WORKING PAPER SERIES

A Network-based View of the U.S. Energy Sector

Vipin Arora | Elizabeth Sendich | Julia Teng

February 2016

U.S. Energy Information Administration
Washington, DC 20585

Summary

We describe portions of the U.S input-output tables through the tools of networks analysis—focusing on either energy intensive industries or those that are part of the separate and distinct energy sector. We first represent both energy intensive and energy sector industries visually through network diagrams for the years 1997, 2002, and 2007. Next, we show that the energy sector is generally more densely connected than either energy intensive industries or all industries over those years, and is more likely to have groups of three sub-sectors all linked as well.

We then move to the level of individual industries within the broad sectors and find that energy intensive industries have the most in-coming connections on average for these tables. Energy sector ones have fewer, but the number grows over time, as do outgoing connections. Other measures of centrality—closeness and betweenness—vary over time for both the energy sector and energy intensive industries. Specifically, petroleum refining and electricity generation stand out for their centrality, drilling oil and gas wells for its lack of centrality.

1. Introduction

Studying U.S. input-output tables—as surprising as it may seem—doesn't exactly get the heart-racing. There is something very sleep-inspiring about trudging through hundreds upon hundreds of multipliers or makes and uses. Import matrices sound more exotic, but are even less exciting. And who would ever want to miss the thrill of trying to figure out the differences between commodities and industries, converting codes between industry classification systems, or using bridge tables?

Most of us just don't think about the companies, products, and relationships embodied in the input-output tables as a big glob of numbers—we visualize them within a network. Yes, petroleum refineries used $375 billion worth of output from the oil and gas extraction sector in 2007. But what that really means is that big U.S. refineries, primarily on the Gulf Coast, received lots of oil from domestic producers, many in Texas and Louisiana, via truck, pipeline, rail, and even barge. The refining and oil extraction industries are two elements within a network—the U.S. industrial network—and the goods and services traded between them are their links. Yet, the U.S. input-output tables have rarely been either analyzed or represented using the tools of network analysis.[1]

In this paper we describe portions of the U.S input-output tables through such tools—focusing on industries that are either energy intensive or part of the energy sector. We represent both energy intensive and energy sector industries visually through network diagrams, and provide summary statistics of each over time. We also characterize the links between industries as a whole, and the industries themselves using different measures of prominence.

2. Network Representation of the U.S. Input Output Tables

The U.S. Bureau of Economic Analysis (BEA) publishes tables on the value of production across industries in a given year, along with the value of trade between industries, in the input-output accounts. The tables are published annually for about 65 U.S. industries, and every 5 years for roughly 400 sectors. These latter benchmark accounts are currently available through 2007, while the annual tables run through 2013.[2]

The standard input-output tables are the make and use. The make table shows the value of goods and services [commodities] produced by each industry in a specific year, while the use table shows the value of goods and services used by each industry in that year. The terms industry and commodity are not synonymous: commodities are the goods and services produced by industries: i.e., industries use commodities to make other commodities. This allows the input-output framework to handle industries which may produce more than one commodity (Miller and Blair, 2009; 184-185).

The make and use tables provide information about the size, structure, and even the technology used by specific industries—but they are primarily descriptive. Other tables derived from the standard make

[1] See Carvalho (2014) and references therein. Other areas of economics have used network analysis for much longer as described in Jackson (2014).

[2] For more detail on the U.S. input-output tables and the industry accounts more generally, see Streitwieser (2010).

and use—requirements tables—can approximate various economic impacts across and within industries.[3]

The economic impacts are generally summarized by these so-called requirements, which attempt to quantify the effect of a change in spending by households, business, or government on overall and industry-level output, employment, and income.[4] The requirements—also called multipliers—can decompose changes in spending between direct and indirect effects. Direct effects are the value of additional inputs required by industry A to produce an additional dollar of its primary commodity. Indirect effects are the value of additional inputs required by industries other than A to produce that dollars' worth of commodity from industry A.

For example, if an oil producer in Texas' Permian Basin wants to extract another barrel to meet increased demand for oil, the direct effects are the value of any additional inputs required to make that happen [i.e. more electricity, construction of structures, etc.]. The indirect effects are the value of any additional inputs used by non-oil industries—more bulldozers for the construction industry or coal consumption by electric utilities.[5]

Total effects sum the direct and indirect effects and are found in the BEA's total requirements tables. The industry-by-industry variant of this table lists industries both across columns and down rows [Figure 1]. The value in each cell shows the production required, both directly and indirectly, from the industry represented by that row to produce a dollar of output for the industry represented in that column.

Going back to the example of additional oil production, Figure 1 shows that the total requirements coefficient for utilities in the mining column is 0.0076. This means that increasing mining production [which includes oil extraction] by a dollar requires the utility sector to produce an additional 0.7 cents worth of its primary commodity. And summing all of the coefficients down the mining column shows

[3] These approximations vary, but all require strong assumptions about the structure and technology of individual industries. See Gretton (2013) for an overview.
[4] This spending is called final demand, but multipliers can also be adjusted so that they reflect the impact of changes in another industry's output as well. See Steinbeck (2004).
[5] Technically, the direct effects of such a change are the differences in commodity inputs required by an industry to produce a dollar of the industry's output. Indirect effects broaden this to include changes in commodity inputs required by all other industries to produce that dollar of output.

that the dollar increase in demand for oil leads to total increases in the value of production throughout the economy of around $1.57.

Figure 1: Sample BEA total requirements table for 2013

Industry-by-Industry Total Requirements, After Redefinitions

(in producers' prices)

2013

	Industries/Industries	11	21	22	23
IOCode	**Name**	**Agriculture, forestry, fishing, and hunting**	**Mining**	**Utilities**	**Construction**
11	Agriculture, forestry, fishing, and hunting	1.2681500	0.0103165	0.0086749	0.0253938
21	Mining	0.0513216	1.1179600	0.1148640	0.0547786
22	Utilities	0.0133362	0.0075510	1.0072300	0.0066163
23	Construction	0.0115953	0.0130253	0.0118680	1.0043500
31G	Manufacturing	0.3730150	0.1642710	0.1421000	0.3876430
42	Wholesale trade	0.0900044	0.0267751	0.0243131	0.0655058
44RT	Retail trade	0.0026814	0.0020080	0.0024163	0.0663582
48TW	Transportation and warehousing	0.0552191	0.0336601	0.0527895	0.0400826
51	Information	0.0136611	0.0139783	0.0126195	0.0192947
FIRE	Finance, insurance, real estate, rental, and leasing	0.0904621	0.0578832	0.0486660	0.0687903
PROF	Professional and business services	0.0765449	0.1029200	0.0742622	0.1050560
6	Educational services, health care, and social assistance	0.0012978	0.0001771	0.0004212	0.0005731
7	Arts, entertainment, recreation, accommodation, and food services	0.0063776	0.0056026	0.0082078	0.0074635
81	Other services, except government	0.0069160	0.0039598	0.0038384	0.0085595
G	Government	0.0167490	0.0101169	0.0122947	0.0138473
	Total Industry Output Requirement	**2.0773327**	**1.5702036**	**1.5245666**	**1.8743096**

Source: BEA

2.1. Diagrams

This industry-by-industry total requirements table can be mapped into a network by treating the industries as vertices [sometimes called nodes] and the requirements coefficients as connections between the industries [also called edges or links]. We use both weighted links, where connections incorporate the values of requirements coefficients, or unweighted, in which case connections are equal to one if there is a link and zero otherwise.

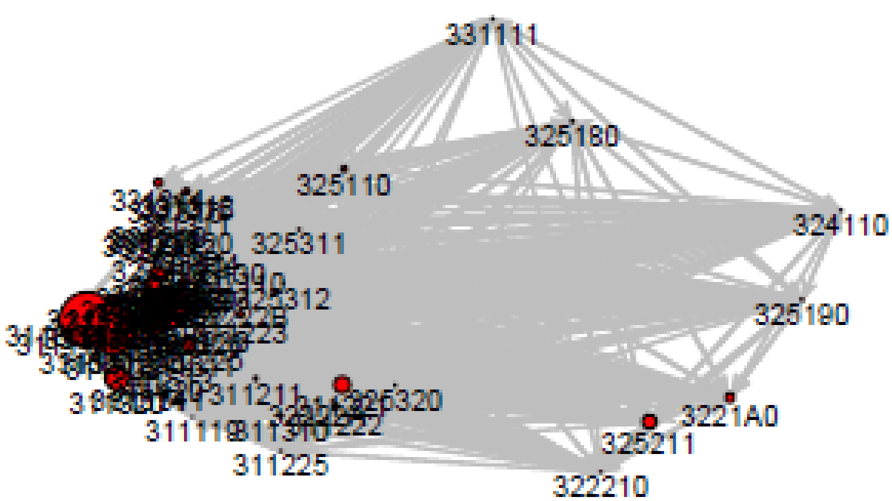
We also represent the network as directed—spelling out that industries on the rows of the total requirements table provide inputs to industries on the columns. In the example above with the mining sector, the weight of the edge directed from utilities to mining is 0.0076, and this is represented in our diagrams by an arrow from utilities to mining.

Figure 2 and Figure 3 show sample network diagrams of energy intensive industries and the energy sector for the total requirements tables that were published by the BEA in 1997; Figure 12 and Figure 13 in the appendix show diagrams for the years 1997, 2002, and 2007. The definition of an energy intensive industry follows the U.S. Energy Information Administration's Annual Energy Outlook (AEO), and generally includes those industries which are heavy users of energy.[6] The energy sector is much narrower, as we only include those industries which produce or supply energy. Our specific definition follows the Department of Energy's Quadrennial Energy Review (QER).[7]

The match between our definition and that from the QER is not exact because portions are captured in other industry definitions. For example, oil and gas pipeline related construction [23712] is not included in our definition of the energy sector because it is part of the much larger construction sector, neither are gasoline stations [447], petroleum and products wholesalers [4247], or any of the energy-related transportation contained in industries with codes between 441 and 446. See Table 4 and

Table **5** in the appendix for a list of industries in each of these groups.

Figure 2: Network plot of the U.S. total requirements table for energy intensive industries in 1997



Source: BEA, Author calculations

Both diagrams are directed, although the arrows are not shown for energy intensive industries for clarity; they also use a multidimensional scaling (MDS) method for generating the plots (see Kolaczyk

[6] See http://www.eia.gov/forecasts/aeo/overview/industrial.html.

[7] See http://energy.gov/epsa/downloads/quadrennial-energy-review-full-report.

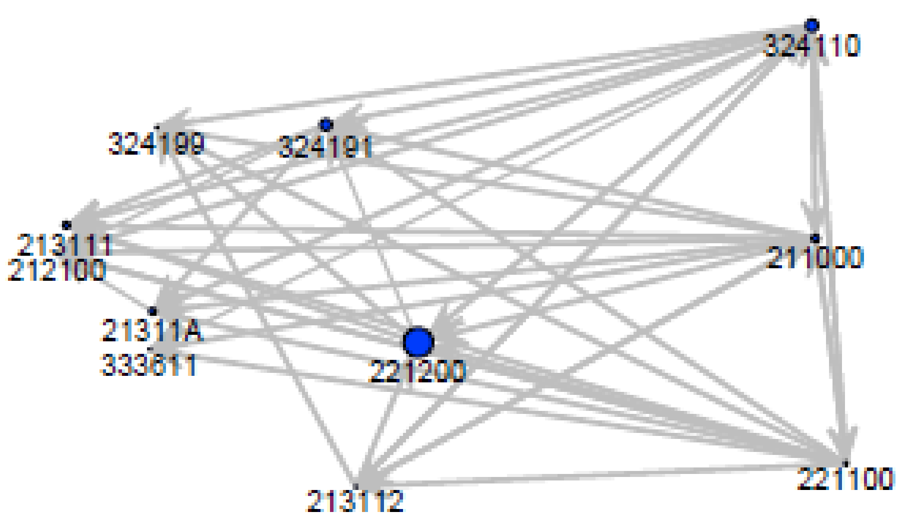
and Csárdi, 2014; 32). Figure 14 and Figure 15 in the appendix display the respective network diagrams using the method of Fruchterman and Reingold (FR).

The industries are color coded [red for energy intensive industries, blue for energy sector], the weight of connecting lines are thicker for larger requirements coefficients, and the size of the nodes representing industries is scaled to match the production in that industry [the greater is production, the bigger the node]. Finally, we have excluded very small connections in the appendix diagrams—the plots on the left-column of each figure only show industries that have total requirements coefficients larger than 0.005, while those in the right columns only show those larger than 0.03.

In general the plots for energy intensive industries are too large to provide much information—either using the FR or MDS plots. But the MDS plots in the appendix are arranged in such a way that certain industries stand out. Although these can vary over time, and depending upon how large connections must be, four industries are consistently grouped apart: petroleum refineries [324110], paperboard container manufacturing [322210], other basic inorganic chemical manufacturing [325180], and other basic organic chemical manufacturing [325190].

