Skatteforsk
Centre for Tax Research

Working Paper No. 13

A Comparison between Traditional and Knowledge Input Output Tables

July 2024

Ronald B. Davies
University College Dublin

Guohao Yang
University College Dublin

Skatteforsk's Working Paper Series is financed through support from The Research Council of Norway (Grant number 341289). The views expressed in this Working Paper are those of the author(s) and do not necessarily represent the views of Skatteforsk - Centre for Tax Research.

School of Economics and Business

Norwegian University of Life Science

A Comparison between Traditional and Knowledge Input Output Tables*

Ronald B. Davies† Guohao Yang‡

February 6, 2024

Abstract

Input-Output tables have become the workhorse data structure when considering global supply chains since, by definition, they measure how production in one country-sector is linked to that in another via trade in intermediate inputs. What traditional input-output tables miss, however, is the role of knowledge. In this paper, we use the well-known World Input-Output Database and construct a corresponding "Knowledge" input-output dataset based on patenting information from PATSTAT. Thus, for 44 countries over 20 sectors for 15 years, we are able to compare linkages via intermediate goods to those created by patent citations. We illustrate that these two networks share strong similarities in terms of the size of activity and linkages yet important differences – most notably the interaction between Asia and the rest of the world – are found. We then conclude by illustrating the link between patents and a measure of intangible capital that can be constructed from the World Input-Output Database, a relation that is a function of tax haven status. This then suggests that there is a likely direct connection between the data found in the production data and those derived from the patent data.

JEL classification: F12; O33; O34; R12.

Keywords: Input Output Table; Innovation; Networks; Centrality; Tax Havens.

*This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No. 101061123. Financial support from The Research Council of Norway (Grant 325720) is gratefully acknowledged. All errors are our own.

†Corresponding author. University College Dublin, School of Economics. Email: ronald.davies@ucd.ie

‡University College Dublin. Email: guohao.yang@ucd.ie

1 Introduction

The importance of global supply chains in the modern economy is not just of academic interest but has become increasingly recognized as a critical feature of the economic system by both policy makers and the general public. Linkages between industries – both across sectors and across countries – have become increasingly visible, not least in part because of the observed consequences when they become disrupted due to pandemics, wars, and blocked shipping routes.[1] This calls for the need for data to describe global supply chains, data that includes both trade in goods and in knowledge. This paper provides a step in this direction by constructing and then comparing input-output tables for trade in intermediate goods and patent citations.

Research on global supply chains has been very active, with Baldwin (2016) providing an excellent overview. These studies range across many topics, including firm-level performance (e.g. Chor et al. (2021), Halpern et al. (2015), Bernard et al. (2012), and Kasahara and Rodrigue (2008)), their role in developing economies (e.g. (Kugler and Verhoogen, 2009, 2011) and Goldberg et al. (2010)), their resilience to shocks (e.g. Bonadio, Huo and Levchenko (2021), Ahmed et al. (2016), and Vandenbussche et al. (2022)), labour market effects (e.g. Kasahara et al. (2016), Parteka and Wolszczak-Derlacz (2018), Lu et al. (2019), and Bonadio, Huo, Levchenko and Pandalai-Nayar (2021)), and the interaction between them and policy (e.g. Amiti and Konings (2007), Bas and Strauss-Kahn (2015), Blanchard et al. (2016), and Brandt et al. (2017)). Underlying the large bulk of this literature is the use of input-output tables, i.e. datasets that map out the links between countries and industries via trade in intermediates. By purchasing inputs from other country-sectors, a given dyad draws from and builds upon them. Likewise, by selling inputs to others, it acts as a upstream provider of value-added. By measuring these linkages, databases such as the World Input-Output

[1]For example, see the recent work on the trade impacts of Covid (e.g. Hayakawa and Mukunoki (2021) and Coquidé et al. (2022)), the Ukraine war (Steinbach (2023) and Orhan (2022)), and the blocking of the Suez canal by the ship Ever Given (Wu and Dong (2022) and de Bodt et al. (2023)).

Database (WIOD) have become the workhorses for the study of global supply chains.[2]

Trade in intermediate physical products, however, is just one part of the network of global production. As evidenced by the rise in services trade over the past forty years, there has undoubtedly been a concurrent rise in trade in intangible intermediates.[3] Measuring services trade is often difficult because it does not go through traditional customs channels. Furthermore, not all services trade is actually paid for and therefore need not show up in national accounts. This is most clearly found in the production of knowledge, something commonly proxied by patenting activity.

When filing a patent application, the applicant must cite the prior art, that is, the innovations already in existence in order to demonstrate the novelty of their innovation. Thus, when one patent cites another (a backward citation) this is an implicit recognition that the first draws and builds on knowledge embodied in the second. This forward citation by the second innovation also indicates its input into further knowledge creation. Thus, citations are links between innovations – and the country-sectors creating them – in the same spirit that trade in intermediates are. Nevertheless, despite the recognized importance of knowledge in production, these unpaid inputs are missing from physical input-output (PIO) tables such as the WIOD.

In this paper, we take steps to bridge the gap between trade in intermediates (embodied in the PIO) and trade in knowledge by constructing a "knowledge input-output table" (KIO) that mirrors the WIOD. We are able to do so for 44 countries for twenty sectors for the years 2000 to 2014 using data from PATSTAT.[4] We then compare the PIO and KIO by recognizing that each can be viewed as a network with country-sectors as individual nodes and sales of intermediates or patent citations acting as edges (links) between them. In the terminology of graph theory, we treat the PIO and KIO as separate weighted digraphs (directional graphs). We then compare these two networks in terms of node size, edge weights

[2]The WIOD can be found at `https://www.rug.nl/ggdc/valuechain/wiod/?lang=en`,
[3]See Francois and Hoekman (2010) and Heuser et al. (2017) for discussion on services trade.
[4]PATSTAT is found at `https://www.epo.org/en/searching-for-patents/business/patstat`.

(the amount of intermediates trade/citations between one country-sector and another), and various centrality measures that discuss the importance of nodes.

This reveals some strong similarities – and notable differences – between the PIO and KIO. First, in large part, there is a strong correlation between the size of nodes (either the value of total sales or the number of patents) with large countries unsurprisingly dominating rankings of node size. The statistical correlation, however, is somewhat sensitive to the fact that economic and innovation activity is skewed by these large outliers. When correcting for this, we find a strong positive correlation between node size and edgeweights across the PIO and KIO. This does not, however, mean that there are no differences. In particular, compared to the PIO where Asian countries such as China dominate, the KIO is dominated more by Western nations. Overall, it seems that Asian countries are somewhat isolated in terms of knowledge creation. In a similar fashion, some nodes that feature heavily in intermediates trade (such as construction) tend to be smaller in knowledge trade.

A second feature that this exercise highlights is the not only is trade and innovation concentrated, but that for many country-sectors, a sizable amount of its inputs and backward citations come from itself (a loop in graph theory). Further, trade in intermediates is far more national (77% of inputs come from the same country) than trade in knowledge (where only 33% of citations are local). Conversely, citations are far more concentrated within the same sector – regardless of location – than is trade in intermediates. This suggests that international borders and other geographic factors likely matter more for physical trade than intangible trade.

A third lesson that comparing the size of nodes and edges teaches is the role of tax havens in the KIO where they rank much higher than in the PIO. This may be suggestive of intellectual property (e.g. patents) being located in tax havens as a means of profit shifting.[5]

Beyond the size of economic/innovation activity and the amount of trade/citing taking

[5]As Tørsløv et al. (2022) discuss, multinationals shift a significant amount of income to tax havens. The latest estimates of profit shifting indicate that it likely exceeds 1 trillion US dollars per year (Annette Alstadsæter and Zucman (2024)). See `https://atlas-offshore.world/` for useful resources on profit shifting.

place, one can use tools from graph theory to measure “centrality”, that is, a statistic suggesting the importance of a given node in terms of its place in the network and the size of its links to other nodes. We specifically calculate five measures of centrality: Degree (measuring direct links), Katz (capturing direct and indirect links), PageRank (also computing direct and indirect links), Closeness (evaluating how “easy” it is to reach another node), and Betweenness (indicating a nodes use as a ”bridge” between others). Further, Degree, Katz, and Closeness centrality can be measured as inbound (purchases of intermediates/backward citations) or outward (sales of intermediates/forward citations). Each of these statistics explores one facet of a node’s role in the network and we therefore begin by comparing a node’s score across these within the PIO and the KIO. We find fairly strong positive correlations across the Degree, Katz, and PageRank measures, i.e. when a country-sector is central under one measure of its links it tends to be central according to another. The Closeness and Betweenness measures, however, provide very different pictures, suggesting that the aspects of networks they are meant to capture may not be especially useful in the context of global supply chains.

We then compare centrality across the PIO and KIO networks. Here, in terms of ranking of countries and sectors, we find lessons comparable to those noted above, namely that Asian countries tend to rank lower in centrality and tax havens rank higher. Unlike the size of nodes and edges, the correlation in the centrality measures is much weaker (albeit generally positive and marginally significant).

Thus, while there are similarities in the network of intermediates trade and that for knowledge, meaningful differences are found. This then raises the question of how one can link the two networks. Put differently, how does knowledge act as an input into the production process? We make some inroads into this queston by using the method of Chen et al. (2021) and Karabarbounis and Neiman (2018) to construct a measure of intangible capital’s value-added from the WIOD.[6] We then regress this on the data underlying the

[6]As a preview, this is roughly a residual from a production function accounting for capital and labour using national accounts data at the country-sector-year level.

KIO: the stock of patents, the flow of new patents, and citations. This exercise shows that a large share of intangible value-added can be explained by the easily observed patenting data, suggesting that patenting activity may be a useful proxy for intangible capital when the necessary data for constructing intangible capital is unavailable. Further, our estimates suggest that tax havens tend to have a higher value-added from intangible capital, again suggesting that they may be hosting intellectual property as a means of profit shifting to those low-tax jurisdictions. Thus, by combining the differences between the PIO and KIO via intangible goods and intellectual property, this provides insight into the fragmentation of a global value chain that includes both physical production and innovation.

Ours is not the only study to compare trade in intermediates with citations. Both Liu and Ma (2021) and Ayerst et al. (2023) use an input-output table and citation data to construct networks for trade in goods and knowledge, albeit less for a detailed comparison than to measure the technological content of trade to measure how trade in intermediates diffuses knowledge. Additionally, their data, both in terms of source and level of aggregation, differs somewhat from ours. In particular, Ayerst et al. (2023) relies heavily on US data for their analysis whereas ours is global from the outset. That said, their findings regarding the ranking of countries in terms of sales, patenting, and citations strongly mirrors ours. Nevertheless, there are some rather important differences. First, both suggest that there is at best a weak correlation in the size of nodes across their PIO and KIO. This is something we also find when using levels of activity, but when using the inverse hyperbolic sine to deal with skewness, zeros, and outliers, we find a strong correlation between the node characteristics in the PIO and KIO. Thus, their lack of a correlation may be driven by such data issues. A second difference is in our examination of centrality. Liu and Ma (2021) uses Eigenvector centrality (which they refer to as innovation centrality). This measure, however as detailed in Njotto (2018) among others, Eigenvector centrality performs poorly in weighted digraphs which are not strongly connected (i.e. many nodes lack direct edges). Ayerst et al. (2023), meanwhile, uses both Degree centrality and Authority centrality (a measure encompassed as part of PageRank centrality) for the US.[7] In any case, both find that there is no clear pattern in the centrality measures for goods trade and citations. We, on the other hand, do find some correlation with the strength depending on the centrality measure. This finding is again aided by dealing with the skewness and outliers in the data. Third, although Liu and Ma (2021) uses taxation as feature in their innovation decision, they do not consider the role of tax havens and profit shifting in the structure of the KIO. Finally, neither attempts to link patent data and intangible capital in the manner that we do. Thus, our paper complements their work by adding further exploration of the data, providing alternative centrality measures, and consideration of potential factors driving the network differences.

Further, our paper contributes the the very large literatures considering either global supply chains in goods trade or the diffusion of knowledge. The work focusing on the structure of global supply chains includes Schott et al. (2017), Koopman et al. (2014), Johnson and Noguera (2012), Hummels et al. (2001)and many others. This then links to the multitude of papers describing the drivers of trade in intermediates and the impacts of such trade on firms, workers, and countries, a small selection of which was cited above. Likewise, our analysis of citation flows links our work to the other studies examining the pattern of patent citations as well as its determinants. Here too there is a wealth of literature, including a significant amount of work outside of economics. We suggest Breschi and Lissoni (2005), MacGarvie (2005), Alcácer and Gittelman (2006), Carlo Giglio and Palmieri (2021), or Jie Cai et al. (2022) as entry points into the broader literature on the topic.

The rest of the paper proceeds as follows. In Section 2 we discuss some important differences when thinking about linkages in knowledge production and linkages via trade in tangible intermediates for goods production. In particular, this considers the role of time and non-exclusivity. Section 3 discusses the data we use and the construction of the network measures we use to compare the PIO and KIO networks. Section 4 provides a discussion of the two networks when collapsing the entire dataset from 2000-2014. Section 5 then considers

[7] In unreported discussion, we also used Authority centrality which yielded similar patterns as what we find with PageRank.

how the PIO data can be used to construct measures of intangible capital and how this then correlates with the KIO data. This is done to indicate the extent to which easily obtainable patent data can be used to proxy for harder-to-obtain measures of intangible capital. Finally, Section 6 concludes.

2 Theoretical Considerations

Since both the PIO and KIO tables represent weighted digraphs, the parallels between them are on the surface, rather straightforward. That said, there are important differences in the nature of intermediate goods trade and patent citations that must be considered in depth before making any comparisons. Here, we go through some of these in turn.

Before proceeding further, it is useful to define some key terms now so as to avoid confusion. First, although we tend to use the term sector, other studies refer to these as industries. We prefer the first because, given the level of aggregation in the WIOD data, some sectors are comprised of multiple industries. Nevertheless, we consider the two terms interchangeable. Second, as discussed in more detail below, we use the word patent to indicate a patent family which may or may not contain a granted patent.

2.1 Rival Goods versus Non-rival Knowledge

When a country-sector produces and sells an intermediate input, this is a rivalrous transaction, i.e. it can sell the same unit of production to one and only one customer. Further, when one customer uses that unit, this eliminates the ability for any other customer to use that same unit for their own production.[8]

This is not true for knowledge. Indeed, one patent can generate multiple forward citations (or none at all).[9] This follows the idea of knowledge being a "joint input" as characterized by Markusen (1984). In his formulation of horizontal multinational firms (MNEs), the entire

[8]The same applies to final consumers.

[9]Indeed, in our data, roughly half of patents are never cited.

basis for such firms arising is that, by using the same knowledge in multiple locations without additional cost, MNEs have an efficiency gain over repeated, national firms each of which much incur the same cost.

2.2 Time

Goods production is relatively easily placed in time – when a unit of production takes place in year t, then it is either sold in year t or placed in inventory. Indeed, inventory is the sole way in which there is a dynamic aspect to most PIO tables, the WIOD included. A new patent in year t, however, can be cited in year t but can also be cited in the future. Thus, a new patent in t contributes to the stock of patents in t and beyond, with this stock of patents being cited both presently and in subsequent years.

Thus, whereas once a good is sold in year t it cannot (except via inventory changes) contribute to the stock or sales in year $t+1$, this is not true for patents. Instead, the current stock of patents should be measured as the (potentially weighted) sum of current new patents and past ones so that even if a patent is cited, this itself does not diminish its capacity to contribute to citations in the future. An implication of this is that in our analysis where inventories are fairly minor, the nodes in the PIO digraph will be fairly independent across years. In contrast, the KIO node sizes will be highly correlated across time, differing primarily by the flow of new patents in each year.

2.3 Size of Edges, Size of Nodes, and Summations

For goods production, in each year, total production must equal total sales plus net inventory changes. Likewise, total sales must equal the sum of inputs (including capital and labour) plus value-added (a feature exploited when constructing value-added from intermediates in Section 5). Thus, in the full PIO table, not just the portion we use to construct the weighted digraph, the rows (costs) must equal the columns (output). Thus, the size of a node (total sales) cannot be smaller than either the sum of the size of its outbound edges (total intermediate sales) nor the sum of the value of its inbound edges (total intermediate purchases).

These restrictions do not hold for the KIO. This is for five reasons. First, whereas the PIO tables provide detail on non-intermediate goods inputs (capital and value-added, e.g. labour), there is no such parallel obtainable for patenting. While one can imagine collecting information on lab costs (capital) and research scientists (labour), these data are not available across countries, sectors, and time in sufficient detail to do so. Second, PATSTAT only provides information on patent usage by other patent creators. This misses the licensing of technologies (i.e. their use in non-patent generating ways) and final consumers (those who consume the embodied knowledge simply for its own sake).

Third, the size of a node is the stock of patents whereas the size of an edge is the number of citations. Thus, these are not comparable units of measurement. Finally, as noted above, patents can be cited multiple times, so that a single patent can generate more than one citation or no citations at all. This differs from goods production where a single dollar of production must generate one dollar of use, i.e. a dollar of sales to other industries, final consumers, or inventory.

This difference then implies that some of the summation restrictions used in consideration of PIO tables cannot be used for KIO tables. This is why we restrict ourselves to the network analysis of the weighted digraph arising from the intermediate input part of the PIO table, something that is more along the lines of what is obtainable for patents.

A further item – and one we highlight in our analysis of the data – is how connected networks are, i.e. how many nodes have direct edges between them. Sparsity is a feature of both the PIO and KIO because there are many nodes where there are no sales/citations between them. This must be considered when constructing some centrality measures such as Eigenvector centrality and, as such, what may be an appropriate statistic in one network may not be so for the other.

2.4 What is Innovation?

In our discussion, we tend to equate innovation with patents. In truth, however, this is not true. In particular, not all innovations are patented because this is both a costly, uncertain process (see Davies et al. (2020)) but also because then this publicly reveals the knowledge contained in the innovation to one's competitors. As a result, not all innovations will be patented, a decision discussed in detail by Hall et al. (2012). While other proxies for innovation, such as spending on research and development exist, these are also somewhat unsatisfactory because there is not a clear, non-stochastic mapping between spending on innovation and generating useful – much-less profitable – ideas. Furthermore, innovation cost data would not indicate the spillover nature via citations needed for constructing a KIO. In contrast, trade in goods does not suffer quite the same issues (although value measures are certainly manipulable for trade and tax reasons as discussed by Davies et al. (2018)). Nevertheless, patent data is arguably the most widely used measure of innovation, particularly in a cross-country setting where work includes Blundell et al. (1995), Stiebale (2016), and much of the other work cited above.

3 Data

In constructing our PIO and KIO tables, we are creating weighted digraphs where a node is a given country-sector. In addition to measuring the size of nodes, we must create a weighted, directional adjacency matrix across 44 countries for 20 sectors. Based on the WIOD data, we do so for each year from 2000 to 2014. While it is feasible to construct these by country-sectory-*year*, this results in a very sparse adjacency matrix for the KIO that is volitile year to year. Therefore, we aggregate across time except as noted otherwise. Here, we detail the construction of these as well as the different node/network descriptors we use for our comparison.

3.1 PIO Data

Our data trade in intermediates comes from the 2016 version of the World Input-Output Database. For details on the WIOD's construction, see Timmer et al. (2015, 2016). In the WIOD data, there is information on 56 sectors, including primary, manufacturing, and some services industries. Likewise, it provides data on 44 countries, one of which is a catch-all "Rest of the World" (ROW) group.

