



UNIVERSIDAD DE BUENOS AIRES
FACULTAD DE CIENCIAS EXACTAS Y NATURALES
MAESTRÍA EN DATA MINING & KNOWLEDGE DISCOVERY

**Análisis empírico del comercio
internacional a partir de la segunda
mitad del siglo XX.
Propuestas metodológicas basadas en
teoría de grafos y modelos generativos
bayesianos**

Tesis de Maestría

Diego Kozlowski

Director: Dra. Viktoriya Semeshenko

Buenos Aires, 2019

**ANÁLISIS EMPÍRICO DEL COMERCIO INTERNACIONAL A
PARTIR DE LA SEGUNDA MITAD DEL SIGLO XX.
PROPUESTAS METODOLÓGICAS BASADAS EN TEORÍA DE
GRAFOS Y MODELOS GENERATIVOS BAYESIANOS**

Abstract
pendiente

Palabras claves: Comercio Internacional, Grafos, Ciencia de Datos, Dirección Latente de Dirichlet, Aprendizaje no Supervisado.

**EMPIRICAL ANALYSIS OF INTERNATIONAL TRADE FROM
THE SECOND HALF OF THE TWENTIETH CENTURY.
METHODOLOGICAL PROPOSALS BASED ON GRAPH THEORY
AND BAYESIAN GENERATIVE MODELS**

abstract pendiente

Keywords: World Trade, Graphs, Data Science, Latent Dirichlet Allocation, Unsupervised Learning.

AGRADECIMIENTOS

pendiente

Índice general

1. introducción	1
1.1 Fuentes de información	2
1.2 Software utilizado	3
2. Comercio a nivel países	4
2.1 Introducción	4
2.2 Análisis Exploratorio de Datos	4
2.3 Metodología	5
2.3.1 Construcción del grafo	5
2.3.2 El comercio internacional como una red compleja	10
2.3.3 Determinación del punto de corte	12
2.4 Resultados	14
2.4.1 Correlación entre la representación de las exportaciones y las importaciones	14
2.4.2 Evolución de mediano y largo plazo del comercio internacional	17
Análisis mediano plazo	17
Análisis de largo plazo	18
2.5 Conclusiones	23
3. Comercio a nivel producto	26
3.1 introducción	26
3.2 Análisis exploratorio de datos	26
3.3 Metodología	31
3.3.1 Latent Dirichlet Allocation Models	34
Definiciones	34
Proceso generativo e inferencial	35
3.3.2 Grafo Bipartito	38
3.4 Resultados	40
3.4.1 Latent Dirichlet Allocation Models	40
Análisis de Componentes	40
Análisis de países	44
3.4.2 Grafo Bipartito	61
Proyección a nivel países	61
Espacio de productos	62
3.4.3 Conclusiones	68
4. Conclusiones	71
Anexo	73

A	Tabla de referencia: Descripción cadenas	73
B	Tabla de referencia: Descripción subcadenas	73
C	Tabla de referencia: Descripción Usos	75
D	Medioides PAM clustering. K=50	76
E	Figuras Complementarias	78

Índice de figuras

2.1	Interacciones Según porcentaje de importaciones que representa al país importador. Destacado par Importador-Exportador.1996-2016	6
2.2	Interacciones. Según valor comercial. Destacado par Importador-Exportador. 1996-2016	7
2.3	Interacciones. Según valor comercial. 1996-2016	8
2.4	Grafos. 2016, según punto de corte	13
2.5	Medidas de centralidad. Importaciones, 2016	14
2.6	Grado total. Grafo exportaciones respecto grafo importaciones. Año 2016 .	15
2.7	Importaciones. Umbral 1%. 1996-2016	17
2.8	Distribución medidas de centralidad según año. umbral 1%	19
2.9	Distribución autovalor ponderado, exportaciones según año. umbral 1% .	20
2.10	Evolución de la estructura de la red. Importaciones. Umbral 1%. 1948-2000	21
2.11	Evolución de la centralidad promedio por continente. Impo. Umbral 1% . .	22
2.12	Argentina, Brasil y México. Importaciones. Umbral 1%	24
2.13	Autovalor. Importaciones. Umbral 1%	25
3.1	Treemaps por tipos de productos. Exportaciones. 1996-2016. Total mundial	28
3.2	Treemaps por Continentes y países. 1996-2016. Total mundial	29
3.3	Treemaps de Cadenas y Subcadenas.2016. Total Latinoamérica	30
3.4	Treemaps de Cadenas y Usos.2016. Total Latinoamérica	32
3.5	Treemaps según socio comercial. 2016. Total Latinoamérica	33
3.6	fuente: Blei et al. (2003)	36
3.7	fuente: Blei et al. (2003)	37
3.8	fuente: Blei et al. (2003)	37
3.9	Grafo bipartito, países y productos	38
3.10	Distribución de los componentes según Lall. K=2	42
3.11	Distribución de los componentes según Lall. K=2	43
3.12	Distribución promedio de los componentes en los países, varios valores de k	44
3.13	Distribución componentes: Irán, Irak, Kuwait, Arabia Saudita y Venezuela.	46
3.14	Distribución componentes: Canadá, México y Estados Unidos.	48
3.15	Distribución componentes: Bolivia, Chile y Perú.	49
3.16	Distribución componentes: Argentina, Brasil y Uruguay.	50
3.17	Distribución componentes: Australia y Nueva Zelanda.	51
3.18	Distribución componentes: Hungría. Polonia y Rumanía.	52
3.19	Distribución componentes: Austria, República Checa y Alemania.	53
3.20	Distribución componentes: Dinamarca, Finlandia, los Países bajos y Suecia.	54
3.21	Francia, Reino Unido e Irlanda.	55
3.22	Distribución componentes: España, Italia y Portugal.	56

3.23	Distribución componentes: China, Japón y Filipinas.	58
3.24	Distribución componentes: Hong Kong, Corea del Sur, Singapur y Taiwan. . .	59
3.25	Distribución componentes: Bangladesh, India y Pakistán.	60
3.26	Proyección a países del grafo bipartito. Clustering Louvain. Exportaciones .	63
3.27	Grafo a nivel países. Clustering Louvain. Exportaciones	64
3.28	Heatmap proximidad. Espacio de productos. Exportaciones. Promedio 1996-2017	65
3.29	Componente gigante. Grafo de proximidad y clustering por K-medoides. Exportaciones. Promedio 1996-2017. 2 medoides	67
3.30	Componente gigante. Grafo de proximidad y clustering por K-medoides. Exportaciones. Promedio 1996-2017. 10 medoides	68
3.31	Componente gigante. Grafo de proximidad y clustering por K-medoides. Exportaciones. Promedio 1996-2017. 50 medoides	69
4.1	Grafos. 2016, según punto de corte. Ponderado por Valor comercial. Importaciones	79
4.2	Distribución acumulada del grado. 2016, según punto de corte. Importaciones	80
4.3	Grado total. Grafo exportaciones respecto grafo importaciones. Año 2016 .	81
4.4	Importaciones. Umbral 1 %. 1996-2016	82
4.5	Evolución de la estructura de la red. Importaciones. Umbral 1 %.1948-2000	83
4.6	Evolución de la estructura de la red. Importaciones. Umbral 1 %.1948-2000	84
4.7	Proyección a países del grafo bipartito. Clustering Walktrap. Exportaciones	85
4.8	Grafo a nivel países. Clustering walktrap. Exportaciones	86

1. INTRODUCCIÓN

El análisis del comercio internacional es una de las áreas de estudio más importantes de la investigación económica. Desde los comienzos de la economía política clásica constituye un tema de preocupación por el efecto del mismo en el desarrollo económico de los países Ricardo (2007). Por su parte, el registro de la información referente al comercio entre países también se remonta en el tiempo.

La cantidad de vínculos comerciales que se establecen entre entidades ubicadas en distintos países implica la imposibilidad de estudiar el fenómeno de forma directa, y plantea la necesidad de elaborar medidas de resumen que permitan apropiar la información subyacente al conjunto de los contratos comerciales existentes.

Sin embargo, el análisis tradicional de la información generada carece de las herramientas necesarias para hacer frente a los grandes volúmenes de datos generados por el comercio internacional en la actualidad. Históricamente, los indicadores sintéticos del área se resumen en volumen y masa de dinero comerciada, partiendo del total mundial, hacia desagregaciones por región, país y sector económico en cuestión WTO (2017). De aquí se desprende la potencialidad del análisis que se basa en técnicas que logren captar en indicadores sucintos la complejidad del objeto de estudio.

A la vez que aumenta el volumen de información persistida respecto del comercio internacional, también se facilita el acceso a técnicas de análisis de datos que requieren un mayor poder de cómputo, y que por lo tanto eran impensadas como herramientas de estudio en épocas anteriores. En este sentido, surge la posibilidad de complementar el análisis tradicional del comercio mundial con técnicas de mayor riqueza en términos cuantitativos.

El presente trabajo se propone utilizar las nuevas técnicas que provee el análisis de grafos para caracterizar el comercio internacional. Su modelización como una red compleja permite la construcción de medidas de resumen que, sin abandonar una mirada holística de la problemática, logren dar cuenta de una mayor complejidad que las métricas tradicionales de dicha área temática.

Una de las posibles formas en que se puede desarrollar dicho análisis es considerar a cada país como la unidad fundamental a partir de la cual se forma una red basada en sus relaciones comerciales con otros países-nodos. Este tipo de análisis permite tener una mirada general de la posición de los países en el comercio internacional, caracterizar su importancia relativa, así como también inferir propiedades estructurales del comercio internacional en su conjunto. En el capítulo dos se desarrolla el estudio siguiendo esta premisa.

Sin embargo, existe aún un nivel mayor de complejidad que puede ser analizado. Es posible descomponer el comercio bilateral entre dos países en los múltiples contratos comerciales individuales entre entidades de los diferentes países. Si bien ese nivel de desagregación es el máximo posible en el objeto de estudio, dicha información no se encuentra disponible al público. Lo que sí es accesible al público es la información de los volúmenes comerciados de cada tipo de producto entre cada par de países, siguiendo un nomenclador estandarizado como el 'Standard International Trade Classification' (ONU, 2006) o el 'Harmonized System' (Weerth, 2008). Este nivel de desagregación implica que si se continúa considerando a los países como la unidad fundamental de análisis, entre cada par

Nivel	Definición
Sección II	Productos Vegetales
Capítulo 10	Cereales
Partida 10.06	Arroz
Subpartida 1006.30	Arroz semiblanqueado o blanqueado, incluso pulido o glaseado

Tab. 1.1: Ejemplo de desagregación para el Sistema Harmonizado

de nodos puede existir una multiplicidad de aristas, potencialmente una por cada producto definido por el nomenclador. Esto implica la construcción de un grafo bipartito entre países y productos. Dicho modelo se analiza en el capítulo tercero de esta tesis.

Por su parte, los nomencladores mencionados tienen la propiedad de ser sistemas jerárquicos en los cuales un mismo producto final pertenece a una serie de categorías intermedias, la tabla 1.1 muestra un ejemplo de los distintos niveles de desagregación que presenta el arroz. Esta estructura de datos permite realizar un análisis exploratorio de la información a través de los diferentes niveles pero a su vez plantea la cuestión de cual es la mejor estructura jerárquica para el análisis de la información. Frente a esto, la multiplicidad de nomencladores, tanto oficiales como realizados ad-hoc para análisis específicos (Molinari & de Angelis, 2016) plantean que no existe un único sistema de estandarización de la información. De aquí surge la posibilidad de utilizar técnicas de reducción de dimensionalidad que utilicen la información disponible para agrupar los productos según grupos de pertenencia. En el tercer capítulo de este trabajo también se avanza en esta dirección proponiendo la utilización del modelo de Latent Dirichlet Allocation (Blei et al., 2003) para explotar la información disponible en la elaboración de un sistema de clustering difuso.

A su vez, tomando el concepto de Ventajas Comparativas Reveladas (Balassa, 1965) se puede calcular un grafo bipartito entre países y productos. La proyección al grafo de países permite utilizar técnicas de detección de comunidades definidas a partir de la semejanza de su estructura exportadora. Finalmente, también en el capítulo tercero, se toma la metodología propuesta por Hidalgo & Hausmann (2009) para caracterizar el espacio de productos a partir del concepto de *similarity*. Esto permite tomar la matriz de similitud resultante como una matriz de adyacencia y construir con la misma un grafo.

1.1. Fuentes de información

La modelización del comercio internacional como un grafo requiere de los datos del flujo anual de comercio bilateral para el total de las mercancías, para la mayor cantidad posible de países, idealmente todos ellos. Dado que los datos del comercio bilateral son realizados por cada país involucrado, es necesaria una base de datos en la cual se haya recolectado la información provista por cada país, y que a su vez haya sido consolidada frente a las posibles, y de hecho abundantes, contradicciones que se presentan entre los reportes oficiales de los países involucrados. En particular, dado que toda importación para un país es una exportación de otro, y que el registro de los datos se realiza de forma nacional, existe una duplicación formal de la información, que no siempre es coincidente. Por esto, se recurrió a una base previamente consistida por un organismo internacional oficial, la Organización Mundial del Comercio (de aquí en más OMC)¹, y la API que dicho

¹ <https://comtrade.un.org/>

organismo provee para descarga de datos. La información analizada en la sección segunda, del período 1996-2016, proviene dicho organismo.

Los datos utilizados para el análisis de largo plazo, entre los años 1948 y 2000 proviene del trabajo realizado por Gleditsch (2002), que también es utilizado en otros trabajos de modelización del comercio internacional mediante redes complejas Fagiolo et al. (2010).

En la sección tercera, donde se analiza el comercio a nivel producto, requiera de las exportaciones de cada país desagregadas por un nomenclador. Para ello, se utilizó la información provista por el ?, la cual se basa en la información de comtrade, consistiendo los datos faltantes a partir de los flujos en espejo, y los distintos nomencladores utilizados en el tiempo. Esta base nos permite realizar un análisis histórico, y es la utilizada por la literatura especializada (Hidalgo et al., 2007; Hidalgo & Hausmann, 2009; Hidalgo, 2009). Sin embargo, dado que la normalización implica un cierta pérdida de información, para la sección 3.4.2, se decidió reconstruir la base directamente a partir de los datos de *comtrade*.

1.2. Software utilizado

Para el presente trabajo se utilizó el lenguaje de programación estadística R(R Core Team, 2017), junto con sendas extensiones del mismo. Aquellas que resulta importante mencionar son *igraph*(Csardi & Nepusz, 2006) para la construcción de los grafos y las medidas de resumen de los mismos; *tidyverse*(Wickham, 2017) para la manipulación de la información y elaboración de los gráficos. Para esto último, también se hizo uso de (Wilke, 2017; Arnold, 2017; Neuwirth, 2014; Slowikowski, 2017; Vu, 2011), como complementos de *ggplot*. Por su parte, los códigos de los países provienen de (Arel-Bundock, 2017).

Para el análisis del capítulo 3 también se utilizó el lenguaje Python (van Rossum, 1995), y específicamente la librería *scikit-learn* (Pedregosa et al., 2011) para la estimación de los modelos. A su vez, para esta sección también se elaboraron dos tableros dinámicos para el análisis de resultados, sobre la base de la librería *shiny* (Chang et al., 2018), y se hostearon en www.shinyapps.io/ (Allaire, 2019)

2. COMERCIO A NIVEL PAÍSES

2.1. Introducción

En el presente capítulo se analiza la estructura comercio internacional a partir de las relaciones comerciales bilaterales. A partir de la elaboración de una red compleja de relaciones entre países, es posible captar la posición relativa de los mismo en el agregado mundial, tanto en un momento del tiempo, como en la evolución. A su vez, a partir de utilizar tanto la información de la exportaciones como de las importaciones, el posible analizar los diferentes roles que juegan los países en el mercado mundial. Junto con esto, analizar las propiedades emergentes de la red, y su evolución en el tiempo, nos permite caracterizar los cambios en el mercado mundial, y buscar nuevas evidencias respecto de la Nueva División Internacional del trabajo (Fröbel et al., 1978).

