a higher $\frac{w_H}{w_L}$) may also contribute to the strength of the pros channel, i.e., amplifying the gains from reaching beyond homophily. Taken together, these insights highlight the background conditions specific to the medical crowdfunding context that shape our empirical results, while also providing a broader perspective on how the tradeoff between the two channels may shift under different circumstances.

[15] r is related to $\phi(k)$ in the expression of $\mathscr{D}_i$. By Proposition 1, if we assume that $\#H = \#L$, then $\phi(\cdot)$ is uniquely determined by r.

6 Education and inequality in medical crowdfunding: The role of network structure

The empirical and theoretical analyses presented in previous sections demonstrate that network structure plays a crucial role in determining medical crowdfunding success. Specifically, "narrow but far" sharing networks, characterized by more dispersed network depth distributions, are associated with significantly better fundraising outcomes. This relationship appears to operate primarily through the channel of reaching beyond homophily, whereby fundraisers with higher economic connectedness through the sharing network can access donors with potentially higher socioeconomic status despite larger social distances.

These findings may offer important insights into the mechanisms underlying well-documented disparities in crowdfunding success. Previous studies have identified several key determinants of medical crowdfunding performance, including fundraiser characteristics such as age, education, race, and geographic location (van Duynhoven et al., 2019; Kenworthy et al., 2020; Zheng et al., 2023; Chen, 2025; Zhang et al., 2025). However, the channels through which these characteristics influence fundraising outcomes remain poorly understood. Our framework suggests that network structure may serve as a critical mechanism, explaining how individual characteristics translate into differential crowdfunding outcomes.

To explore such a role of network structure empirically, we focus on education—one of the most robust predictors of crowdfunding success as identified in prior studies. Educational attainment is strongly associated with social capital accumulation (Glaeser et al., 2002), making it an ideal case for examining how individual characteristics shape network structure and, ultimately, influence fundraising outcomes. To establish the causal effect of educational attainment, we employ the implementation of the Compulsory Education Law (CSL) in China to construct an instrumental variable (IV). We then investigate whether network structure can statistically account for the causal relationship between fundraisers' education and crowdfunding performance, thereby revealing the extent to which social networks serve

as a key driver of education-based disparities in crowdfunding success.

6.1 Empirical strategy

We first estimate the effect of fundraisers' education on crowdfunding outcomes by exploiting the implementation of the CSL in China as a quasi-experiment. We then examine the role of network structure variables in explaining such an effect statistically.

Measuring fundraisers' education Since our original dataset does not contain direct information on fundraisers' educational attainment, we construct a measure of education using microdata of Chinese population censuses in 2010 and 2015.[16] Specifically, using the census data, we calculate average years of schooling for each demographic cell defined by birth year, gender, and prefecture of hukou registration, and then assign these values to fundraisers based on their birth year, gender, and prefecture of hukou registration.

The cell average years of schooling exhibit strong predictive power for individual educational outcomes, with a correlation coefficient of 0.66 in the census data, which lends support to the validity of our imputation method. In terms of the Chinese context, educational attainment exhibits substantial variations across cohorts and regions, as driven by major policy shocks—such as the Cultural Revolution's disruption of schooling and subsequent educational expansion policies—as well as substantial disparities educational resources by gender and across regions (Zhang and Kanbur, 2009; Li et al., 2017; Du et al., 2021), which suggests that cohort-gender-prefecture cell average levels of education effectively captures local educational opportunities available to difference demographic groups.

[16]Since 1990, China has conducted a census every 10 years and an inter-census population survey (also called a "mini-census") at the midpoint year between two censuses, with a sampling fraction of 1%. We use a pooled sample derived from (mini-)census in 2010 and 2015—a 0.35% sample of census 2010 and 15% of mini-census 2015. Pooling microdata of the two censuses helps increase the sample size and reduce the measurement error when imputing education using average levels in cohort-prefecture-gender cells. In addition, in the analysis of education effects, we exclude crowdfunding campaigns involving beneficiaries younger than 15 years (who have typically not completed compulsory education), because for these young beneficiaries, it is more likely the education level of their parents rather than themselves that influences social capital and fundraising outcomes. The final sample for the analysis involving education contains 4,131 campaigns.

Estimating the effect of education We estimate the following equation:

$$Y_{jcg} = \beta Edu_{jcg} + \mathbf{X}_j\gamma + \delta_c + \mu_g + \varepsilon_{jcg}, \tag{3}$$

where j denotes campaigns, and c and g stand for the prefecture and birth year of the fundraiser, respectively. $\mathbf{X}_j$ represents a set of campaign-level covariates, defined in the same way as in equation (1). δ_c and μ_g represent prefecture and cohort fixed effects, respectively. The key explanatory variable in this equation is Edu_{jcg}, which is the defined as average years of schooling in the gender-prefecture-cohort cell for the fundraiser of campaign j as constructed above using census data. While this imputation approach introduces classical measurement error that may attenuate our estimate of β in equation (3), it does not compromise the validity of our instrumental variable strategy based on the CSL reform in China that we introduce below, which creates quasi-experimental variations in education across cohorts and regions. ε_{jcg} is the error term. Standard errors are clustered at the prefecture-cohort level.

To address potential endogeneity concerns regarding Edu_{jcg} in equation (3), we exploit exogenous variation in educational attainment generated by the Compulsory Schooling Law (CSL) reform. The nationwide CSL in China was passed on April 12, 1986, and officially took effect on July 1, 1986. This landmark legislation mandated 9 years of compulsory education for all children aged 6 and above, making education free of charge and prohibiting the employment of school-age children. A key feature of the CSL implementation was that provinces were allowed to have different effective dates due to varying resource constraints and development levels. Most provinces implemented the law between 1986 and 1987, while some provinces in less developed regions implemented it as late as the early 1990s. This staggered implementation creates temporal variation in exposure to the reform across provinces.

Following the literature (Du et al., 2021; Huang et al., 2025), we construct our instrumental variable (IV) of individuals' exposure to the CSL reform, $Exposure_{cg}$. The variable is constructed based on one's age at the reform, which is determined by their birth year and the effective year of the CSL in their province of hukou registration. The CSL exposure

measure equals 1 for individuals fully eligible for compulsory education (aged 6 or below when the law took effect), 0 for those completely ineligible (aged 15 or above), and varies linearly for those partially affected. We use this variable of CSL exposure as the instrument for Edu_{jcg} in equation (3).

Validity of the CSL instrument using an event-study approach To provide evidence in support of the validity of our IV approach, we examine the dynamic effects of CSL exposure using an event-study specification. This approach allows us to assess whether the reform's impact on education and crowdfunding outcomes follows the expected pattern based on individuals' CSL eligibility status. Note that our IV approach is essentially akin to a cohort difference-in-differences (DID) setting, and the event-study specification helps speak to the parallel trend assumption required for the DID identification.

Figure 6 presents the event-study results for both schooling years (Panel A) and log donation amount (Panel B). The horizontal axis represents individuals' age (in 3-year bins) when the CSL became effective in their province of birth, with negative values indicating younger ages. The two vertical dashed lines divide cohorts into three distinct groups: individuals who were completely ineligible for the reform (aged 16 or above, to the left of the left line), those who were partially eligible (aged 7-15, between the two lines), and those who were fully eligible (aged 6 or below, to the right of the right line). In the event-study specification, those who were aged 16-18 when the CSL was implemented are treated as the omitted group. In Panel A, we find that for ineligible cohorts, the coefficients are indistinguishable from zero, which aligns with the absence of differential pre-trends and lends support for our identification. For partially eligible cohorts, the coefficients show a smooth upward trend, indicating larger of longer exposure to the CSL. The trend in coefficients becomes relatively flat for completely eligible cohorts. This pattern aligns precisely with the theoretical expectations based on the eligibility status, lending credibility to our instrumental variable approach. Furthermore, in Panel B, we observe a highly similar pattern when using log fundraising amount as the dependent variable, suggesting that the impact of the

CLS exposure is also less likely to be confounded by unobserved factors that systematically influence the fundraising performance.

Exploring the role of network structure To examine whether network structure serves as a mechanism through which education affects crowdfunding performance, we perform our analysis in two steps. First, we investigate whether individuals with higher educational attainment indeed exhibit network structure characteristics that are more conducive to successful fundraising—namely, "narrow but far" sharing networks—as documented in our previous analyses. To this end, we estimate equation (3) with our measures of network structure—λ^1, HHI, entropy—as dependent variables. If education facilitates access to social capital that enables individuals to reach beyond their immediate social circles, we expect higher education to be associated with lower λ^1 and HHI, and higher entropy.

Second, we assess the extent to which network structure can statistically account for the causal relationship between education and crowdfunding outcomes. We incorporate the network structure variables as additional controls in the regression of total fundraising amount on educational attainment. If network structure serves as a key mechanism underlying the impact of education, we would expect the coefficient on education to become substantially smaller in magnitude and potentially statistically insignificant when controlling for these network measures.

6.2 Results and discussion

We now discuss our estimation results about the relationship among education, network structure, and crowdfunding performance. We first show how education shapes network structure and influences fundraising outcomes, and then investigate the role of network structure underlying the education effect on fundraising success.

The effect of education on network structure and fundraising success Table 4 presents our results based on equation (3). To begin with, Panel A reports the relationship

between education and network structure attributes. For each dependent variable, the first column shows the OLS estimates, while the second column shows the 2SLS estimates using the CSL exposure IV. The OLS results show that an increase in schooling years is associated with a decreased proportion of views from depth 1 (λ^1) and HHI, and an increased value of entropy. Figure 7 presents binscatter plots illustrating the correlations between our network structure measures and schooling years, which reveal linear relationships consistent with our regression specification.

The IV estimations confirm a considerable effect of education. The OLS results underestimate the impact of schooling years in absolute value, suggesting that some uncontrolled factors in the error term are positively correlated with Edu_{jcg} and hence dampens the magnitude of β in equation (3) in the OLS estimation. This result is probably due to a measurement error in Edu_{jcg} (potentially because of the imputation method), or an omitted variable bias. For example, areas with stronger Confucius cultural tradition may invest more in formal education in response to economic development while maintaining dense kinship-based social structures with high social network concentration.

Taking the IV estimates in columns (2), (4), and (6) as the baseline, an additional year of schooling would induce a 5.037 percentage point (21 percent relative to the sample mean) decrease in the proportion of views from depth 1, a lower HHI by 0.039 (15 pecent), and an increase in the value of entropy by 0.122 (5 percent). These effects are of high economic significance. Additionally, in the IV regressions, the first stage F-statistic is well above 10, alleviating the concern about weak instruments.

We then delve into the effect of education on medical fundraising performance. In Table 4, Panels B and C present the OLS and IV estimates in column (1), respectively. The OLS result shows a positive and statistically significant association between schooling years and log total donation amount. The IV estimate is larger in magnitude compared to the OLS, potentially due to the measurement error in the imputed schooling years variable. In the IV estimation, an additional year of schooling is associated with an approximately 19.0 percent

($\approx \exp(0.174) - 1$) increase in the total amount raised. This finding aligns with prior studies that education serves as a key determinant of medical fundraising success (Zheng et al., 2023; Chen, 2025).

The role of network structure in explaining the education effect Most importantly, our results demonstrate that network structure variables possess remarkable statistical explanatory power for the education effect on crowdfunding outcomes. When we include the proportion of views from depth 1 (λ^1) as an additional control in column (2), the coefficient on education drops dramatically from 0.174 to -0.052 in the IV estimation and becomes statistically insignificant. Similarly, when controlling for the HHI in column (3), the education coefficient becomes -0.071 and loses statistical significance. When entropy is included in column (4), the education coefficient reduces to -0.018, again statistically indistinguishable from zero. These results provide compelling evidence that network structure appears to serve as a primary channel through which education affects crowdfunding success. The substantial reduction in both the magnitude and statistical significance of the coefficient on schooling years suggests that our network structure measures capture much of the mechanism underlying the impact of education on fundraising outcomes. Additionally, the OLS results in columns (2)–(4) of Panel B reveal similar patterns: the inclusion of network structure variables make the coefficient on education become much smaller in magnitude and statistically insignificant. Figure 8 illustrates the role of network structure using binned scatterplots.

Network structure versus network scale To further confirm that our findings reflect the role of network structure attributes rather than their correlations with the network scale, we further examine whether other network-related variables can similarly account for the education effect. In columns (5) through (8) of both Panels B and C, we control for total number of views, total shares to Group Chats, and total shares to WeChat Moments—variables that capture the network scale and intensity of sharing activity but not the structural characteristics of the sharing network.

The results reveal a striking contrast. Unlike our network structure measures, these variables related to the network scale fail to substantially diminish the effect of education. In the OLS results (Panel B), the inclusion of these network scale measures merely changes the estimated coefficient on schooling years, in terms of either coefficient size or statistical significance. The IV estimations in Panel C reveal consistent patterns. When controlling for total views in column (5), the education coefficient remains at 0.139 and stays statistically significant at the 10% level. When controlling for shares to group chats in column (6), the coefficient is 0.109, and when controlling for shares to WeChat Moments in column (7), it remains 0.146 and maintains statistical significance. Even when controlling for all three network scale variables simultaneously in column (8), the education coefficient is 0.131 and remains statistically significant.