For our purposes, in a given year t, the node size is the total sales by a given country-sector (cs). As discussed above, this includes the sale of intermediates, sales of final goods, and net changes to inventories. Edge weight, that is, the strength of the link between nodes, is measured by the share of all intermediate input sales in year t sold by node cs to node $c's'$. This is then the outward edge weight for cs and the inward edge weight for $c's'$. These are then aggregated across the period 2000-2014. Note that the sum of all edges is one by definition.

3.2 KIO Data

Our patent data comes from the Autumn 2022 version of PATSTAT. This data set contains several key pieces of information. First, for each patent (again, a term by which we mean both granted patents and applications), it links to a "family" identifier that is unique to a particular innovation. Since a given innovation can lead to multiple patents, either across patent offices and/or to the same office (such as when an application is denied, revised, and then resubmitted), this is crucial to eliminating double-counting.

For each patent within a family, we use three pieces of information from PATSTAT. First, to allocate the family to a year, we use the earliest filing date within the family. By using the earliest filing date, this arguably pins the patent to the point in time closest to when its research was carried out.

Second, to allocate a patent to a given country, we follow standard practice and use fractional apportionment. This is a two-step process. First, for all of the assignees (owners) listed on the patents in the family, we assign a share of the patent to each country based on the share of assignees. If no assignees are listed, we do this using inventors.[10] If no inventors are listed, then the family is dropped from the sample.

Third, we make use of PATSTAT's matching of the technologies contained within a patent it to different industries (giving each industry a share where the sum across industries is one). Note that PATSTAT only links patents to twenty manufacturing industries which are a subset of the those found in WIOD. Thus, for both the PIO and KIO, we are utilizing data on twenty industries for 44 countries (including ROW).

With these in hand, we can then allocate each patent i across country-sectors where, when α_{ic} is the share allocated to a country and α_{is} is the share given by PATSTAT to a given sector, $\alpha_{ic}\alpha_{is}$ of a given patent is allocated to country-sector-year cst where the year t is again determined by the earliest filing date.

Where $I(t)$ is the set of patents filed in year t, $P_{cst} = \sum_{i \in I(t)} \alpha_{ic}\alpha_{is}$ is the (fractional) number of new patents in a country-sector in a given year, that is, the flow of new knowledge generated. To construct the size of a node in the KIO, this is aggregated across time. In Section 5, where we need annual data, we instead construct the stock of patents as of year t as $Stock_{cst} = \sum_{j=0}^{10} .9^{t-j} P_{cst-j}$. In words, this is sum of the current flow of new patents and the discounted sum of new patents over the prior decade, applying a depreciation rate of 10 percent. Note that in order to match the start date of the PIO (which is 2000 and limited by the WIOD data), to construct the stock of patents in 2000, we use data on the flow of patents from 1990 to 2000. Again, to be clear, we use the sum of the flow of new patents from 2000 to 2014 for node size, not the sum of the stock (as that would double count patents).

The final piece of information we use from PATSTAT is the data it provides detailing which patents cite which others. Again, we operate on the family-level so that there is at most one citation from one family to another, with that citation spread across country-

[10] This differs from Ayerst et al. (2023) who use inventors in the fractional apportionment. However, given the small number of cross-border applications, this does very little to the totals used in the analysis.

sectors following the fractional apportionment method above. Note that some families cite other patents from within the same family; these are dropped from the sample.

As noted above, the issue of time requires consideration when working with patent data. Whereas the WIOD and other input-output tables treat sales from one country-sector to another as implicitly contained within a year (excepting how these may affect net inventories), patent data tends to highlight the fact that the cited and citing patents come from different points in time. Further, an especially when operating at the family level, one can find situations in which citing patents pre-date cited patents. This can happen for two reasons. First, there may be issues in the underlying PATSTAT data (e.g. variation across patent offices in registering filing dates and simple data entry errors). The second comes from using patent families. To see this, suppose that the first patent in family A does not cite something from family B, but a subsequent member of A does. Then, if the first filing in A predates that in B, by using the earliest filing date to allocate a family to a single point in time this can look as if A is citing research from the future.

With these in mind, to construct our edge weights in the KIO, for each year t, we sum the (fractional) number of citations made by the flow of one country-sector (P_{cst}) from those patents in the current stock ($Stock_{c's't}$ of another country-sector. Thus, this captures the number of times new patents in one country-sector cites the most-recent eleven years worth of new patents from another. This eliminates any citing of future patents.[11] After aggregating over time, this citation count is the forward citation count for cs and the backward citation count for $c's\prime$ in t and is the outbound edge weight from cs to $c's^n\prime$. Note that although forward citations equal backward citations in a single direction (i.e. from cs to $c's'$) it is *not* necessarily the same as from $c's'$ to cs. As with the PIO, we use the share of citations so that the edge weights sum to one to match the PIO. Thus, we ae constructing the closest parallel that we can to the current sales from one country-sector to another.

Thus, for the PIO and the KIO, the size of a node is a measure of overall activity (the

[11]This is because, in the example above, the patent in family B would not be in any stocks when A is a new patent.

total sales or the total stock of patents) and the direction edge weights indicate the strength of a connection from one node to another (either the share of sales of intermediates or the share of current new citations). Finally, although it may seem obvious to those already familiar with the source data, each of these two networks includes loops, that is, where a node sells to or cites itself.

3.3 Network Measures

As our goal is to compare the PIO and KIO, an appropriate metric must be used. Here, we discuss these, including their caveats and limitations.

The first way to compare the networks is by node size, i.e. to ask whether country-sectors with larger total sales also tend to generate more patents. Likewise, edge weights, the sales of intermediates or citations, can be compared. While such comparisons may be useful, they have clear limitations. First, there is the obvious fact that sales are measured in monetary values (millions of constant US dollars to be precise) whereas citations are a count variable. Second, total sales includes changes to inventories and sales to final consumers. The first of these has no parallel in patenting and, in the patent data, we do not have any information on those who may draw benefit from a patent without citing it. Third, there are the conceptual matters noted above including non-rivalrous knowledge which impacts edge weights in the KIO but not the PIO. Thus, comparisons between total sales and patents (or between total sales and total citations), must be considered in light of these issues.

The second way to compare nodes within a network – and then across them – is to use a measure of their importance in the network, that is, their centrality.[12] Graph theory has produced a number of different measures of centrality which complement one another. Indeed, centrality has led the international trade literature to construct notions such as upstreamness (building from Katz centrality) as defined by Antràs et al. (2012). While we cannot follow thier lead because there is no data on final knowledge users to parallel final goods

[12] There are many excellent resources available on graph theory. For economists, we particularly recommend Jackson (2008) and Sargent and Stachurski (2008).

consumers, we can nevertheless construct several other centrality measures. In particular, for each country-sector, we construct the node's Degree, Katz, Closeness, Betweenness, and PageRank measures.[13] Further, as the edge weights vary by the direction between a given pair of nodes for Degree, Katz, and Closeness, these have in and out values. As these centrality measures may be unfamiliar, here we provide specific detail on the construction of each and what aspect of the KIO or PIO they capture.

Degree Centrality

To start, we calculate a country-industry's "Degree centrality" which considers direct links only. This can be unweighted (so edge weights are equal) or weighted (as we do since our goal is to measure strength of connections, not just their existance). For in Degree centrality, we sum the weights of the edges pointing into a given node, i.e. its purchases of intermediates or its backward citations. Out Degree centrality is the converse of this, i.e. the sum of sales of intermediates or forward citations. For the PIO, these then give a measure of how important a given country-sector is as a supplier/buyer of intermediates. Similarly, the Degree centrality measures in the KIO suggest how important a given node is for using existing knowledge or providing it to others.

Katz Centrality

Whereas degree centrality captures the number of direct connections, an important feature of both the PIO and KIO we wish to consider are the indirect connections. Indeed, the central idea of global supply or knowledge *chains* is that stages of production or innovation feed into another that itself feeds into a further stage. In graph theory, "Katz centrality" (sometimes referred to as alpha centrality) is designed to get at this idea by not just considering directly connected nodes but also those indirectly connected. In this, indirect connections are discounted by an exogenously chosen factor α (hence the alternative name for this measure).[14]

[13] In unreported results, we also constructed Eigenvector, Authority, and Hub measures.
[14] See Katz (1953) and Hubbell (1965).

With this chosen, Katz centrality for node i is expressed as:

$$C(i) = \alpha \sum_{j=1}^{N} A_{ij} C(j) + \beta \tag{1}$$

where A_{ij} is the corresponding cell from the adjacency matrix (equalling the edge weight from i to j, which is 1 in an unweighted graph). The constant β is typically set to 1 and meant to representing the initial centrality; $\alpha = 0.85$ in many settings. We follow these conventions in our analysis. Thus, this measure is higher for nodes that themselves have more, stronger connections and that are connected to other nodes that have more, stronger connections.

PageRank Centrality

The benefit of Katz centrality over degree centrality is that a given node's "importance" depends on the importance of those it connects to and so on. A similar notion underlies the PageRank centrality measure, an algorithm developed by Google to measure the importance of webpages.[15] In contrast to Katz centrality, PageRank is a recursive algorithm (which builds from from the Hyperlink-Induced Topic Search (HITS) algorithm, developed by Kleinberg (1999)). As such, we relegate the details on its construction to B but provide an intuitive understanding here.

The underlying idea of PageRank is to ask how likely it is that an internet user would randomly end up on a given website (node). Nodes that have a large number of inbound edges (known as authorities) are more likely to be reached and are therefore more central to the network.[16] Further, when those inbound edges come from other authorities, this increases the chance that, starting from a random node and following random edges, one

[15] Eigenvector centrality, used by Liu and Ma (2021) is another recursive approach to measuring importance. However, it does not work well in sparse adjacency matrices such as the PIO and KIO, nor does it perform especially well in directed graphs. See Njotto (2018) for discussion. Note that PageRank incorporates aspects of both Authority centrality (used by Ayerst et al. (2023)) and Hub centrality.

[16] Hubs, in contrast, are nodes with many outbound edges.

is more likely to end up at the node in question. This then makes the node in question "important" because it is likely to be stumbled across. Taking the concept to the PIO, it asks how likely it is that a country-sector will be used as an input – directly or indirectly – in another. PageRank in the KIO works similarly. Note that unlike Degree or Katz centrality, this measure does not have a direction.

Closeness and Betweenness

While being directly connected to many other nodes can indicate the importance of a given node, this is not the only way of capturing influence. For example, suppose that, starting from a given node, many others can be reached without having to traverse many transitory nodes, i.e. it is "close" to many other nodes by virtue of its short paths to them.

This then motivates measuring a node's "closeness". In a connected graph, closeness centrality (or closeness) indicates how close a node is to all other nodes in the network.[17] with this in mind, between two nodes i and j, let "distance" $d(j,i)$ be the number of edges one must traverse to get from i to j along the shortest path between them.[18] For the PIO, distance $d(j,i)$ captures how "far" one would have to go to link inputs sold by node i to the production in node j with the KIO distance similarly interpreted.

Then, for node i, we have two closeness measures, one in and one out:

$$C^{in}(i) = \frac{1}{\sum_{j \to i} d(j,i)} \tag{2}$$

and

$$C^{out}(i) = \frac{1}{\sum_{j \leftarrow i} d(i,j)}. \tag{3}$$

[17] A bipartate graph is one made up of two or more subsets of nodes where there is no edge between any of the nodes in the subgraphs. In this case, closeness is measured using the distance only between connected nodes, i.e. where there is *some* path to get from one to the other.

[18] Thus, if i and j are directly connected, $d(j,i) = 1$, if one must go through a third node k to do this, $d(j,i) = 2$, and so forth.

These can – and in our analysis are – weighted by the edge weight of each link in the path. In this, note that i is the end point for each path in the in Closeness measure but the start point for out Closeness. Since our networks are directed, meaning that there can be an edge from A to B but not vice versa, these need not equal each other. For each of these, if distances are small, then the corresponding Closeness figure is larger.

Alternatively, suppose that, for any other two nodes, a given node is often used as a "bridge" in that it serves as a transit node on the shortest path between them. Thus, it is "between" many other node pairs. With this in mind, Betweenness centrality considers i's role in bridging the path between to other nodes s and t. For each s and t, let σ_{st} be the total number of shortest paths between them. Of these, let $\sigma_{st}(v)$ be the number of shortest paths between s and t which pass through i.[19] The betweenness of i is then calculated as

$$g(i) = \sum_{s \neq i \neq t} \frac{\sigma_{st}(i)}{\sigma_{st}}, \tag{4}$$

i.e. the sum of the fraction of shortest paths i lies on. A higher betweenness indicates that i is more useful on bridging the gaps between nodes which are not directly connected, and thus serves as another measure of the node's importance in the graph. Again, this can be weighted by the weights for the various edges in the path. Note that since in Betweenness i is a transit node only, this does not have in and out versions.

While Closeness and Betweenness are common centrality measures, it is not intuitively obvious what aspect of the PIO and KIO they would capture. This is because they are built on the number of edges that must be traversed to get from one node to another. It is not clear why (except perhaps when considering transport costs) the number of stages in a global supply chain would matter. Thus, these measures may not be as useful when describing the PIO and KIO. Indeed, as shown below, we find rather different patterns when using these as compared to the more inuitive measures above..

[19] One can also construct a weighted betweenness by using the sum of weights rather than the number of edges in construction of these.

4 Comparison between PIO and KIO

We now compare the PIO and KIO networks. We begin by comparing the size of nodes and edges and then turn to our various measures of centrality.

4.1 Node Sizes and Edge Weights

We start by comparing the size of nodes (total sales/patents), outbound edge weights (total intermediate sales/total forward citations), and inbound edge weights (total intermediate purchases/total backward citations).

Rankings

To begin with, in Table 2 we aggregate across sectors to focus on the variation across countries in terms of their rank in node size/edge weight for the PIO (left columns) and KIO (right). Perhaps unsurprisingly, the world's dominate economies – namely the US, China, Japan, and Germany – dominate the top positions in both the PIO and KIO rankings. This matches what is found by Liu and Ma (2021) and Ayerst et al. (2023). Despite the similarity between our PIO and KIO, there are also some rather telling differences. For example China holds the number one spot it terms of the PIO, selling more in total and in intermediates and buying more intermediates than any other country. Likewise, it holds the number two spot in terms of patenting. However, when it comes to citations, both in and out, China only ranks sixth. Likewise, although Japan patents more than any other nation, its citation count is lower than that of the US and Germany. Korea, meanwhile ranks eighth in total sales but sixth in trade in intermediates, i.e. it is perhaps more intensively connected to global supply chains than most countries. In the KIO, however, it is fourth in the number of patents but only ninth in forward citations. The opposite is found for Canada, however, whose PIO edge weight ranks are lower than their total sales rank with the reverse found in their KIO rankings. These patterns suggest that while Asian countries may tend to be more linked to global supply chains for production than their western counterparts, the opposite is true when it comes to global knowledge chains.

Another interesting feature Table 2 shows is that some countries rank well in one half of the table but not the other. For example, the UK, India, and Brazil do fairly well in the PIO but is only middling in the KIO. Conversely, Ireland, the Netherlands, Switzerland, and Luxemborgh all rank much higher in the KIO measures than in the PIO. These latter four are notable as all are commonly considered tax havens and have patent boxes, a policy which further reduces the corporate tax rate on income derived from patents (see Davies et al. (2020) for discussion on patent boxes and innovation). This may suggest that such policies may encourage not just patenting, but the development of influential knowledge.[20]

In a similar way, Table 3 focuses on sectors' rankings across the PIO and KIO by aggregating across countries. Here, two patterns again emerge. First, there is some variation across sectors in terms of their connection intensity, that is, how their total sales compares to sales and purchases of intermediates in the PIO (or the number of patents versus citations in the KIO). For example, although the construction sector is the largest in total sales, it is in the middle for sales of intermediates (due to the amount of final goods sales in this sector). Similarly, while Food and Tobacco is very large in total sales, it is much lower in terms of the purchases and sales of intermediates.[21] In contrast, Chemicals is sixth in total sales but second in intermediate sales. This then mirrors what the country data showed.

Second, and also reminiscent of the country data, there is a difference in the ranking of a sector in the PIO and the KIO. For example, despite its dominance in the PIO, construction does not break the top ten in the KIO. Food and tobacco also slides in comparison. Electronics, machinery and, somewhat surprisingly, furniture, however, do well in the KIO rankings in comparison to the PIO. This suggests that there is significant variation in the knowledge intensity across industries, with some patenting often per dollar of sales and others less

[20]Note that this can arise both because of increased innovation activities or the relocation of such activities within multinationals. Schwab and Todtenhaupt (2021) provide evidence consistent with the latter.

[21]The first of these is somewhat due to the exclusion of the primary sectors as they do not appear in the patent data.

frequently.

Beyond providing a first comparison of the supply chains for knowledge and production, these tables illustrate some key features that must be accounted for in further investigation. In particular, Table 2 indicates that country size is going to be an important feature of the level of sales, patenting, and thus citations. Likewise, Table 3 suggests important differences in knowledge intensity. These will feature in our regression analysis below.

Concentration

Although Tables 2 and 3 provide a relative comparison of countries and sectors, these do not themselves indicate the absolute value of the various measures. To provide some insight into this (while not burying the reader with the full dataset), in Table 4, we indicate the share of each of the six measures comprised by the top five and top ten countries (top panel) and sectors (bottom panel). Note that the top five/ten is not quite the same comparison between countries (of which there are 44) and sectors (where there are twenty). Starting with the PIO, we see that the top five countries (11.4% of countries) make up just under 15% of total sales, purchases, and sales of intermediates, i.e. they are slightly over-represented in the absolute numbers. Moving to the top ten countries, we see that roughly one-fifth of countries make up more than three-fourths of the monetary value. Thus, economic activity is highly concentrated in around 25% of the countries. Looking to the sector results in the bottom panel, one-fourth of sectors make up one-half of the activity. Thus, again, economic activity is concentrated in 25% of industries. Ayerst et al. (2023) make a similar observation for the US.

Turning to the PIO, we again see concentration of activity, however, in comparison to the PIO's economic activity, it suggests innovation efforts are far more concentrated. Beginning with the countries, we see that the top five countries make up a staggering 90% of innovation activity. Although this concentration is less pronounced when looking across sectors, the data shows that patenting activity, and the citations in particular, are done in just half of the sectors.

These figures tell us that, as one might have anticipated, activity is concentrated among a handful of major players. What is unexpected, however, is how much greater that concentration is for knowledge creation than production. One potential interpretation of this is that, in practice, global supply chains are comprised not just of trade in intermediates (the PIO) but how that interacts with the creation and spread of knowledge (KIO). If innovation – regardless of industrial application – tends to be done in countries well-suited to such activity, then we might expect a fragmentation of the tangible and intangible stages of the innovation process. Such geographic separation lies at the heart of the earliest models of foreign direct investment such as Helpman (1984) (vertical FDI) and Markusen (1984) (horizontal FDI). This then raises questions over what might drive this fragmentation and concentration. In particular, based on the above observations, one might expect country factors such as high-skill labour, size, and innovation-friendly tax policies to matter.

As an alternative take on concentration, Table 5 shows the percentage of sales of intermediates (in the PIO) and citations (for the KIO) across four groups: same country and sector, same country but a different sector, different country but same sector, and different country and sector. Here two notable features arise. First, 35.1% of intermediate sales and purchases happen within the same country-sector (an edge that in graph theory is known as a “loop” in that it begins and ends with the same node). Likewise, 27.3% of citations are made to patents within the same country-sector. Thus, loops are a significant feature of the PIO and KIO networks, something that will be seen graphically below.

