Classroom Exercise To Motivate Input-Output and Matrix Based Methods

The next two chapters motivate quantitatively-driven methods to support life cycle studies. Both input-output and process matrix methods rely on linear algebra and matrix-based methods to assist computational efforts. Armed with the introduction to life cycles and process flow diagram approaches, the reader is prepared to learn about these matrix-based methods.

A team of researchers at Carnegie Mellon University has created a game-like simulation to assist with this effort. It involves small groups simulating the production of four goods, each of which has a small number of input and output flows. However the four goods have interdependent flows. A key learning objective of the simulation is to realize how interdependent flows lead to process flow diagrams which are dependent on each other, and how addressing that dependency requires additional demand and estimation of effects upstream. Through the exercise, the underpinnings of the matrix approach are revealed. In the end, using matrices to solve such problems is shown to be much more straightforward, and avoids various potential math errors.

This exercise has been designed and tested with audiences ranging from middle school students through corporate executives, all of which are in the process of learning about LCA. Ideal group sizes are about 4 persons, although slightly smaller or larger groups work well. **We strongly suggest that this exercise be done, ideally in a classroom or small group setting, before proceeding with subsequent chapters.** Post-assessments have shown a significant increase in understanding of these topics amongst those exposed to this simulation.

The whole exercise, including the introduction and motivation, as well as blank copies of the various purchase order and tracking forms, was previously published (Hawkins 2009) and has been made freely available through the generous support of the *Journal of Industrial Ecology*. A direct link is available via the textbook website, under E-resources for Chapter 8.

Chapter 8 : LCA Screening via Economic Input-Output Models

The construction and use of economic input-output based LCA models, and the powerful screening capabilities they provide are introduced in this chapter. As described in Chapter 5, process-based LCA models are 'bottom up' in type, and are defined by the scope laid out by the analyst. Economic input-output LCA models can be thought of as being a 'top-down' type because they give a holistic estimate of resources needed for a product across the economy. We recommend our separately published classroom simulation on input-output LCA models (Hawkins 2009). This simulation exercise has been developed explicitly to make these kinds of models understandable and to help you appreciate their strengths and limitations. We will also describe the mathematical structure of the EIO-LCA model. Those intending to use economic input-output tables in their own LCA studies are highly encouraged to read the Advanced Material at the end of this chapter.

Learning Objectives for the Chapter

At the end of this chapter, you should be able to:

1. Describe how economic sector data are organized into input-output tables
2. Compute direct, indirect, and total effects using an input-output model
3. Assess how an input-output model might be used as a screening tool.

Input-Output Tables and Models

In the 1930s, economist Wassily Leontief developed an **economic input–output table** of the United States economy and a system of equations to use them in models (Leontief, 1986). His model represented the various inputs required to produce a unit of output in each **economic sector** based on surveyed census data of purchases and sales of industries. By assembling a table describing all of the major economic sectors, he was able to trace all of the economic purchases needed to produce outputs in each sector, all the way back to the beginning when raw materials are extracted. The result was a comprehensive model of the U.S. economy. For this work, Leontief received the Nobel Prize in Economics in 1973.

What are economic sectors?

Sectors are groups of companies with similar products. Given the model, there may be a single sector for all manufactured goods, or many separate sectors, for everything from electricity (large economic output) to tortillas (relatively small output). There are various national and global systems for categorizing sectors, leading to differences in the number of sectors used and reports in the IO tables.

An economic input–output (EIO, or just IO) table divides an entire economy into distinct economic sectors. The tables can represent total sales from one sector to others, purchases from one sector, or the amount of purchases from one sector to produce a dollar of output.

Input-output models were popular in the mid-20th century for high-level economic planning purposes. They were used so that governments could better understand the requirements of, and plan for, activities like war planning, procurement, effects of disarmament, or economic requirements for building infrastructure such as roads. As will be seen below, economic input-output models are vital for developing national economic accounts, and can also be used for environmental life cycle assessment.

Vectors, matrices, and notation

A vector is a one-dimensional set of values (called elements) referenced by an index. If a vector is named X then X_1 is the first element, X_2 is the second element, ... , and X_n is the last element. If implemented in a spreadsheet, a vector could be arranged as elements in a row or in a column. In this book, we use upper case *italicized* letters to represent vectors. Individual elements (a row/column entry) are upper case, *italic*, and denoted by an index.

A matrix is a two-dimensional array of values referenced by both a row and column index. (The plural of matrix is matrices.) In a spreadsheet, a matrix would have rows and columns. One of the most popular matrices is an identity matrix (**I**), which has the number one for all elements on the diagonal (where the row and column indices are equal), and zeroes in all other cells. We use upper case **bold** letters to represent matrices. A two-dimensional identity matrix is defined as having these elements:

$$\mathbf{I} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Note in equations multiplying vectors and matrices together, the multiplication sign is omitted.

Figure 8-1 shows the structure of an IO **transactions table**. Each entry Z_{ij} represents the input to column sector j's production process from row sector i. The final column, total output of each sector X, has n elements, each the sum across the Z_{ij} inputs from the other sectors (a.k.a. the intermediate output, O_i) plus the output supplied by the sector directly to the final demand Y_i of consumers. To help explain these different components, the intermediate outputs O are being sold to other producers to make other goods, while final demand is sales to users of the product as is. For example, a tire manufacturer might sell some of its output as intermediate product to an automobile producer for new cars and as final product to consumers as replacement tires. For each of the n columns of the transactions table, the column sum represents the total amount of inputs X to each sector from other sectors. The two X values for each sector are equal. A transactions table is like a spreadsheet of all purchases, thus values for a large economy could be in billions or trillions of currency units.

	Input to sectors				Intermediate output O	Final demand Y	Total output X
Output from sectors	1	2	3	n			
1	Z_{11}	Z_{12}	Z_{13}	Z_{1n}	O_1	Y_1	X_1
2	Z_{21}	Z_{22}	Z_{23}	Z_{2n}	O_2	Y_2	X_2
3	Z_{31}	Z_{32}	Z_{33}	Z_{3n}	O_3	Y_3	X_3
n	Z_{n1}	Z_{n2}	Z_{n3}	Z_{nn}	O_n	Y_n	X_n
Intermediate input I	I_1	I_2	I_3	I_n			
Value added V	V_1	V_2	V_3	V_n		GDP	
Total output X	X_1	X_2	X_3	X_n			

Figure 8-1: Example Structure of an Economic Input–Output Transactions Table

Notes: Matrix entries Z_{ij} are the input to economic sector j from sector i. Total (row) output for each sector i, X_i, is the sum of intermediate outputs used by other sectors, O_i, and final demand by consumers. Total (column) output can also be defined as the sum of intermediate input purchases and value added.

Intermediate input, I, in an IO table is the sum of the inputs coming from other sectors (and is distinct from the identity matrix $\mathbf{I}$). Value Added, V, is defined by economists as the increase in value as a result of a process. In an IO model, it is the increase associated with taking inputs valued in total as I and making output valued X. While value added serves a purpose in ensuring consistency between total inputs and total output, it includes some real aspects of industrial production such as direct labor expenses, profits and taxes. While not shown here, some IO frameworks have a sector representing household activities.

The typical process of making a transactions table involves acquiring data on transactions of an economy through periodic surveys of companies in the various sectors to assess how much economic output they are producing, and which other companies (and from which sectors) they buy from. As you might imagine, these data collection activities can be very time and resource intensive. The methods involve only a sampling of companies surveyed rather than surveying every company. Like methods used for counting population, additional statistical analyses are done to check the results and to ensure representative totals have been estimated. A fundamental concern thus relates to deciding how many sectors to divide the economy into. With fairly little effort one could develop a very coarse model of an economy, e.g., with 10 sectors where agriculture, mining, manufacturing, etc. represent various very aggregated sectors of activity. But such models have very low resolution and answer only a limited number of questions. Thus, all parties have incentive to invest resources in generating tables with higher numbers of sectors so that more detailed analyses are possible.

Since the effort needed to make a table is measured in person-years, often the highest resolution tables (greater number of sectors) of an economy are not made annually. The fact that economic production does not change quickly, i.e., the production recipe for the sectors does not change much from year to year, further supports only periodic need for detailed tables. There are typically annual and **benchmark** tables, where lower resolution (fewer sectors) tables are made annually and higher resolution (more sectors) benchmark tables are made less frequently, such as every 5 years as done in the U.S.

Gross Domestic Product (GDP) is an indicator of output of an economy, measured by the sum of final demands or value added across sectors. Consider the alternative – if final demands and intermediate outputs were both components of the output of the economy, then much of that output would be "double counted". Extending the example above, such an economic measure would count both the value of production of intermediate tires in new cars as well as the value of the new cars that came with those tires. Such an outcome would be undesirable, thus only final demands are counted.

Economic input-output models are developed for all major countries, usually by government agencies such as the US Department of Commerce's Bureau of Economic Analysis (BEA). Their primary use is to help develop national accounts used to estimate economic data such as the Gross Domestic Product (GDP). Most nations routinely develop input–output models with 50-100 sectors, although few are as detailed as the current benchmark year 2002 428-sector model of the United States. (Given the processing time required, the IO tables reporting 2002 values were released in 2007. Tables with data collected in 2007 were published in 2013.) A recurring criticism of using EIO-based models is that they rely on relatively old data due to these lags in release of the economic data. However, as shown in Figure 5-14, available process data tends to be fairly old as well. Not all IO tables are made by government employees. In various developing countries, where expertise in government agencies may not exist to do such work, these same activities could be done by other parties like academic researchers in the home or a foreign country.

For calculation purposes, it is helpful to normalize the IO table to represent the proportional input from each sector for a single dollar of output. This table is calculated by dividing each Z_{ij} entry by the total (column) output of that sector, X_j. We denote the resulting table - with all entries between zero and one - as matrix **A** showing the requirements of other sectors directly required to produce a dollar of output for each sector. When done in this way, **A** is called the **direct requirements table (or matrix)**. It is called "direct" because these purchases happen at the highest level of decision making – i.e., the direct purchases needed to produce automobiles are windshields, tires, and engines.

Example 8-1 illustrates the transformation of an IO table into its corresponding **A** matrix. An economic input–output model is linear, so the effects of a $1,000 purchase from a sector will be exactly ten times greater than the effects of a $100 purchase from the same sector.

Example 8-1: In this example, we will use the methods defined above to develop an **A** matrix.

Assume a transactions table for a simple 2-sector economy (values in billions):

	1	2	Y	X
1	150	500	350	1000
2	200	100	1700	2000
V	650	1400		
X	1000	2000		

Question: What is the direct requirements matrix (**A**) for this economy?

Answer: We use the Z_{ij}/X_j formulation as described above. For example, the 150 and 200 values in column 1 are divided by 1000 (the X value in column 1) and the 500 and 100 are normalized by 2000. Thus:

$$\mathbf{A} = \begin{bmatrix} .15 & .25 \\ .2 & .05 \end{bmatrix}$$

The **A** matrix (and a Leontief model in general) thus represents a series of "**production recipes**" for all of the sectors in the economy. A production recipe is just like a recipe for cooking food, where you are told all of the ingredients needed to prepare a meal and in which numerical amounts. Soon after Leontief won the Nobel Prize, he was quoted as saying "When you make bread, you need eggs, flour, and milk. And if you want more bread, you must use more eggs. There are cooking recipes for all the industries in the economy."

In **A** matrix terms, since all values in a column are normalized by the total sector output, each of the coefficients in the production recipe is fractional. As a hypothetical and simple example, imagine a small economy with only two sectors, like in Example 8-1, which are for electricity generation and coal mining. The production recipe for making \$1 worth of electricity would involve purchasing a fraction of a dollar from the coal mining sector (as well as some electricity). Likewise the production recipe for making \$1 worth of coal would involve purchasing some electricity and coal. This interdependence between sectors is common, and a critical reason why EIO models are so useful in representing systems.

A key benefit of using EIO models is not the organization of the economy into tabular form. It is that the direct requirements table can be used to trace out everything needed in the manufacture of a product going back to the very beginning of the life cycle of that product. One can envision this by considering what direct purchases are needed, then what purchases are needed to produce those direct purchases, and continuing back through the purchasing levels to the initial raw materials obtained through mining or farming (n levels back). If considering the manufacture of windows, a window manufacturing column may show purchases of glass and wood (or metal) framing pieces. The glass manufacturing column would show purchases of sand and other minerals, and the wood framing sector would show purchases of forestry products. If given enough time, one could piece together the total requirements by iteratively looking up the **A** matrix chain for such information.

Algebraically, the required economic purchases in all sectors of the economy required to make any desired output Y is simply a vector of the various sectors' final demand inputs. Thus, the total purchases, X, needed to generate output Y can be calculated as:

$$X = [\mathbf{I} + \mathbf{A} + \mathbf{A}\times\mathbf{A} + \mathbf{A}\times\mathbf{A}\times\mathbf{A} + \ldots]\, Y = \mathbf{I}Y + \mathbf{A}Y + \mathbf{A}^2Y + \mathbf{A}^3Y + \ldots \qquad (8\text{-}1)$$

where X is the vector (or list) of required production, **I** is the identity matrix, **A** is the direct requirements matrix (with rows representing the required inputs from all other sectors to make a unit of output for that row's sector) and Y is the vector of desired output. For example, this model might be applied to represent the various requirements for producing electricity purchased by residences. In Equation 8-1, the summed terms represent the production of the desired output itself, electricity ($\mathbf{I}Y$), as well as contributions from the first tier suppliers, e.g., coal or natural gas ($\mathbf{A}Y$), the second tier suppliers, e.g., equipment used at coal mines ($\mathbf{A}^2Y$), etc. In input-output terminology we refer generally to the $\mathbf{I}Y$ and $\mathbf{A}Y$ terms (i.e., $[\mathbf{I} + \mathbf{A}]Y$) as the **direct purchases** (because those are everything related directly to the decisions made by the operators of the final production facility) and all other $\mathbf{A}^2\mathbf{Y}$, $\mathbf{A}^3\mathbf{Y}$, etc., terms as **indirect purchases** (since those production decisions are made beyond the direct operators). The sum of the direct and indirect purchases is the **total purchases**. Note that the terminology used in IO models may differ from that of other modeling domains (and as introduced in Chapter 4) but is emphasized for consistency. In other domains, direct purchases may only refer to $\mathbf{I}Y$.

IO models that estimate direct and indirect purchases use Equation 8-1 to combine the various production recipes across the supply chain into a total supply chain. That is, a final demand of $20,000 into the automobile manufacturing sector will determine all of the direct ingredients (in dollars) needed to produce the car. One of these direct requirements may be a $2,500 engine. Therefore, the IO model also then (in the $\mathbf{A}^2Y$ term) estimates the ingredients needed to produce the $2,500 engine. In the end, the thousands of overall ingredients needed to produce the car are all included in the total purchases estimate. And all are aggregated into the relevant sectors, even if they occur at different tiers of the supply chain (i.e., purchases of any direct or indirect electricity are all added into a single sectoral total for purchases of electricity). Since IO models by default represent flows across the entire economy, they can be classified as a 'top down'

> **How are economic input-output and process based methods similar?**
>
> Think of each of the sectors of the IO model as a process. In each sector's process, the "inputs" are the economic inputs (as purchased) from all of the other sectors and the "output" is the product of the sector. An IO model is thus a linear system of individual economic process models.

approach. Such methods give high level perspectives that can subsequently be decomposed into pieces.

All of the linear algebra or matrix math needed is easily done in Microsoft Excel for small models and there are many resources available on linear algebra as well as Excel matrix arrays and functions should you decide to use them. In the Advanced Material at the end of this chapter (Section 6), we describe how to do these operations in Excel and MATLAB ®,[19] a popular scientific analysis tool used by many academics and researchers that specializes in matrix based computation.

For those of you familiar with the mathematics of infinite series (or matrix math), you will recognize that the series in Equation 8-1 can be replaced by $[\mathbf{I}–\mathbf{A}]^{-1}$, where the $^{-1}$ indicates the multiplicative matrix inverse, following an infinite geometric series approximation. Thus, Equation 8-1 can be simplified to Equation 8-2 (see Advanced Material Section 1 at the end of this chapter for more detail):

$$X = [\mathbf{I} - \mathbf{A}]^{-1} Y \tag{8-2}$$

Two important observations can be made from the use of Equations 8-1 and 8-2 above. First, since summing all direct and indirect purchases results in all of the required production needed, $[\mathbf{I} - \mathbf{A}]^{-1}$ is called the **total requirements table (or matrix)**.

Using the model in Equations 8-1 or 8-2, we can estimate all of the economic outputs required throughout an economy to produce a specified set of products or services. The total of these outputs is often called the **supply chain** for the product or service, where the chain is the sequence of suppliers. For example, an iron ore mine supplies a blast furnace to make steel, a steel mill supplies a fabricator that in turn ships their product to a motor vehicle assembly plant. To make an automobile with its 20,000–30,000 components, numerous chains of this sort are required. An input–output model includes all such chains within the linear model in Equation 8-1.

In the Advanced Material at the end of this chapter (Section 2), we further describe the underlying data sources for the production recipes and transactions tables discussed here. With their whole economy or whole supply chain type, IO models provide very large scopes that can overcome some of the limitations observed with process flow diagram-based models in Chapter 5. In the next chapter, we will see how to apply similar top-down matrix methods to process data.

[19] MATLAB is a registered trademark of The MathWorks, Inc., and will be referred to as MATLAB in the remainder of the book.

Example 8-2: Building on Example 8-1 we calculate direct and total requirement vectors.

Question: Find the direct and total requirements of a $100 billion final demand into each of the two sectors separately.

Answer: Using the direct requirements matrix (**A**) found above, we know that:

$$\mathbf{A} = \begin{bmatrix} .15 & .25 \\ .2 & .05 \end{bmatrix}, \mathbf{I} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, Y_1 = \begin{bmatrix} 100 \\ 0 \end{bmatrix}, Y_2 = \begin{bmatrix} 0 \\ 100 \end{bmatrix}$$

By definition, the direct requirements are [**I+A**]Y, and the total requirements are [**I-A**]^{-1}Y. Using Excel or a similar tool, we can find the (rounded off) inverse matrix, which is:

$$[\mathbf{I} - \mathbf{A}]^{-1} = \begin{bmatrix} 1.254 & .33 \\ .264 & 1.12 \end{bmatrix}$$

Thus, the direct requirements for Y_1 and Y_2 are:

$$[\mathbf{I+A}]Y_1 = \begin{bmatrix} 115 \\ 20 \end{bmatrix} \text{ and } [\mathbf{I+A}]Y_2 = \begin{bmatrix} 25 \\ 105 \end{bmatrix},$$

meaning that a $100 billion demand from sector 1 requires $115 billion in purchases from sector 1 and $20 billion in purchases from sector 2. Similarly, the total requirements are:

$$[\mathbf{I\text{-}A}]^{-1} Y_1 = \begin{bmatrix} 125.4 \\ 26.4 \end{bmatrix} \text{ and } [\mathbf{I\text{-}A}]^{-1} Y_2 = \begin{bmatrix} 33.0 \\ 112.2 \end{bmatrix}, \text{ considering significant figures.}$$

The supply chain perspective gives a basis for considering effects that happen before or after a product is manufactured. We typically refer to points and decisions made before a product is manufactured as **upstream** and those made after a product is manufactured as **downstream**. Building on the previous example, from the perspective of the blast furnace, the iron ore mine is upstream and the vehicle assembly plant is downstream. The upstream and downstream terminology also can apply to process-based models.

