The Distribution of Outcomes for a Networked Economy

Janelle Schlossberger

Harvard University

December 2018

JOB MARKET PAPER

Click here for most recent version

Abstract

This work develops a set of mathematical tools that map the topology of an economic network to a probability distribution of possible outcomes for the economy. We can apply these tools to analyze complex economic systems in closed form and to construct error bounds about the paths of aggregated networked economies. To generate this mapping from network topology to probability distribution, we focus on a class of economies that has the following three features: (1) a population of $N$ agents, each with a binary-valued attribute, (2) a network on which these $N$ agents are organized, and (3) decision-making by each networked agent that depends on the local relative frequency of the attribute. This class of economies also has an aggregate feature: the global relative frequency of the attribute. Given the system’s aggregate feature, underlying network, and population size, we construct in closed form the distribution of possible local relative frequencies of the attribute. The topology of the network determines the extent to which the local relative frequency of the attribute can deviate from its global relative frequency, thereby determining the extent to which the outcome of the economy can deviate from a benchmark outcome. Given this distribution and agents’ decision-making behavior, we then construct the distribution of possible outcomes for the economy. For realistic agent interaction structures featuring a very large population of agents, the distribution of outcomes is meaningfully non-degenerate. We adapt the theoretical framework and mathematical tools developed in this work to study locally formed macroeconomic sentiment and how agents’ interaction structure shapes the capacity for there to exist non-fundamental swings in aggregate sentiment, with implications for election outcomes and for our understanding of animal spirits.

JEL Classification: D85, C40, C60, E00, D72

Key Words: network, economic system, configuration, uncertainty modeling, distribution, higher-order moments, macroeconomic sentiment, animal spirits, voting

*Contact: jschlossberger@hbs.edu. The Supplementary Materials are available on the author’s webpage, https://scholar.harvard.edu/janelle-schlossberger/publications. I thank my committee members, Emmanuel Farhi, Xavier Gabaix, David Laibson, and Tomasz Strzalecki, for helpful comments, and I also thank seminar and conference participants at Harvard University, MIT LIDS & Stats Tea, the 2018 Royal Economic Society Annual Conference, the 2018 North American Summer Meeting of the Econometric Society, the 2018 SIAM Workshop on Network Science, and the 9th International Conference on Complex Systems.
1 INTRODUCTION

In this work, we study an economy that has a population of agents organized on a network. Given the features of the economy’s agents, the topology of the underlying network, and agents’ decision-making behavior, we are interested in the distribution of possible outcomes for the economy. The existence of a network structure introduces complications into the economic system; the outcome of the economy is inherently dependent on the topology of the underlying network, but it is not immediately apparent how the topology shapes the distribution of outcomes. This work provides that mapping from network structure to probability distribution.

The main contribution of this work is that it introduces mathematics that enables us to derive in closed form the probability distribution of possible outcomes for the economy from the topology of agents’ interaction network. The high-level innovation of this work is that it develops a set of tools that mathematically links two fields: networks to statistics, or more precisely, networks to probability distributions. The topology of the network directly affects the shape of the probability distribution of outcomes, and this work makes that relationship explicit. We can carry out this mapping from network to probability distribution for all feasible network topologies. The economy in which the network is embedded also affects the mapping, and we take the features of the economy into account.

To develop this mapping, we focus on a class of economic systems that has three distinguishing characteristics. First, the economy has a population of \( N \) agents, each of whom has a binary-valued attribute. Second, these \( N \) agents are organized on a network. Third, each networked agent’s decision-making depends on the local relative frequency of the attribute. Since agents have different positions on the network, they potentially have different local relative frequencies of the attribute arising from different network neighborhoods, and this can lead the agents to make different decisions. Each economy in this class also has an aggregate feature. That aggregate feature is the global relative frequency of the attribute. There are \( n \leq N \) agents with the attribute’s unit value, so the global relative frequency of the attribute is \( f = \frac{n}{N} \). The exogenous objects in this work are the population size, the underlying network structure, the global relative frequency of the attribute, and agents’ decision-making behavior, while our endogenous object of interest is the probability distribution of possible outcomes for the economy.

In this work, we are mapping the topology of agents’ interaction network to a distribution of outcomes for the economy. In general, we have a non-degenerate probability distribution because there is more than one possible outcome for the economy. This multiplicity of outcomes arises because there are combinatorially many possible configurations, or
arrangements, of the binary-valued attribute among agents consistent with the attribute’s
global relative frequency. When $f = \frac{n}{N}$, there are $\binom{N}{n}$ possible configurations. We can
imagine that the outcome of the economy changes with the particular configuration of the
attribute. As the configuration changes, a different subset of agents has the attribute’s unit
value, which generates a potential adjustment to the local relative frequency of the attribute
for each agent. Agents choose actions based on that local relative frequency of the attribute,
so as the configuration changes, we potentially have a shift in agents’ actions, which leads
to a different outcome for the economy.

The two objects that we focus on in this work are the local relative frequency of
the attribute and the outcome for the economy. Holding fixed the attribute’s global relative
frequency, there are combinatorially many possible configurations and for each configuration,
there is an associated local relative frequency of the attribute. Therefore, given the attribute’s
global relative frequency and the structure of the underlying network, we can construct an
entire probability distribution of possible local relative frequencies of the attribute. We
refer to the distribution of possible local relative frequencies of the attribute as a precursor
distribution because its construction precedes our construction of the distribution of possible
outcomes for the economy. Once we have computed the precursor distribution, we can
then construct the distribution of possible outcomes for the economy given agents’ decision-
making behavior. Our precursor distribution characterizes the extent to which the attribute’s
local relative frequency deviates in either direction away from its global relative frequency.
The capacity for such variation in the attribute’s local relative frequency depends on the
underlying network structure, and it determines the extent to which the outcome of the
economy can deviate from a benchmark outcome. If we have sufficient variation in the local
relative frequency of the attribute, then there is an entire non-degenerate distribution of
possible outcomes for the economy.

We are interested in characterizing the distribution of possible outcomes for the econ-
omy, but we have not yet made explicit what an outcome is exactly. The outcome for an
economy is situational. For example, it might be the aggregate action taken by all agents in
the population, or it might instead be the action of a single agent of interest. Alternatively,
the outcome of the economy might follow from the outcome of an event; for instance, it might
follow from the outcome of an economy-wide political election with two possible candidates.
In such a setting, there would be two possible outcomes for the economy, and the probability
that each outcome occurs is equal to the probability that the corresponding candidate wins
the election.

---

1 The benchmark outcome is the one that results if we ignore the underlying configuration and only take
into account the aggregate properties of the system, namely the attribute’s global relative frequency.
The main technical contributions of this work are as follows. We characterize the shape and properties, including the higher-order statistical features, of our precursor distribution for every feasible network structure, population size, and global prevalence of the binary-valued attribute. We determine those network topologies for which the local relative frequency of the attribute is invariant to configuration, which makes the precursor distribution degenerate. More generally, we statistically characterize the precursor distribution when every configuration is equally likely and when every configuration occurs with some arbitrary probability. Once we have characterized this distribution of possible local relative frequencies of the attribute in full generality, we then study the distribution of possible outcomes for the economy. For certain classes of agent actions, we can provide a closed-form representation of this distribution of possible outcomes. We can characterize this distribution of outcomes for all feasible network structures, population sizes, and global prevalences of the attribute in the population. To the extent that there is variation in the local relative frequency of the attribute across configurations, there can then be significant variation in the economy’s outcome, holding \( f \) fixed.

When our probability distribution of possible outcomes is non-degenerate, meaning that the outcome of the economy varies with configuration, we consider the economic system to be configuration dependent. Configuration dependence enables the existence of phenomena that would otherwise not occur if we only considered the system’s aggregate features. It adds richness to our models of the economy because there is an entire distribution of possible outcomes consistent with our uniquely valued aggregate feature. If we ignored the inherent configuration dependence of the economy, then there would only be one possible outcome.

The tools that we develop in this work enable us to form insights into: (1) complex economic systems and (2) aggregated economies. First, the mathematical machinery allows us to unpack complex systems, and more specifically, complex economic systems featuring network-based agent interaction. Economies with agent-based interaction are quite complicated, as there are myriad ways that these systems can possibly evolve. The tools developed in this work enable the closed-form analysis of such complex economic systems. We can collapse the complexities of agent-based interaction into a simple probability distribution that characterizes how the system will evolve.

Second, these tools allow us to assess aggregate treatments of economic systems and quantify their incompleteness. The class of models that we study in the present work has features that enable direct comparison with aggregate models of the economy. Here, we have a population of \( N \) agents whose decision-making behavior can be aggregated, and there is an aggregate feature, \( f \), built up from the attributes of the underlying set of agents. Aggregate models of the economy similarly feature a population of \( N \) agents whose actions
can be aggregated, and the corresponding representative agent makes decisions based on the aggregated characteristics of the system. In such aggregate models of the economy, the action taken by the representative agent given the system’s aggregate characteristics is unique; there is a single outcome. However, for the class of models in the present work, even though the economy has a parallel structure, there is an entire non-degenerate distribution of possible outcomes for the economy that is consistent with the system’s uniquely valued aggregate feature, $f$.

This work shows how aggregate treatments of economic systems can lead to characterizations that are, in general, incomplete. Rather than the aggregate economic system having a unique outcome determined by the system’s aggregate features, there is instead an entire distribution of possible outcomes centered about that original benchmark outcome. Using the tools developed in this work, we can introduce a configurational error bound and place that error bound about the benchmark outcome of the aggregate economy to account for the multiplicity of possible outcomes. In particular, we construct this error bound for aggregated systems with networked agents who make local decisions. The size of this error bound depends on the underlying network’s topology. By incorporating this error bound, we allow for a more complete and a more nuanced understanding of the phenomena that aggregate models of the economy seek to study. As a result, two systems with the same aggregate features can evolve differently due to differences in their underlying configurations; the configurational error bound that we construct accounts for this variation in the two economic outcomes relative to each other and relative to the benchmark outcome.

We use these theoretical findings to study locally formed macroeconomic sentiments, election outcomes, and animal spirits. In our applied setting, the binary-valued attribute denotes employment status, and each agent makes a voting decision that depends on his or her local unemployment rate. This local unemployment rate is a proxy for individual macroeconomic sentiment and the average local unemployment rate is a proxy for aggregate macroeconomic sentiment. In our model, the fundamentals of the economy, namely the economy’s global unemployment rate, alone favor the election of one candidate with certainty. However, if the overall level of macroeconomic sentiment in the economy sufficiently varies with the underlying configuration of unemployment, then there can be more than one possible election outcome and more than one possible outcome for the economy. In a setting with 137.5 million agents, which is the number of voters in the 2016 U.S. presidential election, we find that the distribution of average local unemployment rates is strongly non-degenerate. The variation in this average local unemployment rate is sufficiently large that it can actually mimic variations in business cycle conditions. As a result, the election outcome depends on the particular configuration of unemployment. Such non-degeneracy of the distribution of
outcomes for very large $N$ emerges from both high variance of in-degrees and heavy-tailedness of weighted in-degrees for the calibrated social observation network.

By assuming that agents form macroeconomic sentiment from their local unemployment rates, we are able to quantify the extent to which aggregate sentiment is positive or negative given the economy’s fundamentals; aggregate sentiment is positive, that is, there are waves of optimism, when the average local unemployment rate is less than the global unemployment rate, while aggregate sentiment is negative, that is, there are waves of pessimism, when the average local unemployment rate is greater than the global unemployment rate. We can quantify this deviation in the average local unemployment rate from the actual unemployment rate and therefore quantify deviations in aggregate sentiment away from a level that is commensurate with the economy’s fundamentals. We show how the underlying interaction structure among agents in the economy shapes the capacity for there to exist these non-fundamental swings in aggregate sentiment for all population sizes. We thus offer a mechanism for the formation of individual and aggregate macroeconomic sentiment that essentially microfounds animal spirits.

1.1 Relation to the Literature

This work interfaces with four different strands of the literature: (1) complex economic systems, (2) networks, (3) aggregation, and (4) macroeconomic sentiment. Research in the area of complex economic systems includes Granovetter (1978), Brock and Durlauf (2001), Bisin et al. (2004), and Horst and Scheinkman (2004). These works all feature some form of agent interaction, namely agents choosing actions that depend on the actions of other agents. These works take great care to establish the existence of equilibria in such settings. For these works, the equilibrium outcome is the object of study. The present work meanwhile has a different focus; the object of interest in the present work is the distribution of possible outcomes for the system.

This work contributes to research on networks by developing a set of tools that allows us to mathematically link the field of networks to the field of statistics, and in particular, the area of probability distributions. This work is innovative in that regard. We can also relate the present work to recent research on network-based social learning. Recent papers in that area include Gale and Kariv (2003), Golub and Jackson (2010), Acemoglu, Ozdaglar, and ParandehGheibi (2010), Acemoglu, Dahleb, Lobel, and Ozdaglar (2011), Banerjee, Breza, Chandrasekhar, and Mobius (2016), Harel et al. (2017), and Chandrasekhar, Larreguy, and Xandri (2018). The present work provides theoretical results that enhance our understanding of DeGroot learning. Through the tools developed in the present work, we are able to
construct the entire distribution of possible consensus learned values, including the higher-order features of this distribution, for any population size, and we can determine how the topology of the network shapes the capacity for there to be learning and mis-learning.

Research in the area of aggregation tends to examine whether aggregate fluctuations in output can arise from micro-level shocks, or if an aggregate parameter is needed in models of the macroeconomy to generate sufficiently sizable aggregate fluctuations. Papers include Bak, Chen, Scheinkman, and Woodford (1993), Scheinkman and Woodford (1994), Horvath (1998, 2000), Gabaix (2011), and Acemoglu, Carvalho, Ozdaglar, and Tahbaz-Salehi (2012). These papers offer different mechanisms that slow the rate at which the law of large numbers applies. They study the role of sectoral production networks, granularity in firm size, and non-convexities in production technologies coupled with local interaction among sectors to generate adequate fluctuations in output that persist even as the economy becomes increasingly disaggregated. Recently, research in this area has also been trying to study how microeconomic shocks shape the higher-order features of the output distribution; recent work includes Acemoglu, Ozdaglar, and Tahbaz-Salehi (2017) and Baqee and Farhi (2018). The present work tackles this issue of aggregation as well. It examines the extent to which the distribution of possible outcomes for the economy remains non-degenerate in a large-$N$ setting. We develop theoretical results that characterize the variance and the CDF of this distribution, including its higher-order features, for every possible population size and network topology, including the limit as $N \to \infty$. We can explicitly examine how the topological features of the interaction network shape the capacity for the distribution of outcomes for the economy to remain approximately non-degenerate, even for large $N$.

Recent research in the area of macroeconomic sentiment includes Barsky and Sims (2012), Angeletos and La’O (2013), Benhabib, Wang, and Wen (2015), Huo and Takayama (2015), Acharya, Benhabib, and Huo (2017), Angeletos, Collard, and Delias (2017), and Milani (2017). In the theoretical domain, these works define mathematically what it means for economies to have sentiment, consumer confidence, and/or animal spirits. In the empirical domain, the literature has tried to determine the extent to which realistically calibrated shocks to sentiment and/or consumer confidence can impact and generate reasonable fluctuations in macroeconomic aggregates. The present work interacts with the existing literature by providing a simple mechanism for the formation of macroeconomic sentiment among agents. This mechanism allows us to quantify the extent to which individual sentiment and aggregate sentiment deviates from a level that is commensurate with economic fundamentals. The present work also shows how the underlying interaction structure among agents shapes the capacity for there to exist non-fundamental swings in aggregate sentiment.
1.2 Outline of Paper

Section 2 provides notation and definitions, introduces the class of problems that we later mathematically solve, and works through two illustrative examples. Section 3 applies this class of problems towards understanding macroeconomic sentiments and political election outcomes. It studies how there can be sizable configuration-induced variations in macroeconomic sentiment in a large-\(N\) economy for fixed economic fundamentals and the resulting impact on election outcomes. After exploring this application, Section 4 begins to develop the mathematics that enables us to solve our class of problems. Section 5 first characterizes the null setting in which the particular configuration of the attribute among agents is irrelevant. It identifies those conditions for which the distribution of possible local relative frequencies of the attribute is either degenerate or invariant to configuration. Sections 6 and 7 then present the tools that enable us to characterize the distribution of possible local relative frequencies of the attribute when it is non-degenerate and every configuration is either equally or not equally likely to occur. Section 8 studies the distribution of possible outcomes for the economy given agents’ decision-making behavior, and Section 9 concludes.

2 Model

We begin by introducing the notation and definitions that will be used throughout this paper. We then proceed to develop our guiding theoretical framework, highlighting the objects of interest that emerge. We conclude this section by working through two examples that make the theoretical framework and the objects of interest even more precise in an applied setting.

2.1 Notation and Definitions

The cardinality of a set \(\mathcal{X}\) is \(|\mathcal{X}|\). A multiset is an object similar to a set, but it allows for multiple instances of each of its elements. Vector \(\mathbf{x}\) is a column vector by default. The \(i^{th}\) element of vector \(\mathbf{x}\) is \(x_i\) or \([\mathbf{x}]_i\). The \(ij^{th}\) element of matrix \(\mathbf{X}\) is \([\mathbf{X}]_{ij}\), the \(i^{th}\) row of \(\mathbf{X}\) is \([\mathbf{X}]_i\), and the \(j^{th}\) column of \(\mathbf{X}\) is \([\mathbf{X}]_j\). The identity matrix is \(\mathbf{I}\), the column vector whose elements all equal 1 is \(\mathbf{1}\), and the unit vector \(\mathbf{e}_i\) has \([\mathbf{e}_i]_j = 1\) for \(i = j\) and \([\mathbf{e}_i]_j = 0\) otherwise. Matrix \(\mathbf{X}\) is row-stochastic if \(\mathbf{X}\mathbf{1} = \mathbf{1}\) and all matrix elements of \(\mathbf{X}\) are non-negative. Matrix \(\mathbf{X}\) is doubly stochastic if it is both row-stochastic and column-stochastic, that is, \(\mathbf{X}\mathbf{1} = \mathbf{1}, \mathbf{X}^T\mathbf{1} = \mathbf{1}\), and all matrix elements of \(\mathbf{X}\) are non-negative. Non-negative matrix \(\mathbf{X}\) is primitive if there exists an integer \(q \geq 1\) such that \([\mathbf{X}^q]_{ij} > 0\) for all matrix elements in \(\mathbf{X}^q\). \(x(N) \sim y(N)\) w.h.p. (that is, \(x(N)\) is asymptotically equivalent to \(y(N)\)
with high probability) if \( \Pr \left( \frac{x(N)}{y(N)} \to 1 \right) \to 1 \) as \( N \to \infty \). \( x(t) = o(y(t)) \) if and only if, for every \( \alpha > 0 \), there exists a real-valued constant \( t_0 \) such that \( |x(t)| \leq \alpha |y(t)| \) for all \( t \geq t_0 \). \( x(t) = \omega(y(t)) \) if and only if, for every \( \alpha > 0 \), there exists a real-valued constant \( t_0 \) such that \( |x(t)| \geq \alpha |y(t)| \) for all \( t \geq t_0 \). Graph \( G \) is an ordered pair \( G = (V, E) \) consisting of a set of vertices (nodes) \( V \) and a set of edges \( E \). \( (x, y, e_{x,y}) \in E \) is an edge between nodes \( x \) and \( y \) with weight \( e_{x,y} \). If the graph is directed, the edge is oriented from node \( x \) to node \( y \); otherwise, the edge is not oriented. \( G(X) \) refers to an unweighted graph with unweighted adjacency matrix \( X \), whose non-zero elements are \( [X]_{ij} = 1 \), and \( G(\bar{X}) \) refers to a weighted graph with weighted adjacency matrix \( \bar{X} \), whose non-zero elements are \( [\bar{X}]_{ij} = e_{i,j} \).

### 2.2 Theoretical Framework

We now develop the theoretical framework that motivates and guides this paper. Consider an economic system with \( N \) total networked agents. Each agent \( i \) has a binary-valued attribute, \( b_i \), with either \( b_i = 0 \) or \( b_i = 1 \); \( n \leq N \) agents have \( b_i = 1 \), the attribute’s unit value. The global relative frequency of the attribute’s unit value in the total population of agents is \( f = \frac{n}{N} \). This quantity, \( f \), is the economy’s aggregate feature. In the rest of this work, we refer to \( f \) as the attribute’s global relative frequency. Given \( f \), there is a particular configuration, or arrangement, of the binary-valued attribute among agents in the economy. We define such a configuration as follows:

**Definition 1** A configuration \( \mathbf{b} \equiv \mathbf{b}(N, n) \) of a binary-valued attribute in a population of \( N \) agents is an allocation of the attribute so that \( b_i \in \{0, 1\} \) for all \( i \in \{1, \ldots, N\} \) and \( \mathbf{b}^T \mathbf{1} = n \).

A configuration \( \mathbf{b} \equiv \mathbf{b}(N, n) \) of the binary-valued attribute among agents in the population is an allocation such that every agent has the attribute’s zero or unit value, and the global relative frequency of the attribute in the population is \( f = \frac{n}{N} \). We construct the \( N \times 1 \) configuration vector by taking each agent’s attribute, \( b_i \), and stacking this value for all agents in the population. From this vector, we can identify those agent indices with \( b_i = 1 \). Two configurations \( \mathbf{b}, \mathbf{b}' \) are distinct if and only if \( \mathbf{b} \neq \mathbf{b}' \) because the agent indices with \( b_i = 1 \) differ across these two configurations. We denote \( \mathcal{B}(N, n) \) as the set of all possible configurations consistent with \( f = \frac{n}{N} \), and the cardinality of this set is \( |\mathcal{B}(N, n)| = \binom{N}{n} \).

Agents in this setting interact, and a network and its accompanying adjacency matrices capture these patterns of interaction. The \( N \times N \) unweighted adjacency matrix \( \mathbf{A} \) captures the existence of linkages among agents; \([\mathbf{A}]_{ij} = 1\) if there is an edge from agent \( i \) to agent \( j \). Meanwhile, the \( N \times N \) weighted adjacency matrix \( \bar{\mathbf{A}} \) captures the weights that

---

2In this work, networks can feature self-loops but not multiple edges.
agents assign to these linkages. $[\tilde{A}]_{ij} = e_{i,j}$ if the network has edge $(i, j, e_{i,j})$ from agent $i$ to agent $j$ with edge weight $e_{i,j}$. Agents allocate non-negative weight to each of their linkages, with the total weight allocated by a particular agent summing to 1; for an agent $i$, we therefore have $\sum_{j=1}^{N} [\tilde{A}]_{ij} = 1$.\(^3\) This summation holds for all agents $i \in \{1, \ldots, N\}$, which makes matrix $\tilde{A}$ row-stochastic.

We are interested in the local relative frequency of the attribute’s unit value, $x(\tilde{A}, b, N, n)$, for every configuration $b(N, n) \in \mathcal{B}(N, n)$. In the rest of this work, we refer to $x(\tilde{A}, b, N, n)$ as the attribute’s local relative frequency. This quantity depends on agents’ interaction structure, $\tilde{A}$, it depends on the global frequency, $n$, of the attribute’s unit value in the population of size $N$, and it depends on which subset of agents on the network actually has that unit value, $b$. Holding $f$ fixed, as the configuration of the attribute adjusts and a different subset of agents has the attribute’s unit value, we can imagine that $x(\tilde{A}, b, N, n)$ changes as well. Figure 1 plots the global and local relative frequencies of a binary-valued attribute across three different configurations. For all of these configurations, the attribute’s global relative frequency is the same: $f = 0.50$. However, the local relative frequency of the attribute varies; it depends on which subset of agents actually has the attribute’s unit value. For each configuration, this local relative frequency can also meaningfully deviate from the attribute’s global relative frequency. For example, when agents 3 and 4 have the attribute, the local relative frequency of the attribute deviates positively, while when agents 1 and 2 have the attribute, the local relative frequency of the attribute deviates negatively. Scalar quantity $x(\tilde{A}, b, N, n)$ is an abstract object; we are only able to directly compute this object once we assign to it a specific interpretation.

Given $x(\tilde{A}, b, N, n)$ for every configuration $b(N, n) \in \mathcal{B}(N, n)$, we then construct the distribution of possible local relative frequencies of the attribute. We define random variable $X(\tilde{A}, N, n)$ whose realizations are the configuration-specific quantities $x(\tilde{A}, b, N, n)$. We are interested in the distributional features of $X(\tilde{A}, N, n)$. $G_X(\tilde{A}, N, n)(t)$ is the CDF of $X(\tilde{A}, N, n)$ and $g_X(\tilde{A}, N, n)(t)$ is the PMF of $X(\tilde{A}, N, n)$. If every configuration is equally likely to occur:

$$G_X(\tilde{A}, N, n)(t) = \frac{1}{|\mathcal{B}(N, n)|} \sum_{b(N, n) \in \mathcal{B}(N, n)} 1_{x(\tilde{A}, b, N, n) \leq t},$$

where $G_X(\tilde{A}, N, n)(t)$ represents the fraction of configurations for which $x(\tilde{A}, b, N, n) \leq t$. $G_X(\tilde{A}, N, n)(t)$ is the precursor distribution from which we then proceed to construct the distribution of possible outcomes for the economy.

