Graph Network Analysis of Idiosyncratic Supply-Side Shocks

Andrew Ma
andrew_c_ma@berkeley.edu
Area: Economic Graph Networks
11 May 2022

Abstract

The permeation of macroeconomic shocks through economic networks results in large macroeconomic fluctuations, whose effect can be measured through Input-Output relationships between industries. We first implement and validate the economic network model utilizing Leontief Inverses proposed by Acemoglu et. al (2015) before investigating the short run propagation of two externally induced industry-level supply shocks: drastic drops in mining industry and the effects of the 2020 Covid pandemic. In each case, we hope to see associated output decreases in strongly associated industries and attempt to connect our results to general macroeconomic intuition and theory – the results of preliminary tests conducted imply that transmissions of different shocks through economic network linkages have significant implications for the macroeconomy; these effects can be quantitively calculated and compared to arrive at concrete conclusions regarding the impact of idiosyncratic supply-side shocks.

Introduction

In a sentence, the objective of our project is to simulate and explore how supply-side shocks travel and amplify through the United States Economy through its representation as a graph network. This will be done by introducing pre-defined perturbations in the supply of industries we are interested in - note the spotlight on disruptions of the supply rather than demand – Acemoglu et.al (2015) have shown that demand-side shocks only result in upstream shocks not felt by consumers. Thus, studying supply-side shocks will provide more practical and interpretable results, as supply side shocks are ultimately felt by consumers.

Initially proposed by economists in the 1970's, the idea of examining small shocks from firms & disaggregated industries in relation to the overall economy was trivialized by American economist Robert Lucas's work on macroeconomic policymaking, where Lucas argues that the **consequences of idiosyncratic shocks are "washed away" through the law of large numbers** and irrelevant for serious economic research [1]. Recently, theoretical interest in this area has spiked due to the availability of computing resources for the large matric calculations involved; economists have discovered that the force of large numbers is not necessarily present as simulated shocks hitting vital supplier industries have resulted in aggregate and amplified fluctuations in the macroeconomy [1].

Motivation for this study is numerous – originally intended to provide a coherent and comprehensive picture of the nation's economy, I-O tables are the building blocks for economic accounts such as gross domestic product (GDP). One important use of I-O tables is the estimation of direct and indirect network effects that changes in industry supply & demand cause; I-O tables have found themselves as the framework for preparing the national business strategy & fiscal policy through simulated stress testing & system experimentation.

Data

To accomplish this task, we will be working with Input-Output tables over the period of 1997-2020 containing 71 industries. Specifically, we are interested in the Total Requirements Table, which shows **total output of a product (row) generated by a unit increase in the final demand of a product (column)**. Note that the Total Requirements Matrix here is analogous to $(I-A)^{-1}$, the Leontief Inverse, and can be constructed from the basic Use and Supply Tables. The first five rows of the Total Requirements Table alongside their Total Output Requirement (sum of requirements

Table 1. BEA Total Requirements Table, Industry-by-Industry (U.S. Economy, 2020)

Industries	**Farms**	**Forestry, Fishing, & Related Activities**	**Oil & Gas Extraction**	**Mining (except Oil & Gas)**	**…**
Farms	1.2340	0.023236	0.001630	0.002204	**…**
Forestry, Fishing, & Related Activities	0.072096	1.110935	0.000673	0.001968	**…**
Oil & Gas Extraction	0.016900	0.005253	1.126379	0.019905	**…**
Mining (except Oil & Gas)					
…	**…**	**…**	**…**	**…**	**…**
Total Industry Output Requirement	2. 400916	1.471436	2.324612	2.062204	**…**

from all industries) are shown below – the complete Total Requirements Matrix is found in **Appendix A and** a detailed derivation of the Leontief Inverse matrix from basic I-O Use/Supply Tables can be found in **Appendix D**.

