Stochastic Input-Output Analysis and Extensions: A Case Study of the United States

Jalal Ali 1; Joost R. Santos 1,\*

\*Corresponding author: Phone: 202-994-1429, Email: joost@gwu.edu

1. Department of Engineering Management & Systems Engineering

School of Engineering & Applied Sciences (SEAS)

The George Washington University

**ABSTRACT**

The input-output (I-O) model’s capability to provide macroeconomic policy insights on interdependent economic systems has recently been extended in the field of quantitative risk analysis. As with any quantitative models, estimates of input data and associated parameters are inevitably prone to some kind of error or bias. The same statement can be said about the susceptibility of the I-O technical coefficients to imprecision originating from various sources of uncertainty. Hence, this paper provides a methodology based on stochastic I-O analysis to address these issues and subsequently measure the uncertainty when using the I-O model. The research uses the supply and use tables from the US Bureau of Economic Analysis for a period of 14 years (1998-2011) to estimate the probability distributions of the technical coefficients. The coefficients are assumed to follow the Dirichlet distribution, and their moments are evaluated by using a Monte-Carlos Simulation of 10,000 iterations. The simulation methodology is implemented in MATLAB and the results are used to generate key sector analysis. Probability distributions can be established to measure the backward and forward linkages for each economic sector. In addition, we used the eigenvalue method to determine the key sectors based on their contribution to the economy and to assess the sensitivity of the sectors to economic disruptions. In sum, this research develops a stochastic model based on historical I-O data and the results are envisioned to contribute positively to strategic economic planning and macroeconomic risk analysis.

**Keywords:** Multipliers, Stochastic I-O analysis

# INTRODUCTION

Input-output (I-O) analysis involves an estimation of industrial interdependency based on goods, workforce, products and monetary factors. The major component of input and output accounts is the Leontief inverse matrix (*I – A)-1* that consists of input-output coefficients. The coefficients provide an estimate of multiplier effects on industry or economic sector output and final demand. Leontief [[1951](#_ENREF_11)] explains the I-O economics as; “individual industry sectors are interconnected with commodity transactions”, essentially infers that total production output equals, commodity flows among interdependent sectors in addition to the output intended to satisfy final demand. In Leontief’s model, Eq.(1), the total production output of a sector is the total sum of what it propagates as input to other sectors and the sector’s output end-users consume directly.

 $x=Ax+c$ (1)

Where, *x* = output vector

 *c* = consumption vector

*A* = interdependency matrix

If n sectors are given, then x and c are column vectors each with *n* elements and *A* is a $n ×n $ matrix of technical coefficients. Coefficient, *aij*, of the interdependency matrix *A* indicates the proportion of the total production requirement of sector *j* that is provided by sector *i*. Sum of elements under the column *j* for a sector represents the total input required from the other *n-1* sectors. Lastly, element, *ci,* of the column vector *c* is the final demand by the end user from sector *i*.

The commodities values required to produce a particular economic sector’s output are obtained across an aggregate of industries, firms and thus the economic sector itself. The first step is to determine the commodity value inputs and outputs of the various economic sectors being considered within the system. This data is collected and specified as Use ‘U’ and Make ‘V’ tables. The combinations of the two tables are used to calculate the coefficients of matrix *A* [[Miller and Blair, 2009](#_ENREF_12)].

The coefficients measure, inputs per units of outputs, and then divided by the output; conventionally this is a non-linear process. The Leontief inverse has convex functional properties and in a multi-sector economy the increasing degree of terms in the inverse function is attributed indirect requirements. Stochastic analysis of the inverse has been explored by [Jansen [1994](#_ENREF_7)], who assumes that the coefficients are independent. He argues that according to the Young’s Theorem the expected value of the Leontief inverse exceeds the expected value of the inverse’s coefficients, and thus in a way it is an under-estimate.  [Dietzenbacher [1995](#_ENREF_4)] analyses aggregation of input-output (I-O) accounts and determines that the expected value is usually over-estimated. The degree of precision of the I-O coefficients remains arbitrary. Stochastic modeling of I-O accounts was initially used to analyze backward and forward linkages by Rasmussen; his analysis provides a stepping stone within key sector analysis [[Beynon and Munday, 2008](#_ENREF_2)].  [Beynon and Munday [2008](#_ENREF_2)] furthers the work of [Dietzenbacher [1995](#_ENREF_4)] and  [Lahr and Stevens [2002](#_ENREF_10)] to provide a framework for regional stochastic input-output analysis using the support of beta distribution for determining surrogate technical coefficients.