The consistent patterns across both OLS and IV estimations demonstrate that the explanatory power we observe is specifically attributable to network structure characteristics rather than general network activity levels. While network scale variables such as total views and shares are certainly important for fundraising success (as evidenced by their positive and statistically significant coefficients), they do not appear to be the primary mechanism through which education influences crowdfunding outcomes. Instead, it is the specific structural property of the sharing network—how far the information flow can reach along the online social network—that accounts for the impact of education on fundraising performance. These findings underscore the importance of social capital "quality" over "quantity" in medical crowdfunding contexts.

7 Conclusions

TBA.

References

Andreoni, James, "Giving with Impure Altruism: Applications to Charity and Ricardian Equivalence," *Journal of Political Economy*, 1989, *97* (6), 1447–1458.

_ , "Impure Altruism and Donations to Public Goods: A Theory of Warm-Glow Giving," *Economic Journal*, 1990, *100* (401), 464–477.

_ , "Philanthropy," *Handbook of the Economics of Giving, Altruism and Reciprocity*, 2006, *2*, 1201–1269.

Bakija, Jon and Bradley T Heim, "How does charitable giving respond to incentives and income? New estimates from panel data," *National Tax Journal*, 2011, *64* (2), 615–650.

Bernhard, Helen, Urs Fischbacher, and Ernst Fehr, "Parochial altruism in humans," *Nature*, 2006, *442* (7105), 912–915.

Chan, Tat Y., Li Liao, Xiumin Martin, and Zhengwei Wang, "Avoiding peer information and its effects on charity crowdfunding: A field experiment," *Management Science*, 2024, *70* (4), 2272–2293.

Chen, Junhao, "Is deservingness merit-based or need-based? College selectivity and medical crowdfunding outcomes," 2025. Working paper.

Chen, Yan and Sherry Xin Li, "Group identity and social preferences," *American Economic Review*, 2009, *99* (1), 431–457.

Chetty, Raj, Matthew O. Jackson, Theresa Kuchler, Johannes Stroebel, Nathaniel Hendren, Robert B. Fluegge, Sara Gong, Federico Gonzalez, Armelle Grondin, Matthew Jacob, Drew Johnston, Martin Koenen, Eduardo Laguna-Muggenburg, Florian Mudekereza, Tom Rutter, Nicolaj Thor, Wilbur Townsend, Ruby Zhang, Mike Bailey, Pablo Barberá, Monica Bhole, and Nils Wernerfelt, "Social capital I: Measurement and associations with economic mobility," *Nature*, 2022, *608* (7921), 108–121.

Currarini, Sergio, Matthew O. Jackson, and Paolo Pin, "An economic model of

friendship: Homophily, minorities, and segregation," *Econometrica*, 2009, *77* (4), 1003–1045.

Du, Huichao, Yun Xiao, and Liqiu Zhao, "Education and gender role attitudes," *Journal of Population Economics*, 2021, *34* (2), 475–513.

Enke, Benjamin, "Kinship, cooperation, and the evolution of moral systems," *Quarterly Journal of Economics*, 2019, *134* (2), 953–1019.

_ , "Moral boundaries," *Annual Review of Economics*, 2024, *16* (1), 133–157.

Glaeser, Edward L., David Laibson, and Bruce Sacerdote, "An economic approach to social capital," *Economic Journal*, 2002, *112* (483), F437–F458.

Huang, Junjie, Huawei Shen, Qi Cao, Li Cai, and Xueqi Cheng, "How medical crowdfunding helps people? A large-scale case study on the Waterdrop fundraising," *Proceedings of the International AAAI Conference on Web and Social Media*, 2021, *15*, 220–229.

Huang, Wei, Xiaoyan Lei, Guangjun Shen, and Ang Sun, "Beyond nature and nurture: The impact of China's Compulsory Schooling Law on selection against high-risk fetuses," *Journal of Human Resources*, 2025, *forthcoming.*

Jackson, Matthew O., *Social and Economic Networks*, Princeton University Press, 2008.

_ **, Brian W Rogers, and Yves Zenou**, "The economic consequences of social-network structure," *Journal of Economic Literature*, 2017, *55* (1), 49–95.

Kenworthy, Nora, Zhihang Dong, Anne Montgomery, Emily Fuller, and Lauren Berliner, "A cross-sectional study of social inequities in medical crowdfunding campaigns in the United States," *PLOS One*, 2020, *15* (3), e0229760.

Li, Hongbin, Prashant Loyalka, Scott Rozelle, and Binzhen Wu, "Human capital and China's future growth," *Journal of Economic Perspectives*, 2017, *31* (1), 25–48.

McPherson, Miller, Lynn Smith-Lovin, and James M. Cook, "Birds of a feather: Homophily in social networks," *Annual Review of Sociology*, 2001, *27* (1), 415–444.

Saleh, Sameh N., Ezimamaka Ajufo, Christoph U. Lehmann, and Richard J.

Medford, "A comparison of online medical crowdfunding in Canada, the UK, and the US," *JAMA Network Open*, 2020, *3* (10), e2021684–e2021684.

Sisco, Matthew R. and Elke U. Weber, "Examining charitable giving in real-world online donations," *Nature Communications*, 2019, *10* (1), 3968.

Tsai, Kellee S and Qingyan Wang, "Charitable crowdfunding in China: An emergent channel for setting policy agendas?," *China Quarterly*, 2019, *240*, 936–966.

van Duynhoven, Alysha, Anthony Lee, Ross Michel, Jeremy Snyder, Valorie Crooks, Peter Chow-White, and Nadine Schuurman, "Spatially exploring the intersection of socioeconomic status and Canadian cancer-related medical crowdfunding campaigns," *BMJ Open*, 2019, *9* (6), e026365.

Vesterlund, Lise, "Why do people give?," *The Nonprofit Sector: A Research Handbook*, 2006, *2*, 168–190.

Zhang, Junsen, "A survey on income inequality in China," *Journal of Economic Literature*, 2021, *59* (4), 1191–1239.

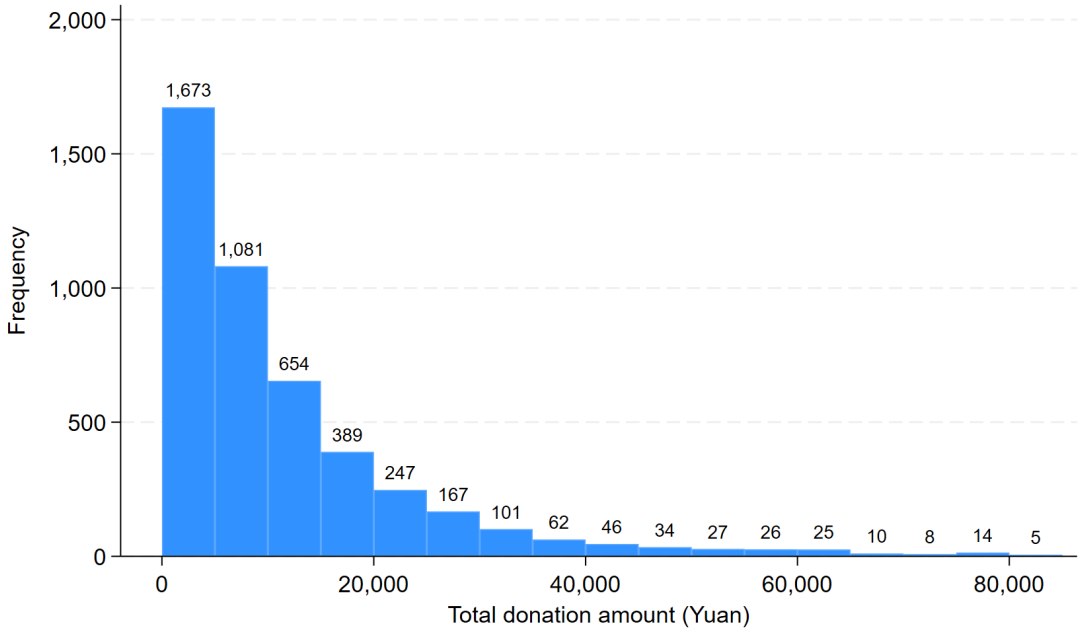
Zhang, Siyuan, Qing Zhang, Mengning Wang, Xin Tang, Xin Lu, and Wei Huang, "Key drivers of medical crowdfunding success: A comprehensive analysis of 84,712 projects," *Humanities and Social Sciences Communications*, 2025, *12* (1), 1–17.

Zhang, Xiaobo and Ravi Kanbur, "Spatial inequality in education and health care in China," in "Regional Inequality in China," Routledge, 2009, pp. 92–110.

Zheng, Xiaoting, Anqi Wu, and Xuejing Wei, "The social support returns to college education in China: Evidence from online medical crowdfunding," *China Economic Review*, 2023, *79*, 101978.

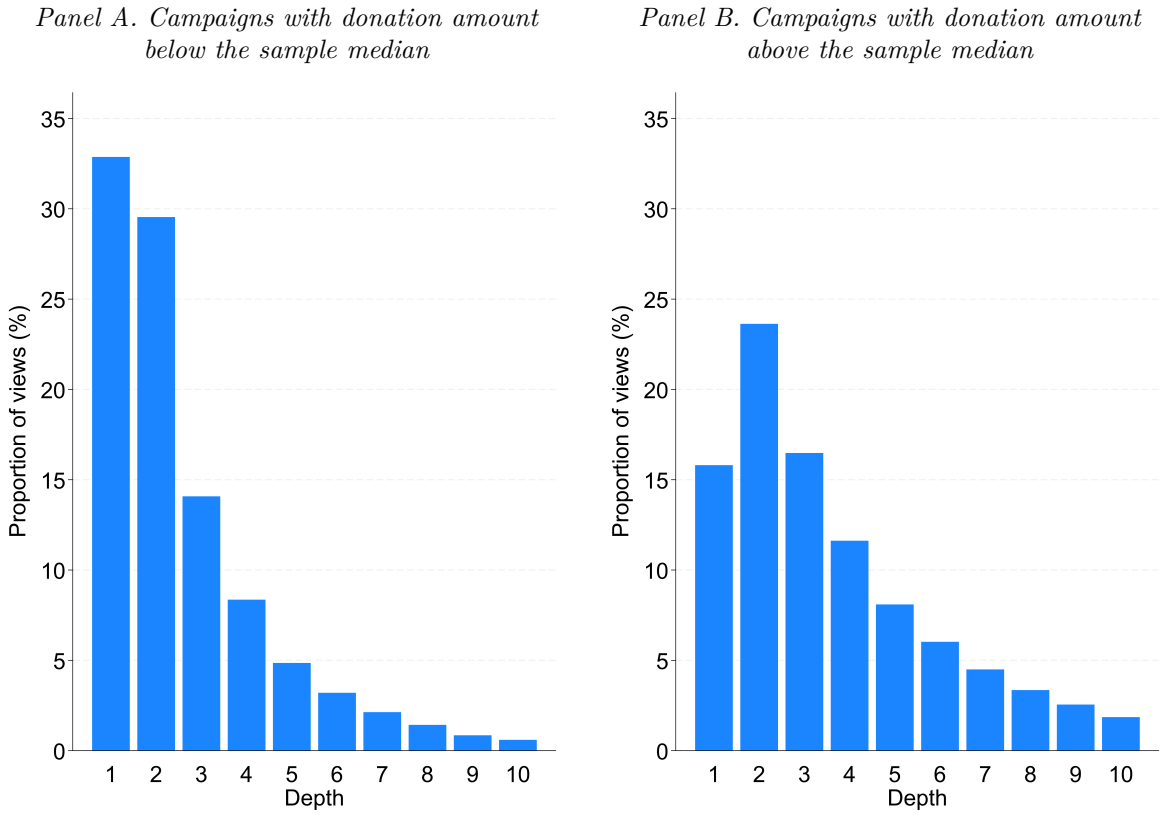
Figures and Tables

Figure 1: The distribution of total donation amounts across campaigns



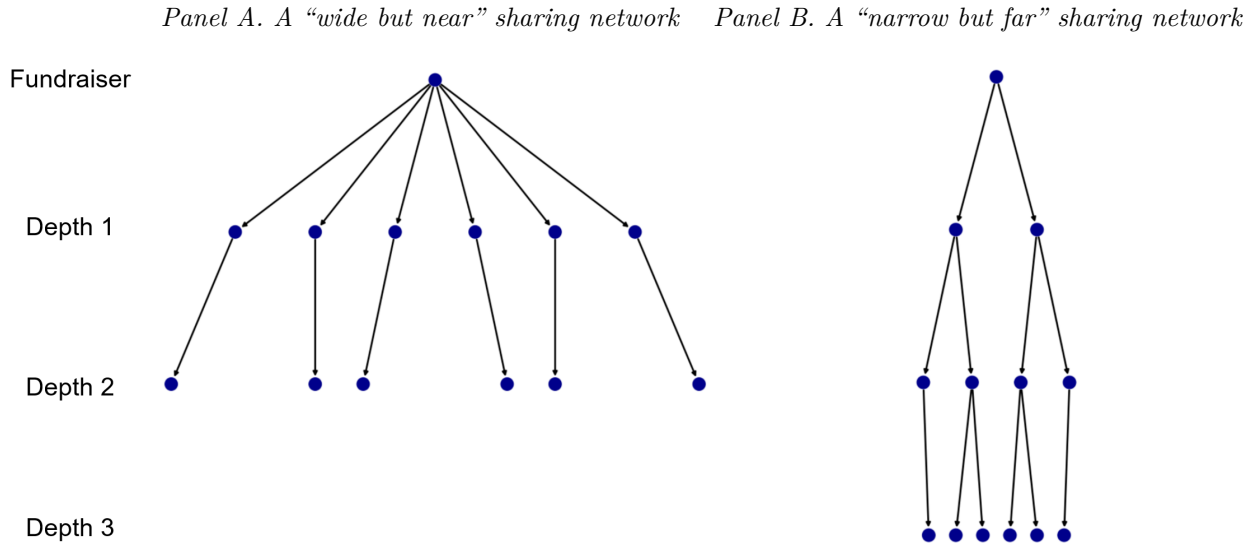
Notes: This figure plots the distribution of total donation amounts in our campaign-level sample. The numbers on the top of each bar indicate the number of campaigns within each range of donation amounts.

