Modelling the first viral bank run: analysis of the spread of information about Silicon Valley Bank on Reddit

Del Frari Elisa, Dupleich Titouan, Massimo Riccardo, Vastola Giorgia
Journal of Behavioral Economics

May 2023

In submitting this assignment:

1. We declare that this written assignment is our own work and does not include (i) material from published sources used without proper acknowledgement or (ii) material copied from the work of other students.
2. We declare that this assignment has not been submitted for assessment in any other course at any university.
3. We have a photocopy and electronic version of this assignment in our possession.

Elisa Del Frari
(Del Frari Elisa)

Dupleich
(Dupleich Titouan)

Riccardo Massimo
(Massimo Riccardo)

Giorgia Vastola
(Vastola Giorgia)

Abstract

The world of social networks holds immense power in shaping the spread of information and ideas, including the potential for viral bank runs that can disrupt the financial sector. In this paper, we investigate the underlying mechanisms of information diffusion on SVB's bank run within the r/Economics subreddit. Through a mathematical model and network analysis, we uncover how Reddit serves as a conduit for the propagation of crucial financial information. The research seeks to explore this new phenomenon in order to enhance regulatory measures to prevent future bank runs, improve crisis communication plans of financial institutions, and uncover the potential of monitoring social media platforms for early indications of bank runs. Our research models information diffusion as an epidemic using compartmental models like S.I.R. and S.E.I.R. However, these models demonstrate to only partially capture the phenomenon. To improve our understanding, we suggest using more advanced models more suitable to capture individuals' behaviour.

1 Introduction

In recent years, social media platforms have emerged as a powerful tool to influence public opinion and shape political discourse. The financial world is not immune to this trend, with many traders and investors turning to social media to gain insights into market trends and sentiment. One noteworthy instance that showcases the connection between social media and the financial market is the GameStop short squeeze: in January 2021 individual retail investors organized on the r/wallstreetbets subreddit to buy GameStop stock, leading to a sharp increase in its price. And now, as exemplified by the Silicon Valley Bank (SVB) bank run, the new frontier might be given by digital bank runs. As the House Financial Services Committee chair Patrick McHenry described it, the SVB bank run was a "Twitter-fuelled bank run", an opinion that was shared by numerous influential newspapers.

In order to understand this emerging phenomenon, this paper examines the information diffusion process that regarded SVB on the biggest economics subreddit using compartmental models, namely SIR and SEIR, together with network science and natural language techniques to analyze the network structure and understand how it functioned as a channel for the propagation of information.

1.1 Background

Silicon Valley Bank (SVB) was a commercial bank based in Santa Clara, California, serving as the main subsidiary of SVB Financial Group. Focusing on the Bay Area region, SVB provided specialized services tailored to meet the needs of the technology industry. Consequently, it emerged as the leading bank in terms of deposits within Silicon Valley and gained popularity as the preferred financial institution for nearly half of all venture-backed tech startups. However, on March 10th, 2023, it was subject to a bank run that led to its collapse. One of the primary contributing factors was inadequate risk management, as in the previous years, when interest rates were low, the bank had invested heavily in long-term U.S. treasuries and when last year the Federal Reserve started to increase interest rates to fight inflation, SVB's bond portfolio's value started to drop. Moreover, when economic factors hit the tech sector, venture capital started drying up, leaving the bank with a shortage of capital. When, on March 8th, the bank announced that it would sell \$2.25 billion in new shares after suffering a \$1.8

billion loss on asset sales, word spread on social media and panic set off among its customers, who started to withdraw money in waves. The bank collapsed on March 10th.
This event was truly unprecedented as it marked the first-ever bank run in the age of social media, one that is characterized by more rapid dissemination of information and higher interconnectedness of individuals. Indeed, the collapse of SVB, the second-largest bank failure ever witnessed in the United States, unfolded in just two days, while the largest one, the downfall of Washington Mutual in 2008, unfolded over a prolonged period of eight months. According to analysts, this rapid escalation was fueled by contagion from anxious Twitter posts, WhatsApp conversations, and the convenience of online banking.