In addition, Table 5 shows a difference between the PIO and KIO when either country *or* sector differs. While 42.7% of intermediate sales are to other sectors in the same country, only 9.9% are to other countries but the same sector. The reverse is found in the KIO, with over half of citations linking the same sector across international lines. Finally, both have very similar shares when both country and sector differ. This pattern suggests that geography matters more than technology for the PIO, since 77.8% of sales are within country but only 45% within ector, with the opposite in the KIO where only 33% of citations are within country but 81.8% are within sector. This speaks to two potential underlying factors helping to drive these global supply chains: trade barriers matter more for physical products than intangibles and innovation spreads best when it has specific industrial applications. Recognizing these patterns shows the value of our more in-depth exploration of the data as a complement to the existing work of e.g. Liu and Ma (2021) and Ayerst et al. (2023).

Correlations

As our last exploration about the size of nodes and the edge weights, we conclude by making a series of bilateral comparisons between our two measures of node size (total sales and patents), our two outbound edge weights (sales and forward citations), and our two inbound edge weights (purchases and backward citations).

We start by making a simple scatter plot for each pair which can be found in Figures 1 (for node size), 2 (for outbound edge weight), and 3 (for inbound edge weight). In the left panel of each, we see that, although there may be some positive correlation, this seems to be masked by a skewed distribution and some large outliers. In the right panel, rather than using the levels, we use the inverse hyperbolic sine (IHS) of the value. This transformation has some of the smoothing properties of taking the log, but does not exclude zero values (see Bellemare and Wichman (2020) for discussion of applications of the IHS in economics).[22] Doing so indicates a clear positive correlation between the node sizes and edgeweights across the PIO and KIO.

To further this analysis, in Table 6, we regress the node size or edge weight from the PIO on its KIO counterpart. Because of the insights from the ranking exercise above, we control for both country and sector fixed effects. We also cluster the errors at both the country

[22]In our data, after aggregating across time, there are relatively few zeros. This is not, however, true when using annual data. While we could use logs here without losing many nodes or edges, we nevertheless use the IHS here to smooth comparison between these data and other, less aggregated data. Finally, note that since zeros can occur in either the PIO or the KIO, in the regression analysis, PPML is not appropriate as we can have zeros in the continuous explanatory variables as well as the dependent variable.

and sector level. In the first three columns, akin to the right panels of Figures 1 to 3, we use levels. This results in at best weakly significant correlations, with the strongest between purchases of intermediates and backward citations.

On the left, we use the IHS transformation. Now we find very robust positive correlations in each case. In column (4), this suggests that country-sectors that have 1% higher patents tend to have sales that are .316% higher. Likewise, the estimates in column (5) those that have 1% more forward citations tend to have sales of intermediates that are .214% higher. Finally, column (6) suggests that a country-sector that cites 1% more often tends to have purchases of intermediates that are .219% higher. Also, note that the fit of the regression is about 50% better when using the IHS transformation, a result that arises from the IHS's reduction in the influence of outliers.

The results in Table 6, however, are still aggregated within a given node, meaning that the edge weights are the totals across all other nodes. Since one of the advantages of using networks is to consider the bilateral nature of the data, we conclude by aggregating just across years at the country-sector pair level. Note that here, since between two nodes A and B the outedge for A is the inedge for B, we no longer need to distinguish between sales and purchases or between forward and backward citations.

In Figure 4, the top two graphs show the scatter plot using the level of intermediate sales and forward citations, for all node combinations on the right and when omitting loops on the left. As with the prior graphs, this indicates a skewed distribution and outliers with no clear correlation. This is the same pattern as Liu and Ma (2021) and Ayerst et al. (2023) find that leads them to suggest that there is little overlap between what transpires in the PIO and the KIO. However, in the bottom two panels, we repeat this exercise but use the IHS transformation. As before, this brings out a clear positive correlation, one which is somewhat stronger when including loops. This is perhaps not surprising given the results of Table 5. Thus, this points to the importance of dealing with outliers and large values when considering the network as such out-sized nodes can act as "black holes" which obscure smaller but nevertheless important patterns in the network.

We conclude this comparison by again turning to regression analysis, now controlling for country and sector fixed effects for both the outbound and inbound node. These results are found in Table 7. Mirroring Figure 4, the first two columns use levels, one with loops and one without, while the latter two columns use the IHS. Now we find a positive, significant correlation for both levels and IHS, although the latter again provides a better fit to the data. The other thing to note from this is that the correlation is much higher when including loops, a result again not surprising given the above discussion.

4.2 Network Structure

Graphical Representations

Before moving to a more data-oriented discussion of the PIO and KIO network structure, it is useful to begin with some graphical representations of each of the networks. A factor when producing such graphical representations is that they rapidly become illegible. Plotting the full network means plotting 880 nodes and, when including loops, up to 774,400 edges resulting in a dense figure in which nothing is visible. We therefore restrict the number of nodes in the following graphs in an attempt to produce something more useful. In these graphs, the size of a node is as above: either total sales (PIO) or total patents generated during the sample period (KIO). Edge weights are likewise share of all intermediate sales (PIO) or share of all forward citations (KIO). When plotting these, the thickness indicates the relative edge weight and the arrow indicates the direction of the edge.[23]

We begin with Figure 5 where we aggregate across sectors. The top panel includes loops whereas the bottom omits them, something done again to reduce the number of edges. Figure 6 repeats this for the KIO. Comparing the two, the standout difference is that the KIO network has far fewer edges than the PIO and for those that are there, they are often

[23]Note that here, edges are represented as straight lines, meaning that in and out edges overlap. While this obscures that variation, we do so in order to increase readability.

weaker. This then indicates that the global supply chain for intermediates is generally more connected than that for knowledge. In other words, and in line with Table 4, the KIO is more concentrated.

Given the difficulty in reading these figures, from this point we omit loops and reduce the number of nodes by focusing on a select group of countries that are driven by the above discussion. For those countries not given individual nodes, they are rolled into "ROW2", a node that includes the ROW from above as well. Again, this comparison shows the relative connectedness of the PIO. Further, it again highlights the differing importance of Asian countries in the PIO as compared to the KIO.

In Figure 8, we plot the networks aggregating not across countries, but across sectors (again omitting loops). Note that nodes are labelled according to the ISIC Rev. 2 code, with the A providing the corresponding description. As with the prior graphs, the key lesson to be learned here is that connections are stronger with the PIO than the KIO.

Although we admit that the figures can often be difficult to read, in Figure 7 we provide the representation of the network where nodes are country-sector, using the US, China, Germany, and Japan while combining all others into ROW2. Likewise, we only include the main fifteen sectors, leaving us with 75 nodes. Even this reduction, however, leaves us with figures that provide little use. As one last attempt, we therefore reduce the nodes further by considering one country at a time: the US (Figure 9), China (10), Germany (11), Japan (12), Korea (13), France (14), the UK (15), and Ireland (16). Although intriguing, because these nations are all very well connected both internally and to the rest of the world, here too, the graphs are less informative than might be desired. We therefore turn to more numerical analysis of the network structure to derive insights.

Centrality

We begin this process by considering the five centrality measures noted above – Degree centrality (direct connections), Katz centrality (direct and indirect connections), PankRank (direct and indirect connections), Closeness (proximity of nodes to other nodes), and Betweenness (use as a bridging node for paths between other nodes) – for the 880 nodes in the network (i.e. 44 countries and 20 sectors). Based on the above discussion of the edge weights, for the PIO we use the inverse hyperbolic sine of intermediate sales as the weight for a given nodes out edge.[24] Similarly, we use the inverse hyperbolic sine of forward citations when calculating the centrality measures of the KIO.[25] In addition, for those measures where inbound and outbound edges can be measured separately (Degree, Katz, and Closeness), we do so for each.

In Table 8, we start our discussion by presenting correlations across the different measures within the PIO (top panel) and within the KIO (bottom panel). This allows us to compare how the different centrality measures compare when applied to the same network. Between the Degree and Katz measures, the correlations are fairly high, especially within the KIO. Closeness and Betweenness, however, are far less correlated with the other measures, including themselves. Finally, PageRank somewhat correlated with the Degree and Katz measures for the PIO but strongly so for the KIO. Thus, at least when using the Degree and Katz measures – as well as PageRank for the KIO – nodes that are more central according to one measure tend to be so for the other.

Next, in Table 9 we list the top ten most central countries according to each centrality measure. When using the PIO in the top panel, regardless of direction, China, the US, Japan, Germany, and Korea are most central according to Degree, Katz, and PageRank centrality. Likewise, the rest of the top ten tend to mirror what was found in Table 2 where we ranked countries according to node size and edge weights. The Closeness and Betweenness measures, however, provide a very different ranking with nations such as Brazil, Cyprus, and Lithuania ranking high. This may suggest that using these measures of centrality, ones which perhaps

[24]Where again the sum of weights was subsequently set to one with the same done in the KIO.

[25]Alternatively, we could use unweighted edges in our calculations so that any positive sales or citations resulted in an edge weight of one. Nevertheless as the purpose of input-output tables is generally to discuss the strength of connections (particularly given the large number of connections within countries) weighted edges seem more appropriate.

may not fit the context of an input-output table, may not yield much insight.

Turning to the KIO in the bottom panel, we again see a pattern reminiscent of Table 2 for Degree, Katz, and PageRank centrality. Again, the US, Germany, and Japan make strong showings, with China slipping in the rankings compared to its PIO performance. Also, we see a much stronger showing of the tax havens, notably Switzerland and the Netherlands. Sweden, well known for a leader in electronics innovation (which as shown momentarily is a central sector), often rounds out the top ten. Thus, compared to the PIO, it suggests that when it comes to importance in the knowledge network, Asian countries are somewhat more isolated and that tax policies may have something important to say not just in the amount of innovation generated, but also it how that feeds into subsequent innovation. Finally, once again the Closeness and Betweenness measures offer a rather different picture than the others, with countries such as Latvia and Australia making an appearance. Additionally, while the UK was not central in the other measures it shows up within the top five for Centrality and Betweenness. It is also worth noting that tax havens again make a strong showing with the Netherlands and Switzerland again appearing, now alongside Ireland and Belgium.

Likewise, Tables 10 and 11 list the top ten sectors according to centrality. Starting with the PIO results and focusing on Degree and Katz centrality in Table 10, we again see a pattern mirroring what was found in the node size and edge weight ranking. In particular, note that Food, Beverages, and Tobacco is ranked far lower than its node size, a result that is reflective of their smaller edge weights. It does, however, do better when using PageRank (Table 11). Also similar to Table 9, the Closeness and Betweeness rankings in Table 11 highlight sectors such as Wood and Paper that do not appear elsewhere. This adds further suggestive evidence that these measures may not be the most appropriate for describing global supply chains.

Similarly, at least for the Degree, Katz, and PageRank measures, the KIO results largely mirror what was seen in Table 3. It is worth recognizing, however, that here there is not such a large difference between these three measures and those for Closeness or Betweenness. As can be seen, even those measures suggest a pattern much more in line with the others. One potential reason for this is that the tangible goods trade in the PIO is direct whereas knowledge can inspire without being cited. Therefore concepts such as inspiration, cross-domain use, and cognitive proximity may make more sense in describing the network behind knowledge creation.[26]

To compare across the PIO and KIO, we again use regressions controlling for country and sector fixed effects. In Table 12 each column uses the specified PIO centrality measure as the dependent variable and the corresponding KIO centrality measure as the control variable. As can be seen, although the centrality measures are positively correlated with one other (excepting inbound Closeness), these correlations are fairly weak. Further, note that the variation explained within the regression is not particularly high (especially when using a measure other than degree centrality). This suggests that, despite the correlations in the size of nodes and edge weights, the role that country-sectors play in the overall global supply chain of intermediates and knowledge are not strongly correlated. This again suggests the possibility of fragmentation of innovation and production may be at play.

As a final examination of the data, Tables 13 and 14 list the top 40 country-sectors according to the Degree, Katz, and PageRank measures for the PIO and KIO respectively.[27] Mirroring the country and sector rankings, we tend to see reasonable consistency across these within the KIO and PIO, with the same countries and sectors found at the top. Also, we again see that the Asian country-sectors tend to rank lower in the KIO than the PIO. Finally, two tax havens – the Netherlands and Switzerland – again feature more in the KIO centrality measures than the PIO.

[26]For a useful discussion of how knowledge and intelligence can be conceptualized in terms of the ability to cross domains and spark innovation rather than imitation, we suggest reading Bostrom (2014).

[27]This is the top 40 out of 880.

5 Intangible Capital

As discussed by Davies and Markusen (2021), there is a long-standing recognition of the role intangible assets play in not just production overall, but especially for multinational firms. This is especially true in the context of profit shifting to avoid taxation. Here, we investigate the feasibility of using available patent data to approximate intangible capital and provide some suggestions about the role of tax policy in this relationship.

5.1 Constructing Intangible Capital

In order to measure the return to intangible capital at the country-industry level, we follow the approach from Chen et al. (2021) and Karabarbounis and Neiman (2018) which attributes the residual value-added after accounting for observed inputs to intangible capital. According to Karabarbounis and Neiman (2018), an industry's gross value-added VA is composed of labor L, tangible capital K, and intangible capital B. Thus, where w is the cost of labour and r is the return on capital (tangible and intangible), for country-sector cs:

$$r_{cs}B_{cs} = VA_{cs} - w_{cs}L_{cs} - r_{cs}K_{cs}. \tag{5}$$

From the national accounts data underlying the WIOD dataset, we observe both value-added and labour compensation. This then leaves us with the need to construct measures of the return on capital and the amount of tangible capital. Hall and Jorgenson (1967) shows that the rental price or user cost of capital consists of depreciation, capital taxes (net of subsidies), expected capital gains, and a net nominal rate of return. Ignoring capital taxes due to data limitations, the rental price of capital for cs is given by:

$$r^K_{cs} = (\delta^K_{cs} + \rho^K_{cs})p^I_{cs}, \tag{6}$$

where $\delta^K_{c,i}$ is the average capital depreciation rate for sector s in country c, $\rho^K_{c,i}$ is the real rate of return (set to 4%), and $\mathrm{p}^I_{c,i}$ is the investment price of the tangible asset.

To measure the tangible capital assets, we follow the asset boundary in the System of National Accounts (SNA) 2008. This includes building, machinery, transport equipment, information technology assets, communication technology assets and other tangible assets. We then employ the EU KLEMS dataset December 2016 release to obtain the tangible asset stocks collected by Jäger (2016).[28] In this, for tangible capital we include items such as computing equipment, communications equipment, transport equipment, other machinery and equipment, and total non-residential investment. We do not, however, include research and development, or other intellectual property assets as those fall under intangible capital. Thus, this allows us to construct the final term in 5 and back out the value added from intangible assets.

Using the above method of constructing the share of value-added attributed to intangibles for a country-sector-year, on can then calculate the US dollar equivalent of the value created by that country-sector for each year, so long as all of the data are available. Unfortunately, this is often not the case. The necessary information is not available at all for some countries and even for those where it is, it is not available for all years. Further, the underlying data aggregate across some of our sectors. Therefore, in the section, note that we are wroking with an unbalanced panel with the following countries: AT, CY, CZ, DE, DK, ES, FI, FR, GB, GR, IT, LU, NL, PT, SE, SK, US; the following sector groups: 10-12, 13-15, 16-18, 19, 20-21; the following years: 2000-2011. All together, this means that for some countries, it is not possible to construct measures of intangible capital via this method for nations where one might desire to have at least an approximation of their importance. With this in mind, for these countries where such data are available, we now examine how the difficult to construct measure of intangible value-added compares to easily obtainable information on

[28]This dataset provides information nominal capital stock, real fixed capital stock (2010 prices), gross fixed capital formation (priced in 2010), real gross fixed capital formation volume (2010), and nominal gross fixed capital formation.

patents.

5.2 Comparing Intangible Capital to Patents

To do so, we now estimate the following equation:

$$VA_{cst} = \beta Innovation_{cst} + Pop_{ct} + Areac + \gamma_s + \gamma_t + \varepsilon_{cst} \tag{7}$$

where VA is the intangible value-added for a country-sector-year (measured in millions of US dollars), Pop is the population of country c in year t, and $Area$ is c's square kilometers. These come from the CEPII database (Gaulier and Zignago (2010)).[29] Our variable of interest, $Innovation$, is a set of measures of patenting activity which are described in Section 3. We additionally control for fixed effects at the sector (γ_s) and year (γ_t) levels. Note that we specifically do not control for country fixed effects because our aim here is to consider how good one can fit the observable data to intangible value-added; for countries where the dependent variable is unavailable, no such country fixed effects can be estimated. Finally, ε is the robust standard error term.

In Table 15, we begin by controlling just for the stock of patents, population, and area (along with the year and sector fixed effects). This strongly suggests that a greater stock of patents implies more intangible value-added even when controlling for the two measures of country size. Column (2) introduces the patent flow, i.e. the number of new patents in year t. While we continue to find a positive coefficient for the stock of patents, patent flow is significantly negative. This suggests that when the stock is primarily comprised of young patents that this does not lead to as much value-added. Columns (3), (4), and (5) add in citations in one direction and then both. These are not significant.

Based on the above observations, we introduce a tax haven dummy for country c in column (6).[30] This coefficient is both large and significant; likewise it lowers the coefficient

[29] They can be found at `http://www.cepii.fr/CEPII/en/bdd_modele/bdd_modele_item.asp?id=37`.
[30] For this sample, this is for the Netherlands, Luxembourg, and Cyprus.

on the stock of patents (although it remains marginally significant). This suggests that some of evidence pointing to a large stock driving higher intangible value-added is driven by tax havens having both larger value-added and higher patent stocks. This is thus suggestive of multinationals in tax havens using patents as a means of shifting profits to avoid taxation.

Finally, column (7) introduces per-capita GDP to the estimation. We only do so at this point since one can reasonably consider that this is endogenous when estimating the value-added of intangible capital. Nevertheless, we find no significant relationship nor a marked effect on the other estimates.

Across the specifications, we find an adjusted R-squared of around 0.66, i.e. two-thirds of the variation can be explained by easily obtainable patent- and country-level data. While this means that there remains a sizable amount of variation that these do not explain, when considering the various assumptions (and noise) that feed into this measure of intangible capital, at the least we believe that this suggests that one can use patent data to provide a fair approximation of a difficult to obtain figure. Further, the explanatory variables are available for countries and years beyond those we use. As such, if one is willing to make the leap to out-of-sample application, this suggests that patent data can be very useful in measuring intangible capital.

Finally, in Table 16, we take the log of our variables (or the IHS in the case of the innovation measures due to zeros) and repeat our exercise. This results in coefficients that are less significant than in Table 15, but still point to a positive link between country-sector patenting and country-sector intangible value-added. In particular, this approach tends to find that both the stock and flow of patents point to higher intangible value-added. Finally, we see that the fit of this specification to the data is somewhat higher.

6 Conclusion

With the increasing recognition of the role global supply chains play in a variety of aspects of economic performance, there is a need for quality data to understand the connections between countries and industries. Although it has long been recognized that knowledge flows between them just as intermediate inputs do, there is little work attempting to compare how trade in intermediate goods parallels that in knowledge flows. In this paper, we do so by constructing and comparing physical and knowledge input-output tables for 44 countries across 20 sectors.

This exercise reveals some strong similarities between the goods network and the knowledge network. In particular, both are dominated by major economies including the US, China, Japan, and Germany. Further, both are fairly concentrated in a small number of countries and sectors. That said, differences are found. In particular, Asian countries are less important in the knowledge network whereas tax havens are generally more important. Furthermore, within country trade (but across sectors) is more important for goods whereas within sector (but across borders) citations matter for knowledge. Finally, while there is some correlation between the centrality measures derived for each, this is somewhat weak and varies by centrality measure.