\(^3\)Schlossberger (2018) relaxes the assumption that edge weights are non-negative and that the total allocated weight must sum to 1.
If the total number of configurations, $|\mathcal{B}(N, n)|$, is small, we can construct $G_X(\bar{A}, N, n)(t)$ given $f$ by computing $x(\bar{A}, b, N, n)$ configuration by configuration. In general, though, for $|\mathcal{B}(N, n)|$ small or large, we can construct $G_X(\bar{A}, N, n)(t)$ by decomposing $x(\bar{A}, b, N, n)$ into two constituent quantities:

$$x(\bar{A}, b, N, n) = [w(\bar{A})]^T b(N, n).$$

The first quantity is a fixed, network-derived vector of agent weights, $w(\bar{A})$, and the second quantity is the particular configuration $b(N, n)$ of the attribute among agents. The topology of the underlying network determines the values of $w(\bar{A})$. We can think of $w(\bar{A})$ as a vector that captures each agent’s effective representation in the population. The higher an agent’s weight, the higher the attribute’s local relative frequency in the population when that agent possesses the attribute’s unit value. To further see how $x(\bar{A}, b, N, n)$ decomposes into $w(\bar{A})$ and $b(N, n)$, note the following brief example: Suppose that we are interested in the relative frequency of the attribute, $x(\bar{A}, b, N, n)$, for agent $j$ in his immediate network neighborhood. We have that

$$x(\bar{A}, b, N, n) = \frac{1}{|\mathcal{N}^+(j)|} \sum_{i \in \mathcal{N}^+(j)} b_i,$$

where $\mathcal{N}^+(j)$ is agent $j$’s out-neighborhood on the network, that is, $\mathcal{N}^+(j)$ is the set of agents $i \in \{1, \ldots, N\}$ for which $[\bar{A}]_{ij} > 0$. Re-writing this expression in Equation 2, we
have:

\[
x(\mathbf{A}, \mathbf{b}, N, n) = \sum_{i=1}^{N} [w(\mathbf{A})]_i b_i,
\]

with \([w(\mathbf{A})]_i = 0\) if \(i \notin N^+(j)\) and \([w(\mathbf{A})]_i = \frac{1}{|N^+(j)|}\) if \(i \in N^+(j)\). Agents not in agent \(j\)'s neighborhood receive zero weight, while agents in agent \(j\)'s neighborhood receive equal positive weight; the total weight allocated across all agents sums to 1.

The decomposition of \(x(\mathbf{A}, \mathbf{b}, N, n)\) in Equation 3 indeed holds more generally, where we assume that each individual agent’s weight is non-negative and \([w(\mathbf{A})]^T 1 = 1\). Depending on the particular setting and the particular interpretation of \(x(\mathbf{A}, \mathbf{b}, N, n)\), \(w(\mathbf{A})\) gets derived differently. However, it is this fixed network-derived vector of agent weights coupled with the combinatorially many possible configurations \(\mathbf{b}(N, n) \in \mathcal{B}(N, n)\), given \(f\), from which we can construct and compute the precursor distribution, \(G_{X(\mathbf{A}, N, n)}(t)\). This decomposition allows us to compute the distributional features of \(X(\mathbf{A}, N, n)\) and the CDF \(G_{X(\mathbf{A}, N, n)}(t)\) even when \(|\mathcal{B}(N, n)| = \binom{N}{n}\) is large.

![Figure 2: Theoretical framework.](image)

Figure 2 illustrates the theoretical framework that forms the basis for this work. We start off with an economic system that has a population of networked agents and an aggregate feature, namely the global relative frequency of the attribute. We assign a weight to each agent in the population and therefore derive from the underlying network a vector of agent weights. There are combinatorially many possible configurations of the attribute consistent with the system’s aggregate feature. We can compute the local relative frequency of the attribute for each configuration, given agents’ weights. We then construct our precursor distribution of possible local relative frequencies of the attribute, and given agents’ decision-
making behavior, we construct the distribution of possible outcomes for the economic system.

### 2.3 Two Examples

We now walk through two examples that make both the theoretical framework and the quantities of interest more precise. Let’s begin by assigning an interpretation to the binary-valued attribute. Let the binary-valued attribute denote employment status, with agent \( i \) unemployed if \( b_i = 1 \), and \( b_i = 0 \) otherwise. The global unemployment rate is \( f = \frac{n}{N} \); in the language from before, \( f \) is the global relative frequency of the unemployment attribute. We are interested in determining the distribution of possible average local unemployment rates. This distribution of average local unemployment rates is a real-world manifestation of the distribution of possible local relative frequencies of the attribute from the previous subsection. We are examining possible local relative frequencies of the unemployment attribute, and more specifically, possible population-averaged local relative frequencies of the unemployment attribute for a given global unemployment rate.

Given the particular configuration of unemployment in the economy, we compute the average local unemployment rate, \( \hat{f}_{\text{avg}}(\bar{A}, b, N, n) \), as follows:

\[
\hat{f}_{\text{avg}}(\bar{A}, b, N, n) = \frac{1}{N} 1^T \tilde{f}(\bar{A}, b, N, n) = \frac{1}{N} 1^T \bar{A} b(N, n) = [d^-_w(\bar{A})]^T b(N, n).
\]

\( \tilde{f}(\bar{A}, b, N, n) \) is the \( N \times 1 \) population vector of agents’ local unemployment rates. Each agent’s local unemployment rate is calculated by determining the weighted relative frequency of the unemployment attribute in that agent’s immediate out-neighborhood. The local unemployment rate for agent \( i \) is therefore \( \hat{f}_i(\bar{A}, b, N, n) = [\tilde{\bar{A}}]_{ii} b(N, n) \), which makes the population vector of local unemployment rates \( \tilde{f}(\bar{A}, b, N, n) = \bar{A} b(N, n) \). The relevant network-derived vector of agent weights is \( d^-_w(\bar{A}) = \frac{1}{N} \bar{A}^T 1 \), the vector of average weighted in-degrees. Note the parallel between the decomposition of \( \hat{f}_{\text{avg}}(\bar{A}, b, N, n) \), the average local unemployment rate, and \( x(\bar{A}, b, N, n) \), the local relative frequency of the attribute:

\[
\hat{f}_{\text{avg}}(\bar{A}, b, N, n) = [d^-_w(\bar{A})]^T b(N, n) \quad \text{and} \quad x(\bar{A}, b, N, n) = [w(\bar{A})]^T b(N, n).
\]

Here, the random variable of interest is \( \hat{f}_{\text{avg}}(\bar{A}, N, n) \) with configuration-specific realization \( \hat{f}_{\text{avg}}(\bar{A}, b, N, n) \), and CDF \( G_{\hat{f}_{\text{avg}}(\bar{A}, N, n)}(t) \) and PMF \( g_{\hat{f}_{\text{avg}}(\bar{A}, N, n)}(t) \), while in the previous subsection, the random variable of interest was \( X(\bar{A}, N, n) \) with configuration-specific realization \( x(\bar{A}, b, N, n) \), and CDF \( G_X(\bar{A}, N, n)(t) \) and PMF \( g_X(\bar{A}, N, n)(t) \). There is an exact
parallel between the computation of \( G_{\hat{F}_{\text{avg}}} (\bar{A}, N, n) (t) \) and \( g_{\hat{F}_{\text{avg}}} (\bar{A}, N, n) (t) \), and the respective computation of \( G_X (\bar{A}, N, n) (t) \) and \( g_X (\bar{A}, N, n) (t) \). When every configuration is equally likely to occur,

\[
G_{\hat{F}_{\text{avg}}} (\bar{A}, N, n) (t) = \frac{1}{|\mathcal{B} (N, n)|} \sum_{b (N, n) \in \mathcal{B} (N, n)} 1_{f_{\text{avg}} (\bar{A}, b, N, n) \leq t}
\]

which exactly parallels the expression for \( G_X (\bar{A}, N, n) (t) \) in Equation 1.

We can now proceed to our first example, in which we study the relationship between agents’ interaction network and the distribution of possible average local unemployment rates:

**Example 1 (Average Local Unemployment Rate, \( N = 4 \))** Consider an economy with \( N = 4 \) agents and an unemployment rate of \( f = 0.25 \). Agents’ social observation network, from which they observe each others’ employment statuses, is depicted in Figure 3. The corresponding row-stochastic weighted adjacency matrix, \( \bar{A} \), is immediately below. Assuming that each configuration of unemployment in the economy is equally likely, we can compute the distribution, \( g_{\hat{F}_{\text{avg}}} (\bar{A}, N, n) (t) \), of possible average local unemployment rates:

\[
\bar{A} = \begin{pmatrix}
1/4 & 1/4 & 1/4 & 1/4 \\
0 & 1/3 & 1/3 & 1/3 \\
0 & 0 & 1/2 & 1/2 \\
0 & 0 & 0 & 1
\end{pmatrix}
\]

\[
g_{\hat{F}_{\text{avg}}} (\bar{A}, N, n) (t) = \begin{cases}
0.25 & \text{if } \hat{f}_{\text{avg}} = 0.0625 \\
0.25 & \text{if } \hat{f}_{\text{avg}} \approx 0.146 \\
0.25 & \text{if } \hat{f}_{\text{avg}} \approx 0.271 \\
0.25 & \text{if } \hat{f}_{\text{avg}} \approx 0.521
\end{cases}
\]

In this example, there are four configurations of unemployment in the economy consistent with \( f = 0.25 \): \( e_1, e_2, e_3, \) and \( e_4 \in \mathbb{R}^4 \). For each configuration, given that \( \hat{f} (\bar{A}, b, N, n) = \bar{A} b (N, n) \), we can compute the average local unemployment rate:

For \( b (N, n) = e_1 \), \( \hat{f} (\bar{A}, b, N, n) = (0.25 \ 0 \ 0 \ 0)^T \) and \( \hat{f}_{\text{avg}} (\bar{A}, b, N, n) = 0.0625. \)

For \( b (N, n) = e_2 \), \( \hat{f} (\bar{A}, b, N, n) = (0.25 \ 0.33 \ 0 \ 0)^T \) and \( \hat{f}_{\text{avg}} (\bar{A}, b, N, n) \approx 0.146. \)

For \( b (N, n) = e_3 \), \( \hat{f} (\bar{A}, b, N, n) = (0.25 \ 0.33 \ 0.50 \ 0)^T \) and \( \hat{f}_{\text{avg}} (\bar{A}, b, N, n) \approx 0.271. \)

For \( b (N, n) = e_4 \), \( \hat{f} (\bar{A}, b, N, n) = (0.25 \ 0.33 \ 0.50 \ 1.00)^T \) and \( \hat{f}_{\text{avg}} (\bar{A}, b, N, n) \approx 0.521. \)

From these values, we can then construct the probability distribution \( g_{\hat{F}_{\text{avg}}} (\bar{A}, N, n) (t) \), which is depicted in Figure 3. This probability distribution is our main object of interest, and most of the present work is devoted to understanding and characterizing this type of object.

The topology of agents’ observation network in Example 1 generates wide variation in individual agents’ local unemployment rates for a particular configuration of unemployment.
in the economy, and it generates variation in individual agents’ local unemployment rates across configurations. The latter variation arises from agents having different effective representations in the population, or in this particular setting, different levels of observability. Agent 1 has a low average weighted in-degree, and therefore a low agent weight and poor observability, so when he is unemployed, agents 1, 2, 3, and 4 respectively have local unemployment rates of 25 percent, 0 percent, 0 percent, and 0 percent; agent 4 meanwhile has a high average weighted in-degree, and therefore a high agent weight and strong observability, so when he is unemployed, agents 1, 2, 3, and 4 respectively have local unemployment rates of 25 percent, 33 percent, 50 percent, and 100 percent. This variation in the observability of agents causes the average local unemployment rate to change with configuration. When agent 1 is unemployed, the average local unemployment rate is 6.25 percent, while when agent 4 is unemployed, the average local unemployment rate is 52.1 percent. These values for the average local unemployment rate also strongly deviate from the actual unemployment
rate of 25 percent. If agents form macroeconomic sentiments from their local rates of unemployment, then depending on the particular configuration of unemployment in the economy, agents on average might feel that the economy is doing much better or worse than its fundamentals would otherwise suggest. If the outcome of the economy somehow depends on this average local unemployment rate, and this average local unemployment rate substantially varies with the particular configuration of unemployment in the economy, then we consider the economy to be strongly configuration-dependent.

![Diagram](image)

Figure 4: Possible average local unemployment rates, $\hat{f}_{\text{avg}} (\bar{A}, b, N, n)$, given the global unemployment rate, $f$, and potential pathways (A-D) for the average local unemployment rate as the global unemployment rate evolves.

We can trace different pathways for the average local unemployment rate as the global unemployment rate evolves. Figure 4 features four potential pathways for the average local unemployment rate; Figure 4 also plots, in the background, the set of all possible average local unemployment rates for each feasible level of unemployment in the economy. Changes in configuration can accommodate different phenomena that would otherwise not emerge from just the aggregate properties of the system. Path (B) illustrates how there can be dramatic swings in sentiment for a small adjustment to the the global unemployment rate, the system’s aggregate feature; as the unemployment rate increases from 25 percent to 50 percent, there is a 72.9-percentage-point increase in the average local unemployment rate. Path (C) illustrates how sentiment can move in a direction opposite to that of fundamentals. Even though the unemployment rate is declining from 50 percent to 25 percent, the average local unemployment rate increases 31.3 percentage points. The decrease in unemployment would suggest that the economy is improving, but the agents in the population on average locally observe the economy to be worsening. Paths (A) and (D) illustrate hysteresis within
the economy. Suppose that the economy takes path (A) as the unemployment rate increases and it takes path (D) as the unemployment rate decreases. Even though the economy is traversing the same set of unemployment rates, it can experience different average local unemployment rates. Configuration dependence of the economic system enables the existence of such phenomena.

In the second example, we study the distribution of possible average local unemployment rates, \( g_{\text{avg}}(\bar{A}, N, n)(t) \), in a setting with a larger sample population:

**Example 2 (Average Local Unemployment Rate, N = 15)** Consider an economy with \( N = 15 \) agents and an unemployment rate of \( f = 0.20 \). Agents’ social observation network, \( G(A) \), is formed from preferential attachment, with a self-loop for every node (see Figure 5). Assuming that agents equally weight each of their observations and each configuration of unemployment in the economy is equally likely, the distribution of possible average local unemployment rates, \( g_{\text{avg}}(\bar{A}, N, n)(t) \), is depicted at the bottom right of Figure 5.

![Figure 5: Graphs G(A) (top left) and G(\bar{A}) (top center) for Example 2. Calculating a node’s weighted in-degree (top right) and the plot of average weighted in-degrees, \( d^w(\bar{A}) \) (bottom left). The distribution of average local unemployment rates, \( g_{\text{avg}}(\bar{A}, N, n)(t) \), for \( f = 0.20 \) (bottom right).](image)
In this second example, there are \( \binom{15}{3} = 455 \) possible configurations of the unemployment attribute among agents in the population consistent with a 20-percent unemployment rate. As in the first example, we compute the average local unemployment rate configuration by configuration, and we then construct the accompanying probability distribution. We observe substantial heterogeneity in agents’ average weighted in-degrees, \( d_{w,\bar{A}}(\bar{A}) = \frac{1}{N}\bar{A}^T \mathbf{1} \), as can be viewed in the bottom left of Figure 5. With such heterogeneity in agents’ weights, individual agents’ contributions to the average local unemployment rate measurably differ when they become unemployed. As a result, the probability distribution of possible average local unemployment rates has sizable variance. For a 20-percent unemployment rate, the average local unemployment rate can vary from 11.9 percent to 33.1 percent. If the local unemployment rate is relevant for agent decision-making, the economy is strongly dependent on the underlying configuration of unemployment in the population.

We can construct such a probability distribution of possible average local unemployment rates for every feasible level of unemployment in the economy. We can conceivably compute the average local unemployment rate configuration by configuration for a given global unemployment rate, and we can then plot the corresponding probability distribution. The left of Figure 6 plots the set of possible average local unemployment rates for a fixed global unemployment rate, and the right of Figure 6 plots the corresponding probability distribution of possible average local unemployment rates for each feasible global level of unemployment in the economy. When one agent is unemployed, there are just 15 possible configurations, and when 7 or 8 agents are unemployed, there are 6435 possible configurations. We continue to observe strong configuration dependence of the system, and the average local unemployment rate can substantially deviate from the actual global unem-

Figure 6: Possible average local unemployment rates, \( \bar{f}_{avg}(\bar{A}, \mathbf{b}, N, n) \), for a given global unemployment rate, \( f \), (left) and the probability distribution of possible average local unemployment rates, \( g_{f_{avg}}(\bar{A}, N, n) \) (right), for each global level of unemployment in the economy (right).
ployment rate depending on the particular configuration of unemployment in the economy. As the population size gets larger, the total number of possible configurations, \( \binom{N}{n} \), grows combinatorially, and it becomes less feasible to construct the distribution of possible average local unemployment rates configuration by configuration. In Section 6, we present a set of theoretical results that allows us to construct this probability distribution and compute its features in closed form for every population size, no matter how large the population size gets.

We can imagine that an agent’s local unemployment rate is a determinant of his or her sentiment about the macroeconomy, and the average local unemployment rate is a population-wide indicator of macroeconomic sentiment. Such macroeconomic sentiments can influence agents’ behavior. In the next section, we consider a setting in which the local unemployment rate impacts agents’ voting decision. Depending on the topology of agents’ observation network, the overall election outcome, and thus the outcome for the economy, can be strongly configuration dependent.

3 MACROECONOMIC SENTIMENT AND ELECTION OUTCOMES

We use the theoretical framework that we just developed to examine election outcomes in a stylized setting. We study a population of voters who must choose between two candidates named Hillary Clinton and Donald Trump. Each voter’s macroeconomic sentiment, formed from that voter’s local unemployment rate, influences his or her voting decision, so configuration-induced variations in sentiment for fixed macroeconomic fundamentals can alter individual agents’ voting decisions, and thereby alter the election outcome and the outcome of the economy. This section essentially shows how we can embed our theoretical framework developed in the previous section into a real-world setting and extract economic meaning.

3.1 MODEL

We begin by considering a population of \( N \) agents, each of whom is a voter in the election. Each agent faces a binary choice problem: either vote for Hillary Clinton or Donald Trump. To make their decisions, agents consider the candidates’ policies. There are \( P \) issues, and the candidates construct a policy for every issue. For each policy put forth by a particular candidate, there is an associated scalar benefit. Scalar benefits are separable across issues. The \( P \times 1 \) vectors of benefits corresponding to the policies of Clinton and Trump are respectively \( \mathbf{x}_C \) and \( \mathbf{x}_T \). Agent \( i \) weights each candidate’s set of policies using the
With the exception of Trump’s, agent assigns to this policy directly depends on his or her local unemployment rate, benefits, where $\epsilon_i$ is the configuration of unemployment in the economy, with $\epsilon_i$ independent across voters. Define $\alpha_i^T$ as the configuration of unemployment in the economy, with $\epsilon_i$ independent across voters. Define $\alpha_i^T$ as the weighted sum of benefits accrued from the policies of a particular candidate. Setting $\eta_i \sim \text{Uniform}(-\beta, \beta)$ for every agent $i \in \{1, \ldots, N\}$,

$$\pi_{iT} = \frac{1}{2} + \frac{\alpha_i^T (x_T - x_C)}{2\beta} \quad \text{and} \quad \pi_{iC} = \frac{1}{2} + \frac{\alpha_i^T (x_C - x_T)}{2\beta},$$

where $\pi_{iT} = \Pr [u_i (p_T ; \epsilon_{iT}) > u_i (p_C ; \epsilon_{iC})]$ and $\pi_{iC}$ are assumed to be both independent of the candidates’ policies and independent across voters. Define $\eta_i = \epsilon_{iT} - \epsilon_{iC}$, which is independent of $\alpha_i^T x_T - \alpha_i^T x_C$. Setting $\eta_i \sim \text{Uniform}(-\beta, \beta)$ for every agent $i \in \{1, \ldots, N\}$,

Let the first policy concern jobs and unemployment. We assume that the weight an agent assigns to this policy directly depends on his or her local unemployment rate, $\hat{f}_i (\bar{A}, b, N, n)$:

$$[\alpha_i]_1 = \alpha_i^{(1)} = \hat{f}_i (\bar{A}, b, N, n);$$

$b \equiv b (N, n)$ is the configuration of unemployment in the economy, with $[b]_i = 1$ if agent $i$
is unemployed and \( b_i = 0 \) otherwise, and the overall unemployment rate is \( f = \frac{n}{N} \). The higher an agent’s local unemployment rate, the more that agent cares about the issue of jobs and unemployment in the economy. This is consistent with the findings of Bisgaard, Dinesen, and Sønderskov (2016), who observe that Danish voters’ dissatisfaction with the national economy increases with the local unemployment rate, defined as the fraction of all unemployed residents within a fixed meter radius from the voter’s place of residence. The work of Healy and Lenz (2017) meanwhile justifies the dependence of agents’ voting decisions on the local unemployment rate, as these authors demonstrate that local unemployment conditions impact the national voting outcome.

The aggregate equation for the expected fraction of votes for Trump is now:

\[
\frac{1}{N} \sum_{i=1}^{N} \pi_{iT} = \frac{1}{2} + \frac{1}{2\beta} \left[ \hat{f}_{\text{avg}} (\bar{A}, \bar{b}, N, n) (x_{T,1} - x_{C,1}) + \sum_{i=1}^{N} \frac{\alpha_i}{N} (x_{T,2} - x_{C,2}) + \cdots + \sum_{i=1}^{N} \frac{\alpha_i P}{N} (x_{T,P} - x_{C,P}) \right].
\]

The overall weight accorded to the issue of jobs and unemployment in the economy is the average local unemployment rate. If this quantity varies enough with the configuration of unemployment in the economy, then we can potentially anticipate different voting outcomes for a particular global unemployment rate.

Let’s assume that Trump’s jobs policy confers a greater benefit to voters than Clinton’s jobs policy: \( x_{T,1} > x_{C,1} \) with \( x_{T,1} - x_{C,1} = 9 \). Let’s also assume that \( x_{C,j} - x_{T,j} = x_{C,\ell} - x_{T,\ell} = 1 \) for all \( j, \ell \neq 1 \), so the policy put forth by Clinton for every other issue yields a benefit that exceeds that of Trump’s corresponding policy. A higher average local unemployment rate favors the election of Trump. The aggregate equation for the expected fraction of votes for Trump becomes:

\[
\frac{1}{N} \sum_{i=1}^{N} \pi_{iT} = \frac{1}{2} + \frac{1}{2\beta} \left[ 10 \hat{f}_{\text{avg}} (\bar{A}, \bar{b}, N, n) - 1 \right].
\]

(4)

If the average local unemployment rate is 10 percent, either candidate is equally likely to win the election. If the average local unemployment rate exceeds 10 percent, the expected vote share for candidate Trump exceeds 50 percent, so the voting outcome favors Trump. If the average local unemployment rate is less than 10 percent, the expected vote share for candidate Trump is less than 50 percent, so the voting outcome favors Clinton. We proceed to construct voters’ observation network so that we can determine the possible values of the average local unemployment rate consistent with the economy’s overall unemployment rate.
Voters compute their local unemployment rates from their nodes on this network. These local unemployment rates affect individual voting behavior, and the average local unemployment rate affects aggregate voting behavior and the election outcome.

3.2 Constructing Voters’ Observation Network

To construct voters’ observation network, we draw on details from the 2016 U.S. presidential election. The observation network has 137.5 million nodes, the total number of voters in the 2016 U.S. presidential election. Each node has a self-loop because voters observe their own employment statuses. We assume that each agent has, on average, 50 reciprocal linkages; these linkages may be formed with relatives, colleagues, acquaintances, and so on. We accordingly construct an undirected Erdös-Rényi graph whose expected degree is 50. We refer to this network of 137.5 million nodes with its self-loops and Erdös-Rényi linkages as the base graph.

In addition to the linkages comprising the base graph, we introduce media-originating directed linkages. During the weeks and months preceding the election, we assume that voters engage with a variety of news/talk media outlets that feature stories about employed and unemployed individuals. These featured individuals can shift voters’ perceptions of unemployment in the economy. We therefore gather statistics on television network viewership, radio show listenership, and newspaper, magazine, business journal, and business publication circulation in the United States.

Appendix A in the Supplementary Materials provides more details about viewership, listenership, and readership statistics for media in the United States and overall data construction. There is a total of 1867 different news/talk media sources. As observed in Figure 7, audience sizes are heavy-tailed. We assume that news/talk media outlets feature an employed or unemployed individual in one story per week for 15 weeks. Therefore, to construct the network of media-originating linkages, for each media outlet, we randomly select the set of audience members and we randomly select the set of 15 individuals that are featured in that news outlet’s stories. Directed edges are then drawn from the set of audience members to each featured agent. This method of network construction is carried out for all 1867 news/talk media sources.