Figure 2: Depth distributions of campaigns with different fundraising outcomes



Notes: This figure compares the average depth distributions between campaigns with total donation amounts below the sample median (Panel A) and above the sample median (Panel B) in our campaign-level sample. The height of each bar represents the average proportion of views from the corresponding depth across campaigns within the subsample.

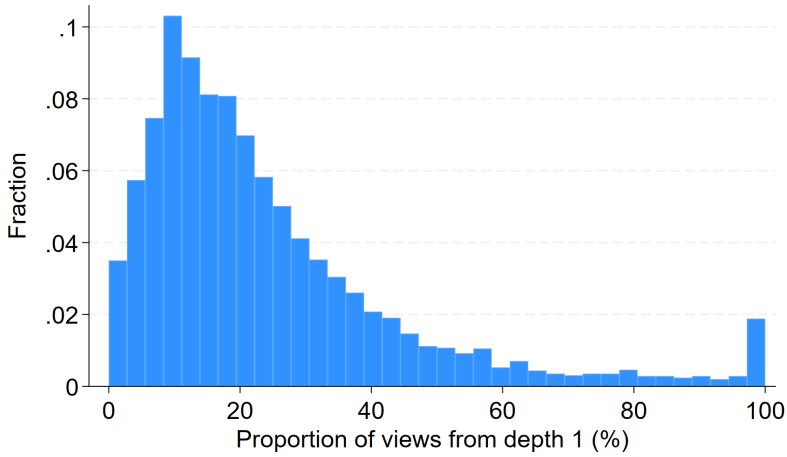
Figure 3: Illustration of two types of sharing network structure (holding #views constant)



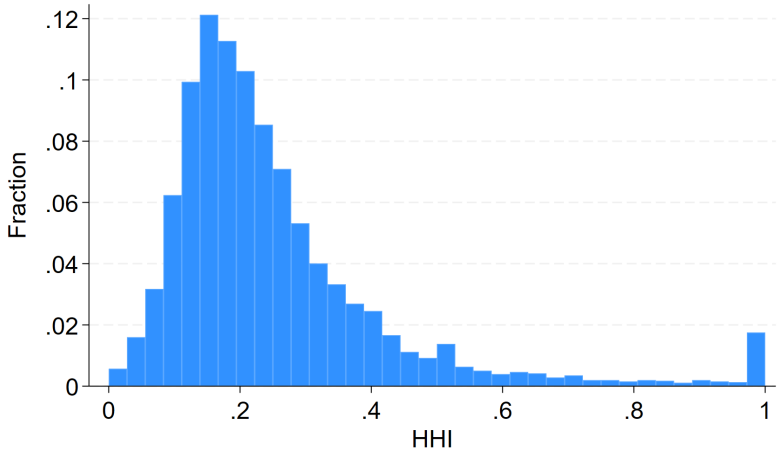
Notes: This figure illustrates two types of sharing network structure, namely the "wide but near" type (Panel A) and the "narrow but far" type (Panel B), holding the total number of views constant. Each node represents a view of the campaign web page from a viewer. The directed connections represent the diffusion path of the campaign's link.

Figure 4: Distributions of measurements of sharing network structure

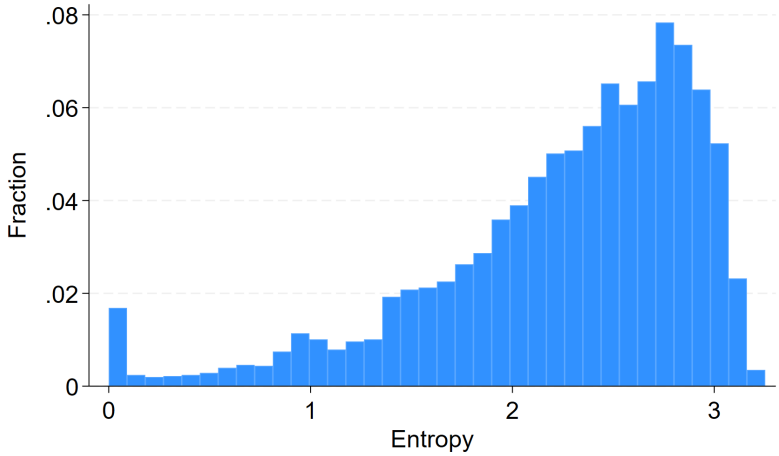
Panel A. Proportion of views from depth 1



Panel B. HHI



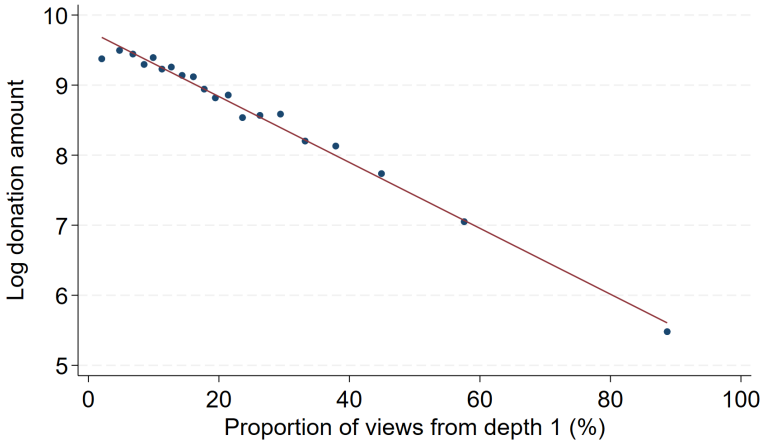
Panel C. Entropy



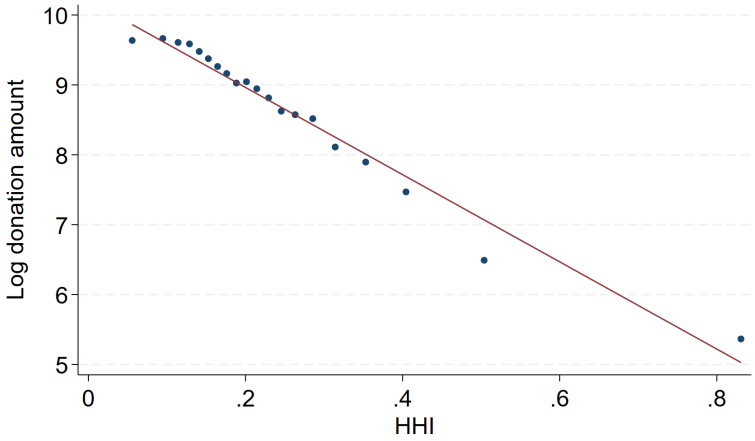
Notes: This figure illustrates the distribution of three key measurements of the sharing network structure across campaigns: the proportion of views from depth 1 (Panel A), the HHI (Panel B), and entropy (Panel C). The height of each bar represents the fraction of campaigns falling within the corresponding bin.

Figure 5: Correlations between medical crowdfunding performance and measurements of sharing network structure (binscatter plots)

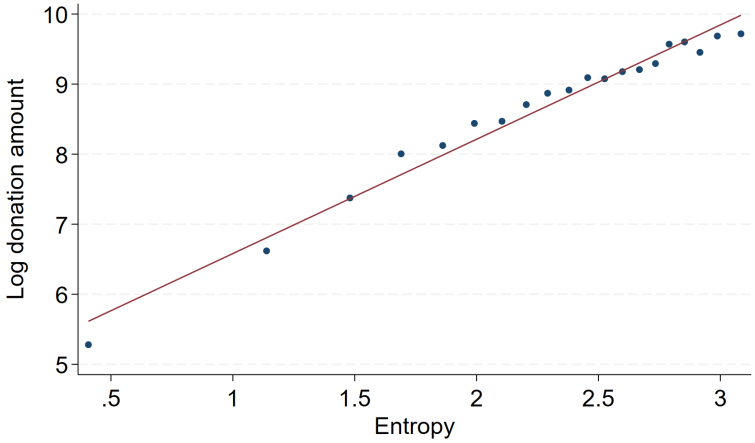
Panel A. Proportion of views from depth 1



Panel B. HHI



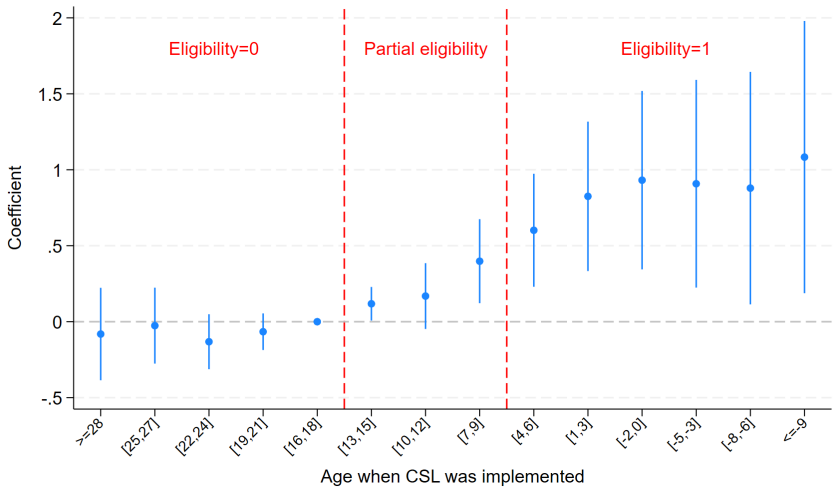
Panel C. Entropy



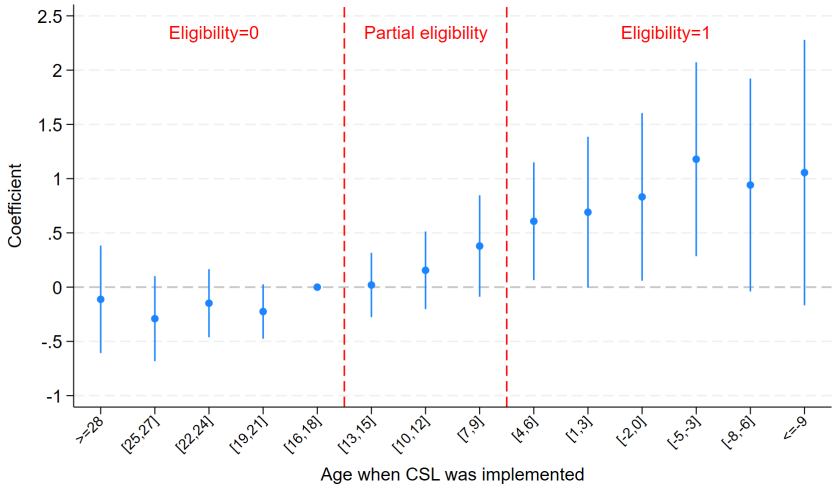
Notes: The figures show binned scatter plots, where we compute the average log total donation amount for each given level of the proportion of views from depth 1 (Panel A), the HHI (Panel B), or entropy (Panel C). In each figure, campaigns are divided into 20 equal-sized bins based on the value of the horizontal line variable, and the red line represents the linear fitted line.