En la literatura reciente se realizaron diversos enfoques desde esta perspectiva, ya sea desde una mirada de la teoría de la información Bhattacharya et al. (2008), como una herramienta de modelización de los fenómenos económicos Fan et al. (2014), para analizar las relaciones de centro-periferia Fagiolo et al. (2010), o bien para realizar una descripción del estado del comercio internacional Chow (2013). Sin embargo, dada la amplitud de los temas a abordar, tanto desde la perspectiva económica, como desde la teoría de grafos, se presenta como un campo abierto para la investigación, con sendas aristas aún por recorrer.

El capítulo se estructura de la siguiente manera: En la sección siguiente se realiza un análisis exploratorio de datos, mientras que la sección tercera describe las fuentes de información utilizadas y realiza un pequeño resumen de los elementos de la teoría de grafos utilizados, así como las particularidades que surgen al aplicar este marco metodológico al comercio internacional. La sección cuarta presenta los resultados, con detalle en el año 2016, así como también un análisis de mediano y largo plazo de la evolución de la red. Finalmente, se presentan las conclusiones específicas a este capítulo.

2.2. Análisis Exploratorio de Datos

En primer lugar, en la figura 2.1(a), analizamos la frecuencia de las interacciones entre países en función de la proporción del comercio que representan para el país importador. La distribución tiene una asimetría a derecha, que asemeja una distribución de chi-cuadrado. En la figura 2.1(a) se puede observar que si bien la gran mayoría de las relaciones bilaterales de comercio representan una baja proporción de las importaciones de un determinado país, existen ciertas relaciones comerciales que se podrían caracterizar como de alta dependencia para el país importador. En particular, se observan cuatro relaciones comerciales de alta dependencia para el año 2016: La venta de mercancías de Estados Unidos a las islas de San Cristóbal y Nieves y Bermudas ; las exportaciones de Dinamarca a Groenlandia y de la India a Nepal.

En la figura 2.1(b) se puede observar aquellas relaciones que representan más de un 70 % de las importaciones del país importador, para el período 1996-2017. Allí se destaca con el tamaño de fuente el Valor comerciado en cada dupla para cada año en particular. La mayoría de las relaciones de tipo dependiente se repiten año a año y constituyen volúmenes bajos de comercio, entre países de tamaños, económicos poblacionales y territoriales, muy

diferentes entre sí. En particular destaca Estados Unidos en su relación comercial con pequeñas islas del Caribe, y Sudáfrica con países de menor tamaño en su mismo continente, como Namibia o Botswana. Otra relación de comercio dependiente que se repite en la serie es la exportación de productos desde Israel a Palestina. Sin embargo, todas estas relaciones implican un monto menor en términos del comercio mundial. La relación de tipo dependiente más importante es la exportación de productos desde Estados Unidos hacia México durante la década del noventa, en el proceso conocido como la "maquila mexicana" (Carrillo & Hualde, 1998)

Por su parte, en la figura 2.2(a) se puede observar la distribución de los montos comerciados en el año 2016, destacando aquellas relaciones de mayor volumen. Al igual que en la figura 2.1(a), también se observa una fuerte asimetría a derecha, siendo las relaciones más importantes en el continente asiático y Americano. En particular, Las importaciones de Estados Unidos desde México, Canadá y China. También destacan las exportaciones de China, además de la ya mencionada, a Japón y Hong-Kong. Estas relaciones no corresponden a caracterizarlas de tipo dependiente en un sentido de intercambio desigual, dado que se trata de volumen, y aquellas que destacan resultan de importancia tanto para exportador como para el importador. Se podría decir que estas son relaciones de alta interdependencia.

En este sentido, la figura 2.2(b) muestra las cinco relaciones más importantes de cada año. Nuevamente lo que se observa es una repetición en las duplas Importador-Exportador entre año y año. Destaca sobre todo el rol de Estados Unidos como importador de China en primer lugar, y Canadá y México en menor volumen. Es importante observar que las exportaciones de China a Estados Unidos toman una importancia creciente en el tiempo y para el final de la serie se encuentran por fuera de la nube de puntos.

Esto también se puede apreciar en relación al total de las relaciones comerciales, en la figura 2.3 donde se incluye la totalidad de las relaciones comerciales para la serie entre 1996 y 2017. Allí se observa que la absoluta mayoría de las relaciones comerciales constituyen montos menores, y que una proporción menor se realiza por montos de hasta cien mil millones de dólares anuales. También se observa un número creciente de operaciones por montos arriba de los cien mil millones de dólares anuales. Finalmente, también cabe destacar nuevamente la anomalía que representa respecto del resto de los datos las exportaciones desde China a Estados Unidos en la segunda mitad de la serie.

2.3. Metodología

2.3.1. Construcción del grafo

Un grafo (o red) se define como un conjunto de nodos y aristas, donde cada nodo representa una entidad, mientras las aristas representan la existencia de una relación entre dos entidades. La representación algebraica del grafo se realiza mediante la denominada matriz de adyacencia, donde las filas representan los nodos de salida, mientras que las columnas representan el nodo de entrada. Por lo tanto, el elemento a_{ij} de dicha matriz representa a la arista que va desde i hacia j .

En el caso de las redes pesadas, una arista, $a_{ij} \in \mathbb{R}$, denota la intensidad de la relación entre los nodos i y j . Por su parte, en las redes binarias, $a_{ij} \in \{0, 1\}$ define cualitativamente la relación, dónde $a_{ij} = 1$ si existe la relación y $a_{ij} = 0$ sino. Este último tipo de grafos sera el utilizado para el presente trabajo, dado que se cuenta con una mayor cantidad de métricas para caracterizar los grafos de este tipo. A su vez, las redes pueden ser tanto dirigidas como no dirigidas. En el primer caso, la relación no tiene un sentido definido,

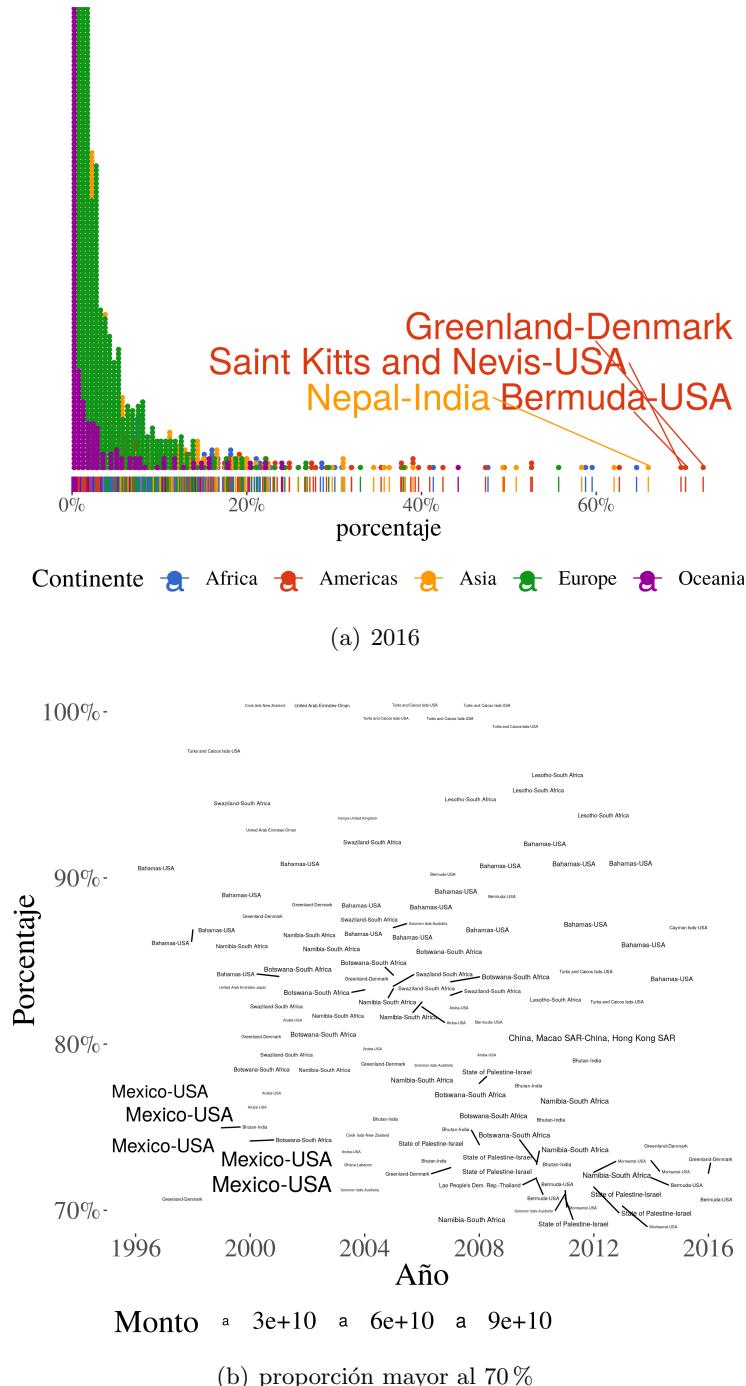


Fig. 2.1: Interacciones Según porcentaje de importaciones que representa al país importador. Des- tacado par Importador-Exportador.1996-2016

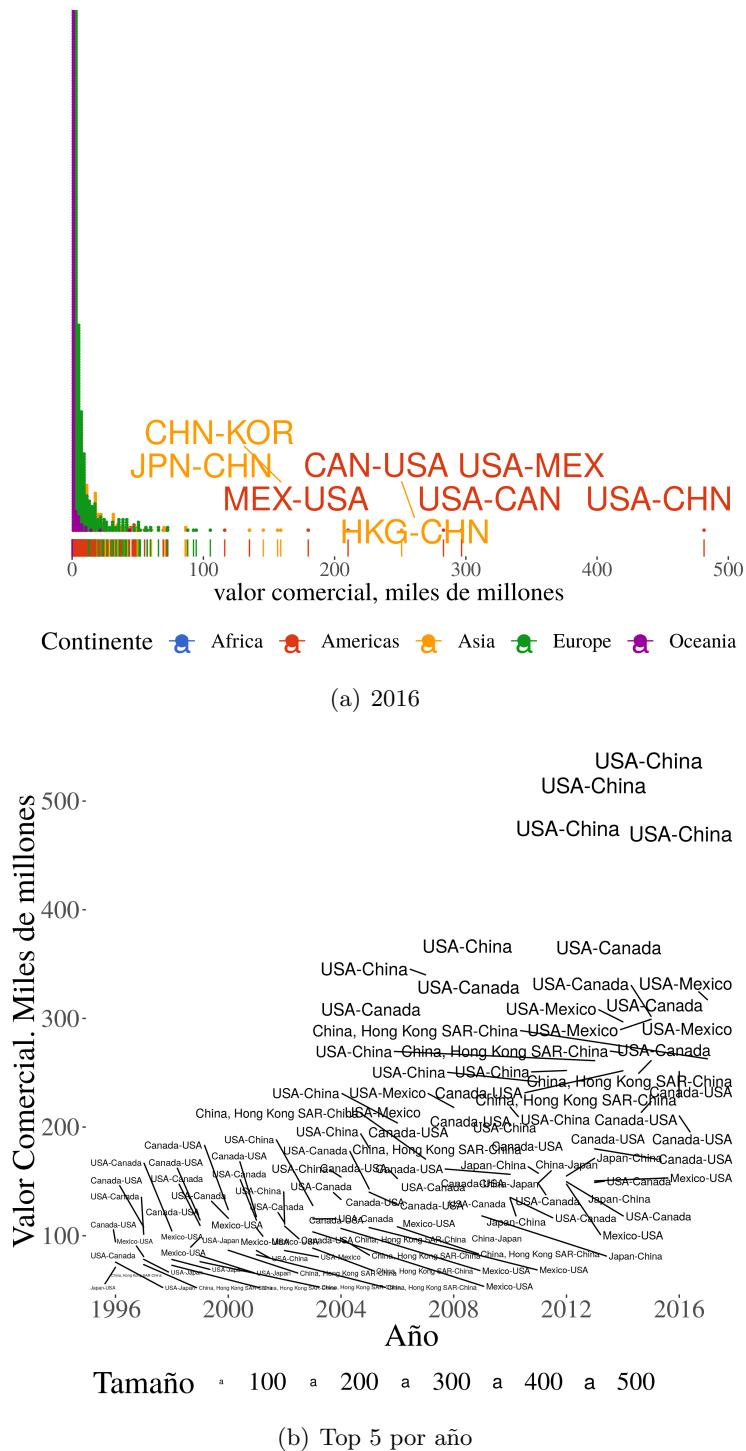


Fig. 2.2: Interacciones. Según valor comercial. Destacado por Importador-Exportador. 1996-2016

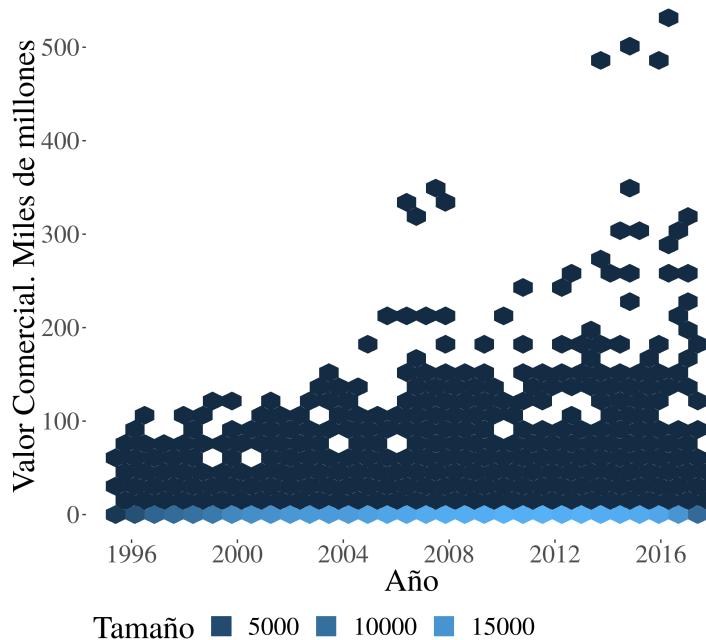


Fig. 2.3: Interacciones. Según valor comercial. 1996-2016

y por lo tanto $a_{ij} = a_{ji}$, mientras que en el segundo caso tal igualdad no es necesaria, y la presencia de una relación tiene un sentido definido. Si bien en la bibliografía existen autores que plantean que la simetría del intercambio comercial es fuerte, y por lo tanto no es necesario utilizar un grafo dirigido para su representación Fagiolo et al. (2007), en el presente trabajo se sostiene que esto depende fuertemente de la forma de construcción de la red, en particular si se trabaja con un grafo no ponderado, y por lo tanto hay una riqueza interpretativa que se pierde si se utiliza una red no dirigida.

Sobre una red se puede calcular una serie de medidas de resumen que permiten caracterizarla. Éstas pueden referirse a un atributo propio del grafo, o bien de cada nodo o arista. Si bien las posibles métricas son muchas, en el presente trabajo nos limitamos a aquellas que se definen a continuación.

Dos nodos se dicen conectados si existe un recorrido entre ellos, es decir, una secuencia de aristas y nodos entre estos. De un grafo se dice que está conectado, si todos sus elementos lo están.

El diámetro d_m de una red es la máxima distancia entre cualquier par de nodos. En teoría de grafos se considera la distancia entre un par de nodos como el camino más corto que se puede realizar entre éstos. Es decir, el diámetro de un grafo no pesado se considera como la cantidad de aristas que hay entre los dos elementos más lejanos del grafo, considerando el recorrido más corto entre ambos.

$$d_m = \max(d_{uv})$$

donde d_{uv} es la distancia entre los nodos u y v, definida como:

$$d_{uv} = \min(L(P))$$

con $P = (v_1, v_2, \dots, v_n)$, un conjunto de n vértices, donde v_i es adyacente a v_{i+1} , y siendo

$L(P)$ la longitud del camino P, definida como:

$$L(P) = \sum_{i=1}^n w_i$$

En el caso de una red no ponderada, $w_i = 1$ y por lo tanto, la longitud de un camino es la cantidad de elementos que lo componen.

