The Linchpins of a Modern Economy

Eric O'N. Fisher*
eric.on.fisher@gmail.com
Cal Poly, San Luis Obispo
Federal Reserve Bank of Cleveland

Fernando Vega-Redondo
vega@merlin.fae.ua.es
University of Alicante
Visiting Cornell University

11 December 2006

Abstract

We show that an input-output matrix is a directed weighted network. We develop and apply a measure of betweenness centrality based on random walks. This notion of centrality has an analogy as a current flow in an electrical circuit. We apply this measure to recent data from the major industrial economies. Wholesale and Retail Trade is the most central sector of many economies. The French and Italian economies have high degrees of centrality. In the United States in 2003, the most central sector was Real Estate and the least central was Hospital and Nursing and Residential Care Facilities.

*The authors will make available the data and Gauss code used in all the computations described in this paper.

1. Introduction

An input-output table A is an $\ell \times \ell$ matrix whose canonical element a_{ij} represents how many units of good i are used in the production of good j. Since the outputs of an industry are not closely related to its inputs, it is natural to think of such a table as a list of directed flows of economic activity. For example, in the United States in 1997, $18.2 billion of rubber and plastic products were used in the production of motor vehicles, but only $46 million of the output of the motor vehicle industry was used in the production of rubber and plastic products. Hence the motor vehicle industry may be severely affected by bottlenecks in the production of rubber and plastic products, but the converse is not true.

Leontief (1986) developed input-output accounting during the Second World War as an attempt to help identify strategic weaknesses in the German economy. The techniques of input-output accounting also have ready applications in economic planning. One of the most important reasons for collecting and constructing economic data at a disaggregated level is to identify the influence of sectors on national economic activity. Fischer Black (1987) hypothesized the business cycle might arise because of the propagation of shocks between the sectors of an economy, and Long and Plosser (1983) developed an elegant analysis of the United States economy based on this idea. Our analysis is an attempt to quantify which sectors are most central in this process.

We are the first to emphasize that an input output matrix is actually a directed network. Its vertices are ℓ different sectors in the economy; the typical level of aggregation used in empirical work allows for about 40 or about 350 different sectors. The edge between vertices i and j is the element a_{ij}, and it is important to emphasize immediately that this network is a weighted network. For example, in the United States in 1997, none of the output of the sector called building and repairing of ships and boats was used in the electricity, gas, and water supply sector, and it is natural to think that there is no directed edge connecting the former to the latter. But only \$1 million of the output of the building and repairing of ships and boats sector was used by the aircraft and spacecraft sector, and it is not appropriate to think that this particular link in the directed network is as strong as the one from building and repairing of ships and boats to the transport and storage sector, which used \$607 million of the former sector's output; in fact, the transport and storage sector uses 81% of all the output produced by shipbuilders and repairers.

Freeman (1977) introduced the notion of centrality in a network; the centrality of a node is the average number of shortest links between pairs of other nodes that pass through it. His definition is not adequate for an economic network in which edges may have different capacities. Freeman, Borgatti, and White (1991) describe a measure of flow for weighted networks that is based upon a maximum capacity of flows between nodes. Their measure ignores the possibility of "parallel processing" whereby information might flow between nodes through many different channels. Addressing this deficiency forthrightly, Newman (2005) defined random-walk betweenness. Our measure builds upon his important work, and it is easy to calculate.

2. Defining Betweenness

The weighted directed edge from sector i to sector j is given by the canonical element a_{ij} of the $\ell \times \ell$ input-output matrix A. The degree of vertex i is the total output of that sector:

$$d_i = \sum_j a_{ij} .$$

Hence a sector that is a large share of the economy has a relatively high degree.

Consider a shock of one extra dollar of output to a source sector s that will work its way to a target sector t. The economic interpretation of this shock is the effect on a marginal increase in the output of the source sector and a simultaneous decrease in the output of the target sector. Such a shock maintains a constant aggregate value for all intermediate inputs produced in the economy. It represents a shift of economic activity from the target sector to the source sector.

We are interested in how this impulse to economic activity works its way through each of the ℓ sectors (nodes or vertices) in the economy. Let us be specific and consider a dollar increase in the source sector of building and repairing of ships and boats that is offset by a dollar decrease in the output of the target sector transport and storage. There will be an important direct link from the source to the target, but there is a secondary link that flows from the source through the construction sector and then into the target sector. The dollar decrease in the output of the target sector will have important primary, secondary, and tertiary effects on the outputs of all the sectors in the economy. These effects will flow through each of the nodes of the network, and we would like to know how central each of these nodes is in the overall pattern of changes in economic activity that follows from this particular impulse.

