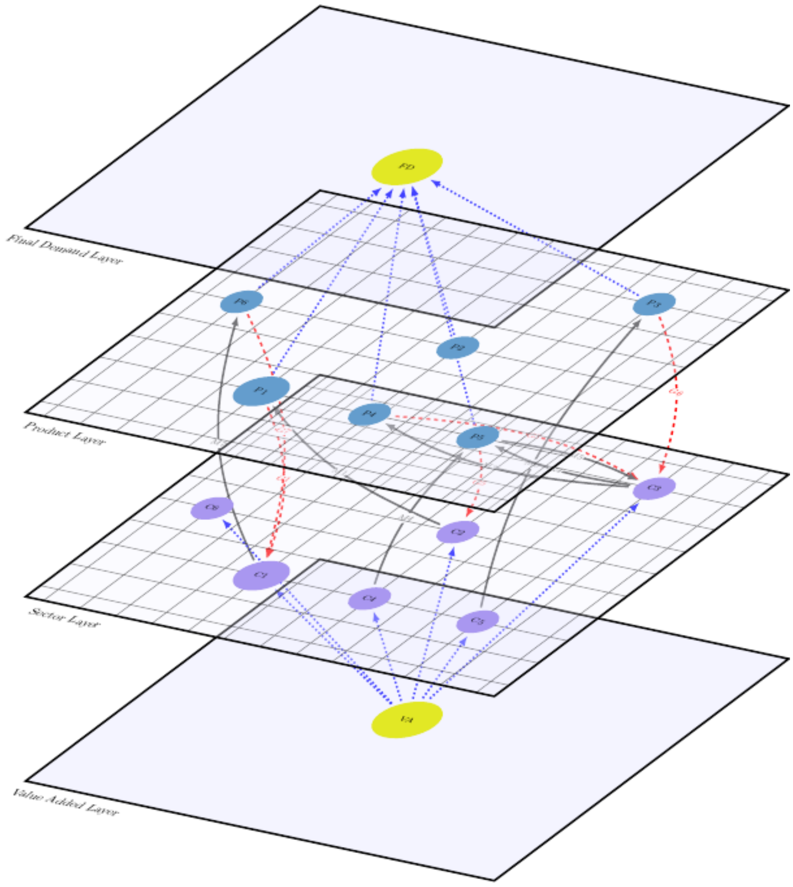
Open Risk White Paper

Follow the Money:
Random Walks on Supply and Use Graphs

Authors: Philippos Papadopoulos

December 16, 2024



www.openriskmanagement.com

The open future of risk management

SUMMARY

We explore how to organize Environmentally-Extended Input-Output frameworks (EEIO), and in particular their Supply and Use Table (SUT) formulation, as graphs. Working directly with SUT systems instead of converting to symmetric IO matrices involves fewer assumptions and (in principle) higher resolution in expressing environmental impacts. We elaborate first on the representation of SUT tables as directed, weighted bipartite graphs. We discuss both closed (circular) and open system configurations, featuring source and sink nodes. These are modeled as regular and absorbing Markov Chains respectively. We outline a probabilistic random walk framework that realizes mathematically the colloquial *Follow the Money* concept. This enables computing a range of various existing and new metrics using the EEIO data. As an illustration, besides the standard environmental footprint metric, we introduce the concept of *footprint variance* or the intrinsic variability of estimates. We illustrate the overall setup using a classic numerical example from the EEIO literature.

Further Resources

- The Open Risk Manual is an open online repository of information for risk management developed and maintained by Open Risk.
- The Open Risk Academy offers a range of online courses around risk and sustainable portfolio management, which utilize the latest in interactive eLearning tools.
- The Open Source Repository contains a variety of open source tools and resources.
- More content: Open Risk White Papers and Open Risk Blog

About Open Risk

Open Risk is an independent provider of training and risk analysis tools to the broader financial services community. Our mission is captured by the motto: The open future of risk management. Learn more about our mission and projects at: www.openriskmanagement.com

1 Introduction

1.1 Background

The need to accelerate efforts to understand and reduce environmental impact of economic activity is increasingly recognized by policy makers, private enterprise and society at large. Many new policies and management tools aim to contain the increase of Greenhouse Gas concentrations in the biosphere, which are unequivocally caused by human activities [1]. Since 2011 and the measurements reported in AR5 [2], concentrations have continued to increase, prompting ever more urgency on the face of significant technical and cultural challenges for adaptation.

Environmental impacts, how they can be measured (accounted for) and, importantly, how they can be *attributed* to human actors is a central aspect of mitigation actions pursued at various levels. Sufficiently accurate, complete and trusted measurement and reporting of environmental impact is a prerequisite for the sustainability transition. Equally important is a transparent, logical, fair and effective attribution that will incentivize the economic actors optimally placed to support the transition.

In government (official) statistics the concept of physical (environmental) accounts is already well established[3],[4]. The UN System of Environmental-Economic Accounting offers an elaborate overarching accounting framework towards describing and analyzing the environment and its interactions with the economy. The physical flows of materials from nature to the economy are described along with any transformation processes and material flows back to nature (consumption, waste, pollution etc). The challenges for such a program (some aspects of it termed as *Industrial Ecology*) are significant. The multitude of products and production processes means that major aggregations are generally required. Currently statistics on physical flows of products are collected at a detail of circa 10000 products in selected areas such as foreign trade and outputs from manufacturing industries[5].

Further, when addressing the environmental impact from complex interlocked production and consumption activities and supply chains, crucial questions arise such as: who is responsible and how is that responsibility to be attributed more precisely? Formal, transparent, comprehensive, mandatory etc. mechanisms of establishing and communicating *responsibility* are part of the toolkit towards the broader sustainability transition, but there are both conceptual questions and more practical challenges. Deep questions such as to whether responsibility should be linked to compensatory or distributive justice ([6]), or the role of individuals, who do not act in isolation but can own, control, shape, and profit from the production processes that create impacts, even while this ownership is not distributed equally[7].

There are several outstanding proposals for environmental impact attribution systems. They are generally framed in terms of linking impacts to specific economic actors (consumers, producers, prime extractors, or other enabling intermediaries and beneficiaries). Most developed in this direction are conceptual frameworks for accounting and attributing Greenhouse Gas (GHG) emissions, in particular the GHG Protocol concepts and proposals [8]. Besides directly involved economic actors, impact accounting and attribution is also important for intermediaries acting as enablers or facilitators. Most prominently, this concerns banking institutions that provide debt and equity finance. A significant current effort in the direction of accounting and attributing emissions in financial portfolios is the Partnership for Carbon Accounting Financials (PCAF) initiative and methodology [9]. An important characteristic of financial portfolios is that they span a diverse fraction of an economy. Hence they require an attribution that is granular, coherent (e.g., does not involve double counting) and comparable across a wide range of financed activities.

The ability to fine tune financial portfolio management methodologies to identify hot-spots, correlations and dependencies, to project and evaluate scenarios into the future, are all elements of an emerging paradigm of *Sustainable Portfolio Management*. In [10] we developed a conceptual framework that synthesized several current approaches to sustainable portfolio management. We discussed the different required information layers that encompass the accounting, attribution and forward-looking limit frameworks that implement *carbon budget constraints*. In this white paper the context is more technical and specialized: we focus on the consistent attribution among disparate portfolio elements, which is one of the key required tools.

The most straightforward attribution methodology attributes environmental impact to the producer (the polluter pays principle). A major alternative concept is consumption-based attribution (commonly referred to as footprinting). It was developed to address a shortcoming of production-based attribution, namely the lack of attribution to international economic exchanges. Trade is a major part of modern economies and ignoring its impact creates significant blind spots and counter-incentives.

Responsibility according to the consumption-based principle rests with the buyers, or the so-called final demand sector (e.g. households, governments and investors as consumers), because the impacts generated by suppliers are induced by consumer demand and are enabled by consumers' purchases. Income-based carbon accounting is yet another strategy to distribute impacts between economic actors[11]. Further proposals and reviews of different types of responsibility attribution are provided in[12],[13],[14, 15].

A strand of academic research and practical economic analysis that is underpinning the above accounting and attribution methodologies concerns data collection and organization in the form of economic *Input-Output* (IO) frameworks. Input-output analysis is a top-down (macro) technique used by economists and statisticians to monitor and account for the interdependency of modern economic systems and guide policy decisions.

The field has a long history since the proposals of economist Wassily Leontief who developed a system of economic analysis in the 1930s and 1940s. The conceptual origins of IO models go even further back: The *Tableau Economique* in early economic theory proposed by François Quesnay in 1758 is being considered the precursor to modern economic Input-Output Models. For a complete overview and references of the IO methodology see [16].

The holistic aspect of IO frameworks is particularly relevant for a consistent and generally applicable view, e.g., in sustainable portfolio management context. In modern IO databases the basic data are organized as an *integrated input-output* framework[5]. This construction involves a large number of steps, with assumptions and adjustments along the way. The complexity grows with the size of the framework: the number of countries covered in detail, the number of sectors and products. The accuracy and fidelity of any attribution is limited by disparities in the collection and standardization of raw data in the different regions and sectors. The impact of data quality on results has been studied extensively[17]. There are further methodological challenges. Comprehensive lists of issues are given in [18] and [19].

An important feature of the vast Input-Output literature from the earliest days is that the distinction between sectors and products is under-emphasized, essentially for computational reasons. Namely the organization and processing of the underlying data is such that the number of recognized products ends up being the same as the number of recognized sectors. This arrangement simplifies the overall system and enables important analytical tools. It comes, though, at the expense of fidelity in representation and it necessitates further methodology assumptions that are not a-priori required. Another important methodological aspect which materially affects attribution outcomes is the choice of structure of an open IO system, namely which sectors are treated as exogenous. For example, after endogenizing labor in global supply chains in[20], they show that the weight of high impact industrial sectors versus service sectors shifts significantly. In other words the manner in which one decides to *cut* the circular graph of the human economy produces different views of the same underlying reality.

Graph Theory is a long-standing field of both mathematics and computer science. There is also a growing related body of work in economic and financial disciplines that makes use of generalized graph structures for *Network Analysis*. Overviews are given in ([21],[22]).

From early on graph theoretic techniques have also been considered in the theory and practice of Input-Output models. At the simplest level the linkage of Input-Output Economic models to Graph Theory goes by the name of Qualitative input-output analysis (QIOA) [16]. More recently the role of directed bipartite graphs to organize the information of symmetric Input-Output frameworks has been recognized in [23].

1.2 In this White Paper

The emphasis on Supply and Use (SUT) approaches as the backbone of EEIO calculations motivates developing graph-theoretic tools further. In particular we are interested in attribution methodologies with minimal methodological assumptions and with potentially significantly higher granularity in products versus sectors.