Compared to process-based methods, IO methods take a more aggregate view of the sectors producing the goods and services in the U.S. economy. IO models are quick and efficient but are not perfect. Some of their key assumptions and limitations are listed below, and lead to various uncertainties, which are discussed below and in the Advanced Material.

Sectors represent average production. All production facilities in the country that make products and provide services are aggregated into a fixed number of sectors (in the US economy models discussed in this book, approximately 400 sectors). Similar production facilities are all assigned by definition into the same sector, and the model assumes identical production in all facilities of these sectors. In short, no facility in a sector produces any differently than any other in the model (even if in fact this is not true). This is the so-called 'average production' assumption.

You can get a sense of why this is true by referring back to Figure 8-1, which simply aggregates all transactions from all of the facilities into the various Z values, then normalizes by the total output of the entire sector, creating average **A** values.

Input-output models are completely linear. That is, if 10% more output from a particular factory is needed, each of the inputs will have to increase exactly 10%. This of course is not generally true, as there could be economies of scale that allow use of inputs to increase less than 10%. However, this assumption is also common in process-based models.

Manufacturing impacts only. Given the data sources available and used, IO models generally estimate total expenditures only up to the point of manufacture; that is, they do not estimate downstream effects from product use (e.g., consideration of the gasoline needed to run the consumer's car) or end-of-life (e.g., disposal costs and impacts). We will describe below ways to use IO models to estimate impacts beyond the manufacturing phase.

Capital investments excluded. Capital inputs to manufacturing are not included in most IO tables. In the US, such transactions are available in a supplemental transactions table and could be added. Exclusion of capital investments is also a typical assumption for process-based models.

Domestic production. An IO model for a single economy is limited to estimating effects within that country. Despite the fact that many inputs are likely sourced (imported) from other countries in today's global economy, imported inputs are assumed to be produced in the same way as in the model's home country. For certain sectors, this may present a problem because there is so little production done within the home country that the data and/or environmental flows represented are not robust, but the model will still treat that production as if done wholly within the home country and with the associated domestic impacts. Models that move beyond this assumption are possible but beyond the scope of this chapter.

Circularity is inherent and incorporated into the model. In the previous chapters, we noted that all interdependent systems have circularities such as steel needed to make steel, etc., and that this complicated the ability to build process models. IO models embrace the existence of circularity, and the effects are included within the basic model and matrix inversion.

Input–Output Models Applied to Life Cycle Assessment

Now that the underlying economic input-output models have been introduced, we can discuss how they are applied to support decisions about LCA. By appending data on energy, environmental, and other flows to the input-output table, non-economic impacts can be predicted. The resultant models are referred to as **environmentally extended input-output models (EEIO).**

Here we differentiate between using economic input-output methods generally to support LCA as **IO-LCA** and the specific method as implemented with our colleagues in the Green Design Institute of Carnegie Mellon University as **EIO-LCA** (see below). This differentiation seeks to emphasize general IO-LCA methods and practices, as well as specific data sources and assumptions for the EIO-LCA model. Within the US, there are various tables and resources related to IO-LCA. Aside from EIO-LCA, there are the CEDA database and OpenIO. There are other similar IO-LCA models outside the US.

The advantage of the process-based approach as described in Chapters 4 and 5 is that it can answer as detailed a question as desired concerning the materials and energy balances of each facility studied, assuming that adequate resources exist to collect and analyze data on the relevant flows. In practice, getting these balances is sufficiently difficult that they rarely are able to answer detailed questions. The disadvantage of the process-based approach is that the required expense and time means that generally a tight boundary must be drawn that excludes many of the facilities relevant to activities within the overall system.

The advantage of the IO-LCA approach is that a specific boundary decision is not required, because by default the boundary is the entire economy of production-related effects, including all the material and energy inputs. Another major advantage is that it is quick and inexpensive. Results can be generated in seconds at no cost other than the time involved. Note though that with respect to modeling in support of an ISO compliant LCA, IO-LCA methods are generally most useful as a **screening tool** rather than as the core model needed to answer the necessary goals of the LCA task. We introduce IO-LCA methods so that the overall LCA task can be improved by gaining an appreciation for where the greatest systemwide impacts occur. Such knowledge can inform choices of scope, boundaries, and data sources for process based models. They can also be used to help validate results from process-based methods, as the more comprehensive IO-LCA boundaries will generally lead to higher estimates of impacts (upper bounds), which can be used to assess whether the process-based results seem reasonable.

The IO-LCA approach has a major disadvantage: it uses aggregate and average data for a sector rather than detailed data for a specific process. For example, an IO-LCA model will yield results for the average production from, say, the sector *iron and steel mills*, rather than from producing particular steel alloys required for an automobile (which one could find in process data). As another example, the US IO table does not distinguish between generating electricity using a 50-year-old coal plant and using a new combined-cycle gas turbine. The former emits much more pollution per kWh than the latter. A process model could compare the different processes to the degree desired. The process models can be specific to particular materials or processes, rather than the output from a sector of the economy. Even with more than 400 sectors available, analysts would often like to disaggregate IO-LCA models, such as dividing the plastics sector into production of different types of plastic. Process models can also handle nonlinear effects.

While we focus on the use of IO-LCA in this chapter, we also describe 'hybrid' models in which IO-LCA and process models are combined to exploit the advantages of both (see Chapter 9). In hybrid models, the results from production of a chemical might be from a process model, while the effects of the inputs to the process might be assessed with IO-LCA. With a hybrid model, the reliance on process models can vary from slight (such as one model) to very extensive (IO-LCA might be used only for a single input such as electricity).

IO-LCA (or EEIO) models work by following the flow chart shown in Figure 8-2. Economic activity generates environmental impacts. Production of steel generates solid waste (slags, air pollution, wastewater) and consumes energy that results in greenhouse gas emissions. These environmental impacts can be assumed to be linear in their magnitude and can also be described as vectors.

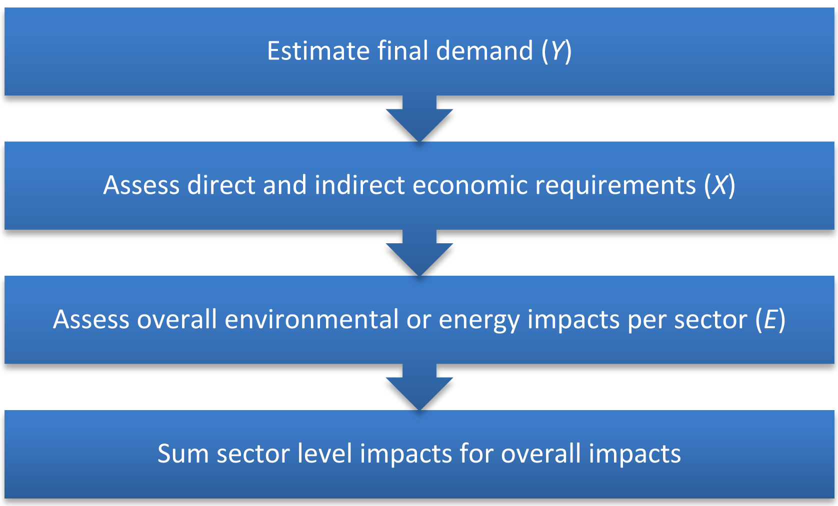


Figure 8-2: Flow chart for IO-LCA models

We have already described above the process needed to complete the first two steps. In this section, we describe the last two steps that allow use of the models for LCA screening purposes. In our discussion, we use 'dollars' as a currency given our own biases, but IO-LCA models can and have been derived around the world in many currencies. Our use of dollars is meant to merely provide a consistent terminology for expression of economic values.

Once economic output for each sector (X) is known, a vector of **total environmental effects** (i.e., the sum of direct and indirect environmental effects) for each sector can be obtained by multiplying the output by the environmental impact per dollar of output:

$$E = \mathbf{R}X = \mathbf{R}\,[\mathbf{I} - \mathbf{A}]^{-1}\,Y \tag{8-3}$$

where E is the vector of environmental burdens (such as toxic emissions or electricity use for each production sector), and $\mathbf{R}$ is a matrix with diagonal elements representing the impact per dollar of output for each sector for a particular energy or environmental burden (Lave 1995, Hendrickson 1998, Leontief 1970). A variety of environmental burdens may be included in this calculation. For example, from estimates of resource inputs (electricity, fuels, ores, and fertilizers), we can further estimate multiple environmental outputs (toxic emissions by media, hazardous waste generation and management, conventional air pollutant emissions, global warming potential, and ozone depleting substances). We find direct and indirect environmental burdens by multiplying the **R** matrix by the direct and indirect purchases. As before, the total environmental burdens are the sum of direct and indirect.

While the matrix math is trivial, the data needs are worth discussing. The **R** matrix has units of burdens per dollar of output (e.g., kg CO_2/$). Multiplying **R** by a vector (X) with unit of dollars for each sector will yield E, with units of burdens by sector (e.g., kg CO_2), removing cost dependence from the final impact tally. Deriving **R** (in units of burdens per dollar) merits additional discussion; understanding the process and its limitations is critical to understanding how and why IO-LCA models can be used to support LCA tasks.

IO-LCA model results are better understood by example. Like the purely economic IO models on which they are built, IO-LCA models can estimate environmental burdens across the supply chain. Revisiting our example of a final demand of $20,000 of automobiles, if our **R** matrix was for emissions, then the $E = \mathbf{R}X$ vector would represent total emissions. Included in these estimated emissions would be not only the emissions from the automobile factory, but also the emissions from the tire factory, the rubber factory, and all other upstream processes (including transportation) that supported production of that $20,000 car. All of this is possible since the simple assumption of the **R** matrix is that it contains emissions per dollar for each sector, and the X vector has already estimated the necessary economic outputs for all of the sectors. As mentioned in the beginning of this chapter, in addition to the inputs from each sector, the output, X, includes value added such as labor or profits. As a result, emissions will be indirectly associated with these activities as well.

Example 8-3: Build on methods above to estimate direct and total environmental burdens by sector.

Question: What are the direct and total emissions of waste for the inputs specified in Example 8-2? Assume emissions of waste per billion dollars of output of sector 1 are 50 g and sector 2 are 5 g.

Answer: The direct emissions are $E = \mathbf{R}\,[\mathbf{I}+\mathbf{A}]\,Y$ and total emissions are $E = \mathbf{R}X = \mathbf{R}[\mathbf{I}-\mathbf{A}]^{-1}\,Y$. Example 8-2 derived $[\mathbf{I}+\mathbf{A}]\,Y$ and $[\mathbf{I}-\mathbf{A}]^{-1}\,Y$ for each of the two sectors. As given above, $\mathbf{R} = \begin{bmatrix} 50 & 0 \\ 0 & 5 \end{bmatrix}$. For Y_1 and Y_2 the direct emissions are:

$$\begin{bmatrix} (50 * 115) + (0 * 20) = 5750 \\ (0 * 115) + (5 * 20) = 100 \end{bmatrix} \text{ and } \begin{bmatrix} 1250 \\ 525 \end{bmatrix}.$$

Thus, the sum of direct emissions for Y_1 are 5,850 g (5.9 kg) and for Y_2 are 1,775 g (1.8 kg). Similarly the total emissions are 6403 g (6.4 kg) and 2211 g (2.2 kg), respectively, with consideration for significant figures. The direct emissions, in general, are a fairly large share of the total emissions in both cases.

Producing an **R** matrix for a burden, e.g., sulfur dioxide emissions, requires a comprehensive data source of such emissions on a sector level. The sector classification of relevance is that of the associated IO model. As mentioned above, IO models typically follow existing classification schemes (e.g., the US 2002 benchmark model generally follows the NAICS classification system used throughout North America). Thus, a data source of total sulfur dioxide emissions broken up by NAICS sector is required. Such a data source is ideally already available and in the US, can be obtained from the Environmental Protection Agency. However, some work may need to be done to translate, convert, or re-classify existing data into a NAICS or IO sector basis to use in an IO-LCA model (see the Advanced Material for this Chapter, Section 4). Once total sulfur dioxide emissions in an economy in a given year are found (ideally the data source would provide emissions in the same year as the IO table used) and allocated to the various IO sectors, each of the sector emissions values is divided by the total output (X_i) of the sector. The result is the **R** matrix of burdens per dollar of output. Typically, IO models are used at the scale of "millions of dollars" (as opposed to dollars or thousands of dollars) so the normalization factor X_i would be scaled to millions (i.e., instead of dividing by \$120 million would be divided by 120). This same process would be repeated for all burdens of interest in the IO-LCA model building process. Example 8-4 provides insight into generating **R** matrix values for the electricity sector.

Example 8-4: In this example, we show how to derive an **R** matrix value for a particular sector.

Question: What is the **R** matrix value for the electric power sector for sulfur dioxide emissions in 2002 (in short tons per million dollars)?

Answer: EPA data for 2002 suggests that the total sulfur dioxide emissions from power generation was 10.7 million short tons. In 2002, the sector output of the power generation and supply sector was \$250 billion. Thus the **R** matrix value for the power generation sector would be (10.7 million short tons) / (\$250 billion) = 42.8 short tons per million dollars.

Uncertainties in IO-LCA Models

Similar to Chapter 7, we revisit several categories of uncertainty for IO-LCA models.

Temporal uncertainty: IO models generally lag a bit more than process based data sources, as it takes considerable time to assemble and report the data required for national input–output tables, and then further to assemble data for the various environmental impacts (i.e., **R** matrices). That said, production processes and IO tables have considerable consistency over time (see Advanced Material for this Chapter - Section 8 for more details).

Linear production: the environmental impact vectors generally use average impacts per dollar of output across the entire sector, even though the effects of a production change might not be incremental or marginal. For example, increasing demand for a new product might be produced in a plant with advanced energy efficiency or pollution control equipment or with brand new technology. If production functions are changing rapidly, are discontinuous, or are not marginal, then the linear approximations will be relatively poor. Reducing errors of this kind requires more effort on the part of the LCA practitioner. A simple approach would be to alter the parameters of the IO-LCA model to reflect the user's beliefs about their actual values. Thus, estimates of marginal changes in **R** vectors may be substituted for the average values provided in the standard model.

Foreign production: Process or IO-based models may assume that manufacture of inputs occurs in the same geographic region as the manufacture of the product. For example, IO-LCA models of a country represent typical production across the supply chain within each sector of the home country, even though some products and services might be produced outside the country and imported. The impacts of production in other regions may be the same, or substantially different. This variation could be the result of different levels of environmental regulation or protection. For example, leather shoes imported from China were probably produced with different processes, chemicals, and environmental discharges than leather shoes produced in the US.

Aggregation uncertainty: Aggregation issues arise primarily in IO-based models and occur because of heterogeneous producers and products within any one input–output sector. However, similar issues may occur in process-based models, such as those that aggregate various production processes into a single average method, e.g., electricity 'grid' processes representing a weighted average of underlying processes.

It is normal to want more detail than is available from a particular model's sectors or processes, but such desires need to be balanced against the level of data available. Even the hundreds of sectors of large IO models do not give sufficiently detailed information on many particular products. For example, we might seek data on lead-acid or nickel-metal hydride batteries, but the model may only have two (primary and secondary) battery sectors. Likewise, a single battery production sector may contain tiny hearing aid batteries as well as massive cells used to protect against electricity blackouts in buildings. $1 million spent on hearing aid batteries will use quite different materials and create different environmental discharges than $1 million spent on huge batteries. But all production within an IO sector is assumed to be uniform (identical), and the economic and environmental effect values in the model are effectively averaged across all products within it.

As one indication of the degree of uncertainty, Lenzen (2000) estimates that the average total relative standard error of input–output coefficients is about 85%. However, because numerous individual errors in input–output coefficients cancel out in the process of calculating economic requirements and environmental impacts, the overall relative standard errors of economic requirements are only about 10–20%.

Introduction to the EIO-LCA Input-Output LCA Model

In this section, we provide specific information about the EIO-LCA model and specific illustrations to aid in your understanding of how such models are built. EIO-LCA was developed by researchers at the Green Design Institute at Carnegie Mellon University and is available both as an Internet web tool and as an underlying MATLAB model environment at http://www.eiolca.net/. Online tutorials and screencasts are available on how to use the web model and are not repeated here. For EIO-LCA MATLAB information, see the Advanced Material for this Chapter, Section 6. EIO-LCA is available free for non-commercial use on the Internet (although even corporate users are able to generate results to get a sense of how it works). Process model databases and software sold by consulting companies can be quite expensive, generally ranging in the order of thousands of dollars, as shown in Chapter 5. Other IO-LCA models likely use very similar processes, but if you intend to use them, you should read their documentation to ensure similarity and limitations.

The calculations required for Equations 8-1 through 8-3 are well within the capabilities of personal computers and servers. The result is a quick and inexpensive way to trace out the supply chain impacts of any purchase. The EIO-LCA website is able to assist in generating IO-LCA estimates for various countries and with various levels of detail. In this chapter, we will use the 428-sector 2002 US economy model to explore a variety of design and purchase decisions. The EIO-LCA model traces the various economic transactions, resource requirements, and environmental emissions required to provide a particular product or service. The model captures all the various manufacturing, transportation, mining, and related requirements to produce it. For example, it is possible to trace out the upstream implications of purchasing $50,000 of reinforcing steel and $100,000 of concrete for a kilometer of roadway pavement. Environmental impacts of these purchases can be estimated using EIO-LCA. Converting such values into the relevant benchmark model year dollar values is discussed in Section 3 of the Advanced Material.

We discuss the various data sources for the model, give an example application of the software, provide a numerical example of the input–output calculations, and provide some sample problems.

The data in the EIO-LCA software is derived from a variety of public datasets and assembled for the various commodity sectors. For the most part, the data is self-reported and is subject to measurement error and reporting requirement gaps. For example, automotive repair shops do not have to report to the Toxics Release Inventory. The level of quality assurance of the public data used varies. The major datasets include:

- *Direct and Total Input–Output Tables*: The EIO-LCA website provides models for the US, Germany, Spain, Canada, and China. US models available include those for the benchmark years 1992, 1997, and 2002. Several of those years have multiple levels of sector detail available. The 428-sector 2002 industry by commodity input–output (IO) matrix of the US economy as developed by the U.S. Department of Commerce Bureau of Economic Analysis is the default model. Economic Impacts are computed from the IO matrix and the user-input change in final demand. While the remaining data sets below are generally available for multiple country-year models, the specific details provided are for the 428-sector 2002 model.