Newman (2005) uses a physical analogy for an adjacency matrix that consists only of zeros and one. But his analogy can be extended to a directed weighted network. Interpreting the impulse as a flow of current between the source and target nodes, Newman shows that Kirchoff's law of current conservation implies that, for a given impulse (s,t), the voltage across each vertex satisfies this system of equations:

$$[D-A]v(s,t) = f(s,t) \tag{1}$$

where $D = diag(d_1,...,d_\ell)$, A is again the input output matrix, and $v(s,t)$ is the $\ell \times 1$ vector of voltages at the nodes in the network. The $\ell \times 1$ vector $f(s,t)$ has elements defined by:

$$f_i(s,t) = \begin{cases} 1 \ if \ \ i = s \\ -1 \ if \ \ i = t \\ 0 \ otherwise \end{cases}$$

The matrix $[D-A]$ is singular because the $\ell \times 1$ unit vector is an eigenvector whose eigenvalue is zero. Hence we use the Moore Penrose pseudo-inverse to find the minimum norm solution to (1):

$$v(s,t) = [D-A]^+ f(s,t)\,. \tag{2}$$

Equation (2) defines the voltage at each of the nodes in the economy for a given impulse (s,t). Let $w_i(s,t) = | v(s,t) - \underset{\ell \times 1}{1} v_i(s,t) |$, where $\underset{\ell \times 1}{1}$ is the conformable unit vector. Newman explains that the current flow through each node proportional to the absolute values of current flowing along the edges of that vertex:

$$I_i(s,t) = Aw_i(s,t)$$

for $i \neq s,t$. In the case where $i = s$ or $i = t$, the current flow is necessarily one unit, and we may write:

$$I_s(s,t) = I_t(s,t) = 1\ .$$

Finally, the definition of betweenness is the average current flow across all pairs of sources and targets:

$$b_i = \sum_s \sum_{t \neq s} I_i(s,t) / \ell(\ell - 1) \,. \qquad (3)$$

where our formulae differ slightly from those in Newman (2005) because we are considering a directed network.

The use of the Moore Penrose pseudo-inverse is an important generalization of Newman's definition of random walk betweenness. First, Newman's measure depends upon the particular computational procedure used. The system of linear equations (1) actually has a redundant equation. The physics of the problem is that Kirchoff's law actually defines only relative voltages; this fact is akin to the notion that a price system needs a numeraire. Newman suggests excluding any arbitrary sector in order to solve the system of linear equations in (1), but his measure of betweenness then depends upon the sector that is excluded. Hence betweenness measures my not be invariant with respect to the ordering of the components of the network. The Moore Penrose pseudo-inverse uses information from every component in the network. Thus our measure can be used to compare betweenness measures across networks with an equal number of nodes.

Also, Newman's measure will not work for a network that is not one connected component, but our always will. This is an important practical and theoretical consideration in the analysis of economic data, where some sectors report mo economic activity because of disparate national conventions for the aggregation of economic data. This fact may be a slight disadvantage for our measure because it will compute a betweenness measure even for a single node that is its own trivially connected component.

3. Random-walk Betweenness

Newman (2005) gives a nice intuition for his definition of betweenness. He imagines sending a message from the source node to the target. The message hits an intermediate vertex and then randomly chooses an edge out of the vertex to continue on to its eventual target. In his formulation, the probabilities assigned to the outgoing edges are all equal, but our definition allows for edges with greater economic activity to be chosen with higher probability. Hence there is a much greater likelihood for an impulse in economic activity that enters into the shipbuilding sector to exit out into the transport and storage sector rather than into the mining and quarrying sector.

It is perhaps appropriate for us to illustrate the usefulness of our measure of betweenness by a series of simple examples. In these examples and in all our subsequent empirical work, before computing our betweenness measure, we normalize the adjacency matrices so that the sum of their ℓ^2 elements is unity. The economic intuition is that we are thinking of the total value of all intermediate inputs in the economy as one unit; in other words, we are looking at probability distributions considered as directed networks in which an edge has a large value if its share of economic activity is high. Hence our empirical measure of betweenness does not depend upon the absolute size of the underlying economy.

Consider first the following adjacency matrix:

$$A_1 = \begin{bmatrix} 1 & 1 & 1 & & & & \\ 1 & 1 & 1 & & & 0 & \\ 1 & 1 & 1 & & & & \\ & & & 1 & & & \\ & & & & 1 & 1 & 1 \\ & 0 & & & 1 & 1 & 1 \\ & & & & 1 & 1 & 1 \end{bmatrix}$$

In this case, the economy consists of three completely connected components: (1) the first three sectors; (2) the singleton consisting of the fourth sector; and (3) the last three sectors. Every edge has equal flow capacity. Our measure of betweenness is: $b_1 = (.44,.44,.44,.29,.44,.44,.44)'$. The fourth sector has a low measure of betweenness because it is connected to neither component of The capacity of the fourth sector, has no effect on b_1 because it is its own component.