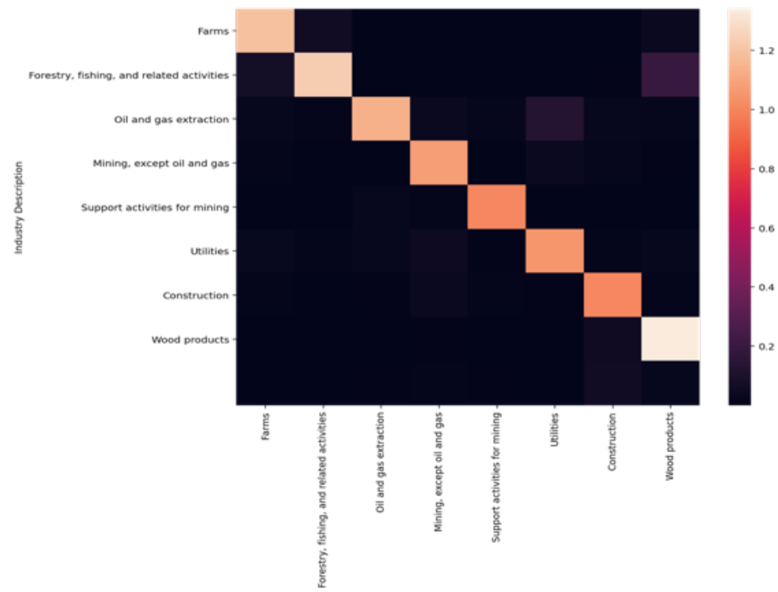
From the construction of the Total Requirements Table (shown in **Appendix D**), there are three main assumptions we must keep in mind:

1. **Homogeneity:** Each industry's output is produced using a unique set of inputs.
2. **Principle of Proportionality**: The ratio of each input to one unit of output remains constant over a wide range of output levels. That is, there are no economies of scale. For example, if the demand for a given product increases by 50 percent, all the inputs required for the product will also increase by 50 percent.
3. **Consistency:** Economic statistics are organized and presented in a uniform manner. The use of a common classification system enables users to effectively compare and analyze data across the broad spectrum of U.S. economic statistics.

Before moving onto the methodology of our study, let's explore two more interesting measures of our economy: cross correlation between industries & network centrality. With respect to the context of our problem, cross correlation is a measure of similarity among input-output relationships between industries; on the other hand, network centrality is calculated by $\frac{num\ edges}{total\ industries}$; i.e., a centrality score of 1 indicates that the industry is connected to every single other industry in the network.

From **Fig. 1** above where warmer colors represent higher levels of cross-correlation, that there is insignificant intra-industry cross correlation - this is good and what we expect because then we know there are no "hidden" shocks or effects we can't catch. **Figure 2.** depicts increasing centrality scores as warmer colors - from this network plot, we notice that the United States Economy has a high level of average centrality in the year of 2020 – perhaps this speaks to the consequences of the massive globalization wave of the current century. These exploratory observations are important to explore as they describe the intra-industry behavior of economic networks.

Figure 1. Cross Correlation Table (U.S. Economy, 2020)



Methodology

The model we will be implementing is directly referenced in Acemoglu et.al [1], and a short summary of it will be detailed here; consider a static perfectly competitive economy with n industries, and suppose that each industry $i = 1, \dots, n$ has a Cobb-Douglas production function of the form:

$$y_i = \exp(z_i) * l_i^{\alpha_i^l} * \prod_j x_{ij}{}^{\alpha_{ij}}$$

where x_{ij} represents the quantity of goods produced by industry j used as inputs by industry, l_i is labor, y_i is the total output of industry i, & z_i is the output of Hicks-neutral productivity shock (representing both technological and other factors affecting productivity). To ensure that the production function of each industry exhibits constant returns to scale, we have the following constraints:

$$\alpha_i^l + \sum_j \alpha_{ij} = 1$$

From the Cobb-Douglas productivity function, we see that the output of each industry is used as input for other industries or consumed in the final good sector. Incorporating the demand from other industries, the market-clearing condition for industry i can be written as:

$$y_i = c_i + \sum_j x_{ji} + G_i$$

where c_i is the final consumption of the output of industry i, G_i is the denotes government purchases of good i. We introduce government purchases to be able to model demand-side shocks in a simple fashion. The preference side of this economy is summarized by a representative household with a utility function:

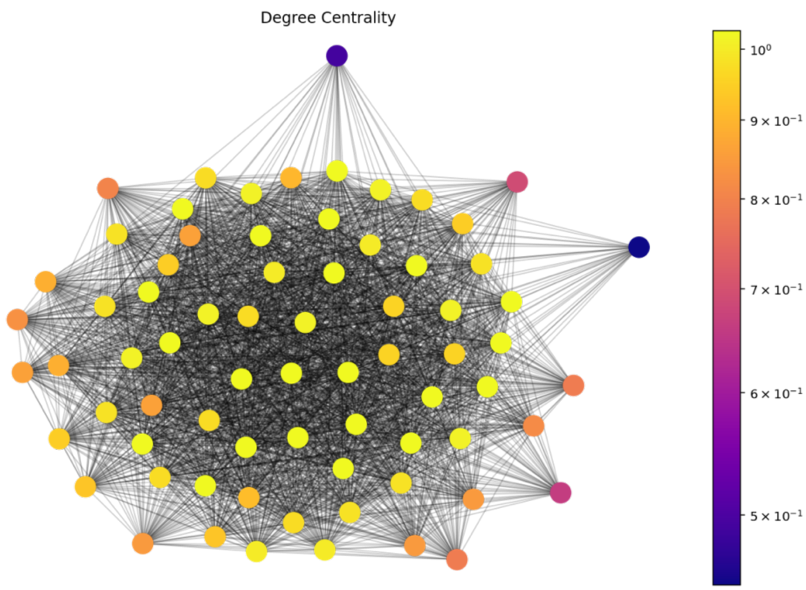
$$u(c_1, c_2, \dots, c_n, l) = \gamma(l) \prod_i c_i \beta_i$$

where $\beta_i \in (0,1)$ designates the weight of good i in the representative household's preferences (with the normalization β_i = 1) & γ(l) is a decreasing (differentiable) function capturing the behavior of labor supply. The government imposes a lump-sum tax, T, to finance its purchases. Denoting the price P of the output of industry i by p_i , this implies T = $p_i G_i$. Since its income comes only from labor, wl, the representative household's budget constraint can be written as

$$\sum_i p_i c_i = wl - T$$

We focus on the competitive equilibrium of this static economy, which is defined in the usual fashion, so that all firms maximize profits, and the representative household maximizes its utility, in

Figure 2. Degree Centrality of U.S. Industries (2020)



both cases taking all prices as given, and the market-clearing conditions for each good and labor are satisfied. The amount of government spending and taxes are taken as given in this competitive equilibrium. The Cobb-Douglas production function (Equation 1) combined with profit maximization, i

$$\frac{p_j x_{ij}}{p_i y_i} = a_{ij}$$

Let A denote the matrix of a_{ij}'s

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}$$

Then define $H \equiv (I - A)^{-1}$ as the Leontief inverse of the input-out matrix A. From the definitions above, we can derive the following proposition, whose complete proof can be found in **Appendix F.**

Proposition 1. The impact of sectoral productivity (supply-side) shocks on the output of sector i is

$$\boldsymbol{dlny = H * dz}$$

where H is the Total Requirements Matrix (Leontief inverse of A), Y is the output vector where y_i are output industries, and Z is the output productivity shock.

Finally, we will propose a methodology to induce and track higher order shocks, i.e., after one simulated shock, how do other interconnected industries react – we want to track the proportion of supply side shocks quantitively.

Algorithm Higher Order Supply-Side Shocks

1. Induce an initial shock into current year with 'dz_0'; **H** given by BEA requirements table
2. **Solve** Proposition 1 for $\mathbf{dlny_0}$
3. Take exponential of $\mathbf{dlny_0}$ to be dz_1
4. **Solve** Proposition 1 for $\mathbf{dlny_1}$ with H_{2020} & $\mathbf{dz_1}$
5. **Repeat** above processes for n-order shocks with $\mathbf{H_{2020}}$, $\mathbf{dz_n}$, and $\mathbf{dlny_n}$.

