

## INFORMATION NETWORKS AND MARKET BEHAVIOR\*

WAYNE E. BAKER and ANANTH. V. IYER

*Graduate School of Business  
University of Chicago  
1101 East 58th Street  
Chicago, IL 60637*

*July 1991*

This paper proposes a mathematical model of financial markets as networks. The model examines the effect of *network structure* on market behavior (price volatility and trading volume). In the model, investors are arrayed in various network configurations through which they gather information to make trading decisions. The basic network considered is a chain graph with two parameters, number of investors ( $n$ ) and the length of time in which information is transmitted ( $k$ ). Closed-form expressions for price volatility and expected trading volume are provided. The model is generalized to more complex networks, focusing on the hub-and-spoke network. The network configurations analyzed do not represent the real (and unknown) communication network among investors, but predictions from the model are consistent with price and volume patterns observed in sociological and economic research on financial markets. The main result is that network structure alone influences price volatility and expected trading volume even though investors are homogeneous and the information introduced into the system is unbiased and random. This result suggests that the structure of the real communication network among investors may influence market behavior.

KEY WORDS: Networks, market, diffusion, graph theory, information, economics.

### 1. INTRODUCTION

The principal tenet of structural sociology is that social behavior and institutions can be explained by analysis of the relations among social entities (natural persons and corporate actors). This tenet is well documented in many areas of social life. The structural perspective has not been applied as often to analyze economic life, but a growing number of studies has clearly demonstrated the promise and utility of doing so (e.g., Baker 1984a, 1990; Berkowitz 1988; Burt 1983, 1988; White 1981, 1988). The purpose of this paper is to advance the structural analysis of economic

---

\*We gratefully acknowledge the helpful comments and suggestions provided by Gary Becker, James Coleman, Edward Laumann, George Loewenstein, John Padgett, Mark Shanley, Arnold Zellner and the editor of this journal. Earlier versions of this paper were presented at the annual meetings of the International Sunbelt Social Network Conference, Tampa, FL (1989), the Academy of Management, Washington, DC (1989), the ORSA/TIMS Joint National Conference, New York City (1989), and James Coleman's seminar on Mathematical Sociology (1990).

institutions by proposing a mathematical model of financial markets as networks. In the model, investors are arrayed in various network configurations through which they gather information to make trading decisions. *The main result is that network structure influences price volatility and expected trading volume even though investors are homogeneous and the information introduced into the system is unbiased and random.* This result suggests that the structure of the real communication network among investors may influence market behavior.

The model is motivated by sociological studies of financial markets that conceptualize and analyze markets as social structures and show that networks influence price volatility (Baker 1984a, 1984b). The premise that markets are social structures (networks) may be palatable to sociologists, but it draws into question a foundational assumption of economic theory: *individualism*. As Coleman (1990), Granovetter (1985), White (1988), Zukin and DiMaggio (1990), and other sociologists criticize, most economic theorists assume an *atomized* market: economic actors are independent, without important or consequential connections among them.<sup>1</sup> To most economists, markets are “interpersonal vacuums” (Frenzen and Davis 1990:1). In contrast, economic and structural sociologists contend that real markets are embedded and enmeshed in social networks of all kinds. As White (1988:232) put it, “The truth is that market activity is intensely social—as social as kinship networks or feudal armies . . . .”

Social dynamics and networks are intrinsic to financial markets (Adler and Adler 1984), even though these markets are considered by many to be exemplars of the atomized market ideal.<sup>2</sup> Baker (1984a, 1984b) found, for example, that network patterns among traders on the floor of a national securities exchange influenced both the direction and magnitude of price volatility. In a survey of investors, economists Shiller and Pound (1989) found that investors were influenced by information received from interpersonal contacts: most institutional and individual investors reported that initial interest leading to the purchase of stock was stimulated by information received by word of mouth.<sup>3</sup> These findings about networks in financial markets are consistent with studies of interpersonal communication, networks, and influence in labor markets (Boorman 1975; Granovetter 1974; Saloner 1985) and consumer goods markets (Frenzen and Davis 1990; Robertson 1971), as well as more generally with studies of diffusion of innovations, rumors, and information

<sup>1</sup>For example, as Hirshman (1982, p.1473) describes, the (idealized) market “... function[s] without any prolonged human or social contact between parties ... [T]he various operators that contract together need not enter into recurrent or continuing relationships as a result of which they would get to know each other well.”

<sup>2</sup>Financial markets are usually assumed to be so “efficient” that information spreads almost instantaneously and prices instantly reflect all information. This view implicitly assumes that communication networks do not exist (e.g., all information is obtained from price histories or public announcements) or that communication networks are so dense and integrated that they do not shape the flow of information and influence prices. Another branch of financial economics analyzes the impact of “incomplete information” on the structure and functioning of markets and firms (Constantinides and Bhattacharya 1988), but such models ignore diffusion mechanisms, especially communication networks, and adhere to the traditional assumption of atomized (nonrelational) markets.

<sup>3</sup>An informal example is The Wall Street Journal (Oct. 27, 1988) report of the “informal investing grapevine” among investors and firms seeking capital in which investors are “part of an informal network of noninstitutional investors who hear of promising start-ups through acquaintances or friends of friends ... ”.

(e.g., Bartholemew 1982; Burt 1987; Coleman, Katz, and Menzel 1966; Hedetniemi, Hedetniemi, and Liesman 1988; Rogers 1979, 1983; Rogers and Kincaid 1981).

Our rationale for developing a formal mathematical model of networks and markets is fourfold:

**1. A mathematical model permits examination of the internal logical consistency of a proposition.**

Because empirical evidence about the proposition that networks influence market behavior is limited to two studies (Baker 1984a; Shiller and Pound 1989), further investigation is required to test the validity of the proposition. A formal model is like an experiment in which the causal variable (e.g., network structure) is manipulated to assess its (hypothesized) effect on the dependent variable (e.g., price and trading volume). If a valid formal model can be constructed in which variation in network structure "causes" variation in market behavior, then the logical consistency of the proposition based on empirical observation is given greater support.<sup>4</sup>

**2. A mathematical model can be validated by comparing its predictions with known patterns.**

Because our model is "stylized"—the network configurations analyzed (see Figure 1) do not represent the real (and unknown) investor communication network—it is critical to compare its predictions with known price and volume patterns. Our model reproduces several known financial market behaviors: (1) Our model can generate the "paradox of large numbers" discovered by Baker (1984a). Baker noted a curvilinear relationship between "market performance" (measured as price volatility) and "size" (number of traders). Market performance improves as the number of traders increases (as economists would predict) but eventually growth in the number of traders impairs market performance. (2) Price volatility and trading volume are generally positively correlated in our model, a pattern observed in real markets (e.g., Karpoff 1987). (3) The "Monday effect," where trading volume on Monday is lower than the weekday average (Mulherin and Gerety 1988), can be reproduced in the model by increasing the length of time in which information diffuses (which represents the longer duration between Friday and Monday). (4) The model can also produce a pattern where price volatility and volume move in opposite directions, which is similar to the so-called "anomalous" empirical findings that price volatility is higher (e.g., Fama 1965) and trading volume is lower (Mulherin and Gerety 1988) over the weekend.

**3. A mathematical model extends the conclusions of empirical research by considering the phenomena over a wide range of parameter values.**

Empirical study is constrained by time, budget, and access to a (relatively) small set of observations. In the laboratory of a formal model, a much wider range of phenomena can be studied. In our model, for example, we can examine the effect of network size (number of investors) on market behavior over any range of sizes. In this way, we found that increasing size does not always improve market performance (consistent with Baker's conjectures). But we also found that the

<sup>4</sup>Furthermore, "by analyzing the internal logical structure of [causal] theories, we can formulate more precisely how one derives empirical consequences from such theories" (Stinchcombe 1987:28).

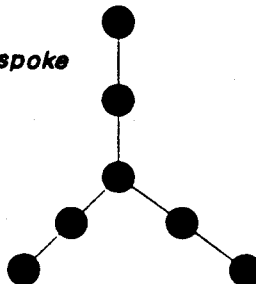
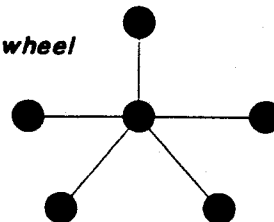
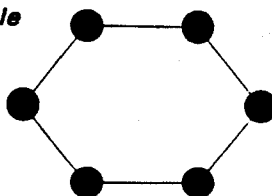
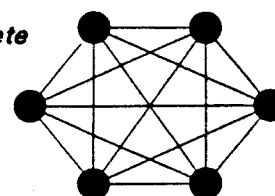
*a. chain**b. hub-and-spoke**c. wheel**d. circle**e. complete*

FIGURE 1. Five classic communication graphs.

effect of size is relative to the length of time in which information is transmitted through the network.

#### 4. A mathematical model can generate new hypotheses.

Baker's (1984a, 1984b) study of financial markets revealed the paradoxical (non-monotonic) effect of size on market behavior. Variation in size has such an effect on market behavior in our model. However, variation in the length of time in which information is transmitted also produces nonmonotonic effects on market behavior. Because prior research has not addressed the effect of transmission time on market behavior, the predictions of our formal model provide new directions for empirical study.

The paper is organized as follows. The basic model is presented in Section 2. It is based on the simple chain graph with two parameters, the number of investors ( $n$ ) and the length of time in which information is transmitted from investor to investor ( $k$ ). Closed-form expressions are presented for price volatility and expected trading volume in chains with different  $n$  and  $k$ .<sup>5</sup>

The basic model is generalized to more complex graphs in Section 3. After discussing circle and complete graphs, we focus on the hub-and-spoke network. Though there are differences, the chain and hub-and-spoke networks produce generally similar market behaviors. The results are discussed in Section 4. Our conclusion is Section 5. Appendix A presents three price adjustment functions used in the model. Appendix B contains proofs.

<sup>5</sup>These equations are provided so that others may analyze the effects of different  $n$  and  $k$  on volatility and volume simply by using a calculator, or to test a simulator for general networks when applied to chain graphs.

