



**Maestría en Explotación de Datos y  
Descubrimiento del Conocimiento**



Facultad de Ciencias Exactas y Naturales  
Facultad de Ingeniería

Facultad de Ciencias Exactas y Naturales - UBA

Maestría en Explotación de Datos y Descubrimiento del  
Conocimiento (DM&KD)

### **Trabajo Final de Especialización**

AUTOR

**Sofía Clara Perini**

TUTOR

**Marcelo Soria**

21 de febrero de 2023

# Cambios en la estructura del espacio de productos de la bioeconomía en las últimas dos décadas

Sofía Clara Perini

21 de febrero de 2023

## Resumen

La estructura productiva de un país tiene suma importancia para su desarrollo, ya que los países en general se mueven a través del espacio de productos desarrollando bienes cercanos a los que producen actualmente. Este trabajo explora cambios en el espacio de productos de la bioeconomía a través del tiempo, haciendo especial énfasis en el caso Argentino. Explotando un conjunto de datos de comercio de las últimas dos décadas, se pretende replicar el estudio de la red de relaciones entre productos propuesto por Hidalgo et al. (2007) y el análisis de redes que plantean Tagliazucchi et al. (2013) para otro caso de uso. Del análisis efectuado, se observó que el espacio de producto en su conjunto es disperso, con una distribución amplia y heterogénea que evidencia presencia de algunos vínculos fuertes y muchos vínculos marginales, donde se percibe mayor presencia de productos manufacturados entre aquellos con mayores conexiones y mayor presencia de agrícolas y/o sin procesar entre los menos conectados. Agrupando los productos por categorías, se pudo corroborar que existe mayor proximidad de productos dentro de las categorías que entre las distintas categorías. Asimismo, la visualización del espacio de producto mostró que la red tiene una estructura de centro-periferia, con dos módulos densamente conectados formados principalmente por Productos manufacturados, Productos químicos, y Alimentos y animales y una periferia formada por el resto de las clases de productos (materias primas, aceites y grasas, bebidas y tabaco, otras manufacturas y combustibles). En la representación del espacio de producto para el caso Argentino se pudo ver que los productos para los que el país presenta ventajas comparativas se ubican más bien en la periferia de la red. Por último, el estudio de los roles que cumple cada producto dentro de las comunidades según las conexiones intra e inter-comunidades demostró que se ha mantenido aproximadamente estable la distribución del número de productos por tipo de rol aunque hay muy pocos productos que preservan el rol de un período a otro. Se probó también que los países tienden a desarrollar ventajas comparativas preferencialmente en bienes cercanos por lo que la estructura del espacio de productos afecta el patrón de especialización de un país. Para promover una transformación estructural de la producción que permita un mayor crecimiento de la bioeconomía argentina, será necesario un esfuerzo mucho más grande si se quiere mover a productos más distantes en el espacio de productos (con mayor nivel de industrialización). No obstante, son precisamente estos saltos los que podrían permitir ese ansiado avance en el desarrollo productivo.

**Keywords** - comercio, espacio de productos, bioeconomía, redes, grafos

## 1. Introducción

Hidalgo et al. (2007) plantean que el espacio de productos condiciona el desarrollo de las naciones [1]. ¿Qué quieren decir con esto? Que la estructura productiva de un país tiene suma importancia para su desarrollo, ya que los países en general se mueven a través del espacio de productos desarrollando bienes cercanos a los que producen actualmente. Es decir, partiendo de la idea que las economías crecen mejorando los productos que generan y exportan, los autores argumentan que los recursos que posee un país (la tecnología, el capital, las instituciones y las habilidades necesarias) para fabricar productos más nuevos se adaptan más fácilmente a partir de algunos productos que de otros.

En este trabajo se procura estudiar, tal como ellos en el paper original, esta red de relaciones entre productos, o espacio de productos para un grupo específico de productos de la Bioeconomía, encontrando que los productos más sofisticados (más elaborados) están ubicados en un núcleo densamente conectado de la red, mientras que los productos menos sofisticados (materias primas u otros productos con poca elaboración) ocupan una periferia menos conectada.

En este sentido, se propone aplicar la metodología del Espacio de Productos propuesta por [1] para identificar cambios en el espacio de productos a través del tiempo y evaluar las posibilidades de diversificación de productos de la bioeconomía Argentina. En particular, se estudian los cambios en la modularidad de la red del espacio de productos para los distintos períodos (1997-99, 2007-09 y 2017-19), y se la compara con una red aleatoria para ver cómo dista el comportamiento del espacio de producto con el de una red con tales características. Asimismo, se estudiarán los roles de los nodos en cada período según diferentes niveles de proximidad.

Cabe destacar, que este documento está basado en los trabajos de Hausmann & Klinger (2006) [2] e Hidalgo et al. (2007) [1] que estudian la red de relación entre productos, o espacio de productos, como también en la definición de productos de la bioeconomía que se establece en Wierny et al. (2015) [3] y el análisis de redes (modularidad, roles) que plantean Tagliazucchi et al. (2013) [4] para otro caso de aplicación (las redes en el cerebro humano).

El objetivo del presente trabajo será replicar algunos de los análisis de [1] y [4] a partir de una muestra de datos de exportaciones mundiales de productos de la bioeconomía. Asimismo, se pretende realizar un abordarje de los diferentes conceptos de teoría de grafos y su aplicación en el espacio de productos. Luego de esta introducción, en la siguiente sección se presenta la metología de análisis implementada en este trabajo. La sección 3 reporta los principales resultados a los cuales se arribó, relacionados al entendimiento del espacio de producto y la medida de proximidad con que se construye, la estructura de las redes, la modularidad, medidas de centralidad de los nodos y sus roles a través de los distintos períodos analizados. Por último, en la sección 4 se esbozan algunas conclusiones del estudio. En el Anexo se adjuntan algunas figuras adicionales, resultantes del estudio que no se incorporaron al cuerpo del documento.

## 2. Metodología de análisis

### 2.1. Bases de Datos

Se utilizaron datos de comercio exterior de fuente [World Integrated Trade Solutions \(WITS\)](#), provenientes de UNCOMTRADE a 4 dígitos del sistema armonizado (SA, o HS, por sus siglas en inglés) para los años 1997 a 1999, 2007 a 2009 y 2017 a 2019. Los 3 períodos de análisis están definidos por un promedio simple del comercio reportado por los países en esos 3 años, distanciados en una década, a saber: promedios 1997-99, 2007-09 y 2017-19. Se decide no utilizar datos de 2020 a 2022 porque al ser años de pandemia y post-pandemia podría tener una comportamiento distinto a la tendencia, y la idea es comparar evolución sin atender a las consecuencias de la pandemia en particular. Podría ser un tema de futura investigación, comparar tendencias pre y post pandemia.

## Selección de Productos

Para definir el universo de productos a abordar en el presente análisis, se descargaron datos de comercio bajo la clasificación (nomenclatura) de productos de WITS, a nivel de partida arancelaria (4 dígitos del sistema armonizado) para 1996 (HS1996 o H1) y se seleccionaron bienes pertenecientes a la bioeconomía de acuerdo a Wierny et. al. (2015) [3], que utilizan la clasificación de las actividades según CIIU Rev3. Para poder trabajar con ambas clasificaciones de productos, se utilizó una tabla de correspondencias entre CIIU Rev3 y HS1996 que se descargó también de WITS. Allí se ofrece información sobre diversas nomenclaturas de productos, lo que contribuye con el mapeo entre varias nomenclaturas distintas. En este caso se descargó la tabla de correlación o correspondencia entre el SA 1996 (o H1) y la ISIC (o CIIU por su sigla en español) Rev 3. En esta tabla los datos de productos del SA se encuentran a 6 dígitos por lo que se realiza un trabajo de conversión adicional a nivel de partida para poder cruzar con los de comercio.

Se excluyeron de la base de datos aquellos productos denominados como código especial de naciones unidas ('UN Special Code'), ya que no poseen distinción de qué productos se trata, quedando 1241 bienes totales con los que se trabajará, de los cuales 584 pertenecen al grupo de la bioeconomía.

## Selección de Países

Dado que no hay información de comercio completa para los 193 países de la base de datos en todos los años a estudiar y con el fin de hacer comparables los análisis por períodos, se decide unificar la cantidad de países en todos los años estudiados. Para eso se optó por tomar sólo los países que poseen información completa para todo el período en análisis. Existen 84 países para los que hay faltante de datos en al menos uno de los períodos considerados. Por lo tanto, se incluyen un total de 109 países para los análisis de este trabajo, para los que efectivamente figuran datos de exportaciones en todos los períodos, y representan en su conjunto aprox. el 82 % de las exportaciones mundiales en 07/09 y 17/19, y 95,7 % en 97/99.

## 2.2. Construcción del Espacio de Productos

Para la construcción del espacio de productos como una red (grafo), donde los productos representan nodos y los enlaces la proximidad entre dichos bienes, se adoptó el enfoque propuesto por [1]. La medida de proximidad utilizada está basada en la idea intuitiva de que la capacidad de un país para producir un producto depende de su capacidad para producir otros. Si dos bienes están relacionados porque requieren similares recursos (instituciones, infraestructura, factores físicos y/o tecnología) tenderán a ser producidos en tandem, mientras que es menos probable que los bienes diferentes se produzcan juntos.

Para incluir todos los productos en la red se generó un Árbol de Expansión Máximo (MST, por sus siglas en inglés) que es el conjunto de N-1 enlaces (siendo N el número de nodos) que conectan todos los nodos de la red y maximiza la suma de las proximidades en ella. Adicionalmente, se consideraron los enlaces más fuertes que no están necesariamente en el MST. Para ello, se incluyeron además del MST todos los enlaces por encima de cierto umbral. Como regla general, según [1] se puede lograr una buena visualización de la red con un grado promedio igual a 4. Esto es cuando el número de enlaces es el doble que el de los nodos, que en nuestro caso consiste en un umbral de 0,53 para los productos de la bioeconomía (un poco inferior al usado en el paper original de 0,55). Se obtuvo una visualización adecuada manteniendo todos los enlaces con un valor de proximidad de 0,53 o mayor. Esto resultó en una red con 584 nodos y aprox. 1100 enlaces (más o menos según el período considerado).

Se probaron distintos valores de proximidad ( $\phi$ ) para generar las redes con diversa densidad de aristas. En el presente trabajo se consideraron proximidades en un rango de  $0,1 \leq \phi \leq 0,85$ . Asimismo, se extrajeron medidas de centralidad promedio, grado máximo, camino mínimo medio y coeficiente de clustering también en función de la proximidad y los distintos períodos considerados. Resulta pertinente destacar que, el camino mínimo se calculó a partir del valor máximo de los promedios de caminos más cortos de los subgrafos que se encontraban fuertemente conectados.

### Cálculo de la Proximidad $\phi$

Se define la proximidad entre los bienes i y j como:

$$\phi_{ij} = \min P(VCR_i \geq 1 | VCR_j \geq 1), P(VCR_j \geq 1 | VCR_i \geq 1)$$

Donde  $VCR_i$  es la Ventaja Comparativa Revelada, que mide si un país c exporta más del bien i, como porcentaje de sus exportaciones totales, que el país “promedio” ( $VCR > 1$ , no  $VCR < 1$ ) y se calcula como:

$$VCR_{c,i} = \frac{\sum_i x(c,i)}{\sum_c x(c,i)}$$

$P(VCR_i \geq 1 | VCR_j \geq 1)$ : es la probabilidad condicional de exportar el bien i dado que se exporta el bien j.

Según los autores al considerar el mínimo entre ambas probabilidades condicionales se resuelve el problema de que en el caso de que un país sea el único exportador de un bien en particular la probabilidad condicional de exportar cualquier otro bien dado ese sería igual a uno para todos los demás bienes exportados por ese país, y asimismo, se logra una matriz de proximidad simétrica.

Simplificando la notación el cálculo de la proximidad quedaría:

$$\phi_{ij} = \min P(i|j), P(j|i)$$

Donde:

$$P(i|j) = \frac{Pconj(i,j)}{P(j)}$$

- $P(i) = I/N$  : probabilidad de exportar i (proporción de exportaciones de i).
- $P(j) = J/N$  : probabilidad de exportar j (proporción de exportaciones de j).
- $Pconj(i,j)$  : probabilidad conjunta de que se exporte i y j al mismo tiempo

Calcular la proximidad en base a la mínima probabilidad conjunta de  $P(i|j)$  y  $P(j|i)$ , es lo mismo que calcular la proximidad en base al que tenga máxima probabilidad individual  $P(i)$  o  $P(j)$  dado que el numerador (probabilidad conjunta) de las probabilidades condicionales es igual para ambas y lo que cambia es el denominador.