Table 2. log of total output change of industries due to 60% decrease in productivity (U.S. Economy, 2020)

Industry	log(Output change)
Primary Metals	-0.845561
Wholesale Trade	-0.122471
Mining, Except Oil & Gas	-0.042747
Fabricated Metal Products	-0.040744
Truck Transportation	-0.031992

Note: Output Change measured by change of Value Added

Model Validity Check

Before attempting to simulate the effects of the 2020 Covid Pandemic, we will present an example where a supply side shock is induced into the 'Primary Metals' industry in the year 2020; we artificially propagate this supply shock into the metals industry through altering the productivity vector dZ = [0, … , 0 , -0.6, 0, …, 0], where Z_{metals}= -0.6 is the **only** non-zero element; i.e., only the metals industry is affected by this 60% decrease in output productivity. From this preliminary experimentation, we hope to validate the results of our model through cross reference with equity pricing data; after applying **Proposition 1.** and with a shock order of 1, we arrive at the results:

1. The industries most affected are those we would automatically associate with metals – however, their absolute difference is hard to measure as we are on the log-scale.
2. Largest change to log Output is to the 'Primary Metals' Industry itself – this pattern seems consistent when applying this experiment to other industries such as Farms, Forestry, Automotive, etc.

In the example above, an induced supply-side shock into the metal industry represented by a sudden drop of output productivity by ~60%, resulted in largest log output losses in the associated industries of trade, mining, fabricated metal products, and truck transportation. These results align with our intuition: industries closely associated with metals would suffer most from unexpected decreases in metal supply.

Experimentation & Results

Before applying our pandemic shock, we must decide both on 'dz' and a choice for H. With respect to the Leontief Inverse for 2020, H, one could choose to utilize data directly given by the BEA or a weighted average of $\mathbf{H_{2019}}$, $\mathbf{H_{2020}}$, & $\mathbf{H_{2021}}$. After testing numerous combinations of weighted averages, we decided to simply utilize $\mathbf{H_{2020}}$ given by the BEA as the brunt of supply side shocks was seen in the year of 2020; moreover, $\mathbf{H_{2021}}$ is not complete as of the publication of this paper and thus not available for analysis.

Table 3. Choice of Leontief Inverse $(I-A)^{-1}$

Method	Equation	Pros	Cons
Weighted Average of ***H***	$\sum_j w_j H_j$	More likely to reflect long-term dynamics	Choice of weights is arbitrary
Individual ***H*** ***(2020)***	H_{2020}	More likely to reflect the Pandemic dynamics	May not capture long-term dynamics

Moreover, we wish to induce a shock that is true representative of the shock that underwent the U.S Economy during the Covid shutdown policies. Below are two proposed derivations of appropriate shock vectors for the pandemic:

Table 3. log of total output change of industries due to 60% decrease in productivity (U.S. Economy, 2020)

Equity-Derived Shock		BEA-Derived Shock	
Industry	**Change of 'Value Added'**	**Industry**	**Change of 'Value Added'**
Oil & Gas	0.1	Air Transportation	0.4
Wholesale Trade	0.06	Real Estate	0.34
Professional Services	0.04	Administrative Services	0.26
Fabricated Metal Products	0.03	Machinery	0.2
Truck Transportation	0.01	Construction	0.15

Note: Output Change measured by change of Value Added

Algorithm Imputation of pandemic 'dz' vector with **BEA Data**

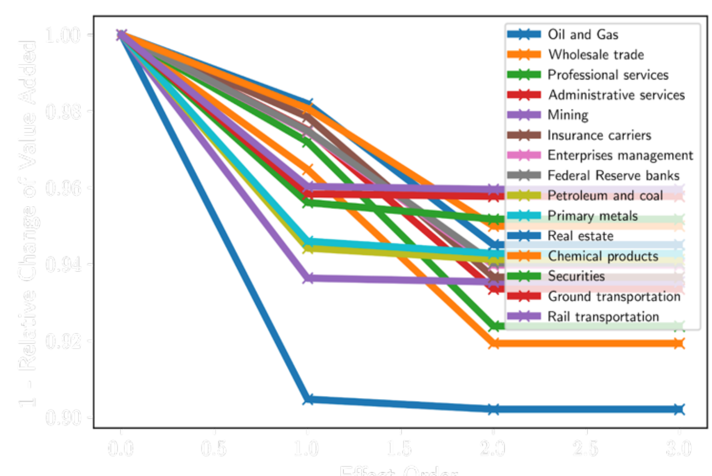
1. **Choose BEA quarterly industry output** data **D, window w, threshold ϵ,** and a **function f**
2. **Roll over** data D with the given window, threshold, and function looking for pandemic onset
3. Once found, **apply function f** to all industries and normalize, setting shock to 0 if above threshold
4. **Return** output as estimated **dz**