One interpretation of their separation is that these industries provide a lot more to the supply chain than they take in: each is a basic commodity industry that produces extremely useful/varied intermediates. Interestingly, the size of these industries in terms of gross output, as indicated by the diameter of the nodes, is not large compared to other energy intensive industries.

Figure 3: Network plot of the U.S. total requirements table for energy sector industries in 1997



Source: BEA, Author calculations

The energy sector provides a clearer picture because it is much smaller. The energy sector industries which stand out are petroleum refineries [324110]; electric power generation, transmission, and distribution [221100]; and oil and gas extraction [211000]. Petroleum refineries are an important input to many manufacturing and retail industries and also stood out across the energy intensive industry diagrams in the appendix. Electric power is used almost everywhere, and oil and gas extraction feeds into refineries [as well as manufacturing plants].

3. Network Statistics

The total requirements tables show substantial differences in size between 1997 and 2007 [Table 1]: industry numbers fell by nearly 100 over that time, from nearly 500 to below 400. Part of this was due to a change in the industry classification system from SIC [standard industrial classification] to NAICS [North American industry classification system]. But there was also redefinition within the NAICS system between 2002 and 2007.

Table 1: Descriptive statistics

Year	Total Number of Industries	Number of EI Industries	Number of ES Industries
1997	491	83	11
2002	426	65	11
2007	389	52	9

Source: BEA, Author calculations

Both energy intensive industries and the energy sector show drops in the number of industries as well—for the energy sector this is marginal, but the number of energy intensive industries falls by over 30 from 1997 to 2007. Much of this had to do with redefinitions, from SIC to NAICS, and then within NAICS. For example, Inorganic Chemicals [NAICS codes 32512, 32513, and 32518] were previously a larger number of SIC codes [2812, 2813, 2816, 2819]. These aggregate numbers, however, don't provide any information on the nature of relationships between industries, and how they may or may not have changed.

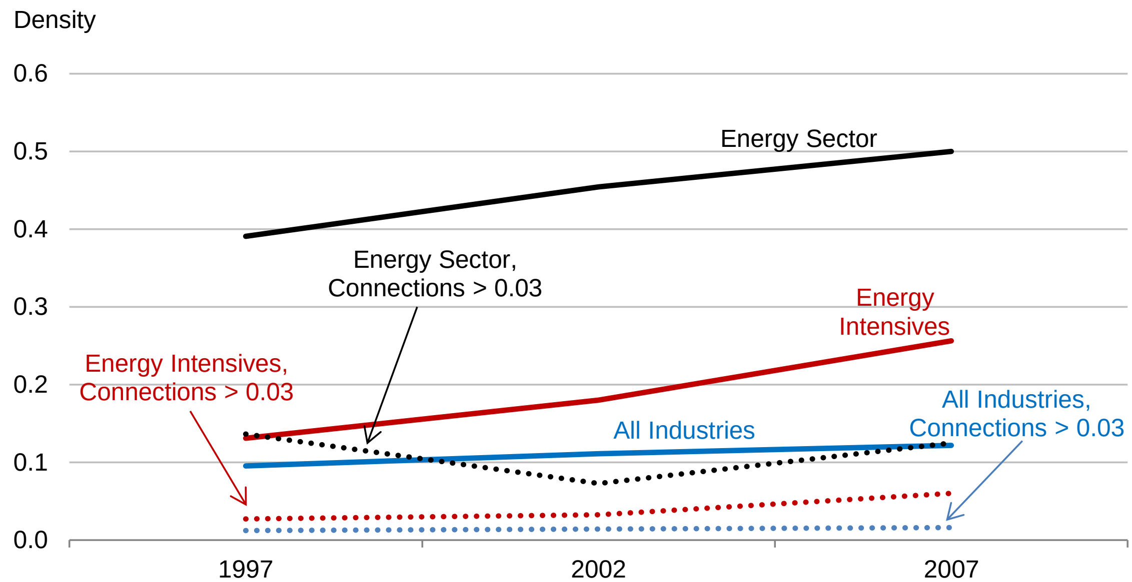
To better understand these relationships at the level of all industries [and the energy intensive and energy sector subsets] we calculate three aggregate network statistics for the total requirements tables in each year: density, transitivity, and reciprocity. We require connections [requirements values] to have a value greater than 0.005 to remove very small links between industries, and also show the statistics when connections must be larger than 0.03.

Density is the proportion of actual connections to potential connections—a measure of average inputs and outputs per industry. It rose for all industries, energy intensive industries, and the energy sector from 1997 to 2007 in the total requirements tables in Figure 4.

The larger networks [all industries > energy intensives > energy sector] have lower density values—because the number of potential connections rises quickly with each new industry that is added. The order is the same when connection size is limited, but in this case the density values are smaller.

Energy sector density rises to almost 50% in 2007—so half of potential connections actually existed—but there are only 9 industries. When small connections are removed, density falls near 10%, which is roughly double the value for energy intensive industries in 2007 when connections are larger than 0.03. Energy intensive industry density in the other case is about 25% in 2007, while that for all industries remains near 10% for each total requirements table, and falls below 5% when connections are over 0.03. Increasing density indicates more connections to other industries—reflecting the exchange of intermediates.

Figure 4: Density of U.S. total requirements tables and subsets in 1997, 2002, and 2007



Source: BEA, Author calculations

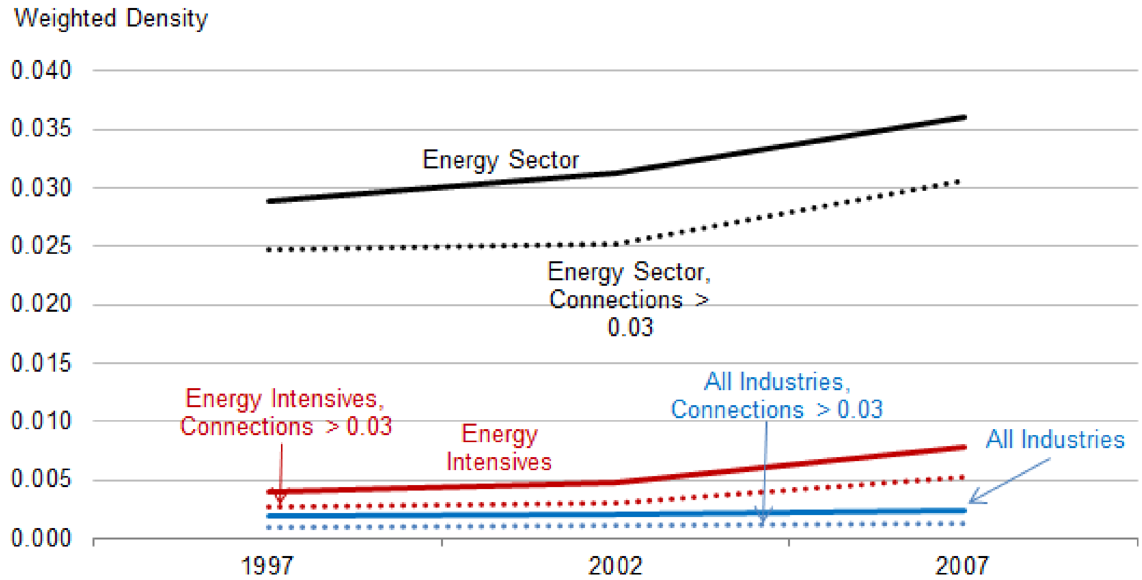
Density calculations do not explicitly account for the size of links between industries. Weighted density does—it is the weighted average of potential connections relative to actual ones [Figure 5]. The pattern here is the same: larger networks have smaller values, and when only large connections are included the density values fall. But the scale is much smaller than with unweighted density, below 0.05 in all cases.

The differences across sectors in weighted density are much smaller than the differences when density is unweighted—indicating that most of the connections are not large. This is also true when links must be above 0.03, but the differences between all industries, energy intensive industries, and the energy sector are even less.

Transitivity—also called the clustering coefficient—divides the network into groups of three industries [triads] and measures the proportion of such groups which are all individually connected. So if industry A connects to industry B, and industry B to industry C, then transitivity tells us the proportion of times that industry A also connects to industry C.

The idea behind transitivity is to see how often two industries that are themselves connected share different links. In an alternative context, it is the same as observing the proportion of two of your friends that are also friends with each other.

Figure 5: Weighted density of U.S. total requirements tables and subsets in 1997, 2002, and 2007



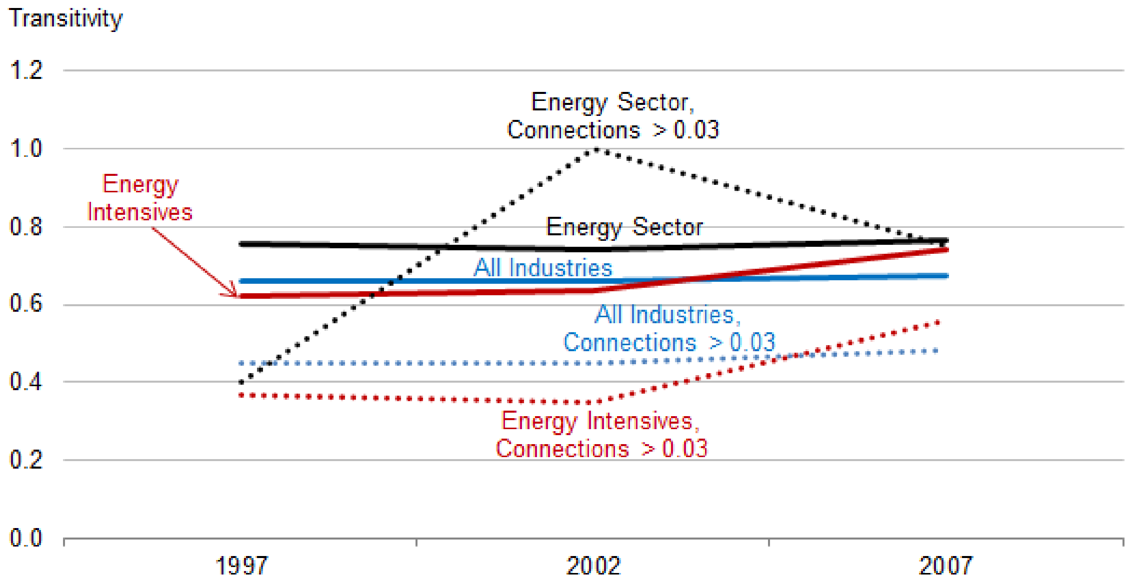
Source: BEA, Author calculations

The energy sector, other than when large connections are removed in 1997, has the largest transitivity values, likely because it is smaller than the other two [Figure 6]. The transitivity of all industries and energy intensive industries when all connections are kept is roughly the same; all industries is slightly higher in 1997 and 2002, energy intensive industries a bit higher in 2007. This same pattern holds when large connections are removed, but the difference between all industries and energy intensive industries is bigger. But the differences between the three groups are not huge—with the exception of the energy sector in 2002.

One interpretation of these results is in terms of isolation—there isn't much isolation in the total requirements tables [most of my friends are also friends with each other], irrespective of whether we look at all industries, energy intensive industries, or the energy sector. Isolation does grow in all industries and energy intensive industries, however, when only large connections are kept. The energy sector may be too small to show an effect from removing the larger connections.

Additionally, one can also think of this as reflecting whether or not commodities represented by these links are upstream or downstream/end use. Because energy is both a base commodity and an end-user product, it should be above the overall average. Energy intensive industries, however, are generally upstream only—because they produce base/bulk commodities, so should be below the average.

Figure 6: Transitivity of U.S. total requirements tables and subsets in 1997, 2002, and 2007



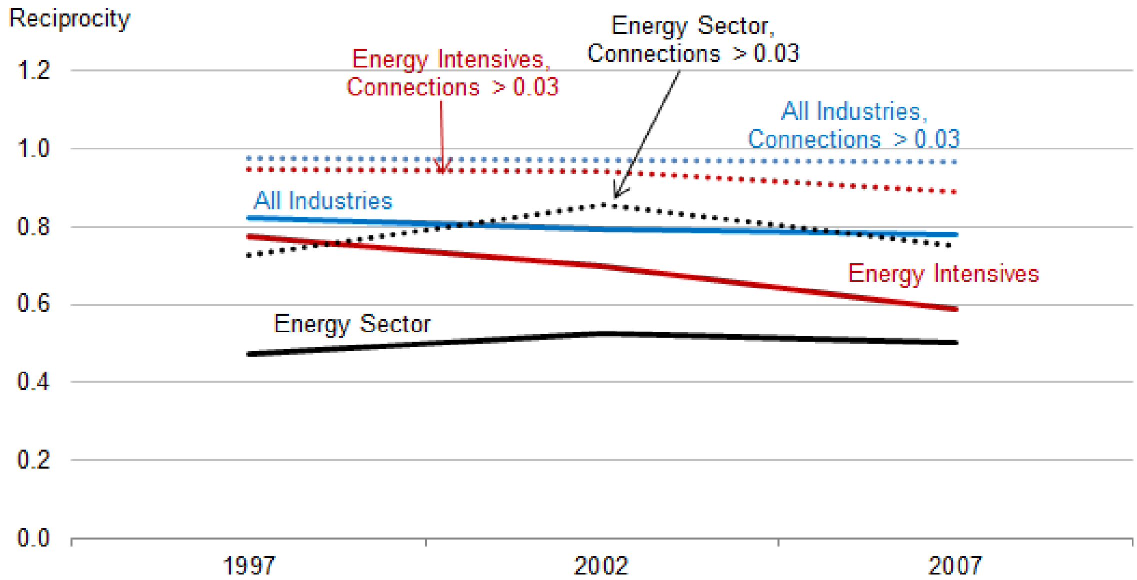
Source: BEA, Author calculations

The concept of reciprocity moves away from groups of three industries and considers the relationships between two—it quantifies the extent to which links exist in both directions. That is, if industry A is an input to industry B [i.e., a non-zero entry in the column of industry B in the total requirements table], reciprocity quantifies the proportion for which industry B is an input to industry A.