In addition to these, we compare how measures of intangible assets constructed from input-output tables for goods compare to the patent data used to describe the knowledge network. We find that a significant share of the variation of the constructed intangible value-added can be explained by country-sector innovation data. In particular, the wide availability of the latter suggests that it may have use as a proxy for intangible assets when data on the latter are unavailable.

While our exercise here has been largely descriptive, we believe that it complements that used elsewhere, in particular by pointing out the issues with outliers when comparing networks. As such, we hope that the methodologies, issues raised, and stylized facts discussed serve as a springboard for future research on global supply chains, their determinants, and the effects they have.

References

Ahmed, S., Appendino, M. and Ruta, M. (2016), 'Global value chains and the exchange rate elasticity of exports', *The B.E. Journal of Macroeconomics* **17**(1).

Alcácer, J. and Gittelman, M. (2006), 'Patent citations as a measure of knowledge flows: The influence of examiner citations', *The Review of Economics and Statistics* **88**(4), 774–779. **URL:** *http://www.jstor.org/stable/40043034*

Amiti, M. and Konings, J. (2007), 'Trade liberalization, intermediate inputs, and productivity: Evidence from indonesia', *American Economic Review* **97**(5), 1611–1638.

Annette Alstadsæter, Sarah Godar, P. N. and Zucman, G. (2024), 'Global Tax Evasion Report 2024', *EU Tax Observatory* .

Antràs, P., Chor, D., Fally, T. and Hillberry, R. (2012), 'Measuring the upstreamness of production and trade flows', *American Economic Review* **102**(3), 412–16.

Ayerst, S., Ibrahim, F., MacKenzie, G. and Rachapalli, S. (2023), 'Trade and diffusion of embodied technology: an empirical analysis', *Journal of Monetary Economics* **137**, 128–145.

Baldwin, R. (2016), *The Great Convergence*, Harvard University Press, Cambridge, MA and London, England.

Bas, M. and Strauss-Kahn, V. (2015), 'Input-trade liberalization, export prices and quality upgrading', *Journal of International Economics* **95**(2), 250–262.

Bellemare, M. F. and Wichman, C. J. (2020), 'Elasticities and the inverse hyperbolic sine transformation', *Oxford Bulletin of Economics and Statistics* **82**(1), 50–61.

Bernard, A. B., Jensen, J. B., Redding, S. J. and Schott, P. K. (2012), 'The empirics of firm heterogeneity and international trade', *Annual Review of Economics* **4**(1), 283–313.

Blanchard, E. J., Bown, C. P. and Johnson, R. C. (2016), Global supply chains and trade policy, Working Paper 21883, National Bureau of Economic Research.

Blundell, R., Griffith, R. and Reenen, J. V. (1995), 'Dynamic Count Data Models of Technological Innovation', *The Economic Journal* **105**(429), 333–344.

Bonadio, B., Huo, Z., Levchenko, A. A. and Pandalai-Nayar, N. (2021), 'Global supply chains in the pandemic', *Journal of International Economics* **133**, 103534.
URL: *https://www.sciencedirect.com/science/article/pii/S0022199621001148*

Bonadio, B., Huo, Z. and Levchenko, Andrei A. and, P.-N. N. (2021), 'Global supply chains in the pandemic', *Journal of International Economics* **133**.

Bostrom, N. (2014), *Superintelligence: Paths, Dangers, Strategies*, Oxford University Press, Oxford, England.

Brandt, L., Van Biesebroeck, J., Wang, L. and Zhang, Y. (2017), 'Wto accession and performance of chinese manufacturing firms', *American Economic Review* **107**(9), 2784–2820.

Breschi, S. and Lissoni, F. (2005), *Knowledge Networks from Patent Data*, pp. 613–643.

Carlo Giglio, Roberto Sbragia, R. M. and Palmieri, R. (2021), 'Cross-country learning from patents: an analysis of citations flows in innovation trajectories', *Scientometrics* **126**, 7917–7936.

Chen, W., Los, B. and Timmer, M. P. (2021), 373Factor Incomes in Global Value Chains: The Role of Intangibles, *in* 'Measuring and Accounting for Innovation in the Twenty-First Century', University of Chicago Press.

Chor, D., Manova, K. and Yu, Z. (2021), 'Growing like china: Firm performance and global production line position', *Journal of International Economics* **130**, 103445. NBER International Seminar on Macroeconomics 2020.

Coquidé, C., Lages, J., Ermann, L. and Shepelyansky, D. L. (2022), 'Covid-19rsquo;s impact on international trade', *Entropy* **24**(3).

Davies, R. B. and Markusen, J. R. (2021), 'What do multinationals do? the structure of multinational firms' international activities', *The World Economy* **44**(12), 3444–3481.

Davies, R. B., Martin, J., Parenti, M. and Toubal, F. (2018), 'Knocking on Tax Haven's Door: Multinational Firms and Transfer Pricing', *The Review of Economics and Statistics* **100**(1), 120–134.

Davies, R., Kogler, D. and Hynes, R. (2020), 'Patent boxes and the success rate of applications', *CES-Ifo Working Paper* (No, 8375).

de Bodt, E., Cousin, J.-G. and Dupire, M. (2023), Does CSR Help Firms to Cope with Supply Chain Disruptions ? Evidence from the Suez Canal Ever Given Obstruction, Technical report.

Francois, J. and Hoekman, B. (2010), 'Services trade and policy', *Journal of Economic Literature* **48**(3), 642–692.

Gaulier, G. and Zignago, S. (2010), Baci: International trade database at the product-level. the 1994-2007 version, Working Papers 2010-23, CEPII.
URL: *http://www.cepii.fr/CEPII/fr/publications/wp/abstract.asp?NoDoc=2726*

Goldberg, P. K., Khandelwal, A. K., Pavcnik, N. and Topalova, P. (2010), 'Imported Intermediate Inputs and Domestic Product Growth: Evidence from India*', *The Quarterly Journal of Economics* **125**(4), 1727–1767.

Hall, B. H., Helmers, C., Rogers, M. and Sena, V. (2012), The choice between formal and informal intellectual property: A literature review, Working Paper 17983, National Bureau of Economic Research.

Hall, R. E. and Jorgenson, D. W. (1967), 'Tax policy and investment behavior', *The American Economic Review* **57**(3), 391–414.

Halpern, L., Koren, M. and Szeidl, A. (2015), 'Imported inputs and productivity', *American Economic Review* **105**(12), 3660–3703.

Hayakawa, K. and Mukunoki, H. (2021), 'The impact of covid-19 on international trade: Evidence from the first shock', *Journal of the Japanese and International Economies* **60**, 101135.

Helpman, E. (1984), 'A simple theory of international trade with multinational corporations', *Journal of Political Economy* **92**(3), 451–471.

Heuser, C., Mattoo, A., Heuser, C. and Mattoo, A. (2017), Services trade and global value chains, Policy Research Working Paper Series 8126, The World Bank.

Hubbell, C. H. (1965), 'An input-output approach to clique identification', *Sociometry* **28**(4), 377–399.

Hummels, D., Ishii, J. and Yi, K.-M. (2001), 'The nature and growth of vertical specialization in world trade', *Journal of International Economics* **54**(1), 75–96. Trade and Wages.

Jackson, M. (2008), *Social and Economic Networks*, Princeton University Press, Princeton, NJ.

Jäger, K. (2016), 'EU KLEMS Growth and Productivity Accounts 2016 release - Description of Methodology and General Notes'.

Jie Cai, J., Li, N. and Santacreu, A. M. (2022), 'Knowledge Diffusion, Trade, and Innovation across Countries and Sectors', *AMERICAN ECONOMIC JOURNAL: MACROECONOMICS* **14**(1), 104–145.

Johnson, R. C. and Noguera, G. (2012), 'Accounting for intermediates: Production sharing and trade in value added', *Journal of International Economics* **86**(2), 224–236.

Karabarbounis, L. and Neiman, B. (2018), Accounting for factorless income, Working Paper 24404, National Bureau of Economic Research.

Kasahara, H., Liang, Y. and Rodrigue, J. (2016), 'Does importing intermediates increase the demand for skilled workers? plant-level evidence from indonesia', *Journal of International Economics* **102**, 242–261.

Kasahara, H. and Rodrigue, J. (2008), 'Does the use of imported intermediates increase productivity? plant-level evidence', *Journal of Development Economics* **87**(1), 106–118.

Katz, L. (1953), 'A new status index derived from sociometric analysis', *Psychometrika* **18**, 39–43.

Kleinberg, J. (1999), 'Authoritative sources in a hyperlinked environment', *Journal of the ACM* **46**, 604–632.

Koopman, R., Wang, Z. and Wei, S.-J. (2014), 'Tracing value-added and double counting in gross exports', *American Economic Review* **104**(2), 459–94.
URL: *https://www.aeaweb.org/articles?id=10.1257/aer.104.2.459*

Kugler, M. and Verhoogen, E. (2009), 'Plants and imported inputs: New facts and an interpretation', *American Economic Review* **99**(2), 501–07.

Kugler, M. and Verhoogen, E. (2011), 'Prices, Plant Size, and Product Quality', *The Review of Economic Studies* **79**(1), 307–339.

Liu, E. and Ma, S. (2021), Innovation networks and rd allocation, Working Paper 29607, National Bureau of Economic Research.

Lu, Y., Lu, Y., Xie, R. and Yu, X. (2019), 'Does global value chain engagement improve firms' wages: Evidence from china', *The World Economy* **42**(10), 3065–3085.

MacGarvie, M. (2005), 'The determinants of international knowledge diffusion as measured by patent citations', *Economics Letters* **87**(1), 121–126.

Markusen, J. R. (1984), 'Multinationals, multi-plant economies, and the gains from trade', *Journal of International Economics* **16**(3), 205–226.

Njotto, L. L. (2018), 'Centrality measures based on matrix functions', *Open Journal of Discrete Mathematics* **8**(4), 79–115.

Orhan, E. (2022), 'The effects of the russia - ukraine war on global trade', *Journal of International Trade, Logistics and Law* **8**(1), 141–146.

Parteka, A. and Wolszczak-Derlacz, J. (2018), 'Global value chains and wages: Multi-country evidence from linked worker-industry data', *Open Economies Review* pp. 1–35.
URL: *https://api.semanticscholar.org/CorpusID:158219570*

Sargent, T. J. and Stachurski, J. (2008), *Social and Economic Networks*.

Schott, P., Pierce, J., Schaur, G. and Heise, S. (2017), Trade policy uncertainty and the structure of supply chains, 2017 Meeting Papers 788, Society for Economic Dynamics.

Schwab, T. and Todtenhaupt, M. (2021), 'Thinking outside the box: The cross-border effect of tax cuts on r&d', *Journal of Public Economics* **204**(C).

Steinbach, S. (2023), 'The russia–ukraine war and global trade reallocations', *Economics Letters* **226**, 111075.

Stiebale, J. (2016), 'Cross-border mas and innovative activity of acquiring and target firms', *Journal of International Economics* **99**, 1–15.

Timmer, M. P., Dietzenbacher, E., Los, B., Stehrer, R. and de Vries, G. J. (2015), 'An illustrated user guide to the world input–output database: the case of global automotive production', *Review of International Economics* **23**(3), 575–605.

Timmer, M. P., Los, B., Stehrer, R. and de Vries, G. J. (2016), An anatomy of the global trade slowdown based on the wiod 2016 release.

Tørsløv, T., Wier, L. and Zucman, G. (2022), 'The Missing Profits of Nations', *The Review of Economic Studies* **90**(3), 1499–1534.
URL: *https://doi.org/10.1093/restud/rdac049*

Vandenbussche, H., Connell, W. and Simons, W. (2022), 'Global value chains, trade shocks and jobs: An application to brexit', *The World Economy* **45**(8), 2338–2369.

Wu, F. and Dong, J. (2022), 'Impact of international freight on trade and supply chains-based on ever given events', *Frontiers in Economics and Management* **3**(9), 330–337.

Appendix

A Sector Codes

Table 1: Industry Codes and Descriptions

ISIC Industry Code	Description
5	Food, beverage, & tobacco
6	Textiles
7	Wood
8	Paper
9	Recorded Media
10	Coke & Petroleum
11	Chemicals
12	Pharmaceuticals
13	Rubber & Plastics
14	Other non-metallic
15	Metals
16	Fabricated Metals
17	Computers & Electronics
18	Electrical equipment
19	Machinery
20	Motor Vehicles
21	Other Transport
22	Furniture
27	Construction
40	Computer Programming

Notes: These industry codes are the ISIC Rev2 codes used in both the WIOD and PATSTAT databases.

B PageRank Calculation

The PageRank value for a node i is calculated as follows:

1. **Initialization**: Assign each node an initial rank. If there are N node, each node i gets a rank $R(i) = \frac{1}{N}$.

2. **Iterative Calculation**: Update the rank of each node based on the ranks of nodes citing it. The formula is given by:

$$R(i) = \frac{1-d}{N} + d \sum_{y \in B_i} \frac{R(j)}{L(j)}$$

Here:

- $R(i)$ is the rank of node i.

- d is the damping factor (α), typically set around 0.85.
- B_i represents the set of nodes for which i has an inbound edge.
- $L(j)$ is the number of outward edges of node j.

3. **Convergence**: Repeat the calculation until the ranks converge.
4. **Normalization**: After convergence, normalize the PageRank values so that the sum of all ranks equals 1.

C R code

In this section, we show the R code used to generate different measures of a country-industry pair's centrality. Our R code provide an easy one-click generating process. In order to use this code, we need to prepare all the data files in the directory folders first. After running this code, the directory will generate the centrality measure in a *.CSV file with named "datafilename_process"

R code to compute the centrality measures

```r

# Install and load the 'igraph' package if you haven't already
install.packages("igraph")
library(igraph)
install.packages("haven")
library(haven)
install.packages("dplyr")
library(dplyr)
# Set the directory containing the CSV files
#directory_path <- "C:\Users\guoha\OneDrive\Desktop\Network analysis with KIO\test"  # Update this with your directory path

directory_path <- "C:\\Users\\guoha\\OneDrive\\Desktop\\Network analysis with KIO\\test"
directory_path_new <- "C:\\Users\\guoha\\OneDrive\\Desktop\\Network analysis with KIO\\test_new"

# List all CSV files in the directory

csv_files <- list.files(directory_path, pattern = "\\.csv$", full.names = TRUE)


for (file in csv_files) {

  # Example of renaming a column
  network_data <- read.csv(file)
  names(network_data)[names(network_data) == "value"] <- "weights"
  names(network_data)[names(network_data) == "citations"] <- "weights"

```

```r
    result_filename <- paste0(directory_path, "/", tools::file_path_sans_ext(basename(file)), "_centrality.dta")

  # Define a new filename for the exported file
  new_filename <- paste0(directory_path_new, "/", tools::file_path_sans_ext(basename(file)), "_modify.csv")

  # Write the data to a new CSV file
  write.csv(network_data, new_filename, row.names = FALSE)

}

csv_files_new <- list.files(directory_path_new, pattern = "\\.csv$", full.names = TRUE)

# Loop over each file
for (file in csv_files_new) {
  # Read the network data from the CSV file
  #  network_data <- read.csv(file)

  edge_data <- read.csv(file)
  edge_data <- edge_data[c("origin", "destination", "weights")]

  # Create a graph from the edge list, considering 'citations' as weights
  g <- graph_from_data_frame(edge_data, directed = TRUE)
  # Set the edge weights (if you want to perform weighted analysis)
  E(g)$weight <- edge_data$weights

  new_weight <- ifelse(E(g)$weight > 0, 1, 0)
  g1 <- graph_from_data_frame(edge_data, directed = TRUE)

  E(g1)$weight <-  new_weight


  # Calculate weighted authority
  authority_weighted <- authority_score(g, weights = E(g)$weight)$vector
  # Calculate unweighted authority
  authority_unweighted <- authority_score(g1, weights = E(g1)$weight)$vector

  # Calculate weighted Eigenvector Centrality (for unweighted, remove the 'weights' parameter)
  Eigenvector_weighted_centrality <- eigen_centrality(g, weights = E(g)$weight)$vector


  # Calculate weighted in-degree and out-degree
  in_degree_weighted <- strength(g, mode = "in", weights = E(g)$weight)
  out_degree_weighted <- strength(g, mode = "out", weights = E(g)$weight)


  # Calculate weighted PageRank
  page_rank_weighted <- page_rank(g)$vector

```

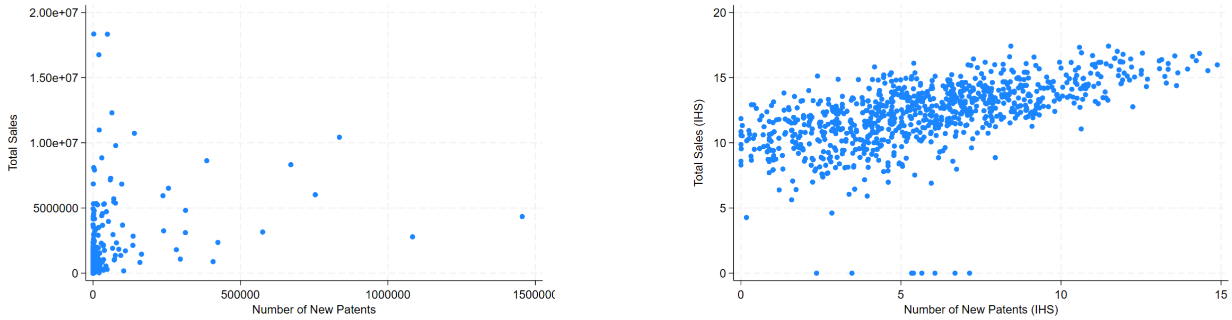
```
# Calculate unweighted in-degree and out-degree
in_degree_unweighted <- strength(g1, mode = "in", weights = E(g1)$weight
 )
out_degree_unweighted <- strength(g1, mode = "out", weights = E(g1)$
 weight)


# Calculate unweighted PageRank
page_rank_unweighted <- page_rank(g1)$vector
Eigenvector_unweighted_centrality <- eigen_centrality(g1, weights = E(g1
 )$weight)$vector


# Calculate weighted and unweighted hubs
hub_unweighted <- hub_score(g1, weights = E(g1)$weight)$vector
hub_weighted <- hub_score(g, weights = E(g)$weight)$vector

# Calculate inward Katz centrality with the specified alpha
adj_matrix <- as_adjacency_matrix(g1, type = "both", attr = "weight",
 sparse = FALSE)
t_adj_matrix = t(adj_matrix)

# Define Katz centrality parameters (you can adjust these as needed)
alpha <- 1/max(abs(Re(eigen(adj_matrix)$values))) -  0.1/max(abs(Re(
 eigen(adj_matrix)$values)))

# Damping factor
beta <- 1.0   # Scaling factor
max_iterations <- 2000  # Maximum number of iterations


# Compute the inward Katz centrality
# Initialize Katz centrality vector
n <- vcount(g)
x <- (alpha) * adj_matrix
y <- rep(1, n)

interation_counter <- 1

# Calculate Katz centrality iteratively
for (iteration in 1:max_iterations) {

  new_x <- (alpha) * adj_matrix %*% x
  y <- new_x %*% rep(1, n) + y
  k <- abs(new_x - x)
  if (all(abs(new_x - x) < 1e-6)) {
    break
    print("Katz centrality converge")
  }
  x <- new_x
  iteration_counter <- iteration
}
```

```
inward_unweighted_Katz_centrality <- y + (alpha) * adj_matrix %*%rep(1,
  n)

# Calculate outward Katz centrality with the specified alpha

# Compute the outward Katz centrality
# Initialize Katz centrality vector
n <- vcount(g)
x <- (alpha) * t_adj_matrix
y <- rep(1, n)

interation_counter <- 1

# Calculate Katz centrality iteratively
for (iteration in 1:max_iterations) {

  new_x <- (alpha) * t_adj_matrix %*% x
  y <- new_x %*% rep(1, n) + y
  k <- abs(new_x - x)
  if (all(abs(new_x - x) < 1e-6)) {
    break
    print("Katz centrality converge")
  }
  x <- new_x
  iteration_counter <- iteration
}

outward_unweighted_Katz_centrality <- y + (alpha) * t_adj_matrix %*%rep
  (1, n)




# Calculate weighted Katz centrality with the specified alpha
adj_matrix <- as_adjacency_matrix(g, type = "both", attr = "weight",
  sparse = FALSE)
t_adj_matrix = t(adj_matrix)

# Define Katz centrality parameters (you can adjust these as needed)
alpha <- 1/max(abs(Re(eigen(adj_matrix)$values))) -  0.1/max(abs(Re(
  eigen(adj_matrix)$values))

# Damping factor
beta <- 1.0   # Scaling factor
max_iterations <- 2000  # Maximum number of iterations


# Compute the weighted inward Katz centrality
# Initialize Katz centrality vector
n <- vcount(g)
x <- (alpha) * adj_matrix
y <- rep(1, n)