La densidad De_m de una red es la relación entre la cantidad de aristas de la red y la máxima cantidad de aristas potenciales de la misma.

$$De_m = \frac{|A|}{|V|(|V| - 1)}$$

Siendo $|A|$ el número de aristas y $|V|$ el número de vértices, o nodos, del grafo. $De_m \in [0, 1]$. Aquellas redes con una densidad cercana a 0 se las considera dispersas, mientras que aquellas con una densidad cercana a 1 se las considera densos.

A su vez, existe una serie de medidas de centralidad de nodos que caracterizan la importancia de los mismos en la red. Las utilizadas en el presente trabajo son la intermediación, cercanía y autovalor.

La centralidad de intermediación mide el rol de un nodo i , $C_{BET}(i)$, como puente entre los demás nodos, considerando sus caminos más cortos. Se define como:

$$C_{BET}(i) = \sum_{j,k \neq i} \frac{b_{jik}}{b_{jk}}$$

donde b_{jk} son todos los caminos más cortos entre los nodos j y k , y b_{jik} son los caminos más cortos entre j y k , que pasan por i .

La centralidad de cercanía es la distancia promedio que tiene un nodo con los demás nodos con los que está conectado. Formalmente:

$$C_{cer}(i) = \frac{1}{\sum_{j \neq i} d_{i,j}}$$

La centralidad de grado se define como la cantidad de aristas de un nodo. En el caso de los grafos dirigidos, se puede considerar un grado de entrada, así como un grado de salida, si las aristas van en dirección al grafo o desde el grafo, respectivamente. La centralidad de autovalor, por su parte, caracteriza la importancia de un nodo considerando la de sus nodos vecinos. Si asumimos linealidad:

$$\begin{aligned} x'_i &= \sum_{j=1}^N a_{ij} x_j \\ x' &= Ax \end{aligned}$$

Siendo A la matriz de adyacencia de la red. si x' es el vector de centralidades de la red, entonces podemos plantear $Ax^* = \lambda x^*$ donde x^* es la solución y λ es el autovalor asociado al mayor autovector. Por lo tanto, podemos definir la medida de centralidad de autovalor como

$$x_i^* = \frac{1}{\lambda} \sum_i A_{ij} x_j^*$$

Esta medida tiene la propiedad de considerar la importancia de un nodo a partir de la importancia de aquellos nodos con los que esta conectado, y con los que éstos están conectados, etc. En el presente trabajo también se considera a la centralidad de autovalor ponderado como una alternativa, donde el peso esta definido por el volumen de dinero comerciado entre los países. Para calcularlo, se multiplica la matriz de adyacencia por una matriz de pesos. Por lo tanto

$$x_i^{pond} = \frac{1}{\lambda} \sum_i W A_{ij} x_j^{pond}$$

Estas medidas de centralidad pueden ser resumidas para el grafo en su conjunto mediante una agregación de algún tipo, como el promedio, máximos y mínimos, etc. En el presente trabajo se utilizan de ambas maneras, tanto para definir la importancia de un país en particular en el grafo, como para caracterizar a la red en su conjunto.

El Coeficiente de clustering estima la cohesividad local, midiendo la probabilidad de que dos nodos que comparten un vecino estén conectados. Para calcularlo, se mide en primer lugar el grado de agrupamiento de cada nodo, tomando la vecindad del mismo, es decir, todos aquellos nodos con los que esta conectado, y se calcula el cociente entre las aristas existentes y la máxima cantidad de aristas posibles:

$$C_i = \frac{|A_i|}{|V_i|(|V_i| - 1)}$$

siendo $|A_i|$ las aristas en el subgrafo de la vecindad de i , y $|V_i|$ los vértices del subgrafo de la vecindad de i . Nótese que esta medida no representa lo mismo que la densidad, si bien las ecuaciones sean similares, dado que en el clustering se analiza solamente la vecindad de un nodo, mientras que la densidad se refiere a toda la red. Si quisieramos caracterizar a toda la red según el coeficiente de clustering, podemos calcular el promedio del mismo para todos los nodos, también conocido como transitividad del grafo, que se define como el promedio del grado de agrupamiento de los nodos. Es decir:

$$\bar{C} = \frac{1}{n} \sum_i C_i$$

Por último, también se considera la correlación de grado entre las aristas. Esta se calcula como la correlación de Pearson entre el grado de nodos adyacentes. Si la correlación es fuerte y positiva, esto implica que la red es de tipo selectiva, lo que indica que los nodos de mayor grado tienden a conectarse con otros nodos de mayor grado. Si la correlación es negativa y fuerte, esto indica que la red es no selectiva o heterogénea, lo que implica que los nodos más importantes de la red tienden a conectarse con nodos poco centrales.

2.3.2. El comercio internacional como una red compleja

La modelización del comercio internacional como un grafo conlleva una serie de simplificaciones de la información original. Nos encontramos en primer lugar que toda compra es a su vez una venta, y que por lo tanto la información se puede interpretar de ambas maneras. Por su parte, dado que lo que se busca representar es existencia de un vínculo

comercial entre dos países, considerar toda compra o toda venta como la existencia de este vínculo resultaría exagerado. Por lo tanto, es necesario establecer un punto de corte a partir del cual se considere que existe una relación comercial entre la dupla de países en cuestión. Sin embargo, tal punto de corte no se debe considerar en término absolutos, como el monto comerciado, ya que dicho monto está sumamente determinado por el tamaño de los países en cuestión. En otros términos, una determinada masa de dinero es significativa para un país en función del valor total comerciado por el mismo. Por lo tanto, es necesario considerar el umbral en términos relativos respecto al tamaño de los mismos.

En este punto se abre una serie de consideraciones respecto al denominador a utilizar. Si se normaliza por el tamaño de la economía, el Producto Bruto Interno (PBI), se está considerando no sólo la importancia relativa de ese vínculo comercial respecto del total comerciado, sino también el grado de apertura de dicha economía. En ese sentido, si se quisiera analizar el rol de un país en el comercio internacional, independientemente del tamaño real de su economía, lo más correcto es normalizar por el volumen total comerciado, y no por su PBI. De esta forma, se puede considerar un punto de corte como un tanto por ciento de las importaciones o exportaciones del país que reporta la transacción.

Dado que se considera que la información de las importaciones suele ser más confiable que la de las exportaciones (Fan et al., 2014), en el presente trabajo se utiliza en la generalidad de los casos el punto de corte como un porcentaje del total de las importaciones del país importador. No obstante, dada la construcción de dicho punto de corte, la interpretación de los resultados varía. Si se utiliza la información provista por el importador, la dirección de la arista resultante es desde el importador hacia el exportador, el sentido contrario al de las mercancías, es decir, el sentido del dinero. Esta arista sólo existe si resulta significativa para el país importador.

Podemos entonces, utilizar un grafo binario dirigido. Dados N países, definimos la matriz de adyacencia NxN, donde:

$$a_{ij} = \begin{cases} 1 & \text{si } \frac{x_{ij}}{x_{i\cdot}} \geq u \\ 0 & \text{sino} \end{cases}$$

Donde u representa el umbral elegido, y x representa el volumen comerciado. Para el caso de las importaciones, x_{ij} representa el monto total comerciado desde el país j hacia el país i , mientras que $x_{i\cdot}$ representa el total de las importaciones del país i . En el grafo de exportaciones, al contrario, x_{ij} representa el monto total comerciado desde el país i hacia el país j , mientras que $x_{i\cdot}$ representa el total exportado por el país i .

Dado que la restricción se construye por la importancia relativa para el nodo de origen, en general sólo existirán una pequeña cantidad de aristas que salgan de cada nodo, aunque no están limitadas las aristas en dirección hacia el nodo. Más precisamente, la máxima cantidad posible de aristas de salida es igual a $\frac{1}{u}$. Esta cantidad se obtiene sólo cuando el país comercia de forma uniforme con exactamente $\frac{1}{u}$ países. Como se verá más adelante, la distribución del comercio dista mucho de ser uniforme, por lo cual la cantidad de aristas de salida de un nodo se verá fuertemente limitada. Sin embargo, las aristas de entrada al nodo sólo se limita formalmente por la cantidad de nodos en el grafo.

Por lo tanto, para el grafo de importaciones, cuando se consideren medidas de centralidad de los nodos, se deberá considerar en particular las aristas que están dirigidas en dirección al nodo, y éstas marcarán la importancia de tal nodo como un país productor del mercado mundial, ya que lo que se mide es si los productos que exporta dicho país resultan significativos o no considerados en las importaciones de los demás países. Por

su parte, si se consideran los datos provistos por los países exportadores, las medidas de centralidad de los nodos, estarán reflejando la importancia de tal país como consumidor global, ya que su consumo resulta significativo como exportación para los países que le venden dichas mercancías. Dado que, por la forma de construcción de los grafos, resulta distinta la interpretación de la información de las exportaciones y las importaciones, se considerarán alternativamente ambas redes según qué resulte más interesante.

2.3.3. Determinación del punto de corte

El punto de corte a utilizar, dada la metodología propuesta para discretizar la información del volumen comerciado, opera fuertemente sobre el grafo resultante. La figura 2.4 muestra el grafo según si se utiliza un punto de corte de 1%, 5%, 10%, 15%, 20% o 25%.¹

Se puede observar como hasta un punto de corte del 10% la red se mantiene unida, aunque disminuye la cantidad de nodos. A partir de este punto, la red deja de estar débilmente conectada y, a la vez que sigue disminuyendo la cantidad de nodos, se observa mejor la importancia de ciertos nodos de alto grado. En particular, uno del continente americano (Estados Unidos) y otro del continente africano (Sudáfrica).

De esto se desprende la importancia de este parámetro en los resultados del modelo. Dado que no se trata de un modelo supervisado, no es posible optimizar el parámetro según alguna medida de bondad de ajuste del modelo. Por el contrario, debemos elegir el punto de corte que resulte más conveniente para los fines del análisis.

En el sentido de esto último, contamos con varias métricas que permiten caracterizar las propiedades del grafo, así como conocimiento previo respecto a propiedades conocidas y ampliamente aceptadas del comercio internacional. De esta forma, para determinar el punto de corte se propone como metodología utilizar una medida de centralidad que describe un atributo previamente conocido del comercio mundial, y buscar aquél punto de corte en que dicho atributo se expresa más plenamente. Una característica ampliamente estudiada del comercio internacional es su regionalismo (Das, 2004), es decir, la tendencia de los países de comerciar en mayor medida con sus vecinos (Head & Mayer, 2014). Si bien no contamos en este trabajo con información respecto de la distancia entre países, de esta característica emerge un fenómeno medible: el nivel de clustering de la red. En este sentido un valor alto en el clustering resulta deseable desde el punto de vista del objeto de estudio. En la figura 2.5(a) se observa el coeficiente de clustering para la red según el punto de corte utilizado, para cada año de la serie entre 1996 y 2016. Allí se observa que dicha métrica tiene una relación negativa no lineal con el punto de corte. Por lo tanto utilizar un punto de corte bajo resulta conveniente para captar el fenómeno. Sin embargo el máximo se encuentra en el punto de corte igual a cero, esto quiere decir que el regionalismo se expresa más cuando consideramos todos los intercambios, independientemente del monto comerciado. Como se observa en la figura 2.5(b), la cantidad de aristas del grafo es significativamente más alta para este punto de corte, y por lo tanto el grafo es sustancialmente más grande y denso, tal como se observa en la figura 2.5(b). En este sentido, si bien en la sección siguiente se continúa estudiando diferentes puntos de corte para ver el efecto que esto produce sobre el grafo y sus posibles interpretaciones, se decidió elegir un punto de corte por defecto de 1% para el resto del trabajo, dado que en dicho valor existe un punto

¹ En el anexo, sección E, se pueden observar el grafo en sus distintos puntos de corte, en una representación circular.