R matrices:

- *Energy use* for the 428 sectors come from a number of sources. Energy use of manufacturing sectors (roughly 270 of 428) is developed from the Manufacturing Energy Consumption Survey (MECS), for mining sectors is calculated from the 2002 Economic Census (USCB 1997). Service sector electricity use is estimated using the IO table purchases and average electricity prices for these sectors.

- *Conventional pollutant emissions* are from the US Environmental Protection Agency, primarily the National Emissions Inventory (NEI) and onroad /nonroad data sources.

- *Greenhouse gas emissions* are calculated by applying emissions factors to fuel use for fossil-based emissions and allocating top-down estimates of agricultural, chemical process, waste management, and other practices that generate non-fossil carbon emissions to economic sectors.

- *Toxic releases and emissions* are derived from EPA's 2002 Toxics Release Inventory (TRI).

- *Hazardous waste*, specifically Resource Conservation and Recovery Act (RCRA) Subtitle C hazardous waste generation, management, and shipment was derived from EPA's National Biannual RCRA Hazardous Waste Report.

- *Water withdrawals* come from various sources, as published in Blackhurst (2010).

- *Transportation use* tracking flow of products through different freight modes comes from multiple data sources, as published in Nealer (201x).

- *Land use* estimates come from various sources, as published in Costello (2011).

Detailed information on how the underlying data sources are used to generate the R matrices in EIO-LCA are available on the EIO-LCA website (at http://www.eiolca.net/docs/).

The EIO-LCA website follows the same workflow as the generic IO model shown in Figure 8-1. As a user, all you need to do is to enter a single value of final demand, select a sector that must produce that final demand, and choose whether you want to see economic (X) or energy-environmental results (**R**). All of the matrix math, data management, etc., is done by a web server and results are shown in tabular form within seconds. In the basic EIO-LCA web model, you can only enter a final demand for a single sector (i.e., you can only enter a value for a single Y_i and all other elements of Y are assumed to be zero). However, in general, IO models can be run with multiple (even many) Y_i final demand entries. It is possible to build a custom model in the EIO-LCA model (see Advanced Material Section 5) where simultaneous purchases can be made from multiple sectors; however, such a model also has limitations on its meaningfulness. If you have not used EIO-LCA before, there are various tutorials, screencasts, and other resources available on the website.

EIO-LCA Example: Automobile Manufacturing

As a specific demonstration of the utility of IO-LCA models (specifically EIO-LCA), this section examines the manufacture of automobiles. As defined by the US Department of Commerce, the *automobile manufacturing* sector for the US 2002 benchmark EIO model is composed of the following NAICS sector:

> **336111 Automobile Manufacturing**
>
> This U.S. industry comprises establishments primarily engaged in one or more of the following manufacturing activities:
>
> * complete automobiles (i.e., body and chassis or unibody) or
>
> * automobile chassis only.

Note that the EIO-LCA server shows the information above when browsing or searching for sectors to make final demand. For example, choosing the related *light truck manufacturing* sector would provide similar but different results below.

We can trace the supply chain for the production of $1 million of automobiles in 2002 using EIO-LCA. This production of $1 million would represent the effects of making roughly 40 automobiles (given an approximate average price of $25,000 each in the year 2002). Figure 8-3 shows the total and direct (including percentage direct) economic contributions of the largest 20 supply sectors within the supply chain for making automobiles in the US.

First, the economic results are considered. From Figure 8-3, a $1 million final demand of automobiles requires total economic activity in the supply chain of $2.71 million. In the total economic output column are the elements of X. EIO-LCA also sums across all of X to present the total. Results for the other 403 sectors are available on the website but are not shown here.

	Total Economic $mill	Direct Economic $mill	Direct Economic %
Total for all sectors	**2.71**	**1.74**	**64.2**
Automobile manufacturing	0.849	0.849	100.0
Motor vehicle parts manufacturing	0.506	0.446	88.1
Light truck and utility vehicle manufacturing	0.150	0.150	99.9
Wholesale trade	0.124	0.057	46.1
Management of companies and enterprises	0.108	0.033	30.9
Iron and steel mills	0.038	0.000	1.60
Semiconductor and related device manufacturing	0.026	0.014	54.7
Truck transportation	0.025	0.009	34.2
Other plastics product manufacturing	0.021	0.010	48.5
Power generation and supply	0.020	0.002	10.7
Real estate	0.020	0.001	5.97
Turned product and screw, nut, and bolt manufacturing	0.017	0.005	30.1
Ferrous metal foundaries	0.015	0.000	1.68
Nonferrous foundries	0.015	0.000	2.14
Glass product manufacturing made of purchased glass	0.015	0.012	84.4
Other engine equipment manufacturing	0.015	0.012	79.8
Machine shops	0.014	0.002	15.9
Oil and gas extraction	0.013	0.000	0.085
Monetary authorities and depository credit intermediation	0.013	0.000	2.18
Lessors of nonfinancial intangible assets	0.013	0.000	3.65

Figure 8-3: Supply chain economic transactions for production of $1 million of automobiles in US, $2002. Top 20 sectors. Results sorted by total economic output.

As discussed above, the change in GDP as a result of this economic activity would be only $1 million, since GDP measures only changes in final output, not of all purchases of intermediate goods (i.e., not $2,710,000). The largest activity is in the *automobile manufacturing* sector itself: $849,000. This includes purchases by the company that assembles vehicles from other companies within the automobile manufacturing industry, like those that make steering wheels, interior lighting systems, and seats.

The economic value of the supply chain is also shown in Figure 8-3. Direct purchases (from the IO perspective, i.e., **I+A**) are $1.74 million, including the $1 million of final demand. Not surprisingly, direct purchases are dominated by vehicle and parts manufacturing sectors. The direct percentage compares the direct purchases for each sector to the total purchases across the supply chain. Sectors with small direct purchase percentages (or, alternatively, large indirect percentages) have most of their production feeding the indirect supply chain of automobiles rather than the automobile assembly factories directly. Many of the top 25 sectors have more than 50% of their total output as direct inputs into making automobiles (e.g., *semiconductor manufacturing* and *glass manufacturing* sectors). Others primarily supply the other suppliers (e.g., *iron and steel mills, power generation and supply*).

EIO-LCA also allows you to generate estimates of energy and environmental effects, using the data sources identified above. Using the same final demand of $1 million, Figure 8-4 shows the energy use across the supply chain for producing automobiles for the top 10 energy-consuming sectors (results available but not shown for the other 418 sectors). We remind you that IO models are linear. Any analysis you do for $1 million of automobiles can be linearly scaled down per vehicle.

EIO-LCA estimates total supply chain energy use of 8.33 TJ per $1 million of automobiles manufactured (or 167 GJ per vehicle assuming 50 vehicles produced at an average cost of $20,000). About 25% of that energy use (2.19 TJ) comes from energy needed in the electricity (*power generation and supply*) sector, and about 15% from *iron and steel mills*. Most of the coal used in the supply chain is for generating power. Natural gas use is fairly evenly split among the top sectors. Similarly, most of the petroleum used is in the various transportation sectors, not all of which are shown in the top 10 list of Figure 8-4. Notice that the top sectors in terms of economic output are not closely associated with the top energy-consuming sectors! IO models will show that generally energy-intensive sectors are an important part of the energy supply chain for every sector but are not always those that have the largest economic input.

Note that specific fuels are not shown in Figure 8-4. Underlying data sources provide information on consumption of diesel, gasoline, and other fuels that are aggregated into a single estimate of "Petroleum" use.

Sector	Total Energy TJ	Coal TJ	NatGas TJ	Petrol TJ	Bio/Waste TJ	Non-Foss Elec TJ
Total for all sectors	8.33	2.56	2.63	1.29	0.435	1.41
Power generation and supply	2.19	1.60	0.467	0.078	0	0.051
Iron and steel mills	1.25	0.743	0.341	0.012	0.005	0.151
Motor vehicle parts manufacturing	0.460	0.005	0.190	0.014	0.024	0.228
Automobile Manufacturing	0.381	0.004	0.190	0.013	0.040	0.133
Truck transportation	0.327	0	0	0.324	0	0.003
Other basic organic chemical manufacturing	0.259	0.032	0.099	0.036	0.078	0.014
Petroleum refineries	0.187	0.000	0.050	0.121	0.009	0.007
Alumina refining and primary aluminum production	0.172	0	0.046	0.001	0.004	0.120
Plastics material and resin manufacturing	0.169	0.007	0.088	0.037	0.018	0.019
Paperboard Mills	0.161	0.015	0.033	0.007	0.095	0.011

Figure 8-4: Supply chain energy requirements for production of $1 million of automobiles in 2002, results for top 10 energy consuming sectors, sorted by total energy.

IO-LCA models must carefully manage fuel and energy data. Fuel use is tracked only in the sector that directly uses it. Many sectors consume electricity, but only the power generation sector consumes the coal, natural gas, petroleum, and biomaterial needed for generation. Also, note that "non-fossil" electricity consumption is estimated in Figure 8-4. While facilities within sectors are assumed to consume average electricity (generated from a mix of fossil and non-fossil sources), if the model tracked total energy use of fossil and non-fossil sourced electricity in TJ, and also tracked the coal and/or natural gas used to generate it, the model would "double count" the energy in the fossil fuel and the electricity. Thus, we only track an average amount of non-fossil electricity (which does not depend on fossil fuels to generate it), avoiding the double counting of energy. To derive the non-fossil share, Department of Energy data on percent non-fossil electricity generation in 2002 (31%) is multiplied by the amount of electricity consumed by each sector.

The case study below helps to describe how an IO-LCA based screening assessment can be used in a corporate setting.

Case Study: Bio-based Feedstocks in the Paint and Coatings Sector

A US company was considering acting on customer requests to provide an alternative product comprised of bio-based feedstocks. These customers were looking to reduce the "carbon footprint" of their products, and to reduce fossil fuel consumption, and had read studies on the net carbon efficiency of bio-based as opposed to petrochemical-based feedstocks. The question is whether such a conversion might be a beneficial substitution for the producer and its customers.

An excerpted screening analysis using EIO-LCA for $100,000 of final demand in the *Paint and coatings* sector demonstrated that the current mix of petroleum based feedstocks – across the entire production chain of making paints and coatings – is a fairly small part of the total purchases as well as only about 5% (6 / 107 tons) of the carbon emissions. On the other hand, supply chain-wide purchases of electricity are comparable in economic value ($2k and $5k) but constitute 25% (25 / 107 tons) of the total carbon equivalent emissions. This screening analysis suggests that the switch to bio-based feedstocks would likely have a modest effect on the burden of the product. In addition, it suggests that a corporate push for more renewable electricity in its supply chain could have substantial benefits. It is this latter strategy that we recommended to our corporate partner.

Figure 8-5: EIO-LCA economic and CO2 emissions results for the 'Paint and Coatings' Sector

Sector	Total ($ thousand)	CO_2 equivalents (tons)
Total across all 428 sectors	266	107
Paint and coatings	100	3
Materials and resins	13	5
Organic chemicals	12	5
Wholesale Trade	10	1
Management of companies	10	< 1
Dyes and pigments	8	17
Petroleum refineries	5	6
Truck transportation	5	8
Electricity	2	25

Beyond Cradle to Gate Analyses with IO-LCA Models

So far, IO tables and models have been represented as capturing the entire upstream supply chain up to the point of manufacture, economically referred to as a **producer price basis.** This means the context of final demand as well as the structure of the production recipes was from the perspective of the producer at their place of business. In other words, the relevant input into the model would be as if the producer was merely trying to recover the costs they incurred. Thus, the perspective (or boundary) of producer-basis models is 'cradle to gate' – the effects estimated by the model end at the point of production. The appropriate final demand input is as measured or 'seen' by the producer.

IO models can also be created on a **purchaser price basis**. In such models, additional stages beyond production are internalized into the production recipe for each sector. For physical goods, typical activities internalized include transportation of product from the producer as well as wholesale and retail operations (so-called **margin activities**). These models have a 'cradle to consumer' boundary. The relevant dollar input into a purchaser price model is the price that a consumer (buyer) would expect to pay, which is generally easier to determine or derive than a producer price. As a simple example, an automobile may have a $20,000 producer price basis, but after transportation and dealer overhead are included may have a purchaser price of $25,000. In such a case, the 'production recipe' of the $25,000 car in a purchaser price model might have $22,000 of the recipe be associated with automobile manufacturing, $1,000 for transportation (e.g., by truck and rail) and $2,000 for retail overhead. If we were able to perfectly separate the pieces, the purchaser price model would have all of the effects as estimated as if a $22,000 input into a producer price model had been used, as well as the additional effects from the $3,000 of other activities. Additional impacts that might be estimated via the wholesale and retail margins are electricity use by computers at the store or emissions from climate-controlled warehouses. On the other hand, if we entered the same value ($22,000 or $25,000) into both models, this correspondence would not occur since the models are linear and the recipes would not be fully inclusive. Additional detail about price systems in producer and purchaser models is available in the Advanced Material (Section 3) at the end of this chapter.

Beyond producer and purchaser basis models, IO-LCA models can be used to estimate effects of even broader life cycles. In the example above, we consider what an EIO model would estimate in terms of effects for a cradle to gate and cradle to consumer boundary for an automobile. With those boundaries, the various requirements of using the vehicle (e.g., purchasing gasoline, insurance, maintenance, etc.) and managing it at end of life would not be included in the model results.

However, these broader considerations can be approximated with a slightly more complicated IO-LCA model final demand input. Instead of entering only a single element of final demand, multiple Y_i elements can be chosen. Alternatively, the model could be run multiple times and the individual results aggregated to a single final result.

If we build on problem 8 from Chapter 3, which compared the life cycle effects associated with two washing machines, the life cycle effects might include various elements of final demand (entered as a series of elements in Y – see Advanced Material Section 5 - or run consecutively through a model as separate individual elements of final demand):

- An input of final demand into the *Household laundry equipment manufacturing* sector
- A final demand input for a lifetime of water used (at an assumed $/gallon cost) from the *Water, sewage, and other systems* sector
- A final demand input for a lifetime of electricity used (at an assumed $/kWh cost) from the *Power generation and supply* sector

A full analysis would include differences pertaining to end-of-life disposal, although the differences are likely to be relatively small for two washing machines.

What do these EIO-LCA results demonstrate?

Keeping in mind that IO-LCA models are a screening tool, they can help us to make and justify our LCA project design decisions. If we were doing an LCA of the energy use of an automobile, our IO-LCA results suggest that the boundary of the manufacturing processes should include electricity, semiconductors, trade, and chemicals needed. Many of the other processes could be ignored with little impact on the results.

IO model frameworks produce results as shown in Figure 8-3 and Figure 8-4. The effects from all facilities within a sector at many levels of the supply chain are "rolled up" into these single value results. For example, from Figure 8-4, the 1.6 TJ of coal used in the power generation sector comes from many individual power plants, some (about 10%) directly, but mostly indirectly. These rolled up results do not allow us to see energy use at specific tiers of the supply chain, or at particular facilities. From such frameworks, the best analysis possible is a comparison of direct and indirect effects. Advanced methods, such as structural path analysis, (Chapter 12) allow one to drill down into specific layers of the supply chain to find specific pathways of connections between requirements.

As discussed in Chapters 2 and 5, referencing of data sources and models is critical in LCA. If using an IO-LCA model in your study, you should be sure to note:

- the name and location of the model,
- its country of focus,
- EIO table year, and

- whether it is a producer or purchaser basis model.

Somewhere in your study, you must also clearly state the input value of final demand, the name of any sectors chosen for analysis, and **R** matrix datasets chosen. For example, the EIO-LCA model suggests the following citation be used:

> Carnegie Mellon University Green Design Institute. (2013) Economic Input-Output Life Cycle Assessment (EIO-LCA) US 2002 (428 sectors) Producer model [Internet], Available from: http://www.eiolca.net/> [Accessed 25 Aug, 2013].

In addition, you could separately provide in the LCA study document a table of final demands, detailed sectors, and impact categories.

So far we have only motivated economic input-output models. However an IO framework can be applied to any type of unit, for example, physical flows. If desired, one could derive a linear system of equations that instead represented the mass quantities needed across an economy to support production. Such models could also be built with multiple or mixed units. All of the same matrix techniques can be used to estimate direct and total requirements (see Chapter 9).

Overall, tracing the supply chain requirements for production has yielded surprises that have raised questions about sole reliance on process-based LCA. The most important suppliers in one dimension (e.g., economic dependence) often are not the most important in another (e.g., energy use). Figure 8-4 showed that some of the largest energy users in the supply chain do not even appear among the top 25 economic supply sectors. A system-wide view is critical in assessing life cycle effects. That said, IO-LCA models provide quick but coarse and average estimates of LCI results, and cannot substitute for detailed process-based analysis. IO-LCA methods can help you draw boundaries, assess which processes are important, and can help validate process-based results. Given the very short time required to generate results, there is little reason not to consult an IO-LCA model in support of an LCA study when setting the SDPs. Your screening analysis could identify whether the choice of sector was critical or not, and also generalize whether placing various processes within the system boundary is critical or not.

Chapter Summary

We have shown several examples that highlight both the ease and utility, along with the complications, of exploring the entire supply chain via IO and IO-LCA models. Process-based models were shown to specifically estimate detailed mass and/or energy balances for specific activities relevant to the life cycle of a product and to link many of these data sources to yield a 'bottom up' model. The process-based method is also generally expensive and time-consuming. The resource intensity can lead to project design decisions that narrow the boundaries around the problem, causing many supply chain aspects to be ignored. IO-LCA methods, on the other hand, have a top-down system boundary of the entire supply chain up to the point of production by default. The benefits of economy-wide comprehensiveness in IO-LCA models are traded off against the reality that the models are built upon average values for sectors and environmental burdens. As such, the utility of IO-LCA models is primarily as a screening tool rather than as a true alternative to a process-based model. For those wishing to read further into the theory and practice of economic input-output models, we can recommend two sources: Miller (2009) and Hendrickson (2006).

References for this Chapter

Blackhurst, Michael, Hendrickson, Chris, Sels I Vidal, Jordi, "Direct and Indirect Water Withdrawals for U.S. Industrial Sectors", *Environmental Science & Technology*, 2010, Vol. 44, No. 6, pp. 2126–2130, DOI: 10.1021/es903147k

Costello, Christine, Griffin, W. Michael, Matthews, H. Scott, and Weber, Christopher, "Inventory Development and Input-Output Model of U.S. Land Use: Relating Land in Production to Consumption", *Environmental Science & Technology*, 2011, Vol. 45, pp. 4937–4943, **DOI**:10.1021/es104245j

Rachael Nealer, Christopher L. Weber, Chris Hendrickson and H. Scott Matthews, "Modal freight transport required for production of US goods and services", Transportation Research Part E: Logistics and Transportation Review, Volume 47, Issue 4, July 2011, Pages 474-489. DOI: 10.1016/j.tre.2010.11.015

Hawkins, Troy R., and Matthews, Deanna H., 2009. A Classroom Simulation to Teach Economic Input-Output Life Cycle Assessment. *Journal Of Industrial Ecology* 13, no. 4 (August): 622-637. doi:10.1111/j.1530-9290.2009.00148.x

Hendrickson, Chris., Arpad Horvath, Satish Joshi, and Lester Lave. Introduction to the Use of Economic Input–Output Models for Environmental Life Cycle Assessment. *Environmental Science and Technology,* 32(7): 184A–191A, 1998.