This paper is provides an alternative stochastic approach in determining the coefficients for the Leontief’s Inverse. This research extends the Beta distribution assumptions to determine surrogate technical coefficients by using Dirichlet distribution to account for the unity property of the technical coefficient matrix *A.* After determining a set of surrogate technical coefficient, key sector analysis is performed based on backward linkages and heightened demand based interdependent sectors. This is primarily done through analyzing the distributions of each sector’s backward linkages, and additionally through an eigenvector analyses determining the ‘power of pull’ and potential growth of the sectors [[Dietzenbacher, 1992](#_ENREF_3)].

# STOCHASTIC INPUT-OUTPUT ANALYSIS

Stochastic analyses were a key component in key sector analysis, additional critique to Rasmussen’s approach (explained in  [Beynon and Munday [2008](#_ENREF_2)]) was performed by [Dietzenbacher [1992](#_ENREF_3)], and  [Lahr and Miller [2001](#_ENREF_9)]. Rasmussen’s key sector analysis took into account the increases in final demand and its effect on the regional multipliers. [Dietzenbacher [1992](#_ENREF_3)] argues that the Rasmussen’s approach over-estimates the scale of linkages; on the other hand his approach provides a linkage measure based on estimating the Gross Domestic Product (GDP) through the use of I-O accounts. For backward linkage measure, the analysis encapsulates all the supplying industries expressed by Rasmussen [[Beynon and Munday, 2008](#_ENREF_2)].

The stochastic foundations used in these studies further encourage, exploration and analysis of surrogate technical coefficients using Monte Carlo simulation. By surrogacy it is implied that a distribution of possible values is determined using the mean and standard deviation of existing I-O accounts. In our analysis we consider I-O accounts of 14 years from 1998 to year 2011.  [Jansen [1994](#_ENREF_7)], [ten Raa and Steel [1994](#_ENREF_15)] and  [Dietzenbacher [2006](#_ENREF_5)] all supported the issue and use of such distributions, including the use of beta distributions and other constructed distributions determined from the coefficients of the aggregated sectors being considered within the I-O accounts.

This research extends the use of such distributions and posits the use of Dirichlet distribution to account for coefficients of the *A* matrix. The proposed distribution functions on the principle of unity which can be broken down to *n* number of pieces [[Frigyik, Kapila and Gupta, 2010](#_ENREF_6)]. This aggregation property of the Dirichlet distribution applies to the row element of a matrix *A*; coefficients determine the normalized input of each industry *j* to that particular sector, and thus sum to unity. Literature identifies that individual coefficients or elements of matrix *A* could still be beta distributed [[Beynon and Munday, 2008](#_ENREF_2)], since Dirichlet process provides a random distribution over distributions, over infinite sample space [[Frigyik, Kapila and Gupta, 2010](#_ENREF_6)]. To further explore the flexibility provided by the distribution’s constraints we decided to generate samples from Dirichlet distribution using Gamma random variables. This process is computationally efficient compared to generating random variables from Beta Distribution [[Frigyik, Kapila and Gupta, 2010](#_ENREF_6)]. This method entails three steps.

## Step 1: Data Analysis

The Leontief I-O model describes how these system interdependencies within economic sectors and regions affect the overall economic system [[Lahr and Dietzenbacher, 2001](#_ENREF_8)]. In the US, the Bureau of Economic Analysis [[BEA, 2013](#_ENREF_1)] is the agency responsible for publishing the I-O accounts for national or regional analysis. BEA devotes significant efforts in compiling and collecting annual industrial data using the North American Industry Classification System (NAICS). We use the aggregated 15 sector accounts in this research from the year 1998-2011. The analysis utilizes the Use ‘U’ and Make ‘V’ tables from these accounts to determine surrogate coefficients of matrix *A.* During matrix manipulation U and V are normalized. Also an extra element (row) needs to be added, the number of rows will now increase from 15 to 16, and this is required to satisfy the aggregation property of the Dirichlet distribution, that row of coefficient is represented by:

For j = 1 to 15, $a\_{16,j}= 1- \sum\_{i=1}^{15}a\_{i,j}$ (1)

## Step 2: Determining Gamma Realization

Mean and Standard deviation for each coefficient within the ‘U’ and ‘V’ matrix is calculated based on all 14 years. These parameters would be utilized in determining the *α* parameter for Gamma random variable generation. The *α* and adj-*α* is calculated using the following equations:

$α\_{i,j}= μ\_{i,j}\*\frac{μ\_{i,j}- μ\_{i,j}^{2}}{σ\_{i,j}^{2}-1}$ (2)

$β\_{i,j}= 1-μ\_{i,j}\*\frac{μ\_{i,j}- μ\_{i,j}^{2}}{σ\_{i,j}^{2}-1}$ (3)

$adjα\_{i,j}= μ\_{i,j}\*\sum\_{j=1}^{15}(α\_{i,j}+β\_{i,j}) $ (4)

Where $α\_{i,j}$, $β\_{i,j}$ are the parameters for a Gamma or Beta distribution, and $adjα\_{i,j}$ is the parameter for Gamma random variable generator, for each element *i* and *j* in the matrix.