```

```
interation_counter <- 1

# Calculate Katz centrality iteratively
for (iteration in 1:max_iterations) {

  new_x <- (alpha) * adj_matrix %*% x
  y <- new_x %*% rep(1, n) + y
  k <- abs(new_x - x)
  if (all(abs(new_x - x) < 1e-6)) {
    break
    print("Katz centrality converge")
  }
  x <- new_x
  iteration_counter <- iteration
}

inward_weighted_Katz_centrality <- y + (alpha) * adj_matrix %*%rep(1, n)


# Calculate outward Katz centrality with the specified alpha

# Compute the outward Katz centrality
# Initialize Katz centrality vector
n <- vcount(g)
x <- (alpha) * t_adj_matrix
y <- rep(1, n)

interation_counter <- 1

# Calculate Katz centrality iteratively
for (iteration in 1:max_iterations) {

  new_x <- (alpha) * t_adj_matrix %*% x
  y <- new_x %*% rep(1, n) + y
  k <- abs(new_x - x)
  if (all(abs(new_x - x) < 1e-6)) {
    break
    print("Katz centrality converge")
  }
  x <- new_x
  iteration_counter <- iteration
}

outward_weighted_Katz_centrality <- y + (alpha) * t_adj_matrix %*%rep(1,
  n)





# Remove edges with zero weight
g_filtered <- delete_edges(g, E(g)[E(g)$weight == 0])
g1_filtered <- delete_edges(g1, E(g1)[E(g1)$weight == 0])

```

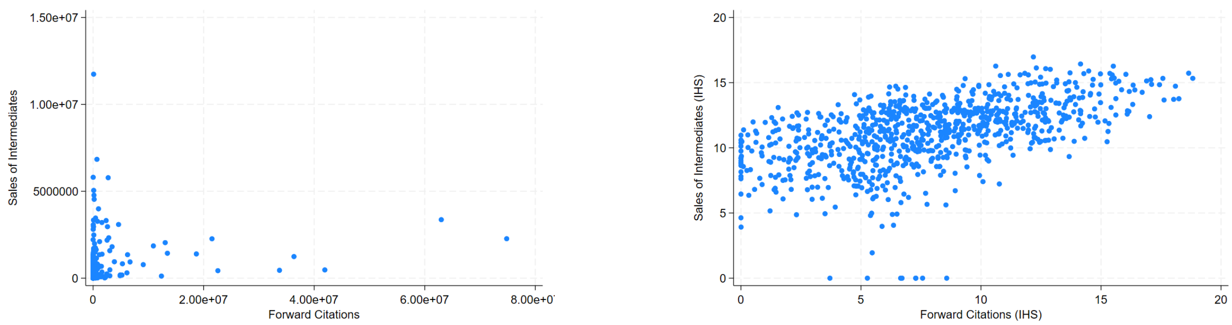
```
  in_closeness_weighted <- closeness(g_filtered, mode = "in")
  out_closeness_weighted <- closeness(g_filtered, mode = "out")
  in_closeness_unweighted <- closeness(g1_filtered, mode = "in")
  out_closeness_unweighted <- closeness(g1_filtered, mode = "out")

  between_weighted <- betweenness(g_filtered, directed = TRUE)
  between_unweighted <- betweenness(g1_filtered, directed = TRUE)


  # Combine centrality measures into a data frame
  centrality_result <- data.frame(
    node = V(g)$name,
    in_degree_weighted = in_degree_weighted,
    out_degree_weighted = out_degree_weighted,
    in_degree_unweighted = in_degree_unweighted,
    out_degree_unweighted = out_degree_unweighted,
    in_close_weighted = in_closeness_weighted,
    out_close_weighted = out_closeness_weighted,
    in_close_unweighted = in_closeness_unweighted,
    out_close_unweighted = out_closeness_unweighted,
    btw_weighted = between_weighted,
    btw_unweighted = between_unweighted,
    hub_unweighted = hub_unweighted,
    hub_weighted = hub_weighted,
    authority_weighted = authority_weighted,
   authority_unweighted = authority_unweighted,
    page_rank_weighted = page_rank_weighted,
    page_rank_unweighted = page_rank_unweighted,
    eigen_unweighted = Eigenvector_unweighted_centrality,
    eigen_weighted  = Eigenvector_weighted_centrality,
    out_weighted_Katz  = outward_weighted_Katz_centrality,
    out_unweighted_Katz  = outward_unweighted_Katz_centrality,
    in_weighted_Katz  = inward_weighted_Katz_centrality,
    in_unweighted_Katz  = inward_unweighted_Katz_centrality
  )


  # Apply your centrality computation code here
  # centrality_result <- compute_centrality(network_data)  # Replace with
   your function

  # Define a new filename for the results
  result_filename <- paste0(directory_path_new, "/", tools::file_path_sans
   _ext(basename(file)), "_centrality.dta")

  # Write the results to a dta file
  write_dta(centrality_result, result_filename )

}

# Your results are now saved in individual files corresponding to each
   input file
```

Figure 1: Total Sales vs. Patents (Node Size)



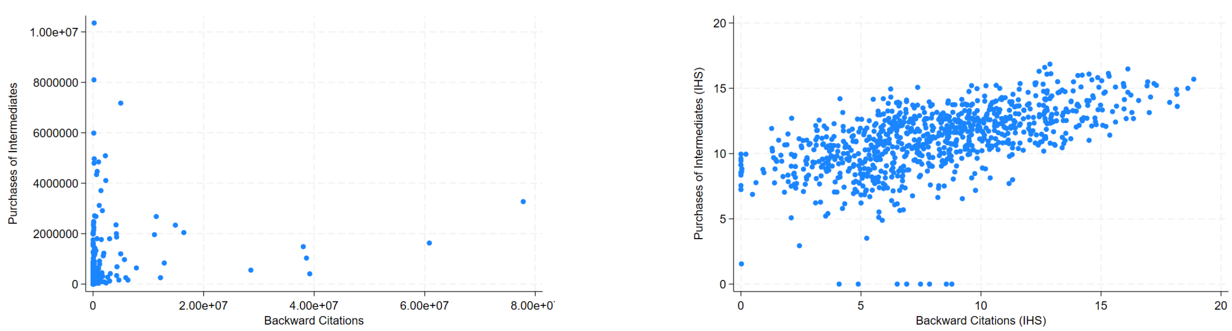
(a) Levels (b) IHS

Figure 2: Sales vs. Forward Citations (Outbound Edge Weight)



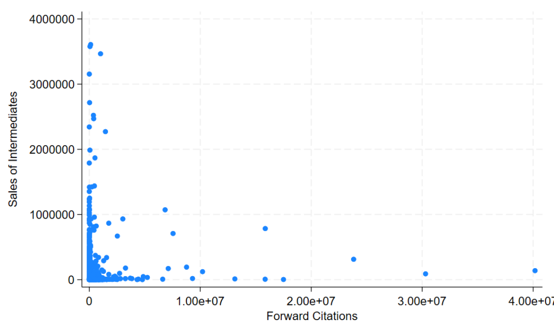
(a) Levels (b) IHS

Figure 3: Purchases vs. Backward Citations (Inbound Edge Weight)

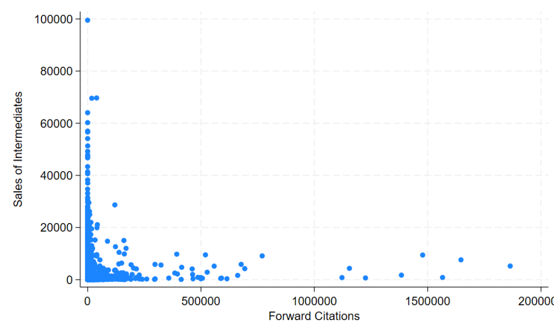


(a) Levels (b) IHS

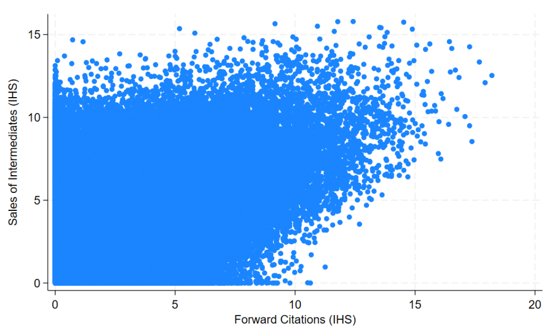
Figure 4: Bilateral Edgeweights



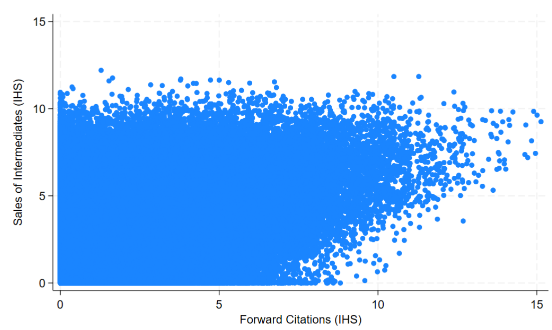
(a) Levels



(b) Levels (No Loops)