Hendrickson, Chris T., Lave, Lester B., and Matthews, H. Scott, "Environmental Life Cycle Assessment of Goods and Services: An Input-Output Approach", RFF Press, April 2006.

Lave, L., E. Cobas-Flores, C. Hendrickson, and F. McMichael, Generalizing Life Cycle Analysis: Using Input–Output Analysis to Estimate Economy-Wide Discharges. *Environmental Science & Technology,* 29(9): 420A–426A, 1995.

Lenzen, M. Errors in Conventional and Input Output-based Life Cycle Inventories, *Journal of Industrial Ecology*, 4(4):127 – 148, 2000. DOI: 10.1162/10881980052541981

Leontief, W., Environmental Repercussions and the Economic Structure: An Input–Output Approach. *Review of Economics and Statistics,* 1970.

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

NAICS 2013 United States Census Bureau. 2013. 'North American Industrial Classification System, http://www.census.gov/eos/www/naics/ (accessed July 10, 2013).

End of Chapter Questions

Objective 1. Describe how economic sector data are organized into input-output tables

Objective 2. Compute direct, indirect, and total effects using an input-output model

1. Use this transactions table (in millions of currency units) to answer the following questions.

	1	2	Y	X
1	**450**	**200**	350	**1000**
2	100	600	**1500**	2200
V	450	1400		
X	1000	**2200**		

   a. Describe in words what the highlighted values in the table represent.

   b. Generate the direct requirements matrix

   c. Generate the total requirements matrix

   d. For a final demand of $50 million in sector 1, find the direct and total requirements.

2. Reconsider the washing machine homework question from Chapter 3 using the 2002 EIO-LCA purchaser price model. Assume a 10-year lifetime of each machine without discounting. Ignore potential impacts from a disposal phase.

   a. Use the $500 and $1,000 purchaser prices as inputs into the EIO-LCA *Household laundry equipment manufacturing* sector to estimate the total energy consumption and fossil CO_2 emissions to manufacture the two machines. Compare direct and indirect effects. What do the results using these inputs suggest?

   b. Use the assumptions about water use to estimate the use-phase energy and fossil CO_2 emissions via input to the *Water, sewage, and other systems* sector.

   c. Use the assumptions about electricity use to estimate the use-phase energy and fossil CO_2 emissions via input to the *Power generation and supply* sector.

   d. Create a table summarizing the results above and find total energy and fossil CO_2 emissions for the two machines. Determine the percent of energy and fossil CO2 emissions associated with manufacturing and use phases. What are some caveats you might want to note if presenting these results given your use of an IO-LCA model?

Objective 3. Assess how an input-output model might be used as a screening tool.

3. You want to do a screening assessment of the energy needed to manufacture two different types of plain white cotton t-shirts, one from a discount store costing $5 and another from a specialty clothing store costing $15. What would the results of using an IO-LCA model's apparel sector suggest about the differences in their energy use for manufacture? What other differences are likely in their manufacturing energy use?

4. Consider the case of a university looking to better manage its greenhouse gas emissions.

a. What are the fossil CO_2 emissions associated with $1 million of *College, university, and junior colleges* in 2002 using the EIO-LCA 2002 benchmark producer price model? What percent of the emissions from each of the top 10 sectors are direct? What are the overall direct emissions? What do these results tell you about the kinds of decisions or activities that might be best managed by an average university in hopes of reducing fossil CO_2 emissions?

b. How could you adjust the average *College, university, and junior colleges* sector fossil CO_2 results from (a) to represent a university with a $200 million annual budget and that is considering the purchase of 10% of its electricity from wind power? As a simplification, assume that no fossil CO_2 emissions are associated with wind power generation and the amount of wind power used in the estimate of part (a) is zero.

c. How could you adjust the average *College, university, and junior colleges* sector fossil CO_2 results from (a) to represent a university with a $200 million annual budget and that is considering the purchase of new equipment to lead to an overall reduction of 20% in fossil CO_2 emissions on site compared to the average university?

d. Comment on the overall feasibility and effectiveness of the university's choice to pursue green power or equipment replacement to achieve GHG emissions targets.

Advanced Material for Chapter 8 - Overview

As with the Advanced Material elsewhere in this book, these sections contain additional detail about the methods and principles discussed in the chapter. They have been moved to the back of the chapter because knowing about them is not vital to understanding the chapter content, but may be necessary if you intend to more substantively use those methods. It is generally expected that an undergraduate course (or casual learner of LCA) would generally focus on the main chapters, and a graduate course (or advanced practitioner) would incorporate elements from the advanced material.

In the advanced material sections of this chapter, you will find more in depth discussion of the theoretical framework of economic input-output models, how price systems change, how the vectors and matrices of IO-LCA models (with specific examples from EIO-LCA) have been constructed, advanced features of the EIO-LCA website, and using software tools to develop IO-LCA models.

Section 1 - Linear Algebra Derivation of Leontief (Input-Output) Model Equations

In the chapter, we showed the format of the transactions table and a general derivation of the round by round purchases and how they become the Leontief inverse equation. In this section, more detail is provided about the system of linear equations that drive IO models. If you will be doing matrix computations in your work using IO-LCA models, it is important to understand the equations in this section. We repeat Figure 8-1 here.

	Input to sectors				Intermediate output O	Final demand Y	Total output X
Output from sectors	1	2	3	n			
1	Z_{11}	Z_{12}	Z_{13}	Z_{1n}	O_1	Y_1	X_1
2	Z_{21}	Z_{22}	Z_{23}	Z_{2n}	O_2	Y_2	X_2
3	Z_{31}	Z_{32}	Z_{33}	Z_{3n}	O_3	Y_3	X_3
n	Z_{n1}	Z_{n2}	Z_{n3}	Z_{nn}	O_n	Y_n	X_n
Intermediate input I	I_1	I_2	I_3	I_n			
Value added V	V_1	V_2	V_3	V_n		GDP	
Total output X	X_1	X_2	X_3	X_n			

Figure 8-1 (repeated). Example Structure of an Economic Input–Output Transactions Table

Notes: Matrix entries Z_{ij} are the input to economic sector j from sector i. Total (row) output for each sector i, X_i, is the sum of intermediate outputs used by other sectors, O_i, and final demand by consumers.

In a typical IO model, output across the rows of the transactions table (Figure 8-1), typically commodity output, can be represented by the sum of each row's values. Thus, for each of the n commodities indexed by i, output X_i is:

$$X_i = Z_{i1} + Z_{i2} + \ldots + Z_{in} + Y_i \tag{8-5}$$

However, I-O models are typically generalized instead by representing inter-industry flows between sectors as a percentage of sectoral output. This flow is represented by dividing the economically-valued (transaction) flow from sector i to sector j by the total output of sector j. Namely,

$$A_{ij} = Z_{ij} / X_j \tag{8-6}$$

In such a system, the A_{ij} term is a unitless technical (or input–output) coefficient. For example, if a flow of \$250 of goods goes from sector 3 to sector 4 (Z_{34}), and the total output of sector 4 (X_4) is \$5,000, then $A_{34} = 0.05$. This says that 5 cents worth of inputs from sector 3 is in every dollar's worth of output from sector 4. As a substitution, we can also see from Equation 8-6 that $Z_{ij} = A_{ij} X_j$. This form is more common since the system of linear equations corresponding to Equation 8-5 is typically represented as

$$X_i = A_{i1} X_1 + A_{i2} X_2 + \ldots + A_{in} X_n + Y_i. \tag{8-7}$$

It is straightforward to notice that each X_i term on the left has a corresponding term on the right of Equation 8-7. Thus all X terms are typically moved to the left hand side of the equation and the whole system of equations written as:

$$\begin{gathered}
(1 - A_{11})X_1 - A_{12} X_2 - \ldots - A_{1n} X_n = Y_1 \\
-A_{21}X_1 + (1 - A_{22})X_2 - \ldots - A_{2n} X_n = Y_2 \\
\ldots \\
-A_{i1}X_1 - A_{i2}X_2 - \ldots + (1 - A_{ii})X_i - \ldots - A_{in} X_n = Y_i \\
\ldots \\
-A_{n1}X_1 - A_{n2}X_2 - \ldots + (1 - A_{nn})X_n = Y_n
\end{gathered} \tag{8-8}$$

If we let the matrix **A** contain all of the technical coefficient A_{ij} terms, vector X all the output X_i terms, and vector Y the final demand Y_i terms, then equation system 8-8 can be written more compactly as Equation 8-9:

$$X - \mathbf{A}X =$$

$$[\mathbf{I} - \mathbf{A}]\, X = Y \qquad (8\text{-}9)$$

where **I** is the $n \times n$ identity matrix. This representation takes advantage of the fact that only diagonal entries in the system are $(1-A_{ii})$ terms, and all others are $(-A_{ij})$ terms. Finally, we typically want to calculate the total output, X, of the economy for various exogenous final demands Y, taken as an input to the system. We can take the inverse of [**I – A**] and multiply it on the left of each side of Equation 8-9 to yield the familiar solution

$$X = [\mathbf{I} - \mathbf{A}]^{-1}\, Y \qquad (8\text{-}10)$$

where $[\mathbf{I} - \mathbf{A}]^{-1}$ is the Leontief inverse matrix or, more simply, the Leontief matrix. As discussed in the main part of the chapter, the creation of this inverse matrix transforms the direct requirements matrix into a total requirements matrix. The total requirements matrix mathematically represents all tiers or levels of upstream purchases (instead of just direct purchases) associated with an input of final demand.

Section 2 – Commodities, Industries, and the Make-Use Framework of EIO Methods

The Leontief model described above is very general, discussing output only in terms of which sectors they apply to. For many readers and users, this is sufficient differentiation. In reality, IO tables and models are generally **commodity by industry**, where commodity production sectors i are in the rows and industry sectors j are in the columns. The distinction between commodities and industries is subtle but important. The traditional definition of a commodity is a basic good that is produced widely but equally, for example, white rice. In this traditional definition all companies and facilities make identical, non-distinct products. Industries, on the other hand, combine various commodity inputs and make a new product. This traditional view of commodities is obsolete as these "commodity sectors" in modern IO tables categorize such complex and distinct products as computers and other electronics. The terminology is the only constant.

The simplified tables shown above (e.g., in Figure 8-1) are generally derived from "make and use" tables. A **make table** organizes economic data related to which industry sectors make which commodities, while a **use table** organizes economic data related to which industries use which commodities. While a full mathematical description of converting from make-use tables to transactions tables is beyond the scope of this chapter (and better left for the teacher or student to implement), the make-use framework forms the foundation of the commodity-by-industry transactions table introduced in the chapter. The matrix math involved in transforming make-use tables into transactions tables internalizes and allocates the multiple productions and uses of commodities into a unified set of production recipes. Matrix math also allows us to generate different formats of transactions tables from the original make-use format, e.g., an industry-industry or commodity-commodity table format.

The columns of a make table show the distribution of industries producing a commodity, while the rows show the distribution of commodities produced by an industry. If you read across the values of a row in the make table, you see all of the commodity outputs that each sector *makes*. The make table might reveal that in fact several sectors are responsible for "making" a particular commodity, e.g., a steel facility and a power plant may both produce electricity. Figure 8A-1 shows an excerpt of actual data from the 2002 Make Table of the US economy.[20] It shows that the vast majority of farm commodities are produced by the farm industry ($197 billion), and that billions of dollars of forestry commodities are produced in both the farm and forestry sectors.

Industries/ Commodities	**Farms**	**Forestry, fishing, and related activities**	**Oil and gas extraction**	**Mining, except oil and gas**	**Support activities for mining**	**Utilities**
Farms	197,334	3,306	...	...	...	...
Forestry, fishing, and related activities	19	38,924	...	...	...	...
Oil and gas extraction	...	...	91,968	133	1,301	...
Mining, except oil and gas	...	...	1	47,270	163	...
Support activities for mining	...	...	33	86	32,074	...
Utilities	...	...	33	...	...	316,527

Figure 8A-1: Excerpted 2002 Make Table of the US economy ($millions)

Use Tables follow the "production recipe" style mentioned earlier in the chapter. If you read down the values of a column in the Use Table, you see all of the economic inputs needed from other sectors, i.e., you see how much the industry sector *uses* from the commodity sectors. The

[20] The Make and Use Tables in Figures 8A-1 and 8A-2 are excerpted from aggregated tables with about 80 sectors of resolution, not the 428 sectors in the benchmark tables. Less than 10 commodity and industry sectors are excerpted, so 70 other columns of data are not shown.

rows show where the commodity outputs of sectors are used. Figure 8A-2 shows excerpted but actual data from the 2002 Use Table of the US economy. It shows that the utilities industry, which includes power generation, uses billions of dollars of oil, gas, and mined (coal) commodities. It also shows that a large share of the production of wood products ($18 billion) were used by the construction industry in 2002.

Commodities/ Industries	Oil and gas extraction	Mining, except oil and gas	Utilities	Construction
Farms	...	0	0	1,098
Oil and gas extraction	25,171	3	45,170	...
Mining, except oil and gas	487	8,028	4,503	8,206
Support activities for mining	5,354	2,378	...	6
Utilities	3,608	3,981	108	2,895
Construction	13,957	2	3,952	733
Wood products	0	9	31	18,573
Nonmetallic mineral products	314	615	52	36,028

Figure 8A-2: Excerpted 2002 Use Table of the US economy ($millions)

Make and Use Tables often exist with classifications of "before and after redefinitions". The figures above are both before redefinitions. While the various methods of redefinition vary by the agency creating the tables, typically the process of redefinition involves carefully remapping secondary activities within established sectors to other sectors. As an example, the hotel industry typically has restaurants and laundry services on site, which are represented by separate sectors in tables. As the data available supports it, activities within the hotel industry are re-mapped into those other sectors (e.g., food purchases are switched from the hotel industry to the restaurants industry). This affects both the make and use tables, and the industry outputs are different between the versions of the tables with and without redefinitions. In the end, some sectors' production recipes are basically unchanged by redefinitions, while others are substantially changed. Since such redefinitions lead to better-linked representations of the activities that could lead to energy and environmental impacts, they are typically the basis of IO-LCA models.

Section 3 – Further Detail on Prices in IO-LCA Models

Adjusting Values to Match Basis Year of EIO Models

As represented in Figure 8-2, one of the critical inputs to an EIO model is an increment of final demand to be studied. The appropriate "unit" of this final demand is a currency-valued input the same year as that of the model. If using a 2002 US EIO model, then a final demand in 2002 dollars is needed. If you are using the model to assess the impacts of automobile production in 2013, then you need to find a method of adjusting from 2013 to 2002 dollars for the final demand, since it is likely that prices in sectors have changed significantly since the year of the model. However, since the intention is to perform a screening-level analysis, you can exploit the fact that production recipes (technologies) do not change quickly, and assume that the only relevant difference between a 2013 and 2002 vehicle is the price (producer or purchaser).

For such conversions, the appropriate type of tool is an economic price index or GDP deflator for a particular sector. These are generally available from national economic agencies (in the US they are provided by the BEA, the same agency that creates the input-output tables, which leads to consistent comparisons). Note that an overall national price index or GDP deflator may be the only such conversion factor available. In this case, it can be used but the adjustment should be clearly documented as using this national average rather than a sector-specific value.

A full discussion of price indices and deflators is beyond the scope of this book, but are typically represented as values where there is a "base year" with an index value of 100, and values for years before and after the base year. Such values could be, e.g., 98 and 102, which if before and after the base year would suggest annual price changes of about 2% per year. It is the percentage equivalent values of the index values that are useful when using indexes to adjust values from present day back to the basis year of EIO models. Note that the base year of the index does not need to match the year of the EIO model; as long as you can use the index values to adjust dollar values back and forth, you can adjust current values back to the appropriate final demand value for the right year (or for any other year you might care about). Equation 8A-1 shows how you can convert values from one basis year to another using a price index (or a GDP deflator represented with a base=100 format):

$$\frac{Value_1}{Value_2} = \frac{Price\ index_1}{Price\ index_2} \tag{8A-1}$$

For example, assume the average retail price of automobiles in 2011 is known to be $30,000, and we want to find the corresponding retail price to be used as the final demand in a 2002 US EIO purchaser basis model so that we could try to estimate the effects of manufacturing a single automobile in 2011. The BEA provides spreadsheets of various economic time-series estimates, such as Gross Output, including price indices, by sector.[21] For example, price index

[21] As of 2014 the file, named GDPbyInd_GO_NAICS_1998-2011.xls, can be found at http://www.bea.gov/industry/gdpbyind_data.htm

values for the *Automobile manufacturing* sector (#336111), of 98.75 for 2002 and 99.4 for 2011, are shown in Figure 8A-3.

Year	Index Value
1998	100.645
1999	100.451
2000	101.488
2001	100.902
2002	**98.750**
2003	98.748
2004	100.286
2005	100.000
2006	97.168
2007	96.280
2008	98.185
2009	99.918
2010	98.653
2011	**99.400**

Figure 8A-3: Price Index Values for Automobile Manufacturing Sector, 1998-2011 (Source: BEA)

Thus, the converted 2002 value can be found by applying equation 8A-1, as shown in Equation 8A-2:

$$\frac{Value_{2011}}{Value_{2002}} = \frac{\$30{,}000}{Value_{2002}} = \frac{99.4}{98.75} \quad (8A\text{-}2)$$

In this case, the adjusted value for 2002 is $29,800. It may be surprising that the price level has been almost unchanged over those ten years! One could instead use the negligible (less than 1%) price level change as the basis of an assumption to ignore the need for adjustment and just use $30,000 directly as the input final demand into the model.

Differences Between Producer And Purchaser Models

It is important to understand the multiple ways in which price bases can be defined in input-output systems. While we primarily discussed and assumed the producer basis in Chapter 8, there are other ways as well, as defined by the UN System of National Accounts (UN 2009):

Basic prices are the amount received by a producer from a purchaser for a good or service, minus any taxes, and plus any subsidies (this is referred to as **net taxes**). This basis typically excludes transport charges that are separately invoiced by the producer. You might more simply consider basic prices as the raw value of a product before taxes or subsidies are considered. **Producer prices** are the amount received by a producer from a purchaser, plus any taxes and minus any subsidies. Producer prices are equivalent to the sum of basic prices and net taxes. Finally, **purchaser prices** are the amount paid by the purchaser and include the cost of delivery (e.g., transportation costs) as well as additional amounts paid to wholesale and retail entities to make it available for sale. These transportation and wholesale/retail components are referred to as margins.

Example: Basic, producer and purchaser prices

To illustrate these different ways of describing prices, consider this example from Statistics New Zealand (2012). Generic currency units of $ are used.