## 2. THE BASIC MODEL

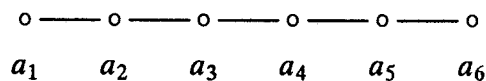
In brief, the model operates as follows. (1) Investors are assumed to be arranged in explicit network configurations (e.g., Figure 1). (2) In the “transmission period,” unambiguous buy and sell signals enter the network at random points from unspecified exogenous sources. The signals are independent and unbiased. (3) The signals are transmitted from investor to investor according to the configuration of the network. Transmission ceases when the transmission period expires. (4) Based on the content of their information sets, investors make trading decisions (buy, sell, or hold). (5) In the “trading period,” investors submit offers to buy or sell to an “auctioneer” (defined in section 2.2) who sets the price and volume on the basis of all offers received.

### 2.1. The Information Network as a Chain Graph

A *network* is a type of graph. The mathematical theory of graphs (e.g., Harary, Norman, and Cartwright 1965) is based on two primitive concepts: points and lines. A point (or node) is usually conceptualized as representing a single actor, such as an individual investor in an information network. A point can also represent an organization, like an institutional investor, or a group of *structurally equivalent* actors (Lorrain and White 1971) who are homogeneous in all respects. A line represents a relationship between two actors, such as a path for interpersonal communication. A graph consists of a finite collection of points  $a_1, a_2, \dots, a_n$  and the set of all lines connecting these points. Points and lines may be arranged in various patterns, yielding different networks or graphs.

The basic model is based on the chain graph. Chain graphs differ in two respects: number of investors ( $n$ ) and the length of the transmission period ( $k$ ). The length of the transmission period is represented by the number of lines through which a given signal may travel. For example, assume a signal is introduced at investor  $a_i$  and  $k = 2$ . All investors who are no more than two lines from  $a_i$  will also get the signal. All points (investors) no more than  $k$  lines from  $a_i$  are members of the same “neighborhood.” Define the “neighborhood” of investor  $i$  in a chain of  $n$  investors as the minimum of  $k$  and  $(i - 1)$  investors to its left and the minimum of  $k$  and  $(n - i)$  investors to its right.

For illustration, consider a short chain<sup>6</sup> with  $n = 6$  and  $k = 2$ :



If a signal is introduced at  $a_1$ , then all points (investors) no more than  $k = 2$  lines to the right ( $a_2, a_3$ ) are members of the neighborhood and also receive the signal. Because  $a_1$  is the left end node, its neighbors can only include investors to the right. If the signal is introduced at  $a_2$ , then all investors no more than  $k$  lines away will get the signal, including one investor to the left ( $a_1$ ) and two investors to the right ( $a_3, a_4$ ). (The complete census of all neighborhoods in this six-investor example is presented in Section 2.2.)

<sup>6</sup>Note that our model can accommodate chains with any number of investors ( $n$ ) and any length of the transmission period ( $k$ ).

We assume that two signals, a signal to buy and a signal to sell, arrive in the transmission period from unspecified exogenous sources.<sup>7</sup> The signals are independent and good.<sup>8</sup> Each signal enters the chain at a randomly selected point (i.e., investor) and diffuses until the transmission period ( $k$ ) expires. Hence, location does not affect the likelihood that an investor gets an exogenous signal; location affects the likelihood that the investor gets the signal as it is passed from investor to investor.

At the end of the transmission period, each investor has complete or incomplete information. Complete information means that the investor possesses all signals that entered anywhere in the system. Incomplete information means that the investor received some or none of the signals. With two signals, there are four possible information sets for an investor: (buy), (sell), (buy, sell), and ( ). Given our assumptions, each investor interprets the four information sets and acts as follows: An investor receiving a buy signal offers to buy *one* unit, an investor receiving a sell signal offers to sell *one* unit, an investor receiving both signals does not make any offers (the signals cancel),<sup>9</sup> and an investor who does not receive any signal does not make any offers. (All offers are made "at the market" by which we mean the price set in the trading period by the auctioneer.) Given our assumptions, two investors with identical information sets interpret their information the same way, make the same trading decision, and take the same trading action.

The following are features of investors in our model:

1. Investors are homogeneous in all respects (e.g., goals, endowments, preferences) except in the content of their information sets, that is, whether they receive all, some, or no signals. Investors receive information in two ways: (1) as signals that enter the network at random points from unspecified exogenous sources, and (2) from their personal contacts, that is, the signals that are "passed on" from investor to investor throughout the network.
2. Investors do not know and cannot infer the structure of the information network. (1) No investor knows the overall configuration of the information network (chain, hub-and-spoke, etc.). (2) No investor knows his specific location in the network (e.g., that he is an end node in a chain). (3) No investor knows how many exogenous signals enter the system (e.g., no one knows that one "buy" and one "sell" always enter the network in the transmission period). These assumptions are made for tractability, but they are consistent with the well known fact that most people neither know nor comprehend the overall networks in which they are engaged.
3. Investors are not strategic in their role as information transmitters. That is, investors freely pass on information; they do not hoard, distort, or otherwise influ-

<sup>7</sup>Diffusion models typically analyze the spread of *one* piece of information, such as a rumor. The few exceptions are discussed in Bartholemew (1982: 316-323) and Hedetniemi, Hedetniemi and Liesman (1988: 341).

<sup>8</sup>We assume that each signal is clear, valid, credible, and high quality. There is no "bad" information, misinformation, bad tips, or noise. The goodness of information does not change as the signals diffuse through the network. We further assume that all signals are equally good. An investor who receives a buy signal and a sell signal will weigh each the same.

<sup>9</sup>If both signals are received by the same investor, we assume they cancel. This is similar to Farley and Schachum's (1983) assumption that two messages arriving together collide and cancel.

ence the information they send to others.<sup>10</sup> Note also that investors in the model only pass information in a very restricted sense: each only passes information to the few investors with whom he is *directly* connected (in a chain, a maximum of only two investors). This behavior is equivalent to passing information only to one's closest friends or associates.

4. Investors cannot free ride. An investor cannot observe another investor's behavior and make better (more informed) decisions. Investors do not "write off" the information received from any investor. Further, investors cannot "improve" their structural positions. For example, an investor at the end of a chain cannot move toward the center.

## 2.2. Example Chain Information Network

This example illustrates the mechanics of the basic model. Consider the chain introduced above ( $n = 6, k = 2$ ). The neighborhoods associated with each of the 6 investors are:

<i>Investor</i>	<i>Neighborhood</i>
1	(1, 2, 3)
2	(1, 2, 3, 4)
3	(1, 2, 3, 4, 5)
4	(2, 3, 4, 5, 6)
5	(3, 4, 5, 6)
6	(4, 5, 6)

Since each investor is a candidate to receive the initial buy signal as well as the initial sell signal, the various possible pairs of investors at which information enters the system from exogenous sources are described by all pairs  $(i, j)$  where  $i = 1, 2, \dots, 6$  and  $j = 1, 2, \dots, 6$ . There are 36  $(i, j)$  combinations in a 6-investor chain, as shown in Table 1. These combinations are labelled "impact sets." In each impact set, the first element receives a buy signal and the second element receives a sell signal. Each signal travels from the impacted investor to all other investors in its neighborhood, i.e., no more than  $k$  lines away, at which point the transmission period expires.

Investors examine the content of their information sets and make trading decisions (offer to buy, to sell, or hold). Offers are submitted to an auctioneer<sup>11</sup> who compares the number of buy offers to sell offers and adjusts the price by  $f(j)$ . We

<sup>10</sup>Why pass on information? Though there are obvious incentives to hoard, there are also incentives to transmit. Interpersonal communication is a type of repeated game in which one investor sends information to another to ensure that he will receive information from that investor in the future; this is consistent with the principle of *reciprocity* commonly observed in networks (e.g., Wellman and Berkowitz 1988; Laumann and Marsden 1982).

<sup>11</sup>The auctioneer is a fiction used to match and set prices according to (2). The auctioneer also intervenes to clear the market by providing  $j = B - S$  units.

TABLE 1  
A list of all possible impact sets and resulting price changes for a six-investor chain<sup>1</sup>

Number	Impact Set	Buy Nodes	Price Change	Sell Nodes
1	(1, 2)	-	Dn	(4)
2	(2, 1)	(4)	Up	-
3	(1, 3)	-	Dn	(4, 5)
4	(3, 1)	(4, 5)	Up	-
5	(1, 4)	(1)	Dn	(4, 5, 6)
6	(4, 1)	(4, 5, 6)	Up	(1)
7	(1, 5)	(1, 2)	Dn	(4, 5, 6)
8	(5, 1)	(4, 5, 6)	Up	(1, 2)
9	(1, 6)	(1, 2, 3)	-	(4, 5, 6)
10	(6, 1)	(4, 5, 6)	-	(1, 2, 3)
11	(2, 3)	-	Dn	(5)
12	(3, 2)	(5)	Up	-
13	(2, 4)	(1)	Dn	(5, 6)
14	(4, 2)	(5, 6)	Up	(1)
15	(2, 5)	(1, 2)	-	(5, 6)
16	(5, 2)	(5, 6)	-	(1, 2)
17	(2, 6)	(1, 2, 3)	Up	(5, 6)
18	(6, 2)	(5, 6)	Dn	(1, 2, 3)
19	(3, 4)	(1)	-	(6)
20	(4, 3)	(6)	-	(1)
21	(3, 5)	(1, 2)	Up	(6)
22	(5, 3)	(6)	Dn	(1, 2)
23	(3, 6)	(1, 2, 3)	Up	(6)
24	(6, 3)	(6)	Dn	(1, 2, 3)
25	(4, 5)	(2)	Up	-
26	(5, 4)	-	Dn	(2)
27	(4, 6)	(2, 3)	Up	-
28	(6, 4)	-	Dn	(2, 3)
29	(5, 6)	(3)	Up	-
30	(6, 5)	-	Dn	(3)
31	(1, 1)	-	-	-
32	(2, 2)	-	-	-
33	(3, 3)	-	-	-
34	(4, 4)	-	-	-
35	(5, 5)	-	-	-
36	(6, 6)	-	-	-

<sup>1</sup>Table 1 does not display the complete set of investors who receive the signals for any of the 36 impact sets. For example, the full set of investors who get signals when the original signals impact (1,2) is (1,2,3,4). Of this four-element set, (1,2,3) have both signals and therefore hold, while (4) only has the sell signal and therefore offers to sell.

restrict our attention to symmetric  $f(j)$  functions, i.e.,  $f(j) = -f(-j)$ . We define  $j$  for a given impact set as

$$j = B - S \quad (1)$$

where  $B$  = number of buy offers and  $S$  = number of sell offers.