---


7. We choose 15 weeks because that is the number of weeks that elapsed from the conclusion of the Republican and Democratic National Conventions until Election Day for the 2016 U.S. presidential election. The Republican National Convention, Democratic National Convention, and Election Day respectively took
Figure 7: Counter-cumulative distribution function (CCDF) of degrees for the base graph (top left). CCDF of audience sizes for 1867 different media sources (top right). CCDFs of out- and in-degrees for the network of media-originating linkages (middle). CCDFs of out- and in-degrees for the composite network (bottom).
Appendix A in the Supplementary Materials presents detailed summary statistics for the base graph, the media-originating graph, and the composite graph that pools both base and media-originating linkages. Figure 7 plots the counter-cumulative distribution function of degrees for the base graph, the counter-cumulative distribution functions of out-degrees and in-degrees arising from the network of media-based linkages, and the counter-cumulative distribution functions of out-degrees and in-degrees arising from the composite network. For the base graph, with its 3,575,017,297 undirected edges, the average degree is 51.0 with a standard deviation of 7.07. We obtain this average degree because the Erdős-Rényi graph has an average degree of 50 and each agent has a self-loop. In the media-originating graph, with its 2,712,493,694 directed edges, the average out-degree is 19.7 with a standard deviation of 17.0, and the average in-degree is 19.7 with a standard deviation of 8,633.3. The counter-cumulative distribution function of out-degrees for the media-originating graph is a step function because agents accumulate 15 out-edges for every media source in which they are an audience member. Therefore, out-degrees for the media-originating graph occur in multiples of 15. Most voters have zero in-degree in the media graph because they are not featured in news/talk media outlets; there are only 28,003 individuals featured in employment-related news stories. The probability of a non-zero in-degree for this graph is $2.04 \times 10^{-4}$, which we can observe in Figure 7 (middle). Since the counter-cumulative distribution function of audience sizes across media sources is heavy-tailed, the counter-cumulative distribution function of in-degrees for the media-originating graph is similarly heavy-tailed.

In the composite graph, the average out-degree is 70.7 with a standard deviation of 18.4. As depicted in Figure 7, we see that the counter-cumulative distribution function of out-degrees for the composite graph takes the same shape as the counter-cumulative distribution function of out-degrees for the base graph, except that the former distribution is shifted to the right. Each agent has accumulated additional out-edges from media-originating linkages, which generates a shift in the distribution function. On average, the total number of media sources to which people are exposed is 1.32. For the composite graph, the average in-degree is 70.7, with a standard deviation of 8,633.3, and the maximum in-degree is 8,020,651. The counter-cumulative distribution function of in-degrees for the composite network (Figure 7, bottom) directly incorporates the distributional features of the counter-cumulative distribution function of degrees for the base graph (Figure 7, top left) and the counter-cumulative distribution function of in-degrees for the media graph (Figure 7, middle). Most agents are not featured by the media, so their in-degree is equal to their degree from the base graph. A

---

8The average out-degree and the average in-degree for any directed graph take the same value since the total number of out-edges equals the total number of in-edges.
small fraction of agents are featured in the media, so their in-degree is equal to their degree from the base graph plus their in-degree from the media graph. The counter-cumulative distribution function of in-degrees for the composite graph therefore becomes heavy-tailed.

### 3.3 When Configurations are Equally Likely: Distribution of Possible Average Local Unemployment Rates and the Expected Voting Outcome

We can now study the distribution of possible average local unemployment rates and the probability that the election outcome favors each individual candidate. We set the unemployment rate in this economy to 9.6 percent. We decide to use the October 2016 U-6 unemployment rate in the United States specified by the Bureau of Labor Statistics. This value is the national unemployment rate that immediately precedes the 2016 U.S. presidential election.\(^9\) Since we are interested in capturing individuals’ macroeconomic sentiments, we use the U-6 unemployment series because it counts those people who are discouraged workers or underemployed for economic reasons as unemployed.

We would like to determine the distribution of possible average local unemployment rates, \(G_{\bar{A}}(\bar{A},N,n,t)\), in the economy given that there is an overall 9.6-percent unemployment rate. For each possible configuration of unemployment in the economy, the average local unemployment rate is computed as follows: 

\[
\frac{\bar{A}}{N}^T \mathbf{b}(N,n) = [d_{w}(\bar{A})]^T \mathbf{b}(N,n).
\]

As in Section 2, the average local unemployment rate depends on the vector of average weighted in-degrees, \(d_{w}(\bar{A}) = \frac{1}{N} \bar{A} \mathbf{1}\), of the underlying social observation network, \(\mathcal{G}(\bar{A})\), and the particular configuration of unemployment, \(\mathbf{b}(N,n)\). The capacity for \(G_{\bar{A}}(\bar{A},N,n,t)\) to be meaningfully non-degenerate depends on the properties of \(d_{w}(\bar{A})\), the latter of which is derived from the 137.5-million-node voter observation network.

Appendix B in the Supplementary Materials considers the case in which agents’ observation network solely consists of the base graph; in that setting, the average local unemployment rate does not meaningfully vary with configuration. Here, we take agents’ observation network to be the more realistic composite graph. We compute each agent’s average weighted in-degree by assuming that agents equally weight each of their observations of employment status. Once we compute this vector of agent weights, we can determine each agent’s effective representation in the population. On average, each agent has an effective weight of 1 agent, which we would expect. The effective minimum weight is 0.249 agents, and the

---

\(^9\) The U-6 unemployment rate is defined as the “total unemployed, plus all marginally attached workers plus total employed part time for economic reasons, as a percent of all civilian labor force plus all marginally attached workers.”
effective maximum weight is 98,733.1 agents. The median agent has an effective weight of 0.756 agents. The left side of Figure 8 plots the counter-cumulative distribution of average weighted in-degrees. This distribution of agent weights is heavy-tailed. There is a relatively small subset of agents in the entire voting population that is particularly influential from being featured in the media.

We observe the distribution of possible average local unemployment rates, $G^*_{\text{avg}}(\bar{A},N,n)(t)$, on the right side of Figure 8. For every distributional feature of $G^*_{\text{avg}}(\bar{A},N,n)(t)$ that we highlight in this section, there is a corresponding theorem presented in later sections of this work that shows how to compute that quantity in closed form. On the right side of Figure 8, the theoretical CDF for $G^*_{\text{avg}}(\bar{A},N,n)(t)$ overlaps an empirical CDF.\[10\] The empirical CDF is constructed by randomly drawing 100,000 configurations of unemployment from the set of all possible configurations consistent with a 9.6-percent unemployment rate, and then computing the associated average local unemployment rate for each configuration. The theoretical and empirical mean of this distribution is 0.096.\[11\] The theoretical standard deviation for this distribution is 0.00266, or 0.266 percentage points, and the size of two standard deviations about the distribution’s mean value is 1.07 percentage points.\[12\] Staying within this two-standard-deviation band, the average local unemployment rate can generally vary from 9.07 percent to 10.1 percent. The lowest possible average local unemployment rate is 5.53 percent, and the highest possible average local unemployment rate is 33.3 percent.\[13\] The average local unemployment rate here exhibits substantial configuration dependence, especially when we consider the magnitude of fluctuations in the actual unemployment rate over the course of a business cycle. For the most recent complete business cycle dated by the NBER, that is, from March 2001 until December 2007 (i.e., from peak to peak), the U-6 unemployment series varied from 7.3 percent to 10.4 percent, a difference of just 3.1 percentage points. Variations in the average local unemployment rate are indeed large enough that they can potentially mimic variations in business cycle conditions.

There are two different ways that we can understand the existence of such strong configuration dependence for this very-large-$N$ economic system. First, strong variation in $\hat{f}_{\text{avg}}(\bar{A}, N, n)$ emerges for a given $f$ because the set of in-degrees for the composite graph (Figure 7, bottom) has a very high variance. Second, strong variation in $\hat{f}_{\text{avg}}(\bar{A}, b, N, n)$ emerges because the distribution of agent weights, that is, the distribution of average weighted in-degrees for the composite graph (Figure 8, left) has a heavy tail; as a result, there is a subset of agents in the voting population that drives variation in the average local unemploy-

---

\[10\] The theoretical CDF is constructed from Theorem 13 in Section 6.
\[11\] The theoretical mean is computed from Theorem 8 in Section 6.
\[12\] The theoretical standard deviation is computed from Theorem 9 in Section 6.
\[13\] Both quantities are constructed from Theorem 10 in Section 6.
ment rate due to their relatively large influence. Theorems 15 and 16 in Section 6 precisely show how these two features of the composite graph generate such a strongly non-degenerate distribution, $G_{\bar{F}_{avg}(\bar{A}, N, n)}(t)$, even for large $N$.

![Graph](image)

Figure 8: Counter-cumulative distribution function of average weighted in-degrees for the composite network, assuming that each agent assigns an equal weight to each of his out-linkages (left). Distribution of the average local unemployment rate, $G_{\bar{F}_{avg}(\bar{A}, N, n)}(t)$, when $f = 0.096$, assuming that configurations of unemployment in the economy are equally likely (right).

The economy here exhibits sufficiently large configuration dependence that its outcome, determined by the outcome of the U.S. presidential election, depends on the actual allocation of unemployment among voters:

**Example 3 (Voting Outcome, Composite Graph)** Aggregate voting behavior is characterized by Equation 4. The unemployment rate is 9.6 percent. Given that voters’ observation network is the composite graph, and voters equally weight each of their observations, there is a probability of 0.0707 that Trump’s expected vote share exceeds 0.5:

$$
Pr \left[ \frac{1}{N} \sum_{i=1}^{N} \pi_{iT} > 0.5 \right] = Pr \left[ \bar{F}_{avg}(\bar{A}, N, n) > 0.10 \right] \approx 0.0707.
$$

In Section 6, we show how to compute this quantity by hand. If we only observe the global 9.6-percent unemployment rate, we might think that the election outcome favors Clinton with certainty. However, since the average local unemployment rate can meaningfully deviate from 9.6 percent, the outcome can favor Trump, thereby setting the economy along a path that differs from the one in which Clinton is elected. The probability that the election outcome favors Clinton is 92.93 percent, and the probability that the election outcome favors Trump is 7.07 percent.

Appendix B in the Supplementary Materials considers the distribution of possible outcomes for the economy for two variants of the composite graph. For both variants, the composite graph is constructed by pooling linkages from the base graph and the media graph,
as is done here. In the first variant, the base graph is modified while the media graph stays the same. The base graph is constructed by assuming that agents form an average of 20 reciprocal linkages with other voters, rather than 50 reciprocal linkages; every voter has a self-loop as well. Given this composite graph and a 9.6-percent unemployment rate, the minimum possible average local unemployment rate is 3.63 percent, the maximum possible average local unemployment rate is 47.6 percent, and the standard deviation of the distribution of average local unemployment rates is 0.433 percentage points, so that a two-standard-deviation band equals 1.73 percentage points. The probability that the election outcome favors Trump is equal to 17.7 percent, and the probability that the election outcome favors Clinton is equal to 82.3 percent. In the second variant, the base graph stays the same, that is, agents each have a self-loop and form 50 reciprocal linkages with other voters on average, while the media graph is modified. The media graph in this second variant is constructed by assuming that each news/talk media source publishes five stories, rather than 15 stories, about the issue of jobs and unemployment. Given this composite graph and a 9.6-percent unemployment rate, the minimum possible average local unemployment rate is 6.55 percent, the maximum possible average local unemployment rate is 21.5 percent, and the standard deviation of the distribution of average local unemployment rates is 0.205 percentage points, so that a two-standard-deviation band equals 0.820 percentage points. The probability that the election outcome favors Trump is equal to 3.27 percent, and the probability that the election outcome favors Clinton is equal to 96.73 percent. For all of the graphs considered in this work, there is inherent randomness; the base graph is constructed by drawing linkages randomly between pairs of individuals, and the media graph is constructed by randomly selecting, for each news/talk media source, audience members and featured individuals. Such randomness serves as a natural benchmark; we would need to separately explore whether deviations from randomness systematically change the properties of the distribution of possible average local unemployment rates.

Each individual’s local unemployment rate essentially serves as a proxy for his or her sentiment about the macroeconomy. The average local unemployment rate is thus an aggregate statistic summarizing sentiment for all agents in the system. Such sentiment, and its fluctuation, is a manifestation of animal spirits at its core. Holding the fundamentals of the economy fixed, there can be configuration-induced variations in sentiment. There can be waves of optimism if the average local unemployment rate is less than its global value, and there can be waves of pessimism if the average local unemployment rate is greater than its global value. Sentiment is moreover quantifiable; the extent to which aggregate sentiment deviates from a level that is commensurate with fundamentals depends on the extent to which the average local unemployment rate deviates from the global rate of unemployment.
The underlying interaction structure among agents in the economy shapes the capacity for there to exist non-fundamental swings in aggregate sentiment. This work therefore provides a microfoundation for animal spirits.

This work buttresses other research that studies sentiment and consumer confidence, and shocks to sentiment and consumer confidence, in the macroeconomic setting and their effects on business cycles and aggregate fluctuations: for example, Farmer and Guo (1994), Barsky and Sims (2012), Angeletos and La’O (2013), Benhabib, Wang, and Wen (2015), Huo and Takayama (2015), Acharya, Benhabib, and Huo (2017), Angeletos, Collard, and Dellas (2017), and Milani (2017). In this work, we provide a simple mechanism for generating fluctuations in sentiment. Fluctuations in sentiment or animal spirits arise here from variations in configuration, with the scope for such fluctuation dependent on the topology of agents’ interaction network. Cross-sectionally, variations in agents’ sentiment arise from differences in agents’ local environments due to differences in network position, holding the economy’s fundamentals fixed.

3.4 When Configurations are Not Equally Likely: Distribution of Possible Average Local Unemployment Rates and the Expected Voting Outcome

Thus far, we have been considering the case in which each configuration of unemployment is equally likely to occur, meaning that each individual in the population is equally likely to be unemployed. We proceed to dispense with this assumption and instead compute the mean and variance of the distribution of possible average local unemployment rates when configurations are no longer equally likely. Now, the mean average local unemployment rate can deviate quite strongly away from the actual global rate of unemployment, \( f \).

We segment the population into two groups: (1) those agents featured by news/talk media and (2) those agents not featured by news/talk media, with agents in the first group relatively more likely to be unemployed. The number of agents in the first group is \( x = 28,003 \), and the number of agents in the second group is \( N - x = 137.5 \times 10^6 - 28,003 \). We re-index the population of agents so that those in group 1 have indices 1 to \( x \) while those in group 2 have indices \( x + 1 \) to \( N \). Attribute \( \gamma_i \) characterizes agents according to whether or not they have been featured by news/talk media. The probability \( \phi_i \) that agent \( i \) is unemployed is \( \phi_i = \rho_1 \) for all \( i \in \{1, \ldots, x\} \) and \( \phi_i = \rho_2 \) for all \( i \in \{x + 1, \ldots, N\} \). The odds ratio for agents in group 1 relative to group 2 is: \( \hat{\psi}_1 = \frac{\rho_1}{1-\rho_1} \). \( \frac{\rho_2}{1-\rho_2} \).

Example 4 (Configurations Unequally Likely, \( \hat{\psi}_1 = 9.42 \)) Suppose that media outlets
engage in “fair and balanced” reporting, providing equal air time (or equal space for hard-copy publications) to those agents who are employed and unemployed. Setting $\rho_1 = 0.50$ and $\rho_2 = 0.096$, $E\bar{F}_{\text{avg}}(\bar{A}, N, n, (\gamma_i)_{i=1}^{N}) = 0.194$ and Std. Dev. $\bar{F}_{\text{avg}}(\bar{A}, N, n, (\gamma_i)_{i=1}^{N}) = 0.00452$.

We show how to compute these first two moments in Section 7, employing Theorem 17. Here we observe extreme bias in the distribution of average local unemployment rates relative to the actual global unemployment rate of 9.6 percent. The mean average local unemployment rate is very high at 19.4 percent, so the voting outcome overwhelmingly favors Trump.

This heightened exposure to unemployment via the media might be a reason why residents of the United States, as well as residents of other countries, grossly overestimate the national unemployment rate. In an August 2014 Ipsos-MORI poll, surveyed Americans stated, on average, that the national unemployment rate was 32 percent, greatly exceeding even the U-6 unemployment rate of 11.9 percent.\(^{14}\) Similarly, in an October 19-22, 2016 survey of 1000 unemployed American adults, about one in three individuals believed that the national unemployment rate was 15 percent or higher.\(^{15}\) The polled individuals were unemployed, so it makes sense that they sense a national unemployment rate that exceeds the actual one. However, the extent of miscalculation is still quite significant, for in some cases, they were even provided with information about the national unemployment rate. The unemployment rate is one of the most salient features of an economy, so individuals’ local perceptions of this statistic directly affect how they perceive the economy’s overall health. Persistently high assessments of the unemployment rate can impact the behavior of various agents in the economy, whether such decision-making concerns voting or something else.

### 4 Sample Network-Derived Vectors of Agent Weights

Beginning with this section, we develop the mathematics that enables us to first construct the precursor distribution of possible local relative frequencies of the attribute and then construct the distribution of possible outcomes for the economy given the topology of agents’ interaction network. To construct the precursor distribution $G_{X(\bar{A}, N, n)}(t)$, we decompose each quantity $x(\bar{A}, b, N, n)$ into a network-derived vector of agent weights, $w(\bar{A})$, and a configuration vector: $b(N, n)$:

$$x(\bar{A}, b, N, n) = \left[w(\bar{A})\right]^T b(N, n).$$

\(^{14}\)“Americans think the unemployment rate is 32 percent,” *Vox*, November 15, 2014.

The vector of agent weights is the object by which the topology of agents’ interaction network shapes the precursor distribution and the distribution of possible outcomes for the economy.

Quantities $x(\bar{\mathbf{A}}, \mathbf{b}, N, n)$ and $\mathbf{w}(\bar{\mathbf{A}})$ are both very useful objects that will allow us to study and characterize the precursor distribution and the distribution of possible outcomes for the economy, but ultimately they are abstract. We refer to $x(\bar{\mathbf{A}}, \mathbf{b}, N, n)$ as the local relative frequency of the attribute; however, without additional details about the specific interpretation of this object, there does not actually exist a method by which we compute this quantity. Similarly, $\mathbf{w}(\bar{\mathbf{A}})$ is a network-derived vector of agent weights, but without additional details about this object, we do not have a method for deriving this vector. In the present section, we study different cases of $x(\bar{\mathbf{A}}, \mathbf{b}, N, n)$ and $\mathbf{w}(\bar{\mathbf{A}})$. For each case, we have enough detail that we can assign a particular interpretation to $x(\bar{\mathbf{A}}, \mathbf{b}, N, n)$, and we can show how the vector of agent weights, $\mathbf{w}(\bar{\mathbf{A}})$, naturally emerges from the underlying network. In the previous two sections, we focused on the pair $(x(\bar{\mathbf{A}}, \mathbf{b}, N, n), \mathbf{w}(\bar{\mathbf{A}})) = \left(\widehat{f}_{\text{avg}}(\bar{\mathbf{A}}, \mathbf{b}, N, n), \mathbf{d}_w(\bar{\mathbf{A}})\right)$; in this section, we introduce additional pairs that can be relevant for other applications. Our sample set of scalar quantities and the vectors of agent weights to which they pair is as follows:

1. $\left(\widehat{f}_i(\bar{\mathbf{A}}, \mathbf{b}, N, n), \mathbf{w}_{a,i}(\bar{\mathbf{A}})\right)$. For every agent $i \in \{1, \ldots, N\}$, $\widehat{f}_i(\bar{\mathbf{A}}, \mathbf{b}, N, n)$ is the configuration-specific weighted local relative frequency of the attribute in agent $i$’s immediate network neighborhood. The accompanying random variable is $\widehat{F}_i(\bar{\mathbf{A}}, N, n)$ with CDF $G_{\widehat{F}_i}(\bar{\mathbf{A}}, N, n)(t)$. The corresponding network-derived vector of agent weights is $\mathbf{w}_{a,i}(\bar{\mathbf{A}}) = \left([\bar{\mathbf{A}}]_{i*}\right)^T$, so $\widehat{f}_i(\bar{\mathbf{A}}, \mathbf{b}, N, n) = \left[\mathbf{w}_{a,i}(\bar{\mathbf{A}})\right]^T \mathbf{b}(N, n)$.

2. $\left(\widehat{f}_i^{(q)}(\bar{\mathbf{A}}, \mathbf{b}, N, n), \mathbf{w}_{a,i}^{(q)}(\bar{\mathbf{A}})\right)$. For every agent $i \in \{1, \ldots, N\}$, $\widehat{f}_i^{(q)}(\bar{\mathbf{A}}, \mathbf{b}, N, n)$ is the configuration-specific weighted local relative frequency of the attribute for agent $i$ following $q$ rounds of repeated linear updating by each agent with his or her immediate neighbors. The accompanying random variable is $\widehat{F}_i^{(q)}(\bar{\mathbf{A}}, N, n)$ with CDF $G_{\widehat{F}_i^{(q)}}(\bar{\mathbf{A}}, N, n)(t)$. The corresponding network-derived vector of agent weights is $\mathbf{w}_{a,i}^{(q)}(\bar{\mathbf{A}}) = \left([\bar{\mathbf{A}}^q]_{i*}\right)^T$, so $\widehat{f}_i^{(q)}(\bar{\mathbf{A}}, \mathbf{b}, N, n) = \left[\mathbf{w}_{a,i}^{(q)}(\bar{\mathbf{A}})\right]^T \mathbf{b}(N, n)$.

3. $\left(\widehat{f}_{\text{avg}}(\bar{\mathbf{A}}, \mathbf{b}, N, n), \mathbf{d}_w(\bar{\mathbf{A}})\right)$. $\widehat{f}_{\text{avg}}(\bar{\mathbf{A}}, \mathbf{b}, N, n)$ is the configuration-specific population-averaged weighted local relative frequency of the attribute in each agent’s immediate neighborhood. The accompanying random variable is $\widehat{F}_{\text{avg}}(\bar{\mathbf{A}}, N, n)$ with CDF $G_{\widehat{F}_{\text{avg}}}(\bar{\mathbf{A}}, N, n)(t)$. The corresponding network-derived vector of agent weights is $\mathbf{d}_w(\bar{\mathbf{A}}) = \frac{1}{N} \bar{\mathbf{A}}^T \mathbf{1}$, so $\widehat{f}_{\text{avg}}(\bar{\mathbf{A}}, \mathbf{b}, N, n) = \left[\mathbf{d}_w(\bar{\mathbf{A}})\right]^T \mathbf{b}(N, n)$.

4. $\left(\widehat{f}_{\text{avg}}^{(q)}(\bar{\mathbf{A}}, \mathbf{b}, N, n), \mathbf{d}_w^{(q)}(\bar{\mathbf{A}})\right)$. $\widehat{f}_{\text{avg}}^{(q)}(\bar{\mathbf{A}}, \mathbf{b}, N, n)$ is the configuration-specific population-averaged weighted local relative frequency of the attribute in each agent’s immediate neighborhood.
averaged weighted local relative frequency of the attribute following \( q \) rounds of repeated linear updating by each agent with his or her immediate neighbors. The accompanying random variable is \( \bar{F}_{\text{avg}}^{(q)}(\bar{A}, N, n) \) with CDF \( G_{\bar{F}_{\text{avg}}^{(q)}}(\bar{A}, N, n)(t) \). The corresponding network-derived vector of agent weights is \( \bar{d}_{w}^{(q)}(\bar{A}) = \frac{1}{N} [\bar{A}^{q}]^{T} 1 \), so \( \bar{F}_{\text{avg}}^{(q)}(\bar{A}, b, N, n) = [\bar{d}_{w}^{(q)}(\bar{A})]^{T} b(N, n) \).

5. \( \bar{f}^{(\infty)}(\bar{A}, b, N, n), w_{\infty}(\bar{A}) \). \( \bar{f}^{(\infty)}(\bar{A}, b, N, n) \) is the configuration-specific consensus local relative frequency of the attribute following infinitely many rounds of repeated linear updating by each agent with his or her immediate neighbors. The accompanying random variable is \( \bar{F}^{(\infty)}(\bar{A}, N, n) \) with CDF \( G_{\bar{F}^{(\infty)}}(\bar{A}, N, n)(t) \). The corresponding network-derived vector of agent weights is \( w_{\infty}(\bar{A}) \), to be later computed in this section, and \( \bar{f}^{(\infty)}(\bar{A}, b, N, n) = [w_{\infty}(\bar{A})]^{T} b(N, n) \).

Observe that each of these scalar quantities, that is, each of these local relative frequencies of the attribute, is computed in a manner exactly parallel to the way that we computed \( x(\bar{A}, b, N, n) \); for example:

\[
x(\bar{A}, b, N, n) = [w(\bar{A})]^{T} b(N, n) \quad \text{and} \quad \bar{f}_{\text{avg}}^{(q)}(\bar{A}, b, N, n) = [d_{w}^{(q)}(\bar{A})]^{T} b(N, n).
\]

Also note that the computation of each CDF parallels the computation of \( G_{X(\bar{A}, N, n)}(t) \). We additionally observe that each of the vectors of agent weights derived above has elements that sum to 1.\(^{16}\)

We would like to now demonstrate how to compute \( w_{\infty}(\bar{A}) \), and we would like to show how the elements of this vector depend on network primitives when graph \( G(A) \) satisfies particular assumptions. First, we define the period of a node and aperiodicity of a graph:

**Definition 2** For a graph \( G(A) = (V(A), E(A)) \), the period \( pd_{i} \) of node \( i \in V(A) \) is \( pd_{i} \equiv \gcd \{ q \geq 1 : [A^{q}]_{ii} > 0 \} \), where \( \gcd \) denotes the greatest common divisor. Node \( i \) is aperiodic when \( pd_{i} = 1 \), and graph \( G(A) = (V(A), E(A)) \) is aperiodic when \( pd_{i} = 1 \) for all nodes \( i \in V(A) \).

To compute the period of a node on a graph, construct a set that contains the lengths of all possible cycles for that node and then identify the greatest common divisor among all integers in that set. A node is aperiodic when the greatest common divisor among all integers in that set equals 1, and a graph is aperiodic when every constituent node is itself aperiodic.

