<h1 style="text-align:center">Data Appendix</h1>

This study utilizes six datasets: *Directory of Directors*, *Who's Who*, IRS Income Tax data, *Poor's Register of Directors*, Citibank officers and directors data from Citibank's *Annual Reports*, and Citibank shareholder data.

The *Directory of Directors* provides a comprehensive list of business executives in New York City, defined to include those who hold directorships or trusteeships in at least one firm. Formally known as the *Directory of Directors in the City of New York*, it is a biannual publication containing the individual's work address, affiliated firms, and positions held in each firm. We use the 1929-1930 volume to construct the board network.

The *Poor's Register of Directors* lists nationally prominent business leaders. Formally called the *Poor's Register of Directors of the United States*, it contains each individual's work and residential address, affiliated firms and positions held in each firm. We use the 1928 volume entries for New York City boroughs and match it with the *Directory of Directors* data.

Who's Who contains a list of nationally prominent individuals across all fields, not just business executives. Individuals are selected based on their accomplishments and official positions. Formally called the *Who's Who in America*, it is a biannual publication that contains a short biography on each individual as well as other information such as social club affiliation, religious denomination, and political affiliation. We use the volume for 1928-1929 to match with the *Directory of Directors* data.

The IRS Income Tax data list wealthy individuals living in the New York City metropolitan area. The dataset contains 3-5% of those who filed individual tax-returns in

the New York City metropolitan region. It is available due to a one-time tax publicity

clause in the 1925 National Revenue Act, which made all income tax returns public for

only 1924. In September 1925, the *New York Times* published income tax information for

individuals whose income tax payments exceeded $500 and for other select prominent

individuals with lower payments. The dataset contains the name, address, and tax payment

for each individual. The dataset is matched with the *Directory of Directors* data.

The Citibank officer and director data covers high-ranking personnel serving at

Citibank or its subsidiaries. Each year, Citibank published an *Annual Report* that listed

high-ranking officers and directors and their exact positions. The list mostly included high-

ranking officers, but sometimes included branch managers and advisory committee

members. We include all listed individuals because some branch managers hold

directorship positions in other firms. We use the *Annual Reports* for December 1922, 1924,

1926, 1928 and match each of them with the *Directory of Directors* data.

The Citibank shareholder data include all shareholders who exercised votes at the

annual meeting for Citibank. The Citibank archive granted us access to the original voting

records for the annual shareholder meeting. The records contain the name, address, and the

number of shares exercised by each individual. We use the shareholder records for January

1925, 1927, and 1929 and match them with the *Directory of Directors* data.

Matching

We match the five datasets (*Poor's, Who's Who*, IRS Income Tax data, Citibank

lists of officers and directors, and the lists of Citibank shareholders) with the *Directory of*

Directors data. We match individuals from different datasets using name similarity. We declare two individuals to be identical people if their names (1) match perfectly or (2) have the same first and middle name initials and the last names match perfectly. All matches that involve different or ambiguous name similarities were verified using other information such as address, occupation, the number of shares owned and firm affiliation.

We match on name similarity for three reasons. First, the benefit of employing a less stringent matching procedure outweighs the costs. By employing a less stringent matching algorithm, we can expect to create a complete network map without omitting anyone. If we employ a more stringent matching procedure, we may have a perfect matching scheme, but risk omitting any individuals and creating an incomplete business network. Second, each dataset already filters individuals based on wealth and prominence. The possibility of any individual being matched to an identically named but different person, mainly one of lower prominence and wealth, is not very high. Third, a substantial number of individuals listed in our datasets have relatively uncommon names. This peculiarity makes mismatching individuals less likely.

Poor's-NYC Directors matching

Each matched entry is marked initially in three ways: 1 (certain matches), 2 (uncertain matches), 3 (possibly omitted matches). For any ambiguous matches, we use address and firm affiliation to verify each of them.

Many certain matches involve identical names and addresses. Uncertain matches involve individuals whose names are not "similar" according to our definition. Although

these matches do not have corresponding addresses and firm affiliations, there remains a reasonable level of confidence that the two individuals are the same. Typically, these matches involve individuals who work in similar industries or regions.

Possibly omitted matches involve nationally prominent directors who have been omitted from the Poor's register for whatever reason. We introduce these matches because the Poor's register explicitly states that it excludes prominent individuals who refuse inclusion. We are able to ascertain their level of prominence because they are listed in *Who's Who* as businessmen. Since the *Who's Who* has a higher selection criterion than the Poor's register, we assume these individuals are omitted from the Poor's register.

Out of 43,030 individuals listed in the *Directory of Directors*, 12,873 individuals are matched with an individual present in Poor's. Among those matched individuals, we are certain about 99% (12,819) of the matches. Only 40 and 14 matches are classified as uncertain and possibly omitted, respectively.