Algorithm Imputation of pandemic 'dz' vector with **Equity Data**

1. **Choose Equity Data D, window w, threshold ϵ, and a function f**
2. **Roll over** data D with the given window, threshold, and function looking for pandemic onset
3. Once found, **apply function f** to all industries and normalize, setting shock to 0 if above threshold
4. **Return** output as estimated **dz**

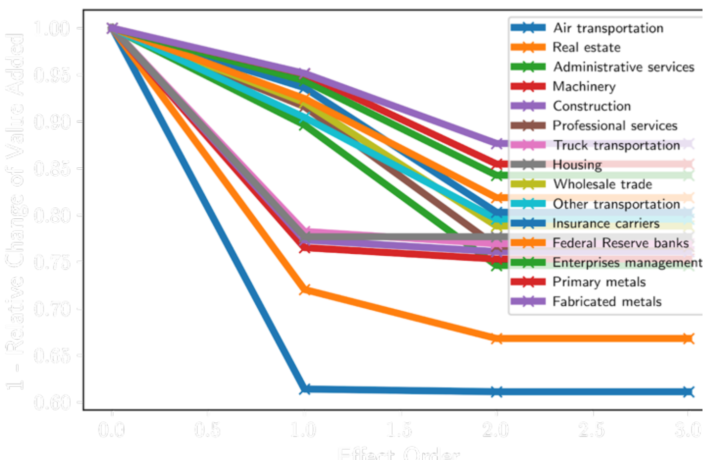
In combination with research already conducted by Acemoglu et.al (2015), the proposed methodology of economic shock simulation has yielded fruitful results in terms of interpretability and connection to macroeconomic knowledge. Note that the choice of shocks $\mathbf{dz}_{\mathbf{equities}}$ versus $\mathbf{dz}_{\mathbf{BEA}}$ requires careful consideration of which results make more sense. For example, the shock induced through equities-derived data, with a rolling window of 18 days, has affected the industries of 'Oil & Gas', 'Wholesale Trade', 'Professional Services', 'Metal Products', & 'Truck Transportation' the most, which we have seen replicated during the shocked industrial years of 2019-2021. Inf act, we are still recovering from this transportation & wholesale trade crisis in 2022 as we can see from the current backlog & supply chain issues – the negative industrial output of machinery, metals, and other wholesale products can also be seen with the skyrocketing price of lumber, catalytic converters, and rare metals which ceased production due to pandemic restrictions.

Figure 3. Equity-Derived Pandemic Shocks



On the contrary, we can re-run our model with BEA-indexed data instead to examine shocks through our macroeconomic graph network. **Figure** 4 below depicts the first industries affected by our supply side shock: 'Air Transportation', 'Real Estate', 'Machinery', and 'Construction'; moreover, each shocked industry suffers a heavier net output loss when compares to the losses of industries under the equities-derived shock. We suspect that the difference in industries and net industrial output losses to be a result of the differences in time indices for equities and BEA data; while equities data is collected daily and our shock vector was derived off a '18' day rolling basis, the BEA data is collected quarterly, resulting in a collection of shocks. In other words, the longer time between BEA's data collection results in larger shocks and output differences being captured in comparison to a equity-pricing methodology of deriving a 'dz' vector.

Figure 4. BEA-Derived Pandemic Shock



Finally, one may cross-reference the above results with an external measure of industrial production stagnation to validate the conclusions of our model:

Table 4. BEA & Equity Index Outputs

Shock Order	BEA Industrial Output	Equity Index Industrial Output	Industrial Output Max Drawdown
0	-0%	-0%	
1	-3.24%	-7.77%	-10.3%
2	-3.38%	-11.78%	

From **Table 4** above, we can see an equity-backed pandemic shock better simulates the true historical industrial downturn felt in 2020; we see a close approximation of the maximum industrial output, implying that our Equity index provides an accurate historical simulation of the 2020 Covid effects on the United States Economy.