\(^{16}\)To see this, note that \( \bar{A} \) is row-stochastic, and row-stochasticity is preserved under matrix multiplication.
To compute $w_\infty (\bar{A})$, one necessary assumption is that the row-stochastic weighted adjacency matrix $\bar{A}$ must be primitive. Primitivity of $\bar{A}$ is equivalent to strong connectedness and aperiodicity of its directed companion graph $\mathcal{G}(\bar{A})$. Defining $w_{ij}^{(q)} \equiv [\hat{A}^q]_{ij}$ as the weight that agent $i$ assigns to agent $j$ following $q$ rounds of linear updating, we now demonstrate the existence of $w_\infty (\bar{A})$ and its computation:

**Theorem 1** If $\bar{A}$ is primitive, then $\lim_{q \to \infty} w_{ij}^{(q)} = [w_\infty^T]_j$ exists. The pair $(w_\infty^T, 1)$ is the unique dominant left eigenpair of $\bar{A}$, $w_\infty^T \bar{A} = w_\infty^T$, and $w_\infty^T 1 = 1$.

Provided that $\bar{A}$ is primitive, we compute vector $w_\infty (\bar{A})$ by solving for the left eigenvector of $\bar{A}$ corresponding to its unit eigenvalue, which happens to be the dominant, or largest, eigenvalue for $\bar{A}$. The proof of Theorem 1 can be found in Appendix K of the Supplementary Materials. As $q \to \infty$, provided that $\bar{A}$ is primitive, the weight that every agent assigns to agent $j$ converges to the same limiting value, $[w_\infty (\bar{A})]_j$. Specifically,

$$\lim_{q \to \infty} \bar{A}^q = \begin{pmatrix} w_\infty^T \\ \vdots \\ w_\infty^T \end{pmatrix}.$$

Since every agent assigns the same weight to each agent $j$ in the population, in this setting, the local relative frequency of the attribute for any configuration, $\hat{f}^{(\infty)} (\bar{A}, b, N, n) = [w_\infty (\bar{A})]^T b (N, n)$, is the same across agents. There is consensus among agents, and we refer to $\hat{f}^{(\infty)} (\bar{A}, b, N, n)$ as the consensus local relative frequency of the attribute. A variant of Theorem 1 is presented in DeGroot (1974), with infinite repeated linear updating and the vector, $w_\infty (\bar{A})$, of agent weights under consensus forming the basis for DeGroot learning.

Vector $w_\infty (\bar{A})$ can be expressed in a closed form provided that $\mathcal{G}(A)$, the unweighted graph that pairs with graph $\mathcal{G}(\bar{A})$, satisfies certain assumptions:

**Theorem 2** If graph $\mathcal{G}(A) = (\mathcal{V}(A), \mathcal{E}(A))$ is undirected, connected, and aperiodic, and all non-zero elements within every row of the corresponding matrix $\bar{A}$ have the same value, then $w_\infty (\bar{A}) = \frac{d_i}{\sum d_i} > 0$, where $[d]_i$ is the degree for agent $i$.

Provided that we satisfy the assumptions of Theorem 2, an agent’s weight under consensus is directly proportional to his or her degree: $[w_\infty (\bar{A})]_i = \frac{d_i}{\sum d_i}$. A self-loop increases an agent’s degree by one unit, which makes the degree vector equal to: $d = A 1$. The proof of Theorem 2 can be found in Appendix K of the Supplementary Materials. A result similar to that of Theorem 2 appears in DeMarzo, Vayanos, and Zwiebel (Theorem 6, 2003).
We can also establish a closed-form solution for \( w_\infty (\bar{A}) \) when \( G (A) \) is directed, aperiodic, and Eulerian. We define a directed graph to be Eulerian when it has the following properties:

**Definition 3** A directed graph \( G (A) = (V (A), E (A)) \) is Eulerian if and only if it is strongly connected and \( d^+ = d^- \).

A directed graph is Eulerian when each node’s in-degree equals its out-degree, and the graph is strongly connected. We can now present the closed-form solution for \( w_1 \bar{A} \):

**Theorem 3** If the graph \( G (A) = (V (A), E (A)) \) is Eulerian and aperiodic, and all non-zero elements within every row of the corresponding matrix \( \bar{A} \) have the same value, then \( w_1 \bar{A} = \frac{d_i^+}{1 + d_i} > 0 \), where \([d^+]_i\) is the out-degree for agent \( i \).

For this class of graphs, an agent’s weight under consensus is directly proportional to his or her out-degree, or equivalently, his or her in-degree: 

\[
\left[ w_\infty (\bar{A}) \right]_i = \frac{d_i^+}{1 + d_i} = \frac{d_i^-}{1 + d_i}.
\]

A self-loop increases both an agent’s out-degree and in-degree by one unit, so \( d^+ = d^- = A1 = A^T1 \).

We can additionally establish a closed-form solution for \( w_\infty (\bar{A}) \) for the family of random digraphs with \( N \) nodes, no self-loops, probability \( p \) of directed edge formation independent across edges, and symmetric edge weights so that 

\[
\left[ \bar{A} \right]_{ij} = \frac{1}{d_i^+} \text{ if there exists a directed edge from node } i \text{ to node } j.
\]

**Theorem 4** If \((\alpha (N) - 1) \log N \to \infty \), where \( Np = \alpha (N) \log N \), then w.h.p. \( w_\infty (\bar{A}) \) is unique and \( w_\infty (\bar{A}) \sim \frac{d_i^-}{E[|E|]} \), where \( [\bar{A}]_i = \max_{j \in N^- (i)} \frac{d_j}{d_i} \) and \( N^- (i) \) is the in-neighborhood of node \( i \). w.h.p. \( w_\infty (\bar{A}) \sim \frac{d_i^-}{E[|E|]} \) for \( N = o (N^{1/4}) \) nodes. If \( \alpha (N) = 1 + \kappa \), \( \kappa > 0 \), or \((\alpha (N) - 1) \log N = \omega (\log \log N) \), then w.h.p. \( w_\infty (\bar{A}) \sim \frac{d_i^-}{E[|E|]} \).

The statements of this theorem are made w.h.p. relative to the family of random digraphs. We take \((\alpha (N) - 1) \log N \to \infty \) so that each random digraph is strongly connected w.h.p.

We can therefore adapt some of their mathematics to provide insight into the behavior of \( w_\infty (\bar{A}) \) over this class of random digraphs.
These newly introduced closed-form expressions for \( w_\infty (\bar{A}) \), in addition to serving as sample network-derived vectors of agent weights, can also be used in research on DeGroot learning. The next example computes the set of network-derived vectors of agent weights introduced at the beginning of this section for a network with 15 nodes formed from preferential attachment:

**Example 5 (Network-Derived Vectors of Agent Weights)** Consider an economy with \( N = 15 \) agents whose interaction structure \( \mathcal{G}(A) \) is depicted in the top left of Figure 9. Assume that agents equally weight each of their linkages. Figure 9 plots the following vectors of agent weights for \( i = 1 \) and \( q = 5 \):

\[ w_{a,i} (\bar{A}), w_{a,i}^{(q)} (\bar{A}), d^-_w (\bar{A}), d^-_w (q) (\bar{A}), \text{ and } w_\infty (\bar{A}). \]

Since graph \( \mathcal{G}(A) \) in Example 5 is undirected, connected, and aperiodic, and agents equally weight each of their linkages, by Theorem 2, each agent’s weight under consensus is directly proportional to his or her degree. We observe this relation between an agent’s degree and his or her weight under consensus in the bottom right plot of Figure 9. With \( \lim_{q \to \infty} \bar{A}^q = \mathbf{1} [w_\infty (\bar{A})]^T \) and convergence of \( \bar{A}^q \) fast in this particular setting, \( w_{a,i}^{(q)} (\bar{A}) \approx w_\infty (\bar{A}) \) and \( d^-_w (q) (\bar{A}) \approx w_\infty (\bar{A}) \), which we also observe in Figure 9.

Appendix C in the Supplementary Materials characterizes \( w_{a,i}^{(q)} (\bar{A}) = ([\bar{A}^q]_{i,1})^T \) and \( \hat{f}_{i}^{(q)} (\bar{A}, b, N, n) = [w_{a,i}^{(q)} (\bar{A})]^T b (N, n) \) for all finite \( q \) and in the limit as \( q \to \infty \) in terms of the fundamental features of \( \bar{A} \). It also studies the rate of convergence of \( \hat{f}_{i}^{(q)} (\bar{A}, b, N, n) \) to \( \hat{f}^{(\infty)} (\bar{A}, b, N, n) \).

5 **When Configuration is Irrelevant: The Degeneracy of \( G_{X(\bar{A}, N, n)} (t) \)**

In the previous section, we studied particular cases of \( x (\bar{A}, b, N, n) \) and \( w (\bar{A}) \). For each case, we showed how to derive the vector of agent weights from the underlying network, and we examined how agents’ weights depend on the topological features of the network. This list of weighting vectors is certainly not exhaustive; the relevant vector of agent weights, in general, emerges naturally from the setting under study.

We now focus on the relationship between the network-derived vector of agent weights and the precursor distribution of possible local relative frequencies of the attribute. For this section, we first study the null case in which the probability distribution, \( G_{X(\bar{A}, N, n)} (t) \), is degenerate, meaning that the support of \( X (\bar{A}, N, n) \) is uniquely valued. We characterize the necessary and sufficient restrictions on both the vector of agent weights and the underlying network structure for the precursor distribution to be degenerate. Degeneracy of this
probability distribution can lead to the distribution of possible outcomes for the economy also being degenerate. After studying the case in which the probability distribution is degenerate, we then determine the necessary and sufficient conditions for the population-wide cross-sectional distribution of local relative frequencies of the attribute to be invariant to configuration. In such a setting, the population set of actions becomes invariant to configuration, which makes the outcome of the economy also invariant to configuration and the distribution of possible outcomes degenerate. We find that the conditions for degeneracy are quite restrictive. Most economic systems with interacting agents therefore tend to have probability distributions that feature some level of non-degeneracy, which makes the outcome of the economy dependent on the particular configuration of the attribute among agents. Section 6 studies the more general case of non-degeneracy.

We begin by focusing on the precursor distribution $G_{X(\mathbf{A},N,n)}(t)$ of possible local relative frequencies of the attribute. Degeneracy of this distribution arises when quantity $x(\mathbf{A},\mathbf{b},N,n)$ is fixed for all possible configurations $\mathbf{b} \in \mathcal{B}(N,n)$. When this condition holds
for every feasible $f$, we say that quantity $x(\bar{A}, b, N, n)$ is invariant to configuration:

**Definition 4** Quantity $x(\bar{A}, b, N, n)$ is invariant to configuration when $x(\bar{A}, b, N, n) = x(\bar{A}, b', N, n)$ for all configurations $b, b' \in \mathcal{B}(N, n)$, and this property holds for all feasible $n$.

When $x(\bar{A}, b, N, n)$ is invariant to configuration for each feasible level $f$, the support of $X(\bar{A}, N, n)$ takes one value and the variance of $X(\bar{A}, N, n)$ is zero. In the next theorem, we determine the necessary and sufficient restrictions on the corresponding vector of agent weights, $w(\bar{A})$, for $x(\bar{A}, b, N, n)$ to be invariant to configuration, and we solve for the support of $X(\bar{A}, N, n)$:

**Theorem 5** Scalar quantity $x(\bar{A}, b, N, n) = [w(\bar{A})]^T b(N, n)$ is invariant to configuration if and only if $[w(\bar{A})]_i = \frac{1}{N}$ for all $i \in \{1, \ldots, N\}$. When $x(\bar{A}, b, N, n)$ is invariant to configuration, $x(\bar{A}, b, N, n) = \frac{n}{N}$.

When $w(\bar{A}) = \frac{1}{N}1$, every agent has the same effective representation in the population. As a result, regardless of which agents have the attribute’s unit value, the overall contribution by those agents to that attribute’s local relative frequency is the same. The local relative frequency of the attribute is always equal to the attribute’s global relative frequency, $f$. The support of $X(\bar{A}, N, n)$ is thus $f$. The configuration of the attribute among agents in the system is irrelevant for how the system evolves, and the outcome of the economy only depends on the system’s aggregate feature, $f$.

We can similarly establish the necessary and sufficient restrictions on specific network-derived vectors of agent weights for their corresponding local relative frequencies of the binary-valued attribute to be invariant to configuration:

**Corollary 1 (to Theorem 5)** (1) Scalars $\hat{f}_i (\bar{A}, b, N, n)$ and $\hat{f}^q_i (\bar{A}, b, N, n)$ are respectively invariant to configuration if and only if, for all $j \in \{1, \ldots, N\}$, $[w_{a,i}(\bar{A})]_j = \frac{1}{N}$ and $[w^q_{a,i}(\bar{A})]_j = \frac{1}{N}$.

(2) Vectors $\hat{f}(\bar{A}, b, N, n)$ and $\hat{f}^q(\bar{A}, b, N, n)$ are respectively invariant to configuration if and only if, for all $i, j \in \{1, \ldots, N\}$, $[w_{a,i}(\bar{A})]_j = \frac{1}{N}$ and $[w^q_{a,i}(\bar{A})]_j = \frac{1}{N}$.

(3) Scalars $\hat{f}_{avg}(\bar{A}, b, N, n)$ and $\hat{f}^q_{avg}(\bar{A}, b, N, n)$ are respectively invariant to configuration if and only if, for all $i \in \{1, \ldots, N\}$, $[d^-_w(\bar{A})]_i = \frac{1}{N}$ and $[d^-_w^q(\bar{A})]_i = \frac{1}{N}$.

(4) Scalar $\hat{f}^{(\infty)}(\bar{A}, b, N, n)$ is invariant to configuration if and only if, for all $i \in \{1, \ldots, N\}$, $[w^{\infty}_w(\bar{A})]_i = \frac{1}{N}$.

When $x(\bar{A}, b, N, n) \in \left\{ \{\hat{f}_i (\bar{A}, b, N, n)\}_{i=1}^N, \{\hat{f}^q_i (\bar{A}, b, N, n)\}_{i=1}^N, \hat{f}_{avg}(\bar{A}, b, N, n), \hat{f}^q_{avg}(\bar{A}, b, N, n), \hat{f}^{(\infty)}(\bar{A}, b, N, n) \right\}$ is invariant to configuration, $x(\bar{A}, b, N, n) = \frac{n}{N}$. When
Theorem 6

(1) Scalars \( \hat{f}_i (\bar{A}, b, N, n) \) and \( \hat{f}^{(q)}_i (\bar{A}, b, N, n) \) are respectively invariant to configuration if and only if \( [\bar{A}]_{i*} = \frac{1}{N} 1^T \) and \( [\bar{A}^q]_{i*} = \frac{1}{N} 1^T \). (2) Vectors \( \hat{f} (\bar{A}, b, N, n) \) and \( \hat{f}^{(q)} (\bar{A}, b, N, n) \) are respectively invariant to configuration if and only if \( \bar{A} = \frac{1}{N} 11^T \) and \( \bar{A}^q = \frac{1}{N} 11^T \). (3) Scalars \( \hat{f}_{\text{avg}} (\bar{A}, b, N, n) \) and \( \hat{f}^{(q)}_{\text{avg}} (\bar{A}, b, N, n) \) are respectively invariant to configuration if and only if \( \bar{A} \) is doubly stochastic and \( \bar{A}^q \) is doubly stochastic. (4) Scalar \( \hat{f}^{(\infty)} (\bar{A}, b, N, n) \) is invariant to configuration if and only if \( \bar{A} \) is doubly stochastic.
When these restrictions on $\bar{A}$ or $\bar{A}^q$ are satisfied, the distributions $G_{\bar{F}_f(\bar{A},N,n)}(t)$, $G_{\bar{F}^{(q)}(\bar{A},N,n)}(t)$, $G_{\bar{F}_{avg}(\bar{A},N,n)}(t)$, $G_{\tilde{F}_{avg}(\bar{A},N,n)}(t)$, and $G_{\bar{F}(\infty)(\bar{A},N,n)}(t)$ are all degenerate for every feasible $f$.

Let us first discern the relationship between the double stochasticity of $\bar{A}$ ($\bar{A}^q$) and the invariance of $\tilde{f}_{avg}(\bar{A},b,N,n)$ to configuration. The sum of the $j^{th}$ column of $\bar{A}$ ($\bar{A}^q$) represents the total weight that every agent in the population accords to agent $j$; it is the effective representation of agent $j$ in the population. When $\bar{A}$ ($\bar{A}^q$) is doubly stochastic, every agent in the population has the same effective representation, that of one agent, so configuration becomes irrelevant. We now discern the relationship between the double stochasticity of $\bar{A}$ and the invariance of $\tilde{f}^{(\infty)}(\bar{A},b,N,n)$ to configuration. When $\bar{A}$ ($\bar{A}^q$) is both primitive and doubly stochastic, $\lim_{q \to \infty} \bar{A}^q = 1 [w_\infty(\bar{A})]^T$ is also doubly stochastic. Since $\left[1w_\infty^T(\bar{A})\right]_i = [w_\infty(\bar{A})]_j$ for each $i \in \{1,\ldots,N\}$, $\lim_{q \to \infty} \bar{A}^q = \frac{1}{N}11^T$ with $w_\infty(\bar{A}) = \frac{1}{N}1$; every agent has the same weight under consensus when $\bar{A}$ is doubly stochastic, which makes $\tilde{f}^{(\infty)}(\bar{A},b,N,n)$ invariant to configuration. The necessary restrictions on $\bar{A}$ for these various local relative frequencies of the attribute to be exactly invariant to configuration are very limiting. Many underlying graphs thus generate some level of non-degeneracy and dependence on configuration.

In the next example, we illustrate how $\tilde{f}_{avg}(\bar{A},b,N,n)$ becomes invariant to configuration and $G_{\bar{F}_{avg}(\bar{A},N,n)}(t)$ becomes degenerate for all $f$ when $\bar{A}$ is doubly stochastic:

**Example 6 (Invariance of $\tilde{f}_{avg}(\bar{A},b,N,n)$ to Configuration)** Consider an economy with $N = 15$ agents whose interaction network $G(\bar{A})$ (Figure 10) is a directed 4-regular graph with self-loops for every node. Assuming that agents equally weight each of their out-edges, $\tilde{f}_{avg}(\bar{A},b,N,n) = \frac{n}{N}$ is invariant to configuration and distribution $G_{\bar{F}_{avg}(\bar{A},N,n)}(t)$ is degenerate for all feasible $n$.

When the underlying graph $G(\bar{A})$ is regular and agents assign an equal weight to each of their linkages, $\bar{A}$ becomes doubly stochastic, so $d_w(\bar{A}) = \frac{1}{N}1$ (Figure 10, top right) and $\tilde{f}_{avg}(\bar{A},b,N,n)$ does not change with configuration (Figure 10, bottom). For all feasible $f$, the population-averaged local relative frequency of the attribute equals its global relative frequency.

We have finished examining settings in which the precursor distribution of possible local relative frequencies of the attribute is degenerate. This degeneracy of $G_X(\bar{A},N,n)(t)$ and its counterparts is often a necessary prerequisite for the distribution of possible outcomes for the economy to also be degenerate. Degeneracy of these probability distributions means that the economy is not configuration dependent; rather, the way that the economy evolves depends on $f$, the global relative frequency of the attribute and the economy’s aggregate feature, and not on the particular configuration of the attribute among agents. It is very
Figure 10: Corresponding to Example 6, directed 4-regular graph with self-loops $G(A)$ (top left), a plot of the average weighted in-degree for each agent, $d_w^\infty(A)$ (top right), and the average local relative frequency of the attribute, $\hat{f}_{\text{avg}}(\bar{A}, b, N, n)$, for every possible configuration $b \in B(N, n)$ and for all feasible global relative frequencies of the attribute, $f$ (bottom).

difficult to exactly satisfy the conditions that make these probability distributions degenerate, and therefore, most economic systems exhibit some level of dependence on configuration. Modeling the economy as if its evolution only depends on its aggregate features is, in general, incomplete.

We next transition towards characterizing environments in which the population-wide cross-sectional distribution of local relative frequencies of the attribute is invariant to configuration. When the cross-sectional distribution is invariant to configuration, the cross-sectional distribution of agent actions becomes invariant to configuration, which can lead to a unique outcome for the economy. The next theorem provides necessary and sufficient conditions for the unordered multisets, $\left\{ \hat{f}_i(\bar{A}, b, N, n) \right\}_{i=1}^N$ and $\left\{ \hat{f}_{i(q)}(\bar{A}, b, N, n) \right\}_{i=1}^N$, to be invariant to configuration. It places necessary and sufficient restrictions on $\bar{A}$ (respectively
\[ \tilde{A}^q \]:

**Theorem 7** Unordered multiset \( \left\{ \tilde{f}_i (\tilde{A}, b, N, n) \right\}_{i=1}^N \) (respectively unordered multiset \( \left\{ \tilde{f}_i^{(q)} (\tilde{A}, b, N, n) \right\}_{i=1}^N \)) is invariant to configuration if and only if the following two conditions hold:

1. the row sum of any \( n \) column vectors of \( \tilde{A} \) (respectively \( \tilde{A}^q \)) has the same multiset of elements, and this property holds for every \( n \in \{1, \ldots, \lfloor \frac{N}{2} \rfloor \} \subseteq \mathbb{Z}_+ \), and
2. \( \tilde{A} \) (respectively \( \tilde{A}^q \)) is doubly stochastic.

When these restrictions are satisfied, the cross-sectional distribution of agents’ weighted local relative frequencies of the attribute is invariant to configuration for any value \( n \). When \( n = 1 \), every column of \( \tilde{A} \) (respectively \( \tilde{A}^q \)) must have the same multiset of elements. Since all matrix elements in \( \tilde{A} \) (respectively \( \tilde{A}^q \)) must also sum to \( N \), \( \tilde{A} \) (respectively \( \tilde{A}^q \)) is doubly stochastic. It is very difficult to exactly satisfy the restrictions on \( \tilde{A} \) so that the population set of weighted local relative frequencies of the attribute is invariant to configuration. Therefore, in most settings, the cross-sectional distribution of agent actions varies with configuration and the distribution of possible outcomes for the economy is non-degenerate.

Notwithstanding these strong restrictions, there do exist networks \( G (\tilde{A}) \) for which the conditions of Theorem 7 are satisfied. When \( \tilde{A} = \frac{1}{N} 1 1^T \), so that the underlying interaction network is a complete graph with self-loops for every node, or when \( \tilde{A} = I_{N \times N} \), so that the underlying interaction network is a graph with isolates and a self-loop for every node, multiset \( \left\{ \tilde{f}_i (\tilde{A}, b, N, n) \right\}_{i=1}^N \) is invariant to configuration for all feasible \( n \). In the former case, multiset \( \left\{ \tilde{f}_i (\tilde{A}, b, N, n) \right\}_{i=1}^N = \left\{ \frac{n}{N}, \frac{n}{N}, \ldots, \frac{n}{N} \right\} \), and in the latter case, multiset \( \left\{ \tilde{f}_i (\tilde{A}, b, N, n) \right\}_{i=1}^N = \left\{ \{1\}_{i=1}^n, \{0\}_{i=n+1}^N \right\} \).

### 6 When Configuration Matters: The Non-Degeneracy of \( G_X (\tilde{A}, N, n) \) (\( t \))

We just finished studying the null setting in which the particular configuration of the attribute among agents is irrelevant and the probability distribution \( G_X (\tilde{A}, N, n) \) (\( t \)) is degenerate. We now leave this null setting behind; in this section, we develop the mathematics that enables us to fully characterize the distribution of \( X (\tilde{A}, N, n) \) when it is non-degenerate. These theoretical results hold for all possible population sizes, network topologies, and prevalences of the binary-valued attribute in the population. These theoretical findings allow us to directly map the topology of agents’ interaction network to the distributional features
of $X(\tilde{A}, N, n)$. Specifically, we map the topology of agents’ interaction network, $G(\tilde{A})$, to a network-derived vector of agent weights, $w(\tilde{A})$, and we then map the network-derived vector of agent weights to the probability distribution, $G_X(\tilde{A}, N, n)(t)$. The theoretical findings presented in this section allow us to collapse the complexities of network-based agent interactions into a simple probability distribution, $G_X(\tilde{A}, N, n)(t)$, that we can then use to construct the probability distribution of possible outcomes for the economy.

We begin by presenting results concerning the distributional features of $X(\tilde{A}, N, n)$, and we then explore how $G_X(\tilde{A}, N, n)(t)$ can remain approximately non-degenerate even for very large $N$. For several findings, we assume that each configuration, $b(N, n) \in \mathcal{B}(N, n)$, of the binary-valued attribute among agents in the system is equally likely.

### 6.1 Distributional Features of $X(\tilde{A}, N, n)$

We characterize all notable features of the distribution of possible local relative frequencies of the attribute, $G_X(\tilde{A}, N, n)(t)$.

However, before doing so, we introduce a few more pieces of notation that simplify certain expressions later on in this section. Define random variable $W(\tilde{A})$ with realization $[w(\tilde{A})]_i$, the weight for agent $i$. In this section, we are interested in the population moments of $W(\tilde{A})$. We construct these moments from the elements of the network-derived vector of agent weights, $w(\tilde{A})$; unless otherwise specified, there is no randomness in the population set of agent weights. We can similarly introduce random variables for particular cases of agent weights. Define $W_{a,j}(\tilde{A})$ for all $j \in \{1, \ldots, N\}$, $W_{a,j}^{(q)}(\tilde{A})$ for all $j \in \{1, \ldots, N\}$, $D_w^-(\tilde{A})$, $D_w^{-(q)}(\tilde{A})$, and $W_{\infty}(\tilde{A})$, whose realizations are respectively agent weights $[w_{a,j}(\tilde{A})]_i$, $[w_{a,j}^{(q)}(\tilde{A})]_i$, $[d_w^-(\tilde{A})]_i$, $[d_w^{-(q)}(\tilde{A})]_i$, and $[w_{\infty}(\tilde{A})]_i$.