The price change associated with a given impact set is defined as

$$P_2 = P_1 + f(j) \tag{2}$$

where  $P_2$  = price in (present) period,  $P_1$  = price in previous period. Note that a price change occurs only when there is an imbalance in the numbers of buy and sell offers. If  $B = S$  for a particular impact set, the price change is zero ( $j = 0$ ).

Table 1 displays price effects for each of the 36 scenarios (impact sets) in the example network. The price effects are summarized as follows:

<i>Price Effect</i>	<i>Number of Scenarios</i>	<i>Probability</i>
Price Up	12	12/36
Price Down	12	12/36
Price Hold	12	12/36

Each of the 36 scenarios (impact sets) is equally likely, due to the randomized entry of exogenous buy and sell signals. Hence, in this example chain network,  $\text{Prob}(\text{price up}) = \text{Prob}(\text{price down}) = \text{Prob}(\text{hold}) = \frac{1}{3}$ . Imbalances occur in 24 scenarios (12 where  $B > S$ , 12 where  $B < S$ ), and buy offers and sell offers are equal in the remaining 12 scenarios. (Note that all chains do not result in equally probable price effects.)

Because the model is symmetric with respect to signals that enter the system in the transmission period (i.e., the exogenous signals are equally likely to impact any investor), and two opposite but equally good exogenous signals are introduced, the probability of a price increase is always equal to the probability of a price decrease. Hence,

$$E(P_2) = P_1 \tag{3}$$

Though (3) is always true in our model, the volatility of prices (defined in Section 2.3) varies across networks.

### 2.3. Price Volatility and Trading Volume

Price volatility in the model is defined as the variance of prices in the trading period. Price volatility is often defined over time; in our model, however, we define volatility as the variance of the distribution of price effects across all possible scenarios (impact sets) for a chain with a specified  $n$  and  $k$ . If the chain network remains unchanged across multiple time periods (i.e.,  $n$  and  $k$  stay the same, signals enter at random points in the network, and there is no memory or learning), then the price path across time will have the same variance as the one-period price variance.

In this section, our objective is to derive closed form expressions that link price volatility to the network parameters,  $n$  and  $k$ . We define  $\text{Prob}(j) = P(B = S + j)$ . For symmetric  $f(j)$  functions with  $E(P_2) = P_1$ , we have

$$\text{Price Volatility} = \sum_j \text{Prob}(j)(f(j))^2 \tag{4}$$

Our next step is to link  $\text{Prob}(j)$  to  $n$ ,  $k$  and  $j$ . Proposition 1 considers all cases where buy and sell signals are transmitted to  $2k + 1$  investors. Proposition 2 generates closed form expressions for all possible  $\text{Prob}(j)$ . Proposition 3 provides the probability that buy offers exceed sell offers and vice versa. Proofs of all propositions are provided in Appendix B. An example illustrating the counting procedure used in Proposition 2 is also provided in Appendix B.

**PROPOSITION 1** *If the buy and sell signals are each transmitted to  $2k + 1$  investors (i.e., the maximum neighborhood), then the number of buy offers is equal to the number of sell offers. (Proposition 1 only considers the number of offers; it does not imply that all or some investors necessarily get both the buy and sell signals.)*

**PROPOSITION 2** *For a chain with transmission period  $k$  ( $0 < k < n - 1$ ) and  $n$  ( $> k + 1$ ) investors, the probability that the number of buy offers ( $B$ ) is equal to the number of sell offers ( $S$ ) +  $j$  is equal to*

$$P(B = S + j) = 0 \quad j > \text{Min}(k, n - k - 1) \quad \text{or} \quad j < -\text{Min}(k, n - k - 1)$$

$$P(B = S + j) = \frac{2(n - 2j)}{n^2} \quad 0 < j \leq \text{Min}(k, n - k - 1)$$

$$P(B = S + j) = \frac{2(n - 2j)}{n^2} \quad 0 > j \geq -\text{Min}(k, n - k - 1)$$

**PROPOSITION 3** *For a chain with transmission period  $k$  ( $> 0$ ) and  $n$  ( $> k + 1$ ),*

1. *Probability that buy offers exceed sell offers =  $2k(n - 1 - k)/n^2$*
2. *Probability that buy offers equal sell offers =  $(n - 1 - 2k)^2 + 2n - 1/n^2$*

The auctioneer adjusts the price as defined in (1). The auctioneer can use various price adjustment functions. The simplest function assumes the auctioneer adjusts the price upward or downward by some constant ( $c$ ) when the number of buy offers and the number of sell offers are not equal, such that

$$\begin{aligned} f(j) &= 0 & \text{if } j &= 0 \\ &= c & \text{if } j > 0 \\ &= -c & \text{if } j < 0 \end{aligned} \tag{5}$$

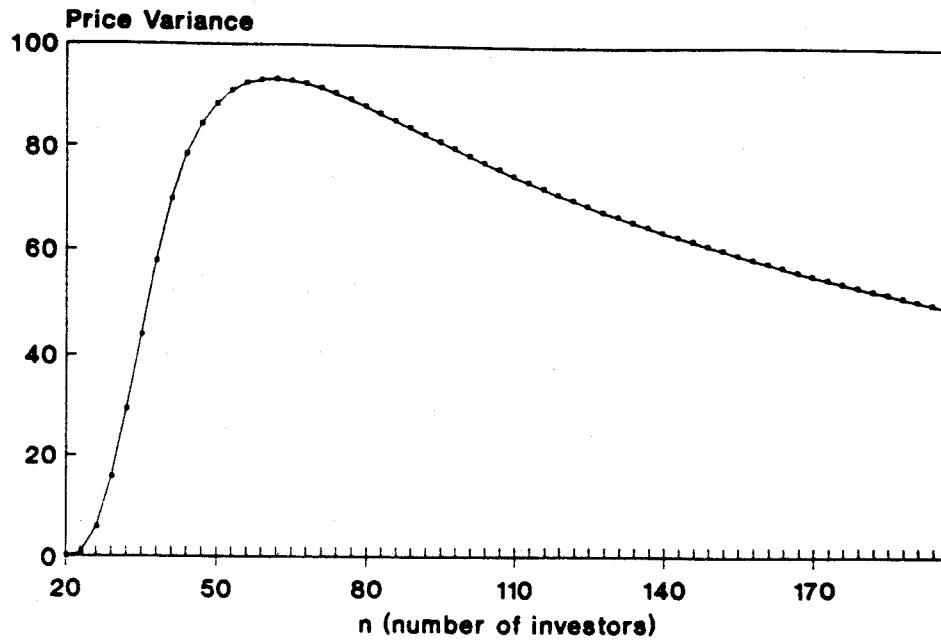
Many other  $f(j)$  are possible. Three functions (constant, linear, and exponential) are defined and illustrated in Appendix A.

Because the model is designed so that the probability of an upward price adjustment equals the probability of a downward price adjustment, the variance of prices is proportional to the probability of an upward (downward) price adjustment. Assuming the constant price adjustment function (5), it can be verified that

$$\text{Price Variance} = 2 \times \text{Prob}(\text{Price Up}) c^2 \tag{6}$$

where

$$\text{Prob}(\text{Price Up}) = \sum_{j>0} \text{Prob}(j) \tag{7}$$



Note:  $k=20$

FIGURE 2A. Price volatility by number of investors.

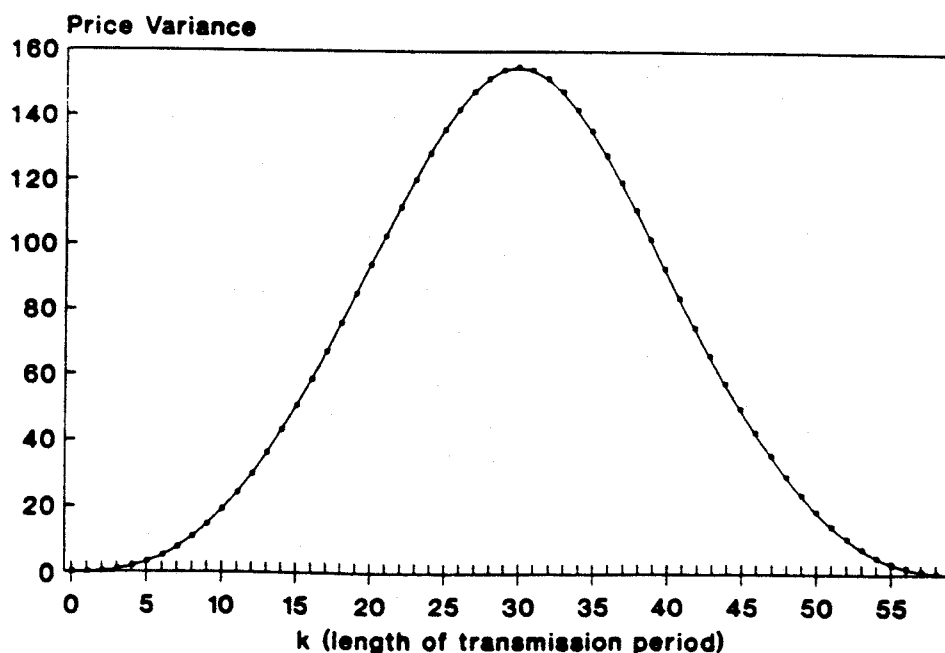
Increasing the number of investors ( $n$ ) has a nonmonotonic effect on price volatility, even when the length of the transmission period ( $k$ ) is held constant. For illustration, Figure 2A shows the effect of varying  $n$  from 21 to 201 with  $k = 20$  (with the linear price adjustment function). As  $n$  increases from 0, the variance of prices rises until  $n = 2k + 1$  (the maximum variance) but as  $n$  increases beyond  $2k + 1$ , the variance declines.