Figure 6: The impact of exposure to the Compulsory Schooling Law: Event-study results

Panel A. Schooling years



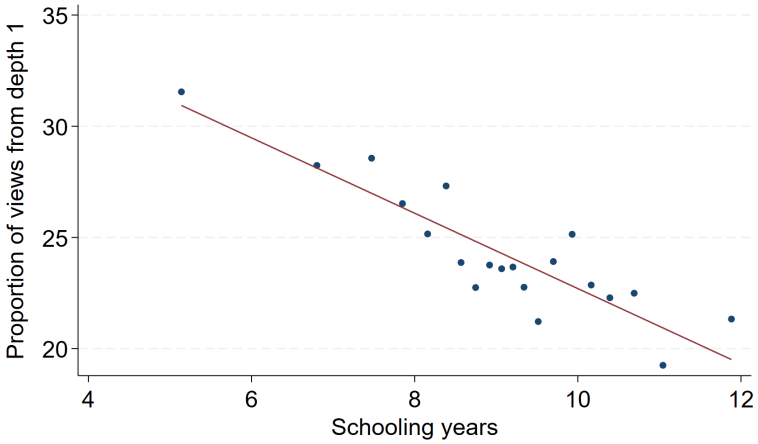
Panel B. Log donation amount



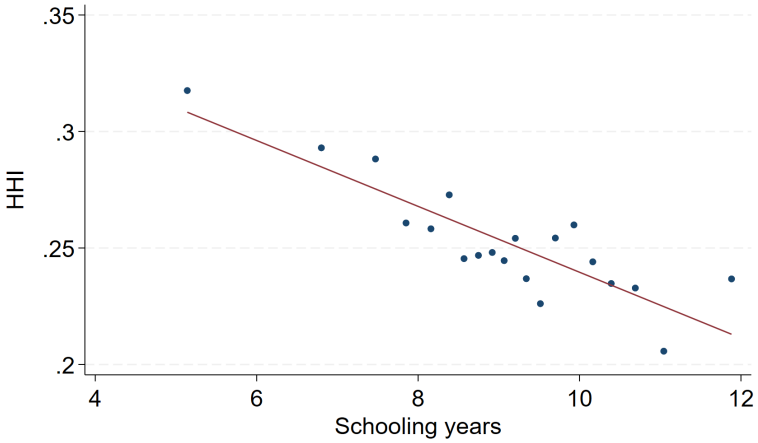
Notes: In the figures, the x-axis represents the beneficiary's age when CSL was implemented in her province of residence. The vertical dashed lines separate groups with different levels of eligibility for CSL based on their age at implementation. The y-axis represents the coefficients and 95% confidence intervals from regressions of schooling years (Panel A) or log donation amount (Panel B) on dummy variables indicating different age groups. The regressions omit beneficiaries aged 16 to 18 when CSL was implemented as the comparison group. Both campaign fixed effects and birth year fixed effects are controlled. Confidence intervals are based on standard errors clustered at the campaign level.

Figure 7: Correlations between schooling years and measurements of sharing network structure (binscatter plots)

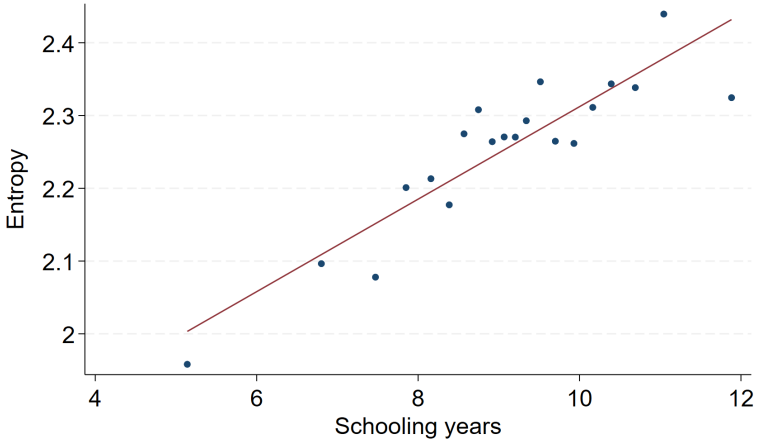
Panel A. Proportion of views from depth 1



Panel B. HHI



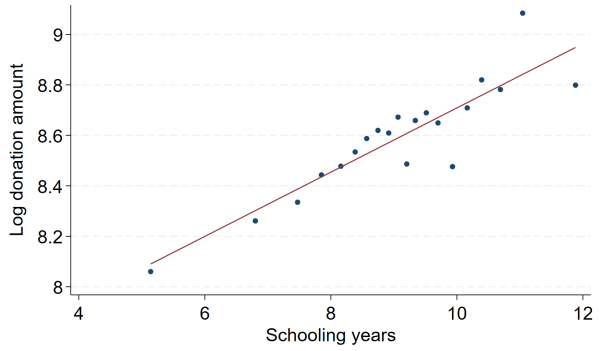
Panel C. Entropy



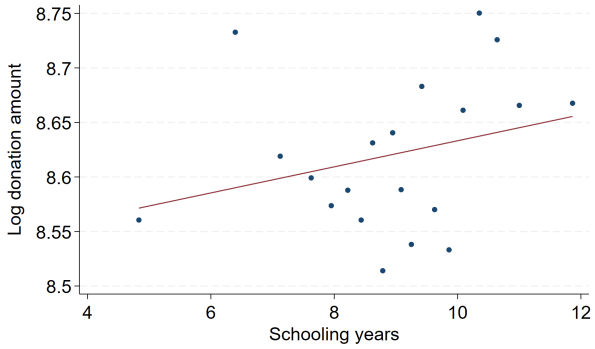
Notes: The figures show binned scatterplots that compute the average proportion of views from depth 1 (Panel A), HHI (Panel B), or entropy (Panel C) for each given level of the beneficiary's schooling years, after controlling for city fixed effects. The bins are constructed such that each dot represents the same number of campaigns. The red lines represent the fitted linear regression line.

Figure 8: Correlation between medical crowdfunding performance and schooling years: The role of network structure (binscatter plots)

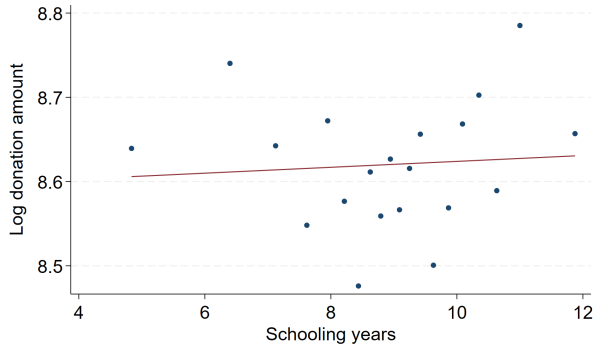
Panel A. Baseline correlation



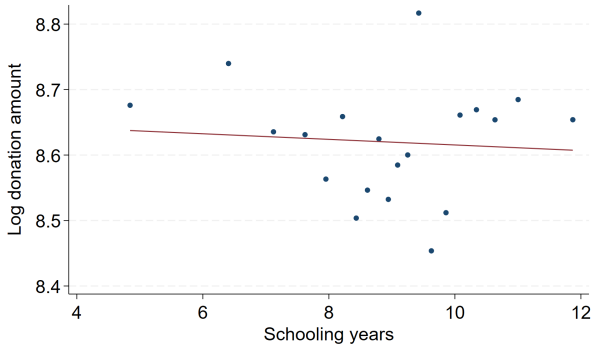
Panel B. Controlling for λ^1



Panel C. Controlling for HHI



Panel D. Controlling for entropy



Notes: TBA.

Table 1: Summary statistics

Variable	Obs.	Mean	Median	SD	Min	Max
Panel A. Medical crowdfunding performance						
Total donation amount (yuan)	4,569	11,522.9	7,706	12,739.1	1	81,562
Log total donation amount	4,569	8.633	8.950	1.551	0	11.309
Completion progress (%)	4,569	5.765	3.980	6.390	0.000	97.512
Panel B. Measures of network structure						
Proportion of views from depth 1 (%)	4,569	24.332	18.556	20.383	0	100
HHI	4,569	0.253	0.208	0.174	0.000	1
Entropy	4,569	2.258	2.419	0.669	0	3.252
Panel C. Other campaign characteristics						
Total number of views	4,569	4,144.1	2,730	4,698.5	2	48,961
Total shares to Group Chats	4,569	50.176	33	63.823	0	1,293
Total shares to WeChat Moments	4,569	158.752	101	196.490	0	2,662
Target fundraising amount (yuan)	4,569	216,441.2	200,000	111,322.3	10,000	600,000
# pictures on the web page	4,569	8.301	8	3.032	0	16
Length of content on the web page	4,569	927.128	886	247.704	103	2,988
Length of title of the web page	4,569	23.928	24	1.671	5	32
Age	4,569	46.786	51	18.632	0	92
Female (yes=1)	4,569	0.360	0	0.480	0	1
Cancer (yes=1)	4,569	0.319	0	0.466	0	1
Commercial insurance (yes=1)	4,569	0.263	0	0.440	0	1
Medical insurance (yes=1)	4,569	0.893	1	0.309	0	1

Notes: This table presents the summary statistics of our campaign-level sample. SD stands for standard deviations. In Panel A, completion progress $= \frac{\text{total donation amount}}{\text{target fundraising amount}} \times 100\%$. In Panel C, the length of contents on the web page and the length of title are in terms of the number of Chinese characters.

Table 2: Relationships between medical crowdfunding performance and measurements of sharing network structure

Dependent variable	Log donation amount								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
λ^1	-0.047***	-0.043***	-0.033***						
	(0.001)	(0.002)	(0.002)						
HHI				-6.238***	-5.862***	-4.636***			
				(0.176)	(0.177)	(0.177)			
Entropy							1.632***	1.521***	1.199***
							(0.039)	(0.041)	(0.041)
# Views			0.014***			0.013***			0.012***
			(0.001)			(0.001)			(0.001)
# Shares to Group Chats			0.275***			0.247***			0.290***
			(0.046)			(0.050)			(0.043)
# Shares to Moments			-0.028			-0.057**			-0.047**
			(0.027)			(0.025)			(0.024)
Dependent variable mean	8.633	8.633	8.633	8.633	8.633	8.633	8.633	8.633	8.633
Baseline controls	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
City fixed effects	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Observations	4,569	4,569	4,569	4,569	4,569	4,569	4,569	4,569	4,569
R^2	0.382	0.504	0.642	0.488	0.592	0.697	0.496	0.591	0.691

Notes: This table presents the results from the regressions of log donation amount on our three measures of the sharing network structure. Baseline controls include the beneficiary's age, gender, commercial and medical insurance status, the campaign's target fundraising amount, and the number of pictures, content length, and title length of the campaign on the platform. Total number of views, total shares to Groups Chats, and total shares to Moments are in hundreds. Standard errors in the parentheses are clustered at the city level. Asterisks indicate the significance level: * $p<0.1$, ** $p<0.05$, *** $p<0.01$.

Table 3: Results of the Monte Carlo experiment (#N=1,000)

#H	p	q	ν	θ_1 (%)	θ_2 (%)	θ_3 (%)	θ_4 (%)	θ_5 (%)
300	0.3	0.1	100	13.20	13.69	15.09	15.84	16.52
300	0.3	0.1	500	13.10	14.05	14.70	15.33	16.47
300	0.3	0.1	1,000	13.09	13.98	14.85	15.54	16.48
300	0.3	0.3	100	29.35	29.30	30.41	30.09	30.10
300	0.3	0.05	100	7.55	7.63	8.58	8.96	9.30
300	0.5	0.1	100	8.75	8.96	9.86	10.41	10.68
300	0.5	0.3	100	20.78	20.87	23.02	23.52	24.03
300	0.5	0.05	100	4.79	4.82	5.49	5.57	5.89
300	0.1	0.05	100	18.41	18.79	20.52	21.11	21.56
500	0.3	0.1	100	25.96	29.61	33.85	36.85	39.85
500	0.3	0.1	500	26.34	30.49	32.75	36.17	40.25
500	0.3	0.1	1,000	26.46	30.40	32.99	36.53	40.28
500	0.3	0.3	100	49.31	49.33	50.68	49.86	49.82
500	0.3	0.05	100	15.42	18.60	22.02	26.00	30.16
500	0.5	0.1	100	17.86	21.31	25.17	29.10	32.60
500	0.5	0.3	100	37.62	40.46	43.60	44.35	46.38
500	0.5	0.05	100	10.40	12.49	14.73	18.55	22.06
500	0.1	0.05	100	33.91	37.19	40.25	42.17	44.57

Notes: This table presents the results of the Monte Carlo experiment aiming to investigate the relationship between depth distribution and the proportion of views from high-SES individuals. #N is the population size, which is fixed at 1,000. Four variable parameters #H, p, q, and ν, are respectively the number of high-SES individuals, the probability that two individuals of the same group are connected, the probability that two individuals of different groups are connected, and total number of views. We specify 18 sets of parameters. For each set of parameters, we generate 100 underlying social networks, and simulate 100 sharing networks on each underlying social network. θ_1 to θ_5 are the average proportion of views from high-SES individuals across all 10,000 simulated networks with depth distribution $\mathbf{P}_1$ to $\mathbf{P}_5$, where $\mathbf{P}_1 = \frac{1}{100}(90, 10, 0, 0, 0, 0, 0, 0, 0, 0)$; $\mathbf{P}_2 = \frac{1}{100}(65, 20, 10, 5, 0, 0, 0, 0, 0, 0)$; $\mathbf{P}_3 = \frac{1}{100}(50, 30, 10, 5, 5, 0, 0, 0, 0, 0)$; $\mathbf{P}_4 = \frac{1}{100}(35, 30, 10, 10, 5, 5, 5, 0, 0, 0)$; $\mathbf{P}_5 = \frac{1}{100}(20, 30, 10, 10, 5, 5, 5, 5, 5, 5)$. Most sets of parameters will generate homophilous network structures. Notably, there are two sets of parameters in which p=q, and their consequent simulated networks are equivalent to random networks without homophilous structure.