Consider second this adjacency matrix:

$$A_2 = \begin{bmatrix} 1 & 1 & 1 & & & & \\ 1 & 1 & 1 & & & 0 & \\ 1 & 1 & 1 & 1 & & & \\ & & 1 & 1 & 1 & & \\ & & & 1 & 1 & 1 & 1 \\ & 0 & & & 1 & 1 & 1 \\ & & & & 1 & 1 & 1 \end{bmatrix}$$

In this economy, economic activity can flow from the first connected component into the second one only through the fourth sector. Now the corresponding measure of betweenness is $b_2 = (.44,.44,1.08,1.14,1.08,.44,.44)'$ The third and fifth sectors have above-average connectedness because each is the only egress from its component into the rest of the economy. Now the economy consists of one connected component.

What happens if we increase the capacity of the linchpin fourth sector? Consider this weighted matrix:

$$A_3 = \begin{bmatrix} 1 & 1 & 1 & & & & \\ 1 & 1 & 1 & & & 0 & \\ 1 & 1 & 1 & 1 & & & \\ & & 10 & 10 & 10 & & \\ & & & 1 & 1 & 1 & 1 \\ & 0 & & & 1 & 1 & 1 \\ & & & & 1 & 1 & 1 \end{bmatrix}$$

Now our measure of betweenness is $b_3 = (.47,.47,1.12,8.9,1.12,.47,.47)'$. The high carrying capacity of the fourth sector makes it an important intermediary for the economy, but it also increases the measure of betweenness for all the other sectors. In this sense, the economy in the third example is more closely integrated than the first two.

What happens if we make the fourth sector a sink for the economy? In this case, all economic activity eventually leads into the fourth sector, but its outputs are used nowhere else in the economy. The adjacency matrix is

$$A_4 = \begin{bmatrix} 1 & 1 & 1 & & & & \\ 1 & 1 & 1 & & & 0 & \\ 1 & 1 & 1 & 1 & & & \\ & & & 10 & & & \\ & & & 1 & 1 & 1 & 1 \\ & 0 & & & 1 & 1 & 1 \\ & & & & 1 & 1 & 1 \end{bmatrix},$$

and the measure or betweenness is $b_4 = (.44,.44,.1.08,.29,1.08,.44,.44)'$. In this case, the largest sector in the economy has the lowest betweenness because it absorbs the economic activity of all the other sectors. The two sectors that are bridges from the connected subcomponents into the fourth sector have the largest betweenness because all economic activity must eventually flow though one of them.

We complete our series of examples by considering a completely connected economy where all the sectors have equal capacity. In this case, our measure of betweenness is $b_5 = (.49,.49,.49,.49,.49,.49,.49)'$. If the fourth sector has a has capacity ten times larger than all the others and the economy is still one completely connected component, then our measure of betweenness is $b_6 = (.46,.46,.46,2.33,.46,.46,.46)'$. Thus having a large capacity increases a sector's betweenness, even when the economy is one completely connected component.

4. Measures of Random-Walk Betweenness for some Modern Economies

We use two sources of data. The first is the standard cross-country comparison of input-output matrices for a wide array of mostly advanced industrial countries; we chose it because it permits fairly accurate cross country comparisons. The second is the most current data on disaggregated macroeconomic activity in the United States.

The first source is the OECD Stan Database that is found at: http://www.oecd.org/sti/stan/ . These data are constructed from national sources, and the OECD tries to make the cross-country comparisons as consistent as possible; for example, each table consists of the same $\ell = 40$ sectors. The data are input-output accounts in local currency from middle of the last decade. The list of the twenty countries covered is: Australia, Brazil, Canada, China, Czech Republic, Denmark, Finland, France, Germany, Greece, Hungary, Italy, Japan, Korea, Netherlands, Norway, Poland, Spain, United Kingdom, and United States. This list of countries is a nice mix of economies, including some of the most advanced countries in the world, some formerly centrally planned economies, a few upper middle income countries, and a developing economy. The Appendix lists the dates at which these economies' data were sampled.

The second source is the United States Department of Commerce's Bureau of Economics Analysis's input output accounts for the United States. These data are available here: http://bea.gov/bea/dn2/i-o_annual.htm. They describe the United States economy in 2003, and they are disaggregated into $\ell = 61$ different sectors. Thus they are more current and slightly more detailed than the data from the OECD. Their units are dollar values, not physical inputs.