Similar nonmonotonic results are obtained by varying the length of the transmission period ( $k$ ), holding the number of investors ( $n$ ) constant. Consider Figure 2B, which shows the effect of varying  $k$  from 0 to 61, with  $n$  held constant at 61. Price volatility is the highest for  $k = (n - 1)/2$  and declines as  $k$  moves away from  $(n - 1)/2$  in either direction. Thus,  $n$  and  $k$  are complementary in the sense that either could be varied to change price volatility.

Because the auctioneer intervenes to clear the market by providing  $j (= B - S)$  units, trading volume in a given scenario (impact set) is defined as

$$\text{Trading Volume} = \text{Max}(B, S). \tag{8}$$

Our objective in this section is to generate closed-form expressions for expected trading volume as a function of the network parameters,  $n$  and  $k$ . In Proposition 4, we link the expected volume for a chain with parameters  $n$  and  $k + 1$  to the expected volume for a chain with parameters  $n$  and  $k$ . In Proposition 5, we use the recursive relationship from Proposition 4 to develop a closed form expression for expected volume in a chain with parameters  $n$  and  $k$ . Proofs for these propositions are provided in Appendix B. (The probability mass functions for volume have



Note:  $n=61$

FIGURE 2B. Price volatility by transmission period.

also been derived, which may be used to determine other moments, but are not presented in this paper to conserve space.)

PROPOSITION 4  $E(\text{Volume}(n, k + 1)) = E(\text{Volume}(n, k)) + 2n^2 - 8nk - 8n + 6k^2 + 12k + 6/n^2$

PROPOSITION 5 For an  $n$  investor chain with transmission time  $k$ ,

$$E(\text{Volume}(n, k)) = \frac{n^2 - n - 4nk^2 + 2kn^2 - 4kn + 2k^3 + 3k^2 + k}{n^2}$$

As illustrated in Figure 3A, increasing the length of the transmission period ( $k$ ) for a fixed number of investors ( $n$ ) has a nonmonotonic effect on expected trading volume. Varying  $k$  for a fixed  $n$  generates nonmonotonic effects on both price volatility and expected volume, but the value of  $k$  that yields *maximum* expected volume is lower than the value of  $k$  that yields *maximum* price volatility, for the same fixed  $n$  (for all  $n > 3$ ). Let  $k_p$  be the value of  $k$  that generates maximum price volatility and  $k_v$  be the value of  $k$  that generates maximum expected trading volume. From Propositions 3 and 5, it can be seen that

$$k_p = (n - 1)/2 \quad (9)$$

$$k_v = \frac{(8n - 6) - \sqrt{16n^2 + 12}}{12} \leq \frac{n}{3} \quad (10)$$

It is easily verified that  $k_v < k_p$  for  $n > 3$ .

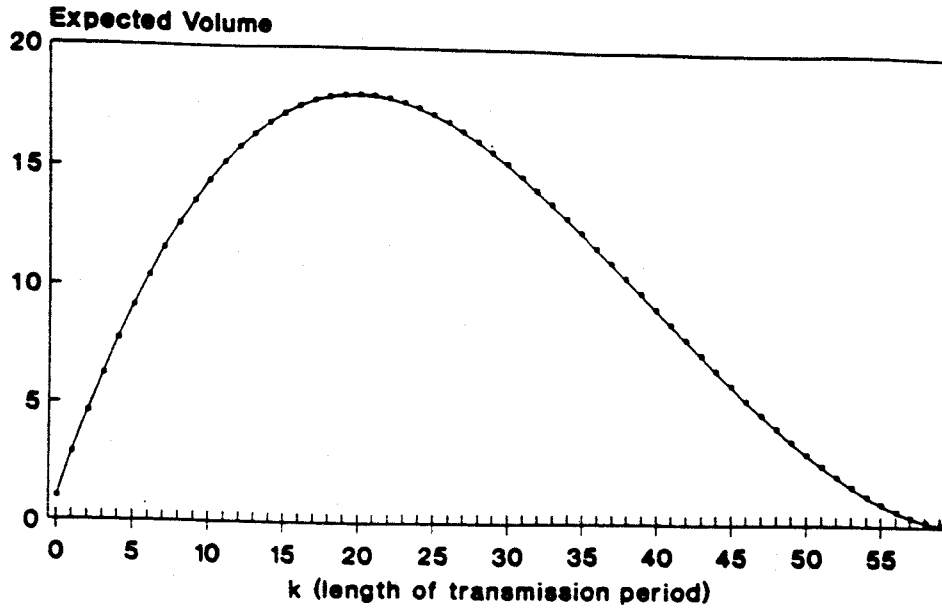


FIGURE 3A. Expected volume by transmission period for  $n = 61$ .

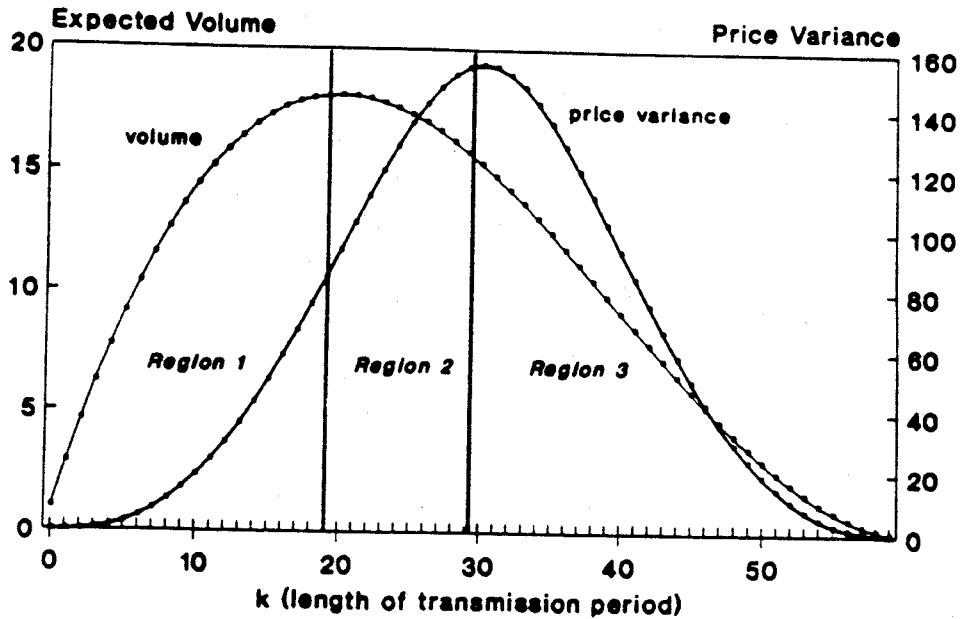


FIGURE 3B. Expected volume and price volatility by transmission period.

Given (9) and (10), the model generates three different "regions" with different effects on market behavior: Region 1, where price volatility and expected volume are positively correlated and increase as  $k$  increases; Region 2, where price volatility increases and expected volume declines as  $k$  increases; and Region 3, where price volatility and expected volume are positively correlated and decrease as  $k$  increases (see Figure 3B).

### 3. COMPLEX INFORMATION NETWORKS

Of the many ways in which the model could be extended (see Section 5), consideration of more complex information networks is an important direction. This permits exploration of the intuition that different types of graphs (see Figure 1) have different effects on market behavior (holding other factors constant). In this section, we examine the effects of circles, complete graphs, and hub-and-spoke graphs on market behavior.<sup>12</sup> (The mechanics of the model remain the same as in the chain graph.)

Price movements in the model are generated by an imbalance of buy offers and sell offers ( $B > S$  or  $B < S$ ). An imbalance can only occur when the two neighborhoods associated with an impact set are unequal in size. By Proposition 1 (Section 2), when each of the buy and sell signals is transmitted to  $2k + 1$  investors (i.e., the maximum neighborhood), the number of buy offers equals the number of sell offers and no imbalance occurs. In a circle graph or complete graph, the neighborhoods of every impact set  $(i, j)$  are equal in size; no matter where a buy (sell) signal enters the network, it is always transmitted to  $2k + 1$  investors. Therefore, unlike the chain, in circles and complete graphs the number of buy offers always equals the number of sell offers and prices never change (price variance = 0).<sup>13</sup>

Though prices do not change in circles and complete graphs, there is an important difference between the two types. In a complete graph, if  $k > 0$ , all investors are fully informed (buy, sell) no matter where the exogenous signals enter the system. In a circle, however, some investors may not be fully informed (but prices still would not change).

A hub-and-spoke graph is composed of a central point (hub) connecting multiple chains (spokes). We denote the number of spokes as  $d$  (the degree of the central node), number of investors as  $n$ , and length of each spoke as  $l$ . (For simplicity, it is assumed that all spokes are of equal length  $(l)$ .) The hub-and-spoke graph in Figure 1, for example, is described as  $d = 3$ ,  $l = 2$ ,  $n = 7$  (where  $n = dl + 1$ ).

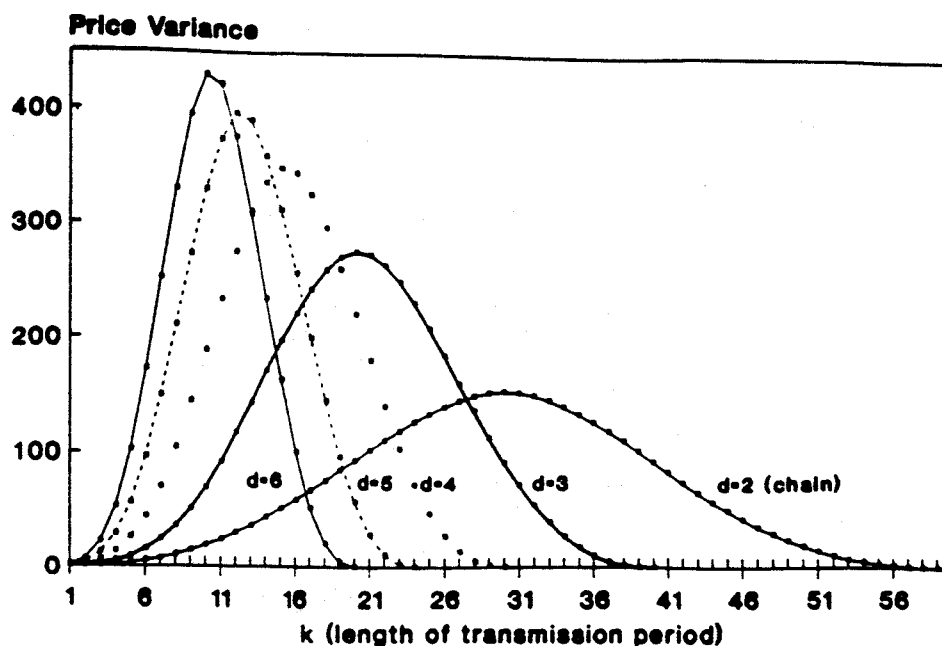
Comparison of the results of the chain graph with the results of the hub-and-spoke graph suggests two main points.