Towards that end we explore how Environmentally-Extended Input-Output data (EEIO) in SUT format relate to the concepts and algorithms of graph theory. We consider in detail the graph interpretation of the standard *supply-use transaction block* as discussed, e.g., in[24]. To the degree possible we view the input-output framework not as a causal economic model, but rather an empirical descriptive tool that registers dependencies between sectors and production related environmental impacts.

On this basis we elaborate on a representation of the SUT tables as a type of *directed, weighted bipartite graph*. IO literature focuses on the so-called open form of the equations and is, in particular, geared towards flexible analysis of changing *final demand*. Our approach is more general: We consider closed and open systems in parallel, as different views of the same system. The most applicable choice for EEIO attribution need not coincide with other economic analysis requirements.

Beyond graphs as static representations of dependencies, stochastic, *Random Walks* on graphs are well established mathematical models with a rich associated toolkit. The discipline has received significant attention in the context of social networks but is applicable far more broadly. The probabilistic interpretation of IO frameworks has attracted some attention both as a means of interpretation and in terms of widening the toolkit of calculations [25, 26, 27, 28].

Here we will formulate the problem of a random walk on a SUT (bipartite) graph for both open and closed IO configurations. With the probabilistic interpretation at hand it is possible to give precise meaning to the phrase **Follow the Money**. Tracing the flow of funds between economic actors and the environmental impact they generate along the way is expressed as an expectation (an average over possibilities). Both direct and indirect impacts can be computed using such expressions.

Furthermore the probabilistic interpretation enables further analytic measures of interest, all derived from the well developed language of conditional probabilities of a multivariate distribution. An immediate example is the computing the intrinsic *variability* of impacts around their expected values. Besides various analytic formulas, Monte Carlo simulations can help derive more elaborate measures.

There are, in summary, several stacked and linked conceptual domains that we must gingerly walk through in the present exercise.

- The basis of all further elaborations and models is formed from a vast array of empirical data (statistical surveys etc.) that are aggregated into the set of Supply and Use tables (SUT) produced and publicized by official statistics agencies. We take these structures as given and do not delve here further in the significant complexities involved. Further, we will only focus on the single region version for simplicity of exposition, but the extension is straightforward..
- From the overall EEIO database, certain numerical elements of the SUT tables are recognized as *mathematical matrices and vectors*. This is the starting point for making connections with graph theory. The SUT matrices can in turn be interpreted as representing SUT graphs, exploiting the well-known matrix / graph correspondence. We characterize the specific category of mathematical graphs (bipartite, directed, etc.) that is applicable in this case.
- As a separate exercise, SUT matrices can also generate transition matrices associated with *Markov Chains*. This a modeling step (an assumption) that is inline with the prevailing interpretation of IO systems. Finally, Markov Chains can being interpreted as Random Walks on the SUT graph. This opens the way for the calculation of various expectations and other metrics.

2 Supply and Use Data as Graphs

2.1 Definitions and Notation

We review first some basic notions from graph theory[1], to fix notation but also link them to the economic concepts underpinning Input-Output frameworks, in particular when those are structured in the form of *Supply and Use Tables* (SUT).

2.1.1 Basic Graphs

A standard mathematical graph is denoted $G = (V, E)$ where V is a set of *Nodes* (also called Vertices). There is a *finite* set of such nodes, v_k, indexed by $k \in [1, \ldots, N]$. The *order* of a graph is the total number $|V|$ of nodes, equal to N. The nodes of a SUT graph will be standing for either economic **Sectors**, which group together homogeneous organizational units comprising the economy, or **Products**, grouping goods or services produced or consumed by organizational units[2]. Maybe useful to note that the type of a node can only be inferred from the connectivity properties of the graph (rather than being an explicit label or attribute).

2.1.2 About Sector Nodes

Following loosely the Eurostat definition [4], a *node* is an aggregation of economic and legal entities characterized by decision-making autonomy in the exercise of their principal economic function. A *Sector Node* ($v_k \in \mathbf{S}$) represents the aggregation of all economic *actors* that exhibit similar patterns of Production

[1] We will use the terms Graph and Network interchangeably.

[2] We will use the terms Industry or Sector interchangeably. Similarly the terms Commodity, Product, Good or Service or Activity.

and Consumption of resources but also similar Environmental Impact through such activities.[3]. A Sector sub-index, i, runs from 1 to n. The number of sector nodes defines the *Sector Resolution* of the graph, namely its ability to resolve the differentiated classes of economic agents making up the economy. The sectoral resolution in turn determines the environmental *Impact Attribution Resolution* that is achievable - the granularity with which one can attribute responsibility. When the desired granularity is lacking, attribution must necessarily make assumptions that can be challenging to validate.

2.1.3 About Product Nodes

Products are the second type of economic concept that needs to be represented in a SUT graph as a node. We will call them **P** nodes. These nodes represent stylized *Product Markets*, where producers and consumers may sell and purchase goods and services respectively. The product market assumption implies that any entity (e.g. company) within a Sector can either purchase or sell Products to any other entity, via market intermediation. Product nodes can represent markets for various manufactured goods, services, energy but also labor.

It is a defining feature of the SUT graph (as opposed to the more commonly encountered Input-Output graphs) that Sector nodes ($v_k \in \mathbf{S}$) are transacting with each other *only* through the intermediation of Product ($v_k \in \mathbf{P}$) nodes. The Product sub-index, p runs from 1 to m Products. The number m of distinct Products, defines the *Product Resolution* of the graph. The Product resolution in turn determines the *Impact Calculation Resolution* that is achievable. The Sector and Product resolution parameters (n, m) need not be the same. Indeed a primary motivation for pursuing SUT graph tools is to enable flexibility in this respect. The order of the graph satisfies $N = m + n$. By convention the overall graph index k will run first over Product nodes, then over Sector nodes.

2.1.4 About Activity Edges

Next to graph nodes representing economic actors and intermediaries, there is a set of *Edges* $E \subseteq V \times V$. Edges denote economic linkages between pairs of nodes. Examining the elements of the edge set $E \subseteq (k, l) : k, l \in V$ indicates whether there is any type of economic interaction between two nodes k and l. The edges of the graph might be representing, for example, economic *transactions* and associated monetary and product flows between nodes. Edges are indexed via the pair of nodes (k, l) but can also be also be indexed as e_q, i.e., using a unique key $q \in [1, \ldots, K]$. The *Size* of a graph is the total number of its edges $|E|$ and is equal to K.

2.1.5 Bipartite Graphs

In SUT graphs the nodes of the graph are naturally split in two sets that are *disjoint and independent*. Graphs that have two disjoint sets V_1 and V_2 of nodes, with edges that connect a node in one set only to nodes in the other set are called *Bipartite Graphs*. The two subsets span the entire set of nodes V:

$$V = V_1 \cup V_2 \tag{1}$$

[3]In IO language sectors sometimes refer to industrial (production) sectors. Here we use the term more generally, as in, e.g., the Household Sector

The node sets V_1 and V_2 are also called the *parts* of the graph. In our case we label the two components as $\mathbf{P}$ and $\mathbf{S}$, thus:

$$V = \mathbf{P} \cup \mathbf{S} \tag{2}$$

The two disjoint sets $\mathbf{S}$ and $\mathbf{P}$ define a so-called *coloring* of the graph with two colors: e.g., if all nodes in $\mathbf{S}$ are cyan, and all nodes in $\mathbf{P}$ are green, then every edge in the edge set E has endpoints of different colors.

2.1.6 Directed Graphs

A *directed graph* (or digraph) is a graph where nodes are connected by *directed* edges. Edges are assigned a *direction*, which means they are deemed to be *emanating* from one node and *ending* on another. Edges are sometimes also called arcs and are represented with arrows. Directed graphs that do not have *self-loops* are called *simple* directed graphs. The no self-loop requirement indicates that there are no arrows that connect nodes to themselves. Mathematically $E \subseteq (k, l) : k, l \in V, k \neq l$.

Input-Output systems are naturally represented as directed graphs: there is an observable directionality of the movement of resources, products, funds etc. from one actor to another. A link from a Sector to a Product does not preclude a link from the Product to the same Sector. The SUT graphs we will work with are thus *not oriented*.

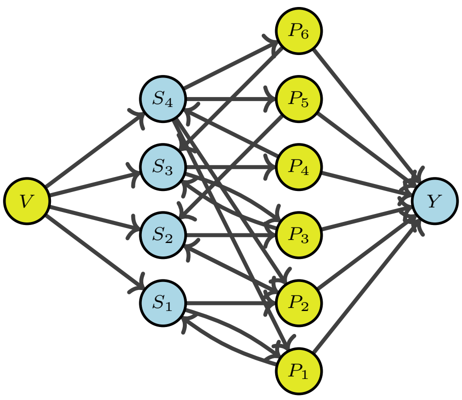


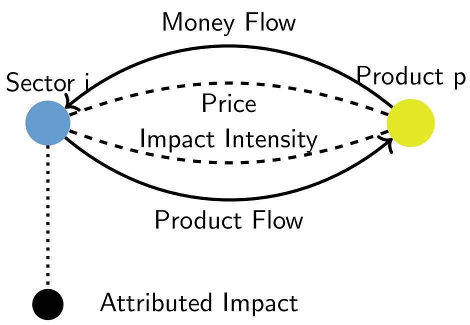
Illustration of a bipartite directed graph. Sector nodes are in cyan, Product nodes in lemon green. As we will see this is an example of an open configuration featuring input and output nodes (V and Y).

2.1.7 Upstream and Downstream Directions

The directionality of the graph implies there are two distinct ways in which one can *navigate* it. One direction is *downstream*, following the sense of the edge arrows. The other direction is *upstream*, or going against the arrow directions. The actual economic meaning of upstream and downstream (i.e., what exactly "flows" up or down) derives from the nature of the *weight functions* associated with the SUT as we will see next.