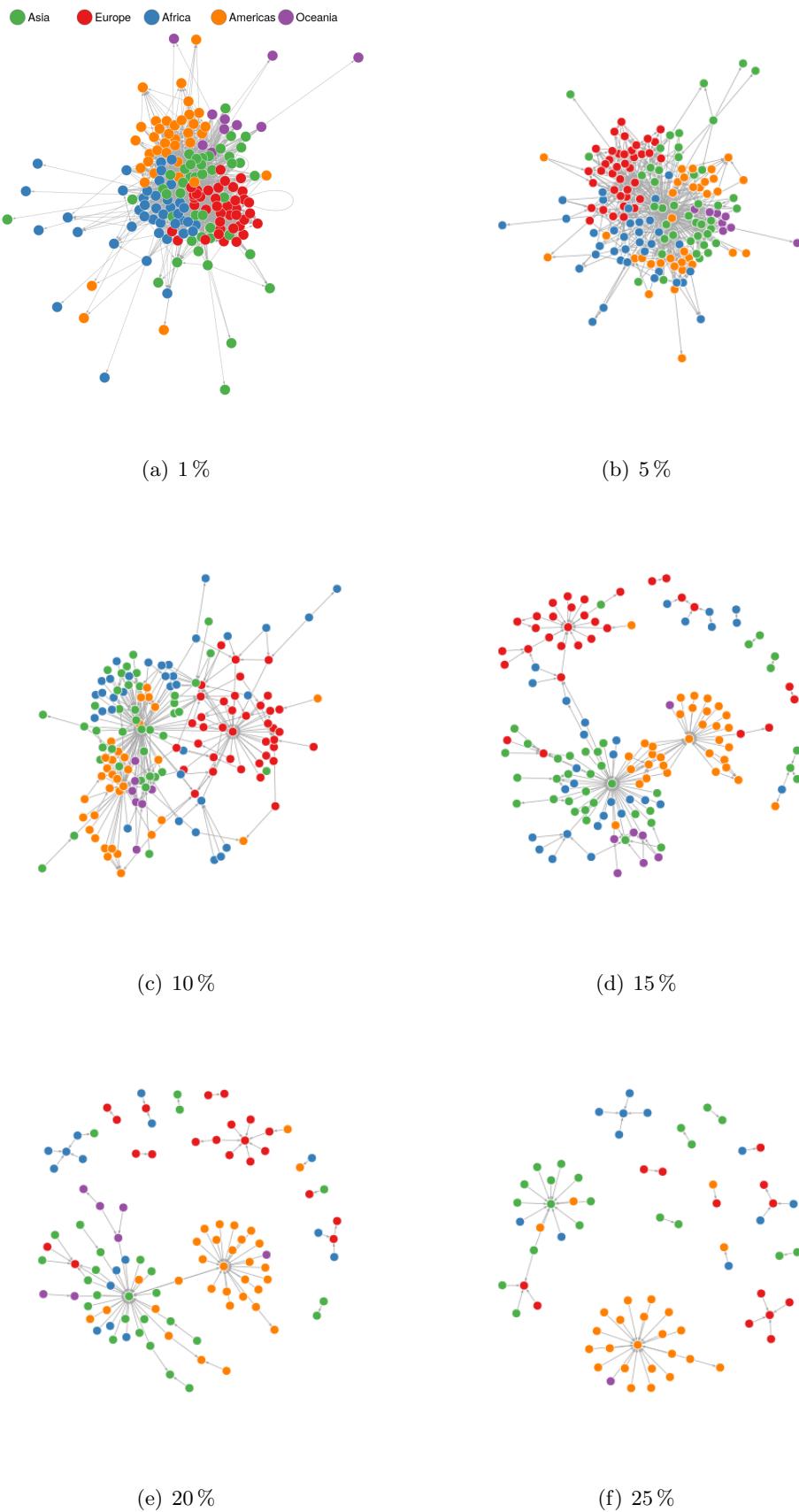


Fig. 2.4: Grafos. 2016, según punto de corte

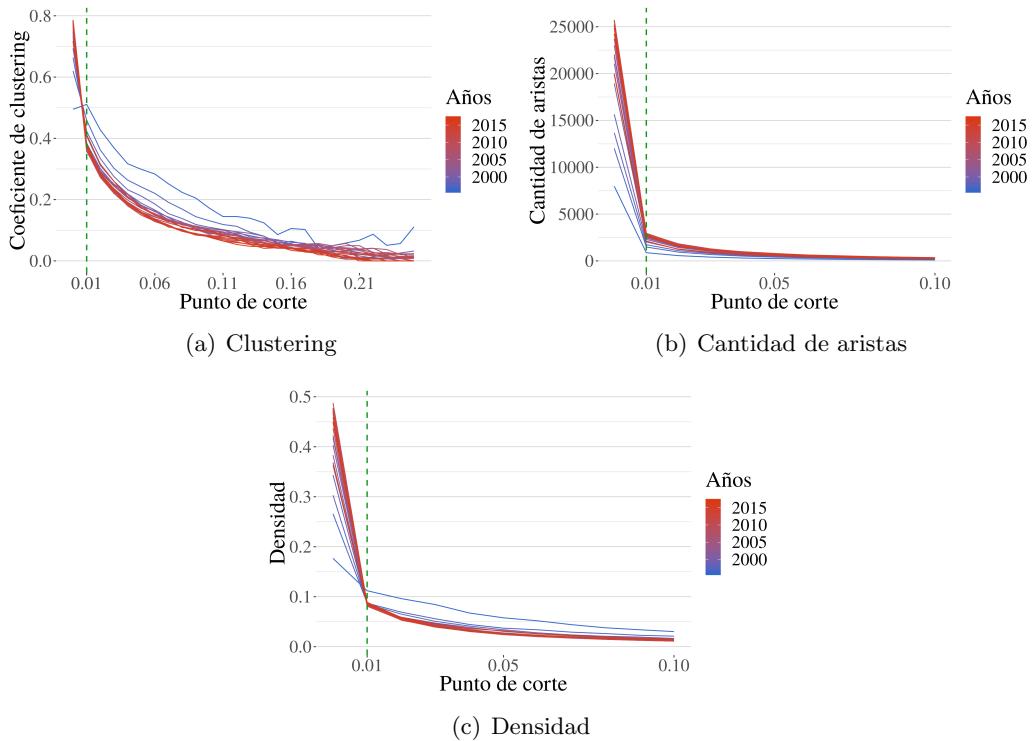


Fig. 2.5: Medidas de centralidad. Importaciones, 2016

de inflexión en la densidad y número de aristas, lo cual permite resaltar aquellas relaciones comerciales más significativas, a la vez que se maximiza el coeficiente de clustering².

2.4. Resultados

2.4.1. Correlación entre la representación de las exportaciones y las importaciones

Dada la potencial diferencia en el análisis si se tiene en cuenta las exportaciones y las importaciones, se incorporó esta disyuntiva como una nueva dimensión de estudio. Para ello, se analizó el grado total de cada nodo, para sucesivos puntos de corte, tanto para el grafo que surge de considerar las exportaciones, como para el grafo de las importaciones. En la figura 2.6 se pueden observar los resultados para el 1 %, 10 % y 20 %³. Vale mencionar que sólo se grafican los nodos de mayor grado, aunque se haya considerado al conjunto de nodos para el cálculo de la tendencia lineal por continente y el coeficiente de correlación de Pearson.

La correlación entre los grafos resultantes oscila entre el 0.72 y 0.96 si se consideran puntos de corte de entre el 1 % y el 25 %, por lo que se puede considerar que al observar las importaciones no se pierde demasiada información. Esto último es especialmente válido para el punto de corte del 1 % que se eligió como estándar.

² En el anexo, en la sección E se presenta a su vez la distribución acumulada de grado para los distintos puntos de corte.

³ En el anexo, sección E se pueden observar estos gráficos para el 5 %, 15 % y 25 %

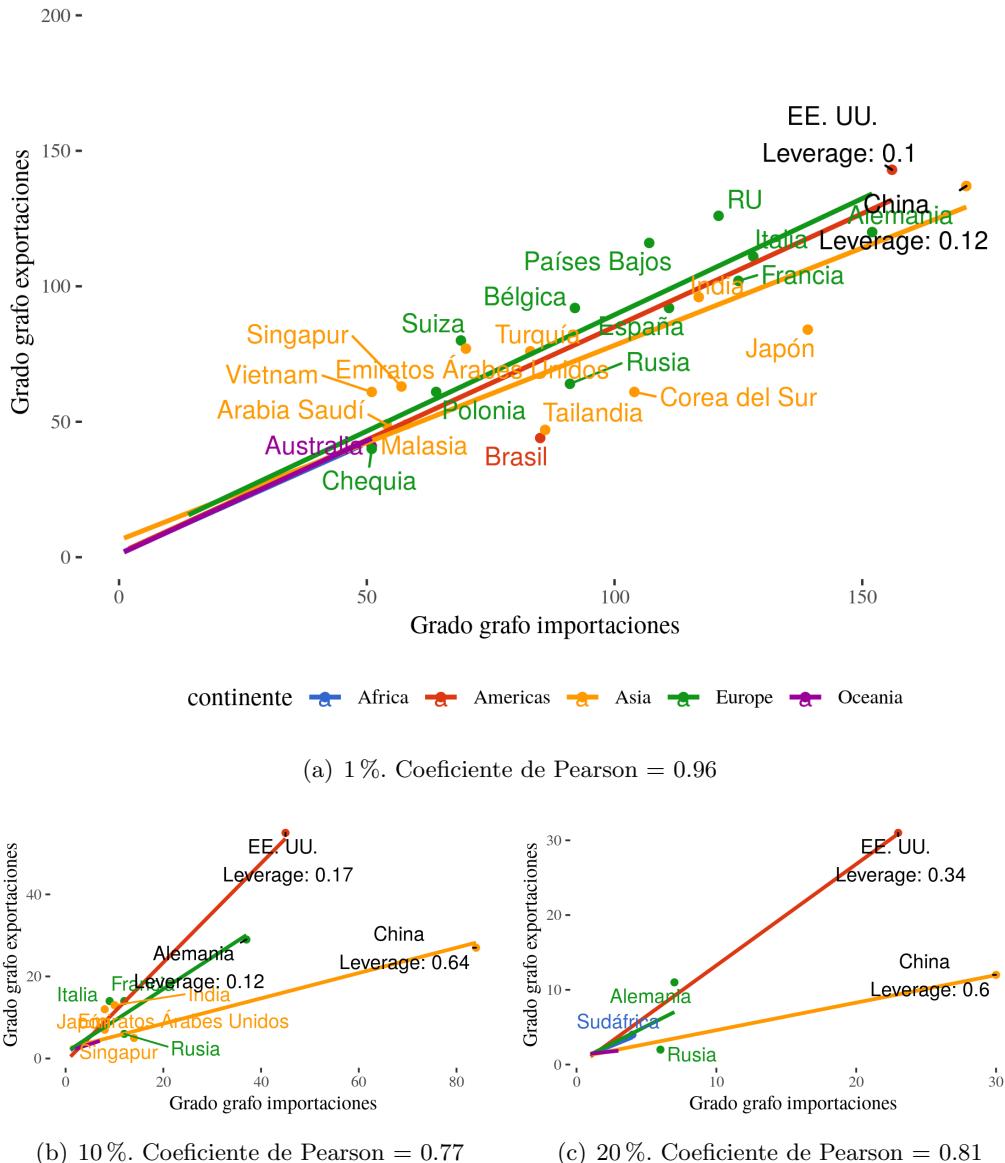


Fig. 2.6: Grado total. Grafo exportaciones respecto grafo importaciones. Año 2016

Tab. 2.1:

Orden	Grafo Importaciones		Grafo Exportaciones	
	País	G °entrada	País	G °entrada
1°	CHN	147	USA	124
2°	USA	139	CHN	114
3°	DEU	130	GBR	103
4°	JPN	117	DEU	98
5°	ITA	109	NLD	98

Sin embargo, surgen algunas diferencias interesantes. La regresión lineal para cada continente, muestra el diferente rol de éstos en el mercado mundial, donde Asia se presenta en el rol de productor de mercancías, mientras que Europa lo hace como un gran consumidor del mercado mundial. El caso americano tiene por un lado el rol preponderante de Estados Unidos, con un comportamiento similar al de Europa (con independencia del umbral del 0.25), mientras que el resto del continente, junto con Oceanía y África, juegan un rol secundario. En el caso europeo, se observa como Rusia tiene un comportamiento más acorde al continente asiático, mientras que en este último se destaca en particular China como gran productora.

Sin embargo, el rol de cada país varía también respecto al umbral considerado. En el caso de Japón, en puntos de corte bajos, donde hay relaciones comerciales de importancia moderada para los países, se presenta jugando un rol de productor más que de consumidor, incluso más que la media asiática. A medida que se aumenta el punto de corte, y por lo tanto se consideran exclusivamente relaciones de carácter más dependiente, Japón se presenta como un país consumidor, incluso más que los países europeos. Esta transición estaría indicando que Japón, productor de productos industriales y tecnológicos, provee a muchos países del mundo de este tipo de mercancías, aunque éstos no dependen en su consumo de Japón, mientras que como consumidor de materias primas, este país establece relaciones comerciales con países que sí dependen de Japón para poder colocar sus mercancías. Más interesante aún, a medida que se consideran relaciones de tipo más dependiente, aparece como destacado el rol de Estados Unidos, relegando a los países europeos y asiáticos. Por su parte, también surge Sudáfrica como un jugador importante del mercado mundial, cuando se consideran este tipo de relaciones de mayor dependencia. Esto implica que estos dos países tienen una forma de interacción en el mercado mundial marcadamente diferente a la de los países europeos y asiáticos.

Por último, en la figura 2.6 también se puede observar el leverage de los países con mayor apalancamiento para la regresión del grado de las exportaciones respecto del de las importaciones, sin diferenciar por continente. Cuando se considera un punto de corte del 1 %, China tiene el mayor leverage, con un valor algo mayor que Estados Unidos, aunque relativamente bajo en términos absolutos. Sin embargo, a medida que se consideran puntos de corte más altos, este país se constituye como el país con mayor apalancamiento en la regresión, a la vez que el valor aumenta fuertemente.

En la tabla 2.1 se presenta el top 5 de países de mayor centralidad de grado para el 2016, en el grafo de importaciones y en el de exportaciones.

La tabla 2.1 muestra como China, Estados Unidos y Alemania son nodos de una centralidad importante, tanto como consumidores, como en su rol de productores. Sin embargo, China se destaca más en este último sentido, dado que su centralidad en el

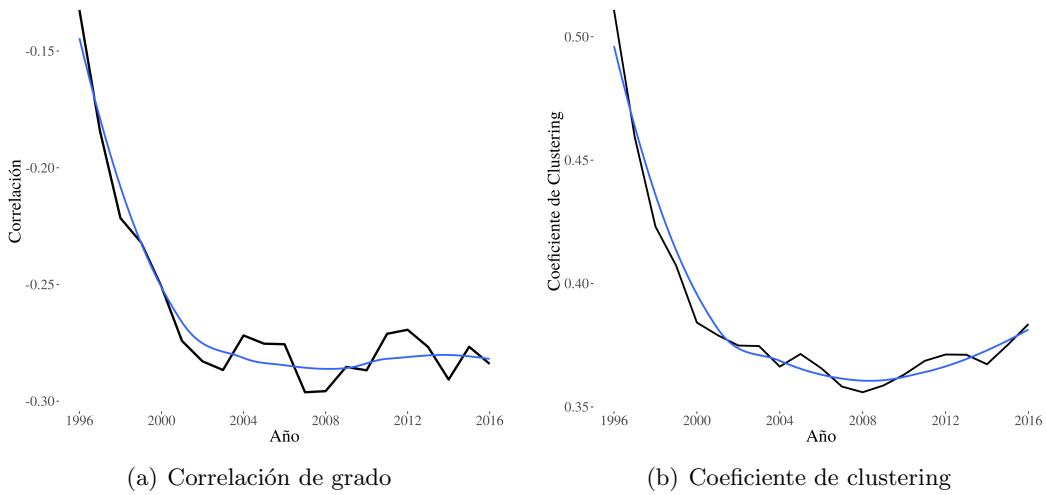


Fig. 2.7: Importaciones. Umbral 1%. 1996-2016

grafo de las importaciones es mayor, respecto a los demás países, que en el grafo de las exportaciones. Por su parte, Estados Unidos marca una importante hegemonía como consumidor respecto de los demás países, tal como se observa en su centralidad en el grafo de las exportaciones, aunque no conserva tal preponderancia como productor global.

2.4.2. Evolución de mediano y largo plazo del comercio internacional

Análisis mediano plazo

Lo que resta del análisis se considerará exclusivamente el punto de corte del 1%, incorporando la dimensión temporal. Para esto, se estudiará en primer lugar el movimiento de algunas medidas de resumen a lo largo del tiempo, para luego analizar la distribución de ciertas medidas de centralidad y el cambio en el rol de los nodos centrales en durante el período.

Medidas de resumen de la red

En la figura 2.7 se presentan los valores de algunas medidas de resumen entre el período de 1996 y 2016, considerando un punto de corte del 1% para las importaciones⁴.

Como se puede observar, las características topológicas del grafo varían fuertemente en el período 2007-2009, durante la gran recesión económica, que se caracterizó por un aumento en el proteccionismo económico de ciertos países centrales, reflejado en el “Buy American”⁵ o “Buy China”⁶.

La figura 2.7(a) muestra la caída de la correlación de grado en el año 2007, que implica una transición circunstancial hacia una economía menos selectiva. Es decir, un comercio internacional donde aumentan las relaciones de nodos de mayor grado con nodos de menor grado, respecto de las relaciones entre nodos de similar magnitud.

Por su parte, la figura 2.7(b) muestra como el coeficiente de clustering tiene una tendencia decreciente hasta 2008, para luego comenzar a crecer nuevamente. Esto se puede

⁴ en el anexo, sección E se presentan otras medidas complementarias

⁵ <http://edition.cnn.com/2009/POLITICS/02/05/senate.buy.american/index.html>

⁶ <https://economix.blogs.nytimes.com/2009/06/29/buy-china/>

interpretar como una de-segmentación del comercio internacional hasta la crisis, con al cual comienza un proceso de re-segmentación.

Distribución de los nodos

A continuación se presenta la distribución de los nodos según ciertas medidas de centralidad a lo largo de los años, destacando aquellos países con mayores valores de centralidad.

En primer lugar, lo que se observa en la figura 2.8 es que la distribución de los nodos de acuerdo a las distintas medidas de centralidad no varía sustancialmente de año en año, ni en su rol como consumidores globales, ni en su rol como productores globales. Por su parte en 2.8(a) se observa la distribución de grado de los nodos del grafo de importaciones, donde se ve como los tres países de mayor grado tienden a alejarse del centro de masa de la distribución. Este podio se disputa entre Alemania, Japón, Estados Unidos y, a partir del 2002, China, quien a su vez tiende a alejarse de éstos posteriormente. Por su parte, la figura 2.8(b) marca el rol de consumidores globales. Aquí, mientras hasta 2007 Estados Unidos y Gran Bretaña ocupan los lugares principales en la centralidad de autovalor, a partir de este año comienza a tener un rol más destacado Francia, los países bajos, y luego China.