1.2 Literature review

Given the social aspect of language, compartmental models have been widely used to study the spread of information on social networks, thus drawing an analogy between the propagation of diseases and the dissemination of ideas.
Network epidemics have been the subject of extensive research due to their relevance in understanding the spread of infectious diseases, information dissemination, and social contagion. Pastor-Satorras and Vespignani's groundbreaking work, "Epidemic Spreading in Scale-Free Networks" (2001), established the foundation for studying epidemics on scale-free networks by highlighting the role of hubs and degree distribution in determining outbreak sizes and epidemic thresholds. Newman's "The Structure and Function of Complex Networks" (2003) explored fundamental network properties such as clustering and community structure, which are crucial for analyzing disease transmission patterns and developing targeted intervention strategies. Vespignani's "Epidemic Processes in Complex Networks" (2009) advanced the field by investigating the interplay between network structure and epidemic dynamics, addressing heterogeneous connectivity and temporal aspects of epidemics. Additionally, Morris' "Network Epidemiology: A Handbook for Survey Design and Data Collection" (2004) provided practical guidance for researchers, emphasizing the importance of high-quality network data and offering insights into survey design and measurement of network ties.

Jin et al. (2013) applied the SIS and SEIZ models to the spread of both true news and rumors on Twitter. In the first model, users could only directly and indefinitely transition from the susceptible department to the infected one, upon tweeting about a topic, while the SEIZ model introduces an exposure delay (exposed compartment, E) and the decision not to tweet after hearing about the news, (skeptic department, Z). The researchers found that the second model is more accurate at capturing the underlying mechanism. Kumar et al. (2020) modelled the information diffusion process on Twitter with a SEI model, where users were classified as susceptible, infected (when they tweeted, retweeted, replied to a tweet, or mentioned someone in a tweet), or exposed (when they saw the information while active on the social network). In their study, Zhao et al. (2013) run a SIR model on a social network through an undirected graph, where the nodes represent individuals and edges represent contacts between them. The population was categorized into three groups: ignorants, spreaders, and stiflers. Ignorants were individuals who were unaware of the rumor and could become either spreaders or stiflers upon contact with a spreader. Stiflers were those who chose not to spread the rumor due to disbelief or lack of interest, while spreaders could become stiflers if they lost interest or forgot about the rumor (forgetting mechanism). The SIHR model, proposed by

Zhao et al. (2012) extends the previous model by adding a remembering mechanism that allows spreaders to forget about the rumor and transition to the hibernator compartment, and later recall the rumor and return to the spreader compartment. Additional compartmental models have been developed to better replicate the information diffusion process, like the SICR, introduced by Zan et al. (2014). In this model, when a susceptible node encounters an infective node, it may transition to either the infected or the counterattack compartment. The latter persuades the infective nodes to refrain from spreading the rumor and become refractory with probability η. Moreover, an infective node may become a refractory individual with a probability of γ upon contacting another infective node or a refractory one.

1.3 Research question

The goal of this paper is to characterise of the underlying mechanism of information diffusion within Reddit, an online social network, during a bank run. Specifically, this paper will:

- Build a mathematical model to explain the information-spreading dynamics
- Analyze the network structure to understand how it serves as a conduit for information propagation

The ultimate objective is to gain a deeper understanding of how to effectively address and manage the emerging phenomenon of viral bank runs and, in particular, to inform and enhance regulatory measures to prevent future bank runs and promote stability in the financial sector, improve financial institutions' crisis communication plans to mitigate panic and restore confidence and assess whether monitoring social media platforms can provide early indications of potential bank runs.

2 Data

We decided to focus on Reddit for the following reasons:
- it allowed us to access a restricted and well-defined community;
- it has a history of impact on financial markets;
- it is easy to grab data;
- the majority of its users are American, which is our target community.

2.1 Data collection

For the collection of data on the subreddit community, we employed the Praw library. This scraping tool allowed us to effectively utilize the official Reddit API. By exploiting it, we were able to generate a dataframe with the necessary information. The mentioned DataFrame had the following columns: 'datetime', 'type', 'id', 'author', 'parent_id', 'parent_author', 'link_id', 'score', 'svb','content'. Each of them specified a feature of a comment/post: the 'type' column specified whether the row was about a comment or a post, the 'author' column revealed the nickname of the writer, 'svb' took the form of a dummy variable, being equal to 1 in the case the comment/post was about SVB, 0 otherwise, 'parent_author', instead, was only not null if the row linked to a comment, and it referred to its parent comment/post. One peculiarity that needs mentioning is the fact that we only considered direct comments as connections. Hence, if user A had replied to user B's comment, which had been itself a reply to user C's comment, although A and B had a connection (just like B and C), A and C did not. The reason for this approach was to construct the edges of the network so that they respected the real connections between users. Then, the column 'score' indicates the number of 'upvotes' that a post or comment has received. Note that this column was not used in the final analysis. In order to gain a deeper understanding of the sentiment and emotions expressed in users' posts and comments and to conduct a more thorough exploration of the topics addressed during that time frame, we decided to also scrape the content of comments and posts (column 'content'). After collecting this data, we decided to filter the dataframe, considering a time frame from the 9th of March at 11:30 a.m. until the end of the 11th of March. The reason why is that, as shown in Figure 1, we wanted to analyze the period before and right after the bank run. During this timeframe, posts and comments about the American bank accounted for 40% of all submissions.