Item	Amount
Basic price	**12**
+ Taxes on product, except sales tax or VAT	1
- Subsidy on products	5
= Producer's price	**8**
+ Sales tax or VAT	2
+ Transport charges and trade margins paid by purchaser	3
= Purchaser's price	**13**

Note: VAT = Value Added Tax (used in many parts of the developed world)

In this example, the seller is actually able to retain $12 for the product (basic price). The sales transaction takes place at $8 (producer's price). The seller gets an additional $4 from the subsidy, less the tax. The purchaser has to pay $13 to take possession of the good (producer's price), with $5 going to non-deductible taxes and transport charges and trade margins.

With respect to LCA, as we discussed, a producer basis is a cradle to gate scope, while a purchaser basis adds in transport and wholesale/retail operations (assuming pickup at store) and is thus cradle to consumer in scope. In some cases, the purchaser and producer price bases are approximately the same. However, if transportation costs or retail markups needed to bring the product to market are significant, there will be a difference.

Figure 8A-4 gives total sectoral output values in producer price and purchaser price bases in the 1997 US benchmark accounts. Note that these values are not provided to help 'convert' values between the models as done above with price indexes, but instead, to demonstrate why the prices and model results are different. Service sectors like barbershops have identical producer and purchaser prices, because the service is produced at the point of purchase. On the other hand, the purchaser price of furniture is roughly split 50-50 between manufacture and the wholesale/retail margin activities needed to market the product.

Item	Producer Price	Transportation Cost	Wholesale and Retail Trade	Purchaser Price	Producer / Purchaser Price Ratio (%)
Shoes	$18,333	$179	$21,748	$40,259	45
Barber shops	$31,246	—	—	$31,246	100
Furniture	$28,078	$235	$27,648	$55,960	50

Figure 8A-4: Differences in Producer and Purchaser Prices (Millions of 1997 Dollars in Sector Output)

When a purchaser price model is used, a dollar of input of final demand to a single sector is distributed into these shares of the various underlying sectors (like creating a final demand vector with multiple entries instead of just a single value for the production sector). This data is often represented on a normalized (percentage basis) instead of using total values. For example, $1,000 of final demand of furniture in a purchaser price model will have a final demand of about $500 to the furniture manufacturing sector, a small amount for transportation, and about $500 to the wholesale and retail trade sectors together. This will be discussed in more detail in Advanced Material Section 6. Depending on the energy or environmental intensity of the various sectors, the results for a given final demand using a producer versus purchaser model may be higher, lower, or about the same. For example, truck transportation (one of the margin sectors included in a purchaser basis model) is fairly carbon intensive. If the purchased product has significant transportation requirements, the purchaser model might have higher emissions than the producer model.

References

Statistics New Zealand, "Introduction to Price Indices", online course materials, http://unstats.un.org/unsd/EconStatKB/KnowledgebaseArticle10351.aspx, posted June 14 2012. Last accessed January 10, 2014.

United Nations, System of National Accounts, 2008. New York, 2009.

Section 4 – Mapping Examples from Industry Classified Sectors to EIO Model Sectors

The organization of data for IO-LCA models is a substantial exercise requiring quality checking and assurance processes. Economic matrices (e.g., an **A** matrix) are typically provided directly by agencies and at worst typically only require minimal conversion or preparation for use in EIO models. **R** matrices, on the other hand, require significant effort. In this section, we focus on explaining how the various different industry classification methods map to each other in support of making these matrices.

In the US, the primary classification scheme for industries (and the businesses within them) is the North American Industry Classification System (NAICS). While the US government has officially decreed that all industry data collection efforts shall use NAICS, some data sources have not yet completely converted to this system. NAICS is a hierarchical classification system with values ranging from 2 to 6 digits. Sectors are broadly categorized by the first two digits, and then sub-classified by appending additional digits. For example, manufacturing sectors start with the first two digits 31-33. Three digit sector numbers (e.g., 311, 312, … , up to 339) further classify manufacturing into activities like food manufacturing and miscellaneous manufacturing. The three digit sector values can be similarly broken up into more specific manufacturing categories which can be described with 4-digits sector numbers. (e.g., 3111, 3112, etc.), Six-digit sectors are the most detailed (and least aggregated) classifications of activity in the economy. For example, the *Automobile manufacturing* sector discussed at various times in this chapter is classified hierarchically in the NAICS system as follows:

NAICS 33 Manufacturing (note 31-33 are all classified in the same way)

NAICS 336 Transportation equipment manufacturing

NAICS 3361 Motor vehicle manufacturing

NAICS 33611 Automobile and light truck manufacturing

NAICS 336111 Automobile Manufacturing

There are of course many other complementary manufacturing subsectors throughout that hierarchy that are not shown. The full official US Census Bureau NAICS classification is available on the Internet (at http://www.census.gov/eos/www/naics/).

While the Census Bureau (via BEA) is also the creator of the input-output tables, they do not simply define the sectors of the input-output table to correspond precisely to 6-digit NAICS industries or commodities. As mentioned in the chapter, they balance available resources against the need to produce a sufficiently detailed input-output table. Thus, of the 428 sectors in the 2002 US input-output model, relatively few correspond directly to 6-digit NAICS codes (though these are mostly in the manufacturing sectors), many IO sectors map to 5-digit NAICS codes, and a significant number map to 3- and 4-digit level NAICS codes.

Beyond the mapping of IO sectors to n-digit NAICS level, many IO sectors are not simple one-to-one mappings, meaning the IO sectors represent aggregations of multiple underlying NAICS codes. Figure 8A-5 shows a summary of how NAICS codes map into the first set of sectors of the 2002 US IO detailed models from BEA. In the left hand column is a subset of the hierarchical classifications of IO sectors.

I-O Industry Code and Title			**Related 2002 NAICS Codes**
11	**AGRICULTURE, FORESTRY, FISHING AND HUNTING**		
1110	**Crop production**		
	1111A0	Oilseed farming	11111-2
	1111B0	Grain farming	11113-6, 11119
	111200	Vegetable and melon farming	1112
	1113A0	Fruit farming	11131-2,111331-4, 111336*, 111339
	111335	Tree nut farming	111335, 111336*
	111400	Greenhouse, nursery, and floriculture production	1114
	111910	Tobacco farming	11191
	111920	Cotton farming	11192
	1119A0	Sugarcane and sugar beet farming	11193, 111991
	1119B0	All other crop farming	11194, 111992, 111998

Figure 8A-5: Correspondence of Crop Production NAICS and IO Sectors, 2002 US Benchmark Model (Source: Appendix A)

In the right hand column are the NAICS level sectors that are mapped into each of the detailed IO sectors. For example, two 5-digit NAICS sectors (11111 and 11112) map into the *Oilseed farming* sector. A single 4-digit sector (1112) maps into the *Vegetable and melon farming* sector. Various 5 and 6-digit level NAICS sectors map into the *Fruit farming* IO sector. The asterisk next to 111336 notes that output from that sector is not 1:1 mapped into a single sector. As you can see, some of NAICS 111336's output is mapped into the *Tree nut farming* sector below it. Fortunately, the names of the IO sectors tend to be very similar or identical to the NAICS sector names (not shown above but available on the Census NAICS URL above), so following the mapping process is a bit easier.

While the discussion above is motivated by how the IO transactions tables are created, it is also critical to understand because of how the classifications and mappings affect the creation of **R** matrices. Since each value of an **R** matrix is in units of effects per currency unit of output for a sector, we need to ensure that data on energy and environmental effects for a sector have been correctly mapped into IO sectors. In other words, we need to ensure that the **R** matrix values (numerators and denominators) have been derived correctly.

Reconsider Example 8-4 from the chapter. If instead of trying to find SO_2 emissions for the power generation sector, imagine you were deriving an **R** matrix of fuel use by sector. As defined in Figure 8A-5, the fuel use of the *Oilseed farming* IO sector (1111A0) would be found by finding data on the fuel use of NAICS sectors 11111 and 11112, and adding them together. Finally, this sum would be normalized by the output of the oilseed farming sector (from the Use Table) and the result would be the entry of the **R** matrix for that sector. As another example, the **R** matrix value for the *Greenhouse, nursery, and floriculture production* (111400) sector requires data on fuel use from just one 4-digit NAICS sector, 1114.

The mapping process for building **R** matrices seems simple, and is conceptually. However, data for the required level of aggregation (4, 5, or 6-digit) is often unavailable. When you have data at one aggregation level, but need to modify it for use for another level, assumptions need to be made and documented.

If you have more detailed data, but need more aggregated data, the process is generally simple. You can aggregate (sum) 6-digit NAICS data into a single 5-digit level. However when the data only exists at an aggregate (e.g., 3 or 4-digit NAICS) level then you need to create ways of allocating the aggregate data into more disaggregated 4, 5, or 6-digit level sectors.

The following example shows the challenges present in mapping available data to the corresponding IO sectors. It represents actual data available for use in the 2002 US benchmark IO model, at the 428 sector level. Figure 8A-6 shows an excerpt of the NAICS to IO mapping for the 29 *Food manufacturing* sectors. For these 29 sectors, the required level of aggregated NAICS data ranges from the 4 to 6-digit levels. Creating the **R** matrix for every type of fuel would require at least 29 different energy values that would then be divided by sectoral output. Figure 8A-7 shows available data on energy use from food manufacturing sectors from the US Department of Energy (MECS 2002).

I-O Industry Code and Title			Related 2002 NAICS Codes
31	**MANUFACTURING**		
3110	**Food manufacturing**		
	311111	Dog and cat food manufacturing	311111
	311119	Other animal food manufacturing	311119
	311210	Flour milling and malt manufacturing	31121
	311221	Wet corn milling	311221
	31122A	Soybean and other oilseed processing	311222-3
	311225	Fats and oils refining and blending	311225
	311230	Breakfast cereal manufacturing	311230
	31131A	Sugar cane mills and refining	311311-2
	311313	Beet sugar manufacturing	311313
	311320	Chocolate and confectionery manufacturing from cacao beans	31132
	311330	Confectionery manufacturing from purchased chocolate	31133
	311340	Nonchocolate confectionery manufacturing	31134
	311410	Frozen food manufacturing	31141
	311420	Fruit and vegetable canning, pickling, and drying	31142
	31151A	Fluid milk and butter manufacturing	311511-2
	311513	Cheese manufacturing	311513
	311514	Dry, condensed, and evaporated dairy product manufacturing	311514
	311520	Ice cream and frozen dessert manufacturing	311520
	31161A	Animal (except poultry) slaughtering, rendering, and processing	311611-3
	311615	Poultry processing	311615
	311700	Seafood product preparation and packaging	3117
	311810	Bread and bakery product manufacturing	31181
	311820	Cookie, cracker, and pasta manufacturing	31182
	311830	Tortilla manufacturing	31183
	311910	Snack food manufacturing	31191
	311920	Coffee and tea manufacturing	31192
	311930	Flavoring syrup and concentrate manufacturing	31193
	311940	Seasoning and dressing manufacturing	31194
	311990	All other food manufacturing	31199

Figure 8A-6: Correspondence of Food Manufacturing NAICS and IO Sectors, 2002 US Benchmark Model (Source: Appendix A)

NAICS Code	Sector Name	Total	Net Electricity	Residual Fuel	Distillate Fuel	Natural Gas
311	Food	1,116	230	13	19	575
311221	Wet Corn Milling	217	23	*	*	61
31131	Sugar	111	2	2	1	22
311421	Fruit and Vegetable Canning	47	7	1	1	36

Figure 8A-7: NAICS Level Fuel Use Data From Manufacturing Energy Consumption Survey, units: trillion BTU (Source: MECS 2002)

As you can see, the immediate challenge is that only 4 different sectors of results (rows) are available from MECS, the best available data source. One of the sectors in MECS is a value of energy use for the entire 3-digit NAICS *Food manufacturing* sector (311). Estimates of energy use are provided for only three more detailed food manufacturing sectors: *Wet corn milling* (311221), *Sugar* (31131), and *Fruit and vegetable canning* (311421). The reasons why only these sectors were estimated is not provided, but presumably again seeks to balance data quality, budget resources, and resulting resolution. Regardless, the MECS data provided for 311221 maps perfectly into IO sector 311221. The MECS data for 31131 needs to be split into values for IO sectors 31131A and 311313. The MECS data for 311421 can be put into IO sector 311420 (but may be missing data for sectors 311422, etc.). So the 5 and 6-digit data from MECS can only at best help us with specifically mapped data into 4 of the 29 sectors. For the remaining 25 sectors, we need to find a method to allocate total energy use data from the 3-digit NAICS level for 311 as shown in the first row of Figure 8A-7. It is not as easy as taking these values for sector 311 and allocating because the values provided (e.g., 1,116 trillion BTU of total energy use) already include the energy use of three detailed sectors below in the table. Thus, the energy of the other 25 sectors to be allocated is the difference between the values provided in the NAICS 311 row and the three detailed rows. In this example, total energy use of the 25 other sectors is 1,116 – 217 – 111 - 47, or 741 trillion BTU (with similarly calculated values for the other fuels).

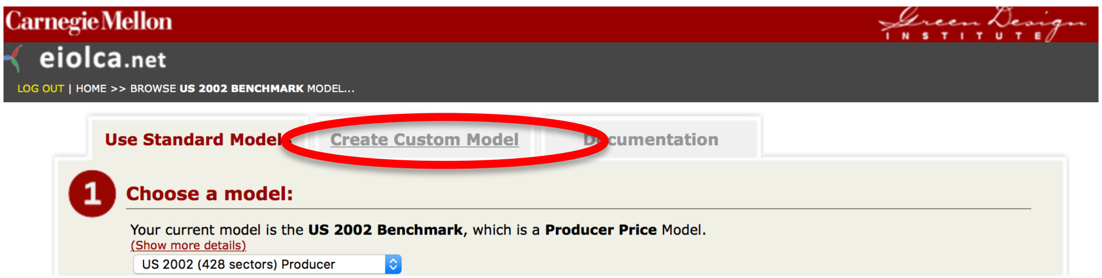
In EIO-LCA, the allocation method used to distribute 741 trillion BTU of energy use into the other 25 food manufacturing sectors is to use the dollar amounts from the 2002 Use Table as weighted-average proxies for consumption of each energy source, which assumes that each sector within a sub-industry is paying the same price per unit of energy. For the case of the 1997 and 2002 Benchmark IO models for the US, more complete documentation of how various effects have been derived is available in the EIO-LCA documentation (http://www.eiolca.net/docs). Other IO-LCA models may make different assumptions to allocate the available data.

Hopefully, this discussion into the lack of consistency in the aggregation and organization of data for IO-based sectoral analysis helps you to appreciate the complexity in creating such models that are, in the end, simple to use!

Section 5 – Modeling Effects of Multiple Final Demand Entries

In this section we briefly demonstrate an advanced feature of the EIO-LCA website which allows a user to estimate the effects of multiple final demand entries into the same model. This is one of two custom advanced models available (the other is described in the Advanced Material of Chapter 9).

From the 'Use the Model' page of EIO-LCA, select the 'Create Custom Model' tab in the top center of the page (as shown in the screenshot below). On the resulting page, select the 'custom product' link.



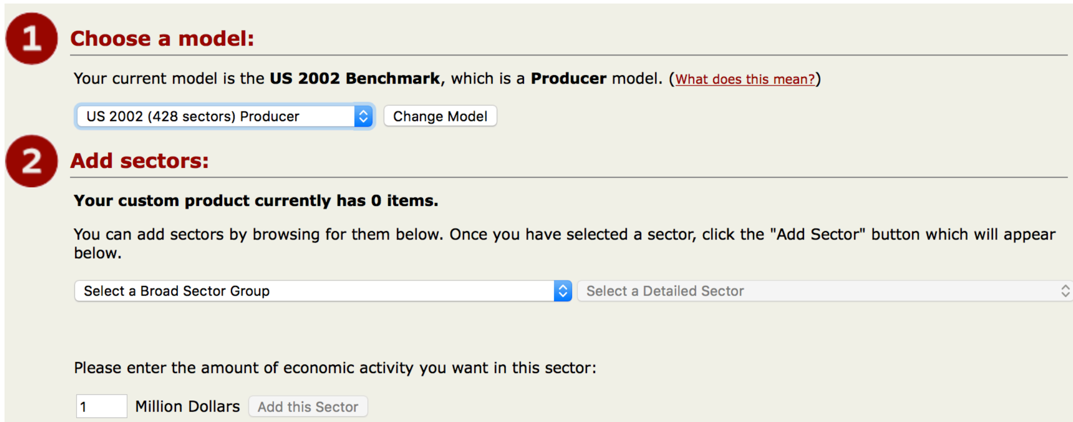
The custom product tool allows you to do an expanded input-output analysis of a more complex production process. Instead of only looking at the effects of demand from a single sector, you can consider effects of increased production in several sectors simultaneously (i.e., as described in the chapter it allows you to consider the effects of multiple final demand entries simultaneously).

There are two primary applications of this tool in terms of types of products to model. Of course, either of these options still has the same advantages and disadvantages of any IO-LCA model in terms of aggregation, representative production, etc.

- Hypothetical products - You can consider the implications of a product that is not currently represented in the input-output model, e.g., a 'recipe' of the requirements needed to build an electronic book reader. These products are similar to laptop computers, but not part of any existing sector production (as of the 2002 model used).

- Improved Analysis of Existing Products - you can perform a more specific analysis of an existing product. For example, instead of just looking at $20,000 of production from the vehicle production sector, you could put together a recipe of similar but different items needed in the supply chain of a hybrid vehicle. This may more closely approximate your desired product.

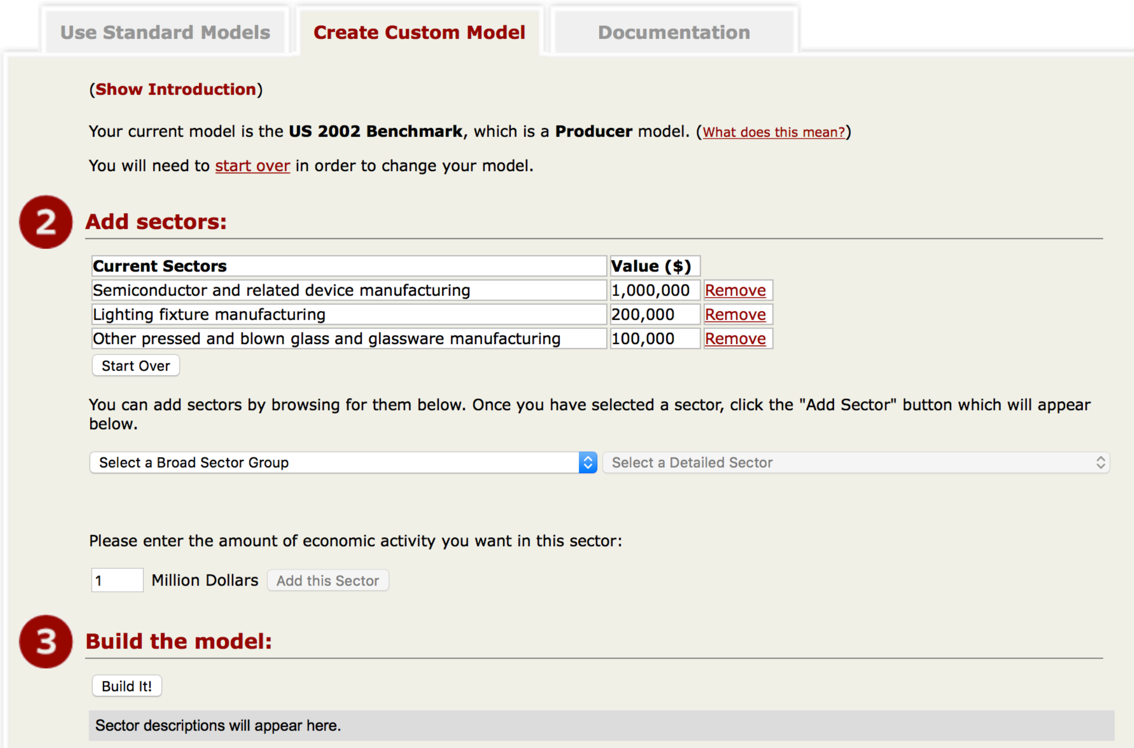
To use the custom product builder, the first step is to choose the EIO-LCA 'model year' to be used (e.g., 1997 or 2002) and click the 'Change Model' button, as shown in the screenshot below.