(1) The results obtained by analysis of the chain graph are “robust” in the sense that both the chain and hub-and-spoke graphs generate nonmonotonic effects on price variance and trading volume.<sup>14</sup> For illustration, consider the examples in Figures 4A and 4B. Suppose the same number of investors ( $n = 61$ ) were arranged in several different hub-and-spoke configurations. These configurations vary by the number of spokes ( $d$ ), such that  $d = 2, 3, 4, 5, 6$ . Because  $n$  is fixed, the length  $(l)$  of each spoke decreases as more spokes ( $d$ ) are added. (Note that a hub-and-spoke graph with  $d = 2$  is a chain.) Price effects for transmission periods of varying duration  $(k)$  and fixed  $n$  are displayed in Figure 4A. In both chain and hub-and-spoke

<sup>12</sup>For circle graphs and complete graphs, we restrict our attention to price effects. Both price and volume effects are analyzed for the hub-and-spoke network.

<sup>13</sup>Circles and complete graphs are cyclic graphs, so the same point could receive the same signal more than once. For simplicity, we assume that multiple receptions of the same signal are interpreted exactly as a single reception.

<sup>14</sup>Our analyses of price variance and expected trading volume for hub-and-spoke graphs are based on an exact counting algorithm (not a simulation).



Note:  $n=61$  for each graph

FIGURE 4A. Price variance for hub-and-spoke graphs.

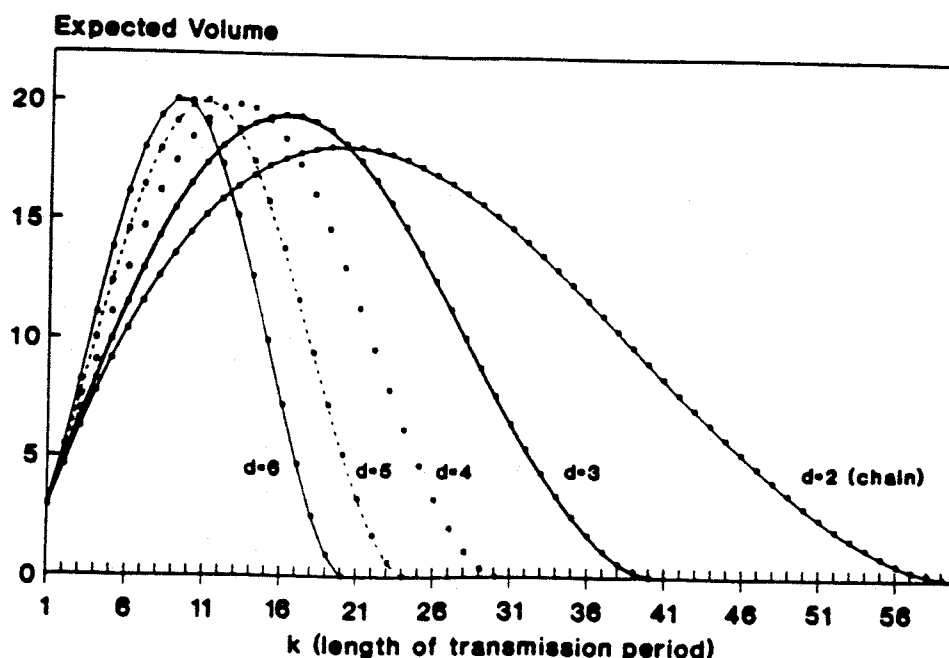
networks, similar nonmonotonic patterns are generated: price volatility increases as  $k$  increases from 1, reaches its maximum at  $k = l$ , and declines thereafter. Similar results are obtained for expected trading volume. The chain ( $d = 2$ ) and various hub-and-spoke graphs produce nonmonotonic effects on expected volume (Figure 4B).

(2) Despite similar nonmonotonic effects, there are important differences. For example, for the same values of  $n$  and  $k$ , the chain and various hub-and-spoke graphs produce different values for price variance and expected volume (Figures 4A, 4B). When  $n$  is held constant, maximum price volatility (obtained in each graph at  $k = l$ ) increases as more spokes are added (Figure 4A). That is, as the same number of investors ( $n$ ) are arranged in networks with more and shorter spokes, the maximum volatility of the system increases.<sup>15</sup>

#### 4. DISCUSSION

The results show that price volatility and trading volume can be generated by the structure of the communication network among investors. Offers are made, prices change, and volume is generated even though the “correct” trading decision—the decision based on the assumption of full information—would be to not make any offers. When investors are fully informed in our model, market activity does not take place. Prices do change, however, because a communication network can cre-

<sup>15</sup>Hub-and-spoke networks do not always generate maximum price volatility greater than a chain. For example, consider a wheel graph (a hub-and-spoke graph with  $d = n - 1$ ) with  $n$  investors and a chain graph with  $n$  investors. The wheel generates lower maximum volatility (at  $k = l$ ) than the chain for  $n < 43$ , but greater than the chain for  $n > 43$ .



Note:  $n=61$  for each graph

FIGURE 4B. Trading volume for hub-and-spoke graphs.

ate an imbalance of buy offers and sell offers ( $B > S$  or  $B < S$ ). Such an imbalance occurs when the two neighborhoods associated with a particular impact set are unequal in size. For example, a chain network always generates a price change when one signal impacts the “central” region (where each investor has  $2k + 1$  investors to the left and to the right) and the other impacts an “end” (where each investor’s neighborhood is smaller than  $2k + 1$ ). In this situation, at least one investor’s information set is incomplete (and not empty) and an imbalance of buy and sell offers occurs. Similarly, volume in the model is generated whenever the neighborhoods associated with each signal do not perfectly overlap. Volume is created because at least one investor’s information set is incomplete (and not empty). But prices would not change and volume would not be generated under the idealized condition of full information (e.g., as when  $k \geq n - 1$ ).

Comparison of the chain and hub-and-spoke networks suggests that the configuration of the communication network itself can explain market behavior, even if the transmission period ( $k$ ) and the number of investors ( $n$ ) are held constant. For almost all  $(n, k)$  combinations, a chain and a hub-and-spoke will generate different price variances and different expected volumes, even when  $k$  is the same and  $n$  is the same in each network (see, e.g., Figures 4A and 4B). Thus, in addition to the effects of size ( $n$ ) and transmission duration ( $k$ ) on market behavior, the pattern of communication among investors also impacts price volatility and trading volume.

Even though the underlying networks in the model do not represent the actual (and unknown) communication network, the model generates price and volume patterns that are consistent with known financial market behaviors. For example, price volatility and trading volume are generally positively correlated in our model (see, e.g., Figure 3B), as in real financial markets (e.g., Karpoff 1987). The observation



that trading volume on Monday is lower than the weekday average is a well-known puzzle in financial theory (Mulherin and Gerety 1988). In our model, this “Monday effect” can be produced by increasing the length of time in which information diffuses (which represents the longer duration between Friday and Monday compared with two contiguous weekdays). This result suggests that the Monday volume effect might only be a simple consequence of the longer duration of time for information to travel throughout the communication network.

Another puzzle is the so-called “anomalous” empirical finding that price volatility is higher (e.g., Fama 1965) and trading volume is lower (Mulherin and Gerety 1988) over the weekend. These “anomalous” market behaviors can be easily produced in the model. They are translated into the model by three inequalities that must be satisfied:

$$k_1 > k_2$$

$$\text{Expected Volume}(k_1) < \text{Expected Volume}(k_2)$$

$$\text{Price Volatility}(k_1) > \text{Price Volatility}(k_2)$$

where  $k_1$  = the Friday–Monday transmission period and  $k_2$  = the transmission period between any two contiguous weekdays. Note that if  $k_1$  is in the neighborhood of  $k_p$  and  $k_2$  is in the neighborhood of  $k_v$ , the three inequalities described above are satisfied.

The Monday volume-volatility effect in the model is illustrated by Region 2 in Figure 3B. The weekend value of  $k$  occurs in Region 2 (which captures Monday volatility-volume effects), while the weekday value of  $k$  occurs in Region 1 (which captures the positive correlation of volatility and volume). This result suggests that both the Monday volume and price volatility effects are simply due to the longer duration of time between Friday and Monday.

Baker (1984a: 804–6) posited a curvilinear relationship between market performance (price volatility) and size (number of traders). The nonmonotonic relationship of  $n$  and price volatility (Figure 2A) in the model is consistent with this expectation. It is often argued by economic theorists that large  $n$  improves market performance. However, large  $n$  does not “wash out” the effect network structure on market behavior. Low price volatility, for example, occurs in a large market only if  $n$  is large *relative* to  $k$ , or  $n$  and  $k$  are comparable in magnitude. Whenever  $k = (n - 1)/2$ , the network system produces maximum price volatility no matter how large (or small)  $n$  may be.

The nonmonotonic relationship of  $n$  and market behavior is consistent with Baker’s observations. But the model also produces nonmonotonic relationships between transmission duration ( $k$ ) and price volatility (Figure 2B) and between  $k$  and expected volume (Figure 3A). Prior research on financial markets has not focused on the relationship between transmission time and market behavior (except for the Monday effect). The nonmonotonic patterns produced by the model, therefore, are provocative hypotheses. They imply, for example, that regulatory manipulations that influence transmission time could increase *or* decrease price volatility and trading volume, even when the number of investors ( $n$ ) and the configuration of the communication network remain the same. Lengthening the trading day or trading week, for instance, could dampen or exacerbate price volatility. Similarly, improving the level of “contact technology” (equivalent to increasing  $k$ ) could impair or improve

market behavior. If  $k$  were initially less than  $(n - 1)/2$ , improving the level of contact technology (increasing  $k$ ) would impair market performance.