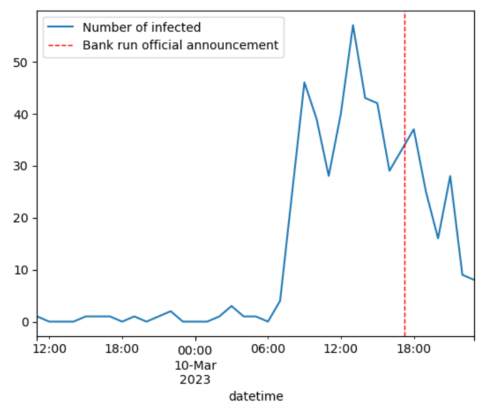

Figure 1: Count of infected users over time

2.2 Data preprocessing

Sentiment analysis was performed on the comments and posts' text using FinBERT, a pre-trained language model specifically designed for financial sentiment analysis by finetuning BERT on a large corpus of financial text data. The sentiment score is calculated as the difference between the counts of positive and negative sentiments, divided by the sum of positive and negative counts. Thus it ranges from -1 to 1, indicating respectively highly negative or positive sentiment. Moreover, emotion detection was performed using Hugging Face's Emotion English DistilRoBERTa-base, a pre-trained BERT-based model specifically designed to accurately detect Ekman's six basic emotions (anger, disgust, fear, joy, sadness, surprise) along with a neutral class.

The information diffusion process regarding the Silicon Valley Bank (SVB) run began on March 9th at 11:49 with a post that stated: 'Silicon Valley Bank Financial Realigns Portfolio–and Blows Up the Banking Sector, Fears of Contagion Send Bank Stocks Down'. This post, which marked the initiation of the information propagation process within the subreddit, exhibits a prevailing negative sentiment, accompanied by an emotion of fear. All of the remaining SVB-related activity on that day consisted of nine comments to this initial post, while throughout the analyzed timeframe, a total of 25 comments were posted in response to it, with some users agreeing on the gravity of the situation and the severe consequences for the economy and some dismissing it as just "some volatility in the market [...] but harshly something new". Overall, the sentiment was highly negative (-0.71). Thus, the first post sparked negative feelings towards SVB's situation.

3 Network analysis

3.1 Structure

By modelling networks, we can gain insights into the relationships and interactions between users, uncovering patterns of influence and information flow and identifying key nodes that play crucial roles.

The network is composed of:

- Nodes: all unique users who either posted or commented within the selected subreddit during the timeframe considered (1302 in total).

- Edges: connects two users if one user commented on a post or comment made by the other user (1788 in total).

As shown in Figure 2, in the latter half of March 9th, the first post and comment between 11:30 and 11:59, the activity level remained quite low (8 submissions). However, there was a significant surge observed in the first segment of March 10th, with a remarkable increase to 282, further escalating to 449 in the second half of that day. Thus, the trend does resemble that of an epidemic.

Figure 2: Evolution of network of infected users every 12 hours from March 9th at 11:30 to March 10th at 23:59

3.2 Key measures

Figure 3 below shows the degree distribution, which indicates the probability that a randomly selected node in the network has degree k, thus showing the pattern of connectivity among nodes.

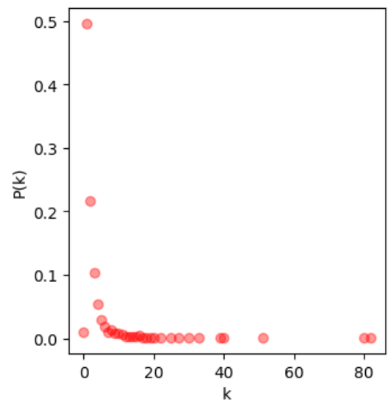

Figure 3: Degree distribution

The degree distribution follows a power-law distribution, indicating the presence of a few highly connected nodes (hubs) and numerous nodes with relatively low degrees. This distribution suggests a scale-free network, where a small number of nodes play critical roles in information flow and influence.

Table 1: Key measures

Average degree	2.75
Average clustering coefficient	0.053
Sparseness	0.0008
Connected components	27

The network exhibits an average degree of 2.75, indicating the average number of comments exchanged by users within the network. With a sparseness coefficient significantly below 1, the network is considered sparse. Furthermore, the network consists of 27 disconnected components, indicating a lack of overall connectivity.