Para ver qué bienes están en una parte densa (más conectados con otros) de la matriz de proximidad y cuáles están en la periferia (menos conectados), se calcularon los paths para cada producto.

$$paths_{i,t} = \sum_j \phi_{i,j,t}$$

### Prueba de Difusión

Para probar si es cierto que los países desarrollan ventajas comparativas preferencialmente en bienes cercanos, se utilizó uno de los enfoques del paper original que consiste en medir la proximidad promedio de un nuevo producto potencial j a la estructura productiva actual de un país, lo que llaman densidad y se define como:

$$w_j^k = \frac{\sum_i x_i \phi_{ij}}{\sum_i \phi_{ij}}$$

Donde  $w_j^k$  es la densidad alrededor del bien  $j$  dada la canasta exportadora del  $k$ -ésimo país y  $x_i = 1$  si  $VCR_{ki} > 1$  y 0 en caso contrario. Un valor de densidad alto significa que el  $k$ -ésimo país tiene muchos productos desarrollados que rodean al  $j$ -ésimo producto.

Para estudiar la evolución de las ventajas comparativas, se consideraron productos de transición y productos no desarrollados. Los primeros están definidos como aquellos con  $VCR_{ki} > 0,5$  en el período anterior al considerado (1997-99 o 2007-09) y  $VCR_{ki} > 1$  en dicho período (2007-09 o 2017-19, respectivamente). Los segundos incluyen aquellos bienes que en los 3 períodos estudiados tuvieron una  $VCR_{ki} < 0,5$  y se descartaron los casos que no cumplieran con ninguno de estos dos criterios. Se comparan las distribuciones de las densidades alrededor de ambos grupos de productos, para analizar sus diferencias.

### Visualización de la red

Una buena visualización de la red requiere un diseño adecuado. Por eso se tuvieron en cuenta las consideraciones del paper original. Se diseñó la red utilizando un algoritmo dirigido por fuerza de Fruchterman-Reingold (llamado diseño de resorte o spring\_layout en NetworkX), donde el diseño de los nodos y los bordes se determina automáticamente, realizando una simulación pseudofísica de los movimientos de los nodos, asumiendo que cada borde es un resorte con una distancia de equilibrio fija. Este algoritmo tiende a acercar entre sí grupos de nodos bien conectados, lo que hace que el resultado de la visualización sea más significativo y estéticamente más prolífico [5].

Al usar una representación de red, se puede observar simultáneamente la estructura del espacio y otras covariables. En este caso, se pintó la red utilizando las categorías de productos que provienen de la clasificación SITC Rev2 y el tamaño de los nodos se generó en base al valor de las exportaciones mundiales del producto para el período correspondiente. Para dar una idea de la proximidad de los enlaces involucrados en la representación de red, se armó un código de colores usando rojo oscuro y azul para enlaces fuertes; y gris y azul claro para los más débiles.

### 2.3. Modularidad y detección de Comunidades

Un módulo o comunidad es un subconjunto de nodos en una red que están más densamente conectados entre sí que con el resto de la red, siendo una partición de la red una asignación de todos los nodos en subconjuntos no superpuestos [6]. En este trabajo se utilizó el algoritmo de Louvain para detectar comunidades, el cual permite encontrar las mejores particiones optimizando la modularidad. En este sentido, se estimó el coeficiente de modularidad ( $Q$ ) y el número de comunidades ( $NC$ ) para los distintos períodos estudiados y los diferentes umbrales de proximidad propuestos. Luego se compararon los resultados con aquellos de una red Random.

La fórmula utilizada para el cálculo de la modularidad ( $Q$ ) fue la siguiente:

$$Q = \frac{1}{2L} \sum \delta(i,j) [A_{ij} - \frac{k_i k_j}{2L}]$$

donde  $L$  es el número total de aristas,  $k_i$  es el grado del nodo  $i$ ,  $k_j$  es el grado del nodo  $j$  (número de aristas que tienen como nodo terminal al nodo  $i$  o  $j$  respectivamente) y  $\delta(i,j) = 1$  si los nodos  $i$  y  $j$  pertenecen a la misma comunidad o 0 en caso contrario.

### 2.4. Rol de los Nodos

Es de particular interés identificar aquellos nodos que contribuyen a las conexiones inter-modulares y aquellos que están muy bien conectados solo dentro de sus propias comunidades. Para definir el rol que cum-

ple cada nodo dentro de las comunidades e identificar cuáles son aquellos que cambian de comunidad o de rol, se clasificaron los nodos para cada período según umbrales definidos para el coeficiente de participación ( $P_i$ ), que mide la conectividad inter-comunidades, y el z-score del grado intra-comunidad ( $z_i$ ), que mide conexiones al interior de las comunidades.

Las fórmulas empleadas para el cálculo de ambas métricas son las siguientes:

$$P_i = 1 - \sum_{N_M} \left( \frac{k_i^{U_j}}{k_i} \right)^2$$

donde  $k_i^{U_j}$  es el número de aristas en el nodo  $i$  y la comunidad  $j$ ,  $N_M$  es el número de comunidades y  $k_i$  es el grado total del nodo  $i$ .

$$z_i = \frac{k_i - \langle k \rangle}{\sigma_k}$$

donde  $k_i$  es el grado intra-comunidad, y  $\langle k \rangle$  y  $\sigma_k$  son el promedio y el desvío estándar del grado intra-modular.

Los roles se asignan atendiendo los criterios propuestos en Tagliazucchi y colaboradores (2013)[4], a saber:

- Hubs:  $P_i > P_c$  y  $z_i > z_c$
- Hubs Provinciales:  $P_i < P_c$  y  $z_i > z_c$
- Nodos Provinciales:  $P_i < P_c$  y  $z_i < z_c$
- Nodos Conectores:  $P_i > P_c$  y  $z_i < z_c$

Los umbrales utilizados en este análisis replican aquellos utilizados por los autores mencionados:  $z_c = 1$  y  $P_c = 0,05$ , dado que no se cuenta con otros de referencia.

## 2.5. Test Estadísticos

A fin de identificar si las diferencias de modularidad promedio entre los distintos períodos son significativas para diferentes valores de proximidad, se utilizó un test t para las comparaciones de a pares. Para la prueba t se utilizó el método `ttest_rel` de `scipy.stats`, que permite probar la hipótesis nula de que 2 muestras relacionadas o repetidas tienen valores promedio (esperados) idénticos. Estos tests se realizaron tanto para el coeficiente  $Q$  como para el  $NC$ . Se aplicaron los mismos tests para medir diferencias entre los roles de los nodos en distintos momentos del tiempo.

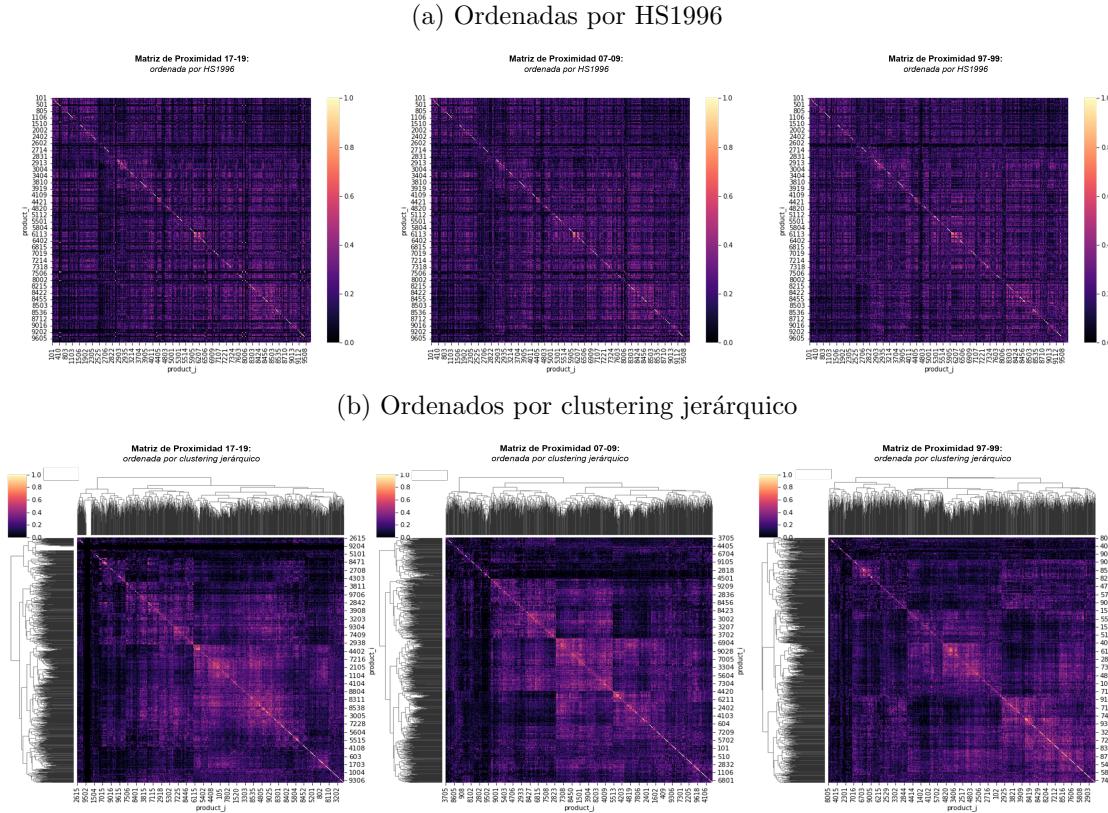
## 3. Resultados

### 3.1. Análisis del Espacio del Producto (Product Space)

El espacio del producto se define como el conjunto de todas las medidas de proximidad. En este trabajo se analizó el espacio de producto construido con datos de flujos comerciales de 3 períodos distintos (1997-99, 2007-09, 2017-19), que constan de una matriz de 1241x1241 para el total de productos (584x584 para los productos de la bioeconomía) cuyas entradas son la proximidad entre productos. Cada fila y columna de esta matriz representa un producto particular y cada elemento fuera de la diagonal representa la proximidad entre un par de productos.

La Figura 1a muestra la matriz de proximidad del total de productos donde las columnas se ordenan usando su clasificación según el código HS1996. La Figura 1b muestra versión agrupada jerárquicamente de la matriz, usando un algoritmo de clustering jerárquico con linkage promedio, revelando su estructura modular y las muchas filas y columnas vacías que pertenecen a productos no comercializados. Estos últimos incluyen alrededor de 24 bienes de las categorías materias primas y productos manufacturados en la matriz de 2017-19.

Figura 1: Visualización de matrices de proximidad para los diferentes períodos



En la Figura 1b un espacio de producto suave y homogéneo implicaría valores uniformes (coloración homogénea) a lo largo de la matriz, mientras que un modelo de escalera de producto sugeriría una matriz con valores altos (o coloración brillante) solo a lo largo de la diagonal [1]. En este caso el espacio de producto parece ser modular, con algunos bienes altamente conectados y otros desconectados. Esto se observa en todos los períodos analizados, aunque los módulos parecen ir cambiando a lo largo del tiempo. En 2017-19 parece haber al menos dos grandes módulos, uno muy grande con mayores conexiones (colores más brillantes) y otro más pequeño con conexiones más débiles, con conexiones no tan fuertes entre módulos; mientras que en 2007-09 se observa un número mayor de módulos y mayores conexiones entre módulos. Por último, en 1997-99 parece haber 3 módulos, con conexiones más o menos altas entre si y bajas entre módulos.

Por otro lado, examinando la correlación entre las matrices de proximidad del total de productos, se pudo notar que existe mayor correlación entre las matrices de 2017-19 y 2007-09 (0.64) que, entre la de estos períodos y la de 1997-99 (0.49 y 0.59, respectivamente). La misma tendencia se observa analizando la matriz reducida de proximidades con solo productos de la bioeconomía, aunque se advierte una correlación levemente superior entre períodos que la observada para el total de productos (0.67, 0.52 y 0.62, respectivamente).

