Financial Contagion: Applying a Networked Disease Model to Great Depression-era Bank Runs

Wesleyan University

Lucas Holman '29
Summer Apprenticeship 2026

Faculty Sponsor: Pavel Oleinikov
Quantitative Analysis Center

Abstract

This study examines modeling financial panic as a contagion using a discrete dynamical system. We construct a variant of the Susceptible-Infected-Recovered (SIR) disease model to capture the spread of financial instability native to pre-FDIC America. We webscrape the Historical Bank Runs Database, assembled by the New York Federal Reserve using LLMs, for historical data of American bank runs in the years 1929-1933. Using this information, we model the banks as an incomplete, undirected, weighted graph, with banks mutually linked by an edge weighted to the inverse-square of the banks' physical distance from each other. Hence, we construct a set of difference equations according to the graph's adjacency matrix, which correspond to the three SIR compartments and represent the probability of the banks' appearances in each.

Upon iterating the dynamical system over many time-steps, we find that our epidemiological model has a strong correlation coefficient (Pearson R = 0.76) with the historical data, though it explains little of the bank runs' variability (R^2 = 0.13). Rather, financial contagion embodies a "slow-burn" takeover of the banking system, and not the sudden spikes in failures characteristic to the Great Depression. We attribute this to the simple implementation of geographic position as a proxy for more sophisticated connections between banks, thus crediting a more complex model of financial contagion.

Introduction

Bank runs are an economic phenomenon which frequently took place in the United States prior to the establishment of the Federal Deposit Insurance Corporation (FDIC) in 1933. Bank runs occurred when bank depositors simultaneously attempted to withdraw large sums of money out of fear of the bank's insolvency, often culminating in bank failure. Banks prefer to lend out the vast majority of their holdings in order to earn interest on the loans and they thus hold little liquid cash on hand, meaning that a large volume of withdrawals within a short timespan can spell disaster for a bank, as they may fail to cough up the cash and collapse. Moreover, because neighboring banks regularly make loans to each other, a run on any one bank (and its possible subsequent failure) may inspire doubt in the solvency of its peer banks, potentially fostering additional runs on those banks. This chain-reaction spreads in a form akin to diseases, and thus can be modeled similarly as a financial contagion.

Methods & Equations

$$S_i(t+1) = S_i(t) - \Lambda_i(t)S_i(t) + \gamma(1-\Phi)I_i(t)$$
$$I_i(t+1) = I_i(t) + \Lambda_i(t)S_i(t) - \gamma I_i(t)$$
$$R_i(t+1) = R_i(t) + \gamma\Phi I_i(t)$$