The second step uses the same two-level sector selection interface as used in the basic EIO-LCA model to iteratively add the sectors required to produce your good or service (i.e., you are selecting the various entries of the final demand vector Y for your model). For each separately required sector (final demand component), you choose the sector, enter a final demand in millions of dollars, and then click the 'Add this Sector' button. The third and final step is done when all components of final demand in your custom model have been entered, and you click the 'Build It' button at the bottom of the screen.

The screenshot below shows what the website would look like when using the 2002 EIO-LCA Producer price model for a custom build model to approximate LED lamp production, including the following three sectoral components of final demand into the model:

- $1 million of *Semiconductor and related device manufacturing* for the LED wafer production.
- $0.2 million of *Lighting fixture manufacturing* for the fixture components of the bulb.
- $0.1 million of *Other pressed and blown glass and glassware manufacturing* for the glass encasing for the bulb.



As above, the last step (#3) would be to click the 'Build It' button, which would return the default EIO-LCA economic result summary shown below. From this screen, you can click the 'Change Inputs' button to instead show results for energy, greenhouse gases, etc., across the economic supply chain.

In the parlance of Chapter 8, this example of the custom builder tool to estimate the effects of LED lamp production is using a final demand as shown below.

$$Y = \begin{bmatrix} 1 \\ 0.2 \\ 0.1 \end{bmatrix}$$

The results generated by the custom tool in this case are identical to those that would be obtained from separately running the 2002 EIO-LCA Producer model three times (once for each of the sectoral final demand inputs), saving the results to a spreadsheet, and then summing the results. The custom tool is merely a way of saving this added manual effort and doing it all at once.

Custom Product
Displaying: Economic Activity
Number of Sectors: Top 10

Change Inputs (Click here to view greenhouse gases, air pollutants, etc...)

Documentation:
The environmental, energy, and other data used and their sources.
Frequently asked questions about EIO-LCA.

This sector list was contributed by Green Design Institute.

Your Custom Product

Ingredients	Weight
Semiconductor and related device manufacturing	1000000
Lighting fixture manufacturing	200000
Other pressed and blown glass and glassware manufacturing	100000

	Sector	Total Economic $mill	Total Value Added $mill	Employee Comp VA $mill	Net Tax VA $mill	Profits VA $mill	Direct Economic $mill	Direct Economic %
	Total for all sectors	*2.61*	*1.28*	*0.691*	*0.051*	*0.541*	*1.99*	*76.2*
334413	Semiconductor and related device manufacturing	1.03	0.495	0.208	0.007	0.281	1.02	99.1
335120	Lighting fixture manufacturing	0.181	0.076	0.041	0.000	0.034	0.181	99.8
550000	Management of companies and enterprises	0.150	0.093	0.079	0.002	0.012	0.110	72.9
420000	Wholesale trade	0.118	0.082	0.044	0.019	0.018	0.081	69.1
327215	Glass Product Manufacturing Made of Purchased Glass	0.059	0.023	0.016	0.000	0.007	0.058	97.9
327212	Other pressed and blown glass and glassware manufacturing	0.052	0.025	0.018	0.000	0.007	0.051	98.4
325188	All other basic inorganic chemical manufacturing	0.039	0.012	0.011	0.000	0.000	0.034	88.0
221100	Power generation and supply	0.034	0.023	0.007	0.004	0.012	0.016	47.8
541700	Scientific research and development services	0.033	0.019	0.018	0.000	0.001	0.025	75.5
531000	Real estate	0.027	0.021	0.002	0.003	0.017	0.006	22.4

If you need additional help with the custom builder tool, see the online EIO-LCA screencasts (at http://www.eiolca.net/tutorial/ScreencastTutorial/screencasts.html).

End of Section Questions

1. Redo question #2 (the washing machine comparison) from the End of Chapter questions, but using the custom model tool described in this section. Again, reflect on the caveats of using the EIO-LCA model for this comparison, and also comment on whether the custom tool introduces any particular new issues in interpreting the results.

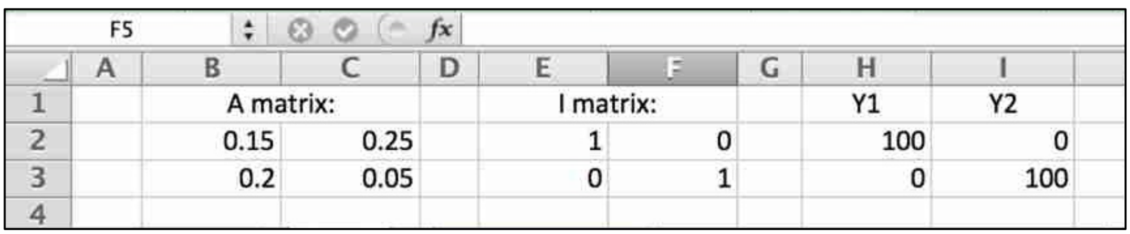
Section 6 – Spreadsheet and MATLAB Methods for Using EIO Models

In this section we overview the specific use of Microsoft Excel and MATLAB in support of linear algebra / matrix math manipulations of IO models. This section is not intended to be an introduction to using these two software tools, or as an introduction to linear algebra.

Modeling Using Microsoft Excel Software

Despite its cost, Microsoft Excel is a ubiquitous spreadsheet software already installed on many computers. Other spreadsheet programs (such as those from OpenOffice) use very similar methods to those described here. For relatively small projects, spreadsheets can be very useful in organizing LCA data and in assisting with matrix calculations.

In Excel, elements of vectors and matrices are easy to enter by hand (for small matrices) and also by pasting or importing data from other sources. A 1 row by 5 column or 5 row by 5 column area of a spreadsheet can be generated quickly. Returning to Example 8-2, the **A** and **I** matrices, and vectors Y_1 and Y_2 could be entered into Excel as in Figure 8A-8.



	A	B	C	D	E	F	G	H	I
1		A matrix:			I matrix:			Y1	Y2
2		0.15	0.25		1	0		100	0
3		0.2	0.05		0	1		0	100
4									

Figure 8A-8: Data Entry in Microsoft Excel for Example 8-2

However, such entries, despite looking like a matrix, would not be treated as such in Excel. All cells in Excel are by default treated individually. To be recognized as a vector or matrix, Excel requires you to create arrays. This can be achieved in one of two ways. The most convenient way of making re-usable arrays in Excel is to highlight the entire area of the matrix (e.g., the 1 by 5 or 5 by 5 series of cells created above) and to use the built-in naming feature of Excel. For example, we can highlight the cell range B2:C3 and then move the cursor to the small box between the "Home" ribbon bar and cell A1 and type in "A", to designate this set of cells as the **A** matrix, as shown in Figure 8A-9. The same can be done for **I**, Y_1, and Y_2, however note that you can not use "Y1" and "Y2" as Excel names because there are already Excel cells with those names (in column Y of the spreadsheet) – you must instead name them something like "Y_1" and "Y_2". These names of specific cells or groups of cells help you to create more complex cell formulas as they act like aliases or shorthand notations that refer to the underlying cell ranges. In practice, instead of having to enter the cell range (e.g., B2:C3) and potentially making typos in formulas, you can instead just use the name you have assigned. This is useful with matrix math because it is easier to ensure you are multiplying the correct vectors and matrices by using their names instead of cell ranges.

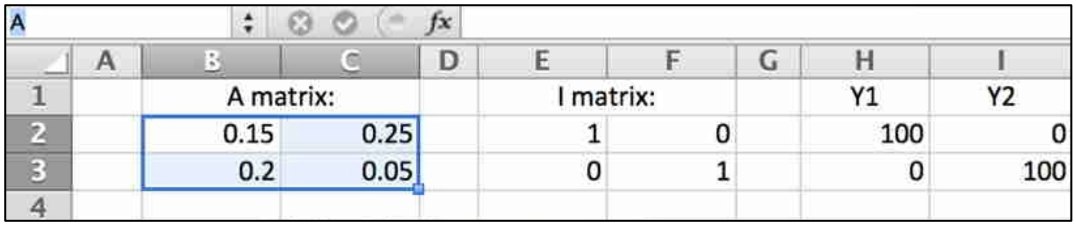


Figure 8A-9: Named A Matrix in Microsoft Excel for Example 8-2

Once you have made names for your data ranges, you can use built-in Excel matrix math functions like multiplication and inversion. Addition and subtraction of vectors or matrices of the same dimensions (m x n) can be done with regular + and – operators. However, you need to help Excel realize that you are making an array and set aside space for it to be created based on knowing its dimensions. To find **I+A** as in Example 8-2, you need to first select an unused cell range in your spreadsheet that is 2 x 2 (and optionally name it, e.g., IplusA, and press enter), then type the equal sign (=), enter the formula (I+A), and then click CTRL-SHIFT-ENTER. This multi-step process tells Excel that you want the results of the matrix operation I+A to be entered into your selected cell range, to add the previously named references I and A, and to generate the result with array formulas (thus the CTRL-SHIFT-ENTER at the end). The screenshots in Figures 8A-10 and 8A-11 show the intermediate and final steps (before and after typing CTRL-SHIFT-ENTER) of this process in Excel. Note that after CTRL-SHIFT-ENTER has been typed, Excel modifies the cell formula such that curly brackets are placed around the formula (not shown), denoting the use of an array function as applied to a cell in the named range.

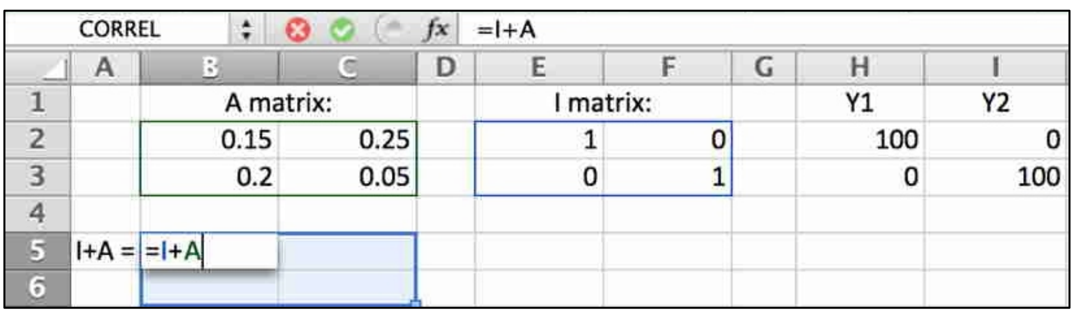


Figure 8A-10: Entering Array Formula for Selected Area in Microsoft Excel for Example 8-2

As shown in the last Advanced Material Section, multiple Y_i entries of final demand can also be modeled, e.g., a final demand Y_3 could have \$100 inputs into both sectors (not shown here).

IplusA

	A	B	C	D	E	F	G	H	I
1		A matrix:			I matrix:			Y1	Y2
2		0.15	0.25		1	0		100	0
3		0.2	0.05		0	1		0	100
4									
5	I+A =	1.15	0.25						
6		0.2	1.05						

Figure 8A-11: Result of Array Formula in Microsoft Excel for Example 8-2

Multiplication and inversion of matrices uses the same multi-step process, but use the built-in functions MMULT and MINVERSE. You can use the MMULT and MINVERSE functions by typing them into the formula bar or by using the Excel "Insert->Function" dialog box helper. As with the example shown above, as long as you first select the cell range of the expected result (with the appropriate m x n dimensions), enter the formula, and click CTRL-SHIFT-ENTER at the end, you will get the right results. You will see an error (or a result in only one cell) if you skip one of the steps. While a bit cumbersome, using array functions in Excel is straightforward and very useful for small vectors and matrices. Figure 8A-12 shows a screenshot where $[\mathbf{I\text{-}A}]^{-1}$, $[\mathbf{I\text{-}A}]^{-1} Y_1$, and $[\mathbf{I\text{+}A}]\ Y_1$ have been created.

E-resource: A Microsoft Excel file solving Examples 8-1 through 8-3 is posted to the textbook website.

B15

	A	B	C	D	E	F	G	H	I
4									
5	I+A =	1.15	0.25		I+A*Y1 =	115		I+A*Y2 =	25
6		0.2	1.05			20			105
7									
8	I-A =	0.85	-0.25		I-A^{-1} =	1.25413	0.33		
9		-0.2	0.95			0.26403	1.12		
10									
11					I-A^{-1}*Y_1=	125.413		I-A^{-1}*Y_2=	33.0033
12						26.4026			112.2112

Figure 8A-12: Result of Matrix Math in Microsoft Excel for Example 8-2

Note that you can perform vector and matrix math without using the Excel name feature. In this case, you would just continue using regular cell references (e.g., B2:C3 for the **A** matrix in the screenshot above). All of the remaining instructions are the same.

Brief MATLAB Tutorial For IO-LCA Modeling

This short primer on using Mathworks MATLAB is no substitute for a more complete lesson or lecture on the topic but will help you get up to speed quickly. It presumes you have MATLAB installed on a local computer with the standard set of toolboxes (no special ones required). MATLAB, unlike Microsoft Excel, is a high-end computation and programming environment that is often used when working with large datasets and matrices. It is typically available in academic and other research environments.

When MATLAB is run, the screen is split into various customizable windows. Generally though, these windows show:

- the files within the current directory path,
- the command window interface for entering and viewing results of analysis,
- a workspace that shows a listing of all variables, vectors, and matrices defined in the current session, and
- a history of commands entered during the current session.

In this tutorial, we focus on the command line interface and the workspace. Despite the brevity of the discussion included here, one could learn enough about MATLAB in an hour to replicate all of the Excel work above.

MATLAB has many built-in commands, and given its scientific computing specialties, is designed to operate on very large (thousands of rows and columns) matrices when installed. Some of the most useful commands and operators for use with EIO models in MATLAB are shown in Figure 8A-11. Many commands have an (x,y) notation, where x refers to rows and y refers to columns. Others operate on whole matrices.

Working with EIO matrices in MATLAB involves defining matrices and using built-in operators much the same way as was done in the Excel examples above. Matrices are defined by choosing an unused name in the workspace and setting it equal to some other matrix or the result of an operation involving commands on existing matrices. MATLAB commands are entered at the command line prompt (>>) and executed by pressing ENTER, or placed all in a text file (called an '.m file') and run as a script. If commands are entered without a semicolon at the end, then the results of each command are displayed on the screen in the command window when ENTER is pressed. If the semicolon is added before pressing ENTER, then the command is executed, but the results are not shown in the command window. One could look in the workspace window to see the results.

Command	Description of Result
zeros(x,y)	creates a matrix of zeros of size (x,y). This is also useful to "clear out" an existing matrix.
ones(x,y)	same as zeros, but creates a matrix of all ones.
eye(x)	creates an identity matrix of size (x,x). Note the command is not I(x), a common confusion.
inv(X)	returns the matrix inverse of X.
diag(X)	returns a diagonalized matrix from a vector X, i.e, where the elements of the vector are the diagonal entries of the matrix (like the identity matrix).
sum(X)	returns an array with the sum of each column of the input matrix. If X is a vector then the command returns the sum of the column.
size(X)	tells you the size of a matrix, returning (number of rows, number of columns). This is useful if you want to verify the row and column sizes of a matrix before performing a matrix operation.
A'	performs a matrix transpose on A, inverting the row and column indices of all elements of the matrix.
A*B	multiplies matrices A and B in left-to-right order and with usual linear algebra.
A.*B	element-wise multiplication instead of matrix multiplication, i.e., A_{11} is multiplied by B_{11} and the results put into 'cell$_{11}$' of the new matrix (A and B must be the same size).
[A,B]	concatenates A and B horizontally.
[A;B]	concatenates A and B vertically.
clear all	empties out the workspace and removes all vectors, matrices, etc. Like a reset.

Figure 8A-11: Summary of MATLAB Commands Relevant to EIO Modeling

In this section, `courier` font is used to show commands typed into, or results returned from, MATLAB. For example, the following commands, entered consecutively, would "clear out" a matrix named "test_identity" and then populate its values as a 2x2 identity matrix:

```
>> test_identity=zeros(2,2)

>> test_identity=eye(2)
```

and the results consecutively displayed would be:

```
test_identity =
     0     0
     0     0
test_identity =
     1     0
     0     1
```

The format of matrices displayed in MATLAB's command window is just as one would write them in row and column format. Matrices are populated with values by either importing data (not discussed here) or by entering values in rows and columns, where columns are separated by a space and rows by a semicolon. For example, the following command would create a 2x2 identity matrix:

```
identity_2 = [1 0; 0 1]
```

which would return the following result in the command window:

```
identity_2 =
     1     0
     0     1
```

The workspace window has a list of all vectors or matrices created in the session. All are listed, and for small matrices, individual values are shown. For larger matrices, only dimensions (m x n) are shown. Display of the dimensions is useful to ensure that you do not try to perform operations on matrices with the wrong number of rows and columns. Double clicking on a vector or matrix in the workspace opens a new window with a tabbed spreadsheet-like view of its elements (called the Variable Editor). It is far easier to diagnose problems in this editor window than in scrolling through the results in the command window, which can be overwhelming to read with many rows and columns.

As discussed above, commands can be run from a text file containing a list of commands. Code is written into such files and saved to a filename with an .m extension. To run .m files, you navigate within the current directory path window until your .m file is visible. Then in the command window, you type in the name of the .m file (without the .m extension) and hit ENTER. MATLAB then treats the entire list of commands in the file as a script and runs it sequentially. Depending on your needs, you may or may not need semicolons at the end of lines (but usually you will include semicolons so the command window does not become cluttered as results speed by in the background). Any commands without semicolons will have their results shown in the command window. If semicolons are always included, the results can be viewed via the workspace.