We also define random variables for the degree of an undirected graph, the out-degree of a directed graph, and the in-degree of a directed graph. When $G(A)$ is undirected, the degree vector is $d(A) = A1$. When $G(A)$ is directed, the out-degree vector is $d^+(A) = A1$ and the in-degree vector is $d^-(A) = A^T1$. Define random variables $D(A)$, $D^+(A)$, and $D^-(A)$ whose realizations are respectively the degree for agent $i$, $[d(A)]_i$, the out-degree for agent $i$, $[d^+(A)]_i$, and the in-degree for agent $i$, $[d^-(A)]_i$.

The distribution $G_X(\tilde{A}, N, n)(t)$ strongly depends on the distributional features of agents’ weights. We can see this relationship most clearly when $n = 1$. For that case, $G_X(\tilde{A}, N, n)(t) = G_W(\tilde{A})(t)$, and the distribution of possible local relative frequencies of the attribute equals the distribution of agent weights. For the remainder of this section, we characterize the distributional features of $X(\tilde{A}, N, n)$ for a general $n$, and we determine how these features of $X(\tilde{A}, N, n)$ relate to network structure via the set of network-derived

---

17These results hold even when the vector of agent weights is not network-derived.
agent weights. We begin by defining the first moment of $X (\bar{A}, N, n)$:

**Theorem 8** $EX (\bar{A}, N, n) = \frac{n}{N} = f$.

The first moment of $X (\bar{A}, N, n)$ is equal to the attribute’s global relative frequency, $f$. The local relative frequency of the attribute can deviate in either direction away from the attribute’s global relative frequency, but in expectation, it must equal this value. The distribution of $X (\bar{A}, N, n)$ is consequently centered about the point in which configuration is irrelevant and only the aggregate feature, $f$, matters. Note that this result follows from assuming that each configuration is equally likely. We will see in the next section that this relationship between $EX (\bar{A}, N, n)$ and $f$ disappears once each configuration is no longer equally likely to occur.

We can extend this result about the first moment to particular cases of $X (\bar{A}, N, n)$:

**Corollary 2 (to Theorem 8)** For every random variable $X (\bar{A}, N, n) \in \left\{ \{ \hat{F}_i (\bar{A}, N, n) \}_{i=1}^N, \{ \hat{F}^{(q)}_i (\bar{A}, N, n) \}_{i=1}^N, \bar{F}_{avg} (\bar{A}, N, n), \bar{F}^{(q)}_{avg} (\bar{A}, N, n), \bar{F}^{(\infty)} (\bar{A}, N, n) \right\}$, $EX (\bar{A}, N, n) = \frac{n}{N}$.

The proof of Corollary 2 can be found in Appendix K of the Supplementary Materials. For the remainder of this section, we omit the corollaries that immediately follow from each of the presented theorems. We can construct such corollaries by substituting for each theorem the triplet $(X (\bar{A}, N, n), W (\bar{A}), w (\bar{A}))$ with one of the following triplets:

$\left( \hat{F}_i (\bar{A}, N, n), W_{a,i} (\bar{A}), w_{a,i} (\bar{A}) \right)$ for every $i \in \{1, \ldots, N\}$,

$\left( \hat{F}^{(q)}_i (\bar{A}, N, n), W^{(q)}_{a,i} (\bar{A}), w^{(q)}_{a,i} (\bar{A}) \right)$ for every $i \in \{1, \ldots, N\}$,

$\left( \bar{F}_{avg} (\bar{A}, N, n), D_w^- (\bar{A}), d_w^- (\bar{A}) \right)$, $\left( \bar{F}^{(q)}_{avg} (\bar{A}, N, n), D_w^{(q)} (\bar{A}), d_w^{(q)} (\bar{A}) \right)$, or $\left( \bar{F}^{(\infty)} (\bar{A}, N, n), W_{\infty} (\bar{A}), w_{\infty} (\bar{A}) \right)$.

We proceed to study the second moment of $X (\bar{A}, N, n)$. Assuming that each configuration is equally likely, the variance of $X (\bar{A}, N, n)$ and its limiting behavior as $N \to \infty$ are:

**Theorem 9** $\text{Var} X (\bar{A}, N, n) = \frac{n}{N} \left( 1 - \frac{n}{N} \right) \frac{N}{N-1} (N \text{Var} W (\bar{A}))$. $\text{Var} X (\bar{A}, N, n) \to 0$ at rate $N^{-1}$ as $N \to \infty$ assuming $\text{Var} W (\bar{A}) < \infty$.

The variance of the local relative frequency of the attribute directly depends on the population variance of agent weights. If there is large heterogeneity in agents’ weights, then the local relative frequency of the attribute strongly varies with configuration, and this gets reflected in the variance of the distribution. This variance is maximal when $f = 0.5$, and it monotonically decreases as $f$ moves away from 0.5. To study the behavior of $\text{Var} X (\bar{A}, N, n)$ as
the population increases in size, introduce replica graphs that both preserve existing agents’ relative weights and maintain the amount of weight accorded to a particular indexed node on the graph. By scaling the population upwards in this manner, \( \text{Var} \ (\vec{A}, N, n) \) halves as the population size doubles.

In specific settings, we can relate the variance of the weighted local relative frequency of the attribute to network primitives. When \( X (\vec{A}, N, n) = \vec{F}_{\text{avg}} (\vec{A}, N, n) \), \( \text{Var} \vec{F}_{\text{avg}} (\vec{A}, N, n) \) directly depends on \( \text{Var} D_{\vec{w}} (\vec{A}) \): the capacity for variation in the average local relative frequency of the attribute directly depends on the variance of average weighted in-degrees for the network. Meanwhile, when \( X (\vec{A}, N, n) = \vec{F} (\vec{A}, N, n) \), \( \text{Var} \vec{F} (\vec{A}, N, n) \) directly depends on the variance of degrees for the network, \( \text{Var} D (\vec{A}) \), given the network’s total number of edges; this relationship holds when graph \( G (\vec{A}) \) is undirected, connected, and aperiodic, and all non-zero elements within every row of \( \vec{A} \) have the same value (see Theorem 2).

The closed-form expression for \( \text{Var} X (\vec{A}, N, n) \) in Theorem 9 provides us with the necessary mathematics to construct a configuration-induced error bound about the outcome of an aggregated economic system. This error bound quantifies the extent to which there can be variation in the outcome of the economy. The error bound gets constructed about the outcome that results from only considering the aggregate properties of the economy and not the underlying configuration of the attribute. Since agents choose actions based on the local relative frequency of the attribute, we can use \( \text{Var} X (\vec{A}, N, n) \) to compute the extent to which the economy’s outcome varies with configuration. The topology of the network determines the size of this error bound.

In the next theorem, we show how to compute the lower and upper bounds on the support of \( X (\vec{A}, N, n) \). These values represent the lowest and highest possible local relative frequencies of the attribute given the attribute’s global relative frequency in the population. They determine the maximal extent to which the local relative frequency of the attribute can deviate from its global relative frequency given the structure of the network. From these values, we are able to bound the distribution of possible outcomes for the economy.

**Theorem 10** Construct the ordered multiset \( \{w_s\}_{s=1}^{N} \) from the elements of \( w (\vec{A}) \) so that \( w_s \leq w_{s'} \) whenever \( s \leq s' \). The lower and upper bounds on the support of \( X (\vec{A}, N, n) \) are respectively:

\[
\min \text{supp} \ X (\vec{A}, N, n) = \sum_{s=1}^{n} w_s \quad \text{and} \quad \max \text{supp} \ X (\vec{A}, N, n) = \sum_{s=N-n+1}^{N} w_s.
\]

The lower bound on the support of \( X (\vec{A}, N, n) \) is equal to the sum of the \( n \) smallest agent
weights in the population, while the upper bound is equal to the sum of the $n$ largest agent weights in the population. All possible weighted local relative frequencies of the attribute given $f$ must then fall within these two bounds.

We proceed to study the limiting behavior of $G_{X(\bar{A},N,n)}(t)$ as $N \to \infty$. To do so, we define the quantity

$$\kappa_N(\epsilon) = \frac{1}{\sum_{i=1}^{N} (\lfloor w_N(\bar{A}) \rfloor_i - \frac{1}{N})^2} \sum_{j \in \{1, \ldots, N\} \text{s.t.} \left| w_N(\bar{A}) \right|_j > N^{\frac{1}{2}} > \epsilon \sigma_N} \left( \lfloor w_N(\bar{A}) \rfloor_j - \frac{1}{N} \right)^2$$

where $\sigma_N = \left( \frac{n}{N} (1 - \frac{n}{N}) \sum_{i=1}^{N} (\lfloor w_N(\bar{A}) \rfloor_i - \frac{1}{N})^2 \right)^{1/2}$. We make the population size, $N$, explicit for the $N \times 1$ vector $w_N(\bar{A})$ of network-derived agent weights because we wish to study the behavior of $G_{X(\bar{A},N,n)}(t)$ as $N$ increases. We establish the following central limit theorem-type result:

**Theorem 11** If $\lim_{N \to \infty} \kappa_N(\epsilon) = 0$ for any $\epsilon > 0$, then $\lim_{N \to \infty} G_{X(\bar{A},N,n) - \frac{X}{\sigma_N}}(t) = \Phi(t)$ for all real $t$, where $\Phi(\cdot)$ is the standard normal CDF.

The requirement that $\lim_{N \to \infty} \kappa_N(\epsilon) = 0$ for any $\epsilon > 0$ is a Lindeberg-type condition. As the population size increases and the number of agent weights increases, the Lindeberg-type condition requires that there cannot be any subset of agent weights as $N \to \infty$ that strongly deviates from the average agent weight. When this condition holds, we informally have that $\lim_{N \to \infty} G_{X(\bar{A},N,n)}(t) \approx \Phi \left( \frac{t - n}{\sigma_N} \right)$. The distribution of weighted local relative frequencies of the attribute is asymptotically normal with mean $\frac{n}{N}$ and variance $\sigma_N^2$, the mean of $X(\bar{A},N,n)$ and the variance of $X(\bar{A},N,n)$. As the population size increases, provided that the set of agent weights is well-behaved, the population variance of agent weights tends to zero, so $\text{Var} \ X(\bar{A},N,n)$ also tends to zero. As $N \to \infty$, the number of nodes on the network indeed grows as well. The only constraint on the underlying network’s growth is that the Lindeberg-type condition continues to be satisfied. From this theoretical result, we see that, as $N \to \infty$, the particular configuration of the binary-valued attribute among agents becomes irrelevant. When every configuration is equally likely, the distribution of possible local relative frequencies of the attribute converges to a degenerate distribution positioned at the attribute’s global relative frequency. The rate at which this central limit theorem-type result applies determines the extent to which configuration is still relevant for a large-$N$ population.

The next theorem provides insight into the rate at which $X(\bar{A},N,n)$ converges to a normal distribution as the population size increases. It places an upper bound on the
maximal distance of the distribution $G_{X(\bar{A},N,n)}(t)$ to a normal distribution with the same mean and variance:

**Theorem 12** For all real $t$, 
\[|G_{X(\bar{A},N,n)}(t) - \Phi \left( \frac{t-\bar{w}_{\bar{A}}}{\sigma_N} \right)| \leq \frac{C}{\sqrt{N} \left( 1 - \frac{\bar{w}_{\bar{A}}}{N} \right)} \left( \frac{\sum_{i=1}^{N} \left| w_N(\bar{A})_i - \frac{1}{N} \right|^3}{\sum_{i=1}^{N} \left( w_N(\bar{A})_i - \frac{1}{N} \right)^2} \right)^{1/2},\]

where $C$ is an absolute constant.

The upper bound depends on $f$ and the normalized third absolute moment for the distribution of agent weights. It is a Berry-Esseen-type inequality that specifies the rate at which convergence to the normal distribution takes place by bounding the maximal error of approximation.

Beyond the statistical features of $X(\bar{A},N,n)$ provided thus far, we are also interested in the CDF of the distribution, $G_{X(\bar{A},N,n)}(t)$. The next result shows, via asymptotic expansion, how we can draw the CDF of our distribution for any feasible population size, network structure, and prevalence of the attribute in the population. Let’s begin by defining the function $J(\bar{A},N,n,t)$:

\[
J(\bar{A},N,n,t) = \Phi(t) - H_2(t) \phi(t) \left( \sum_{i=1}^{N} \hat{w}_i^3 - H_3(t) \phi(t) \left[ C_2 \left( \sum_{i=1}^{N} \hat{w}_i^4 - \frac{3}{N} \right) - \frac{1}{4N} \right] - H_5(t) \phi(t) C_3 \left( \sum_{i=1}^{N} \hat{w}_i^3 \right)^2 \right),
\]

where $\hat{w}_i = \frac{[w(\bar{A})_i - EW(\bar{A})]}{\sqrt{N \text{Var} W(\bar{A})}}$, $C_1 = \frac{1 - \frac{2\bar{w}}{N(1 - \frac{\bar{w}}{N})}}{6(\frac{1}{N} - \frac{\bar{w}}{N})^{1/2}}$, $C_2 = \frac{1 - 6(\frac{\bar{w}}{N})(1 - \frac{\bar{w}}{N})}{24(\frac{1}{N} - \frac{\bar{w}}{N})}$, $C_3 = \frac{(1 - \frac{2\bar{w}}{N})^2}{72(\frac{1}{N} - \frac{\bar{w}}{N})}$, $\phi(t) = \Phi'(t) = \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}}$, and $H_i(t) \phi(t) = (-1)^i \frac{d^i}{dt^i} \phi(t)$. Quantity $\hat{w}_i$ is constructed from the set of agent weights. We can then approximate $G_{X(\bar{A},N,n) - EX(\bar{A},N,n)}(t)$ by the function $J(\bar{A},N,n,t)$:

**Theorem 13** Provided that condition (c) holds,

\[
\left| \frac{G_{X(\bar{A},N,n) - EX(\bar{A},N,n)}}{(\text{Var} X(\bar{A},N,n))^{1/2}}(t) - J(\bar{A},N,n,t) \right| < C_4 \times \sum_{i=1}^{N} |\hat{w}_i|^5
\]

for all $t$, where $C_4$ is only a function of $\frac{n}{N}$.

**Condition (c)** (Robinson (1978)) Given $C' > 0$, there exist $\epsilon > 0$, $C > 0$, and $\delta > 0$ not depending on $N$ such that, for any fixed $t$, the number of indices $j$, for which

\[
1 - \frac{\bar{w}_{\bar{A}}}{N} \approx \frac{n}{N} \sum_{i=1}^{N} \left| w_N(\bar{A})_i - \frac{1}{N} \right|^2 \approx \frac{n}{N} \sum_{i=1}^{N} \left( w_N(\bar{A})_i - \frac{1}{N} \right)^2 = \sigma_N^2.
\]
\(|\hat{w}_j \tilde{x} - t - 2r\pi| > \epsilon, \text{ for all } \tilde{x} \in \left( C' \left[ \max_i |\hat{w}_i| \right]^{-1}, C \left[ \sum_{i=1}^{N} |\hat{w}_i|^5 \right]^{-1} \right) \) and all \(r = 0, \pm 1, \pm 2, \ldots, \text{ is greater than } \delta N, \text{ for all } N.\)

The asymptotic expansion \(J(\bar{A}, N, n, t)\) in Theorem 13 is to order \(1/N\). Condition (c) requires that the multiset \(\{\hat{w}_i\}_{i=1}^{N}\) not be clustered around too few values; it therefore also requires that the multiset of agent weights \(\{[w(\bar{A})]_i\}_{i=1}^{N}\) not be clustered around too few values. This asymptotic expansion is a general result that enables us to very strongly approximate the distribution, \(G_{x(\bar{A}, N, n)}(t)\), of weighted local relative frequencies of the attribute provided that condition (c) holds:

\[
G_{x(\bar{A}, N, n)}(t) \approx J(\bar{A}, N, n, \frac{t - EX(\bar{A}, N, n)}{(Var X(\bar{A}, N, n))^{1/2}}).
\]

The function \(J(\bar{A}, N, n, t)\) is essentially a collection of terms; the first term is the normal distribution, and the other terms represent deviations away from the normal distribution provided that they are non-zero. Note that \(\sum_{i=1}^{N} \hat{w}_i^3 = N^{-1/2} \text{Skew} W(\bar{A})\) and \(\sum_{i=1}^{N} \hat{w}_i^4 - \frac{3}{N} = N^{-1} \times (\text{Excess Kurtosis} W(\bar{A}))\). Accordingly, we can re-write the function \(J(\bar{A}, N, n, t)\) in terms of the higher-order moments of \(W(\bar{A})\):

\[
J(\bar{A}, N, n, t) = \Phi(t) - H_2(t) \phi(t) C_1 N^{-1/2} \text{Skew} W(\bar{A})
- H_3(t) \phi(t) \left[ C_2 \left( N^{-1} \text{Excess Kurtosis} W(\bar{A}) \right) - \frac{1}{4N} \right]
- H_5(t) \phi(t) C_3 N^{-1} \left( \text{Skew} W(\bar{A}) \right)^2.
\]

We can recover the central limit theorem-type result from Theorem 11 by noting that if the skewness and kurtosis of \(W(\bar{A})\) are finite, then \(\lim_{N \to \infty} J(\bar{A}, N, n, t) = \Phi(t)\). The extent to which the higher-order moments of the distribution of agent weights are non-zero determines the extent to which \(G_{x(\bar{A}, N, n)}(t)\) deviates from a normal distribution.

We now demonstrate how skewness of \(W(\bar{A})\) directly generates skewness of \(X(\bar{A}, N, n)\). Take the derivative of \(J(\bar{A}, N, n, t)\) with respect to \(t\) to find an approximating probability
density function to \( g_{X(\bar{A},N,n)-E(\bar{A},N,n)}(t) \):\(^{19}\)

\[
J'(\bar{A},N,n,t) = \frac{\partial J(\bar{A},N,n,t)}{\partial t} = \phi(t) + H_3(t) \phi(t) C_1 \sum_{i=1}^{N} \bar{w}_i^3 + H_4(t) \phi(t) C_2 \left[ \sum_{i=1}^{N} \bar{w}_i - \frac{3}{N} \right] - \frac{1}{4N} + H_6(t) \phi(t) C_3 \left( \sum_{i=1}^{N} \bar{w}_i^3 \right)^2.
\]

The second and fourth terms in the expansion depend on the skewness of \( W(\bar{A}) \). If we expand the second term, we find that it is an odd function. Provided that \( f < 0.5 \) and \( \text{Skew} W(\bar{A}) > 0 \), the second term reallocates mass away from the normal density function \( \phi(t) \) to generate positive skewness. If \( f > 0.5 \) and \( \text{Skew} W(\bar{A}) > 0 \), the second term reallocates mass away from the normal density function \( \phi(t) \) to generate negative skewness. The more heavily skewed the set of agent weights, the more heavily skewed \( X(\bar{A},N,n) \). Even though the fourth term depends on the skewness of \( W(\bar{A}) \), it is an even function, so the reallocation of mass away from the normal distribution has no effect on skewness.

Skewness of \( X(\bar{A},N,n) \) matters, particularly when the distribution is unimodal, because it determines the extent to which the median of the distribution deviates from the mean of the distribution. Then, the probability that the local relative frequency of the attribute is greater than \( f \) does not equal the probability that the local relative frequency of the attribute is less than \( f \). As a result, given a random configuration, the network topology might be such that it favors relatively higher or relatively lower local relative frequencies of the attribute. Depending on the particular setting and the particular real-world interpretation of the binary-valued attribute, this deviation of the mean from the median can be important. In Appendix D, we explore in more detail how the higher-order features of the distribution of agent weights shape the higher-order features of \( G_{X(\bar{A},N,n)}(t) \). The relationship between kurtosis of \( W(\bar{A}) \) and kurtosis of \( X(\bar{A},N,n) \) is a bit more complicated.

We see that the distributional features of \( W(\bar{A}) \) shape the distribution \( G_{X(\bar{A},N,n)}(t) \); that relationship becomes explicit when we examine the function \( J(\bar{A},N,n,t) \). Now, the vector of agent weights is itself network-derived, so ultimately, it is the topological features of agents’ interaction network that shape the distributional features of \( G_{X(\bar{A},N,n)}(t) \). When \( X(\bar{A},N,n) = \bar{F}_{\text{avg}}(\bar{A},N,n) \), for example, the shape of the CDF \( G_{\bar{F}_{\text{avg}}(\bar{A},N,n)}(t) \) depends on the statistical moments for the distribution of average weighted in-degrees. The distribution

\(^{19}\)The distance, \( \left| g_{X(\bar{A},N,n)-E(\bar{A},N,n)}(t) - J'(\bar{A},N,n,t) \right| \), can also be bounded from above. See line (14) of Robinson (1978), for example.
of average weighted in-degrees therefore determines the shape of the distribution of possible outcomes for the economy. Meanwhile, when $X(\bar{A}, N, n) = \tilde{F}^{(\infty)}(\bar{A}, N, n)$, the shape of $G_{\tilde{F}^{(\infty)}(\bar{A}, N, n)}(t)$ depends on the statistical moments for the distribution of degrees, provided that $G(A)$ is undirected, connected, and aperiodic, and all non-zero elements within every row of $\bar{A}$ have the same value (see Theorem 2). For this case, the shape of the degree distribution determines the shape of the distribution of outcomes for the economy.

We can now revisit Example 3 from Section 3 and show how to compute by hand the probability that Trump’s expected vote share exceeds 0.5, or equivalently, the probability that the average local unemployment rate exceeds 0.10:

$$\Pr \left[ \tilde{F}_{avg}(\bar{A}, N, n) > 0.10 \right] = 1 - G_{\tilde{F}_{avg}(\bar{A}, N, n)}(0.10) \approx 1 - J \left( \bar{A}, N, n, \frac{0.10 - E\tilde{F}_{avg}(\bar{A}, N, n)}{(\text{Var} \tilde{F}_{avg}(\bar{A}, N, n))^{1/2}} \right)$$

$$= 0.0707, \text{ where } \hat{w}_i = \frac{[d_w(\bar{A})]_i - ED_w(\bar{A})}{(N \text{Var} D_w(\bar{A}))^{1/2}}.$$

We have a closed-form approximation for $G_{\tilde{F}_{avg}(\bar{A}, N, n)}(t)$. The relevant vector of agent weights here is the vector of average weighted in-degrees, $d_w(\bar{A})$.

Theorem 13 is quite flexible; it allows us to draw the CDF $G_X(\bar{A}, N, n)(t)$ for every feasible population size, network topology, and prevalence of the attribute. It will also later enable us to provide a closed-form expression for the distribution of outcomes for the economy. However, if agents’ weights are clustered over too few values, then condition (c) does not hold and Theorem 13 no longer applies. We therefore proceed to provide a theoretical result that allows us to construct $g_X(\bar{A}, N, n)(t)$ in closed form in certain settings when we are unable to apply the findings of Theorem 13.

We consider an environment in which $k_\omega$ agents have the same non-zero weight $\omega$ and $N - k_\omega$ agents have zero weight. If we define the set $\mathcal{I}$,

$$\mathcal{I} = \left\{ \max \{0, n - (N - k_\omega)\}, \max \{0, n - (N - k_\omega)\} + 1, \ldots, \min \{n, k_\omega\} \right\},$$

we then have the following result:

**Theorem 14** For all $i \in \mathcal{I}$,

$$g_X(\bar{A}, N, n)(i\omega) = \frac{k_\omega \binom{N - k_\omega}{n-i}}{\binom{N}{n}}.$$

For all $i \notin \mathcal{I}$, $g_X(\bar{A}, N, n)(i\omega) = 0$. 

48
In such a setting, the probability mass function is almost everywhere zero. There are only finitely many values for which it is non-zero. Those values are integer multiples of \( \omega \), where the set of allowable integers \( i \) is restricted to those in set \( I \). The probability that the local relative frequency of the attribute equals \( i \omega \) is then equal to the fraction of all possible configurations such that there are \( i \) individuals with weight \( \omega \) and the attribute’s unit value and \( n - i \) individuals with weight zero and the attribute’s unit value. When agent \( j \) equally weights each of his linkages on the network, the resulting vector of agent weights, \( \mathbf{w}_{a,j}(\mathbf{A}) \), takes this exact form and Theorem 14 applies for constructing \( g_{\mathbf{F}_{a,j}}(\mathbf{A},N,n) (t) \).

Appendix D of the Supplementary Materials provides some additional results concerning the properties of \( G_X(\mathbf{A},N,n) (t) \). Appendix E characterizes the statistical features of the multivariate random variable \( \mathbf{F}^{(q)}(\mathbf{A},N,n) \), whose realizations are \( \tilde{f}^{(q)}(\mathbf{A},\mathbf{b},N,n) \), the population vector of weighted local relative frequencies of the attribute. Appendix F identifies those vectors of agent weights and corresponding network topologies for which \( \text{Var} X(\mathbf{A},N,n) \) is maximal. Appendix G characterizes those vectors of agent weights and those matrices, \( \mathbf{A} \), that generate identical distributions \( G_X(\mathbf{A},N,n) (t) \) and therefore identical distributions of outcomes for an economy. Lastly, Appendix H conducts sensitivity analysis, studying how a perturbation to agents’ interaction network affects \( G_X(\mathbf{A},N,n) (t) \) via its effects on the relevant network-derived vector of agent weights.