2.1.8 Weighted Graphs

To nudge the SUT graph structure towards a useful tool, we must introduce at least one mechanism to associate graph nodes (or edges) with *real numbers*. This is done by defining functions over the graph G, for example a function $f : V \to \mathbb{R}$ on graph nodes or a function $h : E \to \mathbb{R}$ on graph edges. A *weighted directed graph* is represented as a triple $G = (V, E, W)$ where V is the node set, $E \subseteq (k, l) : k, l \in V$ is the edge set and $W : E \to (0, \infty)$ is a *positive weight function* that assigns a numerical value to every edge of the graph. Weighted and directed graphs are also known as *directed networks* and edge-weighted graphs. The numerical function W mapping to the edge set E will be an *empirically measured* economic quantity. Three classic weight functions in IO frameworks are: 1) product *volume flows* (in so-called physical units), 2) *prices* of products and 3) monetary flows (the product of volume and price). Here we will focus on monetary flows. For environmental impact calculations we have to further augment the graph to $G = (V, E, W, F)$, where $F : E \to (0, \infty)$ are *impact functions* that assign suitable numerical value to every edge of the graph. In summary, the underlying graph of a stylized Supply and Use Table (SUT) system will be a *weighted, directed, non-oriented bipartite graph without self-loops, with one or more additional defined edge functions.*



The representation of economic exchange is expressed in a SUT Graph building block using two nodes, a Sector node (i) and a Product node (p) and a set of weighted edges. The core weight functions are money flows (in currency units) and product flow volumes (in physical units). Environmental Impact intensities and Prices are further characterizations of the exchange that are expressed as graph functions. The objective of attribution schemes is to use these inputs (across the entire graph) to produce well-defined impact metrics.

More complicated network structures may be required to capture all the relevant economic system details (capital formation, financial linkages etc.), leading naturally to the concept of *property graphs*[29] and graph databases. While computationally very rich network structures can be analyzed with modern computers, the trade-off when introducing more complex structures is that the corpus of mathematical graph theory results might no longer be applicable.

2.1.9 The Linear Algebra Correspondence

The duality between graphs and matrix representations has been a part of graph theory since early on and *Matrix Algebra* has been recognized as a useful tool in graph theory.

Next, we discuss *matrix representations* of the above defined SUT graph class. As a starting point, in the case of non-weighted graphs, an *Adjacency Matrix* A is a matrix of zeros and ones that encodes the connectivity (presence of edges) between nodes:

$$A_{kl} = \begin{cases} 1, & \text{if } v_k \text{ is connected to } v_l \\ 0, & \text{otherwise} \end{cases} \tag{3}$$

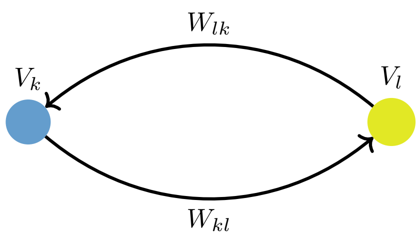
By convention the first index (k) of the adjacency matrix is the row index that denotes the *from* node, while the second index l is the column index that denotes the *to* node.

2.1.10 The Weighted Adjacency Matrix W

Weighted graphs map naturally to real, positively valued, matrices. The weight function W over edges can be represented as W_{kl}, a square matrix of $N \times N$ dimension (N being the order of the graph G). The information contained in the SUT graph can be displayed in matrix form by recording all edges between the **S** and **P** nodes in the W matrix.

2.1.11 The Transposed Weight Matrix W^T

The transposition (flipping along the diagonal) of the weights matrix W is analogous to reversing the direction of the edges between *all* pairs of nodes (switching the upstream and downstream directions). The result is that if an entry W_{kl} indicates that there is a transfer or linkage with a measured value from node k to node l then the transpose weight matrix indicates this linkage is reversed, i.e., node k *received* that quantity from node l.



If the weight matrix $W = W_{kl}$ is associated with downstream product flow then the transpose matrix $W^T = W_{lk}$ expresses the reverse upstream monetary flow.

2.2 The Supply and Use Graph

2.2.1 The Overall Shape

The overall shape of the SUT weight matrix can be arranged to have the typical block structure associated with a bipartite graph:

$$W = \begin{bmatrix} 0 & U \\ V & 0 \end{bmatrix} \tag{4}$$

The zero blocks along the diagonal indicate that there are no links between the two groups of nodes of the bipartite graph. In this representation we order first the product nodes, so the U sub-matrix expresses the weights from **P** nodes to **S** nodes and vice-versa for the V matrix. The transpose weight matrix of a SUT graph is:

$$W^T = \begin{bmatrix} 0 & V^T \\ U^T & 0 \end{bmatrix} \quad (5)$$

We discuss the economic meaning of U and V sub-matrices further below. Economic arguments and additional assumptions will further characterize the SUT graph. Special connectivity patterns can be expressed in terms of the in-degree and out-degree matrices to which we turn next.

2.2.2 The Weighted In-Degree and Out-Degree Matrices D

For every node v_l, the *in-degree* value $D^{in}(v)$ is the sum of assigned weights of all edges ending in the node. This is the same as the column sum of the W matrix, cumulating over the first (row) index.

$$D^{in}_{ll} = \sum_k^N W_{kl} \quad (6)$$

The *out-degree* value $D^{out}(v)$ is the corresponding sum of weights of all edges emanating from the node. This is the row sum of W, cumulating over the second (column) index.

$$D^{out}_{ll} = \sum_k^N W_{lk} \quad (7)$$

The in/out-degree matrices D are diagonal $N \times N$ matrices.

2.2.3 The Use Matrix U

The *Use Table* or Input Table or Absorption matrix is the portion of W matrix that shows the use of goods and services by a Sector. The value of each element of the Use matrix records *purchases* of Product p by Sector i from the corresponding market node. Typically denoted U, the Product-by-Sector Use Matrix, shows the monetary amount U_{pi} of Product p used by Sector i. Mathematically the Use matrix U_{pi} is of dimension $m \times n$ and links the Product nodes $p \in [1, m]$ to the Sectors $i \in [1, n]$. In graph terms the Use matrix assigns edge weights to a subset of the edges in the global graph G.

$D^{in}(S_i)$ is the Sector In-degree, the sum of monetary value of the Products used by sector S_i.

$$D^{in}(S_i) = \sum_p^m U_{pi} \quad (8)$$

$D^{out}(P_p)$ is the Product Out-degree, the value of product P_p usage by all sectors.

$$D^{out}(P_p) = \sum_i^n U_{pi} \quad (9)$$

2.2.4 The Supply Matrix V

Correspondingly, V is a Sector-by-Product *Supply Matrix* (also Output Table), with V_{ip} indicating the volume of output of Product p by Sector i in monetary terms. In bipartite graph terms the Supply matrix is a weight function on the subset of the global graph G that points from Sector nodes to Product nodes. In a SUT graph a Sector i may produce an arbitrary number of Products p. One of those might be classified as a primary product (that characterizes the sector) and the rest as secondary Products. The Supply table in a general SUT will contain values that do not relate to the principal Product of a Sector. As an example, several industries besides the energy sectors may produce various forms of usable energy as byproducts. When the number of Products exceeds the number of industries $m \geq n$, as is the case in general, this phenomenon emerges by construction. In special cases, the number of Products may be assumed to be the same as that of Sectors, but without having a one-to-one association between Sector and Product. If all industries do indeed produce only a unique Product each, then the supply table is *diagonal* and only the diagonal V_{ii} elements will be non-zero. The name *Make Table* historically denotes the transpose of the Supply Table V^T with elements V_{pi} which is sometimes more convenient to depict along U as they have the same ordering of indices. As we will see below, environmental impact functions will only have support over the subset of edges present in the Supply matrix.

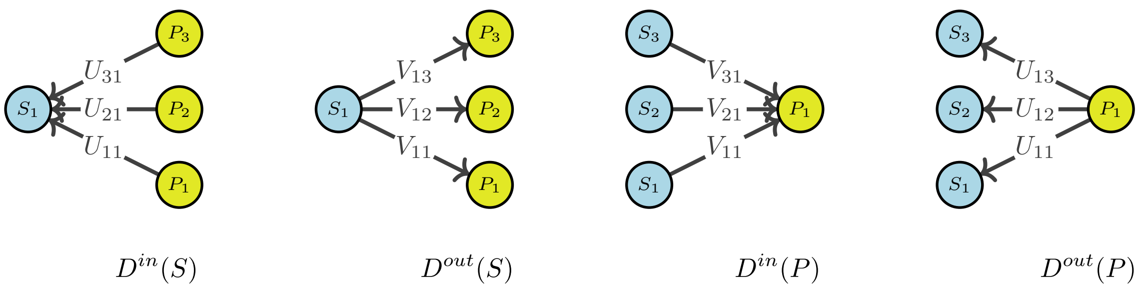
$D^{out}(S_i)$ is the Sector Out-degree, the volume of Products produced by sector S_i.

$$D^{out}(S_i) = \sum_{p}^{m} V_{ip} \tag{10}$$

$D^{in}(P_p)$ is the Product In-degree, the volume of product P_p contributed by Sectors.

$$D^{in}(P_p) = \sum_{i}^{n} V_{ip} \tag{11}$$

Pictorially the in/out degree sums for **S** and **P** nodes are as follows:



2.2.5 About Sources and Sinks

A given node v_k is a *sink node* if it has no outgoing edges, in other words, if the out-degree of v_k is 0. A node v_k is a *source node* if there are no incoming edges, in other words, if the in-degree of v_k is 0. A node is *isolated* if it has zero both in- and out-degrees. We will assume the SUT graph has no isolated nodes.

We will consider in parallel the two important sub-categories of SUT graphs: *Closed Systems* that exhibit no sinks or sources and *Open Systems* that have one or more sink and source nodes.

2.2.6 Open and Closed SUT Graph Characterization

For both closed and open systems there is always at least one Sector using a Product, or $D^{out}(P) > 0$. In matrix terms, the row sum of the first n rows of W cannot be zero. This means that there is no totally useless Production. For both closed and open systems a Sector is using at least one Product, or $D^{in}(S) > 0$. This means that there is no creation ex-nihilo, *some* input is required for any production. The further conditions on the in/out-degree matrices depend on whether we are dealing with a closed on open SUT system.

2.2.7 About Value Added

In a closed system there is at least one edge from *some* Sector into *any* Product node, or $D^{in}(P) > 0$. This means that there is at least one Sector **S** producing a Product **P**. A Product cannot produce itself (no parthenogenesis), so there is always at least one link coming from a Sector node. In matrix terms the column sum of the first m columns of the W cannot be zero.

In open systems there can be Product nodes (will be called *Value Added*) that violate the constraint of no parthenogenesis. Value Added v nodes are source nodes $v \in \mathbf{P}$ that account for ad-hoc (exogenous) inputs to Production. They satisfy $D^{in}(v) = 0$. In realistic systems value added may have several distinct components such as labor, depreciation of capital, indirect business taxes, and imports. Each one of those could be a distinct source node if desired.