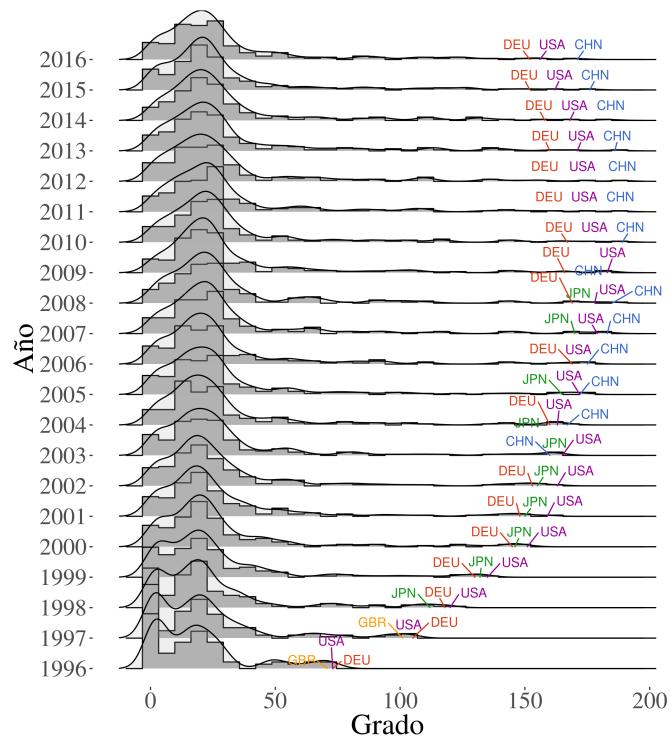
Es de resaltar el aumento en la centralidad de China para el comercio internacional, tanto como productor como consumidor. En particular, este país desplaza a Estados Unidos de ser el nodo de mayor centralidad en muchas medidas. Por ejemplo, la figura 2.9 para el caso del autovalor ponderador por el valor del intercambio, Estados Unidos ocupa el primer lugar hasta el año 2006, para luego ser reemplazado por China, un país que diez años antes tenía un valor de centralidad cuatro veces menor al de Estados Unidos. En ese mismo año China también pasa a ocupar el primer lugar en dicha métrica para el grafo de las importaciones. La medida de centralidad del autovalor es particularmente relevante, sobretodo todo ponderada por el valor de las transacciones, porque considera no sólo la importancia del nodo, sino también la importancia de aquellos nodos con que se relaciona. En este sentido, se destaca el rol de México, Japón, Canadá, junto con China y Estados Unidos como grandes productores mundiales, a la vez que aparece, sobre los últimos años, Corea como una nueva potencia en este aspecto.

Análisis de largo plazo

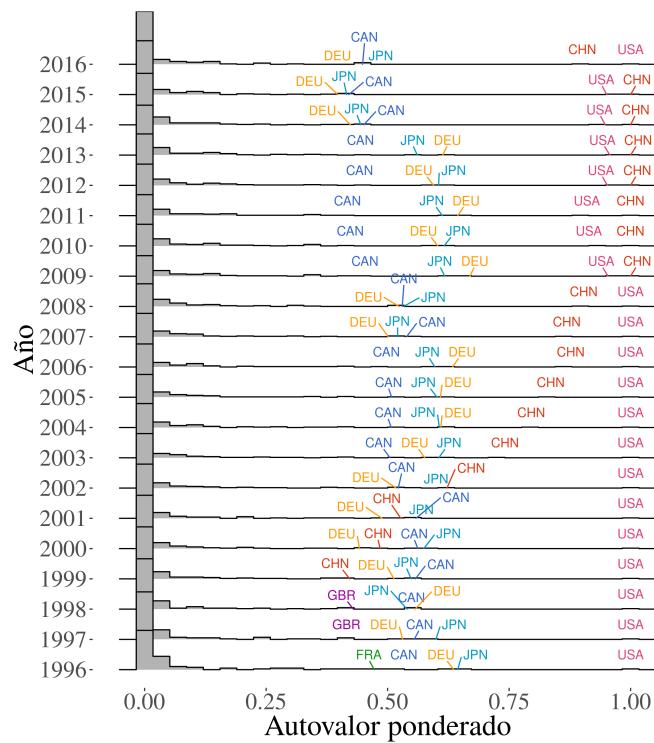
Otro análisis que resulta interesante es la evolución de la red en un período de largo plazo. Esto nos permite encontrar los cambios estructurales del comercio internacional, a la vez que da cuenta de qué elementos de la topología de la red son robustos, y cuales contienen una tendencia histórica. Para esto, tomamos la serie elaborada por Gleditsch (2002), que reúne información del comercio internacional entre 1948 y 2000. A continuación se analizan los elementos estructurales de la red, para luego analizar la evolución de algunos países seleccionados dentro del grafo

Estructura de la red

Como se observa en la figura 2.10, en la red ocurren importantes cambios en el período que va de 1948 al año 2000. La figura 2.10(a) muestra la evolución del coeficiente de clustering de la red. Este, como se mencionó previamente, indica la tendencia de la red a formar vecindarios bien delimitados. La fuerte caída durante las décadas del cincuenta y



(a) Grado, importaciones



(b) Autovalor, exportaciones

Fig. 2.8: Distribución medidas de centralidad según año. umbral 1 %

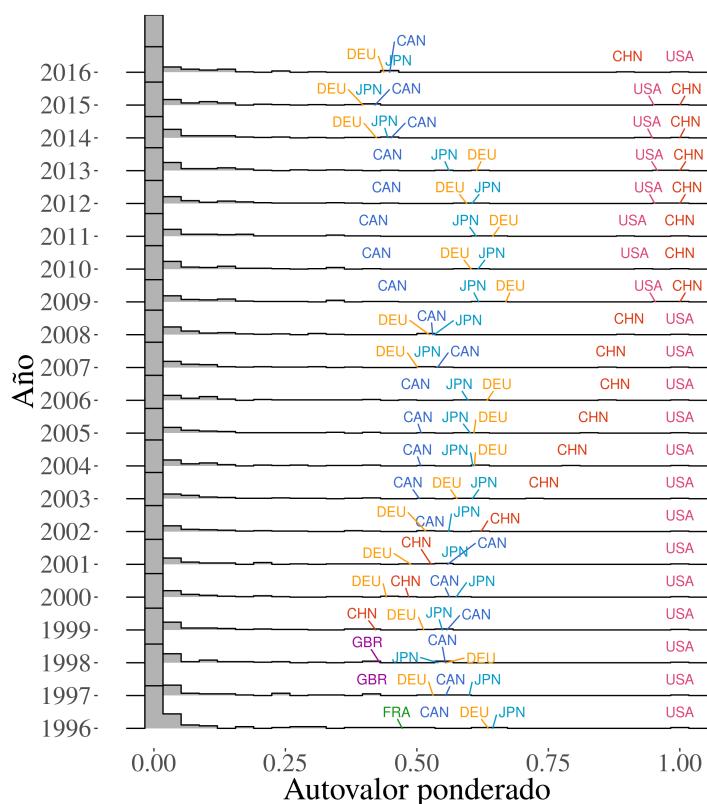


Fig. 2.9: Distribución autovalor ponderado, exportaciones según año. umbral 1 %

sesenta reflejan la nueva división internacional del trabajo, en dónde las relaciones comerciales se alejan del regionalismo, y el comercio inter-continental aumenta en volumen. En concreto, esta expresando la incorporación de Asia como productor de mercancías para el mercado mundial.

La figura 2.10(b) muestra la evolución del coeficiente de correlación de grado. Este indica el grado de homofilia de la red, es decir, la tendencia de los nodos de mayor grado a conectarse con otros nodos de similares características. Recordemos que al tomar valores negativos, esta correlación indica una tendencia de los países de mayor grado a relacionarse con países de menor grado. La caída de este valor en las décadas del cincuenta y sesenta expresa un incremento en esta característica de la red.

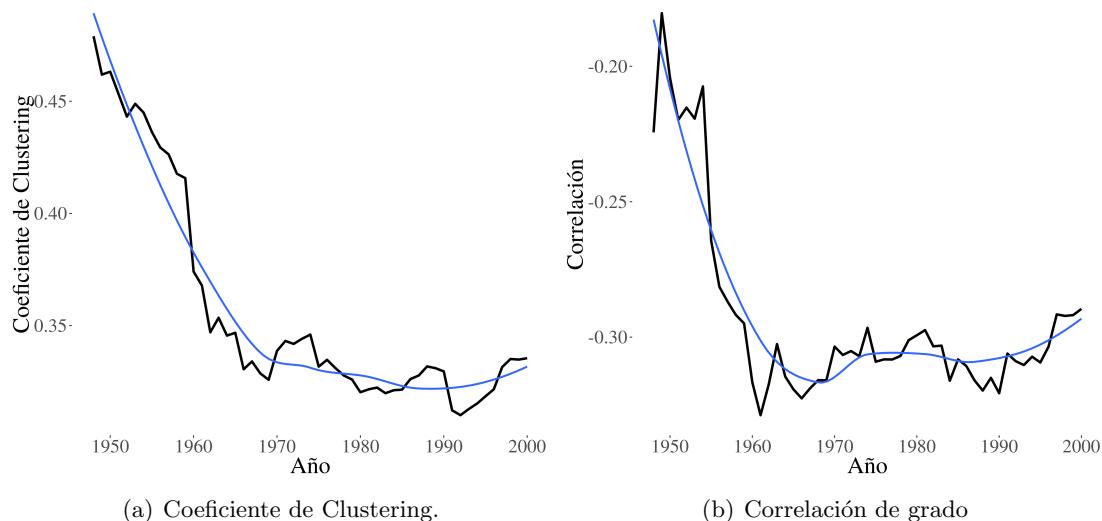


Fig. 2.10: Evolución de la estructura de la red. Importaciones. Umbral 1%. 1948-2000

Estas medidas refieren a características generales del grafo⁷. También es posible analizar, dentro del grafo, el movimiento en la centralidad de subconjuntos del mismo. En particular, en la figura 2.11 se muestra la evolución de la centralidad promedio por continente para las medidas de autovalor, cercanía y grado. En primer lugar, resalta la robustez de la centralidad de cercanía, que evoluciona uniformemente para todos los continentes. La disminución de esta medida indica un crecimiento del tamaño de la red⁸.

La centralidad de autovalor, por su parte, muestra un cambio importante en el rol de los países en la economía mundial. Mientras que Oceanía pasa del segundo al quinto lugar en el período, Asia hace lo opuesto, marcando la incorporación al mercado mundial de este continente como productor de mercancías. África y América, por su parte, si bien siguen un distinto recorrido a lo largo del período, comienzan y terminan el mismo con valores muy similares. Europa mantiene la mayor centralidad promedio a lo largo de todo el período, en las tres métricas. Por último, el recorrido de Asia y Oceanía es similar para la centralidad de grado.

Evolución de los países

⁷ en el anexo, sección E se presentan otras métricas complementarias.

⁸ Que parcialmente puede deberse a la incorporación de nuevas fuentes de información, en particular para los años 1949, 1960 y 1991

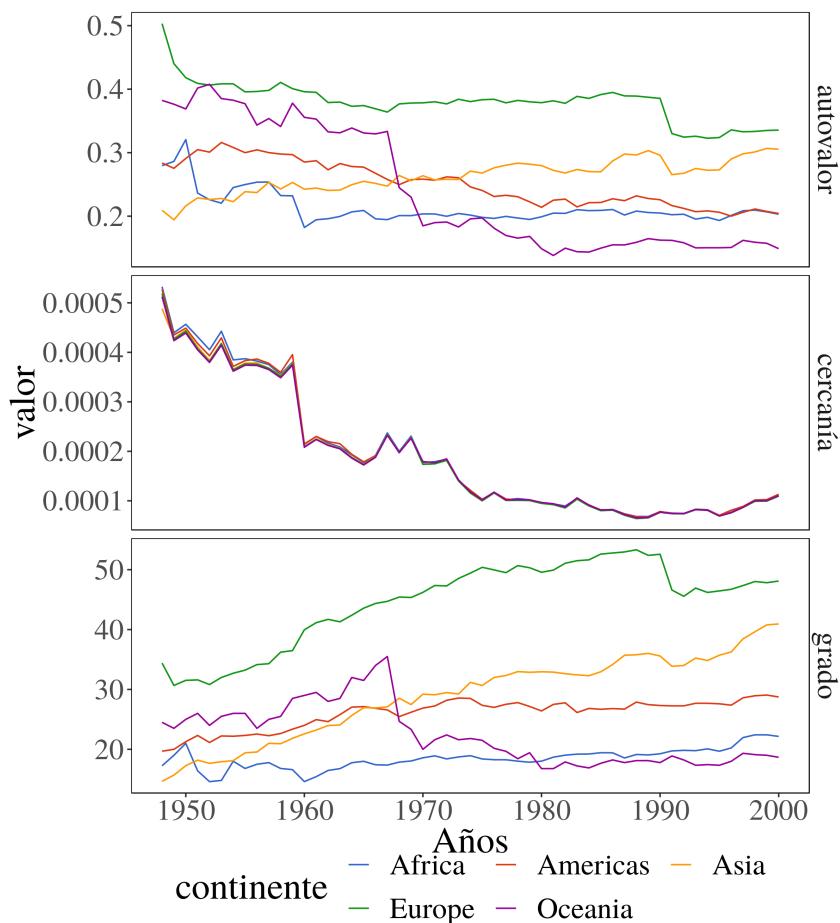


Fig. 2.11: Evolución de la centralidad promedio por continente. Impo. Umbral 1%

Otra característica interesante del análisis de grafos es que permite seguir la evolución de un país en particular dentro del mismo, de forma tal que se puede observar la evolución de la centralidad de algunos nodos en el tiempo. Para ello, se elaboró una aplicación en *Shiny*⁹ en la cual se puede explorar los resultados para los diferentes países, para sus exportaciones e importaciones, según la centralidad de Grado, Intermediación, Autovalor y Autovalor ponderado. A continuación se ejemplifica el uso de la herramienta en el caso de algunos países latinoamericanos, asiáticos, y Estados Unidos.

En la figura 2.12 se muestra la distribución de las medidas de grado y de autovalor ponderado, y se marca la posición de México Brasil y Argentina en dicha distribución. En 2.12(a) observamos la cantidad de conexiones de cada país en el grafo de las importaciones. Aquí se observa como México comienza la serie en la moda de la distribución de grado, mientras que Brasil y Argentina se encuentran a la derecha de la misma. En la evolución se vé como la distribución se convierte en más leptocúrtica, y, por un lado Brasil incrementa la cantidad de nodos considerablemente, mientras que por otro lado Argentina mantiene a lo largo de toda la serie la misma cantidad de socios comerciales. Por su parte, México, pasa a ubicarse en la cola derecha de la distribución, más cercano a la Argentina. En 2.12(b) se observa el autovalor ponderado por el volumen comerciado. Considerando esta

⁹ https://diegokoz.shinyapps.io/Distribucion_nodos_wrdtrade/

medida, la evolución pareciera ser diferente. Mientras Argentina se mantiene en el mismo lugar de la distribución, Brasil se repliega hacia valores similares a los Argentinos, y México se despega de la moda de la distribución a partir de los años noventa. Al considerar no solo el grado, sino también el grado de los vecinos, y el volumen comerciado, se destaca el proceso de la *Maquila Mexicana*, en dónde el comercio entre este país y Estados Unidos, un nodo sumamente importante, es muy difundido.

Por último, resulta interesante complementar el análisis de la figura 2.9, donde se veía como China supera a Estados Unidos en ciertas medidas de centralidad para el año 2006, con un análisis de dicha relación en el período de 1948 a 2000. Para ello, en la figura 2.13(a) se observa la distribución de la centralidad de autovalor, y el detalle de estos dos países.

En todo el período se observa la predominancia de Estados Unidos como un nodo central del grafo. China, por su parte, comienza en la moda o incluso en la cola izquierda de la distribución. Para luego ubicarse del lado derecho de la misma, entre los sesentas y setentas, aunque aún con valores relativamente bajos. A partir de la década del ochenta, comienza, de forma lenta pero acelerada, a escalar posiciones de mayor centralidad dentro de la red. Por su parte, en la figura 2.13(b) se observa la evolución de China junto a Corea y Japón. Aquí se puede ver como no es solo China quién aumenta la centralidad en el grafo, sino Asia en su conjunto. Japón lo hace primero, desde el comienzo de la serie, en los cincuentas, y llega a valores cercanos a 1, el máximo valor posible de esta medida, para los ochentas. Por su parte, Corea también sigue el mismo camino, aunque comienza más tarde, y desde una posición de menor centralidad que los otros países.