$$\hat{A}_{ij} = \frac{A_{ij}}{\sum_k A_{ik}}$$

$$\Lambda_i(t) = \beta \textstyle\sum_j \hat{A}_{ij} I_j(t)$$

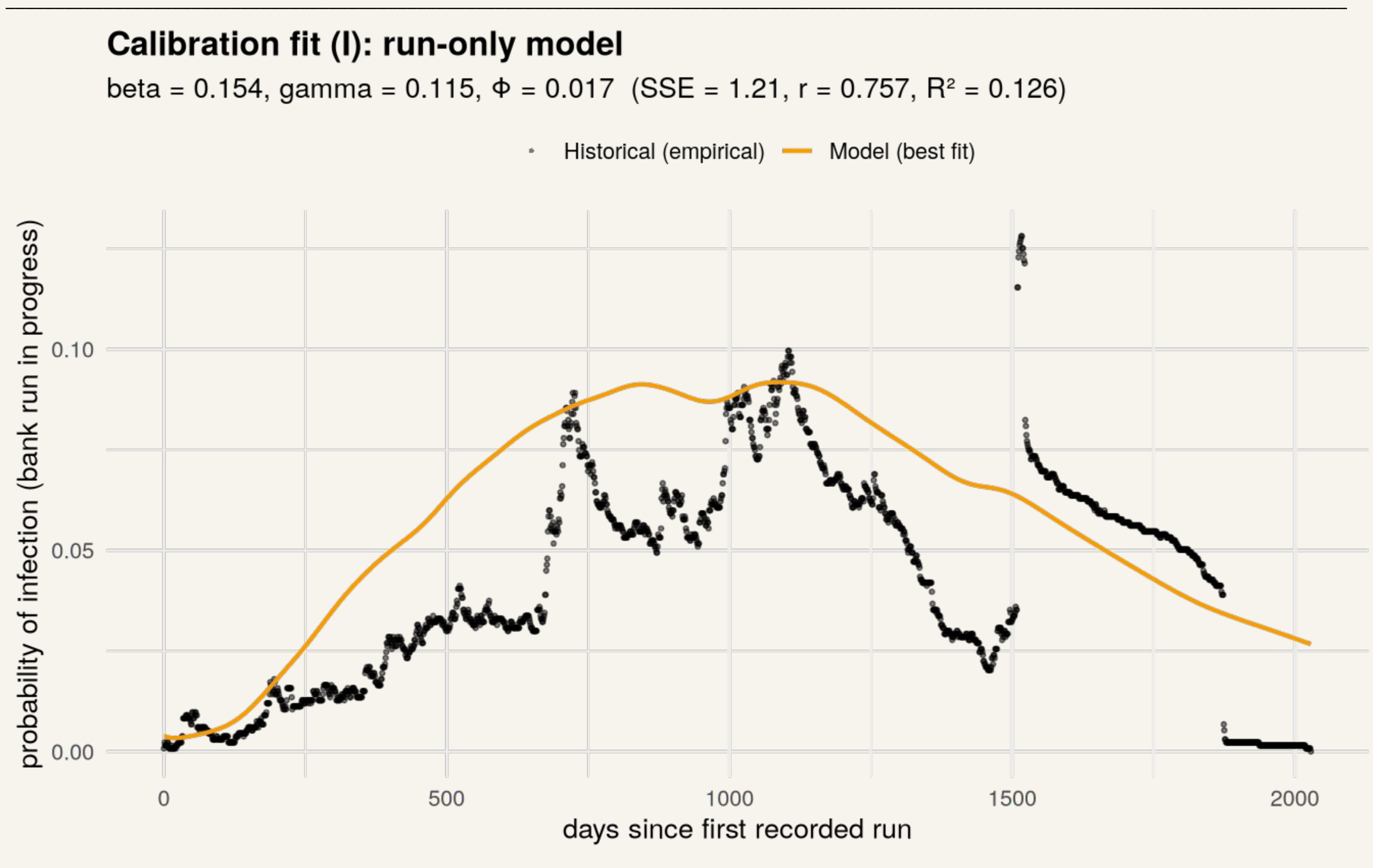
β represents the contagion strength;
γ represents the rate of recovery (i.e. γ^{-1} is the average length of a bank run);
ϕ represents the rate of bank-failure as a result of a run.
$\hat{A}$ is the row-normalized form of A;
Λ_i represents the total "pressure" on the i^{th} bank from its "infected" neighboring banks.

These parameters are calibrated by minimizing squared error between the computed and empirical curves' I-compartments across many runs of the dynamical system.

We define three compartments that a bank can be in, taken from standard epidemiological methods:

- "Susceptible" (S_i) represents the probability that the i^{th} bank is not currently experiencing a run, but is susceptible to one;
- "Infected" (I_i) represents the probability that the i^{th} bank is actively undergoing a run;
- "Recovered" (R_i) is a misnomer (a relic of the model's epidemiological origins), instead representing the probability that the i^{th} bank has collapsed, and is thus no longer susceptible.