The relative magnitude of reciprocity values is the opposite of density—and it appears to decline from 1997 to 2007 [Figure 7]. Bi-directional connections are much more prevalent in all industries and energy intensive industries, as opposed to the energy sector. This is because every industry has to get inputs from somewhere, and the more sectors there are, the easier to form groups of three that are fully connected.

And for all three of our subgroups keeping only larger connections raises the reciprocity values. The fact that removing small links increases reciprocity indicates that if industry A is a large user of industry B's products, B is more likely to use A's products in notable quantities as well. This may have something to do with specialty supply chains and corporate integration.

Figure 7: Reciprocity of U.S. total requirements tables and subsets in 1997, 2002, and 2007



Source: BEA, Author calculations

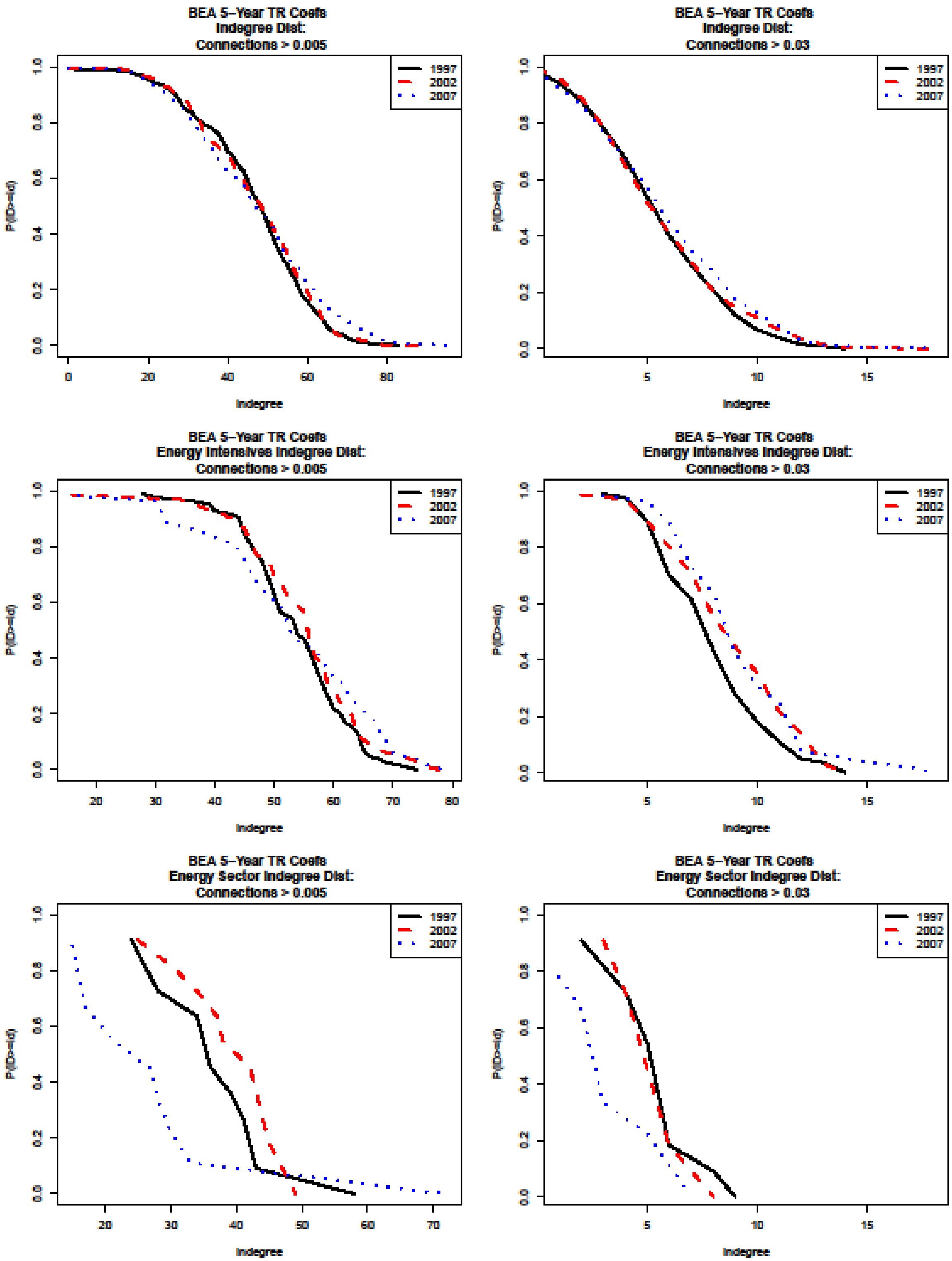
3.1. Industry-level statistics

Statistics below the network level provide more information on connections [degrees] and detail on how prominent or central industries are within the network [closeness and betweenness]. All of the industry-level statistics are shown as distributions which display the value of the specific statistic on the horizontal axis, and the probability that a given industry is less than or equal to that value on the vertical. Distributions are a concise way to summarize such information over the entire network, as each industry has an individual value. Additionally, distributions allow for comparisons of the total requirements tables from different years on the same plot.

One of the most popular statistics is that of in-degrees: the number of connections an industry has coming in. In the total requirements table this is the number of non-zero entries for an industry on a column. The natural corollary to this measure is out-degrees: the number of connections an industry has going out. This is the number of non-zero entries for an industry on a row in the total requirements table. Weighted in or out-degrees are the number of connections times the value of the respective total requirements coefficients.

Figure 8 and Figure 9 show the in and out-degree distributions for the total requirements tables from 1997-2007. The left-column of each figure excludes all connections less than 0.005, while the right column excludes those that are smaller than 0.03. The weighted versions of these diagrams, as well as the total-degree [sum of in and out-degrees] distributions are in Figure 16, Figure 17, Figure 18, and Figure 19 in the appendix.

Figure 8: In-degree distributions of U.S. total requirements tables/subsets in 1997, 2002, 2007



Source: BEA, Author calculations

The distributions of in-degrees for all industries—irrespective of the minimum requirements coefficient—do not change much between 1997 and 2007 [Figure 8]. Industries with the most in-connections have over 80 inputs when only very small links are excluded, and closer to 20 when connections must be larger than 0.03. The central tendencies are similar as well over time with either cutoff for the requirements coefficient size.

Energy intensive industries do show more of a difference across years. In the left-hand column of the figure, it appears that the 2007 tables have a greater concentration of industries with over 60 in-connections, but fewer between 30 and 60. When only larger coefficients are included the 1997 table has a smaller number of in-connections across industries, while 2002 and 2007 are similar.

It appears that, on average over time, energy intensive industries have raised the number of different supplying industries when there are many inputs, but reduced them slightly with a moderate number of inputs. And this is even more pronounced when only larger connections between industries are considered.

This could point to more production diversification or product complexity, where a single factory takes in a greater variety of products to produce one item. For example, the production of higher quality or differing special-needs pet foods might also require the input of many more additional specialty ingredients than was previously demanded by pet owners.

Dog and cat food manufacturing [311111] has the most in-connections of any energy intensive industry in 1997 [see Table 6 in the appendix for this list when connections are greater than 0.03], seasoning and dressing manufacturing the most in 2002 [311940], and fruit and vegetable canning and drying [311420] in 2007 [Table 2].

These are industries where the function or quality of the product is often increasing in complexity to meet consumer demands. An example is the need for salad dressings to have long shelf lives, good taste, low fat or other health benefits, and possibly additional specific characteristics such as “non-separating” or “extra thick”.

This pattern is repeated for the energy sector, but it is the 2007 table where in-connections are concentrated at lower values, and 1997 and 2002 are broadly similar. So energy sector industries had relatively fewer inputs in 2007 than before—both with and without large connections.

This might represent consolidation—a single entity could be the producer and user of its own inputs—or that producers have zeroed in on the most effective or profitable products that can be made based on the resources available. If the product type narrows there are fewer inputs into what is otherwise the same process code.

Petroleum lubricating oil and grease manufacturing [324191] has the most in-connections of this group in 1997, drilling oil and gas wells [213111] the most in 2002, and turbine and turbine generator set units manufacturing [333611] in 2007 [Table 2].

It is conceivable that U.S. production of lubes and greases has narrowed to specific types, which can be motivated by demand, production costs, or sales price. The type of drilling done in the US has changed dramatically over the past few decades as well, so for it to have reduced complexity, given the drive for maximum efficiency through logistics to increase profits, is no surprise. As for turbine production, the demand for ever improving efficiency and increases in gas-based electricity production likely explain why this industry required the most inputs in 2007.

Table 2: NAICS codes of industries with the maximum and minimum in and out-degrees from 1997-2007

Statistic	Energy Intensive			Energy Sector		
	1997	2002	2007	1997	2002	2007
Max Ideg	311111	311940	311420	324191	213111	333611
Min Ideg	325120	311930	324110	221100	221100	213111
Max Odeg	324110	324110	324110	221100	221100	Various
Min Odeg	Various	Various	Various	213111	213111	213111

Source: BEA, Author calculations

Out-degree distributions for all industries have a similar shape to the in-degree ones across time, but are shifted to the right [Figure 9]. The maximum values are above 300, irrespective of the requirements coefficient size included, even though the central tendencies are much lower.

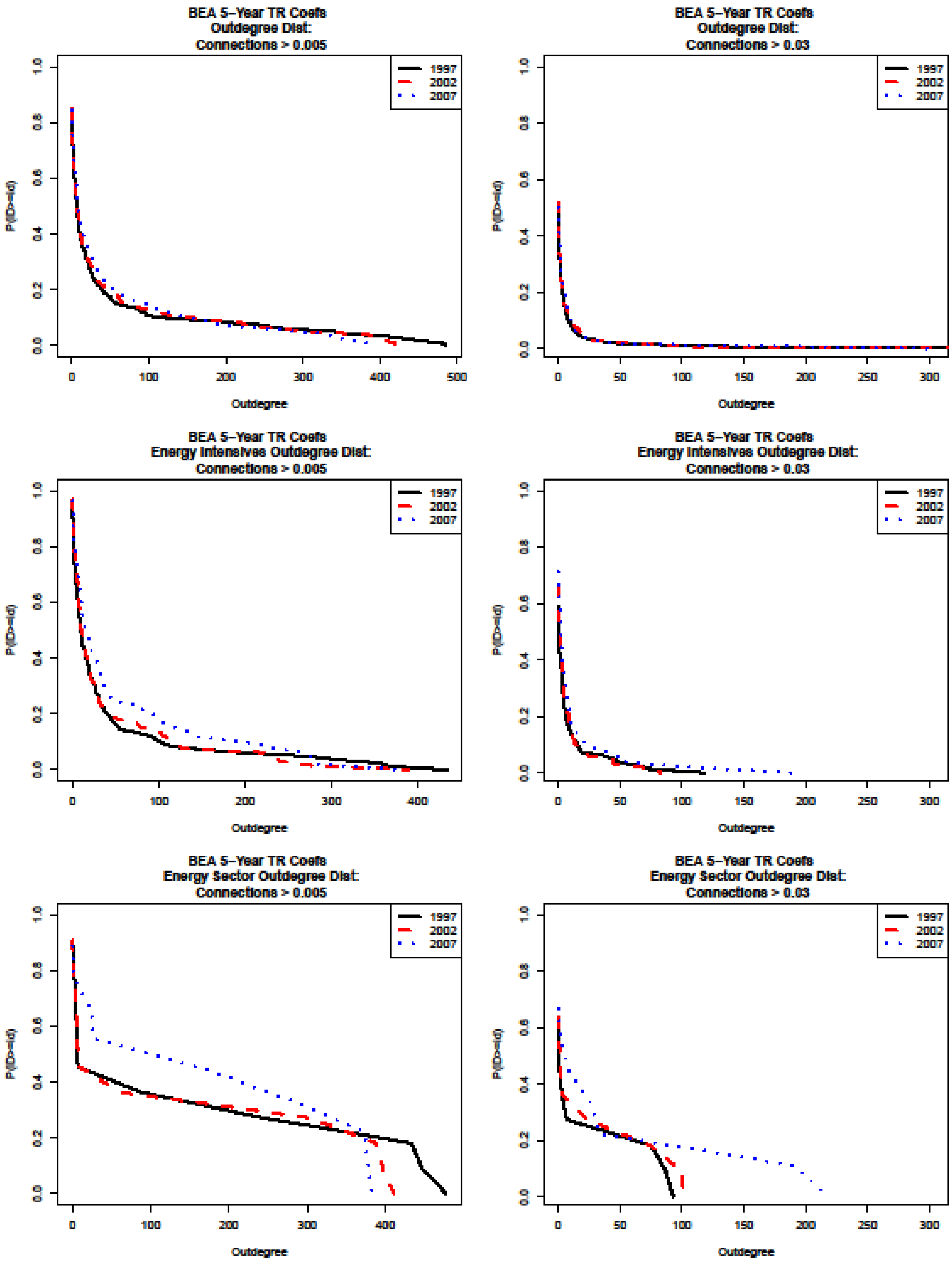
The energy intensive industries show similarity across total requirements tables, although 2007 does appear to have a greater concentration of out-degree connections below 250. This indicates energy intensives supplied more industries on average in 2007 than before. Petroleum refineries [324110] have the most out-connections in 1997, 2002, and 2007 [Table 2].