Who's Who-NYC Directors matching

Each matched entry is marked in two ways: 1 (certain matches), 2 (uncertain matches). All ambiguous matches are verified using address, occupation, and firm affiliations.

Certain matches usually involve similarly named individuals who engage in occupations typical to business executives (e.g. manufacturer, banker, lawyer etc.). Uncertain matches involve individuals in *Who's Who* with occupations in which directorships may have added little value to their accomplishments (e.g. sociologist,

congressman), and therefore, may have been omitted from their biographies. 1,328 individuals were cross-listed in both data sources and we are certain about 99% (1,309) of the matches. Only 19 are uncertain.

IRS Tax-NYC Directors matching

Each matched entry is marked as 1 (certain matches), 2 (uncertain and unverifiable matches). Any matches that involve common names are verified by using the federal census and city directories. Roughly half of the certain matches have identical addresses. Uncertain matches involve individuals who have similar names for whom we are unable to verify a match. The matching process for the IRS Tax data is a bit trickier than for other datasets for three reasons. First, the IRS Tax data only provides name, address and income tax payment. Second, the IRS Tax data include more abbreviations than in other datasets including one letter first name initials and two letter first and middle initials. As a result, we put greater emphasis on verifying each match whenever there is reason to question its accuracy.

Out of 43,030 individuals listed in the *Directory of Directors*, we have 7,682 matches with the IRS data set for individuals with listed addresses from the New York City metropolitan area. In this case, and elsewhere in our study, we employ the 1930 Census definition of locations that are considered part of the New York City Metropolitan Statistical Area. Among these matches, we are certain about 77% (5,917) of them. We are uncertain about 23% (1,765) of the matches. We assume that all uncertain matches are not matches. Alternatively, we test whether treating uncertain matches in the same fashion as

certain matches or dropping uncertain matches altogether affects results and find that all three versions yield nearly identical estimates.

For income tax payments, we typically assign the income tax payment for each matched individual. We assume that all married couples file jointly unless we find records for both individuals. For couples that file separately, we add the income tax payments for both individuals and assign it to the individual that is listed in the *Directory of Directors*. For individuals with multiple tax return records, we take the average of the payments. For individuals who file taxes both individually and through an estate or trust fund, we add the tax payment from the estate or trust fund to the individual income tax payment. Overall, these special cases account for only 57 of the certain matches.

Citibank Shareholder-NYC Directors Matching

Each matched entry is marked as 1 (certain matches), 2 (uncertain matches), 3 (corporate matches). Any ambiguous matches are verified using other information sources such as the federal census records and city directories.

Most certain matches involve identical addresses. Uncertain matches involve individuals who have similar names. However, these matches involve common names and we cannot verify these matches using other sources of information such as Ancestry.com, and phone directories etc. Corporation-related matches involve individuals who serve as officers or directors in firms that own Citibank shares.

The matching procedure for the shareholder data is more complex than for previous datasets because shareholders can be corporations or trust funds. Certain and uncertain

matches involve matching individuals to other individuals, trust funds, and joint ownership entities. There is no uncertainty in matching individuals to corporations. For individual-individual matches, we apply the same rules used for matching the three previous datasets. Additionally, we declare a match if the individual's spouse owns shares or individual holds a guardianship over a different shareholder. There are only two cases involving guardianships, both of which are over the individuals' children.

To match individuals to entities, we develop a set of rules. For trust funds and estates, we declare a match between the individual and the trust fund if the individual is either the designated beneficiary of the fund or a direct family member of the fund's founder (if the fund does not explicitly state its beneficiaries). For joint ownership entities, we declare a match between the individual and joint ownership entity if his name is explicitly stated in the entity's name. Most joint ownership entities involve married couples.

For individual-corporate entity matching, we declare a match if the individual is an officer or director in a firm that owns shares. We mark these matches separately because corporations may have different reasons for purchasing shares than individuals.

When calculating the total number of shares owned by each individual, we add the shares that the individual owns directly through his name and indirectly through other vehicles. For shares owned by trust funds, we assign the shares to each beneficiary in equal proportions. If the fund has no explicit beneficiary, we divide the shares among surviving direct family members. Family-related information is obtained via obituaries and wills. For

guardianships, we assume that the guardian owns all the shares. For joint ownership entities, we divide the shares only when the entity does not involve married couples.