2.2.8 About Final Demand

In a closed system a Sector must have some output $D^{out}(S) > 0$, there are no "parasitic" sectors. In an open system, a Sector may not have any outputs $D^{out}(S) = 0$ if it is a *sink node*. We call these $v \in \mathbf{S}$ nodes *Final Demand* sectors. It is possible for a Product not to have an edge into a final demand node if it is a strictly intermediate product that is only consumed by producing Sectors. In summary:

Connectivity Metric	Closed SUT System	Open SUT System
$D^{in}(P)$	> 0	$= 0$ (Value Added Node)
$D^{out}(P)$	> 0	> 0
$D^{in}(S)$	> 0	> 0
$D^{out}(S)$	> 0	$= 0$ (Final Demand Node)

Summary of in/out degree properties of closed and open SUT graphs.

2.2.9 Weighted Adjacency Matrix for a Closed SUT

In a closed model the weighted SUT matrix reads in more detail as:

$$
W_{kl} = \begin{bmatrix}
0 & 0 & \cdots & 0 & U_{11} & U_{12} & \cdots & U_{1n} \\
0 & 0 & \cdots & 0 & U_{21} & U_{22} & \cdots & U_{2n} \\
\vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & 0 & U_{m1} & U_{m2} & \cdots & U_{mn} \\
V_{11} & V_{12} & \cdots & V_{1m} & 0 & 0 & \cdots & 0 \\
V_{21} & V_{22} & \cdots & V_{2m} & 0 & 0 & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\
V_{n1} & V_{n2} & \cdots & V_{nm} & 0 & 0 & \cdots & 0
\end{bmatrix} \tag{12}
$$

The global column and row indices k, l run over the product and sector sub-indices $[1, \ldots, m, m+1, \ldots n]$. The weighted adjacency matrix of a SUT graph is in general an asymmetric matrix, namely $W^T \neq W$.

2.2.10 Weighted Adjacency Matrix for an Open SUT

An open SUT graph is characterized by source and sink nodes. In the standard arrangement of IO with exogenous households, source nodes are of the Product type. They deliver *value added* such as labor to Sector nodes. Sink nodes are on the other hand of the Sector type (e.g., households that consume, or create *final demand* from Product nodes. In an open system that is fully *closable*, i.e., represents an underlying closed system but in open form, the total value added (source) must be equal to the total final demand (sink). The *weighted adjacency matrix* of such an open SUT system can be organized as follows:

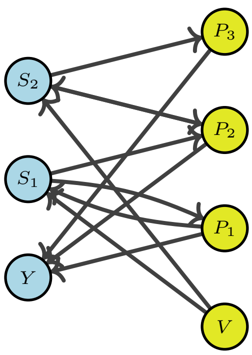
$$
W_{kl} = \begin{bmatrix}
0 & 0 & \cdots & 0 & 0 & U_{11} & U_{12} & \cdots & U_{1,n-1} & y_1 \\
0 & 0 & \cdots & 0 & 0 & U_{21} & U_{22} & \cdots & U_{2,n-1} & y_2 \\
\vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\
0 & 0 & 0 & 0 & 0 & U_{m-1,1} & U_{m-1,2} & \cdots & U_{m-1,n-1} & y_{m-1} \\
0 & 0 & \cdots & 0 & 0 & v_1 & v_2 & \cdots & v_{n-1} & 0 \\
V_{11} & V_{12} & \cdots & V_{1,m-1} & 0 & 0 & 0 & \cdots & 0 & 0 \\
V_{21} & V_{22} & \cdots & V_{2,m-1} & 0 & 0 & 0 & \cdots & 0 & 0 \\
\vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\
V_{n-1,1} & V_{n-1,2} & \cdots & V_{n-1,m-1} & 0 & 0 & 0 & \cdots & 0 & 0 \\
0 & 0 & \cdots & 0 & 0 & 0 & 0 & \cdots & 0 & 0
\end{bmatrix} \tag{13}
$$

The nominal dimension of the W matrix is still the same $(N \times N)$. The i and p indices are running to $n-1$ and $m-1$ respectively spanning the so-called *transient nodes*. The last column captures outflow into the sink Sector node Y while the last row (all zeros) indicates the sink node has no inputs. The middle zero column indicates there are no inflows into the source v product node. The corresponding middle row shows the outflows from the source node into sectors. This W matrix is evidently singular $(\det(W) = 0)$ as it has zero rows and columns.

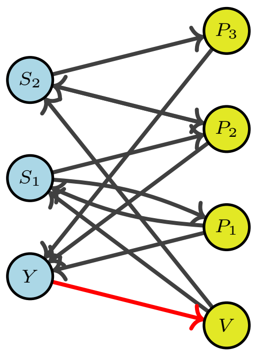
2.2.11 From Closed to Open Systems and Back

An open configuration is the most typical way to arrange a SUT graph as it facilitates "what-if" types of analysis but in contrast to the closed form, it is highly non-unique. A closed SUT system can be converted into an open one by identifying a Sector, (say households) and *disconnecting* their inputs (their consumption of products) from their outputs (their supply of labor to other sectors). A closed system can be "opened" in any number of ways (effectively equal to the number of nodes).

The reverse process of constructing a closed SUT system from an already developed open system[4] involves a rearrangement that is termed *endogenizing* some economic actors. In graph terms this means connecting the source and sink nodes and creating "regular" nodes (whether of the **S** or **P** type). In the most frequent use case, this procedure is applied to household nodes (individuals) and is known as the *closing the model* with respect to households. The transformation eliminates household demand from the total final demand and similarly removes the labor value added from the total value added. Such input–output systems are known as *semi-closed* or extended systems. An open system cannot be closed fully if its not balanced (more on that below). From a graph connectivity and usability perspective these opening / closing processes are non-trivial. They imply additional assumptions and biases that may affect outcomes.



Open SUT System in bipartite form. The Y node has no outgoing edges while the V node has no ingoing edges.



Closing the SUT System by adding an edge between Y and V. Only possible for a balanced system.

It is maybe instructive to introduce a simple means to close a SUT system by changing a single value. In the below weight matrix, notice the value w in the last row. This will be zero for an open SUT, reflecting that the sink node y does not have an edge into any **S** or **P** node and the source node v does not receive and edge from any **S** or **P** node. To close the SUT we set $w = D^{in}(v) = D^{out}(y)$.

[4] One might also construct such systems directly from empirical data, going under the name Social Accounting Matrices (SAM)

$$W_{kl} = \begin{bmatrix} 0 & 0 & \cdots & 0 & 0 & U_{11} & U_{12} & \cdots & U_{1,n-1} & y_1 \\ 0 & 0 & \cdots & 0 & 0 & U_{21} & U_{22} & \cdots & U_{2,n-1} & y_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & U_{m-1,1} & U_{m-1,2} & \cdots & U_{m-1,n-1} & y_{m-1} \\ 0 & 0 & \cdots & 0 & 0 & v_1 & v_2 & \cdots & v_{n-1} & 0 \\ V_{11} & V_{12} & \cdots & V_{1,m-1} & 0 & 0 & 0 & \cdots & 0 & 0 \\ V_{21} & V_{22} & \cdots & V_{2,m-1} & 0 & 0 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ V_{n-1,1} & V_{n-1,2} & \cdots & V_{n-1,m-1} & 0 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & \cdots & 0 & w & 0 & 0 & \cdots & 0 & 0 \end{bmatrix} \quad (14)$$

2.3 Environmental Impact Extensions

2.3.1 Modeling Environmental Impact

In the previous section we expressed key SUT concepts and tools in the language of graphs. Here we expand the toolkit to include the concept of *Environmental Impact* in the spirit that this is usually applied in Environmentally-Extended Input-Output frameworks. In this approach, in order to trace the impact of economic activity, input-output frameworks are extended to account for so-called production factors or *environmental stressors* (such as resources utilized, pollution generated etc.) that are associated with a given amount of economic activity. The working assumption in this approach is that environmental impact can be adequately modeled as a global fungible metric (e.g., the volume of gases that are emitted into the atmosphere) and that this physical measurement can be linked causally to the economic quantities captured in the SUT system.

2.3.2 Impact Factors in IO frameworks

In practice, extended Input-Output models consider environmental impacts as additional rows added to the basic structure of the IO model. This can be done either at Sector or Product level. Widely used official statistics data and EEIO tables provide region or sector-specific average impact factors expressed per economic activity (e.g., tCO2eq/€ of revenue). A vector of impact coefficients f is defined, where each element f_p represents the impact of production (in its own physical units) associated with one monetary unit of total Product output x^p. Correspondingly, F is the vector of per-sector absolute environmental impact. The actual impact of a Product with total output x_p is $F_p = f_p x_p$. The total environmental impact of the economy is the sum $F_T = \sum_p F_p$.

2.3.3 Impact Factors in a SUT system

While in symmetric IO systems impact factors form a vector associated with the Sector/Product clusters, in an asymmetric SUT context where $m \neq n$ the natural container of impact factors is, in principle, a matrix, namely a Sector-by-Product Impact Matrix F_{ip} indicating the absolute impact associated with V_{ip} volume of output of Product p by Sector i. Since in a SUT graph a Sector i may produce an arbitrary number of Products, there is also an arbitrary number of impacts. The general form of the impact weight function is:

$$F_{kl} = \begin{bmatrix} 0 & \cdots & 0 & 0 & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & 0 & 0 & \cdots & 0 \\ F_{11} & \cdots & F_{1m} & 0 & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ F_{n1} & \cdots & F_{nm} & 0 & 0 & \cdots & 0 \end{bmatrix} \tag{15}$$

This matrix can be factored in terms of impact intensities

$$F_{kl} = \begin{bmatrix} 0 & \cdots & 0 & 0 & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & 0 & 0 & \cdots & 0 \\ f_{11}V_{11} & \cdots & f_{1m}V_{1m} & 0 & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ f_{n1}V_{n1} & \cdots & f_{nm}V_{nm} & 0 & 0 & \cdots & 0 \end{bmatrix} \tag{16}$$

The general impact intensity matrix allows more flexibility in modeling varying sectoral impacts.

2.3.4 Single Technology Assumption

Under the assumption that all sectors produce the same product using the *same technology*, and thus have the same environmental impact intensity, the matrix can be factored as:

$$F_{kl} = \begin{bmatrix} 0 & \cdots & 0 & 0 & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & 0 & 0 & \cdots & 0 \\ f_1V_{11} & \cdots & f_mV_{1m} & 0 & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ f_1V_{n1} & \cdots & f_mV_{nm} & 0 & 0 & \cdots & 0 \end{bmatrix} \tag{17}$$

in other words, $f_{ip} = f_p$, the intensity has no dependency on the sector index.

3 Probabilistic Interpretations and Random Walks

Up to this point we did not make any fundamental assumptions that are not already embedded in the production of the official SUT tables. For the weighted graph $G = (V, E, W, F)$ to become a tool in the attribution of environmental impact further assumptions and interpretations are necessary. We start by expressing the *balanced matrix* properties inherent in the SUT system. These properties are also called *accounting identities* or *conservation laws*. Subsequently we will introduce normalizations and interpretations that create a additional structure enabling the attribution task.