Why? Products are more complex because they meet a greater variety of needs, and this may mean selling outputs to diverse customers. For example, the demand for specialty pet foods reflects the general shift in pet ownership, shown in one way by the growth of "doggie daycares", which are also a new customer for purchases of specialty dog foods.

The same pattern holds for the energy sector—the 2007 tables differ from the other two, with a greater concentration at lower values of out-degrees. Electric power generation, transmission and distribution [221100] has the most out-degree connections in 1997 and 2002; multiple industries have the most in 2007 [Table 2].

As discussed above, the electric power sector has shown a trend towards greater use of natural gas, which has been offset by reductions in inputs of coal and oil.

Figure 9: Out-degree distributions of U.S. total requirements tables/subsets in 1997, 2002, 2007



Source: BEA, Author calculations

Figure 10 and Figure 11 display the distributions for two other centrality measures at the industry level: closeness and betweenness. The left-column of each figure excludes all connections less than 0.005, while the right column excludes those that are smaller than 0.03. Closeness attempts to measure how far away one industry is from all others, i.e. how many links on average it takes to get to other industries. It tells us which industries in the total requirements table have the shortest distance to other industries on average. In other words, they are central to the supply chain, or what are often referred to as intermediates rather than base commodities or final, end-use products.

Betweenness centrality is about which industries act as a bridge along the shortest path between other industries. In other contexts it quantifies which node controls knowledge flows; in this context it is akin to which industry is a choke point for inputs and outputs. The distributions for weighted closeness and betweenness are in Figure 20 and Figure 21 in the appendix.

The closeness distributions for all industries do not change much over time, but the shape does when requirements coefficients are restricted to be above 0.03 [Figure 10]. The closeness values are much lower with the higher coefficient limits, with many near zero, whereas they are more spread out otherwise.

The shift when coefficients are restricted indicates larger connections are less central, in that they have a longer distance on average to other industries, or are less central on average, which demonstrates many base commodities are sold in large quantities. This pattern is broadly true for energy intensive industries as well, although the years diverge above a value of 0.4, with 1997 the smallest fraction, then 2002, followed by 2007. As with all industries, this indicates that keeping only larger connections means the remaining industries are less centrally connected than smaller ones, which also reveals that energy-intensive industries in particular are often bulk commodities.

Petroleum refining [324110] has the maximum closeness values across all years in the energy intensive industries [Table 3]. This is as expected, given the prevalence of petroleum use as an intermediate input. Multiple industries have the smallest closeness values in energy intensive industries across the different tables, pointing to the great variety of bulk commodities made and used in the U.S. supply chain. See Table 7 in the appendix for the list when connections are greater than 0.03.

The energy sector is where most differences show up—the 2007 tables have a distribution that is shifted above the other two. And when connection size is larger, 2002 is also shifted above the 1997 table. Over time it appears that energy sector industries, as measured by closeness, have become more central in the overall industrial structure. Another way to say this is that the average distance between energy sector industries and all other industries has gotten smaller between 1997 and 2007

This could be due to the proliferation of non-conventional energy production and the various different products it yields. For example, gas fractionation facilities now produce dry gas [methane] and a suite of hydrocarbon gas liquids [ethane, propane, etc.].

The electric power generation, transmission, and distribution industry [221100] has the maximum values for closeness in 1997 and 2002, and electric power generation and petroleum refining [324110] have the maximum values in 2007 [Table 3]. It is unsurprising that these two industries are the most central considering how important each is to the production processes of others. Drilling oil and gas wells has the smallest closeness value across all years, because it is essentially the start of one of the most critical U.S. supply chains.

Betweenness distributions are relatively similar for all industries and the energy sector, although they do shift up slightly when coefficient values are restricted to be above 0.03. This indicates that industries with larger links are those that are more likely to be bridges between industries. For energy intensive industries, iron and steel mills [331111/331110] have the largest betweenness values—this industry is the most important bridge in the sector [Table 3].

Energy sector industries have more substantial differences in their betweenness distributions over time, and when coefficient values are restricted [Figure 11]. Here, the 2007 total requirements tables have betweenness values in the energy sector that are shifted down and to the left of the other two years, and this is followed by the 2002 tables when coefficient values must be large.

Interestingly, even though closeness centrality appears to increase in the 2007 tables [i.e. shorter distances on average], betweenness centrality is lower on average. Industries within the energy sector are less likely to be choke points than in the past on average.

The maximum betweenness values vary over the tables: oil and gas extraction [211000] in 1997, petroleum refining [324110] in 2002, and coal mining [212100] in 2007 [Table 3]. Drilling oil and gas wells [213111] has the lowest betweenness values across all years.

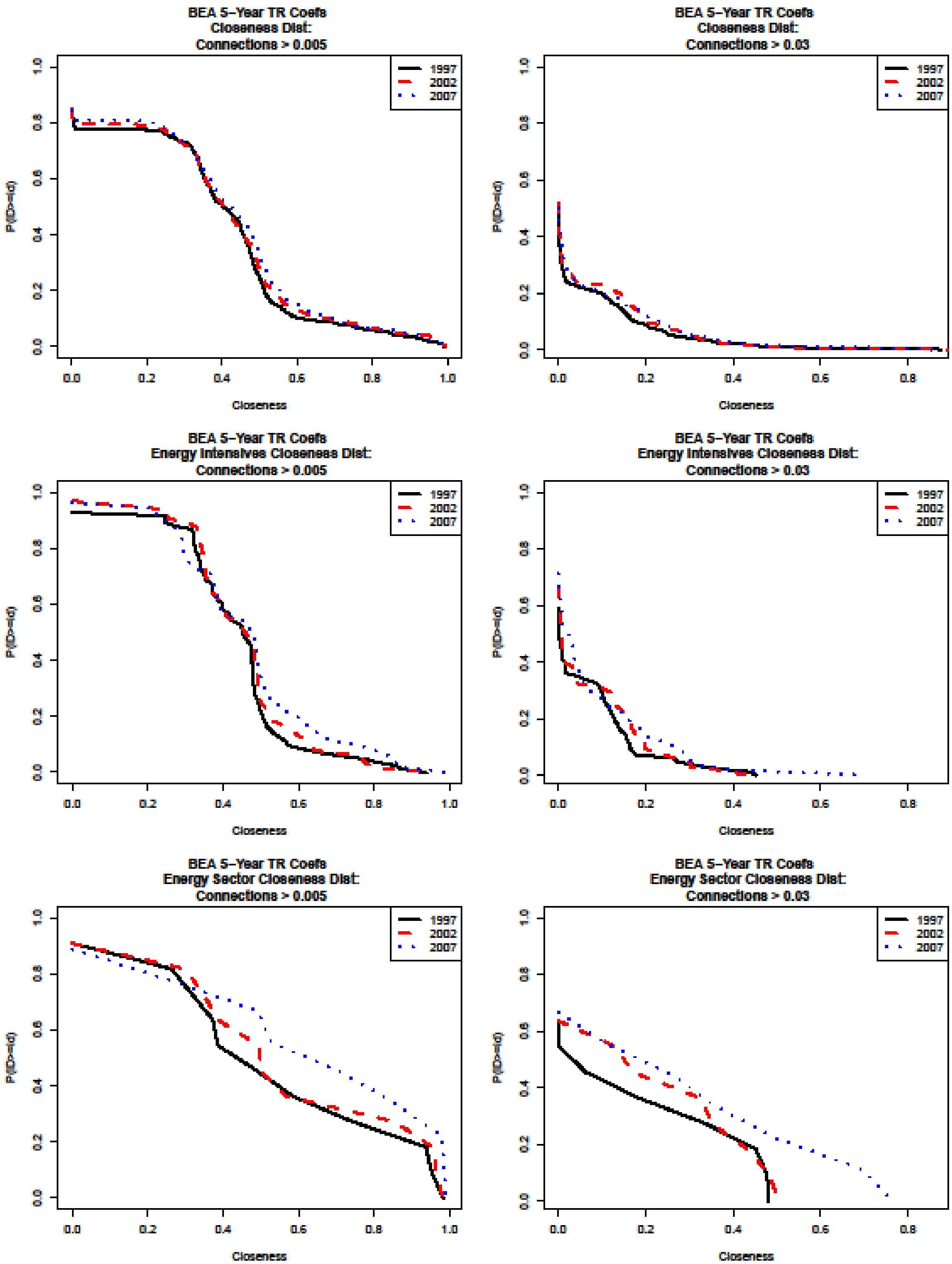
This may again reflect consolidation. Extractive industries were previously supplying products across the supply chain, but after consolidation products are in-house and sold as a single “link” further along the supply chain, conceivably to fewer outside industries.

Table 3: NAICS codes of industries with the maximum and minimum closeness and betweenness from 1997-2007

Statistic	Energy Intensive			Energy Sector		
	1997	2002	2007	1997	2002	2007
Max cl	324110	324110	324110	221100	221100	Various
Min cl	Various	Various	Various	213111	213111	213111
Max bt	331111	331110	331110	211000	324110	212100
Min bt	Various	Various	Various	213111	213111	213111

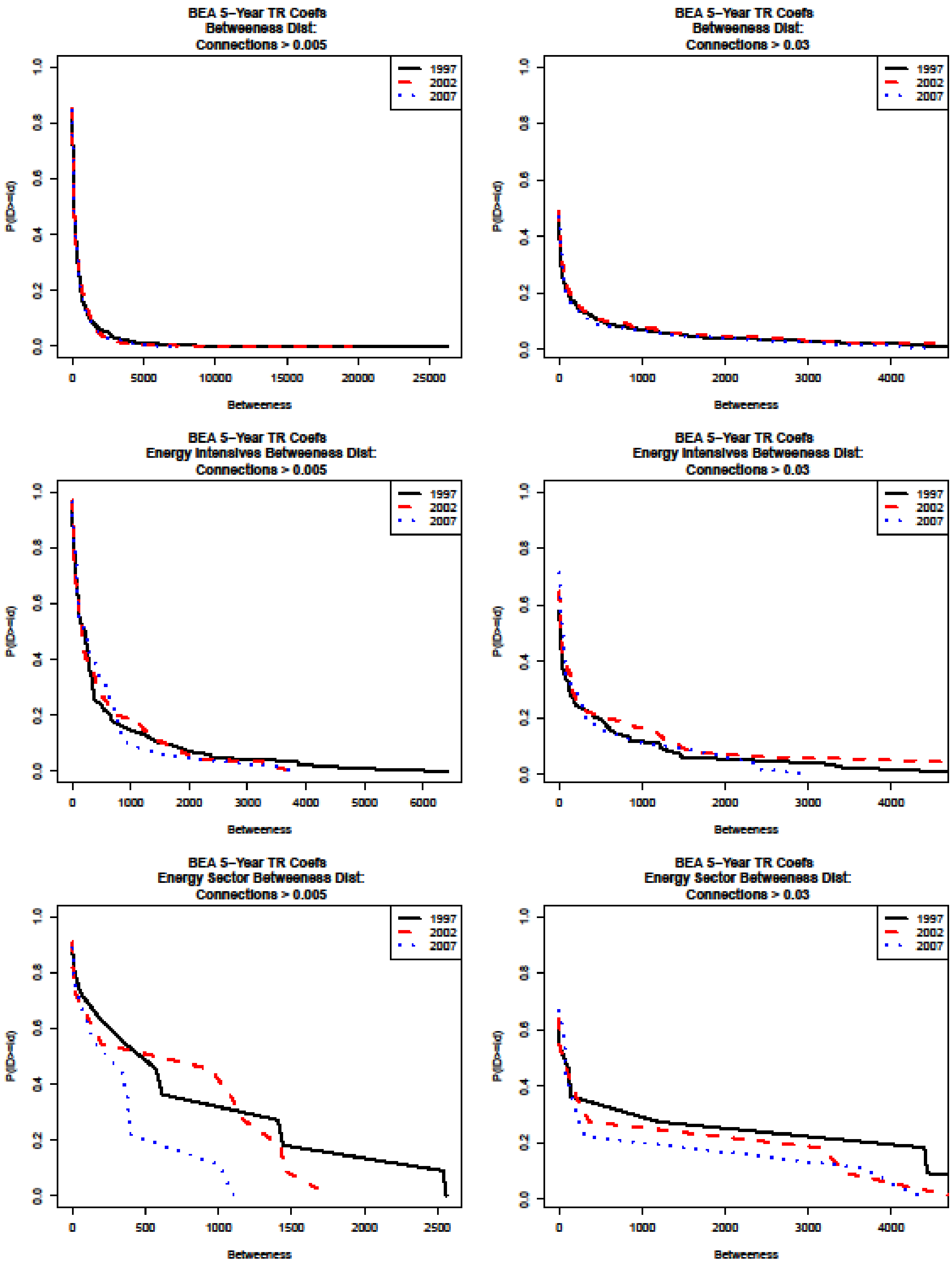
Source: BEA, Author calculations

Figure 10: Closeness distributions of U.S. total requirements tables/subsets in 1997, 2002, 2007



Source: BEA, Author calculations

Figure 11: Betweenness distributions of U.S. total requirements tables/subsets in 1997, 2002, 2007



Source: BEA, Author calculations

4. Conclusion

The density of relationships between industries and their centrality all influence how changes in any one industry impact others (Jackson, 2014; 3-4). And these complex links are well-suited to analysis using the tools of network analysis. In this paper we describe portions of the U.S input-output tables through such tools—focusing on industries that are either energy intensive or part of the energy sector.