For individual-corporate entity matching, we calculate the number of shares owned by each officer and director of the shareholding firm by dividing the number of shares owned by the number of officers and directors. We obtain the number of officers and directors for each firm by using the *Moody's Bank Manual*, *Moody's Industrial Manual*, and *Polk's Bankers Encyclopedia* for January 1929 (March 1929 for *Polk's*). These sources provide information for roughly half of all shareholding firms. The remaining firms consist mostly of small or medium-sized enterprises where the number of officers and directors is not large. For firms not listed in these sources, we rely on the *Directory of Directors*. Using only the *Directory of Directors* to derive the number of officers and directors would produce underreporting of the number because it only lists individuals who hold directorships in New York City-based firms. Therefore, those who only hold officer positions in New York City firms or hold directorship positions in firms based outside of New York City are excluded. To account for this bias, we use the number of the officers and directors derived from these data sources only when that number is greater than four. We follow this practice because the typical firm has at least five officers and directors (one president, one vice-president, one treasurer, one secretary, and one director). For the remaining firms, lacking any other data, we assume that the firm has five officers and directors. Alternatively, we experiment with assuming that the firm has six or seven officers and directors, but the results are very similar.

In total, we have 1,020 matches, including corporation-related matches. Among the matched individual shareholders, we are certain about 684 of them and uncertain about 19 of them. We have 385 corporation-related matches and some individuals own shares both individually or through corporate holdings. Table DA.1 summarizes the shareholding patterns for certain matches.

Year	# of new shareholders	Holding Period
1925	434 (total)	317 (1925-1929)
		86 (1925-1927)
		31 (1925 only)
1927	263	181 (1927-1929)
		82 (1927 only)
1929	323	323 (1929 only)
Total	1020	

Table DA.1: Summary of Shareholders

Variable Explanations

This section defines in detail each variable employed in the regressions.

Baseline variables:

- No. of affiliated firms: the number of firms the individual serves as an officer and/or director in.

- Poors indicator: a binary variable indicating whether the individual is listed in the Poor's National Register of Directors 1928 volume. Certain, uncertain, and possibly omitted matches are included.

49

- IRS Tax indicator: a binary variable indicating whether the individual is listed in the IRS Tax data published by the *New York Times* in September 1925. Certain and uncertain matches are included.

- IRS Tax Indicator * Tax Payment: IRS tax indicator interacted with the actual tax payment. Individuals not listed in the IRS income tax data are coded with 0.

- Who's Who indicator: a binary variable indicating whether the individual is listed in *Who's Who in America 1928-1929*. Certain and uncertain matches are included.

- High position indicator: a binary variable indicating whether the individual holds a high position in at least one of his affiliated firms. A high position is defined as simultaneously holding both officer and director positions in the same firm.

- Financial industry indicator: a binary variable indicating whether the individual is affiliated with financial services firms.

Network centrality variables:

We define each variable assuming a binary marking between any two individuals where 1 indicates that a network connection exists and 0 otherwise.

- In the network literature, the number of network connections an individual has is called degree. Formally, it is $d_i = \sum_{j=1}^{n} W_{ij}$ where W is the adjacency matrix representing the network. In this paper, the number of network connections for individual i is the number of individuals who serve as an officer and/or director in at least one firm with i. Although it intuitively explains the individual's influence within the network, it does not account for whether the individual is connected to

other influential individuals. Eigenvector network centrality and network

betweenness centrality account for this possibility.

- Eigenvector network centrality of individual i is an influence score that measures

   how well-connected i is to other influential individuals. Formally, it is the *ith* entry

   of the *x* where $Wx = \lambda x$ where λ is the largest eigenvalue and x is its

   corresponding eigenvector. The intuition is to consider the individual's influence

   in proportion to the sum of his connection's influence.

- The network betweenness centrality of individual i is an influence score that

   measures how "in-between" i is to other individuals. Formally, it is the proportion

   of shortest paths between all other individuals which pass through i. The shortest

   path between individuals j and k is defined as the path that passes through the

   smallest number of individuals and connects j and k. The intuition is that

   individuals who are along many shortest paths are "between" other individuals and

   exert a greater influence than those who are not.

Network influence variables

We measure network influence in terms of influence exerted by Citibank officers and

directors and other shareholders. We further distinguish each network influence into direct

and once-removed influences.

Influences from Citibank officers and directors

Percentage of Officer/Director Connections measures the direct network influences of

Citibank officers and directors. It is calculated by dividing the number of network

connections to Citibank officers and directors (who have served for at least two year) by the total number of network connections.

Percentage of Once-Removed Officer/Director Connections measures the indirect network influences of Citibank officers. It is calculated by dividing the number of once-removed network connections to Citibank officers and directors (who have served for at least two year) by the total number of once-removed network connections.

Influences from Citibank shareholders

Percentage of Shareholder Connections measures the direct network influences of shareholders. It is calculated by dividing the number of network connections to Citibank shareholders (who have not served as either a Citibank officer or director and have held shares for at least two years) by the total no. of network connection.

Percentage of Once-Removed Shareholder Connections measures the indirect network influences of shareholders. It is calculated by dividing the number of once-removed network connections to Citibank shareholders (who have not served as either a Citibank officer or director and have held shares for at least two years) by the total no. of once-removed network connection.