Table 4: Education, network structure, and medical crowdfunding performance

Panel A. Education and network structure

Dependent variable:	λ^1		HHI		Entropy	
	OLS	2SLS	OLS	2SLS	OLS	2SLS
	(1)	(2)	(3)	(4)	(5)	(6)
Schooling years	-0.777*	-5.037***	-0.008**	-0.039***	0.037***	0.122***
	(0.443)	(1.292)	(0.004)	(0.010)	(0.014)	(0.044)
Dep. var. mean	24.300	24.300	0.253	0.253	2.252	2.252
Observations	4,131	4,131	4,131	4,131	4,131	4,131
R^2	0.175		0.196		0.197	
First-stage F-statistic		136.97		136.97		136.97

Panel B. OLS estimation

Dependent variable:	Log donation amount							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Schooling years	0.069**	0.035	0.020	0.010	0.043*	0.060**	0.059**	0.043*
	(0.031)	(0.024)	(0.021)	(0.021)	(0.025)	(0.028)	(0.027)	(0.026)
λ^1		-0.044***						
		(0.002)						
HHI			-6.226***					
			(0.168)					
Entropy				1.574***				
				(0.041)				
# Views					0.021***			0.020***
					(0.001)			(0.002)
# Shares to Group Chats						1.237***		0.221***
						(0.136)		(0.069)
# Shares to Moments							0.453***	-0.045
							(0.030)	(0.045)
Dep. var. mean	8.590	8.590	8.590	8.590	8.590	8.590	8.590	8.590
Observations	4,131	4,131	4,131	4,131	4,131	4,131	4,131	4,131
R^2	0.242	0.515	0.616	0.604	0.499	0.397	0.449	0.501

Table 4: Education, network structure, and medical crowdfunding performance (continued)

Panel C. 2SLS estimation (IV: CSL reform exposure)

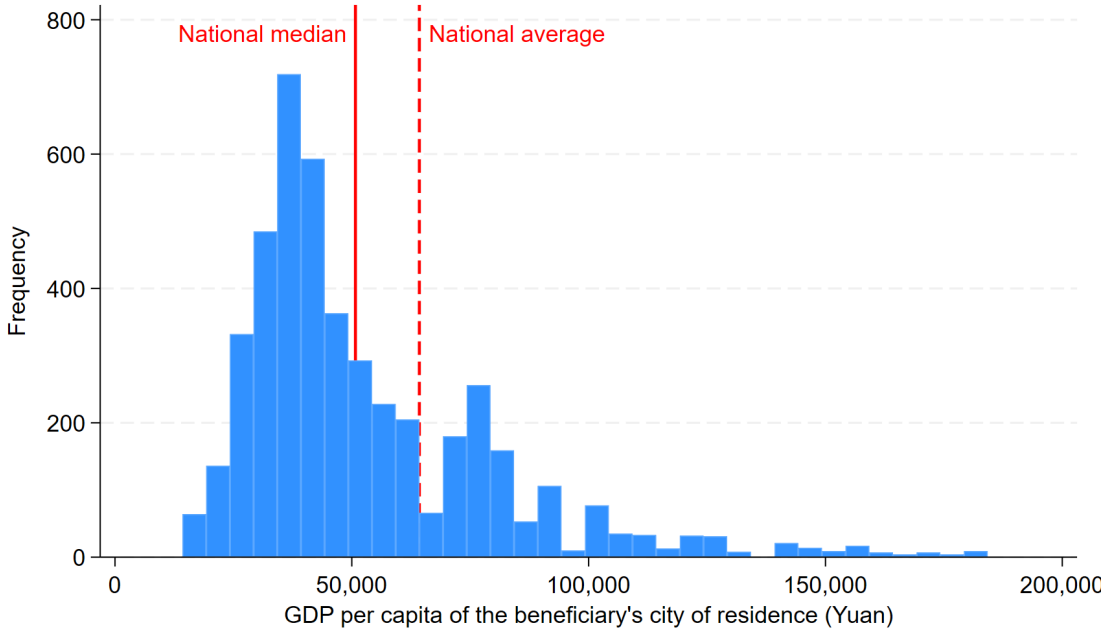
Dependent variable:	Log donation amount							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Schooling years	0.174*	-0.052	-0.071	-0.018	0.139*	0.109	0.146*	0.131*
	(0.093)	(0.077)	(0.071)	(0.069)	(0.071)	(0.082)	(0.077)	(0.071)
λ^1		-0.045***						
		(0.002)						
HHI			-6.274***					
			(0.174)					
Entropy				1.579***				
				(0.043)				
# Views					0.021***			0.020***
					(0.001)			(0.002)
# Shares to Group Chats						1.220***		0.195***
						(0.135)		(0.069)
# Shares to Moments							0.452***	-0.040
							(0.031)	(0.047)
Dep. var. mean	8.590	8.590	8.590	8.590	8.590	8.590	8.590	8.590
Observations	4,131	4,131	4,131	4,131	4,131	4,131	4,131	4,131
First-stage F	136.97	136.65	138.23	138.29	139.15	136.81	137.89	139.66

Notes: This table presents estimates from OLS and 2SLS regressions examining the associations between education, sharing network structure, and medical crowdfunding performance. Panels B and C further compare the explanatory power of network structure measures with that of other important network variables. The control variables include the beneficiary's gender, cancer status, and insurance status, the campaign's target fundraising amount, and the number of images, content length, and title length of the campaign on the platform. We exclude campaigns with beneficiaries younger than 15 years, who have typically not completed compulsory education. We drop four observations with imputed schooling years less than 2. In OLS estimations, city fixed effects and 10-year cohort fixed effects are controlled. Standard errors in the parentheses are clustered at the city-cohort level. In 2SLS estimations, we use exposure to the Compulsory Schooling Laws (CSL) reform as the instrumental variable for schooling years, which is one if the individual is fully eligible for the law (i.e., aged 6 or below when the law took effect) and zero if the individual is ineligible (i.e., aged 15 or above when the law took effect). For those aged between 6 and 15 when the law took effect, we assume that the CSL reform exposure dummy changes linearly with age. In 2SLS estimations, we apply finer cohort fixed effects. Each cohort consists of individuals of a specific age, except that age groups whose maximum CSL exposure is below 0.5 are classified as the control cohort, and those whose minimum exposure exceeds 0.5 are classified as the treated cohort. Standard errors in the parentheses are clustered at city-cohort level. Asterisks indicate the significance level: * p<0.1, ** p<0.05, *** p<0.01.

Online Appendix

A Additional figures and tables

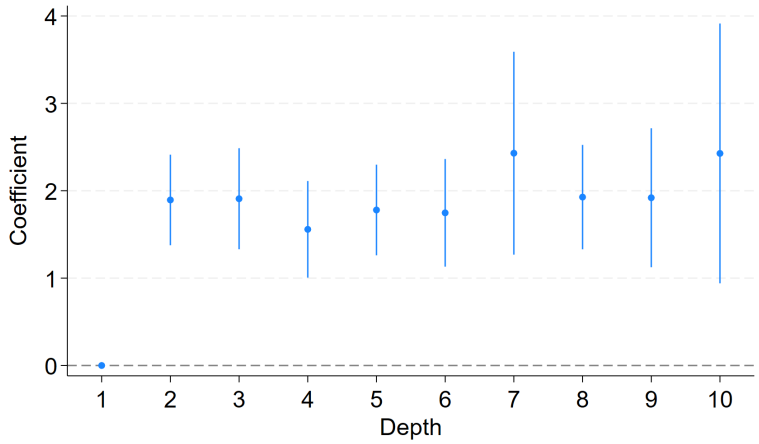
Figure A1: The distribution of GDP per capita in beneficiaries' cities of residence



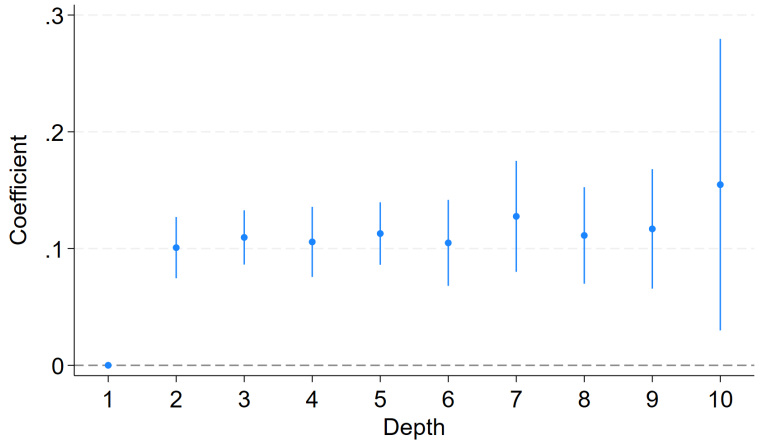
Notes: This figure plots the distribution of 2019 GDP per capita in beneficiaries' hometown prefectures in our campaign-level sample. The vertical lines represent the national median and national average of 2019 GDP per capita weighted by prefecture-level population.

Figure A2: Viewer characteristics at different depths

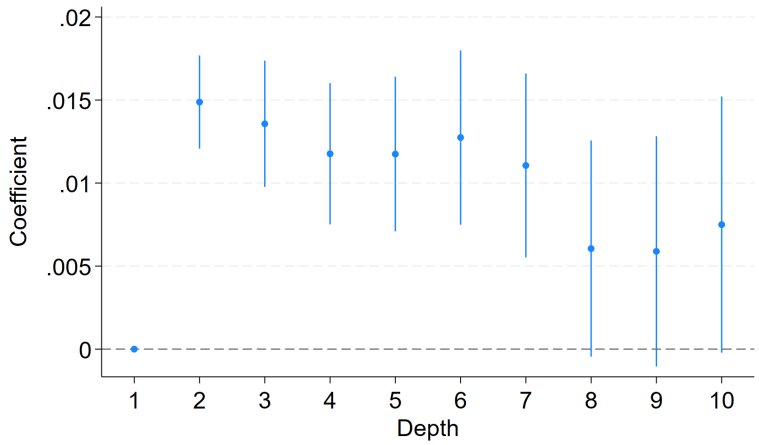
Panel A. Historical donation amount (yuan)



Panel B. Historical donation count

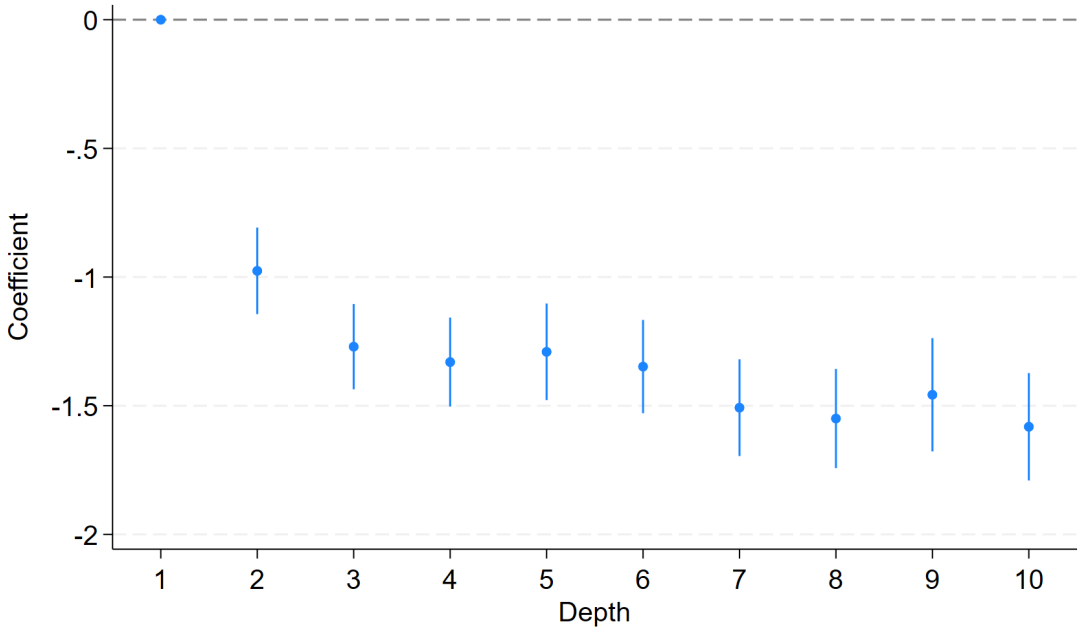


Panel C. Proportion of viewers from provinces different from the beneficiary's



Notes: The figures plot the coefficients and 95% confidence intervals from regressions of different viewer characteristics on dummy variables indicating depths 1 through 10. The sample is at the view level. The regressions omit views from depth 1 as the comparison group. Campaign fixed effects are controlled. In Panel C and Panel D, we also control viewers' city fixed effects. Confidence intervals are based on standard errors clustered at the campaign level.