Finally, though we do not explicitly consider the effect of changes in network structure, a few simple cases illustrate how the model might be used to explain system dynamics over time. High price volatility might cause investors to leave (or enter) the system over time. The results imply that the exit (or entrance) of market participants would influence the volatility of prices. Since a wide range of price volatilities can be reproduced by varying  $n$ , changes in the number of investors could dampen or exacerbate the volatility of prices. Prices could become more or less volatile, for example, if exit caused the chain network to break into disconnected segments.

**CASE 1** Assume a chain with  $n$  investors (where  $n$  is an odd number) and  $k = (n - 1)/4$  that breaks into two equal segments by the exit of the central investor. For example, consider the chain network with  $n=17$  and  $k = 4$ .<sup>16</sup> The variance of prices = .110. The exit of the investor at the precise center of the chain (node 9) causes the 17-node chain to break into two chain segments with 8 investors each. Assume the length of the transmission period ( $k$ ) remains constant and the exogenous buy and sell signals impact one of the chain segments. The price variance of the new system increases to .141.

**CASE 2** Assume a chain network with the same number of investors ( $n = 17$ ) but a longer transmission period,  $k = (n - 1)/2$ . Price variance in this chain = .196. Assume that the same central investor (node 9) leaves the system, again creating two independent chain segments of length 8,  $k$  remains the same, and both exogenous signals impact the same segment. The variance of prices now falls to 0.

## 5. CONCLUSIONS

The principal result of the model is that both price volatility and expected trading volume can be influenced by the structure of the information network. The model produces market behaviors that are consistent with well-known financial market behaviors, even though the networks analyzed in the model do not represent the actual investor communication network. The model supports Baker's (1984a, 1984b) finding that trading networks influence price volatility. It adds to the importance of Shiller and Pound's (1989) discovery of an investor communication network because it demonstrates that the network can influence market behavior. Overall, the model suggests that the real communication network among investors influences market activity.

The study contributes to the structural analysis of economic institutions by explicitly modeling the effect of network structure on market outcomes. It supports the logical consistency and empirical evidence of the proposition that markets are social structures and that social structures have consequences. Markets, like all social institutions, are patterns of roles and relationships that influence the behavior of individual members and the system as a whole.

<sup>16</sup>For simplicity, the cases use the constant price adjustment function (5) with  $c = 1$ .

The model also demonstrates that the conclusions of previous sociological studies (Baker 1984a, 1984b) are applicable over a wide range of parameter values. Variations in size, for example, produce nonmonotonic market behaviors over a wide range of network sizes. In addition, the predictions of the model suggest new hypotheses. Variation in transmission time ( $k$ ), for example, can produce nonmonotonic market behaviors that are very similar to those produced by variation in size ( $n$ ). Moreover, the model shows that the configuration of the communication network itself influences market behavior, even when the number of investors ( $n$ ) and transmission duration ( $k$ ) are held constant. Thus, the effects of transmission time and network configuration on market outcomes are two areas for empirical study.

The model could be extended in various ways. By examining more complex graphs, we have begun to explore an important avenue of further research. Other natural extensions could be investigated. For example, in our model exogenous signals randomly enter the information network (i.e., each investor is equally likely to receive the signal). This design permits, among other things, to "hold constant" the effects of information entry and thereby isolate the effects of network structure. In real information networks, however, each node would not have an equiprobable chance of receiving exogenous signals. Signals might be more likely to enter at some points than others, such as the "center" of a network, or might tend to enter at the same points (e.g., the establishment of a regular communication channel). To account for such factors, the model could be extended by assigning investors different probabilities of receiving the external signal. Further, in our model, every investor in the neighborhood of an investor who receives the exogenous signal also gets the signal (i.e., information is passed on with a probability of one during the transmission period). A natural extension of the model would vary by investor the probability of sending or receiving a signal.

Other extensions of the model would include information of different types, cost, quality, and credibility; more complex investor decision rules; and different  $f(j)$  for the auctioneer (see Appendix A). We suspect, however, that such extensions would only modify, not invalidate, our principal result that communication networks can influence market behavior.

## REFERENCES

- Adler, P., and Adler, P. (eds.) (1984) *Social Dynamics of Financial Markets*. Greenwich, CT: JAI Press.
- Baker, W. E. (1990) Market networks and corporate behavior. *American Journal of Sociology* (November).
- Baker, W. E. (1984a) The social structure of a national securities market. *American Journal of Sociology*, 89: 775-811.
- (1984b) Floor trading and crowd dynamics. In Patricia Adler and Peter Adler (eds.), *The Social Dynamics of Financial Markets*, Greenwich, CT: JAI Press.
- Bartholomew, D. J. (1982) *Stochastic Models for Social Processes*. New York: John Wiley & Sons.
- Berkowitz, S. D. (1988) Markets and market-areas: Some preliminary formulations. Pp. 261-303 in B. Wellman and S. D. Berkowitz (eds.), *Social Structures: A Network Approach*. Cambridge: Cambridge University Press.
- Burt, R. S. (1988) The stability of American markets. *American Journal of Sociology* 94: 356-395.
- (1983) *Corporate Profits and Cooptation*. New York: Academic Press.
- Burt, R. S. (1987) Social contagion and innovation: Cohesion and structural equivalence. *American Journal of Sociology*, 92: 1287-1335.

- Boorman, S. A. (1975) A combinatorial optimization model for transmission of job information through contact networks. *Bell Journal of Economics and Management Science*, 6: 216–249.
- Coleman, J. S. (1990) *Foundations of Social Theory*. Cambridge: Belknap Press of Harvard University Press.
- Coleman, J. R., Katz, E., and Menzel, H. (1982) *Medical Innovation: A Diffusion Study*. Indianapolis: Bobb-Merrill.
- Constantinides, G., and Bhattacharya, S. (eds.) (1988) *Frontiers of Modern Financial Theory*. Totowa, NJ: Rowman and Littlefield.
- Fama, E. F. (1965) The behavior of stock market prices. *Journal of Business* 38: 34–105.
- Farley and Schachum (1983) *Problems in Broadcast-Based Communication Networks*, Manuscript.
- Frenzen, J. K., and Davis, H. L. (1990) Purchasing behavior in embedded markets. *Journal of Consumer Research* 17: 1–12.
- Granovetter, M. S. (1985) Economic action and social structure: The problem of embeddedness. *American Journal of Sociology* 91: 481–510.
- Granovetter, M. (1974) *Getting a Job: A Study of Contacts and Careers*. Cambridge, MA: Harvard.
- Harary, F., Norman, R. Z., and Cartwright, D. (1965) *Structural Models: An Introduction to the Theory of Directed Graphs*. New York: John Wiley.
- Hedetniemi, S. M., Hedetniemi, S. T., and Liesman, A. L. (1988) A survey of gossiping and broadcasting in communication networks. *Networks* 18: 319–349.
- Hirschman, A. (1982) Rival interpretations of market society: Civilizing, destructive, or feeble? *Journal of economic literature* 20: 1463–1484.
- Karpoff, J. (1987) The relation between price changes and trading volume: A survey. *Journal of Financial and Quantitative Analysis*, January: 109–126.
- Laumann, E. O., and Marsden, P. V. (1982) Microstructural analysis in interorganizational systems. *Social Networks* 4: 329–348.
- Lorrain, F., and White, H. (1971) Structural equivalence of individuals in social networks. *Journal of Mathematical Sociology* 1: 49–80.
- Mulherin, J. H., and Gerety, M. S. (1988) Trading volume on the NYSE during the twentieth century: A daily and hourly analysis. Ms. Securities and Exchange Commission and Department of Economics, Clemson University.
- Robertson, T. S. (1971) *Innovative Behavior and Communication*. New York: Holt Reinhardt and Winston.
- Rogers, E. M. (1983) *Diffusion of Innovations*, Third edition. New York: Free Press.
- Rogers, E. M., and Kincaid, D. L. (1981) *Communication Networks: Toward a New Paradigm for Research*. New York: Macmillan.
- Rogers, E. M. (1979) Network analysis of the diffusion of innovations. Pp. 137–164 in P. W. Holland and S. Leinhardt (eds.), *Perspectives on Social Network Research*. New York: Academic Press.
- Saloner, G. (1985) Old boy networks as screening mechanisms. *Journal of Labor Economics* 3: 255–267.
- Shiller, R. J., and Pound, J. (1989) Survey evidence on diffusion of interest and information among investors. *Journal of Economic Behavior and Organization* 12: 47–66.
- Stinchcombe, A. L. (1987) *Constructing Social Theories*. Chicago: University of Chicago Press.
- Wellman, B. and Berkowitz, S. D. (1988) *Social Structures: A Network Approach*. NY: Cambridge University.
- White, H. C. (1981) Where do markets come from? *American Journal of Sociology* 87: 517–547.
- (1988) Varieties of markets. Pp. 226–260 in B. Wellman and S. D. Berkowitz (eds.), *Social Structures: A Network Approach*. Cambridge: Cambridge University Press.
- Zukin, S., and DiMaggio, P. (eds.) (1990) *Structures of Capital*. Cambridge: Cambridge University Press.

## APPENDIX A

This appendix presents and compares three different price adjustment functions that could be used by the auctioneer. (Many other functions are possible.)

### Constant Price Adjustment Function

This  $f(j)$  makes a constant price change when the number of buy offers and the number of sell offers are not equal.

$$\begin{aligned} f(j) &= 0 & \text{if } j &= 0 \\ &= c & \text{if } j > 0 \\ &= -c & \text{if } j < 0 \end{aligned}$$

where  $c$  is some constant.