3.1 Economic Conservation Laws and Stochastic Matrices

3.1.1 Balanced Matrices

The matrix of a directed graph is *weight-balanced* if, at each node, the sum of the weights of the incoming edges is equal to the sum of the weights of the outgoing edges.

$$\sum_k W_{kl} = \sum_k W_{lk} \tag{18}$$

or in terms of the in/out degree matrices:

$$D^{out}_{ll} = D^{in}_{ll}, \ \forall l \tag{19}$$

Introducing the column vector of row sums $Z_k = \sum_l W_{kl}$, the balance expression can be written also as $Z^T = Z$

3.1.2 Closed SUT System Balance

We obtain the total Use of a Product p as a column vector x_p by summing over all sectoral usage (summing over all Sectors i in the use table U).

$$x_p = \sum_i^n U_{pi} = D^{out}(P_p) \tag{20}$$

The total Supply of a Product p sums over all supplies (by all industries i). These totals are a row vector of column sums of the Supply matrix V.

$$x_p^T = \sum_i^n V_{ip} = D^{in}(P_p) \tag{21}$$

The total Input of a Sector i sums over all Product inputs p. These totals form a row vector that involves the column sums of Use matrix U.

$$g_i^T = \sum_p^m U_{pi} = D^{in}(S_i) \tag{22}$$

The total Output of Sector i sums over all Product outputs p. These totals form a column vector of the row sums of Supply matrix V.

$$g_i = \sum_p^m V_{ip} = D^{out}(S_i) \tag{23}$$

All the above can be are organized in a table (not a matrix!) as follows:

$$\left[\begin{array}{cc|c} 0 & U & x \\ V & 0 & g \\ \hline x^T & g^T & \end{array}\right] \tag{24}$$

In a stylized equilibrium economy and a monetary system with a fixed quantity of money[5] monetary input must match monetary output. In these circumstances monetary flows satisfy the constraint that the total amount of flow into any node equals the total amount of flow out of it. Thus, expressed in monetary terms, the total Inputs and Outputs of each Sector node must be equal and, similarly, the total Use output of a Product node must equal its total Supply inputs. For all Sector nodes i, the total Input (left side of the equation) is equal with the total Output.

$$g_i^T = g_i \tag{25}$$

Similarly for Product nodes p

$$x_p^T = x_p \tag{26}$$

Hence,

$$x_p = \sum_i^n V_{ip} = \sum_i^n U_{pi} \tag{27}$$

$$g_i = \sum_p^m V_{ip} = \sum_p^m U_{pi} \tag{28}$$

These equations are similar to Kirchhoff's Law for electric circuits, where they reflect the conservation of charge.

3.1.3 Open SUT System Balance

In the case of an open SUT system the accounting equations, at Product and Sector level are respectively:

$$x_p = \sum_i^n V_{ip} = y_p + \sum_i^n U_{pi} \tag{29}$$

$$g_i = \sum_p^m V_{ip} = v_i + \sum_p^m U_{pi} \tag{30}$$

In an arbitrary open graph the total inflow is mathematically independent (need not be equal) from the total outflow. We focus on closable systems that in addition satisfy the closure constraint:

$$\sum_i^n v_i = \sum_p^m y_p = w \tag{31}$$

where w is the total inflow/outflow. Here we assume this to be a scalar, i.e., only one source and sink node respectively but this can be generalized. This condition can also be written as:

[5] Thus no private or sovereign money creation

$$\sum_p^m \sum_i^n V_{ip} = w + \sum_p^m \sum_i^n U_{pi} \tag{32}$$

The balance equations can be organized in a table where the last column and row represent row and column sums respectively.

$$\left[\begin{array}{ccc:c} 0 & U & y & x \\ V & 0 & 0 & g \\ 0 & v & 0 & w \\ \hdashline x^T & g^T & w & \end{array}\right] \tag{33}$$

3.2 From Weights to Probabilities

3.2.1 Stochastic Matrices

A *stochastic matrix* Q is a square nonnegative matrix whose rows and/or columns satisfy certain conditions. In a right or row stochastic matrix each row sums to 1,

$$\sum_l Q_{kl} = 1 \tag{34}$$

In a left or column stochastic matrix each column sums to 1:

$$\sum_k Q_{kl} = 1 \tag{35}$$

A doubly stochastic (also bistochastic) matrix satisfies both conditions simultaneously. Stochastic matrices are also called *transition matrices* or probability matrices. The raw (empirically derived) weighted adjacency matrix W will in general not be stochastic. Measurements using arbitrary units do not need to sum up to unity. Yet given a weight matrix W, one can, in-principle, derive *two* transition matrices by dividing with either the row or column sums. Of course, to divide by row sums every node has to have at least one out-going edge and to divide by column sums every node has to have at least one in-going edge.

$$Q^{out}_{kl} = \frac{W_{kl}}{z_l} \tag{36}$$

$$Q^{in}_{kl} = \frac{W_{kl}}{z_k} \tag{37}$$

What do these two matrices represent and what are they useful for? Q^{out} has the property that the *outgoing edges* from all nodes are assigned probabilities that sum to one. As we will see below this transition matrix is useful when one wants to start on some node and follow randomly the direction of the graph (downstream) by picking one of the outgoing edges with the assigned probability. Respectively Q^{in} has the property that *ingoing edges* on all nodes are assigned probabilities that sum to one. This version is useful when one wants to start from a node and follow upstream the reverse direction of the graph, picking randomly one of the incoming edges. Maybe worth mentioning that the above mappings are many-to-one, many weighted adjacency matrices map to the same transition matrix Q.

3.2.2 About Markov Chains

The stochastic matrices $Q^{out|in}$, when available, are still effectively processed versions of the empirical economic data. Given their properties though, they can be associated with discrete time Markov Chain *processes*. This association takes us into the realm of *models*. The essential assumption, which is not unlike similar assumptions in the construction of the Leontief and Ghosh IO models, is that the normalized quantities $Q^{out|in}$ are somehow intrinsic to the system, for example invariant under the passage of time.

Markov chains are relatively simple and powerful models but are based on strong assumptions. They posit that the random process being modeled has *no memory*. In other words, a sequence of random events (graph traversing steps) happening in discrete steps is such that the probability of each next event depends only on the state attained in the previous event (but no prior states).

More concretely, a discrete-time Markov chain is a sequence of random variables $X(0), X(1), X(2), X(3), \ldots$ where the possible values of the random variable X form a finite set called the *state space* of the chain. In our case the states of the Markov Chain are in one-to-one correspondence with the *nodes* V of the SUT graph. The variable X takes integer values in $[1, N]$, indicating which node is "occupied" or "visited" at each step of the process.

3.2.3 Closed SUT Stochastic Matrices

For a closed system the normalization is in more detail as follows:

3.2.4 The Row Stochastic Closed SUT Matrix

Dividing by x_p and g_i respectively we get

$$1 = \sum_i^n \frac{U_{pi}}{x_p} = \sum_i^n u_{pi} \tag{38}$$

$$1 = \sum_p^m \frac{V_{ip}}{g_i} = \sum_p^m v_{ip} \tag{39}$$

and the row-stochastic transition matrix Q^{out}:

$$Q^{out}_{kl} = \begin{bmatrix} 0 & 0 & \cdots & 0 & u_{11} & u_{12} & \cdots & u_{1n} \\ 0 & 0 & \cdots & 0 & u_{21} & u_{22} & \cdots & u_{2n} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & u_{m1} & u_{m2} & \cdots & u_{mn} \\ v_{11} & v_{12} & \cdots & v_{1m} & 0 & 0 & \cdots & 0 \\ v_{21} & v_{22} & \cdots & v_{2m} & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ v_{n1} & v_{n2} & \cdots & v_{nm} & 0 & 0 & \cdots & 0 \end{bmatrix} \tag{40}$$

3.2.5 The Column Stochastic Closed SUT Matrix

Using column sums we would obtain alternatively:

$$1 = \sum_{p}^{m} \frac{U_{pi}}{g_i} = \sum_{p}^{m} \tilde{u}_{pi} \tag{41}$$

$$1 = \sum_{i}^{n} \frac{V_{ip}}{x_p} = \sum_{i}^{n} \tilde{v}_{ip} \tag{42}$$

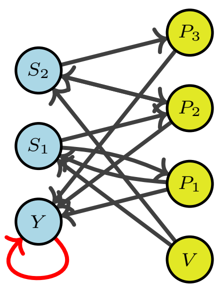
.

which produces a similar expression as 40 replaced with the tilde variables. For closed SUT graphs both Q matrices exist and they retain the simple block structure of the weight matrix:

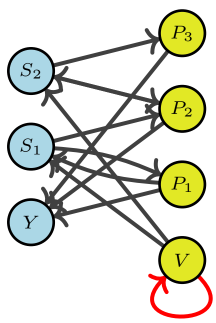
$$Q^{in|out} = \begin{bmatrix} 0 & u^{in|out} \\ v^{in|out} & 0 \end{bmatrix} \tag{43}$$

3.2.6 Open SUT system and Absorbing Markov Chains

In an open SUT system the creation of a stochastic transition matrix from the weighted adjacency matrix involves a subtlety: as we saw, the matrix is singular (does not have full rank). The zero rows and columns encode the fact that sink nodes have zero oud-degree and source nodes have zero in-degree. A simple modification that addresses this is to turn these nodes into *absorbing* nodes. Intuitively, once the state of the system hits a sink node (for a downstream flow) or a source node (for an upstream flow), we want it to stay there for all future steps. Pictorially this is shown as follows:



Absorbing Chain at Final Demand (Sink). Downstream flow reaching Y will stay there.



Absorbing Chain at Value Added (Source). Upstream flow reach V will stay there.