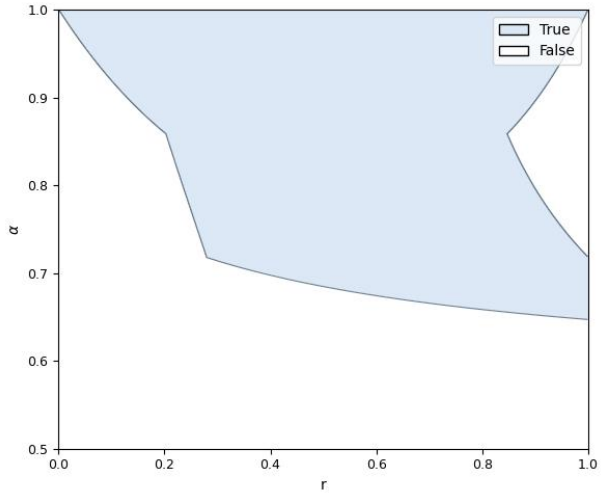
Figure A3: Donation amount across depths



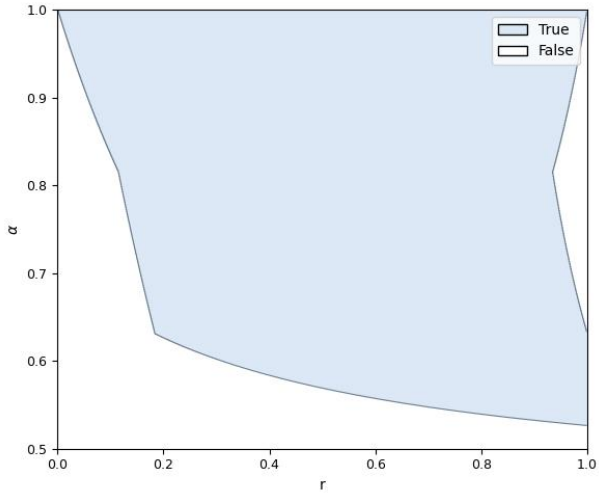
Notes: This figure plot the coefficients and 95% confidence intervals from regression of donation amount on dummy variables indicating depths 1 through 10. The view-level subsample includes only views from viewers who have viewed more than one campaign. The regression omit views from depth 1 as the comparison group. Both viewer and campaign fixed effects are controlled. Confidence intervals are based on standard errors clustered at the viewer-campaign level.

Figure A4: Parameter regions where the pros channel dominates the cons channel

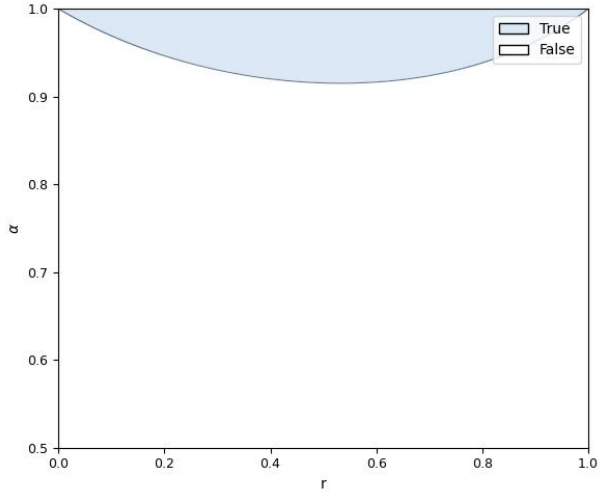
Panel A. $w_L = 1.2D, w_H = 3D$



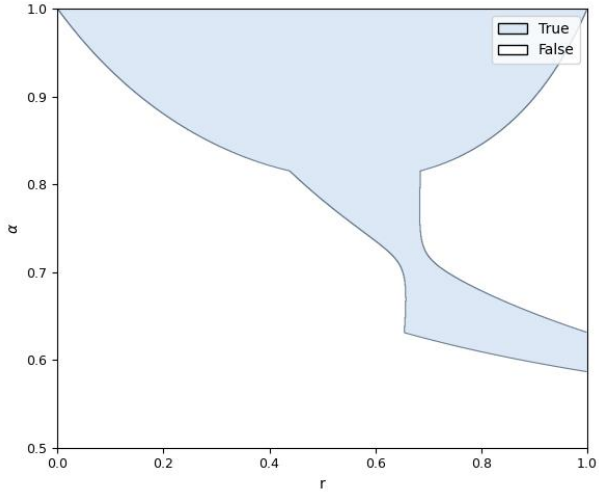
Panel B. $w_L = 1.2D$, $w_H = 5D$



Panel C. $w_L = 1.5D$, $w_H = 3D$



Panel D. $w_L = 1.5D$, $w_H = 5D$



Notes: TBA.

Table A1: Pairwise correlations among key network variables

	λ^1	HHI	Entropy	# Views	# Shares to Group Chats	# Shares to Moments
	(1)	(2)	(3)	(4)	(5)	(6)
λ^1	1.000					
HHI	0.871	1.000				
Entropy	-0.808	-0.899	1.000			
# Views	-0.406	-0.443	0.468	1.000		
# Shares to Group Chats	-0.229	-0.274	0.279	0.696	1.000	
# Shares to Moments	-0.359	-0.408	0.425	0.917	0.657	1.000

Notes: This table presents the pairwise correlation coefficients across key network variables in this paper, including λ^1, HHI, entropy, total number of views, total shares to Group Chats, and total shares to Moments. All correlations are significant with $p<0.0001$.

Table A2: Robustness check

Network variable:	λ^1		HHI		Entropy	
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A. Dependent variable: Log total number of views						
Network Variable	-0.046***	-0.041***	-6.162***	-5.711***	1.602***	1.477***
	(0.001)	(0.001)	(0.123)	(0.125)	(0.029)	(0.032)
Dep. var. mean	7.669	7.669	7.669	7.669	7.669	7.669
Observations	4,569	4,569	4,569	4,569	4,569	4,569
R^2	0.438	0.541	0.581	0.656	0.583	0.651
Panel B. Dependent variable: Completion Progress (%)						
Network Variable	-0.096***	-0.102***	-12.327***	-13.268***	3.456***	3.679***
	(0.004)	(0.004)	(0.456)	(0.514)	(0.123)	(0.131)
Dep. var. mean	5.765	5.765	5.765	5.765	5.765	5.765
Observations	4,569	4,569	4,569	4,569	4,569	4,569
R^2	0.094	0.302	0.112	0.321	0.131	0.336
Panel C. Network variables based on donation distribution across depths						
Network Variable	-0.034***	-0.030***	-5.304***	-4.970***	1.523***	1.434***
	(0.001)	(0.001)	(0.130)	(0.138)	(0.035)	(0.037)
Dep. var. mean	8.633	8.633	8.633	8.633	8.633	8.633
Observations	4,569	4,569	4,569	4,569	4,569	4,569
R^2	0.281	0.420	0.506	0.600	0.521	0.608
Panel D. Excluding campaigns with highest and lowest 5% donation amounts						
Network Variable	-0.028***	-0.026***	-4.132***	-3.873***	1.046***	0.973***
	(0.001)	(0.001)	(0.172)	(0.166)	(0.029)	(0.030)
Dep. var. mean	8.765	8.765	8.765	8.765	8.765	8.765
Observations	4,106	4,106	4,106	4,106	4,106	4,106
R^2	0.216	0.369	0.291	0.431	0.328	0.456
Panel E. Excluding campaigns with highest and lowest 5% values of the network variable						
Network Variable	-0.045***	-0.039***	-7.693***	-7.073***	1.549***	1.403***
	(0.002)	(0.002)	(0.203)	(0.207)	(0.041)	(0.042)
Dep. var. mean	8.769	8.769	8.758	8.758	8.759	8.759
Observations	4,109	4,109	4,110	4,110	4,110	4,110
R^2	0.237	0.397	0.395	0.523	0.376	0.507
Panel F. Excluding campaigns with beneficiary age below 18 or above 70						
Network Variable	-0.050***	-0.046***	-6.730***	-6.375***	1.687***	1.587***
	(0.002)	(0.002)	(0.157)	(0.166)	(0.040)	(0.042)
Dep. var. mean	8.629	8.629	8.629	8.629	8.629	8.629
Observations	3,740	3,740	3,740	3,740	3,740	3,740
R^2	0.411	0.528	0.535	0.629	0.516	0.612
Baseline controls	No	Yes	No	Yes	No	Yes
City fixed effects	No	Yes	No	Yes	No	Yes

Table A2: Robustness check (continued)

Network variable:	P^1		HHI		Entropy	
	(1)	(2)	(3)	(4)	(5)	(6)
Panel G. Excluding campaigns with beneficiaries in coastal prefectures						
Network Variable	-0.047***	-0.043***	-6.182***	-5.801***	1.612***	1.498***
	(0.002)	(0.002)	(0.194)	(0.195)	(0.042)	(0.044)
Dep. var. mean	8.612	8.612	8.612	8.612	8.612	8.612
Observations	3,983	3,983	3,983	3,983	3,983	3,983
R^2	0.379	0.501	0.482	0.586	0.489	0.584
Panel H. Excluding campaigns with beneficiaries in provincial capitals and municipalities						
Network Variable	-0.047***	-0.043***	-6.288***	-5.925***	1.643***	1.535***
	(0.002)	(0.002)	(0.187)	(0.186)	(0.042)	(0.043)
Dep. var. mean	8.637	8.637	8.637	8.637	8.637	8.637
Observations	4,049	4,049	4,049	4,049	4,049	4,049
R^2	0.382	0.506	0.498	0.603	0.501	0.597
Panel I. Excluding campaigns with beneficiaries with commercial insurance coverage						
Network Variable	-0.045***	-0.042***	-6.010***	-5.714***	1.598***	1.509***
	(0.002)	(0.002)	(0.205)	(0.199)	(0.045)	(0.046)
Dep. var. mean	8.510	8.510	8.510	8.510	8.510	8.510
Observations	3,348	3,348	3,348	3,348	3,348	3,348
R^2	0.373	0.504	0.475	0.587	0.490	0.592
Panel J. Excluding campaigns with fundraising target less than or equal to 20,000 yuan						
Network Variable	-0.046***	-0.042***	-6.143***	-5.809***	1.600***	1.498***
	(0.002)	(0.002)	(0.180)	(0.183)	(0.039)	(0.041)
Dep. var. mean	8.662	8.662	8.662	8.662	8.662	8.662
Observations	4,521	4,521	4,521	4,521	4,521	4,521
R^2	0.366	0.489	0.473	0.580	0.483	0.580
Panel K. Excluding campaigns with all views from depth 1						
Network Variable	-0.044***	-0.039***	-6.345***	-5.836***	1.556***	1.421***
	(0.001)	(0.002)	(0.208)	(0.208)	(0.040)	(0.041)
Dep. var. mean	8.704	8.704	8.704	8.704	8.704	8.704
Observations	4,498	4,498	4,498	4,498	4,498	4,498
R^2	0.307	0.453	0.426	0.547	0.441	0.554
Baseline controls	No	Yes	No	Yes	No	Yes
City fixed effects	No	Yes	No	Yes	No	Yes

Notes: This table displays robustness checks on the main results in Table 2. Baseline controls include the beneficiary's age, gender, commercial and medical insurance status, the campaign's target fundraising amount, and the number of pictures, content length, and title length of the campaign on the platform. Standard errors in the parentheses are clustered at the city level. Asterisks indicate the significance level: * p<0.1, ** p<0.05, *** p<0.01.

B Model appendix

B.1 Proofs

B.1.1 Proof of Proposition 1

Proof. Define the transition matrix

$$T = \frac{\beta}{\#H + \#L} M = \frac{\beta}{\#H + \#L} \begin{bmatrix} p\#L & q\#H \\ q\#L & p\#H \end{bmatrix},$$

then $T^k_{1,2}$ captures the probability of the fundraiser reaching a high-SES viewer at depth k, wheareas $T^k_{1,1}$ represents the probability of reaching a low-SES viewer at depth k. Thus, conditional on having reached an individual at depth k, the probability that she is of high SES is

$$\phi(k) = \frac{T^k_{1,2}}{T^k_{1,1} + T^k_{1,2}} = \frac{M^k_{1,2}}{M^k_{1,1} + M^k_{1,2}}.$$

Now we compute M^k. By Cayley-Hamilton theorem, we have $M^2 = tr(M)M - |M|I_2$, where $tr(M) = p(\#H + \#L)$, and $|M| = (p^2 - q^2)\#H\#L$. Hence, for each $k \geq 0$, there exists $a_k, b_k \in \mathbb{R}$ such that $M^k = a_k M + b_k I_2$. For $k \geq 2$, a_k, b_k satisfy the following second-order linear recurrences:

$$\begin{cases} a_k = tr(M)a_{k-1} - |M|a_{k-2}; \\ b_k = tr(M)b_{k-1} - |M|b_{k-2}. \end{cases}$$

The characteristic equations are both $x^2 - tr(M)x + |M| = 0$, whose roots are exactly the two eigenvalues of M, i.e.,

$$\mu_{1,2} = \frac{p(\#H + \#L) \pm \sqrt{p^2(\#H - \#L)^2 + 4q^2\#H\#L}}{2}.$$

Since $\#H, \#L > 0$ and $0 < q < p < 1$, we have $tr(M) > 0$, $|M| > 0$. Thus, $\mu_1 > \max\{p\#H, p\#L\} \geq \min\{p\#H, p\#L\} > \mu_2 > 0$.