6.2 Non-Degeneracy of the Distribution \( G_X(\mathbf{A},N,n) (t) \) FOR VERY LARGE \( N \)

We now examine how \( G_X(\mathbf{A},N,n) (t) \) and its mathematical analogue \( G_{\mathbf{F}_{avg}}(\mathbf{A},N,n) (t) \) can remain approximately non-degenerate even as the population size grows very large. In Section 3, we studied the distribution, \( G_{\mathbf{F}_{avg}}(\mathbf{A},N,n) (t) \), of possible average local unemployment rates for a population of 137.5 million voters, given an October 2016 U-6 unemployment rate of 9.6 percent and an interaction network characterized by the composite graph. We observed strong configuration dependence of the average local unemployment rate for very large \( N \), with a standard deviation of 0.266 percentage points. We identified two different ways for understanding such strong configuration dependence: (1) high variance of in-degrees for the composite graph, and (2) heavy-tailedness in the distribution of agent weights, that is, heavy-tailedness in the distribution of average weighted in-degrees, \( G_{D_{\tilde{\omega}}} \) for the composite graph. We now show how properties (1) and (2) of the composite graph generate a strongly non-degenerate distribution, \( G_{\mathbf{F}_{avg}}(\mathbf{A},N,n) (t) \), even for large \( N \).

We first show how high variance of in-degrees for the composite graph generates an approximately non-degenerate distribution, \( G_{\mathbf{F}_{avg}}(\mathbf{A},N,n) (t) \), of possible average local unemployment rates. To demonstrate this relationship, we must slightly modify the structure of
the composite graph. We require voters to assign an equal weight to each of their observations, so that non-zero elements for each row of \( \tilde{A} \) take the same value; this assumption happens to be one that we were already making when carrying out analysis in Section 3. We also require every agent to have the same out-degree, \( k \), setting \( k = \frac{1}{N} 1^T d^- (A) \). This latter assumption is reasonable given that the standard deviation for the distribution of out-degrees is just 18.4, while the standard deviation for the distribution of in-degrees is 8,633.3. The next theorem makes explicit the relationship between the properties of the in-degree distribution for the slightly modified composite graph and the properties of \( \hat{F}_{\text{avg}} (\tilde{A}, N, n) \):

**Theorem 15** Let \( d^+ (A) = k1 \) and let all non-zero matrix elements in each row of \( \tilde{A} \) take the same value. Then, \( \tilde{d}^- (\tilde{A}) = \frac{1}{Nk} d^- (A) \), so

\[
\text{Var} \hat{F}_{\text{avg}} (\tilde{A}, N, n) = \frac{n}{N} \left( 1 - \frac{n}{N} \right) \frac{N}{N-1} N \left( \frac{1}{Nk} \right)^2 \text{Var} D^- (A).
\]

Construct the ordered multiset \( \{ v_s \}_{s=1}^N \) from the elements of \( d^- (A) \) so that \( v_s \leq v_{s'} \) whenever \( s \leq s' \). The lower and upper bounds on the support of \( \hat{F}_{\text{avg}} (\tilde{A}) \) are respectively

\[
\min \text{supp} \hat{F}_{\text{avg}} (\tilde{A}, N, n) = \frac{1}{Nk} \sum_{s=1}^n v_s \quad \text{and} \quad \max \text{supp} \hat{F}_{\text{avg}} (\tilde{A}, N, n) = \frac{1}{Nk} \sum_{s=N-n+1}^N v_s.
\]

Every quantity of interest in Theorem 15 directly depends on the composite graph’s in-degree distribution. For a fixed number of edges, the greater the variance of the graph’s in-degrees, the greater the variance of \( \hat{F}_{\text{avg}} (\tilde{A}, N, n) \). We compute \( \text{Var} \hat{F}_{\text{avg}} (\tilde{A}, N, n) \) and the bounds on the support for \( \hat{F}_{\text{avg}} (\tilde{A}, N, n) \) using the expressions from Theorem 15. We find that the standard deviation of \( \hat{F}_{\text{avg}} (\tilde{A}, N, n) \) is 0.00307, the minimal bound on the support of \( \hat{F}_{\text{avg}} (\tilde{A}, N, n) \) is 0.0528, and the maximal bound on the support of \( \hat{F}_{\text{avg}} (\tilde{A}, N, n) \) is 0.366. These values are quite similar to the ones that we computed exactly before, with a standard deviation of 0.00266, a minimal bound of 0.0553, and a maximal bound of 0.333. The approximate relation between \( d^- (\tilde{A}) \) and \( d^- (A) \) for the composite graph explains why the shape of \( G_{D^- (\tilde{A})} (t) \) in Figure 8 is very similar to the shape of \( G_{D^- (A)} (t) \) in Figure 7.

With \( \tilde{w}_i = \left[ \frac{d^- (\tilde{A})}{N \text{Var} D^- (\tilde{A})} \left( t \right) \right] \), we can approximate \( G_{F_{\text{avg}} (A, N, n) - \hat{F}_{\text{avg}} (\tilde{A}, N, n)} (t) \) by the function \( J (\tilde{A}, N, n, t) \):

\[\text{We allow } k \text{ to be non-integral.}\]
\[ J(\bar{A}, N, n, t) = \Phi(t) - H_2(t) \phi(t) C_1 \left( N^{-1/2} \text{Skew} D^-(A) \right) \]

\[ - H_3(t) \phi(t) \left[ C_2 \left( N^{-1} \text{Excess Kurtosis} D^-(A) \right) - \frac{1}{4N} \right] \]

\[ - H_5(t) \phi(t) C_3 \left( N^{-1} \left( \text{Skew} D^-(A) \right)^2 \right), \]

with quantities \( C_1, C_2, C_3, \phi(t), \) and \( H_i(t) \phi(t) \) defined in Theorem 13. The distribution of in-degrees shapes the CDF \( G_{\bar{A}_{avg}}(A, N, n)(t) \), and it therefore shapes the distribution of outcomes for the economy.

We now remove the previous two restrictions on the structure of the composite graph. We show, in general, how heavy-tailedness in a distribution of agent weights generates a distribution, \( G_{X(A, N, n)}(t) \), that is non-degenerate even for very large \( N \). We introduce random variable \( \tilde{W}_i(A) \) with non-negative support. \( \tilde{W}_i(A) \) denotes the effective representation of agent \( i \) in the population. This mass equals 1 on average for each agent, hence \( E\tilde{W}_i(A) = 1 \). Here, we assume that random variables \( \tilde{W}_i(A) \) for \( i \in \{1, \ldots, N\} \) are independent and identically distributed.\(^{21}\) The CDF from which the \( \tilde{W}_i(A) \)'s are drawn is \( G_{\bar{W}(A)}(t) = G_{W(A) \times N}(t) \), the distribution of network-derived agent weights scaled by \( N \). \( X(A, N, n) \) is constructed by drawing \( n \) values \( \tilde{W}_i(A) \) from \( G_{\bar{W}(A)}(t) \) and designating those agents as the ones with the attribute’s unit value: \( X(A, N, n) = \frac{\tilde{W}_1(A) + \cdots + \tilde{W}_n(A)}{N} \). We then have the following result:

**Theorem 16** Suppose that \( \text{Var} \tilde{W}_i(A) \) is finite. Then as \( n, N \to \infty \) holding \( f = \frac{n}{N} \) fixed,

\[ n^{1/2} \left( \frac{1}{f} X(A, N, n) - E\tilde{W}_i(A) \right) \xrightarrow{d} N \left( 0, \text{Var} \tilde{W}_i(A) \right) \]

where \( E\tilde{W}_i(A) = 1 \). Next suppose that \( \text{Pr} \left[ \tilde{W}_i(A) > t \right] \sim L(t) t^{-\xi} \), where \( L(t) \) is a slowly-varying function and \( \xi \in (1, 2) \). Then as \( n, N \to \infty \) holding \( f = \frac{n}{N} \) fixed,

\[ n^{1-1/\xi} \left( \frac{1}{f} X(A, N, n) - E\tilde{W}_i(A) \right) \xrightarrow{d} \tilde{S}(\xi, \beta, \gamma, 0; 1), \]

where \( \tilde{S}(\xi, \beta, \gamma, 0; 1) \) is a stable distribution and \( E\tilde{W}_i(A) = 1 \).

When \( \xi \in (1, 2) \), agent weights are power-law distributed with a finite mean and infinite variance.\(^{22}\) Empirically estimating the power-law exponent, we find that \( \xi \approx 1.69 \in \)

\(^{21}\)In reality, the random variables \( \tilde{W}_i(A) \) are not independent, as they are constrained to sum to \( N \): \( \tilde{W}_1(A) + \cdots + \tilde{W}_N(A) = N \).

\(^{22}\)The literature tends to use \( \alpha \) instead of \( \xi \), but the former variable has already been assigned a different interpretation in the text.
This exponent emerges from the right tail of the counter-cumulative distribution function of agents’ effective representations in the population, where \( \tilde{W}(\bar{A}) = W(\bar{A}) \times N \) and an agent’s weight is equal to that agent’s average weighted in-degree for the composite graph. With \( n^{1-1/\xi} < n^{1/2} \) for all \( \xi \in (1, 2) \), the distribution for \( X(\bar{A}, N, n) \) collapses faster to a degenerate distribution as \( N \to \infty \) when agent weights have a finite variance than when they are drawn from a power-law distribution with infinite variance. The rate at which the law of large numbers applies is relatively slower in this latter case, so the variance of \( X(\bar{A}, N, n) \) becomes non-negligible even at extremely large population sizes, \( N \). Accordingly, in the setting of Section 3, distribution \( G_{\bar{A}, N, n}(t) \) is approximately degenerate for \( N = 137.5 \) million when the underlying interaction network is the base graph, but it is approximately non-degenerate for \( N = 137.5 \) million when the underlying interaction network is the composite graph.

We see that economies with all possible population sizes can be configuration dependent. Even in large-\( N \) settings, the aggregate properties of the system are not sufficient to determine how the economy will evolve. We need to account for the configuration dependence of the system so that we can construct the entire distribution of possible outcomes.

7 Features of the Precursor Distribution When Configurations are Not Equally Likely

We continue to characterize the precursor distribution of possible local relative frequencies of the attribute given the economy’s aggregate feature, \( f = \frac{n}{N} \). However, we relax the assumption that every configuration \( b(N, n) \in B(N, n) \) of the attribute is equally likely. Before, when each configuration of the attribute was equally likely, every agent \( i \) had the same probability of \( b_i = 1 \). Now, every agent \( i \) in the system has an arbitrary probability that \( b_i = 1 \), so configurations can occur with any relative likelihood. In this relaxed setting, we solve for the first two moments of the resulting probability distribution of local relative frequencies of the attribute.

We let each agent \( i \) have a vector of characteristics, \( \gamma_i \), that can impact \( \phi_i = \Pr[B_i = 1|\gamma_i] \), the conditional probability that agent \( i \) has the binary-valued attribute’s unit value. \( B_i \) is a random variable whose realization is agent \( i \)’s binary-valued attribute: 0 or 1. We partition agent indices into \( \Theta \) categories according to their conditional probabilities, so that agents \( i, j \) are in category \( \theta \) if \( \phi_i = \phi_j = \rho_\theta \). We define the odds ratio for agents in category \( \theta \) relative to category \( k \) as follows: \( \hat{\psi}_\theta = \frac{\frac{\rho_\theta}{1-\rho_\theta}}{\frac{\rho_k}{1-\rho_k}} \), with \( \hat{\psi}_k \equiv 1 \). Then:
Theorem 17  The first two moments of $X\left(\bar{A},N,n,(\gamma_i)^N_{i=1}\right)$ are:

$$EX\left(\bar{A},N,n,(\gamma_i)^N_{i=1}\right) = \sum_{i=1}^{N} \left[ w(\bar{A}) \right]_i [\mu]_i$$

and $\text{Var} X\left(\bar{A},N,n,(\gamma_i)^N_{i=1}\right) = \left[ w(\bar{A}) \right]^T \Sigma \left[ w(\bar{A}) \right].$

To compute the $N \times 1$ vector $\mu$ and the $N \times N$ matrix $\Sigma$, define the $\Theta \times 1$ vector $\tilde{\mu}$ across the $\Theta$ categories; set $[\mu]_i = [\tilde{\mu}]_\theta$ for each agent $i$ from category $\theta$. Also introduce the $N \times 1$ vector $\zeta$, setting $[\zeta]_i = [\tilde{\mu}]_\theta (1 - [\tilde{\mu}]_\theta)$ for each agent $i$ from category $\theta$. Define the $\Theta \times \Theta$ matrix $\tilde{\Sigma}$ with element $\left[ \tilde{\Sigma} \right]_{\theta_k}$ equal to the conditional covariance $\text{Cov} (B_i, B_j)$ between agent $i$ in category $\theta$ and agent $j$ in category $k$ and element $\left[ \tilde{\Sigma} \right]_{\theta \theta}$ equal to the conditional variance $\text{Var} B_i$ for agent $i$ in category $\theta$. $\mu$ and $\Sigma$ can be approximated by solving the following system of equations:

$$\sum_{\theta=1}^{\Theta} \sum_{\begin{array}{c}i \in \{1, \ldots, N\} \\text{s.t. } \phi_i = \rho_\theta \end{array}} [\tilde{\mu}]_\theta = n$$

$$\hat{\psi}_\theta = \frac{[\tilde{\mu}]_\theta (1 - [\tilde{\mu}]_k) - \left[ \tilde{\Sigma} \right]_{\theta k}}{(1 - [\tilde{\mu}]_\theta) [\tilde{\mu}]_k - \left[ \tilde{\Sigma} \right]_{\theta k}}, \quad \forall \theta \in \{1, \ldots, \Theta\} \setminus \{k\}, \text{ and}$$

$$\Sigma = \frac{N}{N-1} \left( \text{diag} \zeta - \zeta \zeta^T \zeta \right).$$

The $N \times 1$ random vector $\mathbf{B}$, whose $i^{th}$ element is random variable $B_i$, is distributed according to Fisher’s multivariate non-central hypergeometric distribution. $\mu = EB$ is the $N \times 1$ conditional mean vector for $\mathbf{B}$ and $\Sigma$ is the $N \times N$ conditional covariance matrix for $\mathbf{B}$. $\mathbf{B}$, $\mu$, and $\Sigma$ are quantities that correspond to the population of $N$ agents. We can also introduce parallel hatted quantities that correspond to the population of $\Theta$ distinct categories. If we define a $\Theta \times 1$ random vector $\tilde{\mathbf{B}}$, whose $\theta^{th}$ element is random variable $B_i$ from category $\theta$, then $\tilde{\mu} = EB$ is the corresponding $\Theta \times 1$ conditional mean vector for $\tilde{\mathbf{B}}$ and $\tilde{\Sigma}$ is the corresponding $\Theta \times \Theta$ conditional covariance matrix for $\tilde{\mathbf{B}}$; diagonal element $\left[ \tilde{\Sigma} \right]_{\theta \theta}$ is the conditional variance $\text{Var} B_i$ for agent $i$ in category $\theta$. We also introduce the $\Theta \times 1$ vector $\tilde{\Sigma}^{\text{Cov}}; the \theta^{th}$ element $\left[ \tilde{\Sigma}^{\text{Cov}} \right]_\theta$ of this vector is the conditional covariance $\text{Cov} (B_i, B_j)$ for agent $i$ and agent $j$, $i \neq j$, both in category $\theta$.

When the population of agents can be partitioned into $\Theta$ categories, the total number of variables and the total number of equations in the system both equal $2\Theta + \binom{\Theta}{2} + \sum_{\theta=1}^{\Theta} 1_{s_\theta > 1}$, where $s_\theta$ is the number of agents in category $\theta$. There are $\Theta$ variables $[\tilde{\mu}]_\theta$, the conditional
2016 U.S. presidential candidate to elect. Here, we have the first two moments of second economy features different economies with.

simply follow Theorem 17 and make the necessary substitutions. Appendix I studies three cases of we studied in Section 6 when every configuration was equally likely. The corresponding

When \( \phi_i = \Pr [B_i = 1 \mid \gamma_i] \) differs across agents, \( \mathbb{E}X(\mathbf{A}, N, n, (\gamma_i)_{i=1}^{N}) \) no longer must equal the global relative frequency of the attribute, \( f \). The weighted local relative frequency of the attribute can, on average, be either greater or less than \( f \). In the social learning setting, when configurations are not equally likely, it may not be possible for there to be asymptotic convergence to the truth, as the mean of the distribution of \( X(\mathbf{A}, N, n, (\gamma_i)_{i=1}^{N}) \) need not equal \( f \) for all iterations of learning \( q \) and all population sizes \( N \), even as \( q, N \to \infty \). When configurations are no longer equally likely to occur, we can induce bias in \( \mathbb{E}X(\mathbf{A}, N, n, (\gamma_i)_{i=1}^{N}) \) away from the global relative frequency, \( f \).

The next example works through the null case, in which every configuration is equally likely; it computes the first two moments of \( X(\mathbf{A}, N, n, (\gamma_i)_{i=1}^{N}) \) when \( \Theta = 1 \), following Theorem 17:

**Example 7 (First Two Moments of \( X(\mathbf{A}, N, n, (\gamma_i)_{i=1}^{N}) \), \( \Theta = 1 \))** Consider an economic system with \( N \) agents. When \( \Theta = 1 \), so that \( \phi_i = \rho_1 \) for every agent \( i \in \{1, \ldots, N\}, \) \( \mathbb{E}X(\mathbf{A}, N, n, (\gamma_i)_{i=1}^{N}) = \mathbb{E}X(\mathbf{A}, N, n) \) and \( \text{Var} X(\mathbf{A}, N, n, (\gamma_i)_{i=1}^{N}) = \text{Var} X(\mathbf{A}, N, n) \), where \( X(\mathbf{A}, N, n) \) is the random variable of interest when every configuration \( b(N, n) \in \mathcal{B}(N, n) \) is equally likely.

When \( \Theta = 1 \) we recover the first two moments \( \mathbb{E}X(\mathbf{A}, N, n) \) and \( \text{Var} X(\mathbf{A}, N, n) \) that we studied in Section 6 when every configuration was equally likely. The corresponding derivation is in Appendix I of the Supplementary Materials.

In the more general setting, we can characterize these first two moments for particular cases of \( X(\mathbf{A}, N, n, (\gamma_i)_{i=1}^{N}) \), such as \( \hat{F}_{\text{avg}}(\mathbf{A}, N, n, (\gamma_i)_{i=1}^{N}) \) and \( \hat{F}^{(\infty)}(\mathbf{A}, N, n, (\gamma_i)_{i=1}^{N}) \); we simply follow Theorem 17 and make the necessary substitutions. Appendix I studies three different economies with \( \Theta \neq 1 \). The first economy features \( N = 4 \) agents and \( \Theta = 2 \) categories of agents; we compute the first two moments of \( \hat{F}_{\text{avg}}(\mathbf{A}, N, n, (\gamma_i)_{i=1}^{N}) \). The second economy features \( N = 15 \) agents and \( \Theta = 3 \) categories of agents; we compute the first two moments of \( \hat{F}^{(\infty)}(\mathbf{A}, N, n, (\gamma_i)_{i=1}^{N}) \). The third economy revisits the setting of Section 3 and Example 4, in which there are \( N = 137.5 \) million voters deciding which 2016 U.S. presidential candidate to elect. Here, we have \( \Theta = 2 \) categories of agents: (1)
those featured by the media, and (2) those not featured by the media. Setting \( \rho_1 = 0.50 \) and \( \rho_2 = 0.096 \), Appendix I shows how to compute \( E \tilde{F}_{\text{avg}}(\bar{A}, N, n, (\gamma_i)_{i=1}^N) = 0.194 \) and \( \text{Std. Dev.} \tilde{F}_{\text{avg}}(\bar{A}, N, n, (\gamma_i)_{i=1}^N) = 0.00452 \), the first two moments of the distribution of possible average local unemployment rates given that the actual unemployment rate, \( f \), is 0.096. In all settings, the first two moments of the distribution deviate from the ones that would occur when configurations are equally likely.

8 THE DISTRIBUTION OF OUTCOMES FOR THE ECONOMY

Thus far, we have developed the mathematics that enables us to characterize the distributional features of \( X(\bar{A}, N, n) \) and \( X(\bar{A}, N, n, (\gamma_i)_{i=1}^N) \). In this section, we compute the distribution of possible outcomes for the economy. For certain classes of agent actions, we can provide a closed-form expression for the distribution of possible outcomes. We already demonstrated in Section 3 how to compute such a distribution of possible outcomes for the economy. In that setting, agents made a voting decision based on their locally formed macroeconomic sentiments; we computed both the probability that the election outcome favored candidate Clinton and the probability that the election outcome favored candidate Trump. In this section, we show how to construct the distribution of possible outcomes for the economy when agents follow other decision-making rules. The classes of agent decision-making rules that we identify and study in this section are certainly not exhaustive.

Before we can study the distribution of possible outcomes for the economy, we must introduce some additional notation. Constructing this distribution requires us to incorporate agents’ decision-making behavior, so much of the notation concerns agent actions. We assume that each agent \( i \) chooses an action, \( a_i(\bar{A}, b, N, n) \), that depends on his or her local relative frequency of the attribute: \( a_i(\bar{A}, b, N, n) = h_i(x_i(\bar{A}, b, N, n)) \). \( h_i(\cdot) \) is a function that specifies how agent \( i \) responds to his or her weighted local relative frequency of the attribute, \( x_i(\bar{A}, b, N, n) \), where \( x_i(\bar{A}, b, N, n) = [w_i(\bar{A})]^T b(N, n) \). Since agents may respond to different local relative frequencies of the same attribute, we index by agent the quantity \( x_i(\bar{A}, b, N, n) \) and the vector of agent weights \( w_i(\bar{A}) \).

We are interested in the individual action \( a_i(\bar{A}, b, N, n) \) and how it varies with configuration given \( f \), and we are also interested in the population’s aggregate action \( a_{\text{agg}}(\bar{A}, b, N, n) = \sum_{i=1}^N a_i(\bar{A}, b, N, n) \) and how that quantity varies with configuration given \( f \). We define random variables \( A_i(\bar{A}, N, n) \) and \( A_{\text{agg}}(\bar{A}, N, n) \) with respective configuration-specific realizations \( a_i(\bar{A}, b, N, n) \) and \( a_{\text{agg}}(\bar{A}, b, N, n) \). In this section, we study the distributions of \( A_i(\bar{A}, N, n) \) and \( A_{\text{agg}}(\bar{A}, N, n) \). When each configuration of an attribute is equally likely to
outcome in which configuration is irrelevant and only the aggregate feature of the econ-

By computing the mean action for agent $a_i(\bar{A}, b, N, n)$, we can write an
distinct object is the action of agent $i$. Therefore, we study agent $i$’s action, $a_i(\bar{A}, b, N, n)$, and we seek to characterize the distribution $G_{A_i(\bar{A}, N, n)}(t)$. We focus on the case in which agents’ weights are such that they satisfy condition (c) of Theorem 13. We can write an expression for the distribution $G_{A_i(\bar{A}, N, n)}(t)$ provided that agent $i$’s action, $a_i(\bar{A}, b, N, n) = h_i(x_i(\bar{A}, b, N, n))$, is invertible over the domain $x_i(\bar{A}, b, N, n) \in [0, 1]$.\textsuperscript{23} Our expression for $G_{A_i(\bar{A}, N, n)}(t)$ is then:

$$G_{A_i(\bar{A}, N, n)}(t) = G_{X_i(\bar{A}, N, n)}(h_i^{-1}(t)) \approx J \left( \bar{A}, N, n, \frac{h_i^{-1}(t) - EX_i(\bar{A}, N, n)}{(Var X_i(\bar{A}, N, n))^{1/2}} \right),$$

where $EX_i(\bar{A}, N, n) = \frac{n}{N}$, $Var X_i(\bar{A}, N, n) = \frac{n}{N} \left(1 - \frac{n}{N}\right) \frac{N}{N-1} (N Var W_i(\bar{A}))$, and $\hat{w}_j = \frac{[w_{a,i}(\bar{A})]_j - EW_{a,i}(\bar{A})}{(N Var W_{a,i}(\bar{A}))^{1/2}}$. When agent $i$’s action takes an affine form:

$$a_i(\bar{A}, b, N, n) = h_i(x_i(\bar{A}, b, N, n)) = \alpha_i x_i(\bar{A}, b, N, n) + \beta_i,$$

action $a_i(\bar{A}, b, N, n)$ is invertible with respect to $x_i(\bar{A}, b, N, n)$, so we can express $G_{A_i(\bar{A}, N, n)}(t)$ in closed form:

$$G_{A_i(\bar{A}, N, n)}(t) \approx J \left( \bar{A}, N, n, \frac{t - \beta_i - EX_i(\bar{A}, N, n)}{(Var X_i(\bar{A}, N, n))^{1/2}} \right) = J \left( \bar{A}, N, n, \frac{t - EA_i(\bar{A}, N, n)}{(Var A_i(\bar{A}, N, n))^{1/2}} \right).$$

By computing the mean action for agent $i$, $EA_i(\bar{A}, N, n) = \alpha_i \frac{n}{N} + \beta_i$, and we see that the distribution of possible outcomes for the economy, $G_{A_i(\bar{A}, N, n)}(t)$, is centered about the outcome in which configuration is irrelevant and only the aggregate feature of the economy, $f = \frac{n}{N}$, matters. Appendix J of the Supplementary Materials features two examples

\textsuperscript{23}To see why there exists a closed-form expression for $G_{A_i(\bar{A}, N, n)}(t)$ when agent $i$’s action, $a_i(\bar{A}, b, N, n) = h_i(x_i(\bar{A}, b, N, n))$ is invertible, follow the derivation below:

$$G_{A_i(\bar{A}, N, n)}(t) = \Pr \left[ A_i(\bar{A}, N, n) \leq t \right]$$

$$= \Pr \left[ h_i(x_i(\bar{A}, N, n)) \leq t \right], \text{ with supp } X_i(\bar{A}, N, n) \in [0, 1]$$

$$= \Pr \left[ X_i(\bar{A}, N, n) \leq h_i^{-1}(t) \right]$$

$$= G_{X_i(\bar{A}, N, n)}(h_i^{-1}(t)) \approx J \left( \bar{A}, N, n, \frac{h_i^{-1}(t) - EX_i(\bar{A}, N, n)}{(Var X_i(\bar{A}, N, n))^{1/2}} \right).$$
that solve for $G_{\mathcal{A}_i}(\mathbf{A},N,n)(t)$ and its distributional characteristics in closed form. For the first example, agent $i$’s action is an affine transformation of the weighted local relative frequency of the attribute, $a_i(\mathbf{A}, b, N, n) = \alpha_i \tilde{\mathbf{f}}_i(\mathbf{A}, b, N, n) + \beta_i$, and for the second example, agent $i$’s action nonlinearly depends on the weighted local relative frequency of the attribute, $a_i(\mathbf{A}, b, N, n) = \alpha_i \log \tilde{\mathbf{f}}_i^{(q)}(\mathbf{A}, b, N, n) + \beta_i$. In both settings, we can write a closed-form expression for $G_{\mathcal{A}_i}(\mathbf{A},N,n)(t)$, the distribution of possible outcomes for the economy, for every feasible population size, network topology, and prevalence of the binary-valued attribute’s unit value in the population.