**Linear Price Adjustment Function**

This  $f(j)$  adjusts the price in the trading period proportional to the difference between the number of buy offers and number of sell offers.

$$f(j) = j.$$

**Exponential Price Adjustment Function**

This function "penalizes" large differences between number of buy offers and number of sell offers. Let  $c$  be some constant value.

$$\begin{aligned} f(j) &= 0 & \text{if } j &= 0 \\ &= c2^j & \text{if } j > 0 \\ &= c2^{-j} & \text{if } j < 0. \end{aligned}$$

To illustrate the three functions, consider a chain with  $n = 61$  investors and  $k = 30$ . The maximum value of  $j$  for any  $k$  is 30. The following functions generate the same average price adjustment (i.e., each generates the same average price adjustment of  $31/2 = 15.5$ ).

1. *Constant price adjustment,  $c = 15.5$*

$$\begin{aligned} f(0) &= 0 \\ f(j) &= 15.5 & \text{for } j > 0 \\ f(j) &= -15.5 & \text{for } j < 0. \end{aligned}$$

2. *Linear price adjustment*

$$f(j) = j$$

3. *Exponential price adjustment*

$$\begin{aligned} f(j) &= 0 & \text{for } j &= 0 \\ &= 465 * \frac{2^j}{(2^{31} - 2)} & \text{for } j > 0 \\ &= -465 * \frac{2^{-j}}{(2^{31} - 2)} & \text{for } j < 0. \end{aligned}$$

The different price adjustment functions do not affect  $E(P_2)$ . We assume the price adjustment mechanisms are used symmetrically for price increases and decreases. The probability of observing a difference between buy and sell offers equal to  $j$  is symmetric about zero. Thus, the expected price change in the trading period is zero. Note, however, that these functions affect the variance of the prices in the trading period. Figure 5 shows the effect of the three price adjustment functions on the variance of prices as the length of the transmission period ( $k$ ) varies.

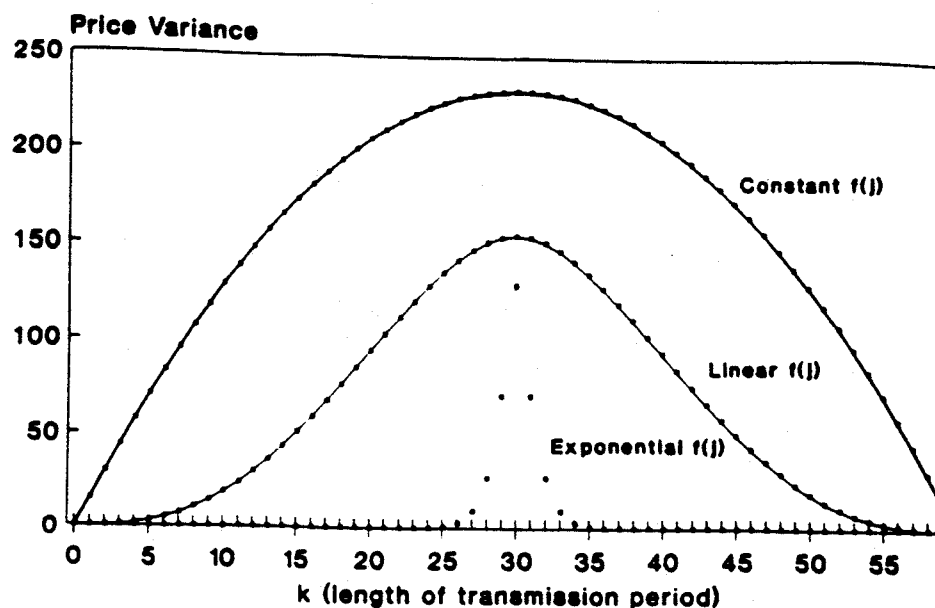


FIGURE 5. Price volatility by transmission period for three price adjustment functions.

In our model, the price volatility always attains its maximum at  $k = (n - 1)/2$ , and the variance is always symmetric around  $k = (n - 1)/2$ , no matter which function is used.

## APPENDIX B

This appendix provides closed-form expressions for price volatility and expected trading volume for  $n$ -investor chain networks with transmission time  $k$ .

### Price Volatility in chain networks

**PROPOSITION 1** *If the buy and sell signals are each transmitted to  $2k + 1$  investors (i.e., the maximum neighborhood), then the number of buy offers is equal to the number of sell offers. (Proposition 1 only considers the number of offers; it does not imply that all or some investors necessarily get both the buy and sell signals.)*

*Proof* If  $t$  investors ( $0 \leq t \leq 2k + 1$ ) receive both buy and sell signals, then the number of buy offers is  $2k + 1 - t$  and the number of sell offers is  $2k + 1 - t$ , and hence the number of buy offers equals the number of sell offers.  $\square$

Proposition 1 implies that for a difference between buy and sell offers to exist, at least one of the signals must impact on the  $k$  investors at either end of the chain and thus be distributed to  $< 2k + 1$  investors. We also note that for any  $j > 0$ , the probability that buy offers exceeds sell offers by  $j$  units is equal to the probability that the sell offers exceed buy offers by  $j$  units because both buy and sell offers impact on an investor with equal probability. We also note that for  $k \geq n - 1$ , the numbers of buy and sell offers are always equal to zero because all investors in the system receive both signals (perfect information).

**PROPOSITION 2** For a chain with transmission period  $k(0 < k < n - 1)$  and  $n(> k + 1)$  investors, the probability that the number of buy offers ( $B$ ) is equal to the number of sell offers ( $S$ ) +  $j$  is equal to

$$P(B = S + j) = 0 \quad j > \text{Min}(k, n - k - 1) \quad \text{or} \quad j < -\text{Min}(k, n - k - 1)$$

$$P(B = S + j) = \frac{2(n - 2j)}{n^2} \quad 0 < j \leq \text{Min}(k, n - k - 1)$$

$$P(B = S + j) = \frac{2(n - 2j)}{n^2} \quad 0 > j \geq -\text{Min}(k, n - k - 1)$$

*Proof* Since the cases  $j > 0$  and  $j < 0$  are symmetric i.e., obtained by exchanging the impact locations of buy and sell signals, we consider only the case  $j > 0$  i.e., Buy offers exceeding sell offers. We prove this result by considering two possible conditions.

*Condition 1:*  $k \leq (n - 1)/2$ .

We note that since each signal impacting the chain is received by at least  $k + 1$  investors and at most  $2k + 1$  investors, the maximum value of  $B - S$  is  $k$ . When  $B - S = j (> 0)$ , we have to count all possible scenarios that can generate exactly  $j$  more buy offers than sell offers. We first note that for every node  $i$  in the chain, there is a corresponding symmetric node located at  $n - i + 1$ . There are *three types* of nodes: left-end nodes, right-end nodes and nodes in the middle. Left-end nodes are node locations  $1, 2, \dots, k$ , right-end nodes are node locations  $n - k + 1, \dots, n$  and nodes  $k + 1, \dots, n - k$  are nodes in the middle. Note that from Proposition 1, we need consider only situations where the sell signal is transmitted to  $< 2k + 1$  investors.

We count all possible scenarios by dividing them up into the following three cases as defined below.

- The sell signal impacts end node locations and the buy signal also impacts end node locations. Thus both buy and sell signals are transmitted to  $< 2k + 1$  investors.

To count these scenarios, we consider the following two sets of scenarios which are symmetric (i.e., obtained by replacing each location  $i$  by  $n - i + 1$ )

**CASE 1** The sell signal impacts location  $m$ , one of the **left end node** locations (i.e.,  $m = 1, 2, \dots, k - j$ ) and the buy signal impacts either a left end node location  $m + j$  (i.e.,  $j + 1, \dots, k$ ) or a right end node location  $n - j - m + 1$  (i.e.,  $n - j, \dots, n - k + 1$ ). Note that the right end node locations are symmetric to the left end node locations i.e., replace location  $i$  by  $n - i + 1$ .

This generates  $2 \times (k - j)$  scenarios.

**CASE 2** This second case is obtained by replacing every location  $i$  in Case 1 above by  $n - i + 1$ .

Thus, we consider the sell signal impacting location  $m$ , one of the **right end node** locations (i.e.,  $m = n, \dots, n + 1 - k + j$ ) and the buy signal impacts either a right end node location  $m - j$  (i.e.,  $n - j, \dots, n - k + 1$ ) or a left end node location  $n - m - j + 1$  (i.e.,  $j + 1, \dots, k$ ).

This too generates  $2 \times (k - j)$  scenarios.

- The sell signal impacts a left end node location and the buy signal impacts a node which is **not an end node** and is transmitted to  $2k + 1$  investors. Symmetrically, the sell signal impacts a right end node location and the buy signal impacts a node which is **not an end node** and is transmitted to  $2k + 1$  investors. This results in the following third case.

**CASE 3** The sell signal impacts a left end node location  $k - j + 1$  and the buy signal impacts nodes  $k + 1, \dots, n - k$ . Symmetrically, the sell signal impacts a right end node location  $n - k + j$  and the buy signal impacts nodes  $n - k, \dots, k + 1$ .

This generates  $2 \times (n - 2k)$  scenarios.

Thus the total number of scenarios with  $B - S = j$  is  $(2 \times (k - j) + 2 \times (k - j) + 2 \times (n - 2k))$  i.e.,  $2 \times (n - 2j)$ . Since the total number of possible scenarios is equal to  $n^2$ , we have Probability  $(B - S = j) = (2 \times (n - 2j))/n^2$ .

*Condition 2:*  $k > (n - 1)/2$ .

In this case, each signal impacting the system is received by at least  $k + 1$  investors and at most  $n$  investors ( $n < 2k + 1$ ); hence the maximum value of  $B - S$  is  $(n - k - 1)$ . It can also be verified that the number of scenarios with  $B - S = j$  is obtained by replacing  $k$  by  $k' = (n - k - 1)$  ( $k' < (n - 1)/2$ ) in (i), (ii) and (iii) in Case 1. Thus, for a given  $n$ , the probabilities that  $B - S = j$  are equal for  $k$  ( $< (n - 1)/2$ ) and  $k' = n - k - 1$ .  $\square$

**Example with  $n = 7, k = 2, j = 1$**

We provide this example to illustrate the counting procedures in Proposition 2. We consider a chain network with  $n = 7, k = 2$  and  $j = 1$ . We classify the scenarios based on the three cases defined above. The total number of possible scenarios with  $j = 1$  is  $2 \times (7 - (2 * 1)) = 10$ .