We begin the empirical analysis by describing the two sectors in each economy that have the highest betweenness. Table 1 reports our statistics. The main conclusion from this comparison across countries is that the same two or three sectors are most central in an advanced industrial economy. These sectors are: (1) wholesale and retail trade; (2) business services; and (3) finance and insurance. A sector's betweenness is correlated with the size of its gross output, but the correlation is not perfect. It seems that a sector has a high betweenness measure if it is large and if it connects with many other important sectors in the economy. It is interesting that China and Korea, the least developed countries in Table 1, measure a high betweenness for the machinery sector. This fact may be a vestige of the transition from a manufacturing to a service economy. The Czech Republic and Poland, two of the economies from the former Soviet bloc, measure a high betweenness for utilities. We find it reassuring that our measure captures high betweenness for similar sectors for economies that are roughly comparable.

Table 1: The Most Central Sectors in Several Economies in the 1990's

Country	Most Central	Second Most Central
Australia	Wholesale and Retail Trade; Repairs	Other Business Activities
Brazil	Finance, Insurance	Wholesale and Retail Trade; Repairs
Canada	Finance, Insurance	Construction
China	Wholesale and Retail Trade; Repairs	Machinery and Equipment, N.E.C.
Czech Republic	Wholesale and Retail Trade; Repairs	Electricity, Gas, and Water Supply
Denmark	Wholesale and Retail Trade; Repairs	Other Business Activities
Finland	Other Business Activities	Wholesale and Retail Trade; Repairs
France	Other Business Activities	Finance, Insurance
Germany	Other Business Activities	Wholesale and Retail Trade; Repairs
Greece	Wholesale and Retail Trade; Repairs	Finance, Insurance
Hungary	Wholesale and Retail Trade; Repairs	Chemicals, Excluding Pharmaceuticals
Italy	Finance, Insurance	Other Business Activities
Japan	Wholesale and Retail Trade; Repairs	Other Business Activities
Korea	Real Estate Activities	Machinery and Equipment, N.E.C.
Netherlands	Other Business Activities	Wholesale and Retail Trade; Repairs
Norway	Wholesale and Retail Trade; Repairs	Finance, Insurance
Poland	Wholesale and Retail Trade; Repairs	Electricity, Gas, and Water Supply
Spain	Finance and Insurance	Other Business Activities
United Kingdom	Other Business Activities	Finance, Insurance
United States	Wholesale and Retail Trade; Repairs	Other Business Activities

We did not develop an overall measure of centrality in the second section, but it is easy enough to report several summary statistics for these twenty networks. The usual statistic that describes centrality is the average of the betweenness of all the nodes. We found it convenient to report the average betweenness and its standard deviation across the 41 sectors in each economy. These statistics are in Table 2.

Table 2: Cross-Country Measures of Network Centrality				
Country	Average	Standard Deviation	Maximum	Minimum
Australia	0.19	0.15	0.73	0.05
Brazil	0.26	0.22	0.76	0.05
Canada	0.18	0.15	0.75	0.05
China	0.25	0.18	0.68	0.05
Czech Republic	0.18	0.11	0.50	0.05
Denmark	0.18	0.13	0.66	0.05
Finland	0.17	0.12	0.60	0.05
France	0.67	0.65	3.71	0.05
Germany	0.19	0.14	0.68	0.05
Greece	0.22	0.21	0.90	0.05
Hungary	0.19	0.14	0.64	0.05
Italy	2.72	4.38	25.07	0.05
Japan	0.66	0.73	4.13	0.05
Korea	0.43	0.39	1.68	0.05
Netherlands	0.19	0.16	0.78	0.05
Norway	0.16	0.12	0.72	0.05
Poland	0.21	0.20	1.22	0.05
Spain	0.20	0.14	0.61	0.05
United Kingdom	0.19	0.14	0.77	0.05
United States	0.23	0.20	0.96	0.05

The Anglo-Saxon economies generally have low centrality, but France and Italy's economies seem to have a high centrality. After all, *dirigisme* is a French word! We are puzzled by the extraordinary betweenness of the finance and insurance sector in Italy; it has the second largest gross output of all sectors, but there are few such large numbers that appear in our statistics.

The minima reported in Table 2 are an artifact of our use of the pseudo-inverse in the calculations. In many of these economies, no economic activity is reported in several sectors, and the solution of minimum norm calculates a notional betweenness for such a null sector. The advantage of our measure is that it is computationally tractable, but its disadvantage is that it assigns a betweenness measure to a singleton component, even a trivial one that has no carrying capacity.