Hence, using those three compartments, we construct a discrete dynamical system to compute the probability of the i^{th} bank's appearance in each. This is achieved through a set of difference equations, a model variant known as an SIRS model; upon a run on the bank (I), banks can either transition back to susceptibility (S) or are absorbed permanently as collapsed banks (R). The difference equations rely on A, an adjacency matrix representing an undirected, weighted graph connecting the nodes (banks). Utilizing K-nearest neighbors, each bank becomes linked to its 8-nearest neighbors, with the edges weighted to the inverse-square of their geographic distance, and the remaining edges weighted by 0. Additionally, we compute the Minimum Spanning Tree (MST) of the graph and union it into our adjacency matrix, ensuring that the graph has no disconnected components and thus avoids containing contagion spread entirely within clusters of banks. Furthermore, we compute Λ_i as the total "pressure" on the i^{th} bank from its neighboring banks. This culminates in the above difference equations, with one-day time-steps.



Model-predicted vs. historical probability of an active bank run over time (Pearson r = 0.757, R^2 = 0.126).

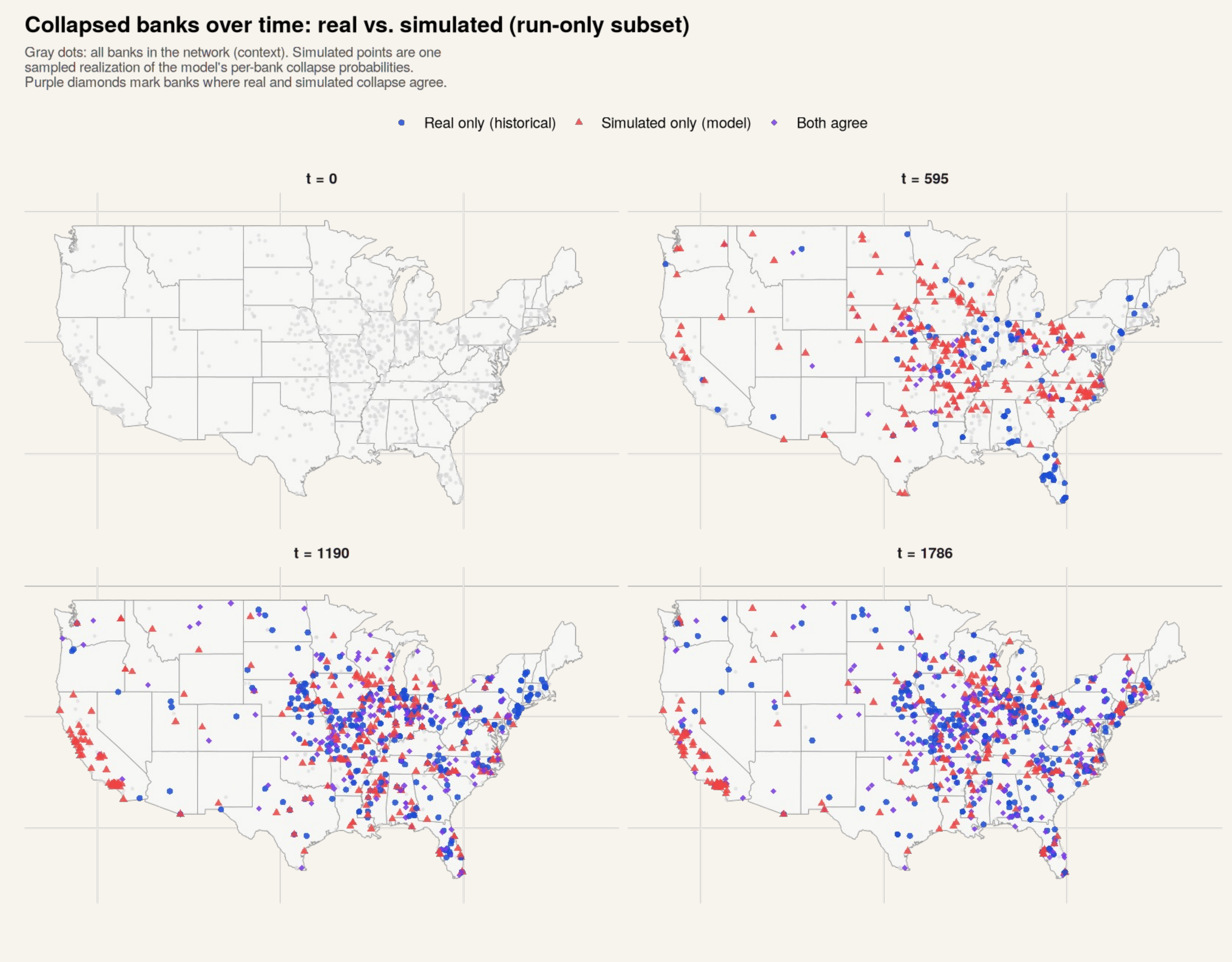
Data

We webscrape the Historical Bank Runs Database, assembled by the New York Federal Reserve using LLMs on the American Stories dataset, for historical data of American bank runs and episodes in the years 1929-1933. We restrict ourselves to a limited subset of events, less than 20% of the total observations scraped, as many events represent episodes other than bank runs (i.e. suspensions) and do not constitute as runs. Some events contain record-keeping artifacts (e.g. geographic coordinates listed as (0,0), missing bank names, etc.), which are also dropped for convenience. This yields us a usable collection of 1,335 historical runs. This includes features such as the bank name, geographic location, date of the run, resolution of the run, etc. which are then used to calibrate the parameters of the SIRS model over many iterations of the dynamical system.

Parameter / Statistic	Value	Description
β	0.154	Fitted, I-curve target
γ	0.115	Implies ~8.68 day mean run-duration
Φ	0.017	Fitted, I-curve target
SSE	1.214	Sum of squared error, I-curve fit
Pearson r	0.757	—
Spearman r	0.772	—
R²	0.126	—
N	1,335	Run-only subset

Findings & Interpretations