We first represent both energy intensive industries and the energy sector visually through network diagrams. Next, we show that the energy sector is generally more densely connected than either energy intensives or all industries over those years, and is more likely to have groups of three sub-sectors all linked as well.

We then move to the level of individual industries within the broad sectors and find that energy intensive industries have the most in-coming connections on average for these tables. Energy sector ones have fewer, but the number grows over time, as do outgoing connections. Other measures of centrality—closeness and betweenness—vary over time for both the energy sector and energy intensives. In terms of specific industries, petroleum refining and electricity generation stand out for their centrality, drilling oil and gas wells for its lack of centrality.

Our analysis in this paper has been purely descriptive: we have shown network diagrams and presented summary statistics for energy-related industries over time. But we have not attempted to reconstruct any portion of the input-output tables using models for random graphs, which is a promising area for future research. These models can be used to test for the importance of certain characteristics in the network, recreate different network properties, or assess the impact of different counterfactual scenarios (Kolaczyk and Csárdi, 2014; 69).

Some might say we have a hammer [network analysis] so now everything looks like a nail. After all, the same information is available directly in the BEA's input-output tables. But looking at the tables differently—in our case through visualization—can be a welcome change of perspective. 'The greatest value of such a picture,' to paraphrase the American mathematician John Tukey, 'is when it forces us to notice what we never expected to see.'

5. References

Carvalho, Vasco M. (2014). “From Micro to Macro via Production Networks,” *Journal of Economic Perspectives*, 28(4), p. 23-48. URL: https://www.aeaweb.org/articles.php?doi=10.1257/jep.28.4.23.

Gretton, Paul (2013). “On Input-Output Tables: Uses and Abuses,” Staff Research Note, Australian Productivity Commission, 2013. URL: http://www.pc.gov.au/research/completed/input-output-tables/input-output-tables.pdf.

Kolaczyk, Eric D. and Gábor Csárdi (2014). *Statistical Analysis of Network Data with R*, 1st Edition, Springer.

Miller, Ronald E. and Peter D. Blair (2009). *Input-Output Analysis: Foundations and Extensions,* 2nd Edition, Cambridge University Press.

Jackson, Matthew O. (2014). “Networks in the Understanding of Economic Behaviors,” *Journal of Economic Perspectives*, 28(4), p. 3-22. URL: https://www.aeaweb.org/articles.php?f=s&doi=10.1257/jep.28.4.3.

Steinbeck Scott R. (2004). “Using Ready-Made Input-Output Models to Estimate Backward-Linkage Effects of Exogenous Output Shocks,” *The Review of Regional Studies*, 34(1), p. 57-71. URL: https://ideas.repec.org/a/rre/publsh/v34y2004i1p57-71.html.

Streitwieser, Mary L. (2010). “Measuring the Nation’s Economy: An Industry Perspective | A Primer on BEA’s Industry Accounts,” BEA Working Paper 58, U.S. Bureau of Economic Analysis. URL: http://www.bea.gov/papers/pdf/an_industry_perspective_a_primer_on_beas_industry_accounts.pdf.

6. Appendix

Tables in the appendix list energy intensive industries in 1997, 2002, and 2007 [Table 4]; list energy sector industries in 1997, 2002, and 2007 [Table 5]; display industries with the minimum and maximum in and out-degrees in 1997, 2002, and 2007 for energy intensive industries and the energy sector when connections are larger than 0.03 [Table 6]; and display industries with the minimum and maximum closeness in 1997, 2002, and 2007 for energy intensive industries and the energy sector when connections are larger than 0.03 [Table 7]. The figures complement the main text.

Table 4: List of energy intensive industries

1997		2002		2007	
NAICS Code	**Industry Name**	**NAICS Code**	**Industry Name**	**NAICS Code**	**Industry Name**
311111	Dog and cat food manufacturing	311111	Dog and cat food manufacturing	311111	Dog and cat food manufacturing
311119	Other animal food manufacturing	311119	Other animal food manufacturing	311119	Other animal food manufacturing
311211	Flour milling	311210	Flour milling and malt manufacturing	311210	Flour milling and malt manufacturing
311212	Rice milling	311221	Wet corn milling	311221	Wet corn milling
311213	Malt manufacturing	311225	Fats and oils refining and blending	311225	Fats and oils refining and blending
311221	Wet corn milling	311230	Breakfast cereal manufacturing	311230	Breakfast cereal manufacturing
311222	Soybean processing	311313	Beet sugar manufacturing	311300	Sugar and confectionery product manufacturing
311223	Other oilseed processing	311320	Chocolate and confectionery manufacturing from cacao beans	311410	Frozen food manufacturing
311225	Fats and oils refining and blending	311330	Confectionery manufacturing from purchased chocolate	311420	Fruit and vegetable canning, pickling, and drying
311230	Breakfast cereal manufacturing	311340	Nonchocolate confectionery manufacturing	311513	Cheese manufacturing

311310	Sugar manufacturing	311410	Frozen food manufacturing	311514	Dry, condensed, and evaporated dairy product manufacturing
311320	Confectionery manufacturing from cacao beans	311420	Fruit and vegetable canning, pickling, and drying	311520	Ice cream and frozen dessert manufacturing
311330	Confectionery manufacturing from purchased chocolate	311513	Cheese manufacturing	311615	Poultry processing
311340	Nonchocolate confectionery manufacturing	311514	Dry, condensed, and evaporated dairy product manufacturing	311700	Seafood product preparation and packaging
311410	Frozen food manufacturing	311520	Ice cream and frozen dessert manufacturing	311810	Bread and bakery product manufacturing
311420	Fruit and vegetable canning and drying	311615	Poultry processing	311910	Snack food manufacturing
311511	Fluid milk manufacturing	311700	Seafood product preparation and packaging	311920	Coffee and tea manufacturing
311512	Creamery butter manufacturing	311810	Bread and bakery product manufacturing	311930	Flavoring syrup and concentrate manufacturing
311513	Cheese manufacturing	311820	Cookie, cracker, and pasta manufacturing	311940	Seasoning and dressing manufacturing
311514	Dry, condensed, and evaporated dairy products	311830	Tortilla manufacturing	311990	All other food manufacturing
311520	Ice cream and frozen dessert manufacturing	311910	Snack food manufacturing	322110	Pulp mills
311611	Animal, except poultry, slaughtering	311920	Coffee and tea manufacturing	322120	Paper mills
311612	Meat processed from carcasses	311930	Flavoring syrup and concentrate manufacturing	322130	Paperboard mills
311613	Rendering and meat byproduct processing	311940	Seasoning and dressing manufacturing	322210	Paperboard container manufacturing

311615	Poultry processing	311990	All other food manufacturing	322220	Paper bag and coated and treated paper manufacturing
311700	Seafood product preparation and packaging	322110	Pulp mills	322230	Stationery product manufacturing
311813	Frozen cakes and other pastries manufacturing	322120	Paper mills	322291	Sanitary paper product manufacturing
311821	Cookie and cracker manufacturing	322130	Paperboard mills	322299	All other converted paper product manufacturing
311822	Mixes and dough made from purchased flour	322210	Paperboard container manufacturing	324110	Petroleum refineries
311823	Dry pasta manufacturing	322230	Stationery product manufacturing	324121	Asphalt paving mixture and block manufacturing
311830	Tortilla manufacturing	322291	Sanitary paper product manufacturing	324122	Asphalt shingle and coating materials manufacturing
311911	Roasted nuts and peanut butter manufacturing	322299	All other converted paper product manufacturing	324190	Other petroleum and coal products manufacturing
311919	Other snack food manufacturing	324110	Petroleum refineries	325110	Petrochemical manufacturing
311920	Coffee and tea manufacturing	324121	Asphalt paving mixture and block manufacturing	325120	Industrial gas manufacturing
311930	Flavoring syrup and concentrate manufacturing	324122	Asphalt shingle and coating materials manufacturing	325130	Synthetic dye and pigment manufacturing
311941	Mayonnaise, dressing, and sauce manufacturing	324191	Petroleum lubricating oil and grease manufacturing	325180	Other basic inorganic chemical manufacturing
311942	Spice and extract manufacturing	324199	All other petroleum and coal products manufacturing	325190	Other basic organic chemical manufacturing
311990	All other food manufacturing	325110	Petrochemical manufacturing	325211	Plastics material and resin manufacturing
322110	Pulp mills	325120	Industrial gas manufacturing	325310	Fertilizer manufacturing

322210	Paperboard container manufacturing	325130	Synthetic dye and pigment manufacturing	325320	Pesticide and other agricultural chemical manufacturing
322225	Flexible packaging foil manufacturing	325181	Alkalies and chlorine manufacturing	327200	Glass and glass product manufacturing
322226	Surface-coated paperboard manufacturing	325182	Carbon black manufacturing	327310	Cement manufacturing
322231	Die-cut paper office supplies manufacturing	325188	All other basic inorganic chemical manufacturing	331110	Iron and steel mills and ferroalloy manufacturing
322232	Envelope manufacturing	325190	Other basic organic chemical manufacturing	331200	Steel product manufacturing from purchased steel
322233	Stationery and related product manufacturing	325211	Plastics material and resin manufacturing	331314	Secondary smelting and alloying of aluminum
322291	Sanitary paper product manufacturing	325212	Synthetic rubber manufacturing	31122A	Soybean and other oilseed processing
322299	All other converted paper product manufacturing	325220	Artificial and synthetic fibers and filaments manufacturing	31151A	Fluid milk and butter manufacturing
324110	Petroleum refineries	325310	Fertilizer manufacturing	31161A	Animal (except poultry) slaughtering, rendering, and processing
324121	Asphalt paving mixture and block manufacturing	325320	Pesticide and other agricultural chemical manufacturing	3118A0	Cookie, cracker, pasta, and tortilla manufacturing
324122	Asphalt shingle and coating materials manufacturing	327211	Flat glass manufacturing	3252A0	Synthetic rubber and artificial and synthetic fibers and filaments manufacturing
324191	Petroleum lubricating oil and grease manufacturing	327212	Other pressed and blown glass and glassware manufacturing	33131A	Alumina refining and primary aluminum production

324199	All other petroleum and coal products manufacturing	327213	Glass container manufacturing	33131B	Aluminum product manufacturing from purchased aluminum
325110	Petrochemical manufacturing	327215	Glass product manufacturing made of purchased glass		
325120	Industrial gas manufacturing	327310	Cement manufacturing		
325130	Synthetic dye and pigment manufacturing	331110	Iron and steel mills and ferroalloy manufacturing		
325180	Other basic inorganic chemical manufacturing	331200	Steel product manufacturing from purchased steel		
325190	Other basic organic chemical manufacturing	331314	Secondary smelting and alloying of aluminum		
325211	Plastics material and resin manufacturing	31122A	Soybean and other oilseed processing		
325212	Synthetic rubber manufacturing	31131A	Sugar cane mills and refining		
325221	Cellulosic organic fiber manufacturing	31151A	Fluid milk and butter manufacturing		
325222	Noncellulosic organic fiber manufacturing	31161A	Animal (except poultry) slaughtering, rendering, and processing		
325311	Nitrogenous fertilizer manufacturing	32222A	Coated and laminated paper, packaging paper and plastics film manufacturing		
325312	Phosphatic fertilizer manufacturing	32222B	All other paper bag and coated and treated paper manufacturing		
325314	Fertilizer, mixing only, manufacturing	33131A	Alumina refining and primary aluminum production		