A Gamma realization can be generated for each *Yi,j* from the Gamma function $Γ(α\_{i,j}, 1)$ (5)

## Step 3: Determining Dirichlet realizations and I-O coefficients

Once all the Gamma realizations are obtained we normalize them to form a pmf:

for $∀ i,j, m\_{i,j}=\frac{Y\_{i,j}}{\sum\_{i}^{}Y\_{i,j}}$ (6)

Thus $m\_{i,j}$ is a realization of Dir ($adjα\_{i,j})$ the theorem has been extensively proven by  [Frigyik, Kapila and Gupta [2010](#_ENREF_6)]. A Monte Carlo simulation of 1000 iterations was performed to determine 1000 sample realizations for the Make matrix. The exact aforementioned process is then implied to calculate the normalized realizations for the Use matrix by calculating$ u\_{i,j}$. Once sample realizations for both Use and Make matrices were determined we can calculate the *A* matrix, since *A = VU* (7)

 [[Miller and Blair, 2009](#_ENREF_12)]*.*

Similarly the three steps above could be repeated as an alternate methodology to produce matrix *A* directly by using total requirement matrices from I-O accounts of 1998-2011. Once the moments are determined using the data, sampling can be performed to determine surrogate coefficients ai,j for matrix *A* from a Dirichlet distribution.

The matrix a can then be transformed into Leontief’s inverse as (I – A)-1 and various key sector analysis can be performed on the surrogate I-O coefficients determined. Throughout the simulation process we ascertained that the standard deviation and mean of the various mechanics remained insignificantly different. This includes the comparison of parameter determined from the actual I-O accounts to the simulated ones. To ascertain process efficacy an alternate methodology was implied, which simulated surrogate I-O coefficients directly for the *A* matrix as mentioned above. The moments of both methodologies had minute difference, which proved to be very promising as a robust methodology. In our case study we will utilize the realizations of matrix *A* that were determined using the moments of I-O tables directly from BEA for the year 1998-2011 [[BEA, 2013](#_ENREF_1)].

# CASE STUDY OF UNITED STATES

This research focused on using BEA data which provided I-O accounts for 15 aggregated sectors of the US economy. Policy makers require extensive information to incorporate decisions whose affects reverb on a macro-economic scale. Thus key sector analysis becomes a critical part of that process. Many accounts have been found in the literature which deals with ‘Key sector analyses’ based on the input-output model. Key sectors analysis identifies and quantifies the economic impact of a particular sector in a given economy. A sector is either an industry or a region, or even a cluster of them [[Beynon and Munday, 2008](#_ENREF_2)]. Methodologies for key sector analysis such as ‘multiplier sums’ and ‘extraction methods’ focus solely on the interdependency affects that result from an input–output structure of the economy. As mentioned earlier inter-industrial backward and forward linkages are considered to analyze the sector’s ability to grow in its economic environment [[Schultz, 1977](#_ENREF_14)]. [Beynon and Munday [2008](#_ENREF_2)] developed stochastic analytical methods to analyze aggregation of regional sector based input-output data; and how it does not provide an accurate identification of sector prioritization. [West [1982](#_ENREF_16)] also argues that while conducting such an analysis, care should be taken about regional trade coefficients, and identifies through his formulation the impact of such sensitive data to ‘multiplier sums’. In our analysis we do not deal with regional coefficients and rely on the already simulated the macro-economic multipliers in the previous section. This research will utilize the concept of multiplier sums to determine the strength of backward linkages of each sector and thus identify key sectors based on those observations.

The Leontief inverse captures both direct and indirect linkages, and is represented as L = (*I-*A)-1 ,the multiplier sums for sector *j* is the column sum of the Leontief inverse for that sector. Similarly the forward linkage is determined by the row sum of the Leontief inverse of the Sector *i*. In our analysis we calculated these column and row sums for each sector for the surrogate technical coefficients determined in the earlier section. Li,j is the coefficients of the inverse matrix for sectors *i* and *j*.