For an absorbing Markov chain with one absorbing state, the transition matrix Q can be arranged to have the following block structure, known as the canonical form:

$$Q = \begin{bmatrix} B & R \\ 0 & 1 \end{bmatrix} \tag{44}$$

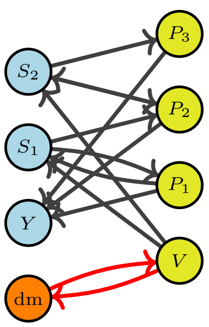
where B is a $(N-1) \times (N-1)$ transition matrix between transient states and R is $(N-1)$-dim column vector of probabilities to transition into the sink node. In our case the matrix downstream transitions would be:

$$Q^{out}_{kl} = \begin{bmatrix} 0 & 0 & \cdots & 0 & 0 & u_{11} & u_{12} & \cdots & u_{1n} & \tilde{y}_1 \\ 0 & 0 & \cdots & 0 & 0 & u_{21} & u_{22} & \cdots & u_{2n} & \tilde{y}_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & u_{m1} & u_{m2} & \cdots & u_{mn} & \tilde{y}_m \\ 0 & 0 & \cdots & 0 & 0 & \tilde{v}_1 & \tilde{v}_2 & \cdots & \tilde{v}_n & 0 \\ v_{11} & v_{12} & \cdots & v_{1m} & 0 & 0 & 0 & \cdots & 0 & 0 \\ v_{21} & v_{22} & \cdots & v_{2m} & 0 & 0 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ v_{n1} & v_{n2} & \cdots & v_{nm} & 0 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & \cdots & 0 & 0 & 0 & 0 & \cdots & 0 & 1 \end{bmatrix} \tag{45}$$

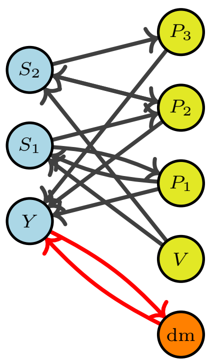
and similarly for upstream transitions.

$$Q^{in}_{kl} = \begin{bmatrix} 0 & 0 & \cdots & 0 & 0 & u_{11} & u_{12} & \cdots & u_{1n} & \tilde{y}_1 \\ 0 & 0 & \cdots & 0 & 0 & u_{21} & u_{22} & \cdots & u_{2n} & \tilde{y}_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & u_{m1} & u_{m2} & \cdots & u_{mn} & \tilde{y}_m \\ 0 & 0 & \cdots & 0 & 1 & \tilde{v}_1 & \tilde{v}_2 & \cdots & \tilde{v}_n & 0 \\ v_{11} & v_{12} & \cdots & v_{1m} & 0 & 0 & 0 & \cdots & 0 & 0 \\ v_{21} & v_{22} & \cdots & v_{2m} & 0 & 0 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ v_{n1} & v_{n2} & \cdots & v_{nm} & 0 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & \cdots & 0 & 0 & 0 & 0 & \cdots & 0 & 0 \end{bmatrix} \tag{46}$$

A peculiarity of the above construction is that the graph is no longer bipartite (and it involves a self-loop). If we want to preserve the bipartite property while still preventing the chain to wander back into transient nodes we can introduce a dummy Sector (resp. Product) that acts as a sort of *reflection chamber*.



Dummy Node that traps upstream flows. Once the flow reaches V it keeps iterating between the dummy node and V.



Dummy Node that traps downstream flows. Once it reaches Y it keeps iterating betweeen the dummy node and Y.

For example, for downstream transitions upon introducing an extra production node, we have an $(N+1) \times (N+1)$ matrix:

$$Q^{out}_{kl} = \begin{bmatrix} 0 & \cdots & 0 & 0 & 0 & u_{11} & \cdots & u_{1n} & \tilde{y}_1 \\ \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & u_{m1} & \cdots & u_{mn} & \tilde{y}_m \\ 0 & \cdots & 0 & 0 & 0 & \tilde{v}_1 & \cdots & \tilde{v}_n & 0 \\ 0 & \cdots & 0 & 0 & 0 & 0 & \cdots & 0 & 1 \\ v_{11} & \cdots & v_{1m} & 0 & 0 & 0 & \cdots & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots & 0 & \vdots & \ddots & \vdots & \vdots \\ v_{n1} & \cdots & v_{nm} & 0 & 0 & 0 & \cdots & 0 & 0 \\ 0 & \cdots & 0 & 0 & 1 & 0 & \cdots & 0 & 0 \end{bmatrix} \tag{47}$$

Thus in the open case too, if desired the transition matrices can be expressed in bipartite form:

$$Q^{out|in} = \begin{bmatrix} 0 & \hat{u}^{out|in} \\ \hat{v}^{out|in} & 0 \end{bmatrix} \tag{48}$$

3.2.7 About Random Walks

A random walk over a graph can be modeled by a Markov chain with probability transition matrix Q. The downstream transitions matrix Q^{out} captures the probability of a transition from node k to node l following the direction of the graph. Thus, conditional on $X(t) = k$ the probability of $X(t+1) = l$ is

$$Q^{out}_{kl} = \Pr(X(t+1) = l \mid X(t) = k) \tag{49}$$

This is the probability of moving forward from node k to node l **if** there is an edge (k, l), following the direction of the graph (downstream). The upstream transition matrix Q^{in} assigns the probability of *having moved* from node l to node k if there is an edge (l, k). It is thus moving at random *against* the direction of the graph (upstream).

$$Q^{in}_{lk} = \Pr(X(t) = k \mid X(t-1) = l) \tag{50}$$

Which version is suitable to use depends on the desired analysis. In either case, for the bipartite SUT graph the random walk will be a sequence of alternating sector and product nodes, e.g., $P_0, S_1, P_2, S_3, P_4, S_5 \ldots$. The walk terminates on a sink or source node if we work with an open SUT system or iterates for ever in a closed SUT system. The starting point of the random walk may equally well be a product or sector node, again depending on the question we ask. Formally a random walk on a bipartite graph represents a Markov Chain with all states having periodicity 2.

3.2.8 The State Probability Vector P

Denote $P_k(t)$ a vector of probabilities ($\sum_k P_k = 1$) that express the likelihood that the system is at a given node k at any step t. One such vector $P_k(0)$ can be taken as the starting state (initial condition) of the chain at time 0. The initial condition can also be taken to be deterministic (fully known) by using a probability vector $P_k(0)$ that has a single non-zero value equal to unity for some k. In a bipartite graph it is natural to consider probability vectors with block structure: $P^T = [p_1, \ldots, p_m, s_1, \ldots, s_n]$.

3.2.9 Kolmogorov Equation

Given $P_k(t)$, the probability of being in node l one step later, is given by:

$$P_l(t+1) = \sum_k P_k(t) Q^{out}_{kl} \tag{51}$$

or in matrix notation

$$P^T(t+1) = P^T(t) Q^{out} \tag{52}$$

Multiplying the probability vector with Q from the right represents one step forward. This is the discrete form of the *forward Kolmogorov equation.*

A similar equation applies for upstream transitions.

$$P_k(t) = \sum_l Q^{in}_{kl} P_l(t-1) \tag{53}$$

which in matrix form shows the multiplication from the left:

$$P(t) = Q^{in} P(t-1) \tag{54}$$

Apply iteratively Equation 51 to link an initial state with a final state after a number of forward transitions we get (dropping the *out* indicator):

$$P^T(t) = P^T(0) Q^t \tag{55}$$

3.2.10 Stationary Probabilities

For a closed SUT system, the vector of stationary probabilities is denoted by π and satisfies $\pi^T Q = \pi^T$ and $\pi^T 1 = 1$.

$$\pi_l = \sum_k \pi_k Q^{out}_{kl} \tag{56}$$

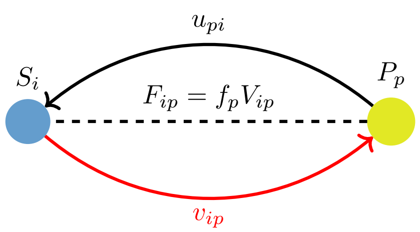
π gives the fraction of the total of visits a random walk has visited every states. In this sense it reflects the *importance* of a node.

3.3 Follow the Money to trace Impact

3.3.1 Monetary Interpretation

The idea of *following the money* on a graph is a mathematical representation of the fact that funds from purchases or sales of diverse economic actors diffuse through the econosphere via the graph edges. With the implicit assumption that the transition probabilities are time invariant, an initial €bill will travel along the graph (forward or backward) following the probabilities indicated by the edge weights. Equation 51 can be interpreted as tracing the origin of funds within an economy, where money is coming from. Equation 53 can be interpreted as tracing the destination of funds with an economy, where money is going to. In an open system money comes from final demand nodes and ends up to value added nodes. In a closed system money keeps going around.

As the initial money amount jumps from node to node, it may be associated with environmental impact.



The two types of random walk steps that transition between **S** and **P** nodes. The black edge is a use-type transition (with probability u_{pi}). The red edge is a supply-type transition with probability v_{ip}. In our formulation this is the transition type responsible for environmental impact and F_{ip} encodes the magnitude of that impact.

3.3.2 Calculating Expectations

With the availability of the propagation equations in principle any quantities concerning the multivariate distribution of $X(1), \ldots X(n)$ can be computed. The most obvious useful metrics concern *marginal* probabilities. For example the expected value of the random walk itself is given by the expression:

$$\mathbb{E}(X(t)) = \sum_k k P_k(t) \tag{57}$$

Given the arbitrary nature of the index k (the ordering of nodes is entirely by convention), this expectation does not carry any intrinsic economic meaning.

3.3.3 Expectations of Node Functions

A node function is a random variable $H(X)$ whose value is solely depending on which node is indicated by $X(t)$ at some step t. As an example of a calculation using node functions, assume that the process $X(t)$ starts in state $X(0) = k \in V$ at step 0, and further

$$H(t) = \begin{cases} H_l, & \text{if } X(t) = l \\ 0, & \text{otherwise} \end{cases} \tag{58}$$

in words, it is only non-zero if at step t the walk is on a certain node l. Then the expected value of $H(t)$ is:

$$\mathbb{E}(H(t)) = \sum_r \mathbb{E}(H(t)|X(t) = r) = H_l P_l(t) \tag{59}$$

If the starting probability distribution is $P_k(0) = u_k$ and $P(t)$ is the probability distribution over nodes t steps later then

$$P^T(t) = u^T Q^t \tag{60}$$

and

$$\mathbb{E}(H(t)) = H_l \sum_r u_r Q^t_{rl} \tag{61}$$

3.3.4 Expected Environmental Impact

For impact calculations we'll be interested in functions that depend on the state of chain in *two successive steps*. This is because impact is not a static attribute of nodes but is generated while specific edges are being traversed.

$$F(t) = \begin{cases} f_{ip}, & \text{if } X(t-1) \in \mathbf{S} \text{ and } X(t) \in \mathbf{P} \\ 0, & \text{otherwise} \end{cases} \tag{62}$$

In other words $F(t)$ is zero when the random walk is moving product from market to sectors and equal to the intensity matrix f_{ip} when moving product from sector to markets. Since $F(t)$ is a function of two random variables $X(t-1), X(t)$ we must condition on all combinations that produce non-zero value.