325320	Pesticide and other agricultural chemical manufacturing	33131B	Aluminum product manufacturing from purchased aluminum
327213	Glass container manufacturing		
327310	Cement manufacturing		
331111	Iron and steel mills		
331112	Ferroalloy and related product manufacturing		
331210	Iron, steel pipe and tube from purchased steel		
331221	Rolled steel shape manufacturing		
331222	Steel wire drawing		
331311	Alumina refining		
331312	Primary aluminum production		
331314	Secondary smelting and alloying of aluminum		
331315	Aluminum sheet, plate, and foil manufacturing		
331316	Aluminum extruded product manufacturing		
331319	Other aluminum rolling and drawing		
31181A	Bread and bakery product, except frozen, manufacturing		
3221A0	Paper and paperboard mills		
32222A	Coated and laminated paper and packaging materials		

32222B	Coated and uncoated paper bag manufacturing
32721A	Glass and glass products, except glass containers

Table 5: List of energy sector industries

1997		2002		2007	
NAICS Code	**Industry Name**	**NAICS Code**	**Industry Name**	**NAICS Code**	**Industry Name**
211000	Oil and gas extraction	211000	Oil and gas extraction	211000	Oil and gas extraction
212100	Coal mining	212100	Coal mining	212100	Coal mining
213111	Drilling oil and gas wells	213111	Drilling oil and gas wells	213111	Drilling oil and gas wells
213112	Support activities for oil and gas operations	213112	Support activities for oil and gas operations	221100	Electric power generation, transmission, and distribution
221100	Power generation and supply	221100	Electric power generation, transmission, and distribution	221200	Natural gas distribution
221200	Natural gas distribution	221200	Natural gas distribution	324110	Petroleum refineries
324110	Petroleum refineries	324110	Petroleum refineries	324190	Other petroleum and coal products manufacturing
324191	Petroleum lubricating oil and grease manufacturing	324191	Petroleum lubricating oil and grease manufacturing	333611	Turbine and turbine generator set units manufacturing
324199	All other petroleum and coal products manufacturing	324199	All other petroleum and coal products manufacturing	21311A	Other support activities for mining
333611	Turbine and turbine generator set units manufacturing	333611	Turbine and turbine generator set units manufacturing		
21311A	Support activities for other mining	21311A	Support activities for other mining		

Table 6: NAICS codes of industries with the maximum and minimum in and out-degrees from 1997-2007 when connections are above 0.03

Statistic	Energy Intensive			Energy Sector		
	1997	2002	2007	1997	2002	2007
Max Ideg	Various	311320	311410	324191	21311A	333611
Min Ideg	311930	311930	324110	211000	Various	213111
Max Odeg	Various	324110	324110	221100	221100	211000
Min Odeg	Various	Various	Various	Various	Various	Various

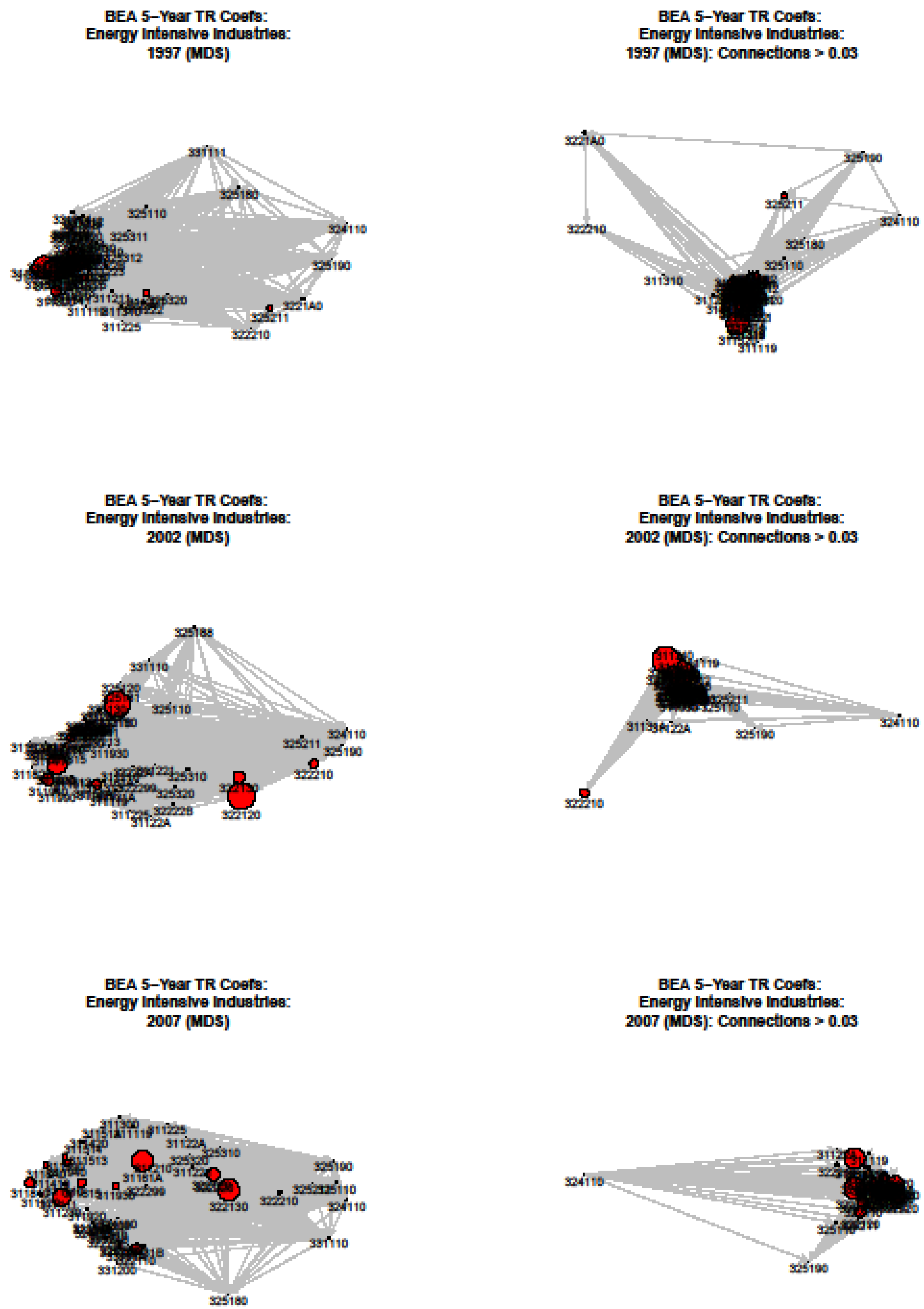
Source: BEA, Author calculations

Table 7: NAICS codes of industries with the maximum and minimum closeness and betweenness from 1997-2007 when connections are above 0.03

Statistic	Energy Intensive			Energy Sector		
	1997	2002	2007	1997	2002	2007
Max cl	324110	324110	324110	221100	221100	211000
Min cl	Various	Various	Various	Various	Various	Various
Max bt	331111	327215	325190	221100	221100	212100
Min bt	Various	Various	Various	Various	Various	Various

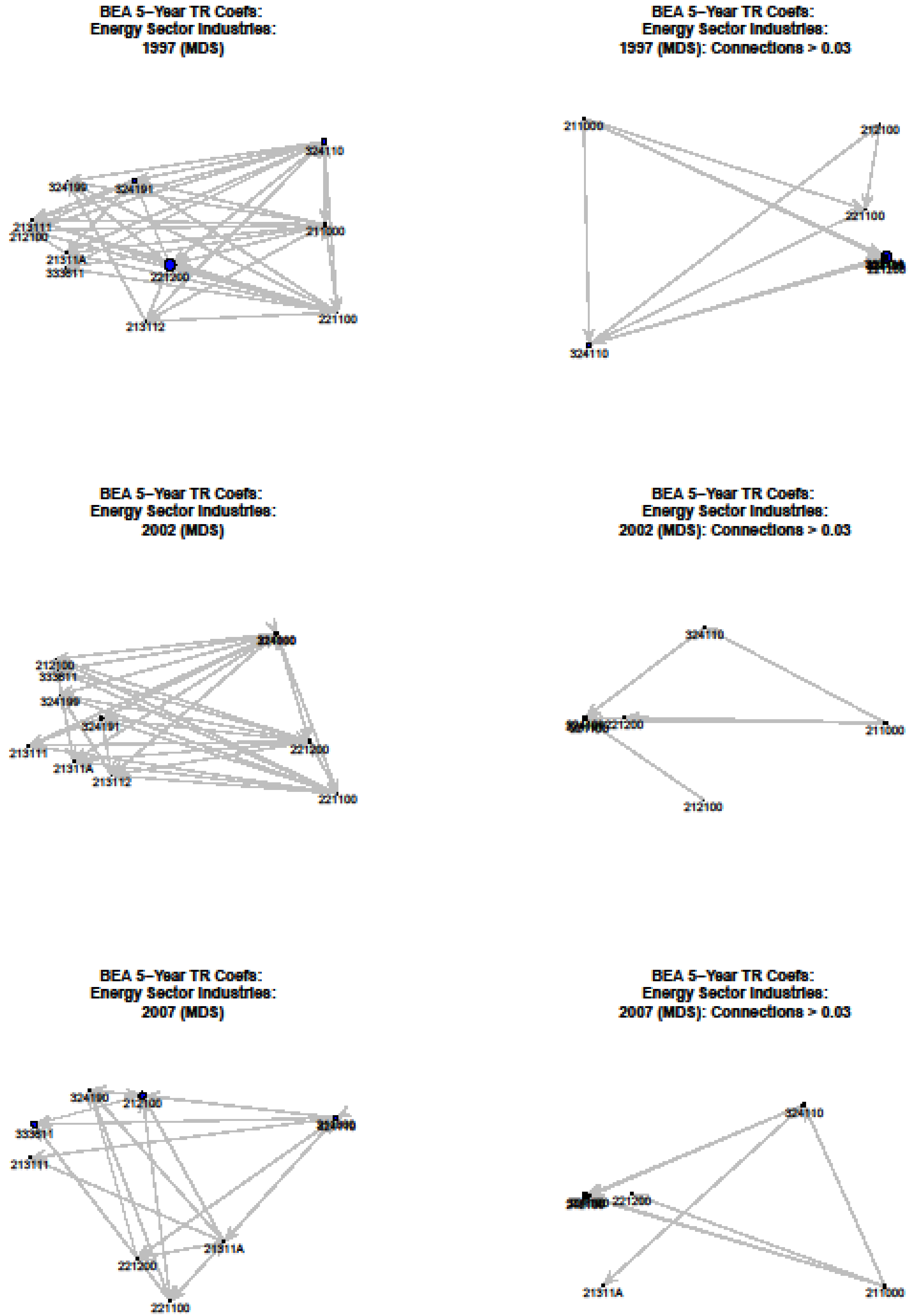
Source: BEA, Author calculations

Figure 12: Network plots of the U.S. total requirements tables for energy intensive industries in 1997, 2002, and 2007 using the MDS algorithm



Source: BEA, Author calculations

Figure 13: Network plots of the U.S. total requirements tables for energy sector industries in 1997, 2002, and 2007 using the FR algorithm