Table 2: Centrality measures

Measure	Mean	Min	Max
Degree centrality	0.0009	0	0.6
Closeness centrality	0.18	0	0.28
Betweenness centrality	0.003	0	0.22

By analyzing the degree centrality measure, we can identify Reddit users who are highly engaged and influential within the network. These users often receive a large number of comments or have interactions with diverse users, indicating their potential impact on discussions and information dissemination. The degree centrality distribution is right-skewed, which indicates that the majority of nodes in the network have relatively low degrees of connectivity, while a small number of nodes possess a disproportionately large number of links (hubs). This suggests an uneven information flow throughout the network. The node with the highest degree centrality on March 10th at 07:54 posted: "Silicon Valley Bank is shut down by regulators, FDIC to protect insured deposits", thus simply presenting the factual information without urging individuals to withdraw their funds. This post garnered considerable attention and sparked a discussion, generating a total of 51 comments. The sentiment is highly negative (-0.91) and the prevailing emotions observed, aside from neutrality, include surprise, disgust, and fear. These emotions collectively signify a spread of unfavorable opinions and criticisms towards the bank. The nodes with the second and third highest degree centrality did not posts on SVB-related matters.

To understand the flow of information and the potential spread of ideas across different communities within the network we studied betweenness centrality, which measures the number of shortest paths between all pairs of nodes in a network that pass through each node. Users with high betweenness centrality act as bridge connectors between different parts of the network. Again, the top three nodes are those with also the highest degree centrality. Such a result comes from the fact that having a high degree centrality means having numerous connections to other nodes, which translates into a higher likelihood to serve as bridges between different parts of the network. Moreover, they act as central hubs or connectors, facilitating the spread of information to other nodes. This increased control over information flow translates into higher betweenness centrality, as these nodes are positioned to regulate and mediate the exchange of information across the network.

Individuals with high closeness centrality, calculated as the reciprocal of the sum of the shortest path lengths between that node and all other nodes in the network, are considered to be socially close to other individuals in the network and have strong connections to different parts of the network. They are well-positioned to receive information from various sources and disseminate it efficiently to others. The account with the highest closeness centrality is again the one that had the top betweenness and degree centralities, while the second user stands out as a consistent contributor, actively engaging with a substantial number of comments on a daily basis, amounting to a total of 35 within the analyzed timeframe. Notably, a significant portion of its comments revolved around SVB, exhibiting a predominantly negative sentiment accompanied by emotions such as fear, surprise, sadness, and neutrality. The third one is a user, who, on the morning of March 10th, posted 4 comments on SVB, all characterized by negative sentiment.

To explore better how the edges were structured we iterated 11 times on the network, eliminating each time the first ten nodes with the highest edge betweenness centrality, since eliminating only one node at a time would have had minimal effect. In the first iterations, as displayed in Figure 4, nothing seems to change much. Only at the 5th or 6th iteration, the network seems to start breaking. This means that probably there is not really the presence of interrelated nodes that serve as crucial connectors.

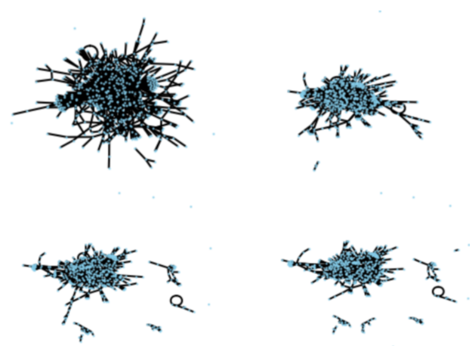

Figure 4: Original graph and graphs obtained by removing nodes with high edge betweenness (iterations 0, 5, 6)

The same has been done for degree centrality. In this case, since degree centrality is a more local measure, at each iteration the 2 nodes with the highest degree centrality were removed, iterating it 11 times. After the 4th and the 5th iteration, it is very clear that the network has started breaking. This suggests that these users play a crucial role in maintaining the overall connectivity and unity of the community by bridging different individuals and conversations, not on the overall network but locally.