Adicionalmente, se analizó qué bienes están en una parte densa de la matriz para las exportaciones promedio 2017-19 y cuáles están en la periferia. De acuerdo a lo observado por Hausmann & Klinger (2006)

[2], la parte más densa tiende a estar dominada por productos manufacturados mientras que los productos con menor densidad tienden a ser productos agrícolas sin procesar. En este caso, se observa una situación similar con bienes manufacturados entre los más densos ya que incluyen en su mayoría las categorías de Productos manufacturados, Maquinaria y equipo de transporte y Otras manufacturas (correspondientes a los códigos 6, 7 y 8 de SITC, respectivamente). De estos productos (top 30) solo 7 pertenecen a la bioeconomía.

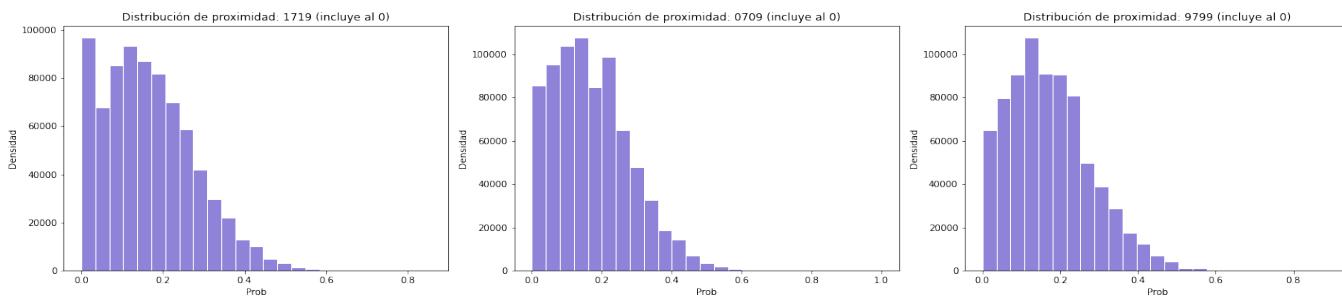
Mientras que se observa mayor presencia de bienes agrícolas y/o sin procesar entre los menos densos (como mate, copra, cerdas o pelos de cerdo o jabalí, productos comestibles de origen animal, perlas, minerales y abonos minerales, entre otros), incluyendo en mayor medida las categorías Materias primas no comestibles y Productos manufacturados (correspondientes a los códigos 2 y 6 de SITC, respectivamente). Cabe destacar, que de los 30 productos con menor densidad, 11 corresponden a bienes de la bioeconomía. Por tanto, se pudo notar mayor presencia de la bioeconomía entre los productos menos conectados, que entre los más conectados.

Estudiando individualmente la matriz de productos de la bioeconomía, se observa la misma tendencia: mayor presencia de productos manufacturados entre aquellos con mayores conexiones y mayor presencia de agrícolas y/o sin procesar entre los menos conectados. Entre los 20 primeros productos con mayor densidad se encuentran Alimentos y animales vivos principalmente para alimentos y Productos Manufacturados y Otras Manufacturas (correspondientes a los códigos 0, 6 y 8 de SITC, respectivamente). Por su parte, entre los menos conectados, se destacan las Materias primas no comestibles (categoría 2 de SITC).

### Distribución de la Proximidad

Al igual que en Hidalgo et al (2007) [1], se observa que el espacio de productos en su conjunto es disperso, con  $\phi_{ij}$  distribuidos según una distribución amplia (ver figura 2) con el 8% de sus elementos iguales a cero para 2017-19, el 28% de ellos menores que 0,1 y el 62% de las entradas tomando valores por debajo de 0,2. El resto de los años presentan distribuciones similares, quizás con algo menos de ceros, como se puede observar en la figura 2, muy similar a la observada por [1] para años anteriores.

Figura 2: Distribución de la proximidad por período analizado



Esta distribución amplia y heterogénea evidencia que el espacio del producto tiene algunos vínculos fuertes y muchos vínculos marginales, que no son significativos y representan el trasfondo de la medida de proximidad. Este número sustancial de conexiones insignificantes requieren una representación de red, lo que nos permite explorar la estructura del espacio del producto junto con la proximidad entre productos y la participación en el comercio mundial (ver más abajo en la sección ??).

Por otro lado, se nota un patrón similar en la proximidad promedio 2017-19 para el total de productos en relación con períodos anteriores. La proximidad media (y mediana) parece haber ido disminuyendo levemente en las últimas dos décadas, pero el modo sigue siendo 0 en todos los casos (ver Cuadro 1). Asimismo, el desvío estándar se mantiene en torno al 0.11 en todos los períodos aunque parece ser algo menor en períodos anteriores. En cuanto a la matriz de proximidades de productos de la bioeconomía, se observa una tendencia similar, con valores levemente más altos de proximidad (media y mediana) que para la matriz total de productos en

cada período analizado y un desvío algo menor.

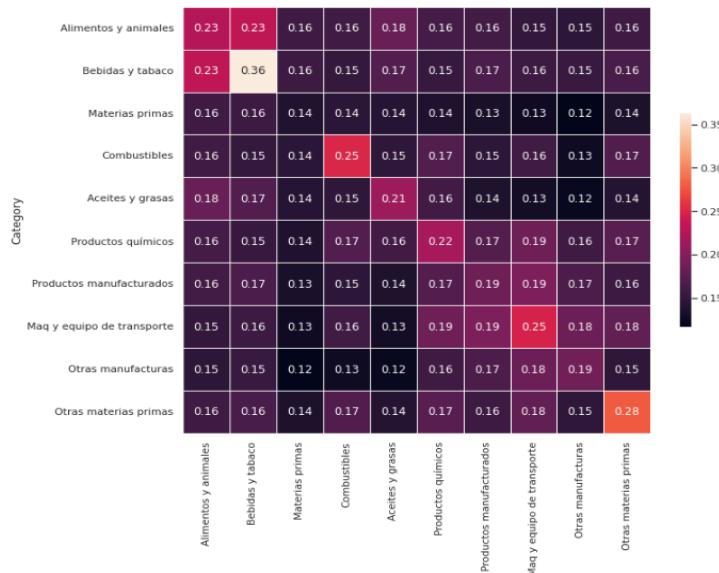
Cuadro 1: Estadísticas descriptivas de la medida de proximidad

products	year	mean	median	mode	std.dev.	obs
total	1719	0.1656	0.1538	0.0	0.1162	1540081
total	0709	0.1703	0.1579	0.0	0.1153	1540081
total	9799	0.1739	0.1613	0.0	0.1116	1540081
bioecon	1719	0.1726	0.1600	0.0	0.1104	170236
bioecon	0709	0.1752	0.1613	0.0	0.1113	170236
bioecon	9799	0.1770	0.1667	0.0	0.1098	170236

### Proximidad promedio intra e inter-grupos

Así como [2] organizan los productos en bloques según la intensidad de sus factores y la proximidad promedio dentro de los bloques y entre los bloques, usando grupos de productos básicos de Leamer (1984), en este caso se calculan los promedios de proximidad para las categorías de productos según la clasificación SITC Rev 2 para el período 2017-19 (ver figura 3). Dicha agrupación sugiere una mayor proximidad dentro de los grupos y una menor proximidad entre los grupos, aunque mayor o menor según el grupo de productos de que se trate.

Figura 3: Detalle proximidad promedio intra e inter-grupos período 2017-19



Por ejemplo, las categorías Alimentos y animales vivos principalmente para alimentos y Bebidas y Tabaco (0 y 1 SITC) presentan **proximidad promedio inter-grupo** bastante alta (0.23) en relación al resto (colores más claros en el mapa de calor). Luego, parece haber categorías de productos manufacturados con proximidades más altas, aunque menores a las recién mencionadas, para Maquinaria y equipo de transporte con Productos químicos (0.19) y con Productos manufacturados (0.19). Las proximidades promedio inter-grupos restantes parecen más bien bajas (colores oscuros, proximidades promedio < 0,18).

En cuanto a la **proximidad promedio intra-grupos**, se puede notar que la mayoría presentan proximidades más altas al interior de las categorías. Algunas de las más destacadas por su alta proximidad intra-grupo

son: Bebidas y tabaco (0.36), Materias primas y transacciones no clasificadas en otra parte (0.28), Maquinaria y equipo de transporte (0.25) y Combustibles minerales, lubricantes y materiales relacionados (0.25). La categoría de Materias primas no comestibles, por su parte, parecería la que menor proximidad intra-grupo presenta (0.14) como también entre grupos. Estas proximidades promedio son similares para los distintos períodos.

Un comportamiento semejante se observa en los restantes períodos analizados, presentando mayor proximidad intra-grupos y menor proximidad inter-grupos (ver comparación de los mismos en el Anexo figura 12).

### 3.2. Visualización del Espacio del Producto

Tal como en el paper original [1], para incluir todos los productos de la bioeconomía (584) en la red se generó un “esqueleto” de la misma, llamado Árbol de Expansión Máximo (MST, por sus siglas en inglés). Este esqueleto es el conjunto de N-1 enlaces (siendo N el número de nodos) que conectan todos los nodos de la red y maximiza la suma de las proximidades en ella. Es decir, que incluye los 583 enlaces que maximizan la proximidad añadida del árbol (ver figura 13 en Anexo). Adicionalmente, se agregaron los enlaces más fuertes que no están necesariamente en el MST. En este sentido, se agregaron todos los enlaces por encima de cierto umbral de proximidad. Se obtuvo una visualización adecuada manteniendo todos los enlaces con un valor de proximidad de 0,53 o mayor (ver figura 14 en Anexo). Según [1] como regla general, se puede lograr una buena visualización de la red con un grado promedio igual a 4. Esto es cuando el número de enlaces es el doble que el de los nodos, que en nuestro caso es con un umbral de 0,53 para los productos de la bioeconomía.

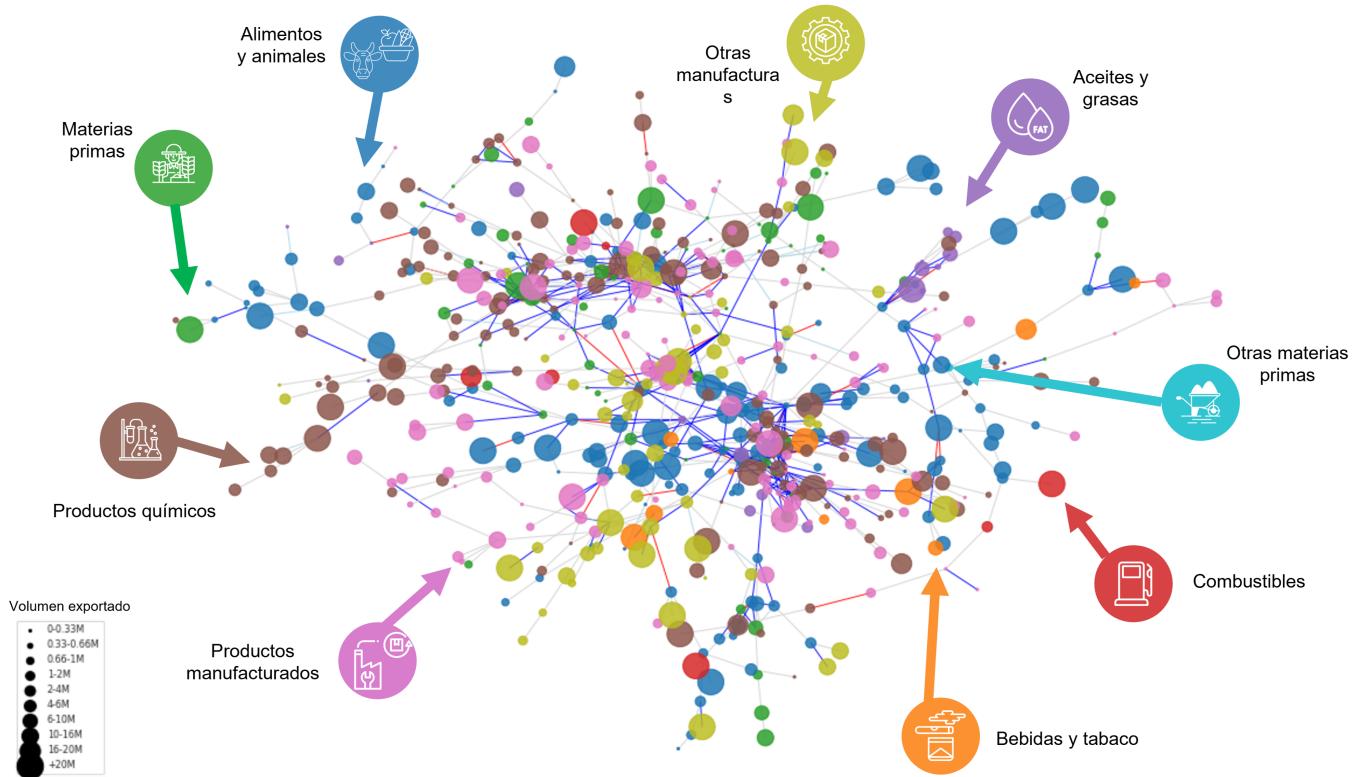
Esto resultó en una red con 584 nodos y 1024 enlaces para el período 2017-19. Este conjunto de 1024 enlaces se utiliza para visualizar la estructura de la matriz de proximidad de productos de la bioeconomía completa, que está lejos de ser homogénea y parece tener una estructura de centro-periferia. Se identifican los enlaces con mayores conexiones en color azul o rojo intenso, mientras que las conexiones más débiles se pintan en colores claros. Los nodos están coloreados según categoría de SITC y el tamaño de los mismos está asociado al volumen comerciado en el período analizado.