8.2 The Distribution of Outcomes Given the Aggregate Action

We now compute the distribution of possible outcomes for the economy when the relevant object is the aggregate action. We therefore study the aggregate action, $a_{agg}(\mathbf{A}, b, N, n)$, with an interest in characterizing the distribution $G_{A_{agg}}(\mathbf{A},N,n)(t)$. We assume that agents’ weights are such that they satisfy condition (c) of Theorem 13. To characterize $G_{A_{agg}}(\mathbf{A},N,n)(t)$ in closed form, we focus on particular classes of agent actions.

Our main class of agent actions is the one in which individual agents’ actions take an affine form: $a_i(\mathbf{A}, b, N, n) = \alpha_i \tilde{\mathbf{f}}_i(\mathbf{A}, b, N, n) + \beta_i$, $\forall i \in \{1, \ldots, N\}$. We can then write a closed-form expression for $G_{A_{agg}}(\mathbf{A},N,n)(t)$ for every feasible population size, underlying network topology, and global relative frequency of the attribute. Observe that the mean aggregate action, $E A_{agg}(\mathbf{A}, N, n)$, and therefore the mean outcome for the economy, is the one that occurs when the system is invariant to configuration and only the aggregate feature of the system matters. The extent to which the aggregate action can deviate away from this mean aggregate action determines how dependent the system is on configuration. When $\alpha_i = \alpha$ for all agents, $G_{A_{agg}}(\mathbf{A},N,n)(t) \approx J \left( \mathbf{A}, N, n, \frac{t - \alpha N E_{avg}(\mathbf{A},N,n) - 1^T \beta}{\alpha N (\text{Var} \ E_{avg}(\mathbf{A},N,n))^{1/2}} \right)$, with $J(\cdot)$ defined in Theorem 13 and $\hat{w}_i = \frac{[d_w(\mathbf{A})]_i - ED_w(\mathbf{A})}{(N \text{Var} D_w(\mathbf{A}))^{1/2}}$. When agents no longer have a common coefficient $\alpha$, $G_{A_{agg}}(\mathbf{A},N,n)(t) \approx J \left( \mathbf{A}, N, n, \frac{t - (1^T \alpha) E_{avg}(\mathbf{A},N,n) - 1^T \beta}{(1^T \alpha) (\text{Var} E_{avg}(\mathbf{A},N,n))^{1/2}} \right)$, with $J(\cdot)$ defined in Theorem 13, $\hat{w}_i = [d_w(\mathbf{A})]_i - ED_w(\mathbf{A}) (N \text{Var} D_w(\mathbf{A}))^{1/2}$, $\hat{\mathbf{A}}_{ij} = \alpha_i [\mathbf{A}]_{ij}$, and $d_w(\mathbf{A}) = \left( \frac{1}{\alpha_1 + \ldots + \alpha_N} \right) \hat{\mathbf{A}}^T \mathbf{1}$. Appendix J of the Supplementary Materials solves for $G_{A_{agg}}(\mathbf{A},N,n)(t)$ and its distributional features for these two settings, when agents have a common coefficient $\alpha$ and when agents no longer have a common coefficient $\alpha$.

When agents’ actions instead depend on the consensus local relative frequency of the attribute, there exists an even larger class of decision rules for which we can write a
closed-form expression for $G_{A_{agg}}(\bar{A}, N, n)(t)$. If agent actions follow a threshold rule:

$$a_i(\bar{A}, b, N, n) = \begin{cases} \beta_i & \text{if } \hat{f}^{(\infty)}(\bar{A}, b, N, n) \geq \alpha \\
0 & \text{if } \hat{f}^{(\infty)}(\bar{A}, b, N, n) < \alpha \end{cases},$$

$$A_{agg}(\bar{A}, N, n) \approx \begin{cases} 1^T \beta & \text{with probability } 1 - J(\bar{A}, N, n, \frac{\alpha - E\hat{F}(\infty)(\bar{A}, N, n)}{(\Var \hat{F}(\infty)(\bar{A}, N, n))^{1/2}}) \\
0 & \text{with probability } J(\bar{A}, N, n, \frac{\alpha - E\hat{F}(\infty)(\bar{A}, N, n)}{(\Var \hat{F}(\infty)(\bar{A}, N, n))^{1/2}}) \end{cases},$$

with $J(\cdot)$ defined in Theorem 13 and $\hat{\mu}_i = \frac{w_{\omega}(\bar{A})}{N \Var W_\omega(\bar{A})} - EW_\omega(\bar{A})$.

Even if we cannot express $G_{A_{agg}}(\bar{A}, N, n)(t)$ in closed form, we can still solve for its lower-order features. If agent $i$’s action, $\forall i \in \{1, \ldots, N\}$, follows a different threshold rule:

$$a_i(\bar{A}, b, N, n) = \begin{cases} \beta_i & \text{if } \hat{f}_i(\bar{A}, b, N, n) \geq \alpha \\
0 & \text{if } \hat{f}_i(\bar{A}, b, N, n) < \alpha \end{cases},$$

we can solve for $EA_{agg}(\bar{A}, N, n)$ in closed form:

$$EA_{agg}(\bar{A}, N, n) \approx \sum_{i=1}^{N} \beta_i \left[ 1 - J(\bar{A}, N, n, \frac{\alpha - E\hat{F}_i(\bar{A}, N, n)}{(\Var \hat{F}_i(\bar{A}, N, n))^{1/2}}) \right],$$

with $J(\cdot)$ defined in Theorem 13 and $\hat{\mu}_j = \frac{w_{a,i}(\bar{A})}{N \Var W_{a,i}(\bar{A})} - EW_{a,i}(\bar{A})$. Appendix J studies in greater detail these two cases in which agent actions follow a threshold rule.

Given agents’ decision-making behavior, we are able to both construct the distribution of possible outcomes for the economy and study the distribution’s statistical features. We can assess the non-degeneracy of the distribution of outcomes and therefore determine the extent to which the economy is dependent on configuration. When the outcome of the economy strongly varies with configuration, the economy’s aggregate feature, $f$, is no longer sufficient for determining how the economy will evolve, and agents’ local interactions and local environments matter for the behavior of the overall economy. For various classes of agent decision-making rules, we can write closed-form expressions for $G_{A_i(\bar{A}, N, n)}(t)$ and $G_{A_{agg}(\bar{A}, N, n)}(t)$ for any feasible population size, network topology, and aggregate feature, and we can study how the statistical features of the distributions $G_{A_i(\bar{A}, N, n)}(t)$ and $G_{A_{agg}(\bar{A}, N, n)}(t)$ emerge from the topology of agents’ interaction network.
9 Conclusion

In this work, we develop a set of mathematical tools that allows us to map network structures to probability distributions. Given agents’ decision-making behavior, we are able to map the specific topology of agents’ interaction network to a specific probability distribution of possible outcomes for the economy. We first map the structure of agents’ network, \( G(\bar{A}) \), to a precursor distribution, \( G_X(\bar{A},N,n)(t) \), of possible local relative frequencies of the binary-valued attribute. Then, given agents’ decision-making behavior and the precursor distribution, \( G_X(\bar{A},N,n)(t) \), we construct the distribution of possible outcomes for the economy.

Our characterization of \( G_X(\bar{A},N,n)(t) \) is complete. For all population sizes, feasible network structures, and possible global prevalences of the attribute in the population, we can solve for the distributional features of \( G_X(\bar{A},N,n)(t) \) in closed form, and we can provide a closed-form expression for the actual shape of this probability distribution. Our mathematical tools show how network primitives and other features of the underlying network directly generate probability distributions with certain statistical features. Meanwhile, for particular classes of agent actions, we can write a closed-form expression for the distribution of possible outcomes for the economy as well; for a larger set of agent actions, we can solve for the lower-order features of this distribution in closed form. When this distribution is non-degenerate, the particular configuration of the attribute among agents matters for how the economy evolves.

The tools and content developed in this work have several implications. First, these mathematical tools enable closed-form analysis of complex economic systems. Such tools allow us to map complex agent interactions into a simple probability distribution characterizing how the system will probabilistically evolve. Second, these mathematical tools allow us to quantify how dependent an aggregate economic system is on the underlying configuration of an attribute among its agents. Aggregate treatments of economic systems put forth a single outcome for the economy, but if the economic system is configuration dependent, there is actually an entire non-degenerate distribution of possible outcomes for the economy consistent with the aggregate properties of the system. The tools of this work allow us to construct in closed form an error bound about the original benchmark outcome of the aggregate economy. Third, the theoretical framework of this work and the accompanying mathematical tools help us to understand locally formed macroeconomic sentiment. This work offers a microfoundation for animal spirits, showing how agents’ interaction structure enables the existence of swings in aggregate sentiment for fixed economic fundamentals that persist even for large-N economies. Schlossberger (2018) utilizes and extends the mathemat-
ical tools and theoretical framework developed in this work for an entirely different economic application. Hopefully the theoretical tools and methodology developed in this work can be broadly used to provide insights in diverse settings.
REFERENCES


APPENDIX: PROOFS

Additional proofs can be found in Appendix K of the Supplementary Materials.

Proof of Theorem 3

Since graph $G(A)$ is directed and all non-zero elements within every row of $\tilde{A}$ have the same value, either $[\tilde{A}]_{ij} = \frac{1}{d^+}$ or $[\tilde{A}]_{ij} = 0$. Conjecture the solution $w_\infty = \frac{d^+}{1^T d^+}$ to $w_\infty^T \tilde{A}$, where $d^+$ is the vector of out-degrees for graph $G(A)$. Then,

$$w_\infty^T \tilde{A} = \left[ \frac{d^+}{1^T d^+} \right]^T \tilde{A} = \left( \sum_{i=1}^N \frac{d^+ [\tilde{A}]_{ji}}{1^T d^+} \right) = \left( \sum_{i=1}^N \frac{d^-_{ji}}{1^T d^+} \right) = \left( \sum_{i=1}^N \frac{d^+_{ji}}{1^T d^+} \right) = \left[ \frac{d^+}{1^T d^+} \right]^T = w_\infty^T,$$

with $d^-_{ji} = d^+_{ji}$ because graph $G$ is Eulerian. $1^T w_\infty = 1^T \left[ \frac{d^+}{1^T d^+} \right] = 1$, so $1^T w_\infty = 1$. □

Proof of Theorem 4

By viewing $w_\infty$ as mathematically analogous to the stationary distribution of a simple random walk on a random digraph, the statements in this Theorem follow from Theorem 2, Lemma 14, and the Remarks of Theorem 2 from Cooper and Frieze (2012). When $\alpha(N) = 1 + \kappa$, $\kappa > 0$, or $(\alpha(N) - 1) \log N = \omega(\log \log N)$, w.h.p. $t_i = o(d^-_i)$, so it follows that w.h.p. $w_\infty \sim \frac{d^-}{E[|E|]}$. Similarly, by Lemma 14 of Cooper and Frieze (2012), w.h.p. $t_i = o(d^-_i)$ for $N - o(1/N^{1/4})$ nodes, so w.h.p. $w_\infty \sim \frac{d^-}{E[|E|]}$ for $N - o(1/N^{1/4})$ nodes. □

Proof of Theorem 5

The scalar $x = w^T b$ is invariant to configuration if and only if $w^T b(N, n) = w^T b'(N, n)$ for all $b(N, n), b'(N, n) \in B(N, n)$, with this relation holding for each integer $n \in [0, N]$. Let $n = 1$, and define $e_i$ to be the $i^{th}$ unit vector whose $i^{th}$ element equals 1 and all other elements equal zero. Then $w^T b(N, 1) = w^T b'(N, 1)$ if and only if $w^T e_i = w^T e_j$, that is, if and only if $w_i = w_j$ for all $i, j \in \{1, \ldots, N\}$. Since $1^T w = 1$, $w_i = \frac{1}{N}$ for all $i \in \{1, \ldots, N\}$. Given that $w = \frac{1}{N} 1$ when $x$ is invariant to configuration, $x = w^T b = \frac{1}{N} 1^T b = \frac{b}{N}$. □

Proof of Theorem 6

This proof employs the statements of Corollary 1. For part (1), $\widehat{f}_i(\tilde{A}, b, N, n)$ is invariant to configuration if and only if $w_{a,i} = \frac{1}{N} 1$. Since $w_{a,i}^T \tilde{A} \equiv [\tilde{A}]_{is}$, $\widehat{f}_i(\tilde{A}, b, N, n)$ is invariant to configuration if and only if $[\tilde{A}]_{is} = \frac{1}{N} 1^T$. Meanwhile, $\widehat{f}^{(q)}_i(\tilde{A}, b, N, n)$ is invariant to configuration if and only if $w_{a,i}^{(q)} = \frac{1}{N} 1$. Since $w_{a,i}^{(q)} \equiv [\tilde{A}]_{is}$, $\widehat{f}^{(q)}_i(\tilde{A}, b, N, n)$ is invariant to configuration if and only if $[\tilde{A}]_{is} = \frac{1}{N} 1$. □
and only if \([\bar{A}^q]_{is} = \frac{1}{N}1^T\). For part (2), \(\hat{f}(\bar{A}, b, N, n)\) is invariant to configuration if and only if \([\bar{A}]_{is} = \frac{1}{N}1^T\) for all \(i \in \{1, \ldots, N\}\), that is, if and only if \(\bar{A} = \frac{1}{N}11^T\). Meanwhile, \(\hat{f}^{(q)}(\bar{A}, b, N, n)\) is invariant to configuration if and only if \([\bar{A}^q]_{is} = \frac{1}{N}1^T\) for all \(i \in \{1, \ldots, N\}\), that is, if and only if \(\bar{A}^q = \frac{1}{N}11^T\). For part (3), \(\hat{f}_{avg}(\bar{A}, b, N, n)\) is invariant to configuration if and only if \(d_w = \frac{1}{N}1\). Since \(\hat{f}_{avg}(\bar{A}, b, N, n)\) is invariant to configuration if and only if \(\frac{1}{N}1\). For part (4), \(\hat{f}^{(\infty)}(\bar{A}, b, N, n)\) is invariant to configuration if and only if \(w_\infty = \frac{1}{N}1\). Since \(w_\infty = \frac{1}{N}1\) if and only if \(1^T = 1^T\bar{A}\), that is, if and only if \(\bar{A}\) is doubly stochastic. □

**Proof of Theorem 7**

Construct the set of \(N + 1\) vectors \(\{b_0, b_1, \ldots, b_n, \ldots, b_N\}\), for which \([b_n]_i = 1\) for \(i \leq n\) and \([b_n]_i = 0\) for \(n + 1 \leq i \leq N\). For each integer \(n \in [0, N]\), also define the set \(S_n\) of all \(N \times N\) permutation matrices \(S_{nx}, S_{ny} \in S_n\) for which \(S_{nx}b_n \neq b_n\) and \(S_{nx}b_n = S_{ny}b_n\).

**Lemma 1**  Unordered multiset \(\left\{\hat{f}_i(\bar{A}, b, N, n)\right\}_{i=1}^N\) (respectively \(\left\{\hat{f}^{(q)}_i(\bar{A}, b, N, n)\right\}_{i=1}^N\)) is invariant to configuration if and only if, for each \(S_{nx} \in S_n\) and for every \(n \in [0, N]\), there corresponds some permutation matrix \(R\) such that \(\bar{A}b_n = R\bar{A}S_{nx}b_n\) (respectively such that \(\bar{A}^q b_n = R\bar{A}^q S_{nx}b_n\)).

**Proof.** Unordered multiset \(\left\{\hat{f}_i(\bar{A}, b, N, n)\right\}_{i=1}^N\) is invariant to configuration if and only if, for every configuration \(b' \in \mathcal{B}(N, n)\), there exists an \(N \times N\) permutation matrix \(R\) such that \(\hat{f}(\bar{A}, b, N, n) = R\hat{f}(\bar{A}, b', N, n)\), or equivalently, \(\bar{A}b = R\bar{A}b'\). Now, generate each configuration in the set \(\mathcal{B}(N, n)\) by introducing the vector \(b_n \in \{b_1, \ldots, b_N\}\) and defining the set \(S_n\) of all \(N \times N\) permutation matrices \(S_{nx}, S_{ny} \in S_n\). For which \(S_{nx}b_n \neq b_n\) and \(S_{nx}b_n = S_{ny}b_n\). Then, for every \(b' \in \mathcal{B}(N, n)\), there exists some \(S_{nx} \in S_n\) for which \(b' = S_{nx}b_n\). Multiset \(\left\{\hat{f}_i(\bar{A}, b, N, n)\right\}_{i=1}^N\) is thus invariant to configuration if and only if, for each \(S_{nx} \in S_n\) and for every \(n \in [0, N]\), there corresponds some permutation matrix \(R\) such that \(\bar{A}b_n = R\bar{A}S_{nx}b_n\). Replace matrix \(\bar{A}\) with \(\bar{A}^q\) to obtain the result that multiset \(\left\{\hat{f}^{(q)}_i(\bar{A}, b, N, n)\right\}_{i=1}^N\) is invariant to configuration if and only if, for each \(S_{nx} \in S_n\) and every \(n \in [0, N]\), there corresponds some permutation matrix \(R\) such that \(\bar{A}^q b_n = R\bar{A}^q S_{nx}b_n\).

From Lemma 1, multiset \(\left\{\hat{f}_i(\bar{A}, b, N, n)\right\}_{i=1}^N\) is invariant to configuration if and only if there exists a permutation matrix \(R\) such that

\[
\bar{A}b_n = R\bar{A}S_{nx}b_n
\]

(5)

for each \(S_{nx} \in S_n\) and for every \(n \in [0, N]\). On the left-hand side of Equation 5, \(\bar{A}b_n\) is the row sum of the first \(n\) columns of \(\bar{A}\). On the right-hand side of Equation 5, \(S_{nx}\) permutes the columns of \(\bar{A}\).
so $\bar{A}S_{n}\cdot b_{n}$ is the row sum of a different set of $n$ column vectors of $\bar{A}$. With matrix $R$ permuting the rows of $\bar{A}S_{n}$, $\bar{A}b_{n} = R\bar{A}S_{n}\cdot b_{n}$ when the set of elements in the first $n$ columns of $\bar{A}$ equals the set of elements in a different group of $n$ columns of $\bar{A}$. Now considering all allowable permutation matrices $S_{n}$, multiset $\left\{ \hat{f}_{i}(\bar{A}, b, N, n)\right\} _{i=1}^{N}$ is invariant to configuration if and only if the row sum of any $n$ columns of $\bar{A}$ has the same set of elements for every integer $n \in [1, N]$.

It turns out that it is redundant to specify that this condition must hold for every $n \in [1, N] \subseteq \mathbb{Z}_{+}$. Since $\bar{A}$ is row-stochastic, the row sum of any $n$ column vectors of $\bar{A}$, $n \in [1, N - 1]$, has the same set of elements whenever the row sum of any $N - n$ column vectors of $\bar{A}$ has the same set of elements. Furthermore, the row sum of $N$ column vectors of $\bar{A}$ is always equal to 1. Therefore, we eliminate such redundancies in the required set of integers $n$ for which this condition must hold and state that multiset $\left\{ \hat{f}_{i}(\bar{A}, b, N, n)\right\} _{i=1}^{N}$ is invariant to configuration if and only if the row sum of any $n$ columns of $\bar{A}$ has the same set of elements for every integer $n \in [1, \lfloor N/2 \rfloor]$.

Next, since $\bar{A}$ is row-stochastic, if the row sum of any $n$ column vectors of $\bar{A}$ has the same set of elements for every integer $n \in [1, \lfloor N/2 \rfloor]$, $\bar{A}$ must be doubly stochastic. For $n = 1$, every column of $\bar{A}$ must have the same set of elements, so every column of $\bar{A}$ must have the same sum $\alpha$. The sum of all matrix elements in $\bar{A}$ is $N$, by the row-stochasticity of $\bar{A}$, so $N\alpha = N$, that is, $\alpha = 1$. □

**Proof of Theorem 8**

*First method of proof:* $EX(\bar{A}, N, n) = E[w^{T}B(N, n)]$, where $B(N, n)$ is a random vector whose elements are $B_{i} \sim \text{Bern}\left(\frac{n}{N}\right), i \in \{1, \ldots, N\}$. Therefore,

$$EX(\bar{A}, N, n) = E(w_{1}B_{1} + w_{2}B_{2} + \ldots + w_{N}B_{N}) = \sum_{i=1}^{N} w_{i}EB_{i} = \frac{n}{N}. \Box$$

*Second method of proof:*

$$EX(\bar{A}, N, n) = \frac{1}{|B(N, n)|} \sum_{b(n, n) \in B(N, n)} [w(\bar{A})]^{T}b(N, n)$$

$$= \frac{1}{|B(N, n)|} \sum_{i=1}^{N} |\{b \in B(N, n) : b_{i} = 1\}| w_{i}$$

$$= \frac{|\{b \in B(N, n) : b_{i} = 1\}|}{|B(N, n)|} \sum_{i=1}^{N} w_{i} = \frac{\binom{N}{n} \times \frac{n}{N}}{\binom{N}{n}} = \frac{n}{N}.$$

By symmetry, $|\{b \in B(N, n) : b_{i} = 1\}| = |\{b \in B(N, n) : b_{j} = 1\}|, \forall i, j \in \{1, \ldots, N\}$. □

**Proof of Theorem 9**

First demonstrating that $\text{Var} X(\bar{A}, N, n) = \frac{n}{N} \left(1 - \frac{n}{N}\right) \binom{N}{n/2} \frac{N}{N-1} \left(N \text{Var} W(\bar{A})\right)$:
Var \( X(\tilde{A}, N, n) = \text{Var}[w^T B(N, n)] \), where \( B(N, n) \) is a random vector whose elements are \( B_i \sim \text{Bern} \left( \frac{n}{N} \right) \), for all \( i \in \{1, \ldots, N\} \). Therefore,

\[
\begin{align*}
\text{Var} \ X(\tilde{A}, N, n) &= \text{Var} (w_1 B_1 + w_2 B_2 + \cdots + w_N B_N) \\
&= \sum_{i=1}^{N} \text{Var} (w_i B_i) + \sum_{i=1}^{N} \sum_{j=1 \atop j \neq i}^{N} \text{Cov} (w_i B_i, w_j B_j) \\
&= \sum_{i=1}^{N} w_i^2 \text{Var} B_i + \sum_{i=1}^{N} \sum_{j=1 \atop j \neq i}^{N} w_i w_j \text{Cov} (B_i, B_j) \\
&= (\text{Var} B_i) \sum_{i=1}^{N} w_i^2 + (E[B_i B_j] - (EB_i)(EB_j)) \sum_{i=1}^{N} \sum_{j=1 \atop j \neq i}^{N} w_i w_j \\
&= \frac{n}{N} \left(1 - \frac{n}{N}\right) \sum_{i=1}^{N} w_i^2 \left[ \left( \frac{n}{N} \right) \left( \frac{n-1}{N-1} \right) - \left( \frac{n}{N} \right)^2 \right] \sum_{i=1}^{N} \sum_{j=1 \atop j \neq i}^{N} w_i w_j \\
&= \frac{n}{N} \left(1 - \frac{n}{N}\right) \times \left( \sum_{i=1}^{N} w_i^2 - \frac{1}{N-1} \sum_{i=1}^{N} \sum_{j=1 \atop j \neq i}^{N} w_i w_j \right). \\
\end{align*}
\]