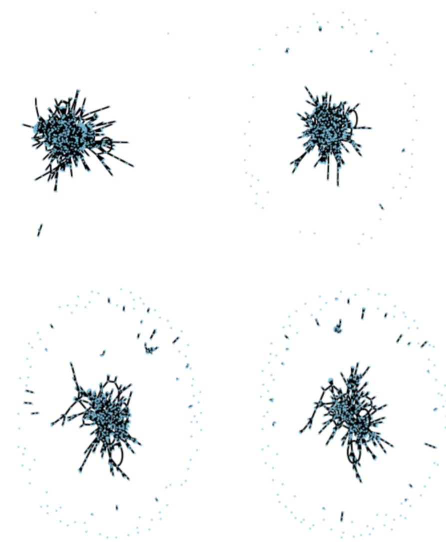

Figure 5: Original graph and graphs obtained by removing nodes with high degree centrality (iterations 0, 4, 5)

The contrasting behavior arises from the fact that degree centrality primarily considers local connectivity (users who have posted or commented to each other), while betweenness centrality takes into account the global structure and flow of information within the network.

3.3 Community detection

By applying community detection algorithms, we can uncover distinct communities or clusters of users based on their interactions. Analyzing the content and characteristics of these communities can provide insights into the formation of interest groups or topic-based discussions on Reddit.
The Louvain best partition algorithm has provided valuable insights by detecting 47 distinct communities within the whole network, as pictured in Figure 6.

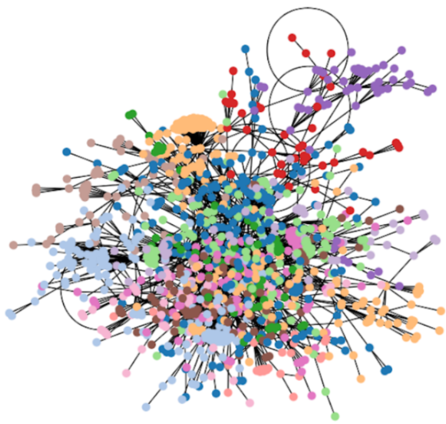

Figure 6: Communities

To gain a deeper understanding of each community's focus, Latent Dirichlet Allocation was applied. Not all communities discussed SVB, despite considering a small time frame. However, among those that did discuss SVB, it was observed that they approached the topic from various angles. Some emphasized the bank's economic stability and the impact of the COVID-19 pandemic, while others focused on political aspects such as the influence of the Federal Reserve, government policies, and prominent political figures. Additionally, certain communities explored how public scandals and societal issues affected the public's trust in banks during potential bank runs. These findings highlight the diverse interests and perspectives within these communities, shedding light on the complex dynamics surrounding SVB Bank and the phenomenon of bank runs.

4 Models

4.1 SIR

Degree-based models operate under the premise that nodes sharing the same degree k are statistically indistinguishable. By adopting this less stringent assumption, these models aim to address the constraints imposed by the strong assumption of perfect mixing of homogeneous models. First, we started with a more basic model, implementing the S.I.R., which provided a simplified framework to examine the spread of fears concerning the crisis. As we progressed, we introduced a refinement by incorporating an Exposed compartment into the model. This addition increased the complexity and was introduced to improve the ability of the model to capture the realistic progression of the crisis.

In the S.I.R. the compartments were divided in the following way:

-Susceptibles: the whole network (1302 nodes)

-Infected: users who have posted or commented something about SVB

-Recovered: members of the subreddit that stopped commenting/posting about the SVB crisis.

The partial equations characterising the model are:

$$\begin{cases} \frac{\partial S_k(t)}{\partial t} = -\beta S_k(t) k \sum_{k'} k' \frac{P(k')}{<k>} \frac{I_{k'}(t)}{N_{k'}} \\ \frac{\partial I_k(t)}{\partial t} = +\beta S_k(t) k \sum_{k'} k' \frac{P(k')}{<k>} \frac{I_{k'}(t)}{N_{k'}} - \gamma I_k(t) \\ \frac{\partial R_k(t)}{\partial t} = R_k(t) + \gamma I_k(t) \end{cases} \tag{1}$$

The initial conditions were expressed by a DataFrame grouped by the degree with 6 columns, referring to the degree 'k', the number of users with that degree 'Nk', the starting 'Sk', 'Ik' and 'Rk' for each degree level and 'pk', the probability to connect to a node of degree k.

Since our focus is on studying the individuals who wrote about SVB (considered as infected in this model), β signifies the rate at which susceptible individuals transition to the infected state by commenting about SVB.

The count of infected users (Figure 1) was computed by only considering the first post related to SVB for each user. This makes sense since subsequent posts are representing the fact that individuals remain infected for some time before recovering.