$$\begin{align} \mathbb{E}(F(t)) &= \sum_k \sum_l \mathbb{E}(F(t)|X(t) = k, X(t-1) = l) \tag{63} \\ &= \sum_i \sum_p \mathbb{E}(F(t)|X(t) = p, X(t-1) = i) \tag{64} \\ &= \sum_i \sum_p f_{ip}\, P(X(t) = p, X(t-1) = i) \tag{65} \\ &= \sum_i \sum_p f_{ip}\, P(X(t) = p\,|X(t-1) = i) P(X(t-1) = i) \tag{66} \\ &= \sum_i \sum_p f_{ip} Q^{in}_{ip} P_p(t-1) \tag{67} \\ &= f \circ Q \cdot P(t-1) \tag{68} \end{align}$$

In terms of the $t = 0$ distribution this is rewritten as

$$\mathbb{E}(F(t)) = \sum_k \sum_l f_{kl} Q^t_{kl} P_l(0) = f \circ Q^t \cdot P(0) \tag{69}$$

3.3.5 Cumulative Impact over Multiple Steps

It is common in EEIO impact analysis to consider so-called 1-st, 2-nd etc. *rounds* of impact. In our context this translates into calculating the expected impact for a sequence of random steps up to the desired number of rounds (which may also be infinite). Given an initial node $X(0)$, the cumulative environmental impact along a random path up to step t will be simply the sum:

$$C(t) = \sum_{s=0}^{t} F(t) \tag{70}$$

While the function $C(t)$ depends on all random variables $X(1), \ldots, X(t)$, the Markovian nature of the random process makes the calculation possible. We can calculate the expectation of the sum up to t to get:

$$\mathbb{E}(C(t)) = \sum_{t} \mathbb{E}(F(t)) = \sum_{t} f \circ Q^{t} \cdot P(0) \tag{71}$$

3.3.6 Variance of Environmental Impact

This calculation is a further example of metrics that can be easily built on top of the probabilistic interpretation of the SUT graph as a random walk. Starting with the definition of variance:

$$\mathbb{V}(F(t)) = \mathbb{E}(F(t)^2) - \mathbb{E}(F(t))^2 \tag{72}$$

The new part is the expectation of the square of the impact function. This proceeds entirely analogously:

$$\mathbb{E}(F(t)^2) = \sum_{k} \sum_{l} \mathbb{E}(F(t)^2 | X(t) = k, X(t-1) = l) \tag{73}$$

$$= \sum_{i} \sum_{p} f_{ip}^2 \, P(X(t) = p, X(t-1) = i) \tag{74}$$

$$= \sum_{i} \sum_{p} f_{ip}^2 Q_{ip} P_p(t-1) \tag{75}$$

$$= f^2 \circ Q \cdot P(t-1) \tag{76}$$

Putting things together:

$$\mathbb{V}(F(t)) = f^2 \circ Q \cdot P(t-1) - (f \circ Q \cdot P(t-1))^2 \tag{77}$$

$$= f^2 \circ Q^t \cdot P(0) - (f \circ Q^t \cdot P(0))^2 \tag{78}$$

Computing the variance metric is the standard way to obtain a sense of uncertainty around a result, in this case the *intrinsic variance* of the calculated environmental impact.

$$F(t) \approx \mathbb{E}(F(t)) \pm \sqrt{\mathbb{V}(F(t))} \tag{79}$$

Note that the uncertainty involved here is not due to *data quality uncertainties* in the underlying data collection and processing but is rather to be understood as the result of the *potential intrinsic fluctuations* in the economic circulation network. In this sense it is a model dependent result that further leverages the fundamental transition probability calculations.

A Worked out Example

As a simple numerical example we will reuse an IO system that has been discussed before [19, 30]. We consider a world with just two production sectors: Agriculture (Ag) and Manufacturing (Ma). These two sectors sell goods and services to each other and also to households who purchase the final, finished products sold by each of the sectors. In classic IO analysis we would introduce here a 2×2 inter-industry sales or transactions Z table (alongside a Final Demand column and a Value Added row).

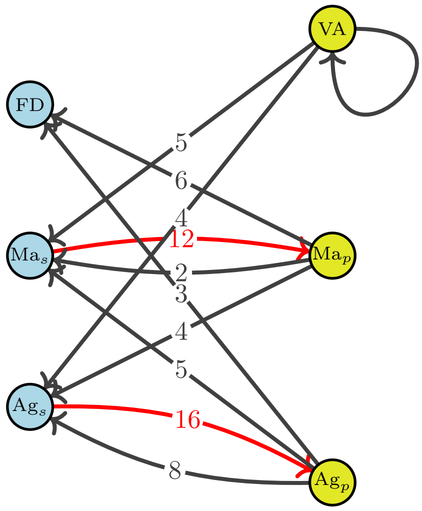
$$\left[\begin{array}{c|ccc} & Ag & Ma & FD \\ \hline Ag & 8 & 5 & 3 \\ Ma & 4 & 2 & 6 \\ VA & 4 & 5 & 0 \end{array}\right]$$

Classic (Symmetric) IO System

This table shows annual transactions between actors in the economy. For example Ag sector purchases 4€ worth of goods and services from the Ma sector and 8€ worth of goods and services from the Ag sector. To place ourselves in a SUT context we will instead imagine three Sectors: The Agriculture Sector (Ag-s), Manufacturing Sector (Ma-s) and Final Demand Sector (FD) but also three corresponding Products: The Agriculture Product (Ag-p), Manufacturing Product (Ma-p) and Value Added (VA). This process might be called a *minimal assumptions* extension from a symmetric IO system to a compatible SUT (there are obviously infinitely more possibilities).

For concreteness we can imagine Final Demand is exclusively from Households (Sector) and the Value Added is their Labor (Product) input into the economy but we keep the notation as-is for comparability purposes. For the Supply matrix S we assume a diagonal structure, namely each Sector produces its own Product. The Use matrix U is verbatim the Z matrix. With these simplest assumptions we can compose the Weights matrix W that characterizes the SUT.

A.1 Preliminaries



An open 3×3 SUT system

$W_{kl} =$

	Ag_p	Ma_p	VA	Ag_s	Ma_s	FD
Ag_p	0	0	0	8	5	3
Ma_p	0	0	0	4	2	6
VA	0	0	0	4	5	0
Ag_s	16	0	0	0	0	0
Ma_s	0	12	0	0	0	0
FD	0	0	0	0	0	0

Weight Matrix of open SUT system. Notice the zero column and zero rows corresponding to the value added product and final demand sector respectively.

The corresponding bipartite graph translates the weight matrix into a visible structure. The supply-side transactions that create environmental impact are indicated with red arrows (the amount of impact is not shown). The Value-Added node at the upper right corner is endowed with a self-loop. In the example we will traverse the graph upstream hence need to create such an absorbing node. The next step is to produce a column stochastic transition matrix Q by dividing all columns (except the one with exclusive zeros) by the column sums:

$Q_{kl} =$

	Ag_p	Ma_p	VA	Ag_s	Ma_s	FD
Ag_p	0	0	0	0.50	0.42	0.33
Ma_p	0	0	0	0.25	0.17	0.67
VA	0	0	1	0.25	0.42	0
Ag_s	1	0	0	0	0	0
Ma_s	0	1	0	0	0	0
FD	0	0	0	0	0	0

Transition Matrix of an Open SUT system (Column Stochastic Version). We insert unity at the diagonal element of Value Added to convert it into an absorbing state.

As discussed in the main document, there are various ways of addressing the singular nature of an open system's weight matrix. Here we illustrate the simplest way. As final step before calculating something we need to construct the environmental impact intensity matrix. We start with the absolute impacts associated with the supply matrix:

$$F_{kl} = \left[\begin{array}{c|cccccc} & Ag_p & Ma_p & VA & Ag_s & Ma_s & FD \\ \hline Ag_p & 0 & 0 & 0 & 0 & 0 & 0 \\ Ma_p & 0 & 0 & 0 & 0 & 0 & 0 \\ VA & 0 & 0 & 0 & 0 & 0 & 0 \\ Ag_s & 8 & 0 & 0 & 0 & 0 & 0 \\ Ma_s & 0 & 4 & 0 & 0 & 0 & 0 \\ FD & 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

The absolute Impact Matrix of the SUT system (this is input data).

We normalize the above matrix by dividing with the transaction values of the supply matrix to obtain:

$$f_{kl} = \left[\begin{array}{c|cccccc} & Ag_p & Ma_p & VA & Ag_s & Ma_s & FD \\ \hline Ag_p & 0 & 0 & 0 & 0 & 0 & 0 \\ Ma_p & 0 & 0 & 0 & 0 & 0 & 0 \\ VA & 0 & 0 & 0 & 0 & 0 & 0 \\ Ag_s & 0.5 & 0 & 0 & 0 & 0 & 0 \\ Ma_s & 0 & 0.33 & 0 & 0 & 0 & 0 \\ FD & 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

Impact Intensity Matrix of the SUT system (derived from W and F).

Now we are ready to compute! We imagine that we (a generic consumer in the final demand sector) purchase a 1€ shirt from a clothing business (part of the Ma sector). A classic EEIO calculation is to find the total upstream emissions associated with this shirt.[6]

A.2 Step 1

In the first step we start at the FD node (consumers of shirts). Mathematically this is described by the the probability mass being one at $k = 6$.

$$P^T(0) = [0, 0, 0, 0, 0, 1] \tag{80}$$

Purchasing the shirt means we move upstream to node Ma$_p$, the product market where the manufacturing sector has deposited the shirt. Here we are *conditioning* on the 1€ being spent on the shirt, so the broader menu of options (purchasing something from the Agriculture sector) does not apply.