Given initial values $a_0 = 0$, $a_1 = 1$, $b_0 = 1$, $b_1 = 0$, we can derive that for $k \geq 0$,

$$\begin{cases} a_k = \frac{\mu_1^k - \mu_2^k}{\mu_1 - \mu_2}; \\ b_k = -\mu_1\mu_2 \frac{\mu_1^{k-1} - \mu_2^{k-1}}{\mu_1 - \mu_2}. \end{cases}$$

Therefore,

$$M_{1,1}^k = a_k M_{1,1} + b_k = \frac{\mu_1^k(p\#L - \mu_2) + \mu_2^k(\mu_1 - p\#L)}{\mu_1 - \mu_2} = \frac{\mu_1^k(\mu_1 - p\#H) - \mu_2^k(\mu_2 - p\#H)}{\mu_1 - \mu_2} > 0;$$

$$M_{1,2}^k = a_k M_{1,2} = q\#H \frac{\mu_1^k - \mu_2^k}{\mu_1 - \mu_2} > 0.$$

As a result, for $k \geq 1$,

$$\phi(k) = \frac{q\#H(\mu_1^k - \mu_2^k)}{\mu_1^k(\mu_1 - p\#H) - \mu_2^k(\mu_2 - p\#H) + q\#H(\mu_1^k - \mu_2^k)}.$$

Define $r = \frac{\mu_2}{\mu_1} \in (0, 1)$, we have

$$\phi(k) = \frac{q\#H(1 - r^k)}{\mu_1 - p\#H + q\#H - (\mu_2 - p\#H + q\#H)r^k}. \qquad (*)$$

Since for $k > 1$,

$$\phi'(k) = \frac{q\#H(\mu_2 - \mu_1)}{(\mu_1 - p\#H + q\#H - (\mu_2 - p\#H + q\#H)r^k)^2} r^k \ln k > 0,$$

it follows that $\phi(k)$ is strictly increasing in k.

Specifically, when $\#H = \#L$,

$$T = \frac{\beta}{2} \begin{bmatrix} p & q \\ q & p \end{bmatrix} = \frac{\beta}{2} \begin{bmatrix} a + e & a - e \\ a - e & a + e \end{bmatrix}.$$

Since

$$T^k = \frac{\beta^k}{2}\begin{bmatrix} a^k + e^k & a^k - e^k \\ a^k - e^k & a^k + e^k \end{bmatrix},$$

it follows that

$$\phi(k) = \frac{1}{2}[1 - (\frac{e}{a})^k],$$

which is consistent with the general form of $\phi(k)$ in (*). ■

B.1.2 Proof of Proposition 2

Proof. By Definition 1, $\widetilde{\lambda}^j = \lambda^j - \varepsilon$, $\widetilde{\lambda}^k = \lambda^k + \varepsilon$, where $j < k$, and $0 < \varepsilon < \lambda^j$. For any positive integer $\ell \neq j$ or k, $\widetilde{\lambda}^\ell = \lambda^\ell$.

(i) $\widetilde{GEC} - GEC = (\widetilde{\lambda}^j - \lambda^j)\frac{\phi(j)}{\#H/\#N} + (\widetilde{\lambda}^k - \lambda^k)\frac{\phi(k)}{\#H/\#N} = \frac{\varepsilon}{\#H/\#N}(\phi(k) - \phi(j))$. By Proposition 1, $j < k$ implies $\phi(j) < \phi(k)$. Since $\varepsilon > 0$, $\widetilde{\text{GEC}} > \text{GEC}$.

(ii) Since $j = 1$ and $\varepsilon > 0$, $\widetilde{\lambda}^1 < \lambda^1$.

(iii-a) *HHI.*

$\widetilde{\text{HHI}} - \text{HHI} = (\widetilde{\lambda}^j)^2 - (\lambda^j)^2 + (\widetilde{\lambda}^k)^2 - (\lambda^k)^2 = \varepsilon(\widetilde{\lambda}^k - \widetilde{\lambda}^j + \lambda^k - \lambda^j)$. Since $\lambda^k - \lambda^j < -\varepsilon$, and $\widetilde{\lambda}^k - \widetilde{\lambda}^j = \lambda^k - \lambda^j + 2\varepsilon < \varepsilon$, we have $\widetilde{\lambda}^k - \widetilde{\lambda}^j + \lambda^k - \lambda^j < 0$. Since $\varepsilon > 0$, $\widetilde{\text{HHI}} < \text{HHI}$.

(iii-b) *Entropy.*

Define $\Delta(\varepsilon) = \widetilde{\text{Entropy}}(\varepsilon) - \text{Entropy}$

$$= -\widetilde{\lambda^j}\log_2\widetilde{\lambda^j} - \widetilde{\lambda^k}\log_2\widetilde{\lambda^k} + \lambda^j\log_2\lambda^j + \lambda^k\log_2\lambda^k$$

$$= -(\lambda^j - \varepsilon)\log_2(\lambda^j - \varepsilon) - (\lambda^k + \varepsilon)\log_2(\lambda^k + \varepsilon) + \lambda^j\log_2\lambda^j + \lambda^k\log_2\lambda^k.$$

The derivative $\Delta'(\varepsilon) = \log_2\frac{\lambda^j - \varepsilon}{\lambda^k + \varepsilon}$. When $0 < \varepsilon < \frac{\lambda^j - \lambda^k}{2}$, $\Delta'(\varepsilon) > 0$, and when $\frac{\lambda^j - \lambda^k}{2} < \varepsilon < \lambda^j - \lambda^k$, $\Delta'(\varepsilon) < 0$. Notice that $\Delta(0) = \Delta(\lambda^j - \lambda^k) = 0$. Therefore, when $0 < \varepsilon < \lambda^j - \lambda^k$, $\Delta(\varepsilon) > 0$, i.e., $\widetilde{\text{Entropy}} > \text{Entropy}$. ■

B.2 Details of the Monte Carlo experiment

This section provides additional technical details of the simulation in the Monte Carlo experiment in Section 5.2, and presents further analysis of the results. As previously explained, we first specify four parameters: $\#H$ (the number of high-SES individuals), p (the probability that two individuals of the same group are connected), q (the probability that two individuals of different groups are connected), and ν (total number of views). For each set of parameters, we randomly generate 100 underlying social networks with 1,000 nodes. Nodes 0 to 499 are designated as low-SES individuals, while nodes 500 to 999 are designated as high-SES. Under each underlying social network, we simulate 100 sharing networks following each of the five depth distributions.

A sharing network on an underlying social network under a certain depth distribution $\mathbf{\Lambda}_i = (\lambda_i^1, \lambda_i^2, \cdots, \lambda_i^{10})$ $(i = 1, 2, 3, 4, 5)$ will be simulated by the following process. Without loss of generality, we start from node 0, a low-SES individual designated as the fundraiser.

We randomly pick $\lambda_i^1 \nu$ of the fundraiser's direct friends (the node's neighbors) with replacement, since each direct friend could contribute more than one views. This simulates the fundraiser's shares attracting views from her direct friends, and it forms a list of views at depth 1 with $\lambda_i^1 \nu$ elements, denoted by V^1.

Then, we simulate the process that some of the viewers at depth 1 share the campaign's link and attract more views at depth 2. We randomly pick a node in V^1, and randomly draw one of this chosen node's neighbors, both with replacement. By repeating this process for $\lambda_i^2 \nu$ times, we generate another list of views at depth 2 with $\lambda_i^2 \nu$ elements, denoted by V^2.

Similar process continues iteratively up to the depth 10, and we will get 10 lists. Each list V^k $(k \in 1, 2, \cdots, 10)$ represents a depth of the simulated sharing network, i.e., all the views at depths k, and its number of elements is $\lambda_i^k \nu$. This completes a simulation of the diffusion process, and create a sharing network with a pre-specified depth distribution.

Finally, we sum the number of views from nodes 500 to 999 in each list and divide by ν, yielding the proportion of views from high-SES individuals for this sharing network. After

the simulation of all 10,000 sharing networks under the certain depth distribution $\mathbf{\Lambda}_i$, we could calculate the average proportion of views from high-SES individuals, denoted by θ_i.

We summarize several facts from the results in Table 3:

(1) Under all sets of parameters that guarantee homophily, the average proportion of views from high-SES individuals is positively correlated with the dispersion level of the depth distribution. Notably, this doesn't hold without the homophilous structure.

(2) The average proportion of views from high-SES individuals is irrelevant to total number of views.

(3) The average proportion of views from high-SES individuals is positively correlated with $\frac{q}{p}$, which is the likelihood of between-group connections relative to within-group connections.

(4) The average proportion of views from high-SES individuals is positively correlated with the number of high-SES individuals.

Fact (1) provides evidence on the validity of the pros channel, which holds true regardless of whether high-SES individuals are fewer or equal in number to low-SES individuals, whether the homophilous structure is strong, or weak, and whether the total number of views is higher or lower. Fact (2) suggests that the scale of the sharing network is unlikely to influence the proportion of high-SES individuals reached, highlighting the distinct role of network structure in the pros channel.

In addition, fact (3) indicates that a reduction of the level of homophily by SES could help fundraisers reach more high-SES individuals, which further emphasizes the role of direct economic connectedness taking the form $\frac{1}{\#H/\#N} \times \frac{q}{p+q}$. The reason behind fact (4) is that an increase in the high-SES population leads to a larger pool of potential donors. These two facts provide further insight into how societal efforts to reduce SES-based homophily and expand the high-SES population can improve medical crowdfunding outcomes.

B.3 Discussion on the Extension Model

We now examine crowdfunding success at the campaign level. For a fundraiser i whose sharing network has a depth distribution $\mathbf{\Lambda}_i$ and total number of views ν, her expected total donation amount is

$$\mathscr{D}_i = \nu \sum_{k=1}^{\infty} \lambda_i^k [\phi(k) \frac{\alpha^k w_H - D}{1+\alpha^k} \mathbf{1}\{\alpha^k w_H - D \geq 0\} + (1-\phi(k)) \frac{\alpha^k w_L - D}{1+\alpha^k} \mathbf{1}\{\alpha^k w_L - D \geq 0\}].$$

Hence, the expected marginal contribution of a viewer from depth k is given by

$$\delta_k = \begin{cases} \phi(k) \frac{\alpha^k w_H - D}{1+\alpha^k} + (1-\phi(k)) \frac{\alpha^k w_L - D}{1+\alpha^k} & \text{if } \alpha \geq \sqrt[k]{\frac{D}{w_L}}, \\ \phi(k) \frac{\alpha^k w_H - D}{1+\alpha^k} & \text{if } \sqrt[k]{\frac{D}{w_H}} \leq \alpha < \sqrt[k]{\frac{D}{w_L}}, \\ 0 & \text{if } \alpha < \sqrt[k]{\frac{D}{w_H}}. \end{cases}$$

In order to compare different depth distributions, we continue to consider a far-reaching transition (j, k, ε) of $\mathbf{\Lambda}_i$. This corresponds to a small shift toward a distribution that is more "narrow but far", which is denoted by $\widetilde{\mathbf{\Lambda}}_i$. As shown in Proposition 2, $\widetilde{\mathbf{\Lambda}}_i$ reflects a higher GEC.

Consequently, the expected total donation amount shifts to $\widetilde{\mathscr{D}}_i = \mathscr{D}_i - \nu\varepsilon\delta_j + \nu\varepsilon\delta_k$. Therefore, $\widetilde{\mathscr{D}}_i > \mathscr{D}_i$ if and only if $\delta_k > \delta_j$. Intuitively, to assess the difference in total donations, it suffices to compare the marginal contributions of the two affected depths. If $\delta_k > \delta_j$ holds, then a sharing network that is closer to the "narrow but far" type tends to generate more donations at the margin.

Proposition 3 *For a low-SES fundraiser $i \in L$ with depth distribution $\mathbf{\Lambda}_i$, we impose a far-reaching transition $(1, 2, \varepsilon)$ on $\mathbf{\Lambda}_i$, which yields a more narrow-but-far depth distribution $\widetilde{\mathbf{\Lambda}}_i$. Consequently, the expected total donation amount $\mathscr{D}_i$ shifts to $\widetilde{\mathscr{D}}_i$.*

Define the relative level of homophily $r = \frac{e}{a}$. Then, a necessary condition for $\widetilde{\mathscr{D}}_i > \mathscr{D}_i$ is:

$$1 + r > \frac{1+\alpha^2}{\alpha(1+\alpha)}.$$

The sufficient condition for $\widetilde{\mathscr{D}}_i > \mathscr{D}_i$ is given as follows.