Then, we defined the recovery time as the period since when an individual posted about SVB until the first time they posted about something unrelated. However, we acknowledge that there could have been instances where individuals posted about SVB, then something else, and then again about SVB. Although this scenario occurred only in a few cases, we made the decision to overlook it for the sake of simplicity and clarity in our analysis. In our case, γ was defined as 1 over the median of the time necessary for an Infected to write about something different from SVB. The median, which we found to be equal to 3.7, was preferred over the average due to its increased robustness to outliers. In the context of tracking the time between SVB-related posts and non-SVB posts, there were cases where unusually long or short intervals occured. By using the median, we mitigated the influence of such outliers and obtained a more representative estimate of the typical time it takes to transition from one compartment to another.

The parameter β was calibrated by minimizing the Root Mean Squared Error (RMSE):

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(I_i - \hat{I}_i)^2}$$

Since $\beta = \frac{R0}{d}$, and d is known, β was selected among a range of values tested for R0. Results were computed with a time-step of 60 minutes.

4.2 SEIR

The second model employed in our study was an S.E.I.R. (Susceptible, Exposed, Infected, Recovered) model. This model introduced an additional compartment — Exposed — which marked an increase in the overall complexity of the model. The Exposed compartment adds an exposure delay by including members of the community who were exposed to infected content within the subreddit but had not yet commented on the SVB crisis. The underlying rationale was that users take some time to intake content related to the crisis and thus, only start posting later. The S.E.I.R. model was then structured according to the following formulas:

$$\begin{cases} \frac{\partial S_k(t)}{\partial t} = -\beta S_k(t) k \sum_{k'} k' \frac{P(k')}{<k>} \frac{I_{k'}(t)}{N_{k'}} \\ \frac{\partial E_k(t)}{\partial t} = +\beta S_k(t) k \sum_{k'} k' \frac{P(k')}{<k>} \frac{I_{k'}(t)}{N_{k'}} - \alpha E_k(t) \\ \frac{\partial I_k(t)}{\partial t} = +\alpha E_k(t) - \gamma I_k(t) \\ \frac{\partial R_k(t)}{\partial t} = +\gamma I_k(t) \end{cases} \tag{2}$$

In the S.E.I.R. model, an additional parameter, α, was introduced. α denotes the rate at which Exposed users become Infected. Determining this parameter involved analyzing the rate at which active subreddit members (who were not yet commenting on SVB) started engaging in discussions about the crisis. By introducing the Exposed state, our hope is to capture the nuanced dynamics within the subreddit community more effectively, by distinguishing between users who were merely members of the subreddit and those who were actively engaging within it. While γ remained the same as in the previous model, β took on a new meaning, specifying the rate at which members of the community (Susceptibles) showed to be active and became Exposed.

Since we did not have data for the number of people being exposed to SVB content before becoming infected themselves, the new α parameter was calibrated with RMSE. Thus, the combination of α and β that yielded the lowest RMSE was chosen.

5 Results

5.1 SIR

Figure 7 below shows the SIR model calibrated on posts and comments in the specified timeframe. The model's best fit was reached for a beta of 0.06, with a RMSE equal to 10.42 This means that every hour, 6% of contacts between an infected I and a susceptible S leads to S posting or commenting about SVB and hence becoming infected. R0 is β times d, which is equal to 0.22. Thus, on average each infected individual creates 0.22 secondary infections.

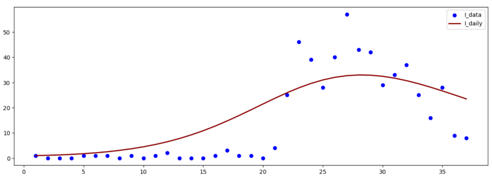

Figure 7: SIR calibration

When looking at the greater picture of the full SIR dynamics (Figure 8), it seems that the infection curve is minimal compared to the number of Susceptible over time. However, it is worth noting that this in no way hinders the importance of the news spread on the subreddit. Indeed, the sharp rise in infections within a few hours in Figure 7 appears here as a clear marker of increased worry right before the bank run's official announcement on the next day.

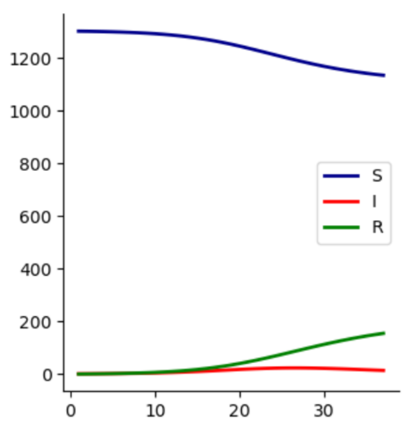

Figure 8: SIR dynamics on the network

Looking at infection curves for each degree level (Figure 9), it is clear that higher degree nodes are subject to much faster and stronger waves of infection. Due to their higher degree, susceptibles have increased exposure to infectious individuals. As expected, posts unrelated to SVB and without any comments (nodes with degree 0) did not and could not suffer from infections. It is worth noting that the proportion of infected individuals with a degree of 6 starts above 0 since this corresponds to the degree of the first SVB post's author.