En la figura 4 se observan dos módulos densamente conectados (enlaces azules) en el centro, formados principalmente por Productos manufacturados y Productos químicos (en el conglomerado superior de la red), y en el segundo (conglomerado inferior) aparecen enlazados además de los ya mencionados algunos Alimentos y animales (en colores marrón, rosa y azul, respectivamente). La periferia está formada por el resto de las clases de productos. Los productos de la parte alta de la periferia pertenecen a Materias primas, Alimentos y animales y Aceites y grasas. La periferia de la parte inferior de la red muestra un gran grupo de Productos manufacturados y Otras manufacturas (rosa y amarillo), como también Materias primas (verde), Bebidas y tabaco (naranja) y Combustibles (rojo).

Visualizando esta misma red para los restantes períodos, se observó que el espacio de productos presenta en 2007-09 una estructura muy parecida a la que observamos para este último período. Con dos grandes módulos muy conectados que incluyen Productos manufacturados, Productos químicos, Alimentos y animales y Otras manufacturas. En el caso de la red del período 1997-99 se observa una estructura un poco diferente, donde hay un módulo más bien grande y otro más pequeño también conectado de Productos manufacturados, Productos químicos y Materias primas. En la periferia figuran Alimentos y animales, Combustibles y Otras manufacturas.

Más allá de esta ilustración gráfica, se analizó si es cierto que los países desarrollan ventajas comparativas preferencialmente en bienes cercanos. Para ello, se midió la proximidad promedio de un nuevo producto potencial j a la estructura productiva actual de un país, llamada densidad [1]. Un valor de densidad alto significa que el país tiene muchos productos desarrollados que rodean al j-ésimo producto. Para estudiar la evolución de la ventaja comparativa, se consideraron “productos en transición” aquellos con  $VCR < 0,5$  en 1997-99 y 2007-09 y  $VCR > 1$  en 2007-09 y 2017-19, respectivamente. Como control, se consideraron “productos no desarrollados” aquellos que en todos los períodos tuvieron  $VCR < 0,5$  y se descartaron aquellos casos que no

Figura 4: Visualización del Espacio de Productos (incluyendo el MST, los enlaces  $> 0,53$ , categorías de productos (colores de los nodos) y nivel de las exportaciones mundiales (tamaño de los nodos))



cumplieran con ninguno de estos dos criterios. La Figura 5a muestra la comparación de cómo se distribuye la densidad alrededor de los productos de transición (celeste) y los productos no desarrollados (naranja). Estas distribuciones son muy distintas, con una mayor densidad alrededor de los productos en transición que entre los no desarrollados.

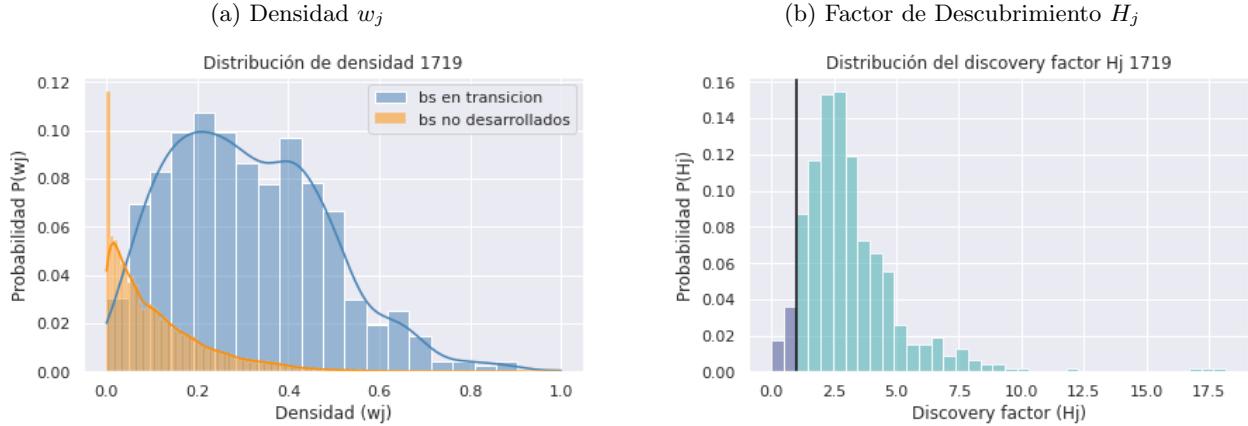
A nivel de producto, se estudió la relación (ratio) entre la densidad promedio de todos los países en los que el  $j$ -ésimo producto era un producto de transición y la densidad promedio de todos los países en los que el  $j$ -ésimo producto no estaba desarrollado, llamado “factor de descubrimiento”  $H_j$ . La figura 5b muestra la distribución de frecuencias de esta relación. En la misma se observa que para el 94.5 % de los productos, esta relación es mayor que 1, lo que indica que la densidad  $w_{jk}$  es mayor en los países que hicieron la transición al  $j$ -ésimo bien que en los que no lo hicieron, habitualmente de manera sustancial. Con ambos análisis se pudo corroborar que en general los países desarrollan VCR preferencialmente en bienes cercanos.

### Caso especial de la Bioeconomía Argentina

Para ver cómo se ubica la Argentina en el espacio de producto de la bioeconomía en base a su estructura productiva, se analizó la evolución de las Ventajas Comparativas Reveladas (VCR) que presenta el país y se generó una nueva visualización de la red del espacio del producto destacando especialmente aquellos bienes con ventajas. La figura 6 muestra la visualización del esqueleto de la red anterior pero identificando en color celeste los productos para los que Argentina presenta VCR y en el tamaño de los nodos el volumen promedio exportado de los mismos.

Examinando los productos que Argentina exporta con VCR para los 3 períodos considerados, se pudo

Figura 5: Distribuciones de Densidad  $w_j$  de productos en transición y no desarrollados y Factor de Descubrimiento  $H_j$  para el período 2017-19



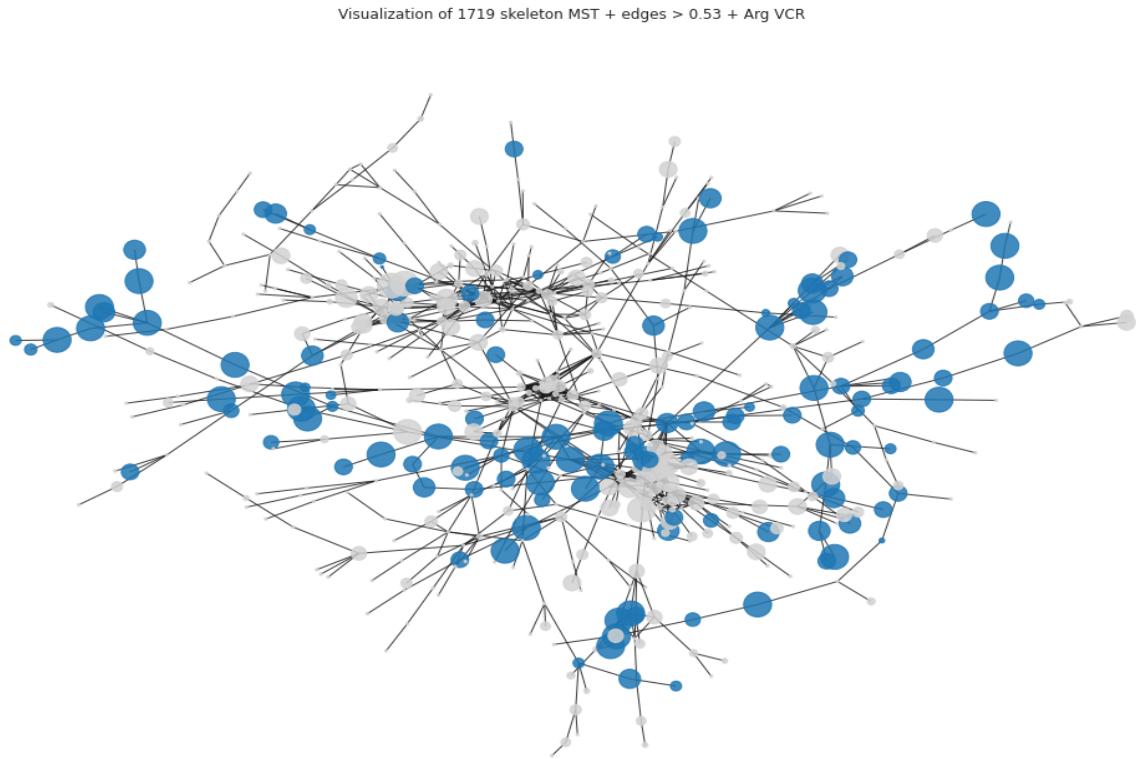
notar que el país ha reducido el número de productos con  $VCR \geq 1$  en la última década. Pasando de 187 en 1997-99 a 199 en 2007-09, pero contrayéndose a 156 ítems con ventajas en 2017-19. Esto coincide con lo advertido en el trabajo de [Palmieri & Perini \(2019\)](#), donde se muestra que las exportaciones han tendido a concentrarse en pocos productos en los últimos años, y que si bien se observa una mayor diversificación de destinos de las exportaciones (mayor número de países que compran productos nacionales) en el último tiempo, esta viene acompañada de mayor concentración de productos. De estos bienes, 110 han presentado VCR sostenidamente durante los últimos 20 años (en los 3 períodos estudiados).

Según [1], los países industrializados ocupan el núcleo del espacio del producto, compuesto en general por productos manufacturados (maquinaria, productos metálicos y productos químicos), mientras que América Latina y el Caribe están más alejados en la periferia en la minería, la agricultura y el sector de la confección. En nuestra representación del espacio de producto corroboramos que, en general, los productos para los que Argentina presenta VCR se ubican más bien en la periferia de la red, con conexiones algo más cercanas al módulo inferior. Estos resultados indican que, tal como decían los autores, cada región tiene un patrón distinguible de especialización claramente visible en el espacio del producto y que la estructura del espacio de productos afecta el patrón de especialización de un país.

Teniendo en cuenta el potencial que tiene Argentina para diversificar los productos, entendiendo que es más factible que un país se mueva hacia bienes más cercanos, se analizó cuáles podrían ser los productos con potencial para los próximos años. En particular, se tuvo en cuenta aquellos bienes que podrían ser "productos en transición" para convertirse en productos con VCR en la próxima década. Es decir, aquellos que presentaron  $VCR > 0,5$  pero menor a  $VCR < 1$  en 2017-19 y que no se mantuvieron en ese rango durante las últimas dos décadas, porque eso implicaría que quizás por alguna otra razón no llegan a convertirse en productos con ventajas. Entre ellos se encontraron bienes pertenecientes en su mayoría a las categorías de Productos manufacturados, Productos químicos y Alimentos y animales. Entre ellos, cabe destacar algunos como frutas y frutos de cáscara congelados, harinas de cereales, aceites, demás azúcares, artículos de confitería sin cacao, chocolate y otras preparaciones con cacao, artículos de panadería y pastelería, salsas y condimentos, aceites de petróleo, gases de petróleo y otros hidrocarburos, ácido nítrico, sulfatos, carburos, hidruros, nitruros, maquinaria agrícola, vehículos, etc. (Ver listado total de productos en nota al pie <sup>1</sup>).