The backward linkages, BLj are determined by the following equation:

$BL\_{j}= n\*\frac{\sum\_{i=1}^{15}L\_{i,j}}{\sum\_{i=1}^{15}\sum\_{j=1}^{15}L\_{i,j}}$(8)

Using the equations above backward linkages for each of the 1000 realizations was calculated. On that basis, we can also determine the average backward linkages for each sector, and thus rank all the sectors based on their strength of backward linkage; implying how extensively the output of that sector is dependent on inter-industry input (or demand of output of other sectors).

 Table 1 provides a ranking of sectors based on the strength of backward linkage for each sector. The analysis identifies 5 key sectors: Manufacturing (S5), Agriculture, Forestry, Fishing and Hunting (S1), Construction (S4), Transportation and warehousing (S8) and Information (S9). The output production of these key sectors is heavily dependent on other sectors. Thus if growth in these sectors is pursued it would cause a ripple effect and would boost demand of products from other sectors these key sectors are dependent on. The boost in demand would motivate other sectors to grow as well and thus enable the entire economy to grow.

Figures 1 and 2 illustrate distributions associated with the multiplier sum of each sector based on its backward linkage. We observe that all the sectors have a fairly similar spread within their distributions; the exception which holds out is the Utilities (S3) sector which has a fair spread across all other sectors. This is fairly intuitive since a measure of this sector, Utilities, is fairly spread across all other sectors, and contributions from other sectors could fairly fluctuate depending on various circumstances.

A complimentary approach to assessing backward and forward linkages is given by  [Dietzenbacher [1992](#_ENREF_3)], which further elaborates the concept of ‘Power of Pull’ within the overall network influences. The idea presented is that a sector would be more powerful if it’s interconnected to other powerful sectors, and

Table - Sector Ranking based on Backward Linkages

|  |  |  |  |
| --- | --- | --- | --- |
| Rank | Sector | Sector Name | Average BLj |
| 1 | S5 | Manufacturing | 1.326304 |
| 2 | S1 | Agriculture, Forestry, Fishing and Hunting | 1.224062 |
| 3 | S4 | Construction | 1.114771 |
| 4 | S8 | Transportation and Warehousing | 1.060625 |
| 5 | S9 | Information | 1.041928 |
| 6 | S2 | Mining | 1.010095 |
| 7 | S13 | Arts, Entertainment, Recreation, Accommodation and Food Services | 1.008763 |
| 8 | S3 | Utilities | 0.952784 |
| 9 | S15 | Government | 0.951998 |
| 10 | S14 | Other Services, Except Government | 0.946515 |
| 11 | S12 | Education services, Health Care, and Social assistance | 0.938831 |
| 12 | S10 | Finance, Insurance, Real Estate, Rental and Leasing | 0.901624 |
| 13 | S11 | Professional and Business Services | 0.873864 |
| 14 | S6 | Wholesale Trade | 0.824414 |
| 15 | S7 | Retail Trade | 0.823421 |



Figure - Box Plot illustrating backward linkage spread



Figure - Ranked distribution of Column Sum for each Sector

thus its influence is aggregated as a result of such interdependence. In this analysis the Leontief Inverse matrix is used input for the Eigen analysis, since the Power of Pull concept is more closely associated with a demand model and the backward linkages method.

In our analysis we determine un-weighted average of coefficients to produce a single L matrix from our 1000 sample realizations. Once that is achieved the Frobenius dominant eigenvalue is calculated for the matrix, and the corresponding eigenvector is observed to determine the ‘power of pull’ of all 15 sectors. This concept asserts that key sectors would have a greater power of pull, and thus be more influential in determining growth gradients within the economy. Figure 3 illustrates the eigenvectors associated with the Frobenius dominant eigenvalue. We observe that Manufacturing (S5) has the highest power of pull confirming the findings of the previous analysis. Other sectors Agriculture, Forestry, Fishing and Hunting (S1), Mining (S2), Transportation and Warehousing (S8) and Information (S9) also show significant power of pull compared to other sectors. On the contrary the second and third highest sectors showing the power of pull are Finance, Insurance, Real Estate, Rental and Leasing (S10), and Professional and Business Services (S11), but they show one of the lowest backward linkages in the previous analysis. Both sectors are potential giants when it comes to comparing sector GDPs and promise economic growth in this day and age. Although intuitively they do not demand outputs extensively from other sectors since they belong to the service industry.