The chain now occupies the state $k = 2$ with certainty:

$$P^T(1) = [0, 1, 0, 0, 0, 0] \tag{81}$$

The shirt purchased has already been produced, so the first transition (buying it) does not entail any material impact in itself, hence:

[6] If the actual shirt is above (or below!) that value we will simply scale the result! This linearity is a fundamental feature of IO models.

$$\mathbb{E}(F(1)) = 0 \tag{82}$$

A.3 Step 2

We must continuing upstream to trace the source of emissions. In this simple case there is only one sector producing shirts so the only direction upstream is to the node Ma_s. That transition is modeled as the matrix-vector product $QP_k(1)$. The result is (with certainty) moving from product sector k=2 to the manufacturing sector k=5:

$$P^T(2) = [0, 0, 0, 0, 1, 0] \tag{83}$$

The associated second step impact is derived by applying equation 63:

$$\mathbb{E}(F(2)) = \sum_k \sum_l \mathbb{E}(F(t)|X(2) = k, X(1) = l) \tag{84}$$

$$= \sum_i \sum_p \mathbb{E}(F(t)|X(2) = i, X(1) = p) \tag{85}$$

$$= \sum_i \mathbb{E}(F(t)|X(2) = i, X(1) = 2) \tag{86}$$

$$= \sum_i f_{i2} P(X(2) = i, X(1) = 2) \tag{87}$$

$$= \sum_i f_2 P(X(2) = i\,|X(1) = 2) P(X(1) = 2) \tag{88}$$

$$= \sum_i f_2 P(X(2) = i\,|X(1) = 2) \tag{89}$$

$$= \sum_i f_2 Q_{i2} \tag{90}$$

$$= f_2 Q_{42} + f_2 Q_{52} + f_2 Q_{62} \tag{91}$$

$$= f_2 Q_{52} = 0.33 \times 1 = 0.33 \tag{92}$$

The first line of the calculation is conditioning on all relevant states of X at steps 1 and 2. The second line focuses on the only non-zero elements (from sector to product). The third line conditions on us being on node=2 with certainty. The fourth line inserts the impact intensity for this path. The fifth line uses again conditioning to bring out the Markov Chain transition matrix and we finally have a sum-product which in this case has only one non-zero value. The impact attributed to $u = 1$€ is thus

$$\mathbb{E}(F(2)) = 0.33 \tag{93}$$

Let us also compute the variance of this result. Starting with the variance formula 77 and going through the same steps we get:

$$\mathbb{V}(F(2)) = (0.33)^2 - (0.33)^2 = 0 \tag{94}$$

This is not terribly surprising as there was no uncertainty as to which path we had to traverse so far. This will not the case in the further steps.

A.4 Step 3

Continuing upstream we now have to trace the ingredients that went into the making of the shirt. There are three inputs, labor from the VA node, and both manufactured products and agricultural products from the respective markets. Since there are thee possible paths the 1€ must follow each one with its own probability:

$$P^T(3) = [0.42, 0.17, 0.42, 0, 1, 0] \tag{95}$$

As for the expected impact in this step going through the same motions we find:

$$\begin{aligned}
\mathbb{E}(F(3)) &= \sum_k \sum_l \mathbb{E}(F(3)|X(3) = k, X(2) = l) &(96)\\
&= \sum_i \sum_p \mathbb{E}(F(3)|X(3) = i, X(2) = p) &(97)\\
&= \sum_i \sum_p f_p P(X(3) = i, X(2) = p) &(98)\\
&= \sum_i \sum_p f_p P(X(3) = i\,|X(2) = p) P(X(2) = p) &(99)\\
&= \sum_i \sum_p f_p Q P_p(2) = 0 &(100)
\end{aligned}$$

or

$$\mathbb{E}(F(3)) = 0 \tag{101}$$

The pattern is that the odd steps of the upstream random walk, moving from sectors to their input product markets does not create impact.

A.5 Step 4

In the fourth step, the transition matrix applied to the step 3 probability vector produces

$$P^T(4) = [0, 0, 0, 0.42, 0.42, 0.17] \tag{102}$$

Going again through the calculation in detail:

$$\mathbb{E}(F(4)) = \sum_k \sum_l \mathbb{E}(F(4)|X(4) = k, X(3) = l) \quad (103)$$
$$= \sum_i \sum_p \mathbb{E}(F(4)|X(4) = i, X(3) = p) \quad (104)$$
$$= \sum_i \sum_p \mathbb{E}(F(4)|X(4) = i|X(3) = p)P(X(3 = p)) \quad (105)$$
$$= \sum_i \sum_p f_{ip} P(X(4) = i, X(3) = p)P(X(3) = p) \quad (106)$$
$$= \sum_i \sum_p f_p Q_{ip} P(X(3) = p) \quad (107)$$
$$= \sum_i f_1 Q_{i1} P(X(3) = 1) + \sum_i f_2 Q_{i2} P(X(3) = 2) \quad (108)$$
$$= \sum_i 0.5\, Q_{i1} 0.42 + \sum_i 0.33\, Q_{i2} 0.17 \quad (109)$$
$$= 0.5\, Q_{31} 0.42 + 0.33\, Q_{42} 0.17 \quad (110)$$
$$= 0.26 \quad (111)$$

With a similar calculation for the variance.

$$\mathbb{V}(F(4)) = (0.123) - (0.26)^2 = 0.05 \quad (112)$$

This means that the volatility intrinsic to the expected impact is substantial. This is due to the fact we have only two sectors.

A.6 Step 5

Step 5 is again an emissionless step and the landing probabilities of the upstream journey are:

$$P^T(5) = [0.28, 0.13, 0.59, 0, 0, 0] \quad (113)$$

Notice that the probability weight of the VA node keeps increasing. This is the action of an absorbing state. Eventually all the probability mass of the chain will be cumulated here.

$$\mathbb{E}(F(3)) = \sum_k \sum_l \mathbb{E}(F(3)|X(3) = k, X(2) = l) \quad (114)$$
$$= \sum_i \sum_p \mathbb{E}(F(3)|X(3) = i, X(2) = p) \quad (115)$$
$$= \sum_i \sum_p f_p P(X(3) = i, X(2) = p) \quad (116)$$
$$= \sum_i \sum_p f_p P(X(3) = i\,|X(2) = p)P(X(2) = p) \quad (117)$$
$$= \sum_i \sum_p f_p Q^{in}_{ip} P_p(2) = 0 \quad (118)$$

Again:

$$\mathbb{E}(F(5)) = \mathbb{V}(F(5)) = 0 \tag{119}$$

A.7 Step 6

In the final step (of our journey, in principle we can go on) we get a probability vector

$$P^T(3) = [0, 0, 0.59, 0.28, 0.13, 0] \tag{120}$$

Repeating the same calculation procedure:

$$\mathbb{E}(F(6)) = 0.18 \tag{121}$$

$$\mathbb{V}(F(6)) = 0.05 \tag{122}$$

and so on and so forth!

References

[1] Intergovernmental Panel on Climate Change. Climate Change 2021: The Physical Science Basis. *A Report of the IPCC*, 2021.

[2] Intergovernmental Panel on Climate Change. Climate Change 2014: Synthesis Report. *A Report of the IPCC*, 2015.

[3] United Nations. Handbook of National Accounting: Integrated Environmental and Economic Accounting 2003, 2003.

[4] Eurostat. European system of accounts (ESA 2010, 2010.

[5] EUROSTAT. Eurostat Manual of Supply, Use and Input-Output Tables, 2008.

[6] K.W. Steininger et al. Multiple carbon accounting to support just and effective climate policies. *Nature Climate Change 6, 35-41*, 2016.

[7] L. Chancel and Y. Rehm. The Carbon Footprint of Capital: Evidence from France, Germany and the US based on Distributional Environmental Accounts. *Preprint*, 2023.

[8] Greenhouse Gas Protocol. Corporate Value Chain (Scope 3) Accounting and Reporting Standard. *Report*, 2013.

[9] PCAF. The Global GHG Accounting and Reporting Standard for the Financial Industry. First edition. *Report*, 2020.

[10] P. Papadopoulos. WP11: Sustainable Portfolio Management: Attribution and Allocation of Greenhouse Gas Emissions. *Open Risk White Papers*, 2021. Online Link.

[11] Manfred Lenzen and Joy Murray. Conceptualising environmental responsibility. *Ecological Economics*, 70:261–270, 2010.

[12] L.J. McDonald et al. Towards a unified carbon accounting landscape. *Phil. Trans. R. Soc. A 382*, 2024.

[13] A. Tukker, H. Pollitt and M. Hnkemans. Consumption-based carbon accounting: sense and sensibility. *Climate Policy, VOL. 20, NO. SUP1, S1-S13*, 2020.

[14] Blanca Gallego and Manfred Lenzen. A Consistent Input-Output Formulation of Shared Producer and Consumer Responsibility. *Economic Systems Research, Vol. 17, No. 4, 365 –391*, 2005.

[15] Alexandra Marques, João Rodrigues, Manfred Lenzen, Tiago Domingos. Income-based environmental responsibility. *Ecological Economics*, 2012.

[16] Ronald E. Miller and Peter D. Blair. *Input-Output Analysis: Foundations and Extensions*. Cambridge, 2008.

[17] João F. D. Rodrigues, Daniel Moran, Richard Wood, and Paul Behrens. Uncertainty of Consumption-Based Carbon Accounts. *Environmental Science & Technology*, 52(13):7577–7586, 2018.

[18] Remond-Tiedrez, Isabelle and Rueda-Cantuche et al. *European Union inter-country supply, use and input-output tables - Full international and global accounts for research in input-output analysis (FIGARO)*. 03 2019.

[19] J. Kitzes. An Introduction to Environmentally-Extended Input-Output Analysis. *Resources, 2, 489-503*, 2013.

[20] D. Horen Greenford, T. Crownshaw, C. Lesk, K. Stadler and H. D. Matthews. Shifting economic activity to services has limited potential to reduce global environmental impacts due to the household consumption of labour. *Environ. Res. Lett. 15 064019*, 2020.

[21] S. Boccaletti et al. The structure and dynamics of multilayer networks. *Physics Reports*, 2014.

[22] M. Kivela et al. Multilayer networks. *Journal of Complex Networks*, 2014.

[23] Pauliuk, S., G. Majeau-Bettez, and D.B. Müller. A general system structure and accounting framework for socioeconomic metabolism. *Journal of Industrial Ecology*, 2015.

[24] Manfred Lenzen and José Manuel Rueda-Cantuche. A note on the use of supply-use tables in impact analyses. *Sort-statistics and Operations Research Transactions*, 36:139–152, 2012.

[25] Florian Blöchl, Fabian J. Theis, Fernando Vega-Redondo, and Eric O'N. Fisher. Vertex centralities in input-output networks reveal the structure of modern economies. *Phys. Rev. E*, 83:046127, Apr 2011.

[26] Ariel L. Wirkierman, Monica Bianchi, and Anna Torriero. Leontief Meets Markov: Sectoral Vulnerabilities Through Circular Connectivity. *Networks and Spatial Economics*, 22(3):659–690, September 2022.

[27] Carlo Piccardi, Massimo Riccaboni, Lucia Tajoli, and Zhen Zhu. Random walks on the world input–output network. *Journal of Complex Networks*, 6(2):187–205, 10 2017.

[28] Olivera Kostoska, Viktor Stojkoski, and Ljupco Kocarev. On the Structure of the World Economy: An Absorbing Markov Chain Approach. *Entropy*, 22(4), 2020.

[29] P. Papadopoulos. WP8: Connecting the Dots: Economic Networks as Property Graphs. *Open Risk White Papers*, 2019. Online Link.

[30] Jonas Bunsen and Matthias Finkbeiner. An Introductory Review of Input-Output Analysis in Sustainability Sciences Including Potential Implications of Aggregation. *Sustainability*, 15(1), 2023.