<sup>1</sup> *Bienes detectados con potencial según los criterios propuestos:* grasa de cerdo y ave, frutas y frutos de cáscara congelados, harinas de cereales distintas de las de trigo, harina, sémola y polvo de hortalizas, aceites, demás azúcares (lactosa, maltosa, glucosa y fructosa), artículos de confitería sin cacao, chocolate y otras preparaciones alimenticias que contengan cacao, artículos de panadería y pastelería, salsas y condimentos, otras preparaciones alimenticias, lías de vino, minerales de molibdeno y concentrados, aceites de petróleo, gases de petróleo y otros hidrocarburos, ácido nítrico, sulfatos, carburos, hidruros, nitruros, alcohol acílico, hormonas, secadores preparados, jabón, peptonas y sus derivados, colas y demás adhesivos preparados, iniciadores de reacción, aceleradores de reacción y preparaciones catalíticas, medios de cultivo preparados para el desarrollo de microorganismos.

Figura 6: Visualización del Espacio de Productos identificando productos donde Argentina presenta VCR



### 3.3. Estructura de la Red

Adicionalmente al grafo final que se mostró en la sección 3.2, se probó construir grafos para distintos umbrales<sup>2</sup> de proximidad que incluyeran diferente número de enlaces (distinta densidad de aristas) para cada período analizado. De acuerdo a un análisis a través de heatmaps de los grafos a distintos umbrales de proximidad (ver Anexo figura 16), se pudo comprobar visualmente que a medida que se incrementa el umbral (se reduce la densidad de aristas) existen menos conexiones entre productos del espacio de productos (se observan menores conexiones en color oscuro), lo cual tiene sentido dado que por definición al aumentar el umbral estamos agregando menor número de enlaces. En contraste, para un número fijo de aristas, en este caso se tomó como ejemplo el mismo umbral de 0.53 que se usó para el grafo final, parecería producirse una reducción de las conexiones a medida que pasan los años. Mientras en 1997-99 se registran 1119 aristas, en 2007-09 1087 y para 2017-19 solo 1024.

---

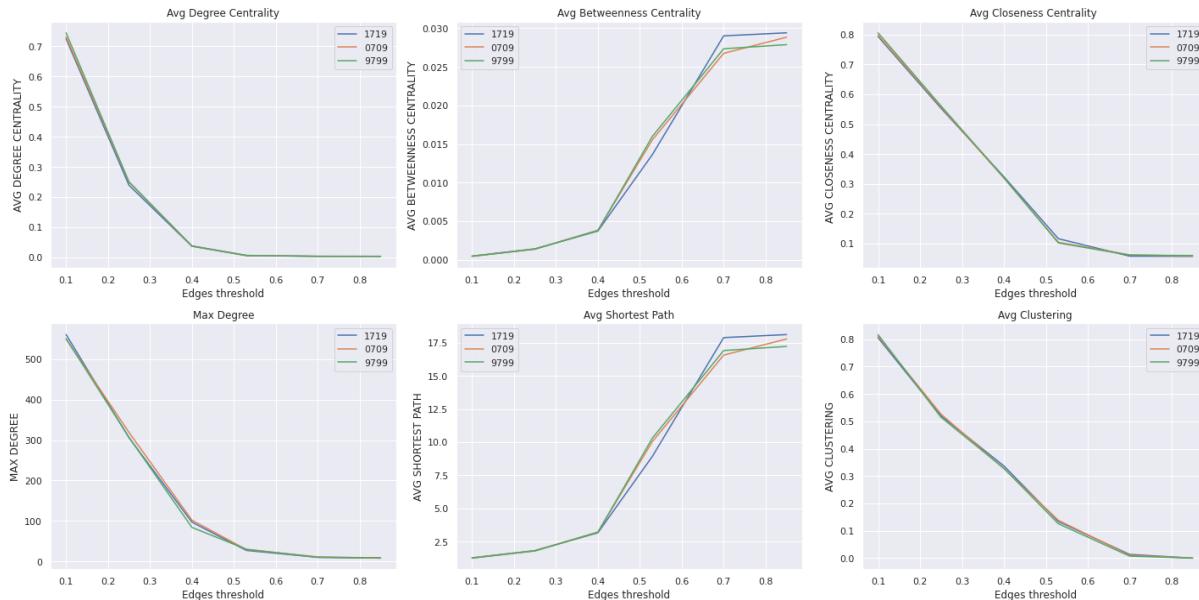
mos, polímeros de etileno, polímeros de cloruro de vinilo o de otras olefinas halogenadas, tubos, tuberías y mangueras, correas transportadoras, pieles, recortes y demás desperdicios de cuero, demás manufacturas de cuero, tableros de fibra de madera, papel y cartón, papeles higiénicos o toallitas, papel de fumar, bobinas, carretes, canillas y soportes similares, libros o folletos, pelo fino, botellas y demás recipientes, productos laminados planos de hierro o acero sin alejar, alambre de los demás aceros aleados, recipientes para gas comprimido o licuado, de hierro o acero, demás manufacturas de plomo, llaves y llaves inglesas manuales, alambre, varillas, tubos, placas, electrodos y productos similares, de metal común, turbinas hidráulicas, ruedas hidráulicas y sus reguladores, aparatos mecánicos (manuales o no) para proyectar, dispersar o pulverizar materias líquidas o en polvo, maquinaria agrícola, máquinas para limpiar, seleccionar o clasificar semillas, cereales u hortalizas, vehículos de motor para el transporte de diez o más personas, incluido el conductor, buques pesqueros; barcos, pinturas, dibujos y pasteles.

<sup>2</sup>Valor de los umbrales seleccionados para el análisis: [0.1, 0.25, 0.4, 0.53, 0.7, 0.85]

Por otro lado, continuando con el análisis exploratorio sobre estos mismos datos se extrajeron medidas de centralidad (de cercanía, intermediación y grado) como también el grado máximo, camino mínimo medio y el coeficiente de clustering en función de los umbrales de proximidad que definimos para cada uno de los períodos estudiados. El cálculo de las mismas se puede visualizar en la figura 7.

Repasando un poco los conceptos de estas medidas [7], recordemos que el grado de un nodo (producto) es el número de conexiones que tiene con otros nodos (productos). En este caso se calculó el **grado máximo**, es decir el número mayor de conexiones presentes en el grafo. Analizando el grado máximo de los nodos por período, se observa que se reduce a medida que aumenta el umbral (ya que se reduce el número de aristas), y en general para todos los períodos es coincidente, salvo para algunos umbrales más bajos en que el período 1997-99 presenta un nivel algo inferior de grado máximo. El grado máximo de la red para el umbral de 0.53 es de 27, por lo que, el máximo número de conexiones de un nodo en la red es de 27. De igual modo se comporta la **centralidad de grado promedio**, que se reduce a medida que aumenta el umbral, es decir, que se incrementa a mayor densidad de aristas (menores valores de umbral) y es idéntica para los 3 períodos analizados. La centralidad de grado define la importancia de un nodo en un grafo a partir de su grado, es decir, cuanto mayor es el grado de un nodo, más importante es en un grafo. En este caso, la centralidad de grado promedio de la red es muy baja (valores cercanos a 0 para un umbral de 0.53), lo que indicaría que la mayoría de los productos tienen una importancia marginal en el espacio de producto.

Figura 7: Comparación de medidas de centralidad (arriba), grado, camino mínimo y coeficiente de clustering (abajo) para todos los períodos analizados según distintos niveles de umbral de proximidad (densidad de aristas)



Por su parte, la **centralidad de intermediación promedio** mide la importancia de un nodo en una red según la cantidad de veces que ocurre en el camino más corto entre todos los pares de nodos del grafo. Esta métrica sí presenta alguna diferencia más visible entre los distintos períodos aunque no estadísticamente significativa, según la densidad de aristas. En particular, para el umbral seleccionado (0.53), la centralidad de intermediación es algo menor en 2017-19 que en los restantes períodos (0.013), mientras que para valores más altos de umbrales (menor densidad de aristas) esta tendencia se revierte. En general, esta centralidad tiende a aumentar a medida que se reduce la densidad de aristas (mayor umbral) pero manteniéndose siempre en valores muy bajos cercanos a 0. Por su parte, el **camino medio mínimo**, que mide el número promedio de pasos a lo largo de los caminos más cortos para todos los pares posibles de nodos de red, crece a medida que se incrementa el umbral (se reduce con la densidad de aristas). Tal como se observó para la centralidad

de intermediación, para un umbral de 0.53 el período 2017-19 presenta menor camino medio mínimo que los restantes períodos, y esta relación se invierte para mayores niveles de umbral (menor densidad), estando el último período por encima de los demás. Para un umbral de 0.53, el camino medio mínimo alcanza un valor de 8.93, es decir, que en este grafo hay que recorrer aproximadamente 9 aristas en promedio para conectar cualquier par de productos.

La **centralidad de cercanía**, por otra parte, mide la importancia de un nodo en un grafo por lo cerca que está de todos los demás nodos (para un nodo, se define como la suma de la distancia geodésica<sup>3</sup> entre ese nodo y todos los demás nodos en la red). Esta métrica muestra una tendencia decreciente con el umbral (creciente con la densidad de aristas), siendo muy similares las curvas para todos períodos a distintos niveles de densidad. Para un umbral de 0.53 asciende a 0.116, valor algo superior que las medidas de centralidad previas, pero cercano a 0 de todas formas.

En tanto que el **coeficiente de clustering promedio** estima la cohesividad local, midiendo la probabilidad de que dos nodos que tienen un vecino común, también estén conectados entre sí. Esta métrica exhibe una reducción a medida que aumenta el umbral (es decir, aumenta a medida que crece la densidad de aristas), manteniéndose, en general, todos los períodos muy parejos para todas las densidades. Para el umbral seleccionado, el coeficiente de clustering promedio ronda un valor de 0.134, lo que indicaría una relativa baja probabilidad de estar conectados entre sí teniendo un vecino común.

Por último, para comparar visualmente las redes a distintos umbrales de proximidad (distintas densidades de aristas) y períodos, se representaron los grafos con layout Spring para 6 umbrales seleccionados que permiten distinguir entre valores bajos, intermedios y altos de aristas ([0.1, 0.25, 0.4, 0.53, 0.7, 0.85]) y otros para el umbral de 0.53 distinguiendo entre períodos. Dichos grafos exhibidos en las figuras 17 y 18, respectivamente, de la Sección 5.1, muestran cómo a umbrales más altos (niveles de densidad bajos) los nodos se distribuyen separadamente y con pocas conexiones en todos los períodos, mientras que a medida que se incrementan las aristas parecerían configurarse comunidades, que si bien varían período a período, parecen mantenerse bastante constantes.

### 3.4. Modularidad de las redes

En la figura 8 se exhiben el coeficiente de modularidad ( $Q$ ) y el número de comunidades ( $NC$ ) para los distintos umbrales de proximidad (densidad de aristas) por cada período, comparando los resultados obtenidos a partir de la detección de comunidades de Louvain con el comportamiento de una red random (aleatoriedad), que preserva la distribución de grados de los nodos. De acuerdo con el análisis, tanto las redes del espacio de producto de Louvain como las aleatorias muestran una estructura modular creciente a medida que aumentamos el umbral (decreciente con la densidad de aristas), solo que la modularidad en las redes aleatorias, en general, es considerablemente menor para todo el recorrido de valores de umbrales contemplados, lo cual es consistente con este tipo de red. Se comprobó que estas diferencias son estadísticamente significativas.