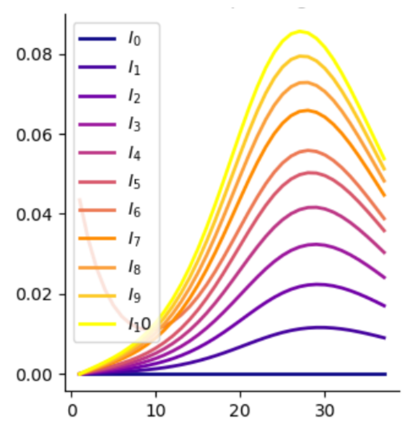

Figure 9: Percentage of infected per degree from SIR

5.2 SEIR

The SEIR model provided a slightly better fit with an RMSE of 9.54. β was equal to 0.09 and R0 was equal to 0.27, thus higher than the one obtained with SIR. The calibrated alpha (0.6) implies that individuals who are first exposed to infectious nodes take on average $1 / \alpha = 1.7$ periods (1.7 hours) before posting about SVB themselves. The very short time spent in the exposed compartment helps to explain the only slight improvement in model fit. Hence, the added complexity, and reduced number of degrees of freedom, caused by the SEIR compartment might not be worth the small decrease in RMSE. Indeed, we can imagine that the panic wave created by the network led to fast worried reactions to infected content, making exposed individuals quickly become infected themselves. Still, it should be noted that due to its increased flexibility, SEIR was better able to capture the sudden drastic increase in cases followed by a rapid fall right before the bank collapse. This better fit suggests that more complex models, better tailored to Reddit communities might lead to even more accurate approximations of information spread.

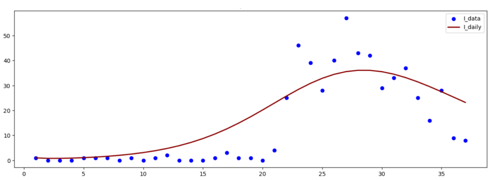

Figure 10: SEIR calibration

In the two figures below are shown the SEIR dynamics and the percentage of infected per degree from SEIR. Similar reasoning as before applies.

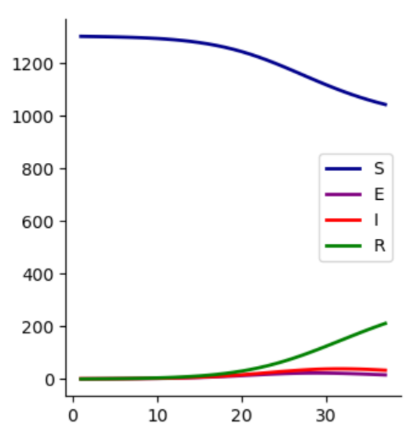

Figure 11: SEIR dynamics on the network

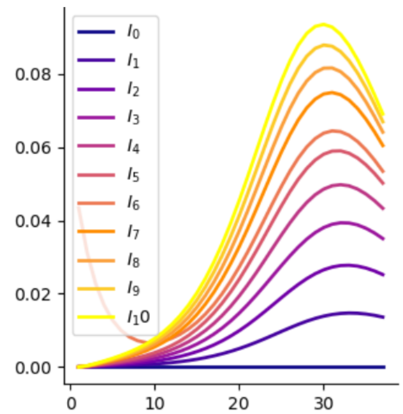

Figure 12: Percentage of infected per degree from SEIR

6 Limitations

There are several limitations to consider in our study. Firstly, it is important to acknowledge that we cannot claim causality, but just a positive correlation, between the diffusion of posts about the precocious conditions of SVB and the bank run itself, as there were likely other factors at play.
In addition, our analysis was limited to a single subreddit, neglecting the influence of other social media platforms where discussions about the bank run may have been more prominent, such as Twitter. However, utilizing Twitter would have required making additional assumptions, since taking all Twitter users as our population was computationally infeasible.
Furthermore, it is important to note that the models used in our analysis were relatively simple, aiming to showcase the feasibility of utilizing compartmental models to study complex phenomena like bank runs. However, we recognize that more advanced compartmental models can be employed to capture a higher level of complexity and accuracy in future research. Models tailored specifically for Reddit communities could potentially be better able to capture the very sharp rise followed by rapid descent