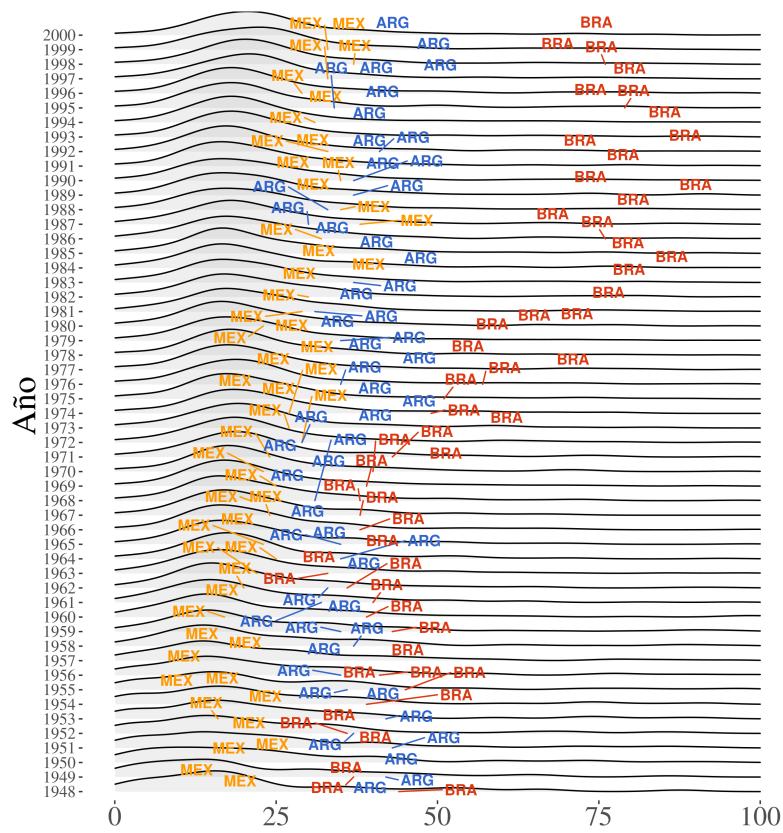
2.5. Conclusiones

En el presente capítulo se propuso una primera aproximación de la teoría de grafos como herramienta para la caracterización del comercio internacional, utilizando los datos agregados a nivel país. Para ello se construyó un grafo dirigido no ponderado. Se utilizó al año 2016 como punto de referencia para determinar determinar los mejores hiperparámetros del modelo. En particular, de este análisis resultó que el 1% constituye un umbral posible para el punto de corte, aunque considerar este hiperparámetro como una dimensión de estudio también permite observar las distintas formas de relación comercial entre los países, y por lo tanto se logra enriquecer la discusión.

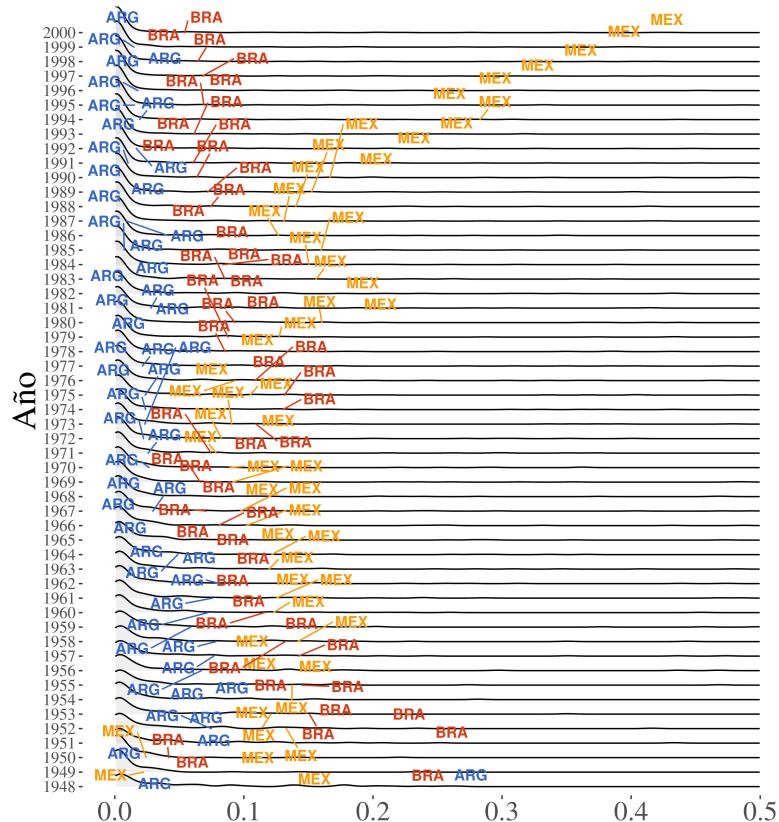
Por su parte, la dimensión temporal en un período de veinte años mostró que las medidas de resumen de un grafo pueden poseer un potencial para la descripción de la crisis del año 2009. El análisis de una serie de largo alcance, por su parte, permite ver los cambios estructurales del comercio internacional.

Se observó también que la distribución de las medidas de centralidad no pareciera variar en el mediano plazo, aunque sí existe un movimiento visible en los actores principales de la red, lo cual muestra una riqueza en el análisis para describir los cambios de la economía mundial, y los roles cambiantes que en esta juegan los distintos recortes nacionales.

Finalmente, el trabajo realizado deja una contradicción planteada respecto del rol de ciertos países como consumidores, y de otros como productores, dado que tal estructura no sería sostenible de forma prolongada.



(a) Grado



(b) Autovalor Ponderado

Fig. 2.12: Argentina, Brasil y México. Importaciones. Umbral 1%

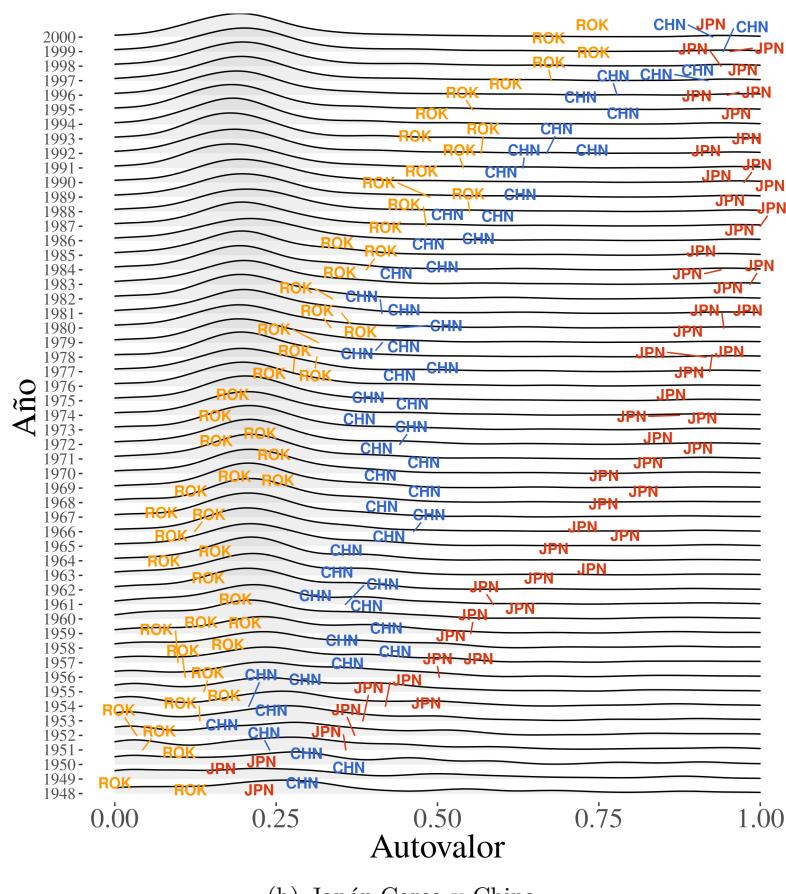
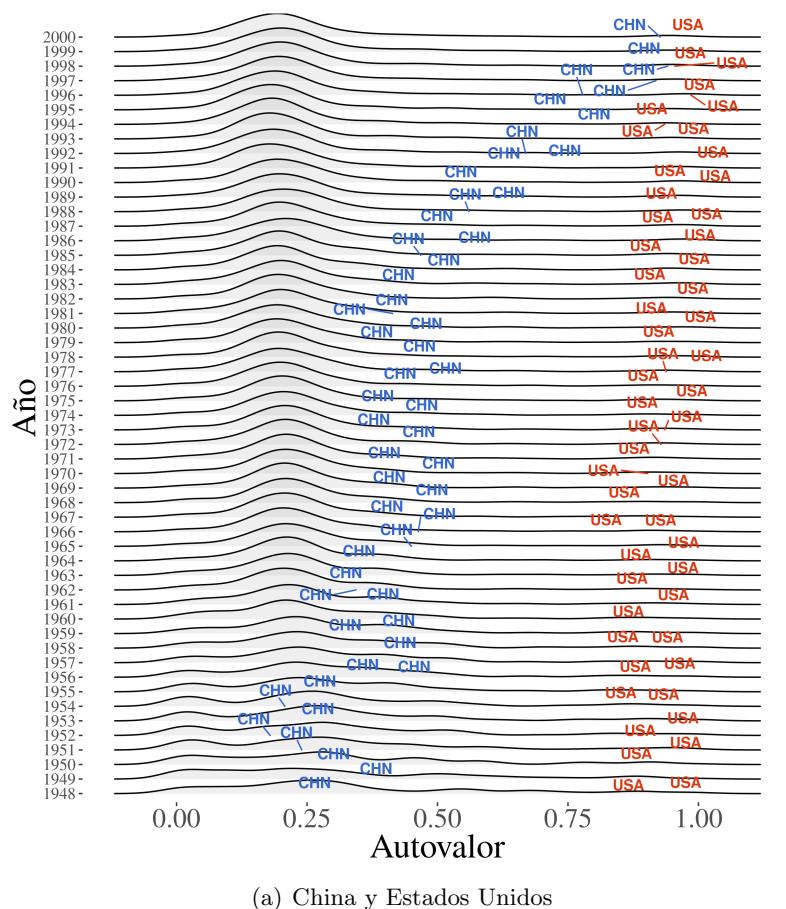


Fig. 2.13: Autovalor. Importaciones. Umbral 1 %

3. COMERCIO A NIVEL PRODUCTO

3.1. introducción

En el presente capítulo¹ se analiza la estructura del comercio internacional a partir de la estructura de la canasta exportadora de los países. El rol que ocupan los países en el mercado mundial está profundamente determinado por cómo se insertan en las cadenas globales de valor, y por ende en los tipos de mercancías que producen para el mercado mundial (Coe et al., 2004; Gereffi et al., 2005; Gereffi, 1994). Por ello, luego de analizar la centralidad de los países desde la topología de la red, resulta de interés interrogarse sobre los determinantes de dicha posición, a partir de su estructura productiva.

Es amplia la literatura que estudia el comercio internacional a partir del nivel producto (Balassa, 1965; Lall, 2000; Lall et al., 2006; Haveman & Hummels, 2004) y en particular existe una extensa bibliografía reciente en la cual se analiza este fenómeno a partir de construcción de un grafo bipartito entre países y productos (Guan et al., 2018; Straka et al., 2017; Ferreira & Others, 2016; Caldarelli et al., 2012). Destaca el trabajo de Hidalgo (2009) (Hidalgo & Hausmann, 2009; Hidalgo et al., 2007), donde se define el espacio de producto en términos de la complejidad de la producción y las influencias que ello tiene sobre el desarrollo de los países.

El presente capítulo propone, luego de un análisis exploratorio de la información a partir de treemaps, dos modelos alternativos para abordar el fenómeno. En primer lugar, inspirado en el modelo propuesto para detección de tópicos en minería de textos (Blei, D. M.; Ng, A. Y.; Jordan, 2003), se propone un modelo generativo bayesiano para detección de dimensiones latentes en el espacio de productos, que permite construir un nomenclador alternativo a los convencionales, construido de forma automática basado en los datos. Luego, con dichas dimensiones latentes, *componentes*, se analiza la participación de las mismas en la canasta exportadora de los países. En segundo lugar, fuertemente inspirado en la propuesta metodológica de Hidalgo & Hausmann (2009), se realizan diferentes técnicas de clustering y detección de comunidades, tanto sobre la proyección del grafo bipartito entre países y productos, como a partir de una matriz de distancias del espacio de productos.

3.2. Análisis exploratorio de datos

El análisis exploratorio de la información desagregada a nivel producto implica un mayor nivel de complejidad, dado que se incorpora una nueva dimensión, de alta cardinalidad, al estudio. Dado que la información a nivel de comercio agregado entre países ya fue realizada en el capítulo previo, en este análisis exploratorio se realizará foco en comprender la composición de las canastas exportadoras e importadoras de los países.

Dado que los nomencladores de productos en su nivel desagregado implican una alta cardinalidad, el estudio de su distribución resulta inabordable. Por ello, tradicionalmente el análisis económico recurre a diferentes niveles de agregación. Para realizar el análisis exploratorio de esta sección recurrimos al nomenclador elaborado por Molinari & de Angelis

¹ Parte del trabajo del presente capítulo se realizó en el marco del proyecto de Investigación Científica y Tecnológica (PICT) 1085-2016. 'Abordando la restricción externa en América Latina a partir de la integración regional: integración productiva, cooperación Sur-Sur y financiamiento para el desarrollo'.

(2016) a partir del concepto de *cadenas productivas*. Este nomenclador tiene dos niveles de agregación: Cadenas y Subcadenas (ver Apéndice).

La figura 3.1 muestra la distribución de las exportaciones según Cadenas, Subcadenas y Usos, para el promedio mundial durante el período 1996-2016. En 3.1(a) se puede ver que las dos cadenas más importantes son las de Bienes de Capital y Otras Manufacturas, dentro de las cuales las subcadenas que destacan son los equipos eléctricos y de transporte. Les sigue la cadena de insumos difundidos, donde destacan los metales y químicos. También son importantes las subcadenas de autopartes, autos, y dentro de la cadena de combustibles, el gas y petróleo. En la figura 3.1(b) las cadenas se subdividen según el uso que se le da a los productos: Productos Primarios, Bienes de Capital, etc (ver Apéndice). Vale mencionar que la cadena Bienes de Capital no incluye exclusivamente productos de este tipo, tal como se observa en la figura, dado que las cadenas hacen referencia a las Cadenas de valor (Humphrey & Schmitz, 2000), es decir a la rama de la producción a la que pertenecen los productos del final de la cadena. Allí podemos observar que en la cadena de Otras manufacturas destacan los productos semiterminados y bienes de consumo, mientras que en la cadena de Bienes de capital destacan su homónimo y las partes y componentes. A su vez en la cadena de insumos difundidos destacan los productos semiterminados ampliamente, mientras que en la industria automotriz se exportan mayoritariamente partes y componentes.

La figura 3.2 muestra la distribución de las exportaciones e importaciones según continente y país exportador, para el promedio 1996-2016. Allí destaca el mayor volumen de exportaciones que importaciones de Asia, y el mayor volumen de importaciones que exportaciones de Estados Unidos, siguiendo las mismas conclusiones que el capítulo 2.

Sin embargo, la riqueza en el análisis a nivel producto se encuentra en comprar la distinta composición de la balanza comercial de los países. Ya sea la diferencia entre la composición de sus exportaciones respecto de sus importaciones, como de estos elementos respecto de otros países o regiones. Dado que dicho análisis implica la comparación de una multiplicidad de treemaps, se decidió elaborar una herramienta interactiva para el análisis. La misma fue elaborada utilizando la librería *shiny* (Chang et al., 2018) y se encuentra disponible en <https://treemaps.shinyapps.io/treemaps/>. Dados los límites del servidor y el interés de estudiar la herramienta para un caso específico, el análisis se centra en el caso latinoamericano, tanto para el subcontinente en conjunto como para los países que lo integran. El objeto de análisis es estudiar las posibilidades de integración regional de la producción, y para ello analizar el comportamiento diferencial de las canastas exportadoras e importadoras de los países latinoamericanos cuando comercian entre sí respecto de su comercio con el resto del mundo. A su vez, dada la importancia del comercio de esta región con China, se decidió dividir la información del resto del mundo exceptuando a China y este país de forma individual.