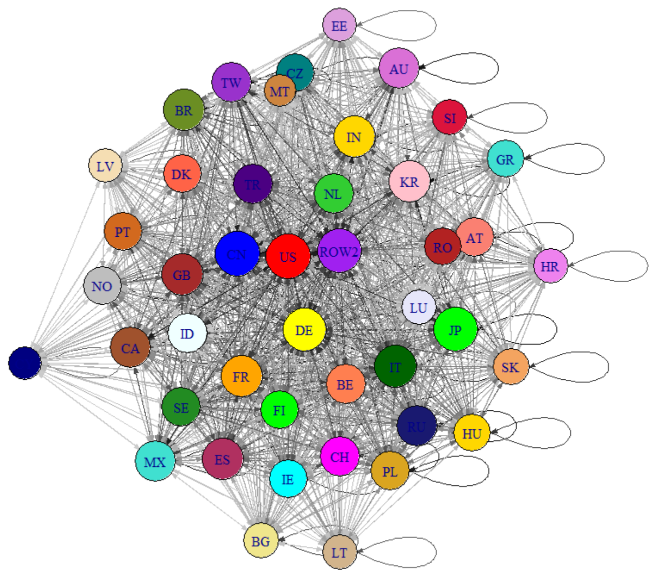


(c) IHS

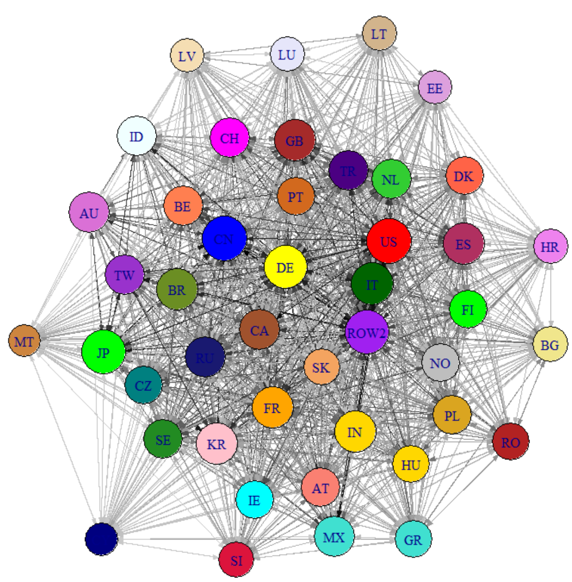


(d) IHS - No Loops

Figure 5: Countries in the PIO Network

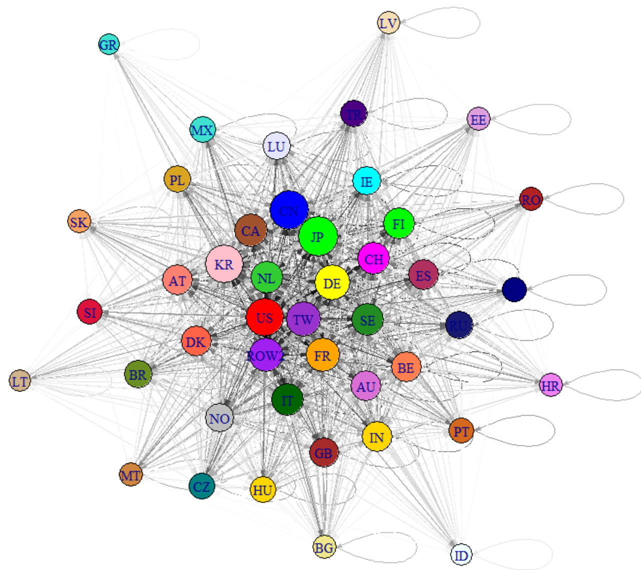


(a) With Loops

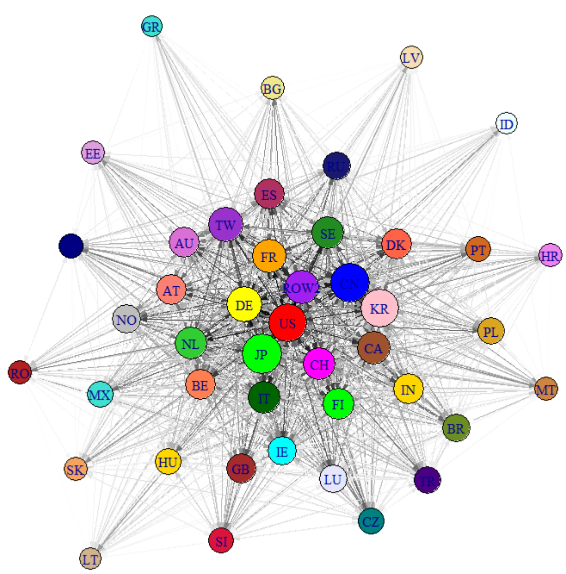


(b) Without Loops

Figure 6: Countries in the KIO Network

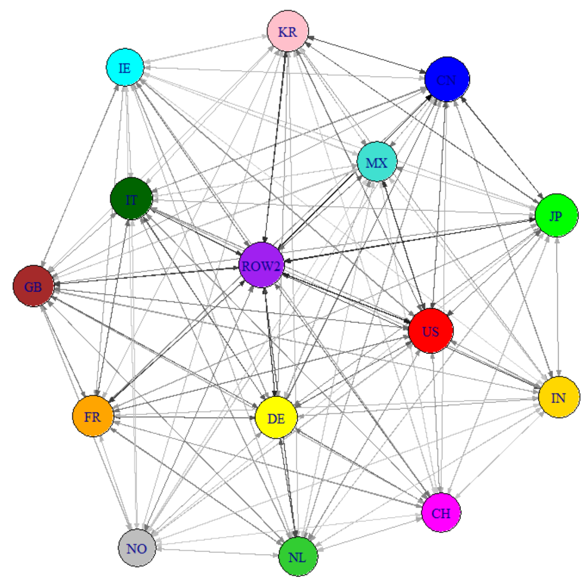


(a) With Loops

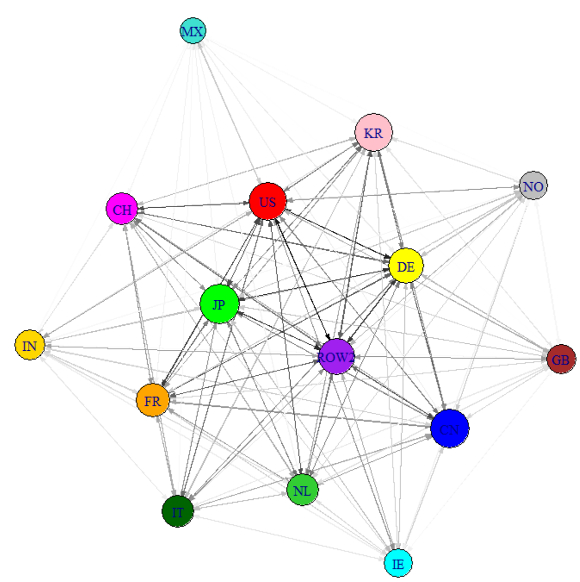


(b) Without Loops

Figure 7: Select Countries

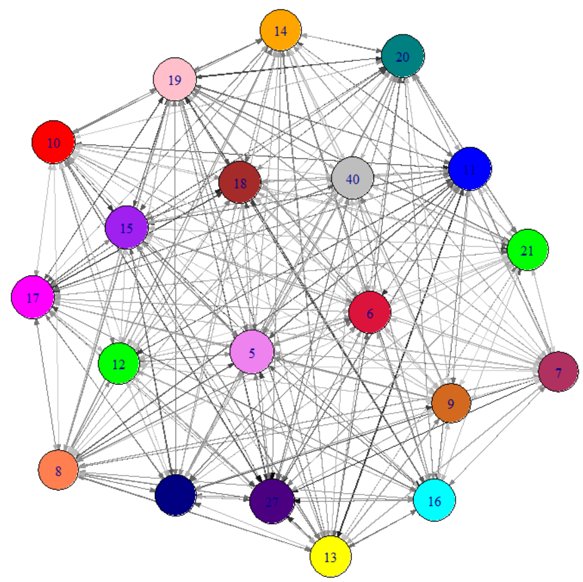


(a) PIO

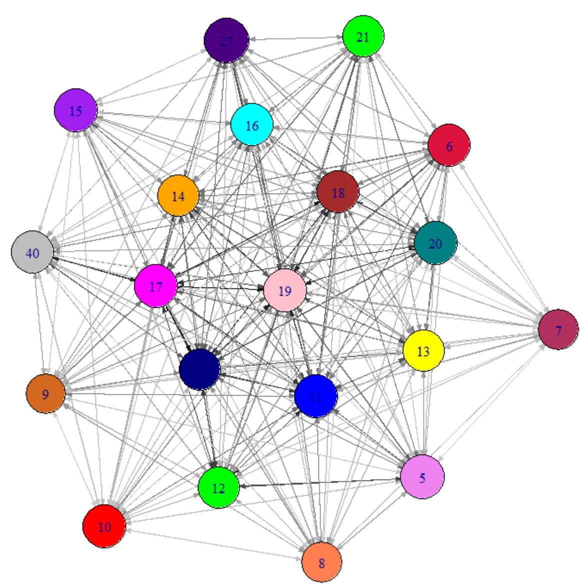


(b) KIO

Figure 8: Comparing the PIO and KIO across Sectors

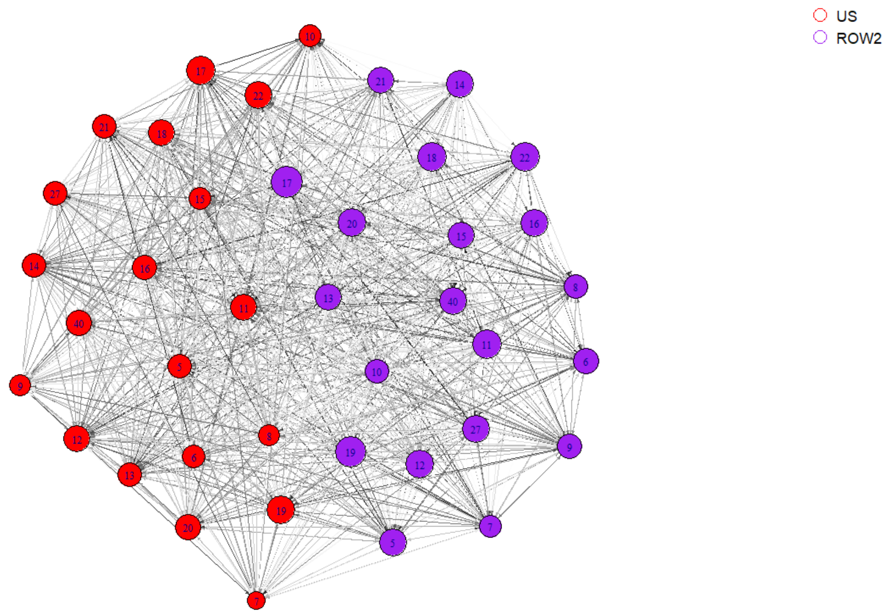


(a) PIO

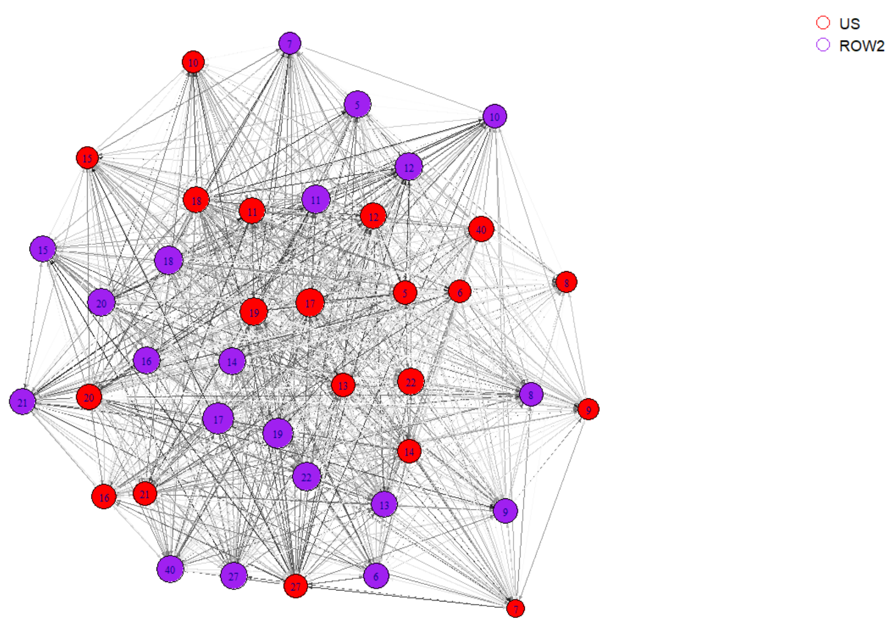


(b) KIO

Figure 9: United States

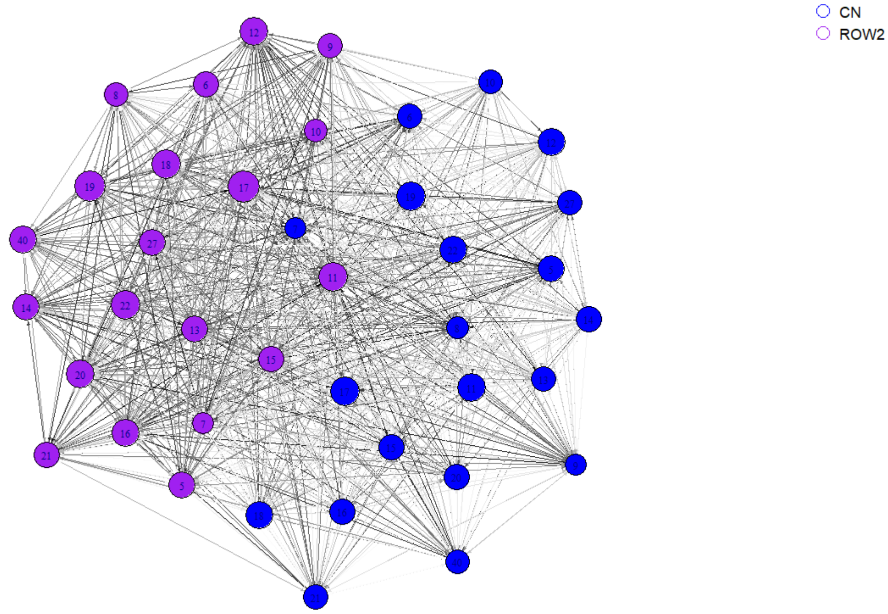


(a) PIO

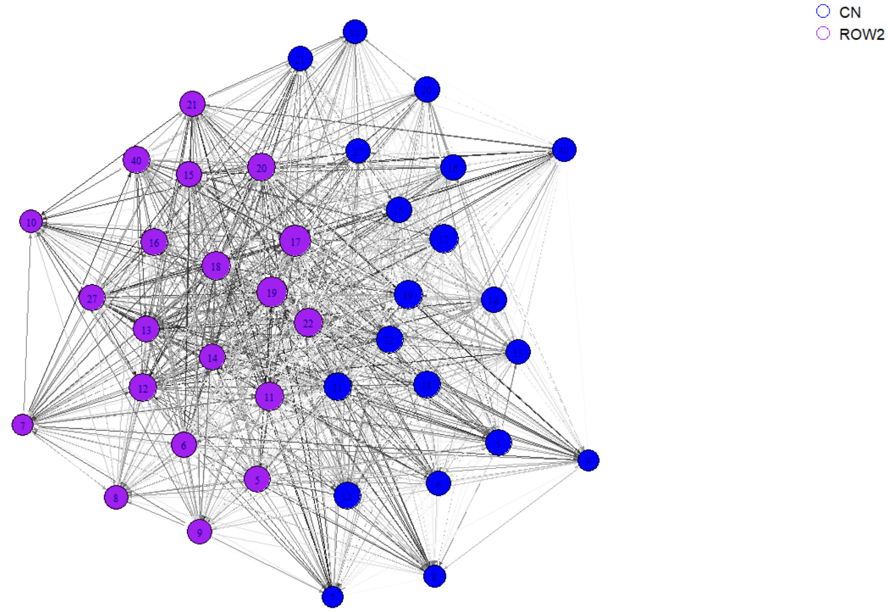


(b) KIO

Figure 10: China

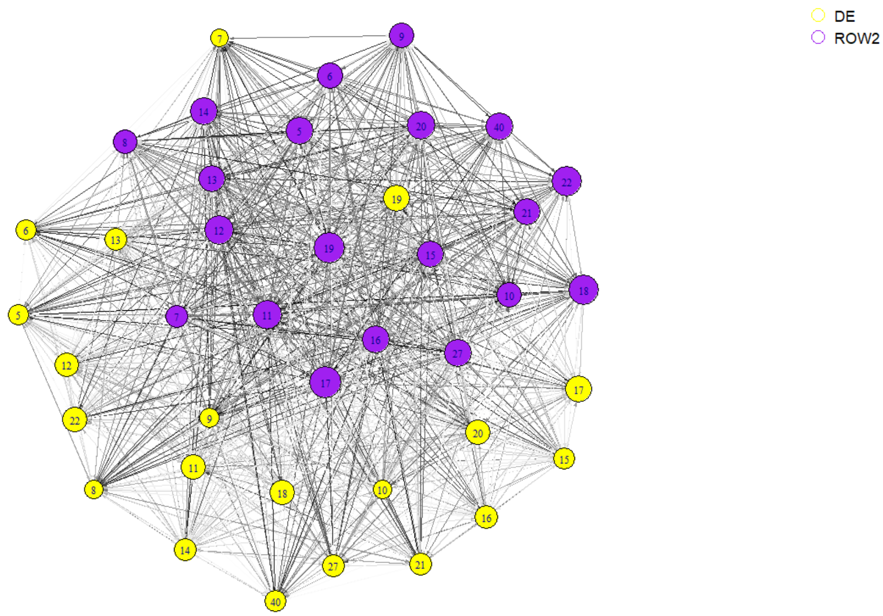


(a) PIO

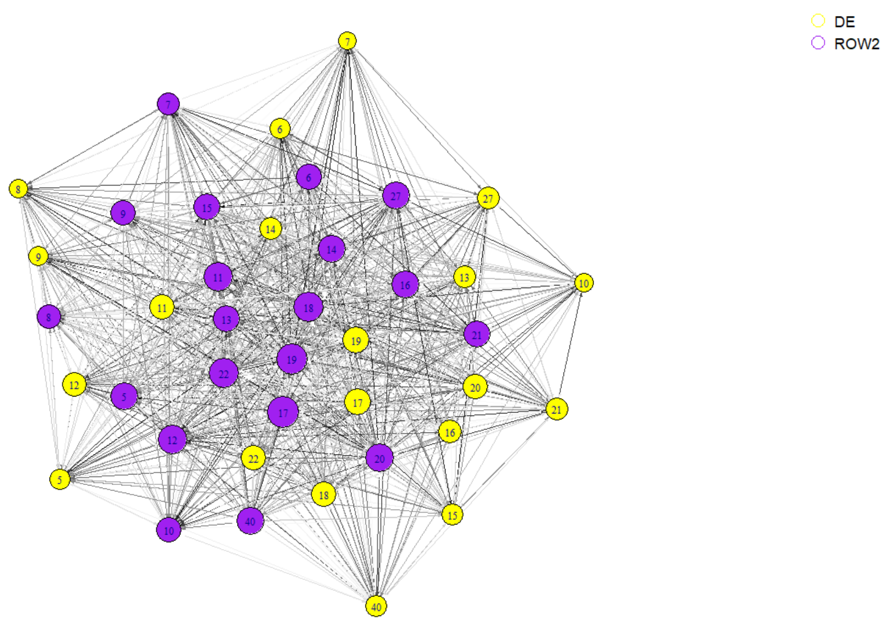


(b) KIO

Figure 11: Germany

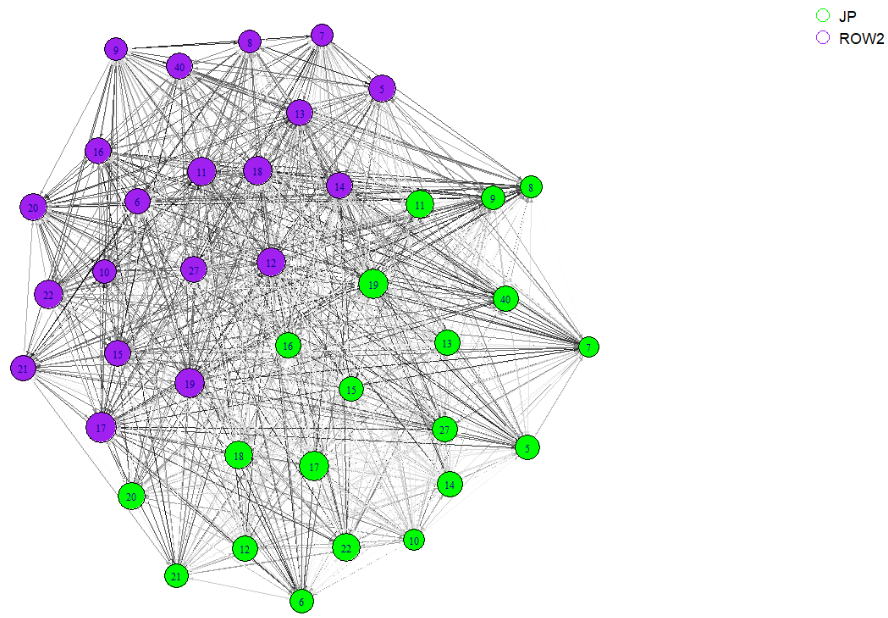


(a) PIO

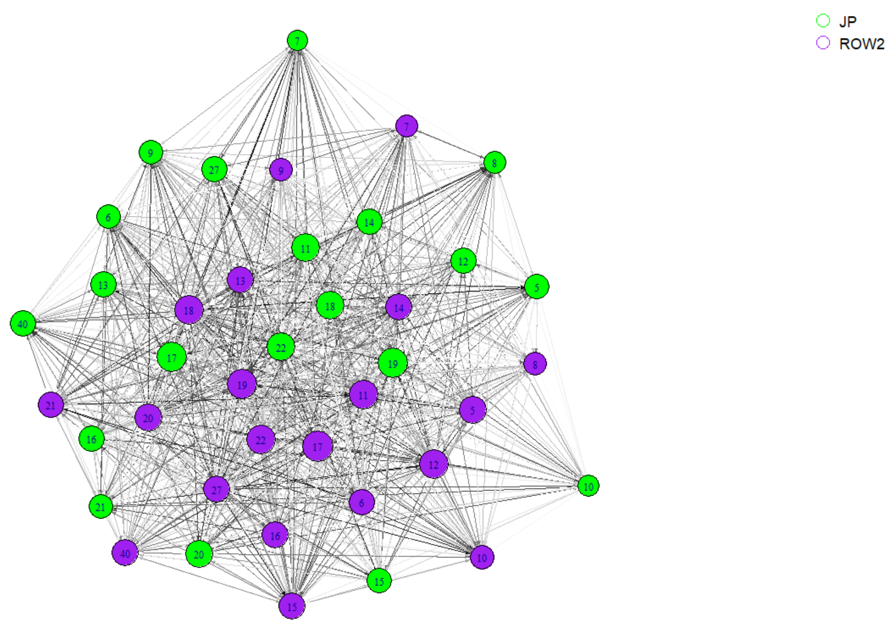


(b) KIO

Figure 12: Japan

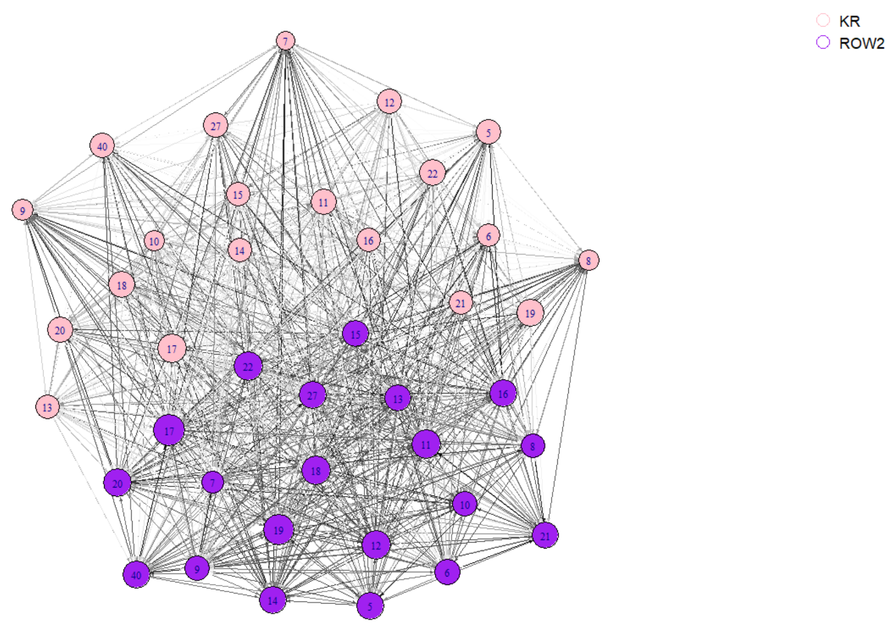


(a) PIO

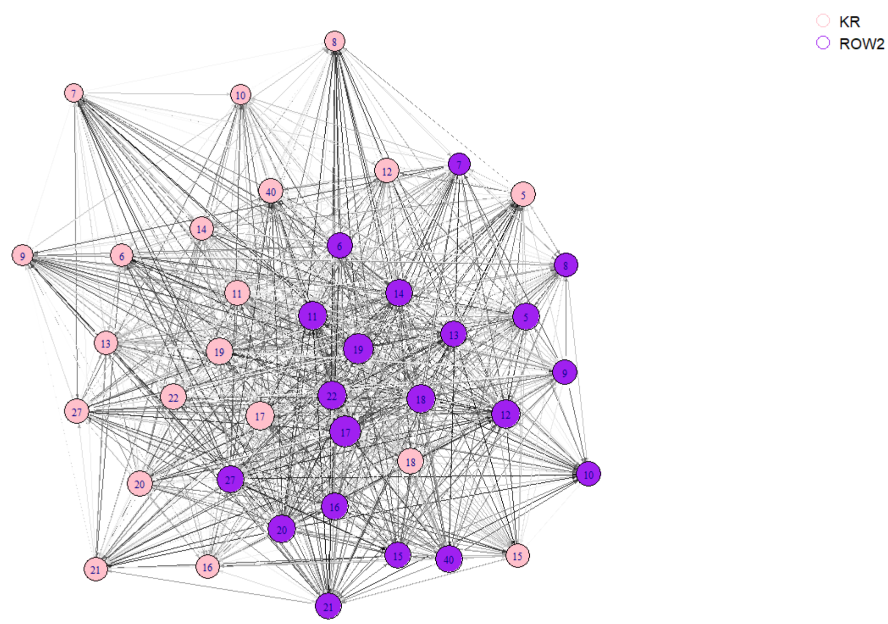


(b) KIO

Figure 13: Korea

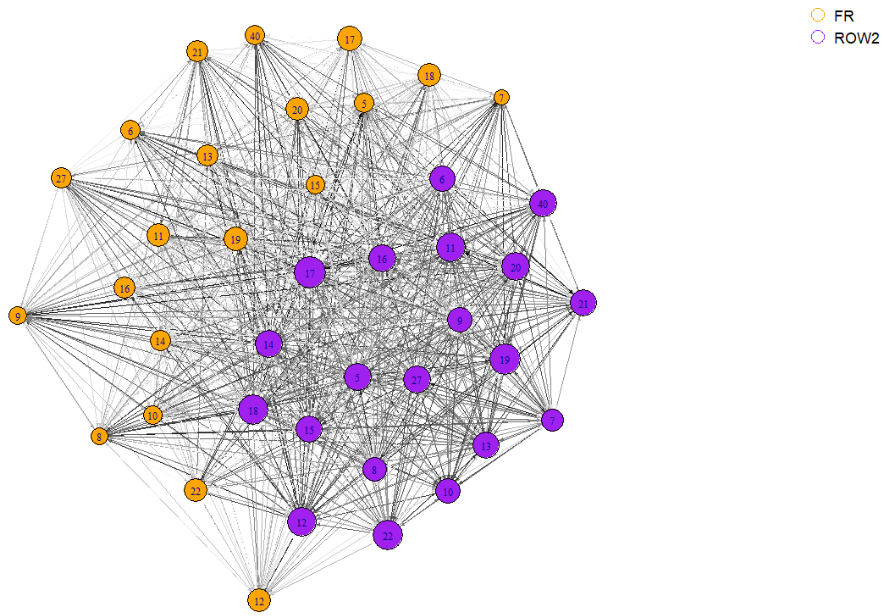


(a) PIO

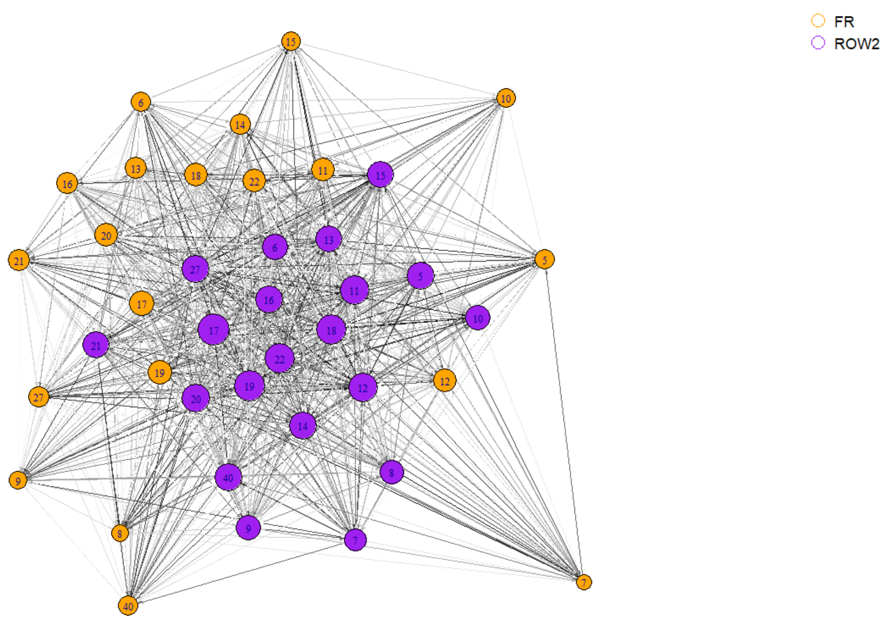


(b) KIO

Figure 14: France

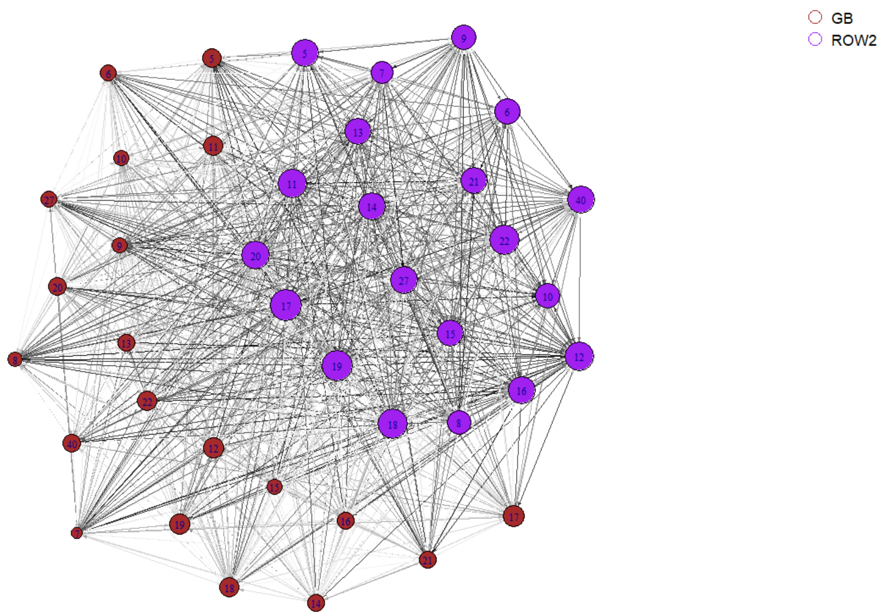


(a) PIO

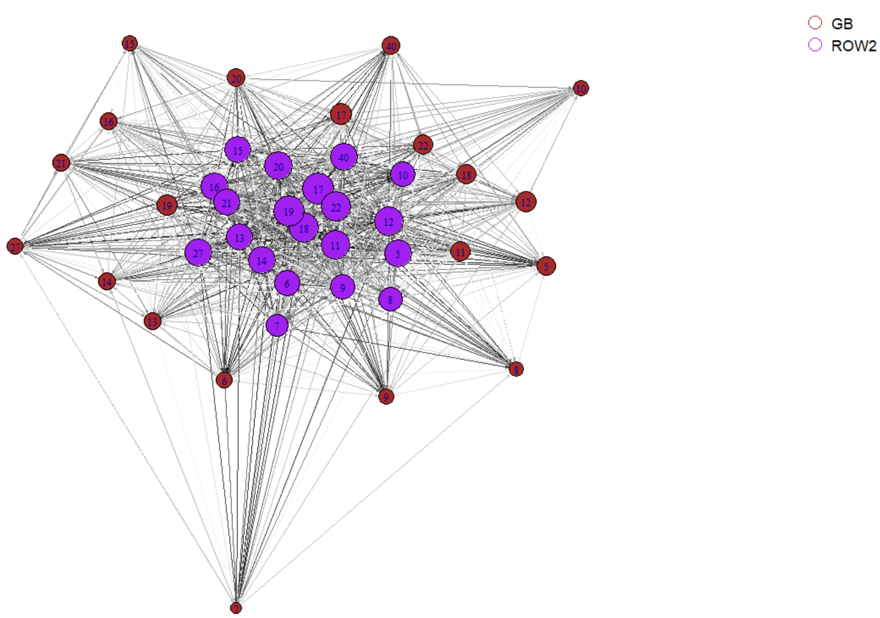


(b) KIO

Figure 15: United Kingdom

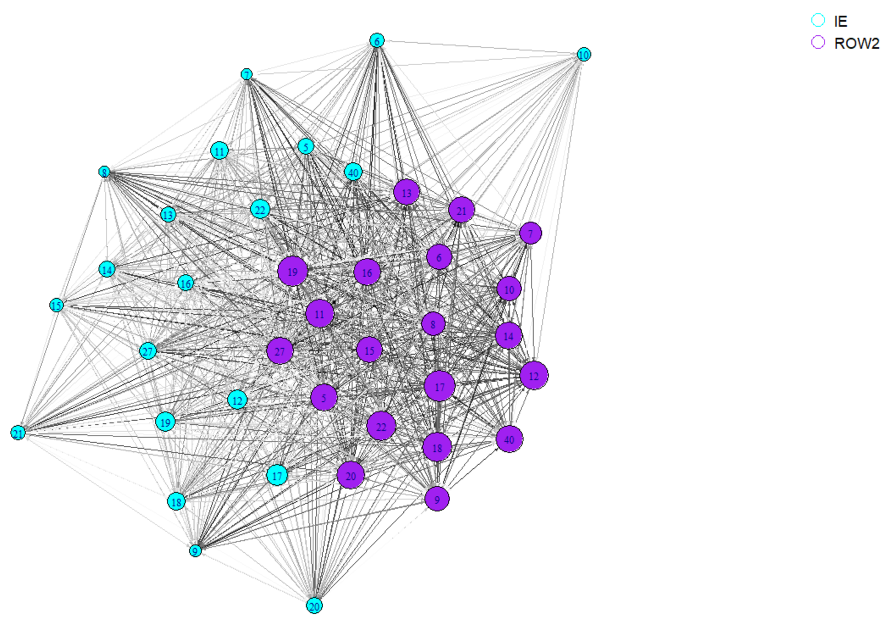


(a) PIO

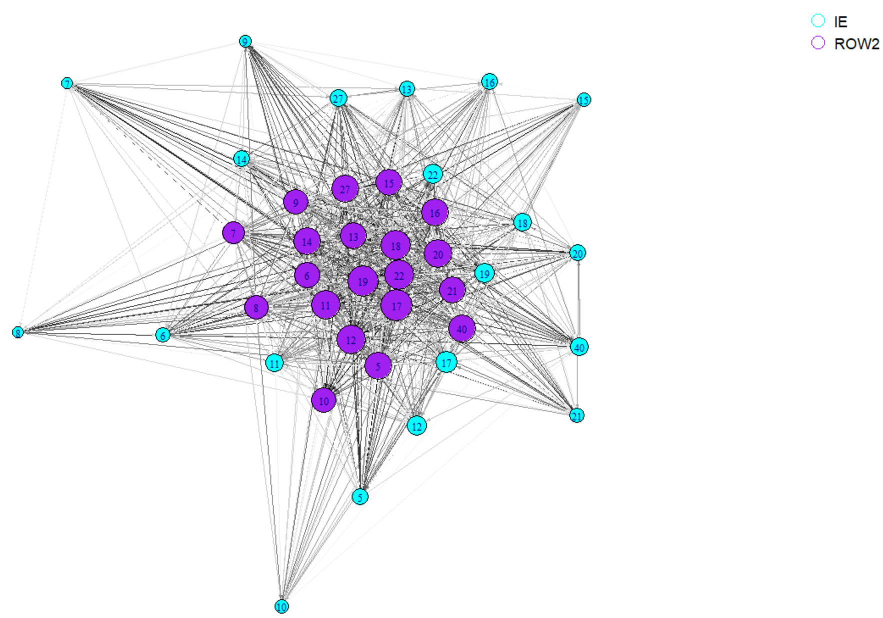


(b) KIO

Figure 16: Ireland



(a) PIO



(b) KIO

Table 2: Country Rankings

Rank	PIO			KIO		
	Total Sales	Purchases	Sales	Patents	Backward	Forward
1	China	China	China	Japan	USA	USA
2	USA	ROW	ROW	China	ROW	Germany
3	ROW	USA	USA	USA	Germany	ROW
4	Japan	Japan	Japan	South Korea	Japan	Japan
5	Germany	Germany	Germany	Germany	France	France
6	Italy	South Korea	South Korea	Taiwan	China	China
7	France	Italy	Italy	ROW	South Korea	Switzerland
8	South Korea	India	India	France	Switzerland	Netherlands
9	India	France	France	Canada	Netherlands	South Korea
10	UK	Spain	Spain	Switzerland	Sweden	Sweden
11	Brazil	UK	UK	Netherlands	Italy	Italy
12	Spain	Brazil	Brazil	Italy	Taiwan	Taiwan
13	Canada	Taiwan	Taiwan	Sweden	Canada	Canada
14	Mexico	Mexico	Russia	Finland	Finland	Finland
15	Russia	Russia	Mexico	Australia	Belgium	Belgium
16	Australia	Canada	Canada	Austria	Austria	Denmark
17	Taiwan	Australia	Australia	Belgium	Denmark	Austria
18	Indonesia	Turkey	Netherlands	India	Spain	Australia
19	Netherlands	Netherlands	Turkey	Denmark	Ireland	Spain
20	Turkey	Indonesia	Indonesia	Spain	Australia	Ireland
21	Switzerland	Switzerland	Belgium	UK	Norway	Norway
22	Belgium	Belgium	Switzerland	Norway	India	UK
23	Poland	Poland	Poland	Ireland	UK	India
24	Sweden	Austria	Sweden	Russia	Luxembourg	Luxembourg
25	Austria	Sweden	Austria	Brazil	Russia	Russia
26	Ireland	Czech Rep.	Czech Rep.	Luxembourg	Poland	Brazil
27	Finland	Finland	Finland	Turkey	Czech Rep.	Hungary
28	Czech Rep.	Portugal	Ireland	Poland	Turkey	Czech Rep.
29	Norway	Ireland	Portugal	Czech Rep.	Brazil	Poland
30	Denmark	Hungary	Norway	Mexico	Hungary	Mexico
31	Portugal	Denmark	Denmark	Hungary	Mexico	Turkey
32	Greece	Norway	Hungary	Slovenia	Portugal	Portugal
33	Romania	Slovak Rep.	Slovak Rep.	Portugal	Cyprus	Cyprus
34	Hungary	Romania	Romania	Cyprus	Slovenia	Slovenia
35	Slovak Rep.	Greece	Greece	Bulgaria	Bulgaria	Bulgaria
36	Slovenia	Slovenia	Slovenia	Romania	Romania	Romania
37	Bulgaria	Bulgaria	Bulgaria	Malta	Malta	Malta
38	Croatia	Croatia	Luxembourg	Slovak Rep.	Slovak Rep.	Croatia
39	Lithuania	Luxembourg	Lithuania	Estonia	Estonia	Slovak Rep.
40	Luxembourg	Lithuania	Croatia	Croatia	Croatia	Estonia
41	Latvia	Latvia	Latvia	Latvia	Lithuania	Indonesia
42	Estonia	Estonia	Estonia	Lithuania	Indonesia	Lithuania