in infectious cases. Such models could for instance make use of the 'score' associated with each post/comment to better assess their impact.
In future research, it would be highly valuable to explore additional bank run cases, such as the Credit Suisse case, and investigate whether the patterns observed in our study hold true across different scenarios. By examining the Credit Suisse incident, we can assess whether the synchronization between social media discussions and the occurrence of a bank run is consistent or if there are variations in the relationship. This comparative analysis would provide a broader perspective on the role of social media in different financial events and allow us to identify commonalities or distinctions in how social media influences the dynamics of bank runs.

7 Conclusions

The network analysis of Reddit user interactions provides valuable insights into the patterns of engagement and community dynamics within the platform. By examining key influencers, bridge users, and community structures, we gained a deeper understanding of how information spreads, communities form, and discussions evolve on Reddit. These findings can be useful for various applications, including content moderation and understanding user behavior on online social platforms.
Compartmental models provide a simplified representation of the dynamics of information propagation processes at the population level, as well as insights into important parameters such as the basic reproduction number (R0), which represents the average number of secondary infections caused by a single infected individual. These models can be used to make predictions about the future course of a viral bank run, such as estimating the number of posts and comments, identifying peak infection periods, and evaluating the effectiveness and impact of different intervention strategies and control measures. By adjusting the model parameters and running simulations, researchers can simulate various scenarios and assess the potential outcomes and make informed decisions to mitigate the spread of rumors.
Viral bank runs can be influenced by underlying bank instability and weak foundations. However, it is possible to mitigate their impact and even prevent them through proactive monitoring of information spread. Indeed, a single negative post can create panic and cause a cascade effect of similar posts, which might push people to withdraw cash from their accounts. To prevent future viral bank runs, it is crucial for managers to recognize the power and significance of social media. They should not underestimate its impact and instead prioritize its role in communication strategies. Monitoring social media conversations about the bank is fundamental, as it allows for the quick identification of negative sentiment and potential concerns. Rapid response is key when it comes to addressing negative information. One approach is to identify influential individuals, in our case such as those with high degree and provide them with accurate and up-to-date information about the bank's financial stability, security measures, and deposit protection mechanisms, trust and credibility can be built in order mitigate panic and reassure users.

References

Jin, F., Dougherty, E. R., Saraf, P., Cao, Y., & Ramakrishnan, N. (2013). Epidemiological modeling of news and rumors on Twitter. In Social Network Mining and Analysis. https://doi.org/10.1145/2501025.2501027

Zhao, L., Cui, H., Qiu, X., Wang, X., & Wang, J. (2013). SIR rumor spreading model in the new media age. Physica A: Statistical Mechanics and its Applications, 392(4), 995-1003. https://doi.org/10.1016/j.physa.2012.09.030

Zhao, L., Wang, J., Chen, Y., Wang, Q., Cheng, J., & Cui, H. (2012). SIHR rumor spreading model in social networks. Physica A: Statistical Mechanics and its Applications, 391(7), 2444-2453. https://doi.org/10.1016/j.physa.2011.12.008

Zan, Y., Wu, J., Li, P., & Yu, Q. (2014). SICR rumor spreading model in complex networks: Counterattack and self-resistance. Physica A: Statistical Mechanics and its Applications, 405, 159-170. https://doi.org/10.1016/j.physa.2014.03.021

Kumar, S., Saini, M., Goel, M., & Aggarwal, N. (2020). Modeling Information Diffusion In Online Social Networks Using SEI Epidemic Model. Procedia Computer Science, 171, 672-678. https://doi.org/10.1016/j.procs.2020.04.073

Yerushalmy, J. (2023, March 18). 'The first Twitter-fuelled bank run': how social media compounded SVB's collapse. The Guardian. https://www.theguardian.com/business/2023/mar/16/the-first-twitter-fuelled-bank-run-how-social-media-compounded-svbs-collapse

Pastor-Satorras, R., & Vespignani, A. (2001). Epidemic Spreading in Scale-Free Networks. Physical Review Letters, 86(14), 3200-3203.

Newman, M. E. (2003). The Structure and Function of Complex Networks. SIAM Review, 45(2), 167-256.

Vespignani, A. (2009). Epidemic Processes in Complex Networks. Nature Physics, 6(7), 487-488.

Morris, M. (2004). Network Epidemiology: A Handbook for Survey Design and Data Collection. Oxford University Press.