**CASE 1** The sell signal impacts node 1 and the buy signal impacts node 2 or node 6. This provides two scenarios.

**CASE 2** Symmetric with case 1, the sell signal impacts node 7 and the buy signal impacts node 6 or node 2. This provides two more scenarios.

**CASE 3** The sell signal impacts node 2 and the buy signal impacts nodes 3, 4 or 5. Symmetrically the sell signal impacts node 6 and the buy signal impacts nodes 3, 4 or 5. This generates 6 more scenarios.

Note that in the scenarios in Case 3, the buy signal is transmitted to  $2k + 1$  i.e., five investors.

Thus the total number of scenarios which generate  $B - S = 1$  is  $(2 + 2 + 6) = 10$ . Therefore  $P(B = S + 1) = \frac{10}{49}$ .

**PROPOSITION 3** For a chain with transmission period  $k (> 0)$  and  $n (> k + 1)$ ,

1. Probability that buy offers exceed sell offers =  $(2k(n - 1 - k))/n^2$
2. Probability that buy offers equal sell offers =  $((n - 1 - 2k)^2 + 2n - 1)/n^2$



*Proof*

1. By definition, the probability that buy offers exceed sell offers is

$$\begin{aligned} \sum_{j>0} P(B - S = j) &= \sum_{j=1}^{\text{Min}(k, n-k)} \frac{2(n-2j)}{n^2} \\ &= \frac{2k(n-k-1)}{n^2} \end{aligned}$$

2. By definition of probabilities,

$$P(B > S) + P(B < S) + P(B = S) = 1.$$

Also, as discussed earlier,  $P(B < S) = P(B > S)$ . Hence,

$$\begin{aligned} P(B = S) &= 1 - \frac{2 \times 2k(n-k-1)}{n^2} \\ &= \frac{(n-2k-1)^2 + 2n-1}{n^2} \end{aligned}$$

Hence the result.  $\square$

**Expected Trading Volume for chain networks**

Let  $E(\text{Volume}(n, k))$  denote the expected trading volume observed in an  $n$  investor chain with transmission time  $k$ . In this appendix we will show that

1. For an  $n$  investor chain with  $k = 0$ ,  $E(\text{Volume}(n, 0)) = (n^2 - n)/n^2$
2.  $E(\text{Volume}(n, k + 1)) = E(\text{Volume}(n, k)) + (2n^2 - 8nk - 8n + 6k^2 + 12k + 6)/n^2$
3. The relationship generated in (i) and (ii) imply that for an  $n$  investor chain with transmission time  $k$ ,  $E(\text{Volume}(n, k)) = (n^2 - n - 4nk^2 + 2kn^2 - 4kn + 2k^3 + 3k^2 + k)/n^2$

We note that (1) is easily verified. To prove (2) and (3) we will use lemma 1, which can be verified, and prove Propositions 4 and 5.

**LEMMA 1** *For an  $n$  investor chain graph, if transmission time increases from  $k$  to  $k + 1$ , trading volume changes only for the following impact sets*

1. For  $i = 1, 2, \dots, n$  and impact set  $(i, i + 2k + 2)$  with the property that at least one signal can spread to  $2k + 3$  investors for transmission time  $k + 1$ , the trading volume increases by one unit.
2. For  $i = 1, 2, \dots, n$  and impact sets  $(i, i + j)$  with  $j \geq 2k + 3$ , trading volume increases by one unit if both signals can spread to  $\leq 2k + 2$  investors for a transmission time of  $k + 1$ , otherwise trading volume increases by two units.
3. For  $i = 1, 2, \dots, n$  and impact sets  $(i, i + j)$  with  $j \leq 2k + 1$ , trading volume  $> 0$  and signals transmitted to both ends of the chain for a transmission time of  $k$ , trading volume decreases by one unit.

**PROPOSITION 4**  $E(\text{Volume}(n, k + 1)) = E(\text{Volume}(n, k)) + (2n^2 - 8nk - 8n + 6k^2 + 12k + 6)/n^2$

*Proof* To prove the recursion, we have to count the number of impact sets that fall into each of the categories in lemma 1 as transmission time increases from  $k$  to  $k + 1$ . This counting process has to be separated into four regions, which depend on the relationship between  $n$  and  $k$ , as follows:

1. Case 1 ( $n \geq 4k + 2$ ): Number of impact sets for which trading volume increases by at least one unit =

$$2 * \sum_{i=1}^{n-2k-2} \sum_{j=i+2k+2}^n 1 = (n-2k-1)(n-2k-2)$$

Number of impact sets for which trading volume increases by exactly two units

$$\begin{aligned} &= 2 * \left\{ \sum_{i=1}^{k+1} \sum_{j=i+2k+3}^{n-k-1} 1 + \sum_{i=k+2}^{n-2k-3} \sum_{j=i+2k+3}^n 1 \right\} \\ &= 2(n-3k-3)(k+1) + 2(n-2k-2)(n-3k-4) \\ &\quad - (n-2k-3)(n-2k-2) \end{aligned}$$

We also note that there are no impact sets for which trading volume decreases, thus

$$\begin{aligned} E(\text{Volume}(n, k+1)) &= E(\text{Volume}(n, k)) \\ &\quad + \frac{(n-2k-1)(n-2k-2) + 2(n-3k-3)(k+1) + 2(n-2k-2)(n-3k-4) - (n-2k-3)(n-2k-2)}{n^2} \end{aligned}$$

This equation simplifies to yield the stated result.

2. Case 2 ( $n \leq 2k + 1$ ): In this situation, there are no impact set pairs whose trading volume increases.

Impact sets with trading volume decreasing by one unit

$$\begin{aligned} &= 4 \sum_{i=1}^{n-k-1} \sum_{j=n-k}^{k+1} 1 + 2 \sum_{i=1}^{n-k-1} \sum_{j=k+2}^n 1 \\ &= -2n^2 + 8nk + 8n - 6k^2 - 12k - 6 \end{aligned}$$

Thus,

$$E(\text{Volume}(n, k+1)) = E(\text{Volume}(n, k)) + -1 * \frac{-2n^2 + 8nk + 8n - 6k^2 - 12k - 6}{n^2}$$

which is the required result.

3. Case 3 ( $2k + 2 \leq n \leq 3k + 2$ ): In this case we have no impact sets which result in trading volume increase of two units.

Impact sets with trading volume increasing by one unit

$$2 \sum_{i=1}^{n-2k-3} \sum_{j=i+2k+3}^n 1 = (n-2k-3)(n-2k-2)$$

Impact sets with trading volume decreasing by one unit

$$= 2 \sum_{i=1}^{n-2k-1} \sum_{j=n-k}^{i+2k+1} 1 + 2 \sum_{i=n-2k}^{k+1} \sum_{j=n-k}^n 1$$

$$= -n^2 + 4nk - 2k^2 + 3n - 2k$$

Thus,

$$E(\text{Volume}(n, k + 1)) = E(\text{Volume}(n, k))$$

$$+ \frac{(n - 2k - 3)(n - 2k - 2) - (-n^2 + 4nk - 2k^2 + 3n - 2k)}{n^2}$$

This simplifies to the required result.

4. Case 4 ( $3k + 3 \leq n \leq 4k + 1$ ): In this case we have the impact sets of types 1, 2 and 3 described in lemma 1.

The number of impact sets of each type are as follows:

Impact sets with trading volume increasing by at least one unit

$$2 * \left\{ \sum_{i=1}^{n-3k-3} \sum_{j=i+2k+2}^n 1 + \sum_{i=n-3k-2}^{k+1} \sum_{j=i+2k+3}^n 1 + \sum_{k+2}^{n-2k-2} \sum_{j=i+2k+2}^n 1 \right\}$$

$$= 2 * \left\{ \sum_{i=1}^{n-3k-3} (n - i - 2k - 1) + \sum_{i=n-3k-2}^{k+1} (n - i - 2k - 2) \right.$$

$$\left. + \sum_{i=k+2}^{n-2k-2} (n - i - 2k - 1) \right\}$$

$$= 2 * \{ (n - 2k - 1)(n - 3k - 3) - 0.5(n - 3k - 3)(n - 3k - 2)$$

$$+ (n - 2k - 2)(4k - n + 4) + (n - 2k - 1)(n - 3k - 3)$$

$$- 0.5(2n - 5k - 4)(k + 1)$$

Impact sets with trading volume increasing by two units

$$4 \sum_{i=1}^{n-3k-4} \sum_{j=i+2k+3}^{n-k-1} 1 = 2(n - 3k - 3)(n - 3k - 4)$$

Impact sets with trading volume decreasing by one unit

$$2 \sum_{i=n-3k-1}^{k+1} \sum_{j=n-k}^{i+2k+1} 1 = -1 * (4k - n + 3)(4k - n + 4)$$

Thus  $E(\text{Volume}(n, k + 1)) = E(\text{Volume}(n, k)) + 1/n^2 * (\# \text{ of Impact sets increasing by at least one unit} + \# \text{ of Impact sets increasing by two units} - \# \text{ of Impact sets decreasing by one unit})$

The above equation simplifies to the required result.

Hence the proof.  $\square$

**PROPOSITION 5** For an  $n$  investor chain with transmission time  $k$ ,

$$E(\text{Volume}(n, k)) = \frac{n^2 - n - 4nk^2 + 2kn^2 - 4kn + 2k^3 + 3k^2 + k}{n^2}$$

*Proof* Given proposition 4 and  $E(\text{Volume}(n, 0))$  from (1), for an  $n$  investor chain network with transmission time  $k$ , we have the following:

$$E(\text{Volume}(n, k)) = E(\text{Volume}(n, 0)) + \sum_{t=1}^{k-1} (2n^2 - 8nt - 8n + 6t^2 + 12t + 6)$$

From (1), we have  $E(\text{Volume}(n, 0)) = (n^2 - n)/n^2$ , which implies

$$\begin{aligned} E(\text{Volume}(n, k)) &= \frac{n^2 - n}{n^2} + \frac{(2n^2 - 8n + 6)k + k(k-1)(2k-1) + 0.5(12-8n)(k-1)k}{n^2} \end{aligned}$$

It can be verified that this simplifies to the statement in proposition 5. Hence the result.  $\square$