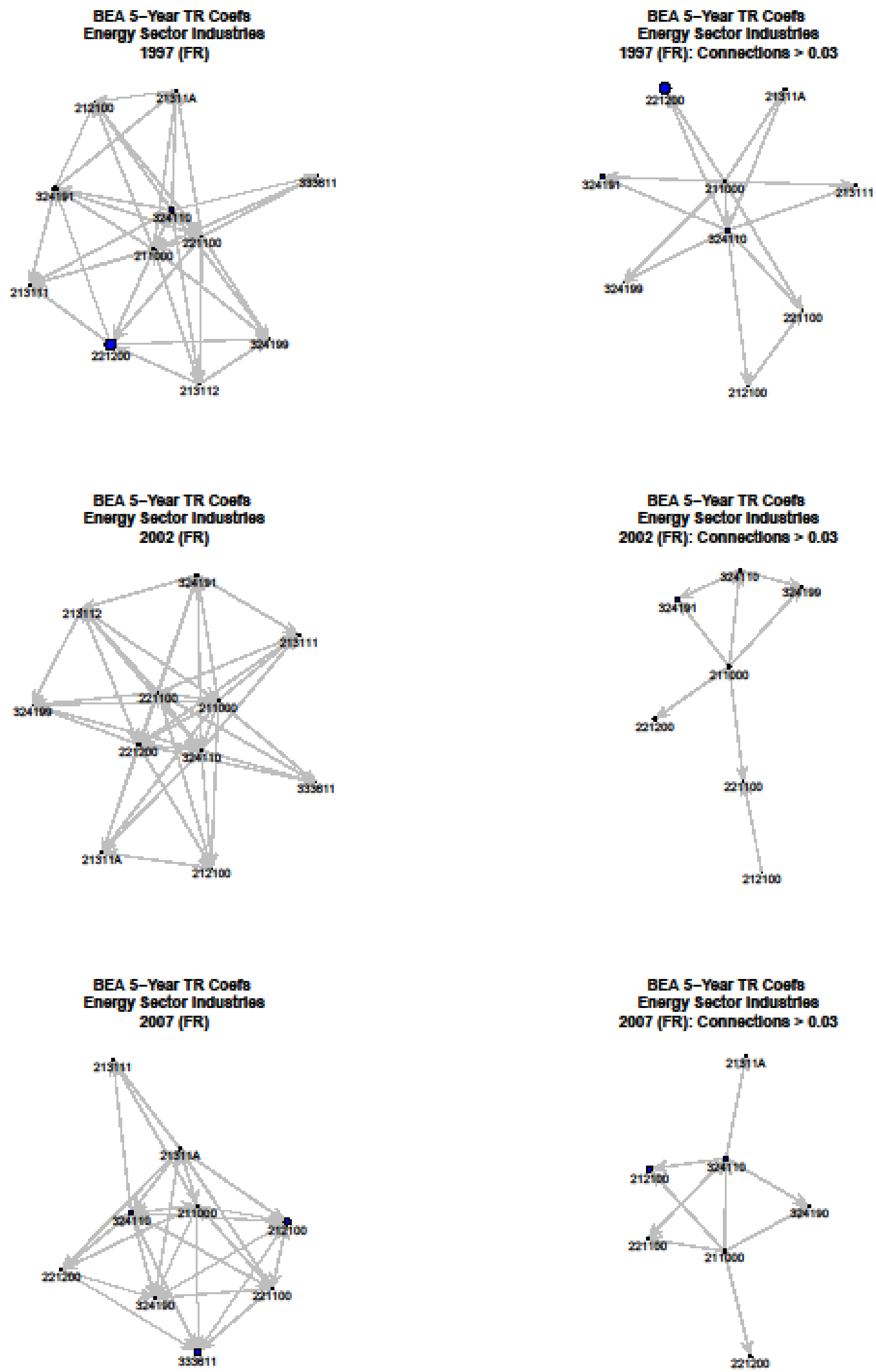


Source: BEA, Author calculations

Figure 14: Network plots of the U.S. total requirements tables for energy intensive industries in 1997, 2002, and 2007 using the FR algorithm

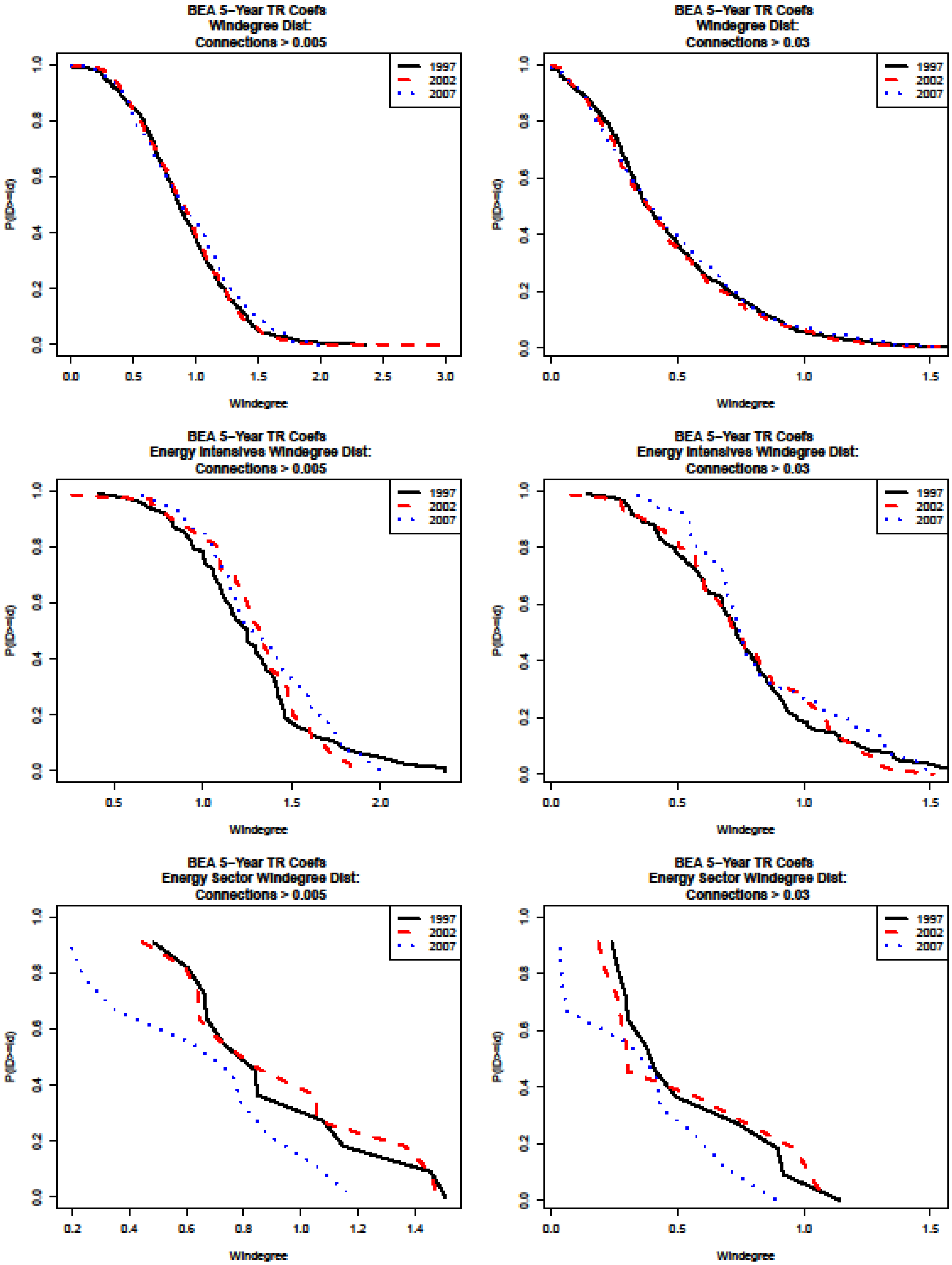
Source: BEA, Author calculations

Figure 15: Network plots of the U.S. total requirements tables for energy sector industries in 1997, 2002, and 2007 using the FR algorithm



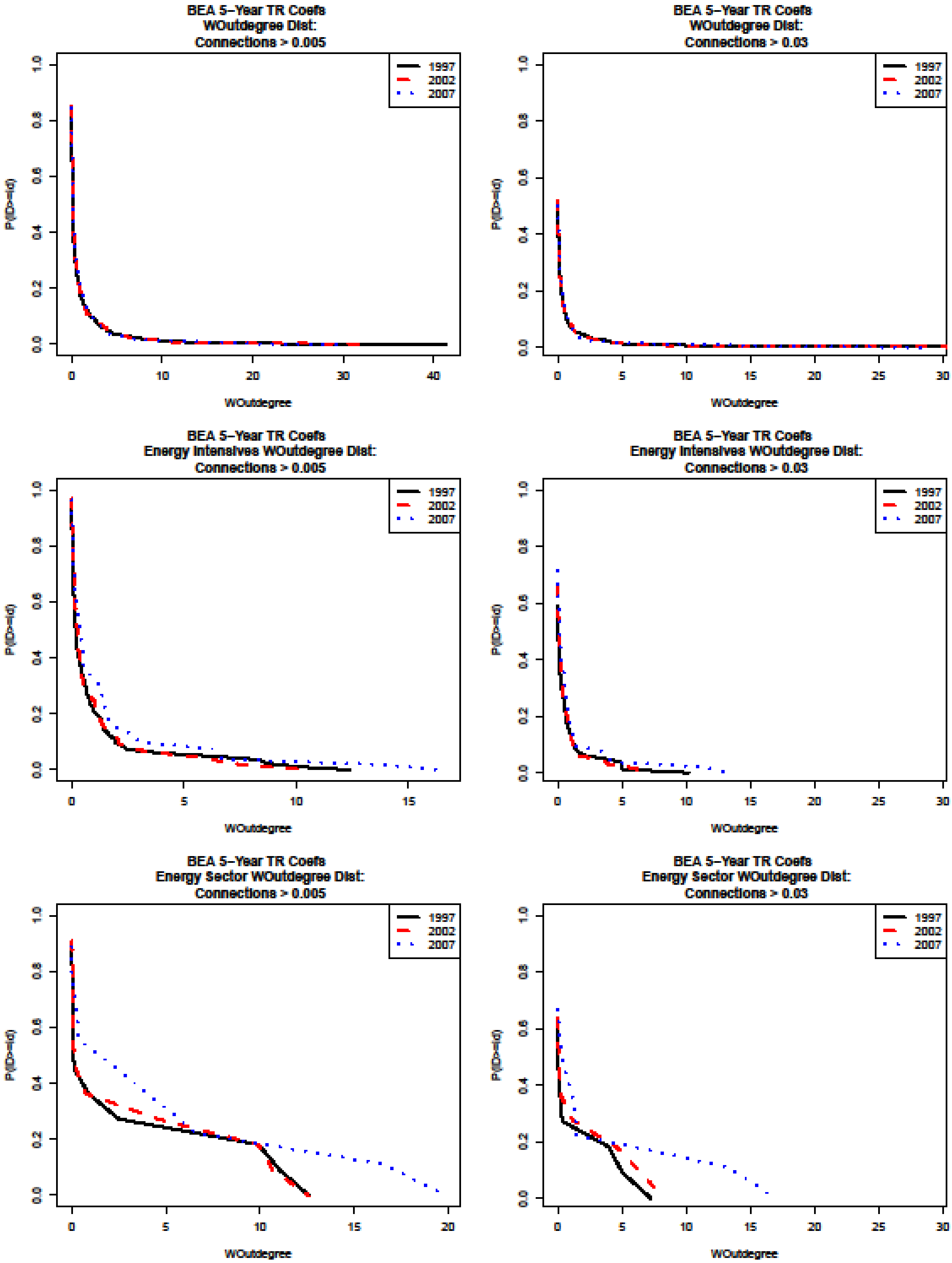
Source: BEA, Author calculations

Figure 16: Weighted in-degree distributions of U.S. total requirements tables and subsets in 1997, 2002, and 2007



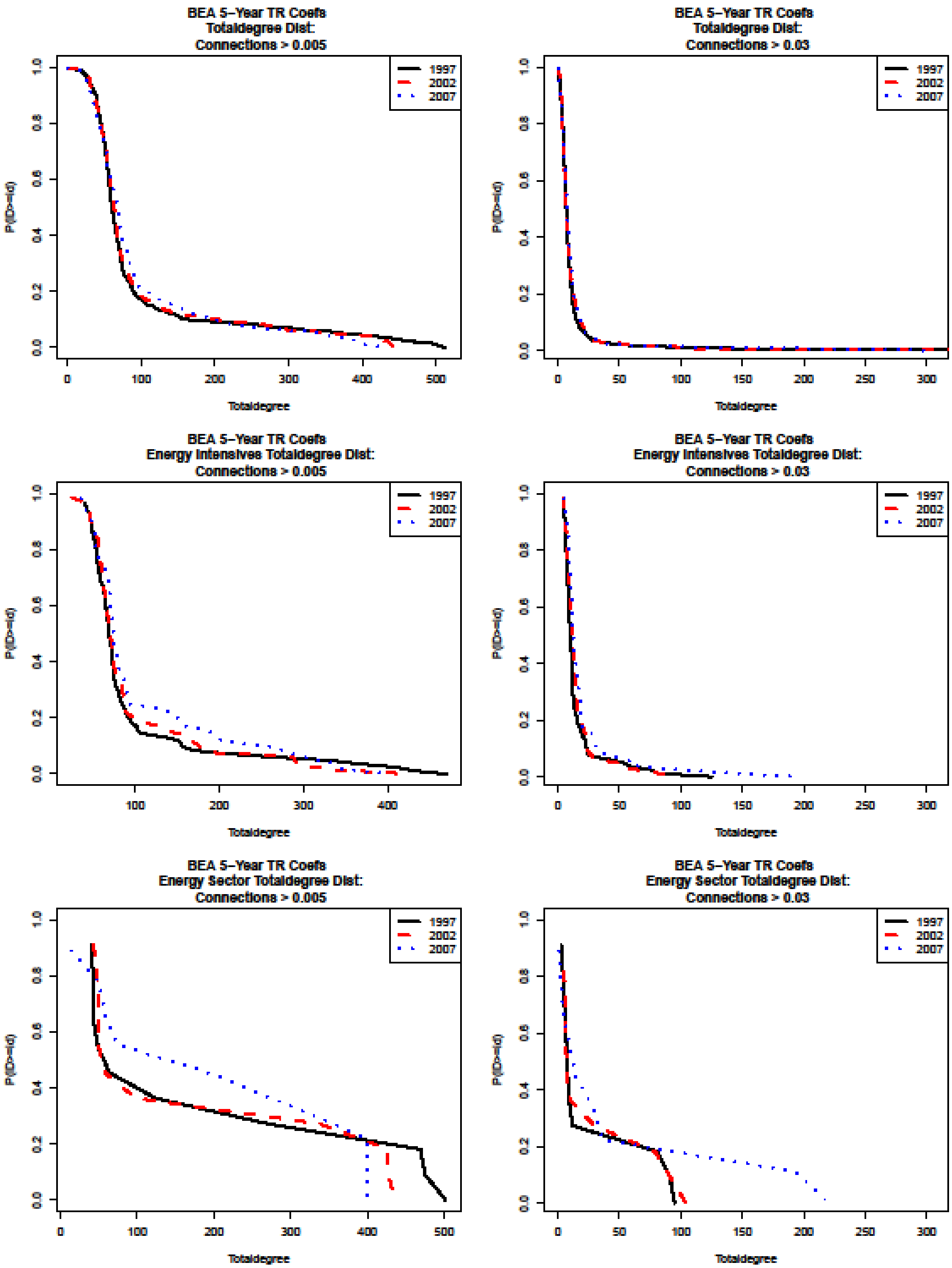
Source: BEA, Author calculations

Figure 17: Weighted out-degree distributions of U.S. total requirements tables and subsets in 1997, 2002, and 2007