Though the calibrated SIRS model returns an extremely low p-value when compared to the historical dataset, this is quickly discarded, as a dynamical system of this granularity and degree of autocorrelation is bound to produce significant p-values, no matter what. Rather, we look at the Pearson correlation ($R = 0.76$) and coefficient of determination ($R^2 = 0.13$). The large difference in the two informs us that the model succeeds in representing a larger trend, but struggles to represent the variability that comes from successive waves of bank runs. A qualitative analysis produces something similar: SIRS succeeds at representing the spread of financial panic as a gradual, expansive force, but suffers in capturing the spikes from the real data. Furthermore, upon holding our parameters constant at $\beta = 0.154, \gamma = 0.115$ and varying $\phi \in [0.05, 0.95]$, the maximum cumulative collapse proportion returns only 45.3%, far from the true cumulative collapse proportion of 53.6%. This suggests that the model-infection's ability to spread quickly is kneecapped, most likely by our reliance on geographic distance as a simple proxy for more sophisticated connections between banks (i.e. loan structures, telegram messaging, rumor circulation, etc.). However, given the constraints, we find that this model is relatively successful at capturing a general, national spread of bank runs. Moreover, our data is a subset of a much larger list of historical banks and banking episodes, which could be used to further optimize the model.



Blue: real historical collapses. Red: simulated collapses (one sampled realization of the model's per-bank probabilities). Purple: real and simulated agree.

References

Correia, Sergio, Stephan Luck, and Emil Verner. 2026. "Bank Runs With and Without Bank Failure." Working Paper.

Cho, Changhee, Okyu Kwon, KyoungEun Lee, and Jae Woo Lee. "Commuter metapopulation models for epidemic spreading in human mobility networks." Chaos: an Interdisciplinary Journal of Nonlinear Science 35, no. 10 (October 1, 2025).

Acknowledgements

I would like to give my special thanks to my faculty sponsor, Prof. Pavel Oleinikov, for his advice, mentorship, and care, as well as Prof. Manolis Kaparakis as Director of the Quantitative Analysis Center and mentor. I would like to give further thanks to the entire faculty of the QAC for their instruction. I would also like to give thanks to my father, Dr. Matthew J. Holman, for his advice and consultation, both technical and inspirational.