(i) When $D < \alpha^2 w_L$:

$$1+r>\frac{1+\alpha^2}{\alpha(1+\alpha)}+\frac{1.172(w_L+D)}{\alpha(1+\alpha)(1-r)w_H};$$

(ii) When $\alpha^2 w_L \le D < \alpha w_L$:

$$1+r>\frac{1+\alpha^2}{\alpha(1+\alpha)}+\frac{1.373 w_L}{\alpha(1+\alpha)(1-r)w_H};$$

(iii) When $\alpha w_L \le D < \alpha^2 w_H$:

$$1+r>\frac{1+\alpha^2}{\alpha(1+\alpha)}+\frac{D}{\alpha^2(1+\alpha)w_H}.$$

(iv) When $D \ge \alpha^2 w_H$, $\widetilde{\mathscr{D}}_i > \mathscr{D}_i$ will never hold.

Proof. Since $\widetilde{\mathscr{D}}_i = \mathscr{D}_i - \nu\varepsilon\delta_1 + \nu\varepsilon\delta_2$, we have $\widetilde{\mathscr{D}}_i > \mathscr{D}_i$ if and only if $\delta_2 > \delta_1$.

Notice that when $\alpha^2 w_H \le D < \alpha w_H$, $\delta_2 = 0$. Thus, $\delta_2 > \delta_1$ cannot hold. And when $D > \alpha w_H$, $\delta_1 = \delta_2 = 0$. Therefore, $\delta_2 > \delta_1$ only if $D < \alpha^2 w_H$.

By Proposition 1, $\phi(1) = \frac{1}{2}(1-r)$, $\phi(2) = \frac{1}{2}(1-r^2)$.

(i) When $D < \alpha^2 w_L$,

$$\begin{aligned}\delta_2-\delta_1 =&\phi(2)\frac{\alpha^2 w_H - D}{1+\alpha^2}+(1-\phi(2))\frac{\alpha^2 w_L - D}{1+\alpha^2}-\phi(1)\frac{\alpha w_H - D}{1+\alpha}-(1-\phi(1))\frac{\alpha w_L - D}{1+\alpha}\\ =&\frac{1}{2}\left(\frac{(1-r^2)\alpha^2}{1+\alpha^2}-\frac{(1-r)\alpha}{1+\alpha}\right)w_H+\frac{1}{2}\left(\frac{(1+r^2)\alpha^2}{1+\alpha^2}-\frac{(1+r)\alpha}{1+\alpha}\right)w_L-\left(\frac{1}{1+\alpha^2}-\frac{1}{1+\alpha}\right)D.\end{aligned}$$

Thus, $\delta_2 > \delta_1$

$$\begin{aligned}\Leftrightarrow&(1-r)\alpha\left(\frac{(1+r)\alpha}{1+\alpha^2}-\frac{1}{1+\alpha}\right)w_H>-\left(\frac{(1+r^2)\alpha^2}{1+\alpha^2}-\frac{(1+r)\alpha}{1+\alpha}\right)w_L+\left(\frac{2}{1+\alpha^2}-\frac{2}{1+\alpha}\right)D\\ \Leftrightarrow&(1+r)\alpha(1+\alpha)-(1+\alpha^2)>\frac{\left((1+r)\alpha(1+\alpha^2)-(1+r^2)\alpha^2(1+\alpha)\right)w_L+2\alpha(1-\alpha)D}{\alpha(1-r)w_H}\\ \Leftrightarrow&1+r>\frac{1+\alpha^2}{\alpha(1+\alpha)}+\frac{\left((1+r)(1+\alpha^2)-(1+r^2)\alpha(1+\alpha)\right)w_L+2(1-\alpha)D}{\alpha(1+\alpha)(1-r)w_H}.\end{aligned}$$

Let $g(r,\alpha) = (1+r)(1+\alpha^2)-(1+r^2)\alpha(1+\alpha)$, where $r, \alpha \in (0,1)$. Since $1+r > 1+r^2 > 0$, $1+\alpha^2 > \alpha(1+\alpha) > 0$, it follows that $g(r,\alpha) > 0$. Hence, $1+r > \frac{1+\alpha^2}{\alpha(1+\alpha)}$ is a necessary condition.

From the necessary condition, we have $\frac{1+\alpha^2}{\alpha(1+\alpha)} < 2$. Thus, $\alpha > \sqrt{2}-1$ should hold in the sufficient condition.

Expressing $g(r,\alpha)$ as a function of α, we have $g(r,\alpha) = r(1-r)\alpha^2 - (1+r^2)\alpha + (1+r)$. Since $r(1-r) > 0$, $g(r,\alpha) < \max\{g(r,\sqrt{2}-1), \lim_{\alpha\to 1} g(r,\alpha)\} < 4-2\sqrt{2} \approx 1.1716$. Besides, $2(1-\alpha) < 4-2\sqrt{2} \approx 1.1716$. Therefore, a sufficient condition is $1+r > \frac{1+\alpha^2}{\alpha(1+\alpha)} + \frac{1.172(w_L+D)}{\alpha(1+\alpha)(1-r)w_H}$.

(ii) When $\alpha^2 w_L \le D < \alpha w_L$,

$$\begin{aligned}\delta_2 - \delta_1 =& \phi(2)\frac{\alpha^2 w_H - D}{1+\alpha^2} - \phi(1)\frac{\alpha w_H - D}{1+\alpha} - (1-\phi(1))\frac{\alpha w_L - D}{1+\alpha} \\ =& \frac{1}{2}\left(\frac{(1-r^2)\alpha^2}{1+\alpha^2} - \frac{(1-r)\alpha}{1+\alpha}\right) w_H - \frac{1}{2}\frac{(1+r)\alpha}{1+\alpha} w_L - \frac{1}{2}\left(\frac{1-r^2}{1+\alpha^2} - \frac{2}{1+\alpha}\right) D.\end{aligned}$$

Thus, $\delta_2 > \delta_1$

$$\begin{aligned}&\Leftrightarrow (1-r)\alpha\left(\frac{(1+r)\alpha}{1+\alpha^2} - \frac{1}{1+\alpha}\right) w_H > \frac{(1+r)\alpha}{1+\alpha} w_L + \left(\frac{1-r^2}{1+\alpha^2} - \frac{2}{1+\alpha}\right) D \\ &\Leftrightarrow (1+r)\alpha(1+\alpha) - (1+\alpha^2) > \frac{(1+r)\alpha(1+\alpha^2)w_L + \left((1-r^2)(1+\alpha) - 2(1+\alpha^2)\right) D}{\alpha(1-r)w_H} \\ &\Leftrightarrow 1+r > \frac{1+\alpha^2}{\alpha(1+\alpha)} + \frac{(1+r)\alpha(1+\alpha^2)w_L + \left((1-r^2)(1+\alpha) - 2(1+\alpha^2)\right) D}{\alpha^2(1+\alpha)(1-r)w_H}.\end{aligned}$$

Since $((1-r^2)(1+\alpha) - 2(1+\alpha^2)) < 0$ and $D < \alpha w_L$, we have

$$\begin{aligned}&(1+r)\alpha(1+\alpha^2)w_L + \Big((1-r^2)(1+\alpha) - 2(1+\alpha^2)\Big) D \\ >& [(1+r)(1+\alpha^2) + (1-r^2)(1+\alpha) - 2(1+\alpha^2)]\alpha w_L \\ =& (1-r)(r+\alpha(1+r-\alpha))\alpha w_L > 0.\end{aligned}$$

Thus, $1+r > \frac{1+\alpha^2}{\alpha(1+\alpha)}$ is a necessary condition. Similarly, $\alpha > \sqrt{2}-1$ should hold in the sufficient condition.

On the other side, since $D \ge \alpha^2 w_L$, we have

$$\begin{aligned}&(1+r)\alpha(1+\alpha^2)w_L + \Big((1-r^2)(1+\alpha) - 2(1+\alpha^2)\Big) D \\ \le& [(1+r)(1+\alpha^2) + (1-r^2)\alpha(1+\alpha) - 2\alpha(1+\alpha^2)]\alpha w_L.\end{aligned}$$

Let $h(r,\alpha) = (1+r)(1+\alpha^2) + (1-r^2)\alpha(1+\alpha) - 2\alpha(1+\alpha^2) = -\alpha(1+\alpha)r^2 + (1+\alpha^2)r + (1-\alpha+2\alpha^2-2\alpha^3)$, where $r \in (0,1), \alpha \in (\sqrt{2}-1, 1)$. The necessary condition implies $0 < \frac{1+\alpha^2}{2\alpha(1+\alpha)} < 1$. Thus, for any given $\alpha \in (\sqrt{2}-1, 1)$, $h(r,\alpha) \le h(\frac{1+\alpha^2}{2\alpha(1+\alpha)}, \alpha)$.

Define $M(\alpha) = h(\frac{1+\alpha^2}{2\alpha(1+\alpha)}, \alpha)$, where $\alpha \in (\sqrt{2}-1, 1)$. By envelope theorem, $M'(\alpha) = h_2(\frac{1+\alpha^2}{2\alpha(1+\alpha)}, \alpha) = \frac{-24a^6-30a^5+7a^4+8a^3-2a^2-2a-1}{4\alpha^2(1+\alpha)^2}$. Repeated differentiation yields $M'(\alpha) < 0$. Therefore, $M(a) < \lim_{a\to\sqrt{2}-1} M(a) = h(1, \sqrt{2}-1) = 2(1-\alpha)(1+\alpha^2) = (4-2\sqrt{2})^2 \approx 1.3726$. As a result, a sufficient condition is $1 + r > \frac{1+\alpha^2}{\alpha(1+\alpha)} + \frac{1.373 w_L}{\alpha(1+\alpha)(1-r)w_H}$.

(iii) When $\alpha w_L \leq D < \alpha^2 w_H$,

$$\begin{aligned}\delta_2 - \delta_1 =& \phi(2)\frac{\alpha^2 w_H - D}{1+\alpha^2} - \phi(1)\frac{\alpha w_H - D}{1+\alpha} \\ =& \frac{1}{2}\left(\frac{(1-r^2)\alpha^2}{1+\alpha^2} - \frac{(1-r)\alpha}{1+\alpha}\right) w_H - \frac{1}{2}\left(\frac{1-r^2}{1+\alpha^2} - \frac{1-r}{1+\alpha}\right) D.\end{aligned}$$

Thus, $\delta_2 > \delta_1$

$$\begin{aligned}&\Leftrightarrow \left(\frac{(1+r)\alpha}{1+\alpha^2} - \frac{1}{1+\alpha}\right)\alpha w_H > \left(\frac{1+r}{1+\alpha^2} - \frac{1}{1+\alpha}\right) D \\ &\Leftrightarrow (1+r)\alpha(1+\alpha) - (1+\alpha^2) > \frac{\left((1+r)(1+\alpha) - (1+\alpha^2)\right) D}{\alpha w_H} \\ &\Leftrightarrow 1 + r > \frac{1+\alpha^2}{\alpha(1+\alpha)} + \frac{\left((1+r)(1+\alpha) - (1+\alpha^2)\right) D}{\alpha^2(1+\alpha) w_H}.\end{aligned}$$

Since $(1+r)(1+\alpha) - (1+\alpha^2) = r + \alpha(1+r-\alpha) > 0$, thus $1 + r > \frac{1+\alpha^2}{\alpha(1+\alpha)}$ is a necessary condition.

Meanwhile, since $(1+r)(1+\alpha) - (1+\alpha^2) < -\alpha^2 + 2\alpha + 1 < 1$, it follows that $1 + r > \frac{1+\alpha^2}{\alpha(1+\alpha)} + \frac{D}{\alpha^2(1+\alpha)w_H}$ is a sufficient condition. ■

Based on our conceptual framework, this requires that the pros channel dominates the cons channel. Since r represents the relative level of homophily, the strength of the pros channel is positively associated with $1 + r$. The term $\frac{1+\alpha^2}{\alpha(1+\alpha)} = \frac{\alpha}{1+\alpha} / \frac{\alpha^2}{1+\alpha^2}$ captures the level of altruism at depth 1 relative to depth 2, thereby reflecting the strength of the cons channel. Therefore, the conditions in Proposition 3 clearly show that when the pros channel is strong enough and dominates the cons channel, a "narrow but far" sharing network would be associated with better medical crowdfunding outcomes

Therefore, our empirical findings are consistent with a high level of overall SES-based

homophily and a relatively low level of decrease in altruism with depth. Moreover, in Proposition 3, the sufficient condition generally requires w_H to be sufficiently larger than both w_L and D, which suggests that higher level of wealth inequality may also contribute to the strengthen of the pros channel, i.e., amplifying the gains from reaching beyond homophily.

C Robustness appendix

C.1 Coarsened exact matching

C.2 Instrumental variable

C.3 Network structure and average donation amount

D Institutional background

D.1 The Compulsory Schooling Law in China

TBA.