En la figura 3.3 se pueden observar los treemaps de cadenas y subcadenas para 2016 del total de los países sudamericanos, según su destino. En la figura 3.3(a) se ve como la canasta exportadora de los países latinoamericanos varía según si su destino se encuentra dentro o fuera de la región, y en particular si son exportaciones hacia China. En particular, el comercio intra-regional tiene un componente importante de bienes de capital, y del sector automotriz, ya sean autos terminados o autopartes. Por su parte, en el comercio con el resto del mundo estos componentes cumplen un rol secundario, mientras que se destacan las cadenas agroindustriales, de combustibles e insumos industriales. Respecto del resto del mundo, el comercio con China resalta por las exportaciones de la subcadena



(a) Cadenas y Subcadenas



(b) Cadenas y Usos

Fig. 3.1: Treemaps por tipos de productos. Exportaciones. 1996-2016. Total mundial

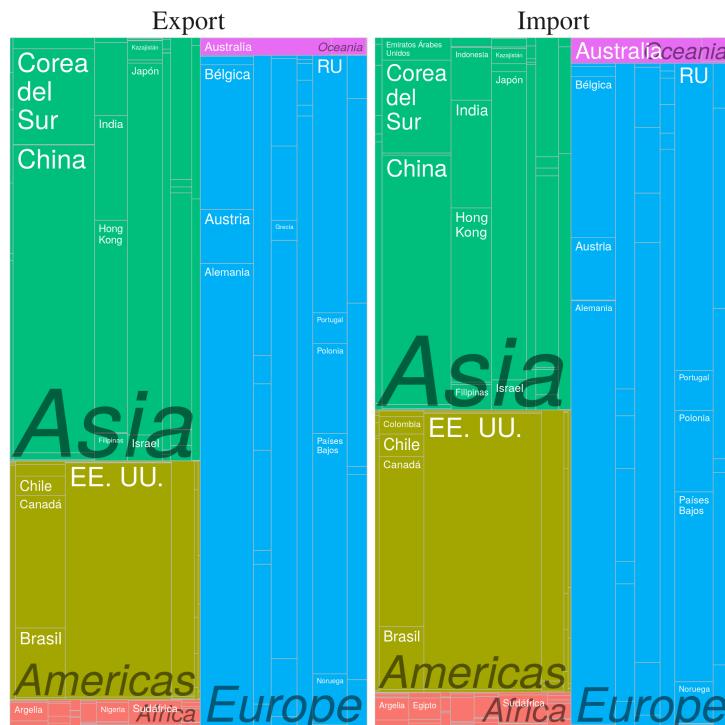
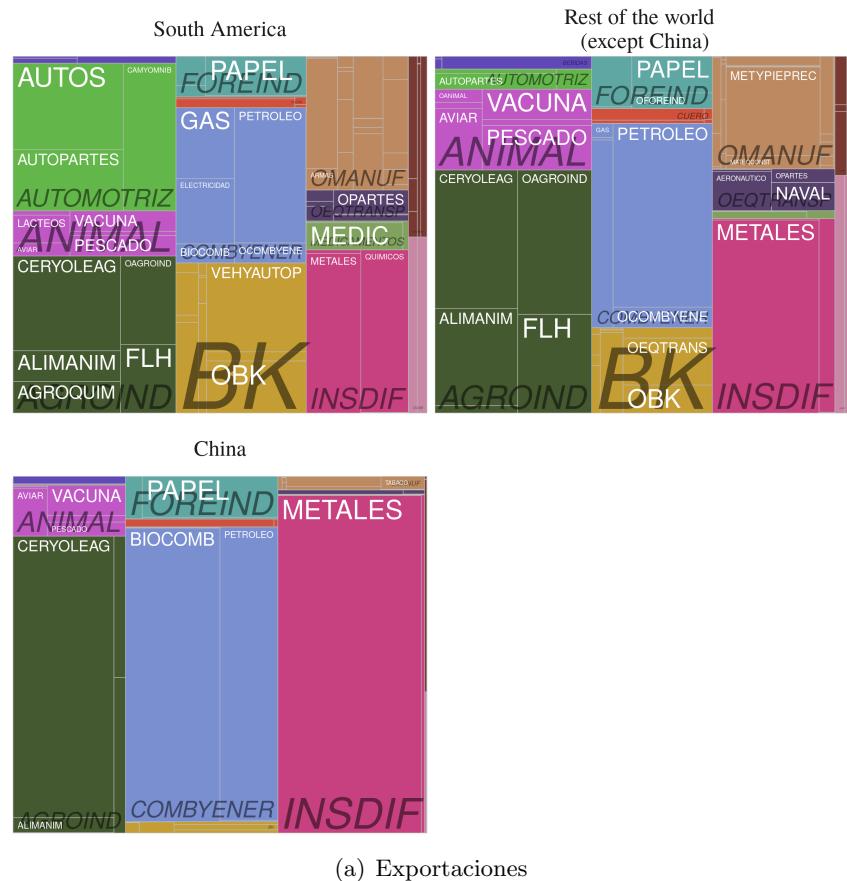


Fig. 3.2: Treemaps por Continentes y países. 1996-2016. Total mundial

de biocombustibles, cereales y oleaginosas y Metales. Las importaciones de estos mismos países se pueden observar en la figura 3.3(b). Naturalmente los treemaps de las exportaciones e importaciones son muy similares, ya que la única diferencia es el país que registra la operación. Sin embargo, en las importaciones del resto del mundo destacan las cadenas automotriz y de bienes de capital. A su vez, resulta de interés el cambio de composición de las cadenas: Mientras que las exportaciones de insumos industriales hacia el resto del mundo son mayoritariamente metales, en las importaciones destacan los productos químicos. Por su parte, mientras que en la cadena de otras manufacturas se exportan hacia el resto del mundo metales y piedras preciosas, se importan dentro de esta cadena equipos eléctricos. A su vez, del resto del mundo se importan medicamentos, aunque este rubro no aparece en el treemap de las exportaciones. Cabe destacar que la región exporta e importa petróleo hacia el resto del mundo, aunque el comercio intrarregional de este producto es menor. Existe, por lo tanto, una potencialidad de integración comercial dentro de la región para este producto en particular. El comercio con China está particularmente orientado a la venta de materias primas y la compra de productos industriales. Destacan especialmente las exportaciones de cereales y oleaginosas, biocombustibles y metales; mientras que se importan equipos eléctricos, bienes de capital, calzados y químicos.

La figura 3.4 muestra los treemaps exclusivamente según los usos de los productos, para el total de la región, según su destino, en el año 2016. En 3.4(a) se puede observar las exportaciones. Los usos de los productos exportados varían fuertemente según su destino. Los productos semiterminados constituyen la mayoría de las exportaciones tanto hacia el interior de la región, como con el resto del mundo. Sin embargo, mientras que los bienes de consumo resultan de particular importancia dentro de la región, los productos primarios



(a) Exportaciones



(b) Importaciones

Fig. 3.3: Treemaps de Cadenas y Subcadenas.2016. Total Latinoamérica

lo son respecto del resto del mundo. El comercio con China destaca esta tendencia, donde la mayoría de las exportaciones provienen de productos primarios. En 3.4(b) se observa la distribución de las importaciones según su uso. Allí se ve que las importaciones de productos semiterminados desde el resto del mundo supera a las exportaciones, si se lo compara con 3.4(a). A su vez, prácticamente no se exportan productos primarios, pero sí partes y componentes, y bienes de capital. El comercio con China se encuentra equitativamente distribuido entre bienes de capital, productos semiterminados, de consumo y partes y componentes. Es notorio que no se importa prácticamente productos primarios, aunque estos constituyen la base de las exportaciones a dicho país.

En la figura 3.5 se observa la distribución de los socios comerciales de Sudamérica en el 2016. Esto es, los países destino de las exportaciones y origen de las importaciones. En 3.5(a) se excluye al resto del mundo y China, para observar la distribución interna del comercio. Naturalmente Brasil es el mayor socio comercial interno, para ambos tipos de flujo comercial, aunque su importancia es mucho mayor como exportador que como importador, dado que es el origen del 40 % de las importaciones de los demás países, y el destino del 25 % de las exportaciones. En la figura 3.5(b) se puede observar el mismo gráfico, pero incluyendo al resto del mundo. Allí se ve que tanto las exportaciones como importaciones hacia el resto del mundo exceptuando a China constituyen más de un 60 % del comercio, y China exclusivamente representa casi un 20 % del comercio de la región, superando ampliamente a Brasil.

De las figuras 3.3, 3.4 y 3.5 se desprende que la región de Sudamérica tiene una baja integración regional en tanto una pequeña proporción de su comercio se realiza entre países de la región. A su vez, mientras el comercio intrarregional se realiza fundamentalmente sobre productos de alta complejidad y valor agregado, el comercio con el resto del mundo se encuentra desbalanceado: mientras las exportaciones se concentran en productos primarios de baja complejidad, aquellos productos que llegan desde el resto del mundo tienen un mayor grado de tecnificación. Estas conclusiones son particularmente válidas para el comercio con China, que constituye una quinta parte del comercio total de la región. Vale mencionar que la importancia de dicho país ha crecido de forma sostenida en el período analizado, siendo que en 1996 representaba menos de 2 % del comercio ²

3.3. Metodología

Del análisis exploratorio se desprende el potencial de la información desagregada a nivel producto para caracterizar la inserción en el mercado mundial de un país o región. Sin embargo, dada la alta cardinalidad de la información y la múltiples dimensiones de estudio, resulta difícil su estudio de forma general, sin hacer éje en un determinado país, o sin agregar la información según un nomenclador. En particular, el uso de agrupaciones jerárquicas de los productos resulta esencial para el análisis de resultados. La elaboración de los mismos constituye una extensa tarea por parte de expertos en las temáticas sectoriales de los diferentes tipos de productos, y el nomenclador resultante está fuertemente determinado por los objetivos con que será utilizado. Es por ello que en la presente sección se propone una elaboración alternativa de niveles jerárquicos de agrupamiento de productos.

agregar grafo
bipartito

² ver <https://treemaps.shinyapps.io/treemaps/>

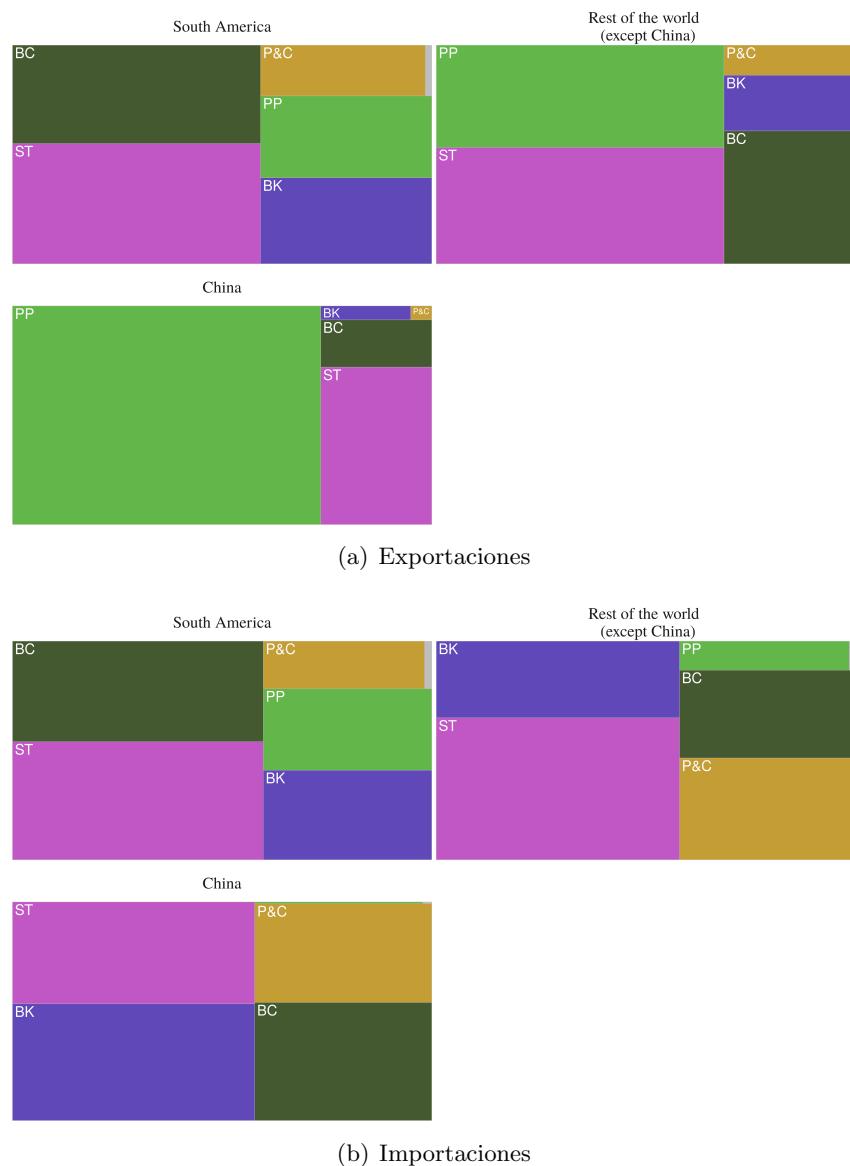
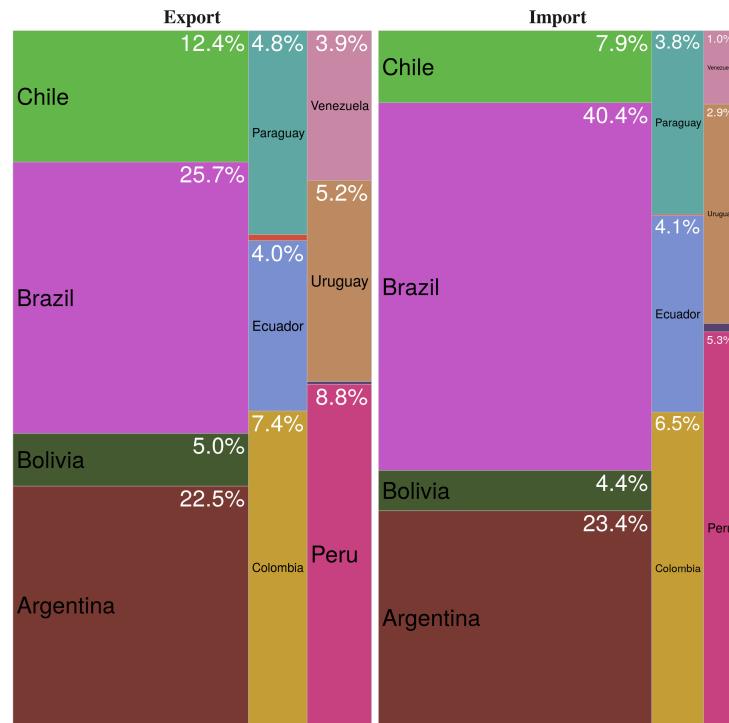
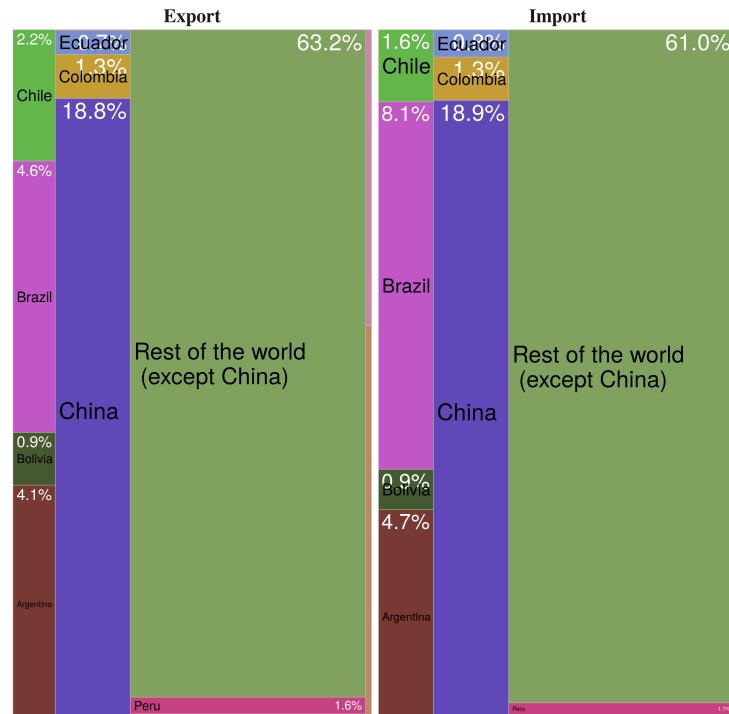


Fig. 3.4: Treemaps de Cadenas y Usos.2016. Total Latinoamérica



(a) Excluyendo Resto del Mundo



(b) Incluyendo Resto del mundo

Fig. 3.5: Treemaps según socio comercial. 2016. Total Latinoamérica

3.3.1. Latent Dirichlet Allocation Models

El objetivo de la presente sección es elaborar un agrupamiento automático de los productos, basado en la información disponible. Este problema puede ser concebido desde dos puntos de vista: por un lado, se puede pensar como un problema de *clustering* donde lo que se busca es crear grupos de productos con similares características. Por otro lado, es un problema de reducción de dimensionalidad, donde lo que se busca es encontrar un espacio de menor dimensión al que existen originalmente los datos. Esto sería posible dado que existe la posibilidad de explotar las similitudes y diferencias entre los distintos productos.