43	Cyprus	Cyprus	Cyprus	Indonesia	Latvia	Latvia
44	Malta	Malta	Malta	Greece	Greece	Greece
Total	619.20	244.02	244.02	12.77	506.33	506.33

Notes: Total in billions constant USD for PIO; millions for KIO.

Table 3: Sector Rankings

	PIO			KIO		
Rank	Total Sales	Purchases	Sales	Patents	Backward Cites	Forward Cites
1	Construction	Construction	Metals	Electronics	Electronics	Electronics
2	Food & tobacco	Motor Vehicles	Chemicals	Machinery	Furniture	Furniture
3	Motor Vehicles	Electronics	Electronics	Electrical equipment	Machinery	Machinery
4	Metals	Chemicals	Fabricated Metals	Furniture	Chemicals	Chemicals
5	Electronics	Metals	Motor Vehicles	Chemicals	Electrical equipment	Electrical equipment
6	Chemicals	Machinery	Other non-metallic	Pharmaceuticals	Pharmaceuticals	Pharmaceuticals
7	Coke & Petroleum	Food & tobacco	Machinery	Motor Vehicles	Motor Vehicles	Motor Vehicles
8	Machinery	Textiles	Food & tobacco	Food & tobacco	Programming	Programming
9	Textiles	Fabricated Metals	Rubber & Plastics	Programming	Food & tobacco	Food & tobacco
10	Fabricated Metals	Electrical equipment	Construction	Fabricated Metals	Fabricated Metals	Fabricated Metals
11	Electrical equipment	Rubber & Plastics	Coke & Petroleum	Other non-metallic	Other non-metallic	Other non-metallic
12	Programming	Other Transport	Textiles	Construction	Rubber & Plastics	Rubber & Plastics
13	Rubber & Plastics	Other non-metallic	Electrical equipment	Rubber & Plastics	Textiles	Textiles
14	Other non-metallic	Furniture	Wood	Metals	Construction	Construction
15	Furniture	Paper	Paper	Other Transport	Other Transport	Other Transport
16	Other Transport	Coke & Petroleum	Other Transport	Textiles	Metals	Metals
17	Pharmaceuticals	Pharmaceuticals	Pharmaceuticals	Recorded Media	Recorded Media	Recorded Media
18	Paper	Wood	Programming	Coke & Petroleum	Coke & Petroleum	Coke & Petroleum
19	Wood	Programming	Furniture	Paper	Paper	Paper
20	Recorded Media	Recorded Media	Recorded Media	Wood	Wood	Wood
Total	619.20	244.02	244.02	12.77	506.33	506.33

Notes: Total in billions constant USD for PIO; millions for KIO.

Table 4: Concentration of Sales and Innovation

		PIO		KIO		
	Total Sales	Purchases	Sales	Patents	Backward	Forward
Top 5 Countries	14.84	14.66	14.55	90.92	89.17	90.84
Top 10 Countries	77.58	79.92	80.04	96.79	96.77	96.91
Top 5 Sectors	50.31	52.61	48.2	77.74	88.1	88.13
Top 10 Sectors	76.52	80.59	74.93	92.19	97.87	97.92

Notes: 44 countries (including ROW); 20 sectors.

Table 5: Intra-Country and Intra-Sector Connections

	PIO			KIO		
	Same Sector	Different Sector	*Total*	Same Sector	Different Sector	*Total*
Same Country	35.1	42.7	77.8	27.3	5.7	33
Different Country	9.9	12.4	22.3	54.5	12.5	67
Total	45	55.1		81.8	18.2	

Table 6: Node Size and Edge Weights Correlations

	(1)	(2)	(3)	(4)	(5)	(6)
		Levels			IHS	
	Total Sales	Int. Sales	Purchases	Total Sales	Int. Sales	Purchases
Patents	2.192*			0.316***		
	(1.184)			(0.0847)		
Forward Cites		0.0125			0.214***	
		(0)			(0.0551)	
Bckward Cites			0.0123**			0.219***
			(0.00541)			(0.0643)
Constant	671,824***	270,101	270,196***	10.40***	9.454***	9.329***
	(9,107)	(0)	(538.7)	(0.493)	(0.430)	(0.525)
Observations	880	880	880	880	880	880
Adj. R-squared	0.615	0.557	0.562	0.725	0.753	0.744

Notes: Robust standard errors clustered by country and sector in parentheses. *** p¡0.01, ** p¡0.05, * p¡0.1.

Table 7: Bilateral Edge Weight Correlations

	(1)	(2)	(3)	(4)
	Levels		IHS	
	With Loops	No Loops	With Loops	No Loops
Forward Citations	0.0147***	0.00497***	0.553***	0.396***
	(0.000196)	(9.79e-05)	(0.00141)	(0.00148)
Constant	305.5***	41.66***	1.204***	1.125***
	(15.59)	(0.587)	(0.00180)	(0.00157)
Observations	774,400	718,960	774,400	718,960
Adj. R-sq	0.012	0.043	0.522	0.556

Notes: Dependent variable is sales of intermediates. Robust standard errors clustered by country and sector in parentheses. *** p¡0.01, ** p¡0.05, * p¡0.1.