Conclusion & Next Steps

Before concluding our discission of economic network analysis, let's consider limitation & challenges associated with future work. The main challenges associated with this project is in the time - aspect of our results; while the affected industries align with general economic intuition, how does one time these shocks? Will these predicted industries crumble in a week, or a month? Building on these shocks, how long should we expect to wait until second order shocks hit? Third order shocks? These questions must be explored further to determine an accurate timetable of our predicted shocks – with an accurate estimate of industry collapses, companies and financial institutions may better prepare fallback plans or hedges in a time-indexed manner.

Nevertheless, this model looks promising as an addition to other concrete economic methods when conserving macroeconomic conditions. Moving forward, one may choose to expand the Acemoglu model to an 'open' form model; this includes taxes, government consumption as industries (defense, public goods, etc.), imports, & exports. The inclusion of these economic variables forms an 'open' model which may more accurately capture the flow of goods and money in our graph network economy.

References

[1] Acemoglu, D., Akcigit, U., & Kerr, W. (2016). Networks and the macroeconomy: An empirical exploration. *NBER Macroeconomics Annual*, *30*(1), 273–335. https://doi.org/10.1086/685961

[2] Bureau of Economic Analysis. (n.d.). *Total Requirements Derivation*. bea.gov. Retrieved March 17, 2022, from https://apps.bea.gov/industry/pdf/TotalRequirementsDerivation.pdf

[3] Choi, Jason and Forester, Andrew, (2017), The Changing Input-Output Network Structure of the U.S. Economy, Economic Review, (Q II), 23-49

[4] Karen J. Horowitz & Mark A. Planting, 2006. "**Concepts and Methods of the U.S. Input-Output Accounts**," BEA Papers 0066, Bureau of Economic Analysis.

[5] Sekhon, Rupinder, and Roberta Bloom.
Applications – Leontief Models. De Anza College,
4 Sept. 2021,
https://math.libretexts.org/@go/page/37851.

Appendix

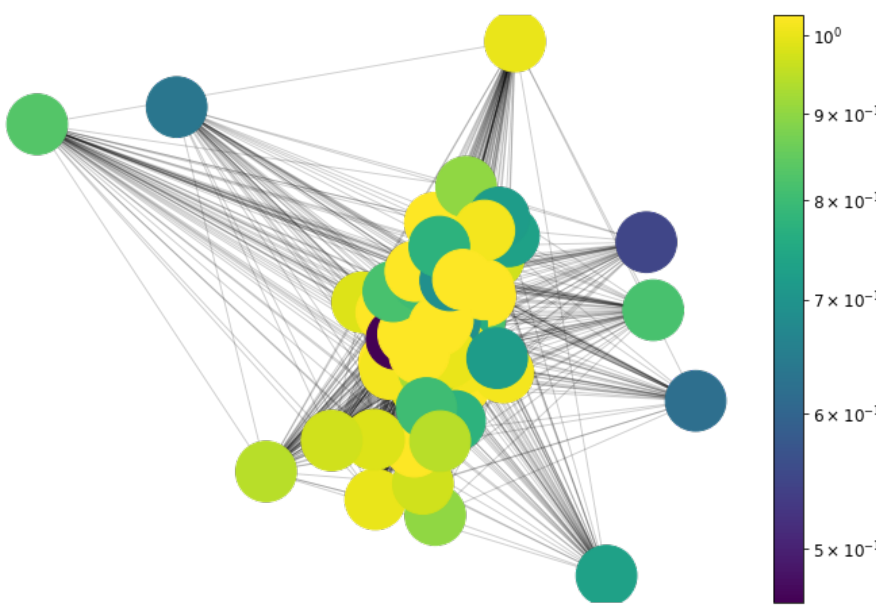
Appendix A: Total Requirements Table

The total size of a single requirements table is ~20 pages when exported to a Latex of PDF document; to this end, I have attached a dynamic copy of this table, viewable via this link. Data is also attached to the end of our presentation and formal replication studies may be conducted with the full dataset contained in our GitHub repository.