As a demonstration, one possible sequence of commands to complete **Example 8-1** (either entered line by line or run as an entire .m file) is:

```
Z=[150 500; 200 100];
X=[1000 2000; 1000 2000];
A=Z./X;
```

A command sequence for **Example 8-2** is (assuming commands above are already done):

```
y1=[100; 0];
y2= [0; 100];
direct=eye(2)+A;
L=inv(eye(2)-A);
directreq1=direct*y1;
directreq2=direct*y2;
totalreq1=L*y1;
totalreq2=L*y2;
```

where the final 4 commands create the direct and total requirements for Y_1 and Y_2. A command sequence for **Example 8-3** is (assuming commands above are already done):

```
R=[50 5];
R_diag=diag(R);
E_direct_Y1 = R_diag*directreq1;
E_direct_Y2 = R_diag*directreq2;
E_total_Y1=R_diag*totalreq1;
E_total_Y2=R_diag*totalreq2;
E_sum_Y1=sum(E_total_Y1);
E_sum_Y2=sum(E_total_Y2);
```

EIO-LCA in MATLAB

The EIO-LCA model in a MATLAB environment is available as a free download from the website (www.eiolca.net). The full 1997 model in MATLAB is available directly for download, and a version of the 2002 model excluding energy and GHG data is available directly for download. The 2002 MATLAB model with energy and GHG data is available for free for non-commercial use via a clickable license agreement on the www.eiolca.net home page (teachers are encouraged to acquire this license and the MATLAB file for local distribution but to make non-commercial license terms clear to students).

Within the downloaded material for each model are .mat files with the vectors and matrices needed to replicate the results available on the www.eiolca.net website, and MATLAB code to work with producer and purchaser models. MATLAB .m files named EIOLCA97.m and EIOLCA02.m are scripts for the 1997 and 2002 models, respectively, to generate results similar to what is available on the website.

For example, running the EIOLCA97.m file in the 1997 MATLAB model will successively ask whether you want to use the producer or purchaser model, which vector (economic, GHG, etc.) to display, and how many sectors of results (e.g., all 491 or just the top 10). Before running this script, you need to enter a final demand into one or more of the 491 sectors in the SectorNumbers.xls spreadsheet file. Results will be saved into a file called EIOLCAout.xls in the 1997 MATLAB workspace directory. **Note: to run the EIOLCA97.m script file, you must be running MATLAB directly in Windows or via Windows emulation software (e.g., Boot Camp or Parallels on a Mac) since it uses Microsoft Excel read and write routines only available on Windows. The vectors and matrices in the 1997 model though are accessible to MATLAB on any platform. Due to these limitations, and the age of the data in the 1997 model, this section focuses on the 2002 MATLAB model (but similar examples and matrices exist in the 1997 model).**

Running the EIOLCA02.m file in the 2002 MATLAB model files will successively ask whether you want to use the producer (industry by commodity basis default) or purchaser model, the name of the vector variable that contains your final demand (which you will need to set before running the .m file), and what you would like to name the output file. Note the 2002 MATLAB model can be run on any MATLAB platform (not just Windows).

Before running this script, you need to create and enter a final demand into one or more of the 428 sectors. The following MATLAB session shows how to use the EIOLCA02.m script to model $1 million of final demand into the *Oilseed farming* sector. All lines beginning with **>>** show user commands (and as noted above the user also needs to choose between the producer and purchaser models, and give names for the final demand vector and a named txt file for output – highlighted in green). Before running this code, you will need to change the current MATLAB directory to point to where you have unzipped the MATLAB code.

```
>> y=zeros(428,1);
>> y(1,1)=1;
>> EIOLCA02
```

```
Welcome to EIO-LCA

This model can be run in 2002 $million producer or
purchaser prices.

For producer prices, select 1. For retail (purchaser)
prices, select 2.

Producer or Purchaser prices?   1


Name of the 428 x 1 final demand vector     y

Output file name? (include a “.txt”)

Filename      xout.txt

Total production input is:  1$M2002, producer prices
```

The resulting xout.txt file shows the total supply chain results across all sectors for $1 million of final demand in all data vectors available in the MATLAB environment (which would match those on the website), all in one place. This file can be imported as a text file into Microsoft Excel with semicolon delimiters for more readable output and for easier comparison to the results on the website. An excerpt of rows and columns from this file is shown in Figure 8A-12 (sorted by sector number and rounded off):

		Total econ, \$M	Total Energy, TJ	Fossil CO_2 Emissions, mt CO_2e
	Total, All Sectors	2.1	16.1	944
1111A0	Oilseed farming	1.1	8.4	476
1111B0	Grain farming	0.0	0.2	13
111200	Vegetable and melon farming	0.0	0.0	0.06
111335	Tree nut farming	0.0	0.0	0.023
1113A0	Fruit farming	0.0	0.0	0.11
111400	Greenhouse and nursery production	0.0	0.0	0.36
111910	Tobacco farming	0.0	0.0	0.14
111920	Cotton farming	0.0	0.2	11
1119A0	Sugarcane and sugar beet farming	0.0	0.0	0.06
1119B0	All other crop farming	0.0	0.0	0.75

Figure 8A-12: First 10 Sectors of Output from EIOLCA02.m script for \$1M of *Oilseed farming*

The script .m files have much useful information in them, should you care to follow the code. For example, you can see the specific matrix math combinations used to generate the producer and purchaser models and their direct and total requirements matrices used in EIO-LCA (summarized in Figure 8A-13). The Readme file within the EIO-LCA 1997 and 2002 ZIP files give additional detail.

Element Name	Description
A02ic	Direct requirements matrices – this particular one is an industry by commodity format (there are similarly-named matrices for industry by industry, etc.
L02ic	Total requirements matrices (others similarly named), using Leontief inverse.
EIOsecs EIOsecnames	List of IO sector numbers, names of sectors in the model.
envect02	List of effects available to be output, starting with direct and total economic effects, then energy (including fuel-specific), then greenhouse gases (specified by gas), water, conventional air pollutant, hazardous waste, transportation by mode, toxic releases, then TRACI impact assessment.
EIvect	The energy and greenhouse gas emission **R** matrix values, including totals and specific to fuels and gases. There are similar matrices for water flows, etc.
W	A modified make matrix (see section on make and use tables). This matrix is used to ensure that all sectors that actually produce an input are considered for overall national average production. Any final demand input is 'split' into all sectors that produce it. For example, many sectors make electricity, not just the power generation sector.
purchtransmat	Matrix that distributes purchaser-valued final demand into the production, transportation, and wholesale-retail margin inputs.

Figure 8A-13: Summary of Vectors and Matrices

E-resource: A Microsoft Excel file with the components of purchaser-valued output (as used in **purchtransmat**) for the 2002 benchmark input-output table of the US economy is posted in the Chapter 8 folder (filename is Purchaser Producer Cost Map 2002.xls). Figure 8A-14 summarizes three values in the file. For example, it shows to convert a final demand of $1 in purchaser-valued output in *Automobile manufacturing* into a more complex final demand of $0.88 of *Automobile manufacturing*, $0.1 in *Wholesale trade*, etc.

			Transportation Margin Sectors						
Sector Name	**Producer Price**	**Wholesale trade**	**Air**	**Rail**	**Water**	**Truck**	**Pipeline**	**Retail trade**	**Purchaser Price**
Automobile Manufacturing	0.88	0.1	0.002	0.00	0	0.01	0	0.01	1
Light Truck and Utility Vehicle Manufacturing	0.69	0.08	0.00	0.01	0	0.01	0	0.2	1
Heavy Duty Truck Manufacturing	0.83	0.02	0.00	0.01	0	0.01	0	0.14	1

Figure 8A-14: Example of Purchaser Model Margin Coefficients

Instead of using the provided .m script files, the MATLAB workspaces for 1997 and 2002 can be used on any MATLAB platform to do tailored modeling using the various vectors and matrices. For example, you may want to generate the total fossil CO_2 emissions for $1 million of oilseed farming in the same EIO-LCA 2002 model. Fossil CO_2 emissions are in the matrix

EIvect, in row 8 (rows 1-6 are the various energy vector values and rows 7-12 are the various GHG emission vector values). The MATLAB code needed is as follows:

```
>> clear all
>> load EIO02.mat
>> y=zeros(428,1);
>> y(1,1)=1;
>> x=L02ic*y;
>> E=EIvect(8,:)*x
```

which returns:

```
E = 944.2073
```

The same value appears in the first row of the last column of Figure 8A-12.[22]

Likewise, you might be interested in generating the total fossil CO_2 emissions across the supply chain for $1 million into each of the 428 sectors:

```
>> allsects=EIvect(8,:)*L02ic;
```

which returns a 1x428 vector containing the requested 428 values (where the data in column 1 is the same as above for oilseed farming). This simple one-line MATLAB instruction works because the 1x428 row vector chosen from EIvect (total fossil CO_2 emissions factors per $million for each of 428 sectors) is multiplied by the column entries in the total requirements matrix for each of the sectors, and the result is the same as finding the total GHG emissions across the supply chain as if done one at a time. The first four values in this vector (rounded) are:

```
[944.2 1123.1 739.8 756.4],
```

representing the total fossil CO_2 emissions for $1 million of final demand into the first 4 (of 428) sectors in EIO-LCA. Much more is possible given the available economic and environmental/energy flow matrices that is not possible on the website or with the included script file. For example, you could do a similar analysis as above for the purchaser-based model to find the results of $1 million in all sectors.

[22] Technically, the EIO-LCA MATLAB code uses the **W** matrix to find direct effects in these equations. See the .m file.

End of Section Questions

E-resource: A Microsoft Excel file with the aggregated 15-sector 2002 benchmark input-output table of the US economy is posted to the textbook website. It contains the **A** matrix and the sector outputs for the 15 sectors.

1. Use the US 2002 benchmark 15-sector spreadsheet to answer the following questions.

a. Find the direct supply chain purchases ($M) from all sectors for a $100 million purchase in Sector 1 (*Agriculture, forestry, fishing, and hunting*).

b. Find the second level supply chain (first level indirect, or Tier 2) purchases ($M) from all sectors for a $100 million purchase in Sector 1.

c. Find the total supply chain purchases ($M) from all 15 sectors for a $100 million increase in Sector 1.

d. What percent of the total purchases in each sector are direct? What percent of the total purchases in each sector are second-tier?

2. Answer the questions below using the EIO-LCA 2002 MATLAB environment, using 2002 producer and purchaser priced industry by commodity matrix models.

Hint: see the EIOLCA02.m code and observe differences in the producer and purchaser models. No deductions for inelegant solutions, but be sure to show equations or code used. You will be able to verify you have done it right because the eiolca.net website for the 2002 US model will show the same emissions for a given sector.

a. Find the fossil CO_2e emissions resulting from $1 million of *producer-valued* final demand for every sector in the model. Which ten sectors, sorted by total fossil CO_2e emissions, result in the highest total tons of CO_2e as a result of this economic activity?

b. Repeat part (a) for *purchaser-valued* final demand (still submitting only the top 10). Show results broken out by production, transportation and margin sectors.

c. Make a visual comparing the producer and purchaser-valued results for your top 10 fossil CO_2-emitting sectors per purchaser $million (i.e., you should have the top 10 from part (a) matched with the results from part (b) even if they are outside of that top 10). Discuss the differences, and why they are greater or less than each other.

Section 7 – IO-LCA-based Uncertainty Analysis: Example with Ranges[23]

Input-output models are generally used for screening tools, for example to identify hot spots within a product system network for subsequent process-based analysis or to focus primary data collection efforts. Traditionally the screening is done on a magnitude basis, meaning that the point of the screening is to identify the likely biggest hot spots. However by considering available data, uncertainty based screening is also possible.

Chapter 8 (including its Advanced Material sections) described how publicly available data is used to generate the data for the **R** matrices used for environmental flows per unit of output. In short, a single 'best' data source is identified for each of the sectors in the model for each needed flow. The overall IO-LCA model then is comprised of many 'best guess' deterministic data points and no representation of uncertainty. Most IO-LCA models are deterministic in nature.

Chen et al (2016) considered the availability of various data sources on energy use for each of the sectors in the 428 sector 2002 US EIO-LCA model, as well as various combinations of different assumptions (e.g., different assumptions on unit prices of fuel) that could be used to generate alternative values for **R** matrices. They also introduce a method for expressing and visualizing the resulting parameter uncertainty in matrix-based models. The results provide additional insight to understand how a screening tool can be used to identify both hot spots as well as uncertainties that could be better understood with additional effort to improve LCA results. Instead of the commonly used simulation method, this study develops a new parameter uncertainty estimation and visualization method focusing on the uncertainty across multiple data sources for each inventory sector. The inventory flow of energy consumption was chosen due to high data availability, and because it is fundamental to estimate other inventory categories such as greenhouse gas emissions. In addition, uncertainty in energy consumption is one of the most important impact categories to stakeholders, due to financial and global scarcity concerns. Industrial stakeholders would want to understand the fuel use expense for each year and thereby make decisions based on the available information. Therefore, the uncertainty of the result is very important for users to make more robust decisions. The general method used for energy consumption in 2002 EIO-LCA model could be applied to other inventory categories, and other matrix-based LCA models. See Advanced Material Section 3 for more detail on the methods used.

From the data sources and with various assumptions, multiple values were calculated for each sector. For the energy consumption category in the 2002 EIO-LCA model, data from 6 publicly available sources from government agencies were used:

- U.S. Bureau of Economic Analysis (US BEA)

[23] This section is excerpted / modified from the published paper Chen (2018). Reference provided at end of section.

- U.S. Census Bureau (Census)
- U.S. Department of Energy, Energy Information Administration (US EIA)
- U.S. Environmental Protection Agency (US EPA)
- U.S. Department of Agriculture (USDA)
- National Transportation Research Center (NTRC).

Figure 8-6 shows a high level summary of different assumptions and data sources used in estimation of energy use (USDA 2002) (Census 2007) (MECS 2002) (NTRC 2014) (EIA 2004). Each distinct color represents an energy/$ result (i.e., a different **R** matrix value for a sector) calculated from different data sources and/or assumptions.

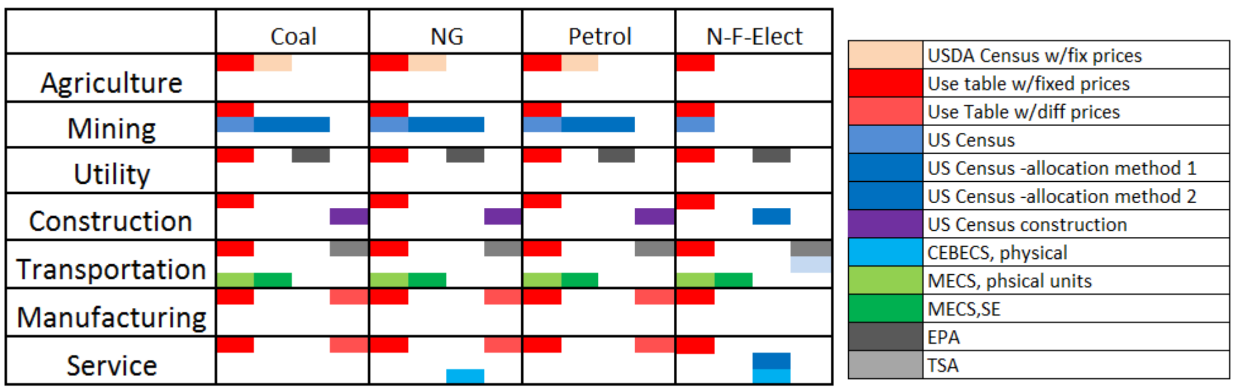


Figure 8-6: summary of values from different assumptions and data sources

EIO-LCA Uncertainty results for single sectors

Sector No. 120, *Petrochemical manufacturing*, is used to demonstrate the uncertainty range results. Figure 8-7 shows the results for total energy consumption as well as separated by fuel. The consumption values are based on $1 million dollar final demand of sector 120. Since IO-LCA models are linear, the relative uncertainty for a different input of final demand would be the same. In this case, the sector's total energy consumption varies from 21 TJ to 62 TJ, resulting in the range of -50% to 40% compared to the default value (42 TJ). The graph shows that the uncertainties of natural gas and petroleum usage are the major contribution to the large discrepancy of total energy consumption (and they are also estimated to be the highest fuel inputs in the total).

The causes of the uncertainties are also visible in Figure 8-7. The red marks are the values calculated from ***R*** matrix with values from the Use Table; all of red marks are larger than the value calculated from the default ***R*** matrix (black asterisk). It indicates that the values in the Use Table are a major contribution to the larger values in the ranges. In addition, the large discrepancy in the total energy consumption are caused by the large range in petroleum usage category; the large range is caused by the Use Table as well. On the contrary, the results calculated from Census data are closer to the maximum value only when maximum values are chosen as substitutes (blue circles in figure), likewise for the minimum values and default values. It suggests that there are only a small proportion of data available from the Census, the majority of the data are substituted, thereby the results are similar to the results calculated from upper, lower bound or default value.

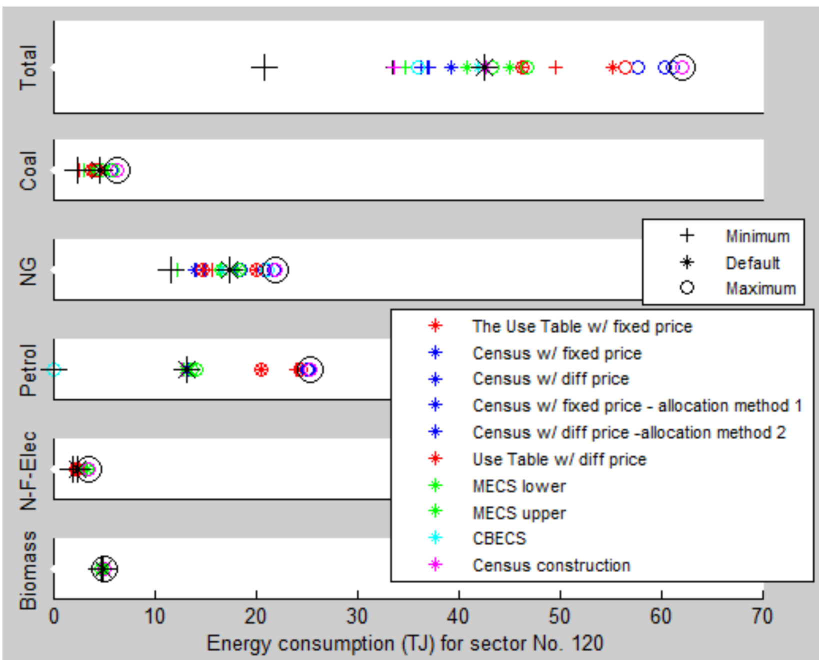


Figure 8-7: bounding results for sector No. 120, based on total energy consumption and separated fuel consumptions. All consumption values are based on 1 million dollar output of sector No.120. NG=natural gas, Petrol=petroleum products, N-F-Elec = non-fossil fuel electricity.