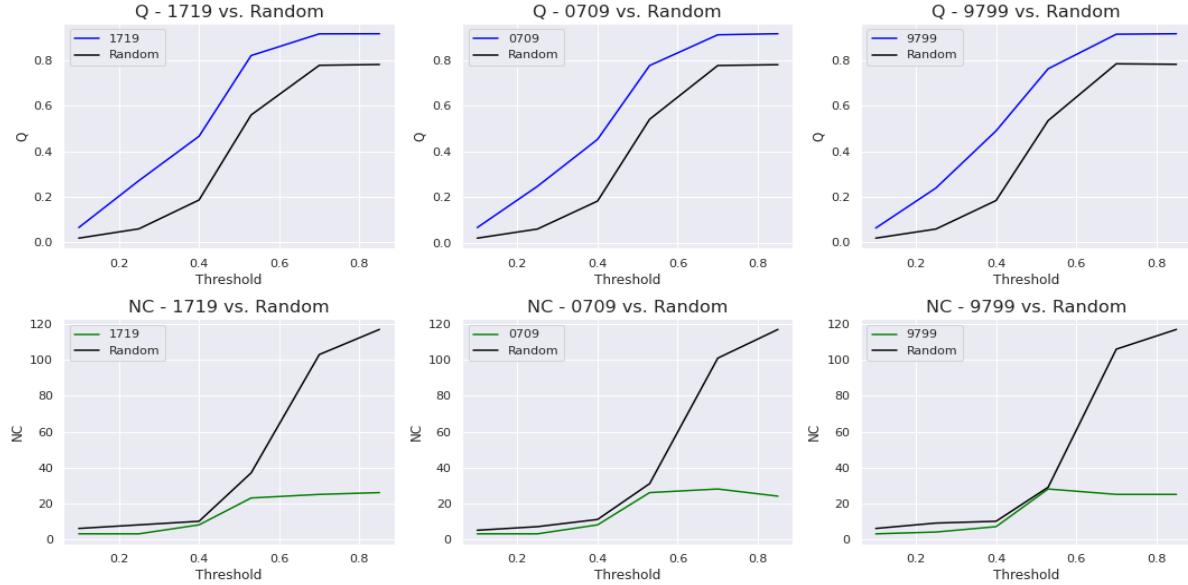
En relación a los períodos analizados, como se puede observar en los gráficos, no parecería haber diferencias sustanciales entre las modularidades para distintos umbrales. De hecho, mediante test estadísticos se corroboró que esas diferencias de modularidad entre los distintos períodos no son estadísticamente significativas.

Por otro lado, se observa que el comportamiento del número de comunidades es bien distinto entre las redes de louvain y las aleatorias. Mientras que en las redes random a medida que se incrementa el umbral (se reduce la densidad de aristas) el número de comunidades se incrementa, en las redes de louvain se puede notar un comportamiento de escalera, con un numero similar de comunidades a valores de umbral menores a 0,5 (menores a 10) y uno más elevado (valores cercanos a 25) pero similar entre valores mayores a ese umbral (en algunos casos incluso podría ser decreciente). En general, en todos los años y para todos los

---

<sup>3</sup>La distancia geodésica  $d$  entre dos nodos  $a$  y  $b$  se define como el número de aristas/enlaces entre estos dos nodos en el camino más corto (camino con el mínimo número de aristas) entre ellos.

Figura 8: Modularidad de la red durante los distintos períodos: Comunidades de Louvain vs. Red Random



niveles de densidad de aristas el número de comunidades parecería ser menor al de una red aleatoria. Otra vez, se observa un comportamiento similar entre los distintos períodos, probado mediante test estadísticos donde las diferencias de  $NC$  entre los distintos períodos no resultan estadísticamente significativas. Tampoco resultaron significativas las diferencias entre en  $NC$  de la red de louvain y la aleatoria.

### 3.5. Cambios en el rol de los nodos

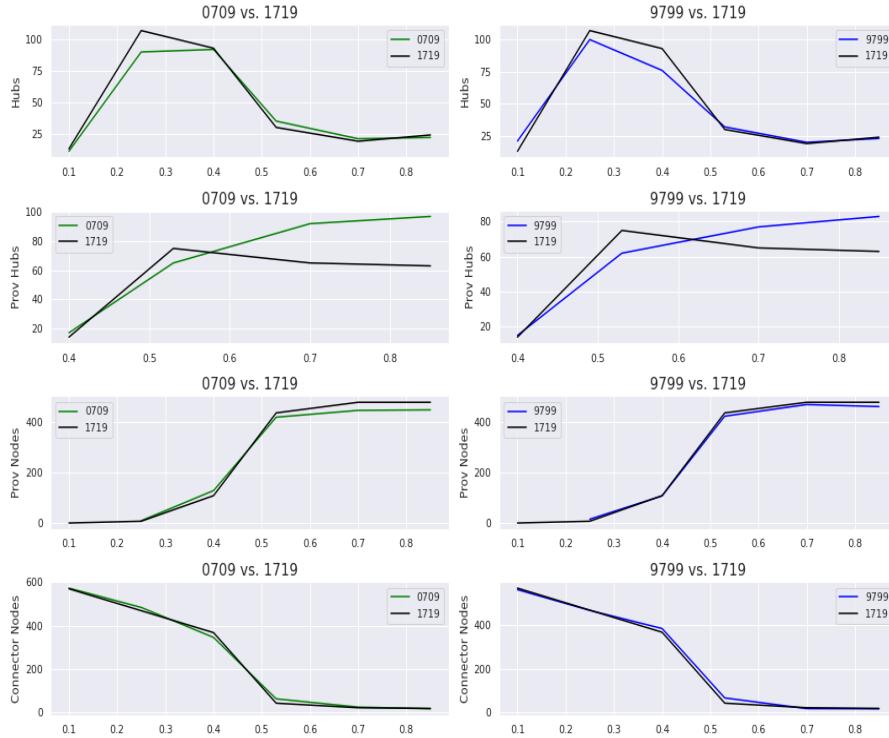
Adicionalmente, para definir el rol que cumple cada nodo dentro de las comunidades e identificar aquellos que cambian de comunidad o de rol, se clasificaron los nodos para cada período según umbrales definidos en la metodología para el coeficiente de participación ( $P_i$ ), que mide la conectividad inter-comunidades, y el z-score del grado intra-comunidad ( $z_i$ ), que mide conexiones al interior de las comunidades. De esta forma se pudo verificar la presencia de nodos en todos los períodos analizados para diferentes umbrales de proximidad (densidades de aristas), que actúan como:

- **Hubs globales (H):** tienen alta conectividad tanto intra ( $z_i$ ) como inter-comunidades ( $P_i$ ).
- **Hubs Provinciales (PH):** tienen alto grado intra-comunidad y baja conectividad inter-comunidades.
- **Nodos Provinciales (PN):** tienen baja conectividad tanto intra como inter-comunidades.
- **Nodos Conectores (CN):** tienen alta conectividad inter-comunidades y baja conectividad intra-comunidad.

En la figura 9 se exhiben los distintos tipos de nodos para los 3 períodos en función de los umbrales de proximidad (densidad de aristas), comparando años anteriores con 2017-19. Se puede notar que el número de hubs y el número de nodos conectores se reducen conforme aumenta el umbral de proximidad (se reduce la densidad de aristas), mientras que los nodos y hubs provinciales ascienden. En particular, los nodos provinciales y conectores, no parecen presentar diferencias entre períodos, mientras que los hubs y hubs provinciales

muestran cierto cambio según el año a ciertos valores de umbral. Por ejemplo, se visualiza un número de hubs inferior en los períodos previos a 2017-19 para valores de umbral menores a 0,5, al tiempo que se observa mayor número de hubs provinciales en 1997-99 y 2007-09 respecto a 2017-19 para valores superiores a 0,6 de umbral. De todas formas, ninguna de las diferencias que se observan resultó estadísticamente significativa.

Figura 9: Roles de los nodos según período para diferentes umbrales de proximidad



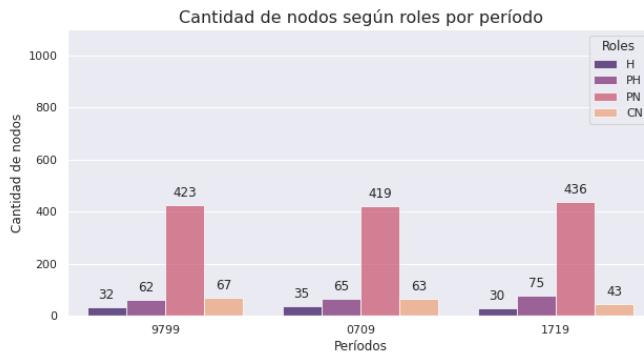
Continuando con el análisis y con el objeto de identificar cuáles son los nodos que cambian entre los distintos períodos, las figuras 10a y 10b representan gráficos de barras del rol de los nodos para un umbral de proximidad dado (de 0.53) para los distintos períodos. Como se puede apreciar, en los 3 períodos analizados se ha mantenido aproximadamente la distribución de los nodos por tipo, siendo los nodos provinciales los que mayor número presentan, muy por encima del resto, seguidos de los hubs provinciales y nodos conectores, y por último los hubs.

El número de **hubs globales** es relativamente similar en todos los períodos y se encuentra en torno a los 30 hubs. Dicho número resulta inferior a todo el resto de los tipos de nodos (nodos provinciales y conectores y hubs provinciales), pero estos nodos tienen el rol más importante de la red ya que tienen alta conectividad con otras comunidades y también al interior de la propia comunidad. No obstante, si bien en número se han mantenido estables, se observa que estos hubs han ido cambiando a través del tiempo, representados por distintos productos. Respecto a los hubs de 2017-19, un 16,6 % de bienes (5 nodos) coinciden con aquellos del período 2007-09 y un 20 % con 1997-99 (6 nodos). Los únicos dos bienes que se mantuvieron como nodos hubs en todos los períodos han sido los demás muebles y sus partes y preparados lubricantes.

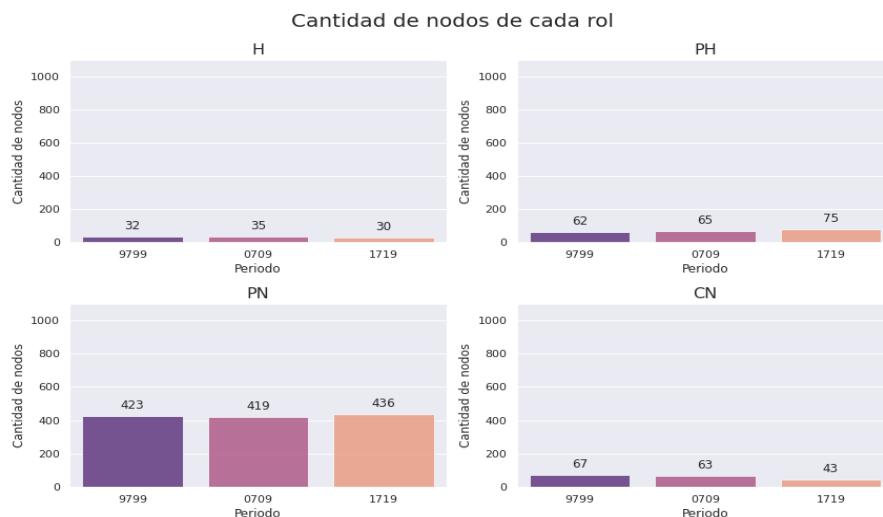
Por otro lado, el número de **hubs provinciales** parece haber crecido levemente en las últimas dos décadas alcanzando un valor de 75 en 2017-19. Recordemos que estos son nodos muy importantes al interior de las comunidades ya que tienen alta conectividad intra-comunidad. En este caso el porcentaje de hubs provinciales que coinciden con períodos anteriores es mayor al de los hubs: existe un 46,6 % de bienes (14 nodos) que coinciden entre 2017-19 y 2007-09, mientras que hay 56,6 % de coincidencia con 1997-99 (17 nodos), y

Figura 10: Roles de los nodos según período y tipo

(a) Segundo Período, por tipo de nodo



(b) Segundo Tipo de Nodo, por período



un 13.3 % (4 nodos) que se comparten en todos los períodos, incluyendo artículos de panadería y pastelería, sangre y vacunas, pescado congelado y ésteres de ácidos inorgánicos y sus sales.