Appendix B: Intra-Industry Cross Correlation Table

	Farms	Forestry	Mining	Mining, No Oil/Gas	Mining Support	Utilities	Construction
Farms	**1.23**	0.02	0.00	0.00	0.00	0.00	0.00
Forestry	0.07	**1.11**	0.00	0.00	0.00	0.00	0.00
Mining	0.02	0.01	**1.13**	0.02	0.01	0.06	0.00
Mining, No Oil/Gas	0.01	0.00	0.01	**1.08**	0.00	0.02	0.02
Mining Support	0.00	0.00	0.04	0.02	**1.01**	0.00	0.00
Utilities	0.03	0.00	0.03	0.04	0.01	**1.06**	0.01
Construction	0.02	0.00	0.02	0.03	0.01	0.02	**1.01**

Appendix C: Degree Centrality of U.S. Economy, 2020 (71 Industries)



Appendix D: Derivation of Total Requirements Table from Make & Use Tables

From make and use tables, the following are defined:

- **^:** A symbol that, when placed over a vector, indicates a square matrix in which the elements of the vector appear on the main diagonal and zeros elsewhere.
- **q:** A column vector in which each entry shows the total amount of each commodity's output.
- **g:** A column vector in which each entry shows the total amount of each industry's output.
- **U: Intermediate portion of the use matrix** in which the column shows for a given industry the amount of each commodity it uses, including noncomparable imports and used and secondhand goods. This is a commodity-by-industry matrix.
- **V: Make matrix** in which the column shows for a given commodity the amount produced in each industry. This is an industry-by-commodity matrix. V has columns showing only zero entries for noncomparable imports and used and secondhand goods.
- **B: Direct input coefficients matrix** in which entries in each column show the amount of a commodity used by an industry per dollar of output of that industry.

This is a commodity- by-industry matrix, which can be created with the formula below.

$$\boldsymbol{B} = \boldsymbol{U}\hat{\boldsymbol{g}}^{-1} \; \mathbf{(1)}$$

- **D:** A matrix in which entries in each column show, for a given commodity, the proportion of the total output of that commodity produced in each industry. In the model, it is assumed that each commodity is produced by the various industries in fixed proportions.

This is an **industry-by-commodity matrix. D is also referred to as the market share matrix or transformation matrix.**

$$\boldsymbol{D} = \boldsymbol{V}\,\hat{\boldsymbol{q}}^{-1} \; \mathbf{(2)}$$

- $\boldsymbol{i}$**:** Unit (summation) vector containing only 1's.
- **I:** Identity matrix, where I = $\hat{\imath}$.
- **e:** A column vector in which each entry shows **the total final demand purchases for each commodity from the use table.**

From the above definitions, the following identities are derived:

$$\boldsymbol{q} = \boldsymbol{Ui} + \boldsymbol{e} \; \mathbf{(3)}$$

$$\boldsymbol{g} = \boldsymbol{Vi} \; (4)$$

The model expressed in equations **(1)** through **(4)** thus involves two constants (*B, D*) and five variables (*U, V, e, q, g*). The model solution is derived as follows: From **(1)** and **(3),** we derive:

$$\boldsymbol{q} = \boldsymbol{Bg} + \boldsymbol{e} \; \mathbf{(5)}$$

From **(2)** and **(4)**, we derive:

$$\boldsymbol{g} = \boldsymbol{Dq} \; \mathbf{(6)}$$

Substituting **(6)** into **(5)** and solving for q gives:

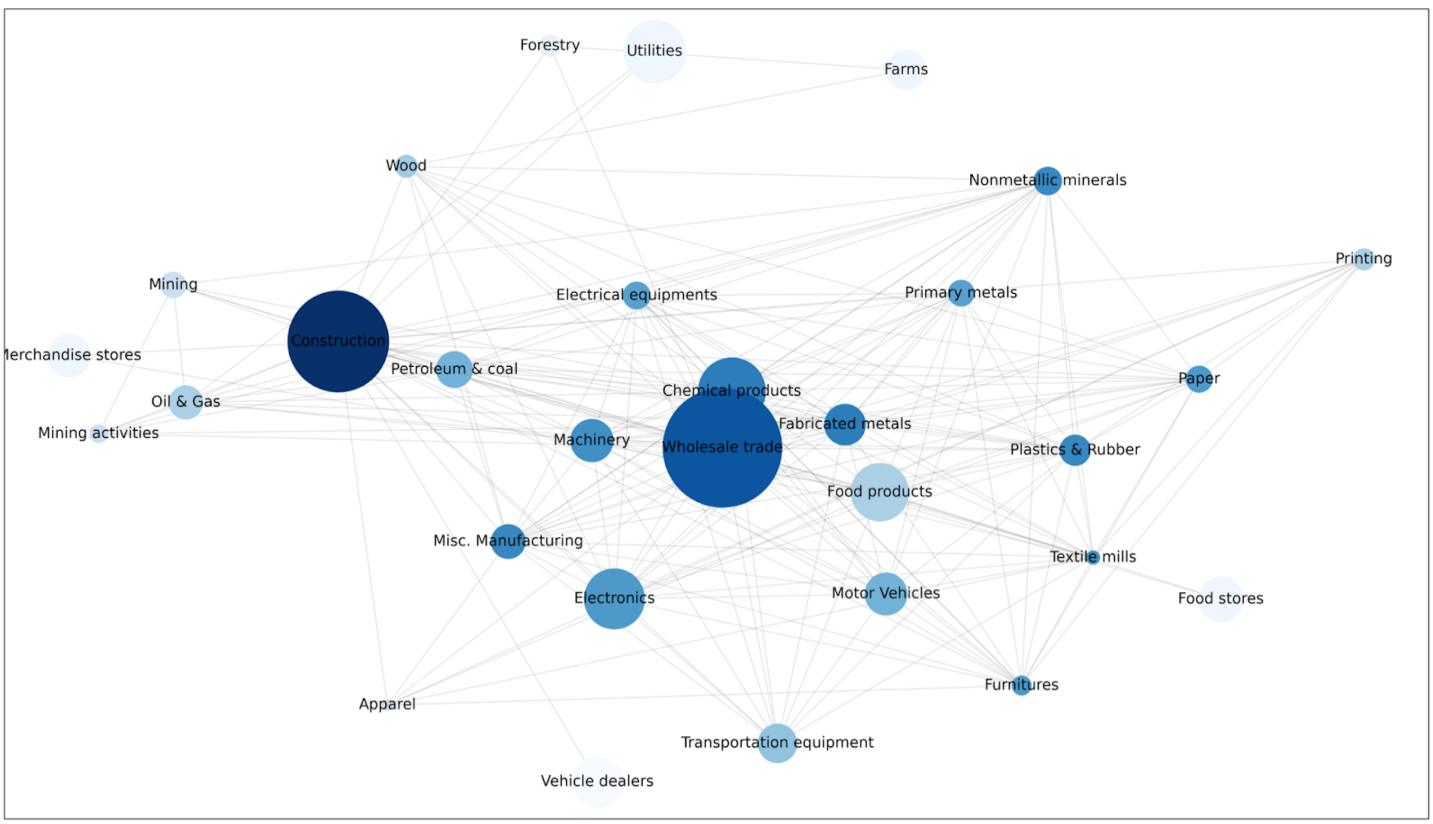
$$\boldsymbol{q = B(Dq) + e\ (I - BD)q = e}$$
$$\boldsymbol{q = (I - BD)^{-1}e} \quad \textbf{(7)}$$

The matrix $(I - BD)^{-1}$ is known as the commodity-by-commodity total requirements matrix and it shows, on a per-dollar basis, the commodity output the economy generates to provide commodities to final users. Substituting **(5)** into **(6)** and solving for g gives:

$$\boldsymbol{g = D(Bg + e)}$$
$$\boldsymbol{(I - DB)g = De}$$
$$\boldsymbol{g = (I - DB)^{-1}De} \quad \textbf{(8)}$$

The matrix $(I - DB)^{-1}$ is known as the industry-by-industry total requirements matrix and it shows, on a per-dollar basis, the industry output the economy generates to provide an industry's commodities to final users. The vector De is a final demand vector where each entry shows the final demand for an industry's output.

Appendix E: Complete Graph Network of U.S. Economy (2020)



Appendix F: Proof of Proposition 1.

Note: This proof has been directly taken from Acemoglu et.al (2015); due to its length of ~4 pages, I have decided not to include it here. For those interested, it can be found under **Appendix A: Proof of Proposition 1.** in the Acemoglu paper.