Another shortcoming of our measure is that it is only as good as the aggregation scheme that is used in constructing the data. It is instructive to delve more deeply into the concordance that the OECD uses to construct these data. The sectors that comprise "Other Business Activities" are described in the Appendix. This amalgam includes information services and administrative services that are central in the operation of most businesses in a modern economy. For example, it includes advertising, legal services, and services to other dwellings and buildings. This level of aggregation may be too coarse to be of much analytical use.

We also calculated the betweenness of the United States economy in 2003. These data are slightly more recent, and they are also more disaggregated, consisting of 61 different sectors. Table 3 reports the three most central and the three most peripheral sectors in that economy.

Table 3: The Most Central and the Most Isolated Sectors in the United States Economy in 2003		
Name	Betweenness	Rank
Real estate	107.5	1
Administrative and support services	57.8	2
Miscellaneous professional, scientific and technical services	51.1	3
Chemical products	37.6	4
Wholesale trade	32.9	5
. . .	. . .	. . .
Water transportation	0.3	57
Funds, trusts, and other financial vehicles	0.1	58
Support activities for mining	0.1	59
Social assistance	0.0	60
Hospitals and nursing and residential care facilities	0.0	61

The real estate sector has high betweenness and so does administrative and support services. The sectors with low betweenness also seem intuitive, especially for a political economy that takes pride in having reduced social assistance in the last decade.

5. Conclusion

We are the first authors to emphasize that an input-output table can be interpreted as a weighted directed network. Economic activity flows from a sector to those that use its outputs. An edge from one sector to another has a large carrying capacity to the extent that a large share of gross economic activity in intermediate inputs flows from the former into the latter.

We have extended a measure of random-walk betweenness to incorporate three important practical considerations: (1) our measure can handle weighted and directed networks; (2) it can deal with networks that do not consist of a single connected component; and (3) it is computationally tractable. Our measure of betweenness will have ready applications in many different kinds of economic, social, and physical networks.

The application of random-walk betweenness to economic data is completely novel. We hope that our first step inspires other researchers to expand this line of inquiry. Our measures are tractable, and they can be applied to networks in international trade, measures of job flows between sectors, and many other empirical aspects of economic activity.

We applied our measure of betweenness to data from a wide array of modern economies in the last decade. Our analysis shows that wholesale and retail trade and other business activities are generally the two most central sectors in an economy that has moved beyond the stage of development where manufacturing is a crucial part of economic activity. The strategic implications of this finding may be important. For example, the surge in suicide bombings in Israel in the early part of this decade had an important macroeconomic effect on the Israeli economy, perhaps in part because these attacks often targeted the retail sector. Another implication is that a shock to the real estate sector in the United States may well have important macroeconomic effects.

References

Black, Fischer. *Business Cycles and Equilibrium*, New York: Basil Blackwell, 1987.

Freeman, Linton C. "Theoretical foundations for centrality measures," *Sociometry* 40 (1977), 35-41.

Freeman, Linton C., Stephen P. Borgatti, and Douglas R. White. "Centrality in valued Graphs: A measure of betweenness based on network flow." *Social Networks* 13 (1991), 141-54.

Leontief, Wassily. *Input-Output Economics* (2nd ed.), New York: Oxford University Press, 1986.

Long, John B. and Charles I. Plosser**.** **"**Real Business Cycles," *Journal of Political Economy* 91 (1983), 39-69.

Newman, M. E. J. "A measure of betweenness centrality based on random walks," *Social Networks* 27 (2005), 39-54.

Appendix

Sample Dates for the Input Output Matrices	
Country	Date
Australia	1994-95
Brazil	1996
Canada	1997
China	1997
Czech Republic	1995
Denmark	1997
Finland	1995
France	1995
Germany	1995
Greece	1994
Hungary	1998
Italy	1992
Japan	1997
Korea	1995
Netherlands	1998
Norway	1997
Poland	1995
Spain	1995
United Kingdom	1998
United States	1997

Aggregation Scheme for “Other Business Activities” OECD Concordance for the IO Accounts of the United States	
Number	Nomenclature
434	Royalties
439	Portrait photographic studios, and other miscellaneous personal services
444	Services to dwellings and other buildings
445	Personnel supply services
447	Detective and protective services
449	Photofinishing labs and commercial photography
450	Other business services
451	Management and public relations services
453	Advertising
454	Legal Service
455	Engineering, architectural, and surveying services
456	Accounting, auditing and bookkeeping, and miscellaneous services, n.e.c.