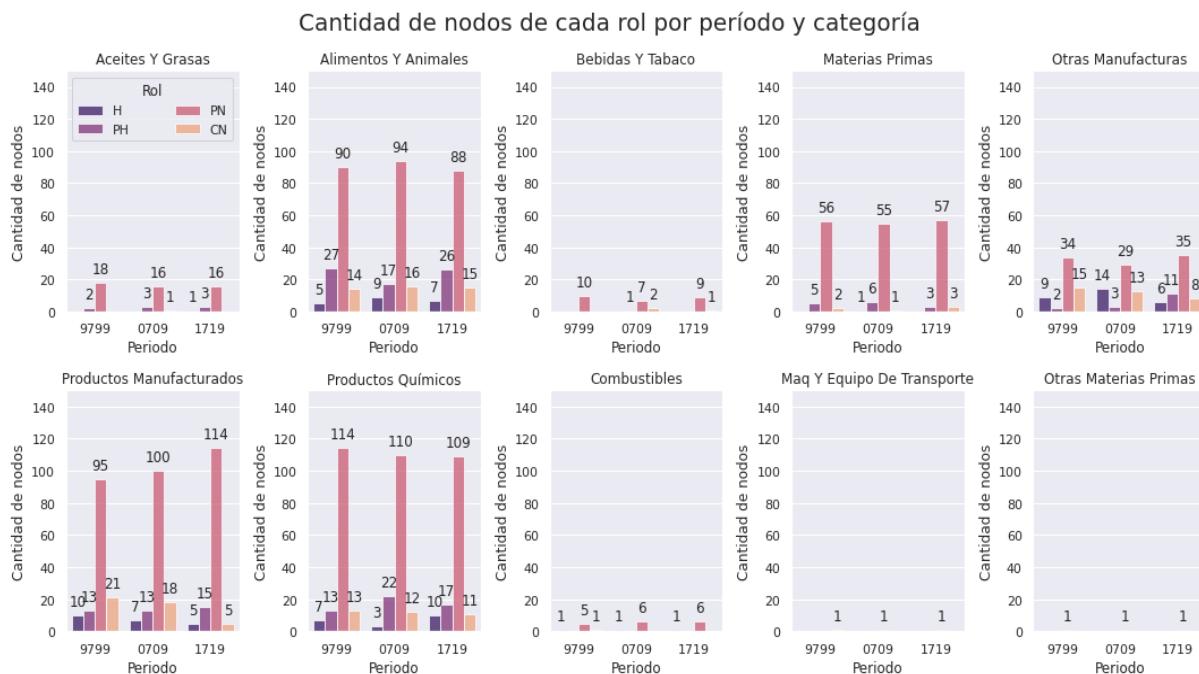
Por su parte, los **nodos conectores** muestran un comportamiento opuesto a los hubs provinciales, se han retraído en los últimos años llegando a un valor de 43 en 2017-19. Estos nodos, que como su nombre lo indican son importantes para conectar entre comunidades, han ido modificándose en el tiempo. Solo un 20,9 % (9 nodos) de los bienes con rol de nodo conector coinciden entre 2017-19 y los períodos previos, al tiempo que un 6.9 % (3 nodos) se han mantenido constantes en todos los períodos, abarcando carnes y despojos preparados o conservados, frutas y frutos de cáscara congelados, y preparados o conservas de pescado.

Por último, se puede notar que los **nodos provinciales** también han crecido levemente desde 1997-99, pero manteniéndose por encima de los 400 en todo el período analizado. La mayoría de los bienes tiene este rol, que consiste en tener alta conectividad con la propia comunidad y baja con las demás comunidades. Estos nodos presentan un comportamiento más sostenido a través del tiempo, donde 78,4 % (342 nodos) de los bienes con rol de nodo provincial coinciden entre 2017-19 y los períodos previos, y un 66.9 % (292 nodos) se han mantenido constantes en todos los años, incluyendo bienes de casi todas las categorías: 'Alimentos y animales', 'Productos químicos', 'Materias primas', 'Aceites y grasas', 'Productos manufacturados', 'Bebidas y tabaco', 'Otras manufacturas', 'Otras materias primas' y 'Combustibles' (ver detalle de par-

ticipación de cada categoría de bienes en cada tipo de rol en la figura 19 de la sección 5.1).

Por último, estudiando la cantidad de nodos y los roles por categorías de productos del espacio de producto, se pudo notar que es muy heterogénea la cantidad de nodos y roles por categoría, aunque relativamente homogénea por períodos (ver figura 11). Cabe destacar, que la presencia mayor o menor de nodos de cada tipo en las distintas categorías responde principalmente a la cantidad de bienes que están incluidos en las mismas, que no es homogénea. Las categorías con más de 130 nodos cada una son Productos químicos, Productos manufacturados y Alimentos y Animales, les siguen con aprox. 60 nodos cada una la de Materias primas y Otras manufacturas, y luego las restantes presentan 20 nodos o menos.

Figura 11: Roles de los nodos por categoría y período



En particular, los nodos provinciales tienen una gran presencia en las categorías de Alimentos y Animales, Productos Manufacturados, Productos Químicos y Materias Primas. Mientras que los nodos conectores se han mantenido en Alimentos y Animales y Productos químicos en las últimas décadas, en Productos Manufacturados y Otras Manufacturas se han reducido sustancialmente. Por su parte, los hubs son pocos y parecen haber sufrido algunos cambios en los últimos años, con menor participación en Productos Manufacturados y mayor en Productos Químicos. En relación a los hubs provinciales, se observa que se han mantenido relativamente estables en categorías de Alimentos y Animales, Productos Manufacturados y Productos Químicos, registrándose un aumento en la de Otras Manufacturas.

## 4. Conclusión

En este trabajo se propuso explorar los cambios en la estructura de redes del espacio de productos de la bioeconomía, para distintos períodos. En este sentido, se observó que el espacio de producto en su conjunto es disperso, con una distribución amplia y heterogénea que evidencia presencia de algunos vínculos fuertes y muchos vínculos marginales. Si bien la proximidad media parece haber ido disminuyendo levemente en las últimas dos décadas, se sigue percibiendo mayor presencia de productos manufacturados entre aquellos con

mayores conexiones y mayor presencia de agrícolas y/o sin procesar entre los menos conectados.

Agrupando los productos por categorías de SITC, se pudo corroborar que existe una mayor proximidad dentro de las categorías y una menor proximidad entre las distintas categorías, siendo las categorías de “Alimentos y animales” y “Bebidas y tabaco” las que mayor proximidad promedio inter-grupo presentan. Mientras que aquellas con mayor proximidad intra-grupo resultaron ser “Bebidas y tabaco”, “Materias primas”, “Maquinaria y equipo de transporte” y “Combustibles”.

La visualización del espacio de productos de la bioeconomía mostró que la red parece tener una estructura de centro-periferia, con dos módulos densamente conectados formados principalmente por “Productos manufacturados” y “Productos químicos”, y algunos “Alimentos y animales”. La periferia está formada por el resto de las clases de productos (materias primas, aceites y grasas, bebidas y tabaco, otras manufacturas y combustibles). En nuestra representación del espacio de producto para el caso Argentino se pudo ver que los productos para los que el país presenta VCR se ubican más bien en la periferia de la red, con conexiones algo más cercanas a uno de los módulos. Estos resultados indican que cada región tiene un patrón distingible de especialización claramente visible en el espacio del producto y que la estructura del espacio de productos afecta el patrón de especialización de un país.

Del análisis de la estructura de redes, se detectó que para un mismo umbral de proximidad se produce una reducción de las conexiones a medida que pasan los años. Es decir, que la reducción de proximidad media observada vino acompañada también de un menor número de enlaces. Por su parte, las medidas de centralidad evidencian que la mayoría de los productos tienen una importancia marginal en el espacio de producto, alcanzando valores cercanos a 0. A su vez, se probó que en esta red hay que recorrer aproximadamente 9 aristas en promedio para conectar cualquier par de productos. Asimismo, se observó que dentro de la red habría una relativa baja probabilidad de que dos productos estén conectados entre sí teniendo un vecino común.

Por otro lado, a partir de la detección de comunidades de la red, se observó que el espacio de producto muestra una estructura modular con una modularidad bastante alta ( $Q = 0,82$  en el período 2017-19) en comparación a una red aleatoria y bastante estable en el tiempo. Si bien se detectó que el número de comunidades parecería ser inferior al de una red aleatoria en todos los años estudiados, no parecería ser una diferencia estadísticamente significativa.

Por último, el análisis de los roles que cumple cada nodo dentro de las comunidades en los diferentes períodos estudiados demostró que en los 3 períodos analizados se ha mantenido aproximadamente la distribución de los nodos por tipo, siendo los nodos provinciales los que mayor número presentan, muy por encima del resto, seguidos de los hubs provinciales y nodos conectores, y por último los hubs globales. Mientras el número de hubs globales se ha mantenido estable en el tiempo, el de hubs y nodos provinciales ha crecido ligeramente y el de nodos conectores se han retraído ligeramente. Los nodos provinciales son los únicos que mantienen mayor proporción de productos con el mismo rol en las dos décadas, para los restantes tipos de roles los productos se han modificado sustancialmente. Hay muy pocos productos que preservan el rol, como por ejemplo los demás muebles y sus partes y preparados lubricantes se mantuvieron como nodos hubs globales con alta conectividad intra e inter-comunidades en todos los períodos. Mientras que las carnes y despojos preparados o conservados, frutas y frutos de cáscara congelados, y preparados o conservas de pescado se mantienen como nodos conectores con alta conexión entre comunidades y los artículos de panadería y pastelería, sangre y vacunas, pescado congelado y ésteres de ácidos inorgánicos y sus sales conservan su rol de hubs provinciales, con alta conectividad intra-comunidad.

En conclusión, no resultaría tan sencillo para un país como Argentina moverse hacia productos más industrializados, ya que hay una distancia importante para llegar a esos productos. Se probó que los países tienden a desarrollar ventajas comparativas preferencialmente en bienes cercanos y se identificaron algunos productos que podrían tener potencial para diversificación en Argentina, convertirtiéndose en productos con VCR en los próximos años. Sin embargo, se observó cómo la estructura del espacio de productos afecta el patrón de especialización de un país y las conclusiones del paper original se sostienen. Para promover una

transformación estructural de la producción que permita un mayor crecimiento de la bioeconomía argentina, será necesario un esfuerzo mucho más grande si se quiere mover a productos más distantes en el espacio de productos. No obstante, son precisamente estos saltos los que podrían permitir ese ansiado avance en el desarrollo productivo.

## Referencias

- [1] Hidalgo, Cesar & Klinger, Bailey & Barabasi, Albert-Laszlo & Hausmann, Ricardo. (2007). The Product Space and its Consequences for Economic Growth. 22006-.
- [2] Hausmann, Ricardo & Klinger, Bailey. (2006). Structural Transformation and Patterns of Comparative Advantage in the Product Space. SSRN Electronic Journal. 10.2139/ssrn.939646.
- [3] Medición de la bioeconomía: cuantificación del caso argentino / Marisa Wierny [et al.]. - 1a ed. - Ciudad Autónoma de Buenos Aires: Bolsa de Cereales de Buenos Aires, 2015. Libro digital, PDF
- [4] Tagliazucchi, E., Von Wegner, F., Morzelewski, A., Brodbeck, V., Borisov, S., Jahnke, K., and Laufs, H. Large-scale brain functional modularity is reflected in slow electroencephalographic rhythms across the human non-rapid eye movement sleep cycle. Neuroimage 70 (2013), 327–339.
- [5] Sayama, H. (2022, July 17). [15.4: Visualizing Networks with NetworkX](#). Binghamton University, State University of New York via OpenSUNY.
- [6] FERNANDEZ URQUIZA, M, LLADA, M., PERINI, S.C. (2020). Cambios en la estructura de redes en el cerebro para distintos estadios de sueño. Trabajo Práctico Materia Data Mining en Ciencia y Tecnología, Maestría DM&KD.
- [7] Bhasin, J., 2019. Graph Analytics — Introduction and Concepts of Centrality. [online] Towards Data Science. Available at: <https://towardsdatascience.com/graph-analytics-introduction-and-concepts-of-centrality-8f5543b55de3>; [Accessed 19 February 2023].