Figure - Dominant Eigenvector illustrating Power of Pull

# CONCLUSION

In this research we performed stochastic analysis to determine sample realizations of technical input-output coefficients. BEA I-O accounts, Use and Make tables for a period of 14 years (1998-2011) were used in our analysis and data’s moments calculated. Each coefficient was assumed to be a random variable of a Dirichlet distribution, and each realization was calculated using the moments calculated from the data [[Frigyik, Kapila and Gupta, 2010](#_ENREF_6)]. 1000 Realizations of coefficients were simulated using a Monte Carlo simulation to produce the *A* matrix.

Two alternate methodologies were utilized; the first using moments of the Use and Make tables from BEA and simulating sample realizations for each type to produce sample normalized Use and Make tables. The Use and Make matrices were then used to calculate the *A* matrix [[Raa and Rueda-Cantuche, 2007](#_ENREF_13)]. The second methodology involved calculating moments directly from the I-O accounts provided by BEA and simulating realizations for matrix *A*. Both methodologies produced similar results, validating the process.

Matrix *A* was then converted to the Leontief inverse to estimate the direct and indirect strength of backward linkages of each sector, key sectors were determined based on that approach. An additional complementary Eigen analysis was conducted to determine the potential of growth within the economy based on the influence of each sectors. The analysis determined the Manufacturing sector (S5) to be the most influential and interdependent sector along with few others. Further analysis could be made to observe Forward linkages and determine a multi-objective criterion of determining key sectors.

Policy maker can focus their efforts on key sectors to stimulate growth in the economy and also be vary that potential inoperability to these sector could hamper growth due to decreased demand within the economic system. This research produced a unique framework for stochastic input-output analysis which can be used in conjunction with improved key sector analysis techniques found in the literature.

# REFERENCES

1. BEA, *Interactive access to industry economic accounts data: Input-output*, (2013),B. o. E. Analysis.Available at: <http://bea.gov/iTable/index_industry_io.cfm>

2. M. Beynon and M. Munday, *Stochastic key sector analysis: An application to a regional input–output framework*, Ann Reg Sci 42 (2008), no. 4, 863-877.

3. E. Dietzenbacher, *The measurement of interindustry linkages: Key sectors in the netherlands*, Economic Modelling 9 (1992), no. 4, 419-437.

4. E. Dietzenbacher, *On the bias of multiplier estimates*, Journal of Regional Science 35 (1995), no. 3, 377.

5. E. Dietzenbacher, *Multiplier estimates: To bias or not to bias?\**, Journal of Regional Science 46 (2006), no. 4, 773-786.

6. B. A. Frigyik, A. Kapila and M. R. Gupta, *Introduction to the dirichlet distribution and related processes*, UWEE Technical Report Number UWEETER-2010-0006 (2010), University of Washington. Available at: <http://www.ee.washington.edu>

7. P. S. M. K. Jansen, *Analysis of multipliers in stochastic input-output models*, Regional Science and Urban Economics 24 (1994), no. 1, 55-74.

8. M. L. Lahr and E. Dietzenbacher, *Input-output analysis : Frontiers and extensions*, Palgrave, Basingstoke, Hampshire, England, New York, 2001.

9. M. L. Lahr and R. E. Miller, *Regional science perspectives in economic analysis : A festschrift in memory of benjamin h. Stevens*, Elsevier, Amsterdam: New York, 2001.

10. M. L. Lahr and B. H. Stevens, *A study of the role of regionalization in the generation of aggregation error in regional input –output models*, Journal of Regional Science 42 (2002), no. 3, 477-507.

11. W. W. Leontief, *Input-output economics*, Scientific American 185 (1951), no. 4, 15-21.

12. R. E. Miller and P. D. Blair, "The commodity-by-industry approach in input–output models," *Input-output analysis : Foundations and extensions*, Cambridge University Press, New York, 2009, pp. 184-196.

13. T. T. Raa and J. M. Rueda-Cantuche, *Stochastic analysis of input–output multipliers on the basis of use and make tables*, Review of Income and Wealth 53 (2007), no. 2, 318-334.

14. S. Schultz, *Approaches to identifying key sectors empirically by means of input-output analysis*, Journal of Development Studies 14 (1977), no. 1, 77.

15. T. ten Raa and M. F. J. Steel, *Revised stochastic analysis of an input-output model*, Regional Science and Urban Economics 24 (1994), no. 3, 361-371.

16. G. R. West, *Sensitivity and key sector analysis in input-output models*, Australian Economic Papers 21 (1982), no. 39, 365.