As a hotspot analysis tool, IO models like the EIO-LCA model deterministically list the highest sectors across the supply chain for an industry. Figure 8-8 shows the ranges of the top 15 energy intense sectors for $1 million of sector 120, sorted by the values in the current deterministic **R** matrix used in EIO-LCA. As can be seen, sector 120 (*Petrochemical manufacturing*), 115 (*Petroleum refineries*) and 126 (*Other basic organic chemical manufacturing*) are the top 3 energy intense sectors, contributing an estimated 64% to the total energy consumption. These three sectors are also major contributors to the uncertainty of total energy consumption; however, the rankings of top 5 energy intense sectors can be changed when uncertainties are considered. The red symbols for the top 3 sectors show that the Use Table is the data source for the high end of the ranges. As these sectors belong to Petroleum Product Manufacturing category, it indicates that the Use Table provides greater energy consumption values for some petroleum products. Similar conclusions can be made for values from CBECS: CBECS provides relatively smaller values for petroleum products.

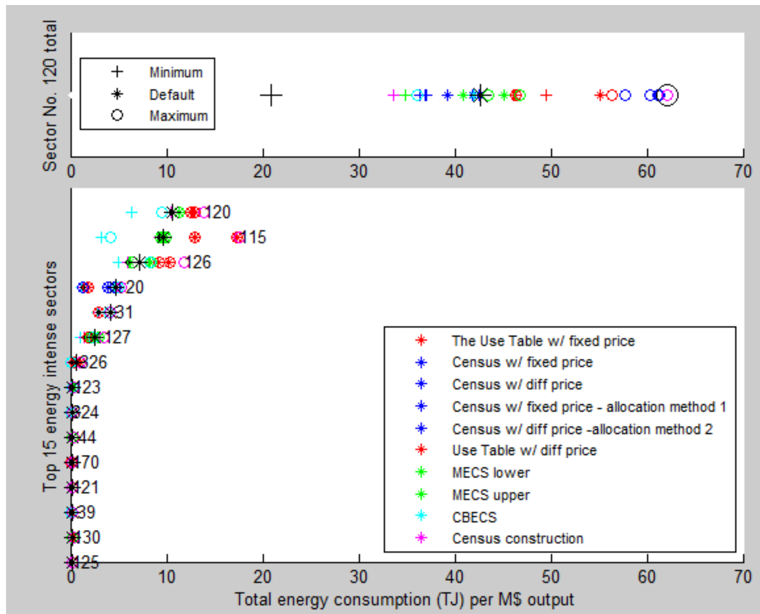


Figure 8-8: bounding results for sector No. 120, based on total energy consumption. Numerical labels represent the 428 sectors. The lower part of the graph shows the top 15 energy intense sectors in terms of $1 million dollars output of sector No. 120.

While Figure 8-8 suggests sectors with high default values also have high uncertainty for Sector 120, that is not a general rule – uncertainty varies across many dimensions. Figure 8-9 shows the uncertainty ranges of all 428 sectors (x axis) in all 12 **R** matrices, based on each sector's percentage change compared to its default value. Note that the positive side of the x-axis is transformed to log scale for legibility. Results suggest that in general, the sectoral uncertainty for direct and indirect energy consumption in the **R** matrix is approximately ±50% overall. In some extreme cases, the values reach over 40 times larger than the default value. For example, the value for *Petroleum lubricating oil and grease manufacturing* (No. 118) varies from 3 to 110 TJ/M$, and *Hospitals* (No. 338) varies from 1 to 33 TJ/M$. This outcome demonstrates how different data sources and assumptions can result in a significantly large range for some sectors, which is not currently expressed in current models.

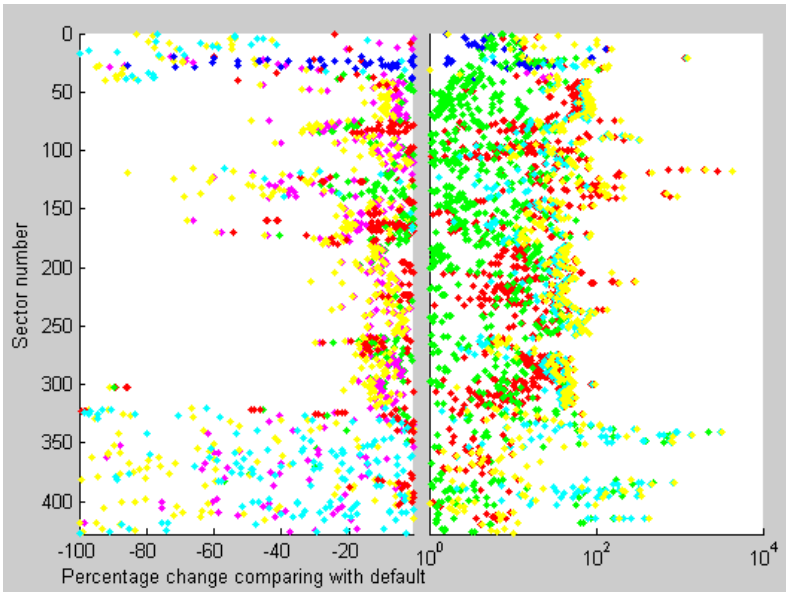


Figure 8-9: bounding results of R matrices for all 428 sectors, based on percentage changes comparing with default value, base 10 log scale is used for positive values.

This uncertainty range analysis suggests that the uncertainty in the **R** matrix can be large. Given the large uncertainties, using only a single data source for a sector in the **R** matrix (as typically done in IO-LCA models) ignores important variation in the system.

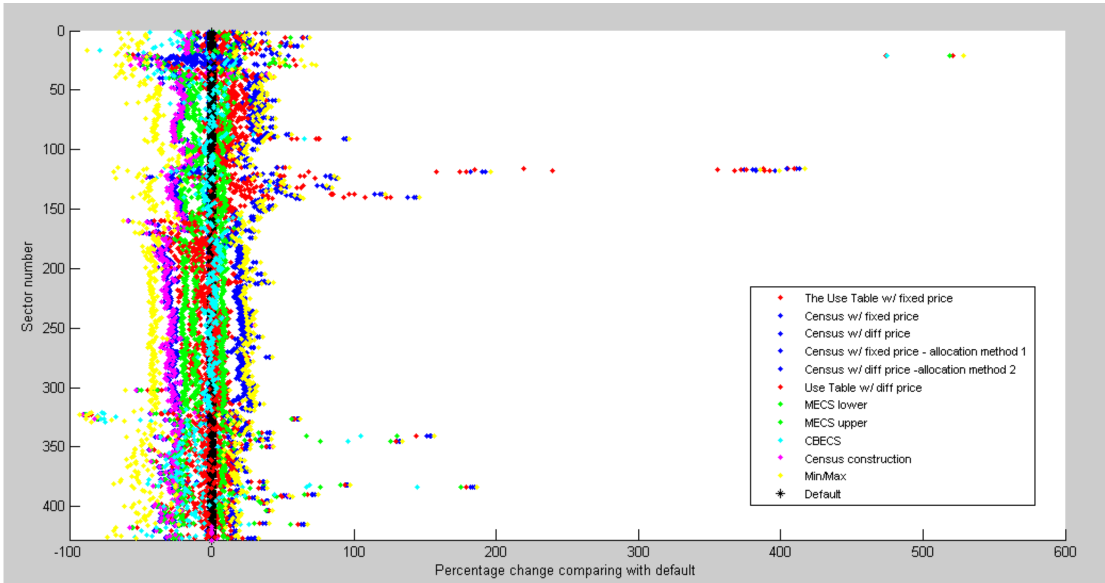


Figure 8-10: bounding results of the B vectors for all 428 sectors, based on percentage changes comparing with default value

Uncertainties of the B matrix

Figure 8-10 shows the uncertainty ranges of the **B** matrices, based on the energy consumption (TJ) per million dollar output (from the **R** matrix range above) for each of the 428 industry sectors. The results were calculated based on the total economic output in 2002 for each sector. The overall uncertainty in total energy consumption for a sector varies about -40% to 40%; however, for the most extreme cases, the result is approximately 5 times larger than the default value. The percentage changes are smaller compared to changes in the **R** matrix, especially for the extreme cases, given the interconnectedness of supply chains and confirms the cancellation effect discussed in (Lenzen 2001): in IO-LCA models, the **R** matrix corresponding relative errors cancel out in the final results. It also shows the combination effect of uncertainty results in matrix-based models- the percentage changes in the **B** matrix are less scattered than the changes in the ***R*** matrix.

These different patterns of uncertainties are caused by the discrepancies of values from different data sources. Investigating the uncertainties of different data sources can give a clearer concept of where the uncertainties occur. Using this information, potential sectors that have large uncertainties could be found and identified or adjusted after further investigation in order to reduce the uncertainty in the model.

Uncertainty values re-evaluation

The existence of big discrepancies for some of the sectors is one of the novel results found from using the uncertainty range method. Tracking the reasons for the big discrepancies can provide decision makers with a clearer concept of how or when to use the various data given the uncertainties. An example is the values shown in Figure 8-8: larger values were calculated from the Use Table (red) while smaller values were calculated form CBECS (light blue). A possible explanation is the limitation of the data source. Values provided by the Use Table were based on the purchase of fuels, rather than energy consumed in production phases. Thus, these purchases could include fuel used as both energy and feedstocks. Petroleum product manufacturing industries have a considerable amount of inputs from petroleum refinery as the feedstocks; however, without additional detail all the purchase values were assumed to be for energy consumed in production. Thus, the results from the Use Table may be larger due to misinterpretation; the consequential uncertainty can be reduced when better-documented data are provided (we note that MECS separates fuel and feedstock usage but is more aggregated). A similar conclusion can be made for the results calculated from CBECS: smaller values were calculated because only energy consumptions related to buildings such as natural gas for heat and electricity were counted.

A more extreme case is the uncertainty in the Coal Mining sector. The coal mining sector uses data from Census and the Use Table, and more than 5 data points were gathered for each fuel type. In the lower and upper bound evaluation, the total energy consumption for the Coal mining sector in the $\boldsymbol{R}$ matrix varies from 1 to 58 TJ/M$, with the default value 4 TJ/M$. The total coal consumption for Coal mining sector from Census, while the value for the upper bound was estimated via the Use Table. Comparing the evaluation from the Use Table with other data sources, a possible reason for the discrepancy could be that coal purchases between coal companies for beneficiation services were counted as an energy source in the Use Table. Creators of IO-LCA databases may use such data sources without adjustment or accommodation, and thus users of the data may not be aware of the reasons that lead to such large deviations. In this case, if the result is used with caution, the decision impact of uncertainty can be reduced.

The two examples show that even reliable data sources can lead to large outliers, giving significantly different results that lead to large uncertainty. As discussed, the large discrepancy between reliable data sources are typically caused by different assumptions used in reporting the data. Using other empirical estimation methods such as the pedigree matrix approach, the uncertainty of Coal mining sector will be determined as less than 40% no matter using which data sources, as they are all from reliable first-hand data sources. The unreasonably large energy consumption value due to the misinterpretation of data would be ignored.

Decision making and screening under uncertainty

The uncertainty ranges estimated for the EIO-LCA model can improve the utility of this source of information to assist with LCA decision making. Adding uncertainty allows for screening tools that 'screen' based on both the magnitude of effect as well as uncertainty, which allows a two-dimensional hotspot analysis. The first dimension relates to which are the likely most important effects, and the second is how relatively uncertain they are. Either or both could be indicators of where additional data (e.g., primary data or process-specific data) could be useful in subsequent effort. Three case studies based on previously published LCA work are shown to demonstrate how the information in the uncertainty range approach can be used. Case study 1 shows how the uncertainty result can help robust decision-making in LCA studies. Case study 2 shows that decisions can be totally overturned when the uncertainty result is considered. Case study 3 shows how rankings provided by the hotspot analysis in LCA tools are changed considering the uncertainty and how decisions can be affected.

Case Study 1 – Energy consumption of plastic cup and paper cup

In the early days of motivating LCA for use in personal decisions, such as the paper vs. plastic debate (Hocking 1991), Lave and colleagues compared the toxic releases and energy use (electricity only) of plastic and paper cups using a 1987 EIO-LCA model (Lave 1995). Using price assumptions, they estimated a plastic cup consumed almost 50% less electricity (4,400 kWh vs. 8,600 kWh) than a paper cup. Using our uncertainty range results, total energy consumption, not just electricity, for a plastic cup and a paper cup varies from 0.3 TJ to 0.7 TJ, and from 0.3 TJ to 0.4 TJ, respectively. When the uncertainty of the comparison is considered, the original conclusions change. Instead of plastic cup production clearly using 50% less energy, the overlap range for possible energy use values suggest high potential for about the same energy use, or for lower energy use for the paper cup. While the lack of a simple conclusion on which of the two materials is better may seem like a step backward, it more appropriately represents the challenge of the decision. The uncertainty may be reduced by finding primary or process specific data.

Case Study 2 – Energy consumption comparison between asphalt and concrete pavement

In an LCA study, decisions could change significantly by considering the uncertainty from the LCA tools. To better illustrate the idea, a LCA study of energy consumption comparison of asphalt and concrete pavement is used and re-evaluated based on the new uncertainty result. Hendrickson and colleagues determined the energy consumption for one km-long asphalt and concrete pavement using the 1992 EIO-LCA model (Hendrickson, Lave, and Matthews 2010). The values were updated for the 2002 EIO-LCA model. From the 2002 EIO-LCA model's deterministic default value, the direct and indirect energy consumption for one km-long asphalt and cement concrete pavement were 4 TJ and 5 TJ, respectively, favoring the choice of asphalt on an energy basis.

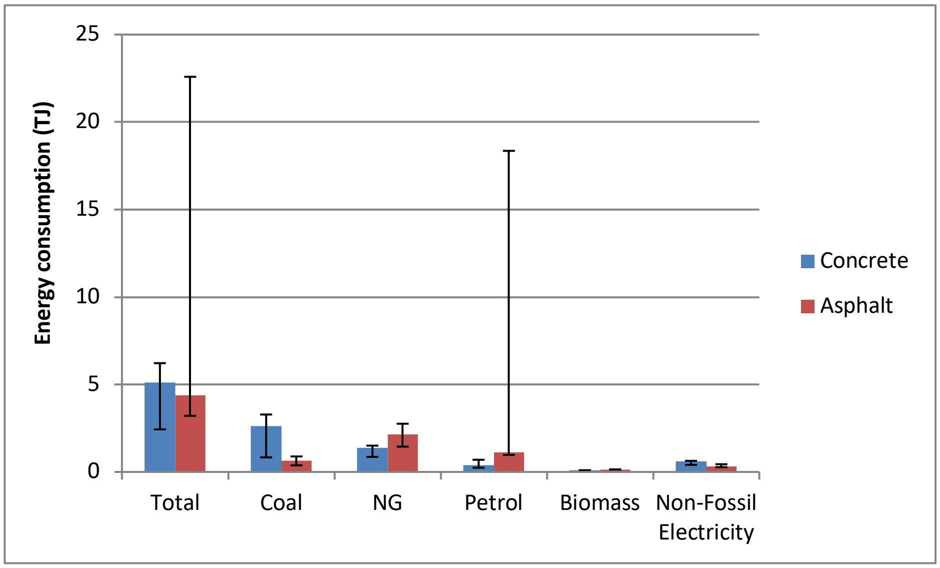


Figure 8-11: Results of direct and indirect energy consumption for 1 km concrete and asphalt pavement, error bars indicates the upper and lower bound values

Considering the uncertainty, as shown in **Figure 8-11** with error bars, the energy consumptions of asphalt and concrete pavement have large ranges. The default values show that one functional unit concrete pavement on average consumes 1 TJ more energy than asphalt pavement, mostly because of the larger coal consumption in producing concrete. However if the uncertainty results are considered, asphalt pavement could be nearly 5 times more energy intense, overturning the simple deterministic and conclusion. The large range in total energy use for asphalt pavement is associated with the uncertainty of petroleum usage. As mentioned previously, the Use Table provides larger petroleum purchase values for petroleum product manufacturing industries due to the feedstock use. The sector used for evaluating asphalt pavement is 'Asphalt paving mixture and block manufacturing', a sector associated with petroleum products, the values from the Use Table provided have the same issue of including feedstock as fuel consumption. The uncertainty can be reduced if better documented data are provided; however, with available information, no firm conclusion can be made regarding the superiority of two pavements.

Case Study 3 – Energy saving from structural steel reuse

Input-output based LCA models (and matrix based LCA models in general) are often used as a hotspot screening tool to help decision makers access quick and simple relative results. The importance of sectors regarding environmental impacts rankings helps determine what and where to look for potential impacts. Yeung and colleagues completed a hotspot analysis of steel reuse (Yeung, Walbridge, and Haas 2015), applying rankings from EIO-LCA to help in evaluating the impacts of reused steel. Due to the purpose of the study, the rankings of energy use for steel and cement manufacturing was crucial to the analysis.

If only deterministic results are considered from EIO-LCA, *iron and steel mills* and *cement manufacturing* sectors were ranked as 3rd and 7th regarding energy consumption. If the uncertainty is considered, the sectors' rankings change, shown as Default column in Figure 8-12. Using the upper bound of the uncertainty range, the iron and steel mills sector changes slightly, but cement manufacturing sector is no longer in the top 10. If the hotspot analysis is used to determine what to include in a detail analysis, then at the default value missing the impacts of the paper mills and lighting fixture manufacturing sector at the upper bound. Although these sectors may turn out to have limited impact on the final results, the analyst considering all possibilities has a better picture and can investigate their potential impacts.

Rank	Higher	Default	Upper
1	Nonresidential manufacturing structures	Nonresidential manufacturing structures	Nonresidential manufacturing structures
2	Power generation and supply	Power generation and supply	Power generation and supply
3	Petroleum refineries	Iron and steel mills	Petroleum refineries
4	Iron and steel mills	Petroleum refineries	Iron and steel mills
5	Oil and gas extraction	Oil and gas extraction	Oil and gas extraction
6	Cement manufacturing	Truck transportation	Clay and non-clay refractory manufacturing
7	Truck transportation	Cement manufacturing	Paperboard Mills
8	Other basic organic chemical manufacturing	Clay and non-clay refractory manufacturing	Other basic organic chemical manufacturing
9	Clay and non-clay refractory manufacturing	Other basic organic chemical manufacturing	Paper mills
10	Petroleum lubricating oil and grease manufacturing	Paperboard Mills	Lighting fixture manufacturing

Figure 8-12: Top 5 Energy Consumption Sectors for $1 Million dollar Output of Nonresidential manufacturing structures sector

The method used in this study can be applied to other matrix-based models for future work; the results will be applied to uncertainty-based screening in matrix-based models. In addition, the current method can be improved by considering all the data points and their impacts to the model rather than only the maximum and minimum $\boldsymbol{R}$ matrices. The data points could be defined with discrete distributions and simulation could be conducted, results with distributions based on real life data could be provided.

Reference for this Section

Xiaoju Chen, W. Michael Griffin, H. Scott Matthews, "Representing and visualizing data uncertainty in input-output life cycle assessment models", Resources, Conservation, and Recycling, Volume 137, October 2018, pp. 316-325. **DOI**: https://doi.org/10.1016/j.resconrec.2018.06.011