## 5. Anexo

### 5.1. Figuras

Figura 12: Proximidad promedio intra e inter-grupos por período analizado

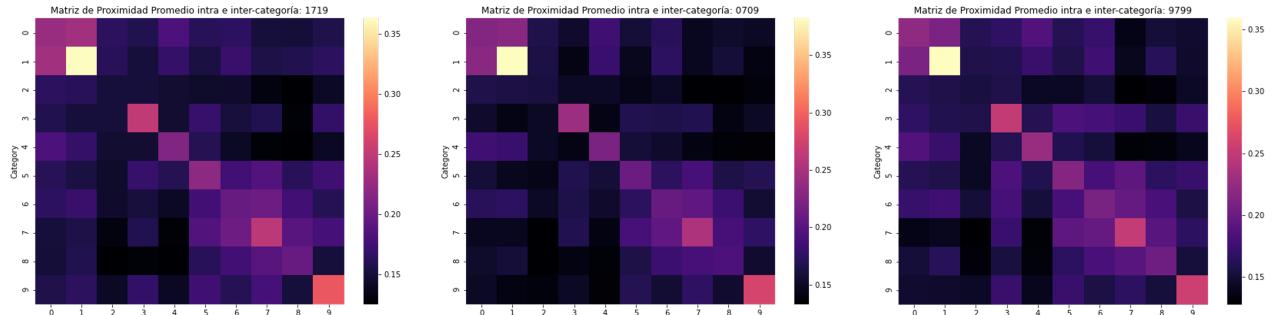


Figura 13: Espacio del Producto representado por su Árbol de Expansión Máxima (MST)



Figura 14: Espacio del Producto representado por su Árbol de Expansión Máxima (MST) y los enlaces más fuertes (por encima de 0.53)

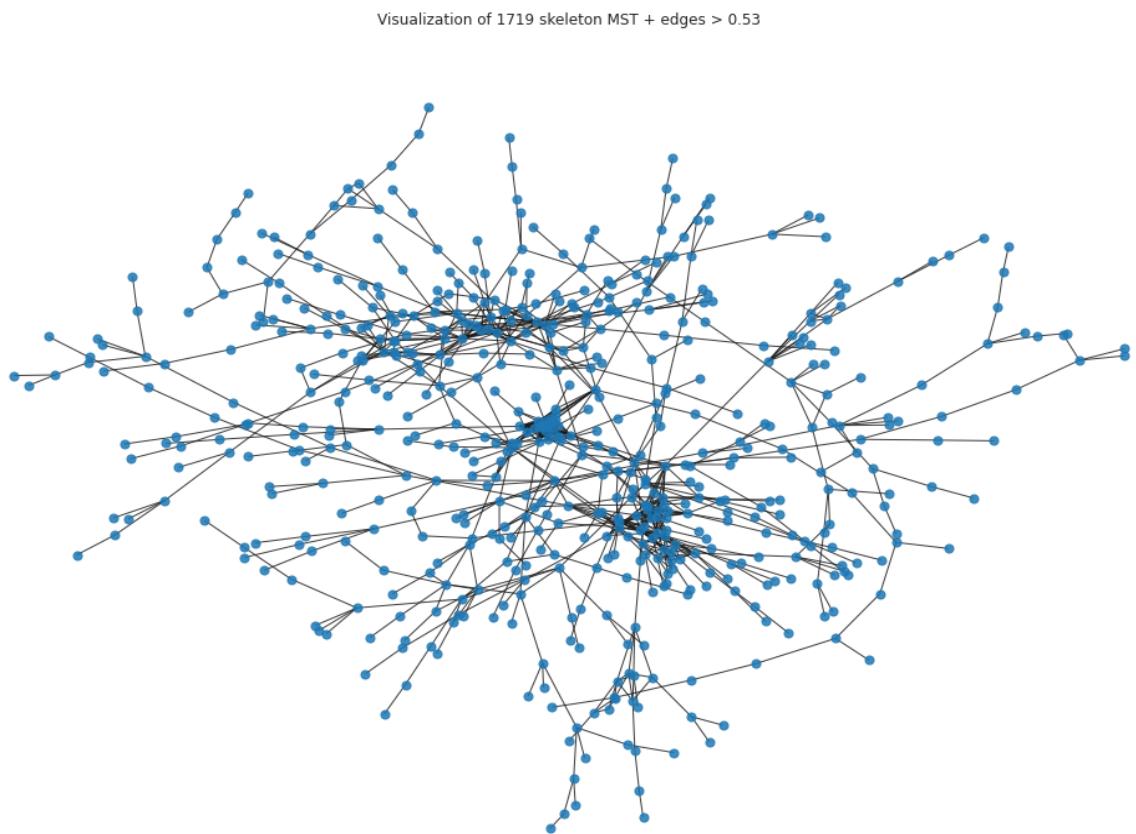


Figura 15: Espacio del Producto representado por su MST, los enlaces más fuertes ( $> 0,53$ ) y los productos por categorías

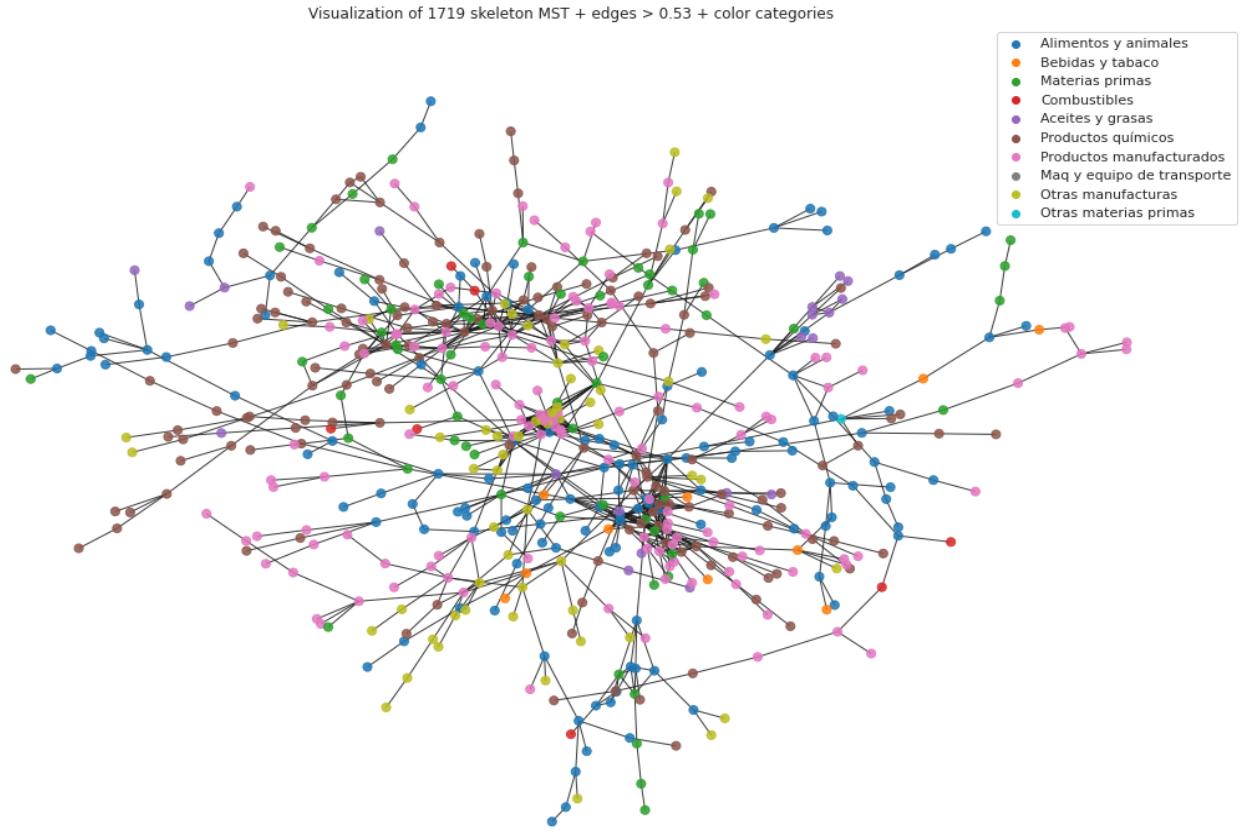


Figura 16: Visualización de grafos a distintos umbrales de proximidad para el período 2017-19

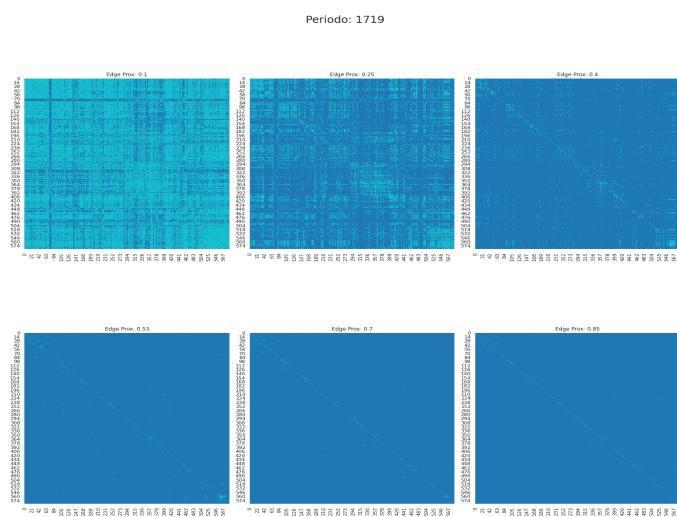


Figura 17: Visualización de grafos para el período 2017-19 en función de distintos umbrales de proximidad seleccionados

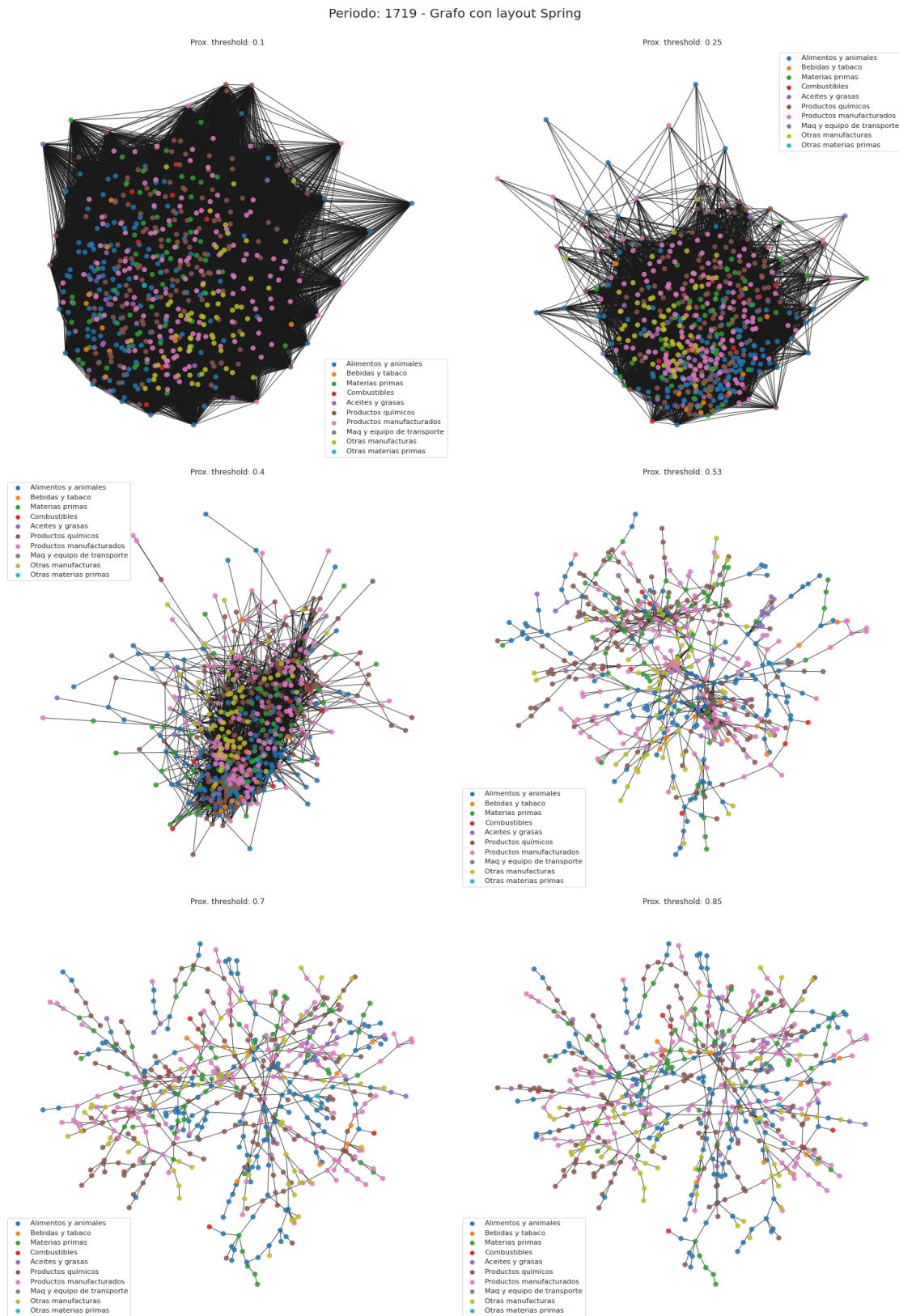


Figura 18: Visualización de grafos para el umbral 0.53 de proximidad según distintos períodos



Figura 19: Participación de las Categorías en cada rol para el período 2017-19