Table 8: Centrality Measure Correlations in PIO and KIO

PIO

	Degree In	Degree Out	Katz In	Katz Out	Closeness In	Closeness Out	Betweenness	PageRank
Degree In	1							
Degree Out	0.7473	1						
Katz In	0.5498	0.8223	1					
Katz Out	0.8273	0.6535	0.6554	1				
Closeness In	0.1595	0.2979	0.1205	0.0571	1			
Closeness Out	0.0364	0.0018	-0.0054	0.0002	-0.0099	1		
Betweenness	-0.0013	-0.0203	-0.0096	-0.0078	-0.0115	0.6668	1	
PageRank	0.7382	0.3839	0.1686	0.4132	0.1089	0.1376	0.0495	1

KIO

	Degree In	Degree Out	Katz In	Katz Out	Closeness In	Closeness Out	Betweenness	PageRank
Degree In	1							
Degree Out	0.9796	1						
Katz In	0.9021	0.9134	1					
Katz Out	0.9155	0.8945	0.9863	1				
Closeness In	0.0261	0.0219	0.0108	0.0133	1			
Closeness Out	0.0252	0.02	0.0119	0.015	0.7167	1		
Betweenness	-0.0013	-0.0064	-0.0011	0.0023	0.1317	0.1534	1	
PageRank	0.9707	0.9669	0.8271	0.8278	0.0357	0.0344	-0.0087	1

Table 9: Centrality Country Rankings

PIO								
Rank	Degree In	Degree Out	Katz In	Katz Out	Closeness In	Closeness Out	Betweenness	PageRank
1	China	China	China	China	ROW	France	Mexico	ROW
2	ROW	ROW	ROW	ROW	Germany	Mexico	Cyprus	USA
3	USA	USA	Japan	USA	UK	Cyprus	Malta	China
4	Japan	Japan	USA	Japan	Netherlands	Hungary	Switzerland	Germany
5	Germany	Germany	Korea	Korea	France	Denmark	Luxembourg	Italy
6	Korea	Korea	Germany	Germany	Belgium	Taiwan	Indonesia	France
7	Italy	Italy	Taiwan	Italy	China	Japan	Latvia	Spain
8	India	India	Italy	India	USA	Brazil	France	UK
9	France	France	India	Taiwan	Sweden	Malta	Hungary	Japan
10	Spain	Spain	France	France	Spain	Lithuania	Lithuania	Russia

KIO								
Rank	Degree In	Degree out	Katz In	Katz out	Closeness In	Closeness out	Betweenness	PageRank
1	USA	USA	USA	ROW	Belgium	UK	Ireland	USA
2	ROW	Germany	ROW	USA	UK	Netherlands	Switzerland	Germany
3	Germany	ROW	Germany	Germany	ROW	ROW	India	ROW
4	Japan	Japan	Japan	Japan	Switzerland	Ireland	France	China
5	France	France	France	France	Canada	Belgium	UK	Japan
6	China	China	Netherlands	China	Netherlands	Canada	ROW	France
7	Korea	Switzerland	Korea	Korea	Australia	Germany	Austria	Korea
8	Switzerland	Netherlands	China	Netherlands	Sweden	Switzerland	Australia	Switzerland
9	Netherlands	Korea	Sweden	Switzerland	Taiwan	Australia	Netherlands	Netherlands
10	Sweden	Sweden	Switzerland	Taiwan	Ireland	France	Latvia	Sweden

Table 10: Centrality Sector Rankings: Degree and Katz

PIO

Rank	Degree In	Degree Out	Katz In	Katz Out
1	Construction	Metals	Metals	Construction
2	Motor Vehicles	Chemicals	Chemicals	Comp & Electronics
3	Comp & Electronics	Comp & Electronics	Comp & Electronics	Motor Vehicles
4	Chemicals	Fabricated Metals	Coke & Petroleum	Metals
5	Metals	Motor Vehicles	Textiles	Machinery
6	Machinery	Other non-metallic	Machinery	Electrical Eqip
7	Food& tobacco	Machinery	Fabricated Metals	Textiles
8	Textiles	Food& tobacco	Motor Vehicles	Chemicals
9	Fabricated Metals	Rubber & Plastics	Food& tobacco	Fabricated Metals
10	Electrical Eqip	Construction	Other non-metallic	Food& tobacco

KIO

Rank	Degree In	Degree Out	Katz In	Katz Out
1	Comp & Electronics	Comp & Electronics	Comp & Electronics	Comp & Electronics
2	Furniture	Furniture	Furniture	Furniture
3	Machinery	Machinery	Machinery	Machinery
4	Chemicals	Chemicals	Electrical Eqip	Electrical Eqip
5	Electrical Eqip	Electrical Eqip	Chemicals	Chemicals
6	Coke & Petroleum	Coke & Petroleum	Computer Programming	Computer Programming
7	Motor Vehicles	Motor Vehicles	Coke & Petroleum	Coke & Petroleum
8	Computer Programming	Computer Programming	Motor Vehicles	Motor Vehicles
9	Food& tobacco	Food& tobacco	Fabricated Metals	Fabricated Metals
10	Fabricated Metals	Fabricated Metals	Food& tobacco	Food& tobacco

Table 11: Centrality Sector Rankings: Closeness, Betweenness, and PageRank

PIO

Rank	Closeness In	Closeness Out	Betweenness	PageRank
1	Chemicals	Construction	Construction	Construction
2	Metals	Computer Programming	Other Transport	Motor Vehicles
3	Coke & Petroleum	Other Transport	Coke & Petroleum	Food& tobacco
4	Machinery	Coke & Petroleum	Motor Vehicles	Comp & Electronics
5	Other non-metallic	Motor Vehicles	Wood	Machinery
6	Comp & Electronics	Recorded Media	Coke & Petroleum	Textiles
7	Fabricated Metals	Wood	Metals	Other Transport
8	Electrical Eqip	Paper	Computer Programming	Chemicals
9	Rubber & Plastics	Textiles	Recorded Media	Metals
10	Coke & Petroleum	Metals	Food& tobacco	Electrical Eqip

KIO

Rank	Closeness In	Closeness Out	Betweenness	PageRank
1	Comp & Electronics	Comp & Electronics	Chemicals	Comp & Electronics
2	Coke & Petroleum	Furniture	Recorded Media	Machinery
3	Furniture	Coke & Petroleum	Comp & Electronics	Furniture
4	Electrical Eqip	Chemicals	Other non-metallic	Chemicals
5	Chemicals	Machinery	Food& tobacco	Coke & Petroleum
6	Machinery	Electrical Eqip	Rubber & Plastics	Electrical Eqip
7	Other non-metallic	Other non-metallic	Furniture	Motor Vehicles
8	Fabricated Metals	Rubber & Plastics	Computer Programming	Food& tobacco
9	Rubber & Plastics	Computer Programming	Textiles	Construction
10	Computer Programming	Fabricated Metals	Machinery	Fabricated Metals

Table 12: Comparing PIO and KIO Centrality Measures

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
PIO Measure of:	Degree In	Degree Out	Katz In	Katz Out	Closeness In	Closeness Out	Betweenness	PageRank
KIO Measure	0.0256*	0.0259*	0.104*	0.0837**	-0.00136*	1.763	0.0481	0.0230
	(0.0141)	(0.0135)	(0.0602)	(0.0416)	(0.000773)	(1.753)	(0.0603)	(0.0234)
Constant	0.00111***	0.00111***	0.00102***	0.00104***	0.00115***	-0.000872	0.00108**	0.00111***
	(7.77e-05)	(7.38e-05)	(7.71e-05)	(5.80e-05)	(1.04e-06)	(0.00167)	(0.000514)	(7.69e-05)
Obs	880	880	880	880	861	856	880	880
Adj. R-sq	0.562	0.557	0.320	0.460	0.208	0.035	0.002	0.416

Notes: Robust standard errors in parenthesis. All specifications include country and sector fixed effects. $***$ p¡0.01, $**$ p¡0.05, $*p < 0.1$

Table 13: Top 40 Most Central Country-Sectors: PIO

Rank	Degree In	Degree Out	Katz In	Katz Out	PageRank
1	China: Construction	China: Metals	China: Metals	China: Comps & Elec	ROW: Construction
2	ROW: Construction	China: Chemicals	China: Chemicals	China: Construction	Germany: Motor Vehicles
3	China: Comps & Elec	ROW: Metals	China: Comps & Elec	China: Metals	USA: Motor Vehicles
4	China: Metals	China: Comps & Elec	China: Coke & Petroleum	China: Machinery	China: Construction
5	China: Machinery	China: Other non-metallic	China: Textiles	China: Motor Vehicles	Australia: Construction
6	China: Textiles	China: Textiles	China: Machinery	China: Electrical Equip	France: Construction
7	China: Chemicals	Japan: Metals	China: Food & tobacco	China: Textiles	UK: Construction
8	China: Motor Vehicles	China: Machinery	China: Other non-metallic	China: Fabricated Metals	USA: Food & tobacco
9	China: Electrical Equip	China: Electrical Equip	China: Rubber & Plastics	China: Chemicals	ROW: Motor Vehicles
10	USA: Construction	ROW: Comps & Elec	ROW: Metals	ROW: Construction	Spain: Construction
11	USA: Motor Vehicles	China: Motor Vehicles	China: Electrical Equip	ROW: Comps & Elec	China: Comps & Elec
12	Japan: Motor Vehicles	ROW: Chemicals	China: Fabricated Metals	China: Other Transport	Italy: Construction
13	ROW: Comps & Elec	USA: Fabricated Metals	ROW: Comps & Elec	China: Rubber & Plastics	ROW: Textiles
14	USA: Food & tobacco	China: Food & tobacco	China: Motor Vehicles	China: Other non-metallic	ROW: Comps & Elec
15	China: Food & tobacco	USA: Chemicals	Japan: Metals	Japan: Motor Vehicles	China: Motor Vehicles
16	Japan: Construction	China: Rubber & Plastics	ROW: Chemicals	China: Food & tobacco	Norway: Construction
17	ROW: Motor Vehicles	China: Fabricated Metals	Japan: Motor Vehicles	USA: Motor Vehicles	USA: Construction
18	Germany: Motor Vehicles	Japan: Motor Vehicles	China: Furniture	USA: Construction	Germany: Construction
19	Japan: Metals	USA: Metals	China: Wood	ROW: Motor Vehicles	USA: Other Transport
20	China: Fabricated Metals	China: Coke & Petroleum	Japan: Comps & Elec	USA: Food & tobacco	Russia: Food & tobacco
21	USA: Chemicals	ROW: Other non-metallic	Korea: Comps & Elec	ROW: Machinery	Germany: Machinery
22	ROW: Machinery	Japan: Chemicals	Taiwan: Comps & Elec	Korea: Comps & Elec	Japan: Motor Vehicles
23	China: Rubber & Plastics	USA: Comps & Elec	China: Paper	ROW: Textiles	Brazil: Construction
24	China: Other non-metallic	Japan: Comps & Elec	USA: Chemicals	Japan: Metals	ROW: Machinery
25	ROW: Textiles	China: Wood	USA: Metals	ROW: Electrical Equip	India: Construction
26	USA: Machinery	USA: Motor Vehicles	Japan: Chemicals	Japan: Comps & Elec	Mexico: Construction
27	ROW: Metals	USA: Food & tobacco	USA: Fabricated Metals	ROW: Metals	China: Machinery
28	Spain: Construction	ROW: Machinery	ROW: Machinery	Germany: Motor Vehicles	ROW: Other Transport
29	ROW: Electrical Equip	Japan: Fabricated Metals	ROW: Other non-metallic	Taiwan: Comps & Elec	China: Textiles
30	Italy: Construction	Germany: Motor Vehicles	USA: Comps & Elec	Japan: Construction	ROW: Food & tobacco
31	Japan: Comps & Elec	Korea: Comps & Elec	Korea: Chemicals	China: Coke & Petroleum	Netherlands: Construction
32	Korea: Comps & Elec	ROW: Motor Vehicles	ROW: Electrical Equip	China: Furniture	Russia: Motor Vehicles
33	USA: Fabricated Metals	ROW: Construction	USA: Food & tobacco	USA: Machinery	USA: Machinery
34	ROW: Chemicals	USA: Rubber & Plastics	ROW: Fabricated Metals	USA: Chemicals	Switzerland: Pharmaceuticals
35	Japan: Chemicals	ROW: Textiles	USA: Motor Vehicles	USA: Comps & Elec	USA: Chemicals
36	UK: Construction	ROW: Fabricated Metals	ROW: Textiles	China: Wood	Germany: Food & tobacco
37	France: Construction	ROW: Electrical Equip	ROW: Coke & Petroleum	ROW: Chemicals	Cyprus: Construction
38	USA: Comps & Elec	Korea: Chemicals	ROW: Rubber & Plastics	China: Paper	China: Other Transport
39	Australia: Construction	USA: Coke & Petroleum	ROW: Motor Vehicles	China: Programming	Portugal: Construction
40	India: Construction	Japan: Rubber & Plastics	Korea: Metals	USA: Fabricated Metals	Spain: Food & tobacco

Table 14: Top 40 Most Central Country-Sectors: KIO

Rank	Degree In	Degree Out	Katz In	Katz Out	PageRank
1	ROW: Comps & Elec	USA: Comps & Elec	USA: Comps & Elec	ROW: Comps & Elec	ROW: Comps & Elec
2	USA: Comps & Elec	ROW: Comps & Elec	ROW: Comps & Elec	USA: Comps & Elec	USA: Comps & Elec
3	Germany: Comps & Elec	USA: Furniture	Germany: Comps & Elec	Germany: Comps & Elec	Germany: Machinery
4	USA: Furniture	Germany: Machinery	Japan: Comps & Elec	USA: Furniture	USA: Furniture
5	Germany: Machinery	Germany: Comps & Elec	USA: Furniture	Germany: Machinery	Germany: Comps & Elec
6	ROW: Furniture	ROW: Furniture	Germany: Machinery	ROW: Furniture	ROW: Furniture
7	USA: Machinery	Japan: Comps & Elec	ROW: Furniture	Japan: Comps & Elec	USA: Machinery
8	ROW: Machinery	USA: Machinery	USA: Machinery	USA: Machinery	Germany: Chemicals
9	Germany: Chemicals	Germany: Chemicals	ROW: Machinery	ROW: Machinery	ROW: Machinery
10	Germany: Furniture	ROW: Machinery	France: Comps & Elec	France: Comps & Elec	Germany: Motor Vehicles
11	Germany: Motor Vehicles	Germany: Furniture	Germany: Furniture	China: Comps & Elec	USA: Pharmaceuticals
12	Japan: Comps & Elec	Germany: Motor Vehicles	Korea: Comps & Elec	Germany: Furniture	Germany: Furniture
13	Germany: Electrical Equip	Japan: Machinery	Japan: Machinery	Korea: Comps & Elec	USA: Chemicals
14	France: Comps & Elec	Germany: Electrical Equip	Netherlands: Comps & Elec	USA: Programming	ROW: Pharmaceuticals
15	ROW: Pharmaceuticals	USA: Pharmaceuticals	Germany: Chemicals	Netherlands: Comps & Elec	Germany: Electrical Equip
16	USA: Pharmaceuticals	Japan: Furniture	USA: Programming	Germany: Motor Vehicles	Japan: Comps & Elec
17	China: Comps & Elec	USA: Electrical Equip	Sweden: Comps & Elec	Germany: Chemicals	China: Machinery
18	Japan: Machinery	ROW: Pharmaceuticals	Germany: Motor Vehicles	Taiwan: Comps & Elec	China: Comps & Elec
19	Germany: Pharmaceuticals	France: Comps & Elec	USA: Electrical Equip	Germany: Electrical Equip	Germany: Pharmaceuticals
20	USA: Electrical Equip	Germany: Pharmaceuticals	China: Comps & Elec	USA: Electrical Equip	USA: Electrical Equip
21	Japan: Furniture	USA: Chemicals	Finland: Comps & Elec	Sweden: Comps & Elec	China: Food & tobacco
22	Korea: Comps & Elec	Japan: Electrical Equip	Germany: Electrical Equip	ROW: Electrical Equip	ROW: Chemicals
23	ROW: Electrical Equip	Korea: Comps & Elec	Japan: Furniture	Japan: Machinery	Japan: Machinery
24	USA: Chemicals	USA: Programming	Canada: Comps & Elec	Canada: Comps & Elec	China: Chemicals
25	USA: Programming	Netherlands: Comps & Elec	Taiwan: Comps & Elec	Switzerland: Comps & Elec	USA: Food & tobacco
26	Netherlands: Comps & Elec	ROW: Electrical Equip	USA: Pharmaceuticals	Finland: Comps & Elec	France: Comps & Elec
27	ROW: Chemicals	Japan: Chemicals	ROW: Electrical Equip	USA: Pharmaceuticals	USA: Textiles
28	France: Machinery	China: Comps & Elec	Switzerland: Comps & Elec	ROW: Pharmaceuticals	USA: Programming
29	Switzerland: Furniture	Japan: Motor Vehicles	ROW: Pharmaceuticals	ROW: Programming	ROW: Electrical Equip
30	USA: Motor Vehicles	USA: Motor Vehicles	ROW: Programming	Japan: Furniture	USA: Construction
31	China: Machinery	France: Machinery	Japan: Electrical Equip	USA: Chemicals	Korea: Comps & Elec
32	Sweden: Comps & Elec	ROW: Chemicals	USA: Chemicals	France: Machinery	China: Construction
33	Taiwan: Comps & Elec	Sweden: Comps & Elec	Germany: Pharmaceuticals	Germany: Pharmaceuticals	USA: Motor Vehicles
34	Japan: Electrical Equip	France: Furniture	France: Machinery	Switzerland: Furniture	Germany: Fabricated Metals
35	France: Furniture	Switzerland: Furniture	Italy: Comps & Elec	Italy: Comps & Elec	France: Machinery
36	Switzerland: Comps & Elec	Finland: Comps & Elec	France: Furniture	ROW: Chemicals	France: Chemicals
37	China: Food & tobacco	Taiwan: Comps & Elec	Switzerland: Furniture	China: Machinery	Japan: Furniture
38	France: Chemicals	Switzerland: Comps & Elec	USA: Motor Vehicles	USA: Motor Vehicles	USA: Fabricated Metals
39	Japan: Chemicals	China: Food & tobacco	Japan: Chemicals	France: Furniture	Germany: Rubber & Plastics
40	ROW: Programming	France: Chemicals	Japan: Motor Vehicles	Netherlands: Machinery	USA: Other non-metallic

Table 15: Intangible value-added and Patenting

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Stock	0.228***	2.750***	2.770***	2.749***	3.018**	2.862*	2.861*
	(0.0847)	(0.884)	(0.896)	(0.892)	(1.432)	(1.466)	(1.466)
Patent Flow		-16.15***	-16.03***	-15.95***	-17.07**	-16.64**	-16.63**
		(5.498)	(5.541)	(5.538)	(7.716)	(7.938)	(7.938)
Forward Cites			-0.00272		-0.0340	-0.0118	-0.0124
			(0.00203)		(0.0894)	(0.0913)	(0.0914)
Backward Cites				-0.00235	0.0280	0.00845	0.00904
				(0.00180)	(0.0798)	(0.0813)	(0.0815)
Tax Haven						1,002***	1,218***
						(335.8)	(428.6)
Population	53.36***	59.34***	67.51***	67.44***	64.84***	67.91***	67.99***
	(12.89)	(12.10)	(11.03)	(10.81)	(11.14)	(11.42)	(11.44)
Area	0.00529***	0.00509***	0.00478***	0.00479***	0.00480***	0.00490***	0.00492***
	(0.000827)	(0.000784)	(0.000885)	(0.000892)	(0.000901)	(0.000963)	(0.000966)
GDP per capita							-0.0140
							(0.00888)
Constant	-797.7**	-639.9**	-701.9**	-698.3**	-718.5**	-912.2***	-489.0**
	(332.2)	(298.0)	(287.8)	(286.6)	(303.2)	(319.3)	(234.5)
Obs	809	809	809	809	809	809	809
Adj R-sqr	0.648	0.677	0.677	0.677	0.677	0.643	0.643

Notes: Robust standard errors in parentheses. All specifications include year and sector fixed effects. $***$ p¡0.01, $**$ p¡0.05, $*p < 0.1$

Table 16: Intangible value-added and Patenting (Logs)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Stock	0.163***	0.0346	0.126*	0.103	0.154**	0.157**	0.117
	(0.0213)	(0.0714)	(0.0757)	(0.0753)	(0.0767)	(0.0768)	(0.0806)
Patent Flow		0.154*	0.223**	0.280***	0.299***	0.326***	0.329***
		(0.0831)	(0.0922)	(0.103)	(0.104)	(0.107)	(0.106)
Forward Cites			-0.115**		-0.0877*	-0.0792	-0.0853*
			(0.0448)		(0.0493)	(0.0497)	(0.0502)
Backward Cites				-0.134**	-0.0980	-0.114*	-0.104
				(0.0571)	(0.0633)	(0.0645)	(0.0660)
Tax Haven						-0.272*	-0.289*
						(0.162)	(0.161)
Population	0.611***	0.606***	0.606***	0.606***	0.606***	0.598***	0.634***
	(0.0571)	(0.0564)	(0.0560)	(0.0564)	(0.0561)	(0.0564)	(0.0601)
Area	0.267***	0.264***	0.254***	0.244***	0.242***	0.192***	0.189***
	(0.0504)	(0.0500)	(0.0505)	(0.0511)	(0.0511)	(0.0560)	(0.0560)
GDP per capita							0.197
							(0.133)
Constant	0.915*	1.067**	1.130**	1.307***	1.292***	1.877***	-0.0233
	(0.475)	(0.477)	(0.480)	(0.490)	(0.487)	(0.572)	(1.448)
Obs	809	809	809	809	809	809	809
Adj R-sqr	0.740	0.740	0.742	0.742	0.742	0.743	0.743

Notes: All non-binary variables logged excepting the innovation variables where the IHS is used. Robust standard errors in parentheses. All specifications include year and sector fixed effects. $***$ p¡0.01, $**$ p¡0.05, $*p < 0.1$