Now, \( \text{Var} W = \frac{1}{N} \sum_{i=1}^{N} \left( w_i - \frac{\sum_{j=1}^{N} w_j}{N} \right)^2 = \frac{1}{N} \left( \sum_{i=1}^{N} w_i^2 - \frac{1}{N} \left( \sum_{i=1}^{N} w_i \right)^2 \right) \\
= \frac{1}{N} \left( \sum_{i=1}^{N} w_i^2 - \frac{1}{N} \left( \sum_{i=1}^{N} w_i^2 + \sum_{i=1}^{N} \sum_{j=1 \atop j \neq i}^{N} w_i w_j \right) \right) \\
= \frac{1}{N} \frac{N-1}{N} \left( \sum_{i=1}^{N} w_i^2 - \frac{1}{N-1} \sum_{i=1}^{N} \sum_{j=1 \atop j \neq i}^{N} w_i w_j \right), \text{ so}
\]

\[
\text{Var} \ X(\tilde{A}, N, n) = \frac{n}{N} \left(1 - \frac{n}{N}\right) \frac{N}{N-1} (N \text{Var} W) .
\]

Next demonstrating that \( \text{Var} X(\tilde{A}, N, n) \rightarrow 0 \) at rate \( N^{-1} \) as \( N \rightarrow \infty \) assuming \( \text{Var} W \) < \( \infty \):

Consider the graph \( G(\tilde{A}_1) = (\mathcal{V}(\tilde{A}_1), \mathcal{E}(\tilde{A}_1)) \), \( |\mathcal{V}(\tilde{A}_1)| = N_1 \), corresponding to the row-stochastic matrix \( \tilde{A}_1 \) with vector \( w_1 \). Construct replica graphs \( G(\tilde{A}_2), \ldots, G(\tilde{A}_K) \), with \( \tilde{A}_1 = \ldots \ldots \)
Lemma 2 (Erdős and Rényi (1959), Theorem 1) with notation modified for the present work: Proof of Theorem 11 is defined so that where

\[
\Var A \overset{\text{w}}{\sim} \ Bern \left( \frac{n}{N} \right) \quad \text{for } i \in \{1, \ldots, N_1\} \quad \text{and } j \in \{1, \ldots, K\}, \quad \text{with } B_{ij}, B_{gr} \text{ independent for all } j \neq r,
\]

\[
\Var X (\bar{A}, N_1K, nK) = \Var \left( \frac{1}{K} (w_{11}B_{11} + \cdots + w_{N_11}B_{N_11} + w_{12}B_{12} + \cdots + w_{N_12}B_{N_12} + \cdots + w_{N_1K}B_{N_1K}) \right)
= \frac{1}{K^2} N \Var (w_{11}B_1 + \cdots + w_{N_1N}B_N) = \frac{N_1}{N} \Var X (\bar{A}_1, N_1, n),
\]

so \( \Var X (\bar{A}, N_1K, nK) \to 0 \) at rate \( N^{-1} \) as \( N \to \infty \). □

Proof of Theorem 10

\[
x (\bar{A}, b, N, n) = [w(\bar{A})]^T b(N,n) = \sum_{i \in \{1, \ldots, N\} \text{ s.t. } |b_i|=1} [w]_i. \quad \text{Therefore,}
\]

\[
\min \sup X (\bar{A}, N, n) = [w(\bar{A})]^T b_*(N,n) = \sum_{i=1}^n w_s,
\]

where \( w_s \) is the \( s^{th} \) smallest element in \( w(\bar{A}) \) in the ordered multiset \( \{w_s\}_{s=1}^N \) and \( b_*(N,n) \) is defined so that \( [w(\bar{A})]^T b_*(N,n) \leq [w(\bar{A})]^T b(N,n) \) for all \( b(N,n) \in B(N,n) \). Meanwhile,

\[
\max \sup X (\bar{A}, N, n) = [w(\bar{A})]^T b^*(N,n) = \sum_{i=N-n+1}^N w_s,
\]

where \( w_s \) is the \( s^{th} \) smallest element of \( w(\bar{A}) \) listed in the ordered multiset \( \{w_s\}_{s=1}^N \) and \( b^*(N,n) \) is defined so that \( [w(\bar{A})]^T b^*(N,n) \geq [w(\bar{A})]^T b(N,n) \) for all \( b(N,n) \in B(N,n) \). □

Proof of Theorem 11

The proof of Theorem 11 makes use of the following result from Erdős and Rényi (1959), with notation modified for the present work:

Lemma 2 (Erdős and Rényi (1959), Theorem 1) Consider the infinite triangular matrix of
real elements

\[
\begin{pmatrix}
  w'_{11} \\
  w'_{21} & w'_{22} \\
  \vdots & \vdots & \ddots \\
  w'_{N1} & w'_{N2} & \cdots & w'_{NN}
\end{pmatrix}
\]

with \( w'_N \) denoting the \( N^{th} \) row of the matrix and \( \sum_{j=1}^{N} w'_{Nj} = 0 \). For any real value \( t \), determine \( T( w'_N, N, n, t ) \), that is, the total number of sums

\[
y( w'_N, N, n ) = w'_{N1} + w'_{N2} + \cdots + w'_{Nn}, \quad 1 \leq i_1 < i_2 < \cdots < i_n \leq N,
\]

whose value does not exceed \( t \sigma ( w'_N, N, n ) \equiv t \sqrt{\frac{n}{N} (1 - \frac{n}{N}) \sum_{j=1}^{N} w'^2_{Nj} / \sigma ( w'_N, N, n )} \). Let CDF \( G_{y( w'_N, N, n )}(t) = \frac{T( w'_N, N, n, t )}{\binom{N}{n}} \). With

\[
\kappa( w'_N, N, n, \epsilon ) \equiv \frac{1}{\sum_{j=1}^{N} w'^2_{Nj}} \sum_{j \in \{1, \ldots, N\} \text{ s.t. } |w'_{Nj}| > \epsilon \sigma ( w'_N, N, n )} w'^2_{Nj}
\]

if \( \lim_{N \to \infty} \kappa( w'_N, N, n, \epsilon ) = 0 \) for any \( \epsilon > 0 \), then \( \lim_{N \to \infty} G_{y}(t) = \Phi( t ) \) for any real \( t \), where \( \Phi( \cdot ) \) denotes the standard normal CDF.

For a given population size \( N \), set \( w'_N = w_N ( \bar{A} ) - \frac{1}{N} \), where \( w_N ( \bar{A} ) \) is the general vector of weights discussed in the text, and subscript \( N \) is added to make the population size explicit. Then \( \sum_{j=1}^{N} w'_{Nj} = \sum_{j=1}^{N} ( w_{Nj} - \frac{1}{N} ) = 0 \). Scalar quantity

\[
x( \bar{A}, b, N, n ) - \frac{n}{N} = [ w_N ( \bar{A} ) ]^T b( N, n ) - \frac{n}{N}
\]

\[
= \left( w_{N1} - \frac{1}{N} \right) + \left( w_{N2} - \frac{1}{N} \right) + \cdots + \left( w_{Nn} - \frac{1}{N} \right)
\]

\[
= w'_{N1} + w'_{N2} + \cdots + w'_{Nn},
\]

where \( 1 \leq i_1 < i_2 < \cdots < i_n \leq N \), given a configuration \( b( N, n ) \in B( N, n ) \). Thus, by Lemma 2,

\[
G_{y(w'_N, N, n)}(t) = \frac{T( w'_N, N, n, t )}{\binom{N}{n}} = \frac{1}{\binom{N}{n}} \sum_{1 \leq i_1 < i_2 < \cdots < i_n \leq N} [ \sum_{1 \leq i_1 < i_2 < \cdots < i_n \leq N} 1_{w'_{N1} + w'_{N2} + \cdots + w'_{Nn} \leq t \sigma ( w'_N, N, n )} ]
\]

\[
= \frac{1}{\binom{N}{n}} \sum_{b(N, n) \in B(N, n)} 1_{x( \bar{A}, b, N, n ) - \frac{n}{N} \leq t \sigma ( w'_N, N, n )} = G_{X( \bar{A}, N, n ) - \frac{n}{N}}(t),
\]
Proof of Theorem 12

The proof of Theorem 12 makes use of the following result from Höglund (1978), with notation modified for the present work:

Lemma 3 (Höglund (1978), Main Theorem) Let $w_1, \ldots, w_N$ be a sequence of real numbers. Let $0 < n < N$ and let $G_Y(t) = \frac{T(w, N, n, t)}{\binom{N}{n}}$, where $T(w, N, n, t)$ is the total number of sums

$$y = w_{i_1} + w_{i_2} + \cdots + w_{i_n}, \quad 1 \leq i_1 < i_2 < \cdots < i_n \leq N,$$

whose value does not exceed $t$. Then, for all real $t$,

$$\left| G_Y(t) - \Phi\left( \frac{t - n\bar{w}}{\left( \frac{n}{N} \left( 1 - \frac{n}{N} \right) \sum_{i=1}^{N} [w_i - \bar{w}]^2 \right)^{1/2}} \right) \right| \leq \frac{C}{\sqrt{\frac{n}{N} \left( 1 - \frac{n}{N} \right) \left( \sum_{i=1}^{N} [w_i - \bar{w}]^2 \right)^{3/2}}} \sum_{i=1}^{N} |w_i - \bar{w}|^3,$$

where $\bar{w} = \frac{1}{N} \sum_{i=1}^{N} w_i$.

For the vector of weights, $w(\bar{A})$, discussed in the text, $\bar{w} = \frac{1}{N} \sum_{i=1}^{N} [w(\bar{A})]_i = \frac{1}{N}$, and

$$G_Y(t) = \frac{T(w, N, n, t)}{\binom{N}{n}} = \frac{1}{|B(N, n)|} \sum_{b(N, n) \in B(N, n)} 1_{x(\bar{A}, b(N, n), t) \leq t} = G_X(\bar{A}, N, n)(t).$$

By Lemma 3, Theorem 12 thus follows.

Proof of Theorem 13

The proof of Theorem 13 makes use of the following result from Robinson (1978), with notation modified for the present work:

Lemma 4 (Robinson (1978), Main Theorem) Let $\{a_{Ni}\}$ be a triangular array of real numbers for $i = 1, \ldots, N, N = 2, 3, \ldots$ and suppose $\sum_{i=1}^{N} a_{Ni} = 0$, $\sum_{i=1}^{N} a_{Ni}^2 = 1$. Let $K_{Ni} = \frac{1}{N} a_{Ni}$, where $(R_{N1}, \ldots, R_{NN})$ is a uniform random permutation of $(1, \ldots, N)$. Let $L_{Nn} = \frac{K_{Ni}}{(\text{Var} K_{Ni})^{1/2}}$ and $G_{Nn}(t) = \text{Pr}(L_{Nn} < t)$. Set $p = \frac{n}{N}$ and $q = 1 - \frac{n}{N}$. If condition (c) holds, then

$$|G_{Nn}(t) - J_{Nn}(t)| < C_4 \times \sum_{i=1}^{N} |a_{Ni}|^5.$$
for all \( t \), where \( C_4 \) is a function of \( p \) only,

\[
J_{Nn}(t) = \Phi (t) - H_2(t) \phi (t) \frac{q-p}{6(pq)^{1/2}} \sum_{i=1}^{N} a_{Ni}^3 \\
- H_3(t) \phi (t) \left[ \frac{1-6pq}{24pq} \left( \sum_{i=1}^{N} a_{Ni}^4 - 3N^{-1} \right) - \frac{1}{4} N^{-1} \right] - H_5(t) \phi (t) \frac{(q-p)^2}{72pq} \left( \sum_{i=1}^{N} a_{Ni}^3 \right)^2,
\]

\( \phi (t) = \Phi' (t) = (2\pi)^{-1/2} e^{-\frac{1}{2} t^2} \) and \( H_i(t) \phi (t) = (-1)^i \left( \frac{d^i}{dt^i} \right) \phi (t) \). Condition (c) is as follows:

**Condition (c)** Given \( C' > 0 \), there exist \( \epsilon > 0 \), \( C > 0 \), and \( \delta > 0 \) not depending on \( N \) such that, for any fixed \( t \), the number of indices \( j \), for which \( |a_{Nj}\hat{x} - t - 2r\pi| > \epsilon \), for all \( \hat{x} \in \left( C' \left[ \max_i |a_{Ni}| \right]^{-1}, C \left[ \sum_{i=1}^{N} |a_{Ni}|^5 \right]^{-1} \right) \) and all \( r = 0, \pm 1, \pm 2, \ldots \), is greater than \( \delta N \), for all \( N \).

Fix \( N \) and substitute \( a_{Ni} \) with the standardized weight \( \hat{w}_i = \frac{[w_i] - EW}{\sqrt{N} \text{Var } W} \), where \( EW = \frac{1}{N} \sum_{i=1}^{N} [w_i] \) and \( \text{Var } W = \frac{1}{N} \sum_{i=1}^{N} ([w_i] - EW)^2 \). Verify that

\[
\sum_{i=1}^{N} \hat{w}_i = \sum_{i=1}^{N} \frac{w_i - EW}{\sqrt{N} \text{Var } W} = 0 \quad \text{and} \quad \sum_{i=1}^{N} \hat{w}_i^2 = \frac{1}{N} \sum_{i=1}^{N} \frac{(w_i - EW)^2}{\text{Var } W} = 1.
\]

With \( B_i \sim \text{Bern } \left( \frac{w_i}{N} \right) \) and \( \text{Var } \hat{W} = \frac{1}{N} \sum_{i=1}^{N} \left( \hat{w}_i - EW \right)^2 = \frac{1}{N} \sum_{i=1}^{N} \hat{w}_i^2 = \frac{1}{N} \),

\[
L_n = \frac{K_n}{(\text{Var } K_n)^{1/2}} = \frac{\sum_{i=1}^{n} \hat{w}_R_i}{(\text{Var } \sum_{i=1}^{n} \hat{w}_R_i)^{1/2}} = \frac{\sum_{i=1}^{n} \frac{(w_{R_i} - EW)}{\text{Var } W^{1/2}}}{(\text{Var } \sum_{i=1}^{n} \hat{w}_R_i)^{1/2}} = \frac{\sum_{i=1}^{n} w_{R_i} - nEW}{(\text{Var } W)^{1/2}} \left( \frac{n}{N} \right)^{1/2} \frac{N}{n} \text{Var } \hat{W}^{1/2} = \frac{X (\bar{\mathbf{A}}, N, n) - EX (\bar{\mathbf{A}}, N, n)}{(\text{Var } X (\bar{\mathbf{A}}, N, n))^{1/2}}, \quad \text{and Theorem 13 follows.} \]

**Proof of Theorem 14**

When \( k_\omega \) agents have the same non-zero weight \( \omega \), and all other \( N - k_\omega \) agents have zero weight, \( g_X (\mathbf{A}, N, n) (t) \) is non-zero only for integer multiples of \( \omega : i\omega \). We now determine the allowable values of \( i \). Given \( N, n, k_\omega \), the smallest possible value of \( i \) is \( \max \{0, n - (N - k_\omega)\} \). If \( N - k_\omega \geq n \), that is, if the number of agents with zero weight is greater than or equal to \( n \), then there exists at least one configuration \( \mathbf{b} (N, n) \) for which \( x (\bar{\mathbf{A}}, \mathbf{b}, N, n) = 0 \) and \( g_X (\mathbf{A}, N, n) (0) > 0 \). If \( N - k_\omega < n \), so that the number of agents with zero weight is less than \( n \), then \( \min \text{ supp } X (\mathbf{A}, N, n) = [n - (N - k_\omega)] \times \omega \). Given \( N, n, k_\omega \), the largest possible value of \( i \) is \( \min \{n, k_\omega\} \). If \( n \leq k_\omega \), then there exists at least one configuration \( \mathbf{b} (N, n) \) for which \( x (\bar{\mathbf{A}}, \mathbf{b}, N, n) = n\omega \) and \( g_X (\mathbf{A}, N, n) (n\omega) > 0 \). If \( n > k_\omega \), then
max \supp X (\bar{A},N,n) = k_\omega \omega. Therefore, we define the set
\[ \mathcal{I} = \{ \max \{0,n-(N-k_\omega)\}, \max \{0,n-(N-k_\omega)\} + 1, \ldots, \min \{n,k_\omega\} \}. \]

For all values \( i \in \mathcal{I} \),
\[ g_{X(A,N,n)}(i\omega) = \binom{k_\omega}{i} \binom{N-k_\omega}{n-i} \binom{N}{n}, \]
which is hypergeometric. For all values \( i \notin \mathcal{I} \), \( g_{X(A,N,n)}(i\omega) = 0 \). \( \square \)

**Proof of Theorem 15**

For all \( i, \ell \in \{1, \ldots, N\} \), \( [\bar{A}]_{i\ell} \in \{0,1/\xi\} \). Then \( d^-_w (\bar{A}) = \frac{1}{N} \bar{A}^T 1 = \frac{1}{N} d^- (A) \) and \( D^- (\bar{A}) = \frac{1}{N^2} D^- (A) \). With Var \( D^- (\bar{A}) = \frac{N}{N^2} \sum_{i=1}^{N} ([d^-_w (\bar{A})]_i - \frac{1}{N})^2 = \frac{N}{N^2} \sum_{i=1}^{N} ([d^- (A)]_i - \frac{1}{N})^2 = \frac{1}{N} \text{Var} D^- (A) \), it follows from Theorem 9 that Var \( \bar{F}_{\text{avg}} (\bar{A},N,n) = \frac{N}{N} (1 - \frac{1}{N}) \frac{N}{N-1} N (\frac{1}{N})^2 \text{Var} D^- (A) \).

Construct the ordered multiset \( \{w_s\}_{s=1}^{N} \) from the elements of \( d^-_w (\bar{A}) \) so that \( w_s < w_{s'} \) whenever \( s \leq s' \). From Theorem 10, \( \min \supp \bar{F}_{\text{avg}} (\bar{A},N,n) = \sum_{s=1}^{n} w_s = \frac{1}{N^2} \sum_{s=1}^{n} v_s \) and \( \max \supp \bar{F}_{\text{avg}} (\bar{A},N,n) = \sum_{s=N-n+1}^{N} w_s = \frac{1}{N^2} \sum_{s=N-n+1}^{n} v_s \). \( \square \)

**Proof of Theorem 16**

\[ X (\bar{A},N,n) = \frac{\bar{W}_1 (\bar{A}) + \cdots + \bar{W}_n (\bar{A})}{n} \times f, \] with \( f = \frac{n}{N} \) fixed as \( N \to \infty \), so \( \bar{W}_1 (\bar{A}) + \cdots + \bar{W}_n (\bar{A}) = \frac{1}{f} X (\bar{A},N,n) \). Suppose that \( \text{Var} \bar{W}_i (\bar{A}) = N^2 \text{Var} W (\bar{A}) \) is finite. Then by the Classical Central Limit Theorem,
\[ n^{1/2} \left( \frac{1}{f} X (\bar{A},N,n) - EW_i (\bar{A}) \right) \xrightarrow{d} \mathcal{N} \left( 0, \text{Var} \bar{W}_i (\bar{A}) \right), \]
where \( EW_i (\bar{A}) = 1 \). Now suppose that \( \text{Var} \bar{W}_i (\bar{A}) \) is infinite and \( EW_i (\bar{A}) = 1 \). Specifically, \( \Pr [\bar{W}_i (\bar{A}) > t] \sim L (t) t^{-\xi} \), where \( L (t) \) is a slowly varying function and \( \xi \in (1, 2) \) since \( \text{Var} \bar{W}_i (\bar{A}) \) is infinite and \( EW_i (\bar{A}) \) is finite. As discussed in Nolan (2014), a specific case of the Generalized Central Limit Theorem is as follows, with notation adapted for the present setting:

**Lemma 5** Let \( \bar{W}_1 (\bar{A}), \bar{W}_2 (\bar{A}), \ldots \) be independent, identically distributed random variables. Suppose that \( \Pr [\bar{W}_i (\bar{A}) > t] \sim C^+ t^{-\xi} \) and \( \Pr [\bar{W}_i (\bar{A}) < -t] \sim C^- |t|^{-\xi} \) as \( t \to \infty \) with \( 1 < \xi < 2 \) and \( C^+ + C^- > 0 \). Set \( \gamma = \frac{C^+ - C^-}{C^+ + C^-} \). Then
\[ \frac{\bar{W}_1 (\bar{A}) + \cdots + \bar{W}_n (\bar{A}) - nEW_i (\bar{A})}{n^{1/\xi}} \xrightarrow{d} \mathcal{S} (\xi, \beta, \gamma, 0; 1), \]
where $\bar{S}(\xi, \beta, \bar{\gamma}, 0; 1)$ is a stable distribution with characteristic function

$$E \exp \left( iu \bar{W}_i(\bar{A}) \right) = \exp \left( -\bar{\gamma}^\xi |u|^\xi \left[ 1 - i\beta \left( \tan \frac{\pi \xi}{2} \right) \times \text{sign} \, u \right] \right)$$

when $\xi \neq 1$.

With

$$\frac{\bar{W}_1(\bar{A}) + \cdots + \bar{W}_n(\bar{A}) - nE\bar{W}_1(\bar{A})}{n^{1/\xi}} = n^{1-1/\xi} \left( \frac{\bar{W}_1(\bar{A}) + \cdots + \bar{W}_n(\bar{A})}{n} - E\bar{W}_1(\bar{A}) \right),$$

we obtain the result

$$n^{1-1/\xi} \left( \frac{1}{f} X(\bar{A}, N, n) - E\bar{W}_1(\bar{A}) \right) \overset{d}{\longrightarrow} \bar{S}(\xi, \beta, \bar{\gamma}, 0; 1).$$

**Proof of Theorem 17**

$$EX\left(\bar{A}, N, n, (\gamma_i)_{i=1}^N\right) = E \left[ w_1(\bar{A}) \cdot B_1 + w_2(\bar{A}) \cdot B_2 + \cdots + w_N(\bar{A}) \cdot B_N \right] = [w(\bar{A})]^T \cdot \mathbf{\mu},$$

and

$$\text{Var} \left[ w_1(\bar{A}) \cdot B_1 + w_2(\bar{A}) \cdot B_2 + \cdots + w_N(\bar{A}) \cdot B_N \right] = [w(\bar{A})]^T \cdot \mathbf{\Sigma} \cdot [w(\bar{A})],$$

with $\mathbf{\mu} = EB$ as the conditional mean vector for $B$ and $\mathbf{\Sigma}$ as the $N \times N$ conditional covariance matrix for $B$. The $N \times 1$ random vector $B \equiv B(N, n, s, \psi)$ is distributed according to Fisher’s multivariate non-central hypergeometric distribution (see McCullagh and Nelder (1989)). Each agent exists in the population with frequency 1, so define the $N \times 1$ frequency vector $s = 1$. Next, let $\phi$ be an $N \times 1$ vector of probabilities whose elements are defined as $\phi_i = \Pr[B_i = 1 | \gamma_i]$; construct the $N \times 1$ vector $\psi$ with element $\psi_i = \frac{\phi_i}{1 - \phi_i} / \frac{\phi_k}{1 - \phi_k}$ relative to some agent $k$ and $\psi_k \equiv 1$. $\mathbf{\mu}$ and $\mathbf{\Sigma}$ can be approximated by solving the following system of equations:

$$\sum_{i=1}^N \mu_i = n,$$

$$\psi_j = \frac{\mu_j (s_k - \mu_k) - \Sigma_{jk}}{(s_j - \mu_j) \mu_k - \Sigma_{jk}}, \quad \forall j \in \{1, \ldots, N\} \setminus \{k\},$$

$$\mathbf{\Sigma} = \frac{N}{N - 1} \left( \text{diag} \, \zeta - \frac{\zeta \zeta^T}{1^T \zeta} \right),$$

with $\frac{1}{\zeta_j} = \frac{1}{\mu_j} + \frac{1}{s_j - \mu_j}$.

Now, the number of equations in the system can be reduced by noting that $\mu_i = \mu_j$ when $\phi_i = \phi_j$, so $\zeta_i = \zeta_j$, $\psi_i = \psi_j$, and $\Sigma_{ik} = \Sigma_{jk}$. Therefore, partition agent indices into $\Theta$ categories according to their conditional probabilities, that is, agents $i, j$ are in category $\theta$ if $\phi_i = \phi_j = \rho_\theta$. Define the odds ratio for agents in category $\theta$ relative to category $k$ as: $\psi_\theta = \frac{\rho_\theta}{1 - \rho_\theta} / \frac{\rho_k}{1 - \rho_k}$, with $\hat{\psi}_k \equiv 1$. Define the $\Theta \times 1$ vector $\hat{\mathbf{\mu}}$ across the $\Theta$ categories, setting $\mu_i = \hat{\mu}_\theta$ for each agent $i$ from category $\theta$, and setting $\zeta_i = \hat{\mu}_\theta (1 - \hat{\mu}_\theta)$ for each agent $i$ from category $\theta$. Define the $\Theta \times \Theta$ matrix $\hat{\mathbf{\Sigma}}$ with element $\hat{\Sigma}_{\theta k}$ equal to the conditional covariance $\text{Cov}(B_i, B_j)$ between agent $i$ in category $\theta$.
and agent \( j \) in category \( k \). The system of equations then collapses to the following:

\[
\sum_{\theta=1}^{\Theta} \sum_{i \in \{1, \ldots, N\} \text{ s.t. } \phi_i = \rho_\theta} \hat{\mu}_\theta = n,
\]

\[
\hat{\psi}_\theta = \frac{\hat{\mu}_\theta (1 - \hat{\mu}_k) - \hat{\Sigma}_{\theta k}}{(1 - \hat{\mu}_\theta) \hat{\mu}_k - \hat{\Sigma}_{\theta k}}, \quad \forall \theta \in \{1, \ldots, \Theta\} \setminus \{k\}, \text{ and}
\]

\[
\Sigma = \frac{N}{N - 1} \left( \text{diag } \zeta - \frac{\zeta \zeta^T}{1^T \zeta} \right).
\]