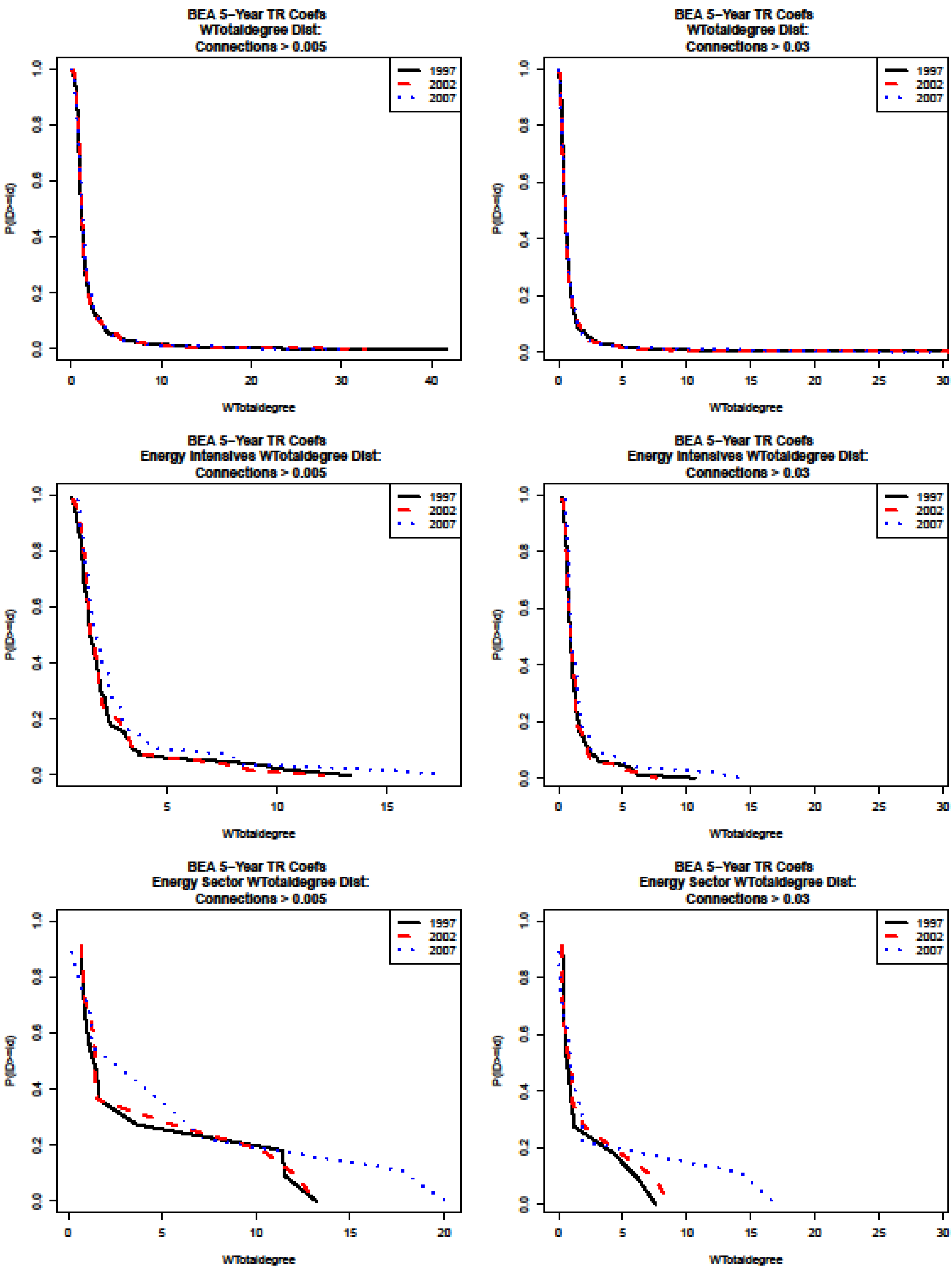
Source: BEA, Author calculations

Figure 18: Total-degree distributions of U.S. total requirements tables and subsets in 1997, 2002, and 2007



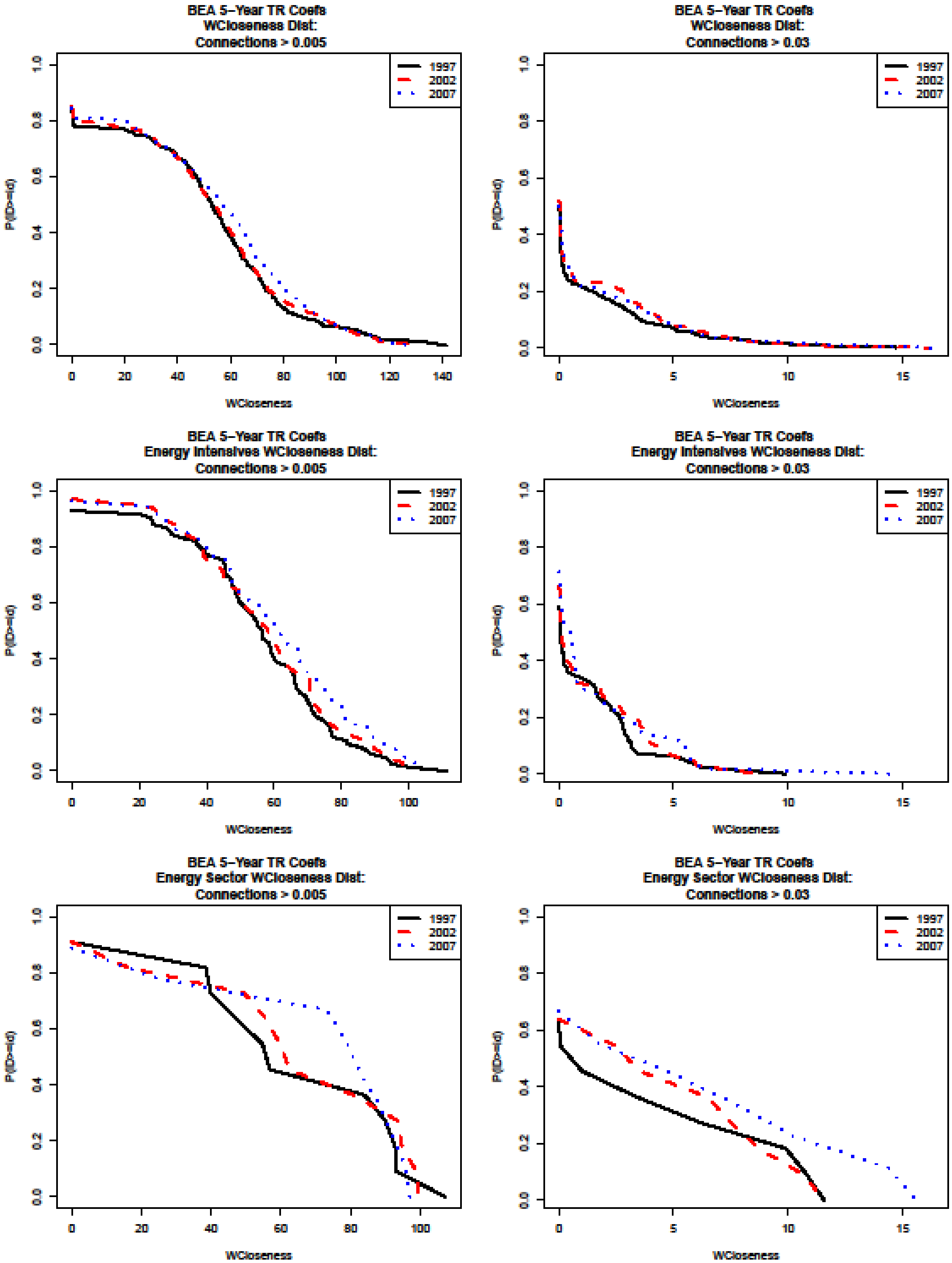
Source: BEA, Author calculations

Figure 19: Weighted total-degree distributions of U.S. total requirements tables and subsets in 1997, 2002, and 2007



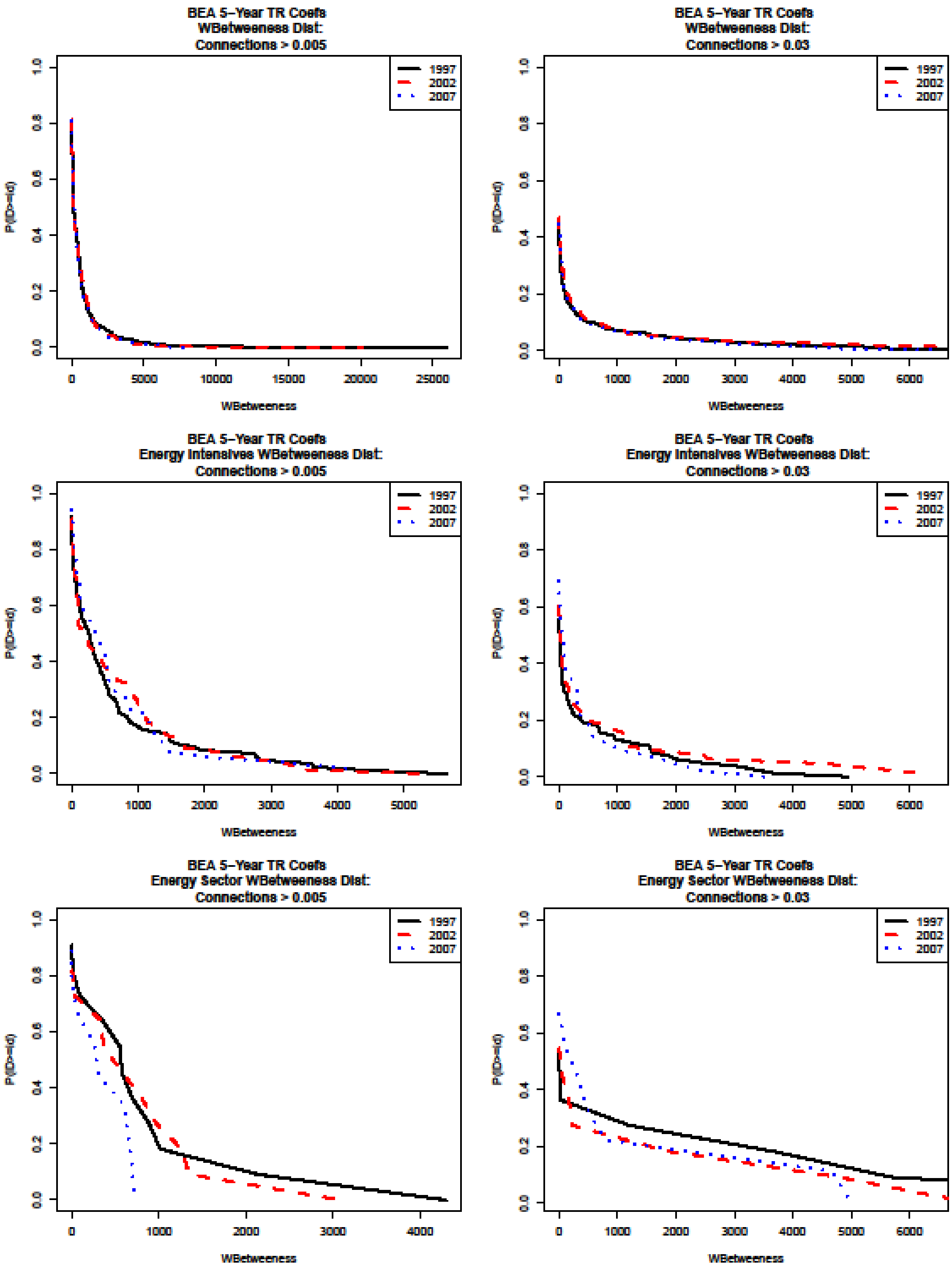
Source: BEA, Author calculations

Figure 20: Weighted closeness distributions of U.S. total requirements tables and subsets in 1997, 2002, and 2007



Source: BEA, Author calculations

Figure 21: Weighted betweenness distributions of U.S. total requirements tables and subsets in 1997, 2002, and 2007



Source: BEA, Author calculations