Sin embargo, las técnicas de clustering tradicionales encuentran problemas en espacios de alta dimensionalidad (Aggarwal et al., 2001). A su vez, siguiendo el trabajo de Molinari & de Angelis (2016), se entiende que los grupos no deberían ser excluyentes, dado que un mismo producto puede ser utilizado en diferentes formas, como producto intermedio o final. En este sentido, el problema se puede especificar como de *clustering difuso*.

En concreto, la dimensionalidad del problema se puede pensar como el siguiente espacio:

$$\mathcal{R}^{N*P*Y*2}$$

Es decir la interacción de N países, P productos, Y años, tanto para las exportaciones como las importaciones. La propuesta es utilizar la técnica propuesta por Blei et al. (2003) conocida como *Latent Dirichlet Allocation Models* o *Topic Modeling*. En su versión original, esta técnica se propone como una forma de encontrar los tópicos presentes en un corpus, y la distribución de dichos tópicos sobre cada texto. Este problema es análogo al que se busca en el presente trabajo: Allí se busca una dimensión latente de k tópicos, embebidos en un diccionario de alta dimensionalidad (las palabras presentes en el corpus), que se distribuyen a lo largo de los textos que componen dicho corpus. En el presente problema, se busca una dimensión latente de k *componentes*, embebidos en un nomenclador de alta dimensionalidad, que se distribuyen a lo largo de los países. Esta técnica a su vez puede ser pensada como un problema de clustering difuso, en tanto cada palabra (en el contexto original) puede pertenecer a más de un tópico.

A continuación se realizará una descripción del modelo propuesto por Blei et al. (2003), adaptado al presente dominio.

Definiciones

- Un **producto** es la *unidad básica discreta de los datos*, se define como un ítem de un nomenclador (SITC). Se representa como un vector unitario. Definimos el superíndice i del vector como el i -ésimo producto del nomenclador y el i -ésimo elemento del vector. El V -ésimo producto del nomenclador es el vector w , tal que $w^v=1$ y $w^u=0$, $u \neq v$
- Un **país&año** es una secuencia de N productos, definido como $W = (w_1, w_2, \dots, w_N)$
- Nuestro **corpus** es la colección de M países, definido como $D = (d_1, d_2, \dots, d_M)$
- Un **componente** es una dimensión latente sobre el corpus, y suponemos una cantidad fija k de los mismos.
- Nuestro objetivo es obtener:

1. Una distribución de componentes sobre cada país&año.
2. Una distribución de los productos sobre los componentes.

Proceso generativo e inferencial

Intuitivamente, suponemos el siguiente proceso generativo de datos:

- Para cada país del corpus, imaginamos que las exportaciones surgen de un proceso de dos etapas:
 - Elegimos aleatoriamente una distribución sobre los componentes
 - Para cada dólar exportado:
 - Elegimos aleatoriamente el componente al que pertenece, dada la distribución definida en el paso anterior
 - Elegimos aleatoriamente un producto de la distribución correspondiente a dicho componente
- 1. Para cada Componente $k \in \{1, 2, \dots, K\}$
 - Generar una distribución sobre los componentes $\beta \sim Dir_v(\eta)$ con $\eta \in \mathcal{R}_{>0}$ un parámetro fijo³
- 2. Para cada país $d \in \{1, 2, \dots, D\}$
 - generar un vector de proporciones de componentes $\theta_d \sim Dir_K(\alpha)$ con $\alpha \in \mathcal{R}_{>0}^K$ un parámetro fijo⁴
 - Para cada dólar exportado:
 - I. generar una asignación del componente $z_{dn} \sim Mult(\theta_d)$
 - II. asignar el producto $w_{dn} \sim Mult(\beta_{zn})$

En la figura 3.6 se puede observar el proceso generativo de forma gráfica. Aquí, cada nodo representa una distribución de probabilidad. Las aristas significan que la distribución de salida definen los parámetros de la distribución de entrada. Los recuadros significan replicación: El recuadro interior representa que el proceso se realiza para cada dólar exportado en el país. El recuadro exterior representa que el proceso se realiza para cada país en el corpus.

Un Proceso de Dirichlet es una familia de procesos estocásticos donde las realizaciones son ellas mismas distribuciones de probabilidad. Es decir, el rango de esta distribución (así como en una normal son los reales) son distribuciones de probabilidad. Para interpretarlo geométricamente, la figura 3.7 muestra un ejemplo de la distribución de densidad para 3 productos y 4 componentes. El triángulo representa todas las distribuciones (multinomiales) posibles sobre los tres productos. Cada uno de los vértices del triángulo es una distribución de probabilidad que asigna una probabilidad de 1 a uno de los productos. El

³ En nuestro caso el parámetro se fijo en $1/K$

⁴ En nuestro caso el parámetro se fijo en $1/K$

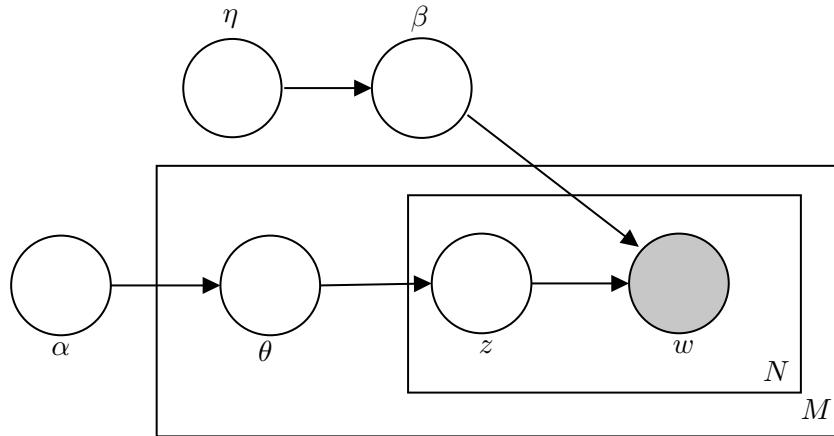


Fig. 3.6: fuente: Blei et al. (2003)

punto medio de cada lado, es una distribución con probabilidad 0.5 a dos componentes. El cuarto componente, el centroide del triángulo, asigna probabilidad de $\frac{1}{3}$ a cada producto. Los cuatro puntos marcado con x son las distribuciones multinomiales de $p(w|z)$ para cada uno de los cuatro componentes. La altura en el eje z es una posible distribución de densidad sobre el simplex, es decir, sobre las distribuciones de densidad multinomiales, dada por LDA. el parámetro que define a la distribución de dirichlet (en nuestro caso, los parámetros η y α) determinan el grado de concentración de las distribuciones resultantes. Para una distribución $Dir(\alpha)$, α define el grado de simetría de las distribuciones multinomiales que el proceso genere. Con valores mucho menores a 1, las distribuciones resultantes estarán muy concentradas sobre algunos elementos, mientras que valores mucho mayores que 1 generarían distribuciones muy uniformes. En términos de nuestro problema, un α muy pequeño generará que en los países haya uno o unos pocos componentes característicos, mientras que un η muy pequeño generará que la distribución que caracteriza a un componente sea muy asimétrica sobre los productos, y por lo tanto que haya unos pocos productos muy importantes, y el resto con probabilidad nula o casi nula.

Cuando observamos los datos, no contamos con los tópicos ni con su distribución, sino con los productos y países. El objetivo es realizar inferencia sobre las variables latentes, mediante el Teorema de Bayes:

$$p(\theta, z|w, \alpha, \beta) = \frac{p(\theta, z, w|\alpha, \beta)}{p(w|\alpha, \beta)}$$

Nuestra función objetivo es:

$$\ell(\alpha, \beta) = \sum_{d=1}^M \log p(w_d|\alpha, \beta)$$

El problema es que esta ecuación es intratable, por la interacción entre θ y β . Por ello, la inferencia se realiza sobre una familia de modelos que se sabe que son una cota inferior de probabilidad, y que son tratables. Estos modelos tienen parámetros variacionales, que se ajustan para obtener el modelo que más se acerca a la cota inferior. La forma de obtener una familia de modelos tratables es considerar algunas modificaciones sobre el modelo gráfico original, removiendo nodos y aristas (Hoffman et al., 2013). En la figura 3.8 se

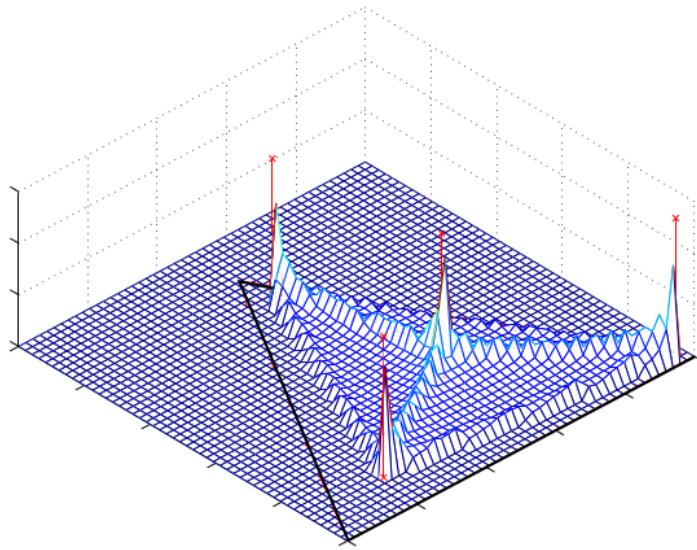


Fig. 3.7: fuente: Blei et al. (2003)

puede observar la solución propuesta por (Blei et al., 2003) que, dado que utilizamos la implementación del modelo que se encuentra en Pedregosa et al. (2011) es también la del presente trabajo.

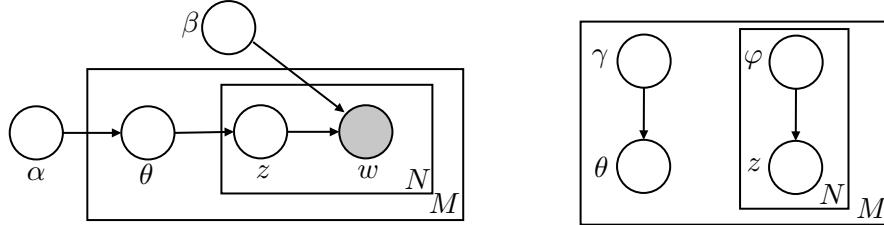


Fig. 3.8: fuente: Blei et al. (2003)

La estimación de los parámetros se realiza a través del proceso de *variational Expectation Maximization* (EM):

- **paso E:** Optimizamos los parámetros variacionales γ, φ
- **paso M:** Para los valores fijos γ, φ , maximizamos la cota inferior respecto a los parámetros del modelo, α, β

Estos dos pasos se alternan hasta converger. Por último, se realiza un suavizado sobre las probabilidades que un componente asigna a un producto, para que sean siempre mayores a 0.

reflexiones finales Por último, es interesante detenerse en la factibilidad del modelo dando el cambio de dominio del problema. La naturaleza tan distinta de los datos usados tradicionalmente en minería de textos y Topic Modelling, respecto de los de comercio internacional, amerita el interrogante respecto de la posibilidad de que el modelo pueda operar en el nuevo dominio. Sin embargo, en términos de la estructura de datos, ambos

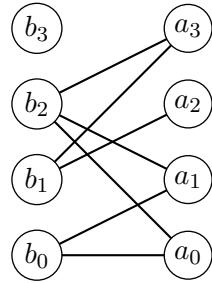


Fig. 3.9: Grafo bipartito, países y productos

problemas tienen más semejanzas de las que aparentan. En primer lugar, La dimensión tradicional del problema es $N \times V$ (N observaciones, en el orden de magnitud de miles, V el vocabulario, también en el orden de magnitud de los miles). En este caso, el problema es aproximadamente de dimensión $N \times P$, donde las N observaciones son los pares año-país, con 180 países y 54 años, y P productos, que en SITC a 4 dígitos son unos 750 aproximadamente. Es decir, estamos en un orden de magnitud similar al de un dataset pequeño en un problema tradicional de Topic Modelling. Respecto de la distribución, en el vocabulario suele operar la ley de Zipf (Powers, 1998), es decir una distribución fuertemente asimétrica de la frecuencia de las palabras. En nuestro problema también existe una distribución asimétrica de los productos exportados por país, aunque no necesariamente del mismo orden de magnitud. Por último, un cambio importante en ambos dominios es la diferencia entre la frecuencia de las palabras en un texto (decenas o centenas, según el tamaño) y los valores exportados de los países (miles de millones). Esta diferencia en principio no debería afectar al modelo, dado que lo que el modelo considera en su optimización son las distribuciones entre los distintos elementos (frecuencias de palabras o valores exportados por producto) y no los valores absolutos.

3.3.2. Grafo Bipartito

Otra manera de analizar el comercio internacional considerando la dimensión *producto* es mediante técnicas de análisis de redes. Es difundido en la literatura el uso de grafos bipartitos para este análisis (Guan et al., 2018; Straka et al., 2017; Ferreira & Others, 2016; Caldarelli et al., 2012). En la figura 3.9 se puede observar un diagrama de un grafo bipartito de países, a , y productos, b . Una arista entre un país a_c y un producto b_i representa en este caso que el país a_c exporta el producto b_i . Dado que tanto los países como los productos se representan como nodos, la estructura de grafo bipartito nos permite mantener la diferencia cualitativa entre ambos tipos de vértices.

Tal como sucedía con el comercio agregado a nivel países, los tamaños relativos de las economías nacionales implican que un mismo monto exportado tenga un significado muy distinto según el país del cual provienen. A su vez, como los volúmenes comerciados de los diferentes productos también varían fuertemente, lo anterior aplica a este otro tipo de nodos. Es por ello que en la literatura se utiliza una normalización de las exportaciones, conocida como *Ventajas Comparativas Reveladas*, o *RCA* por sus siglas en inglés. La misma fue propuesta por Balassa (1965) y tiene la siguiente forma funcional:

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

Dónde $x(c, i)$ es el valor de las exportaciones del *país c* en *producto i*. El numerador indica la proporción que dicho producto representa en las exportaciones totales del país *c*. El denominador indica la proporción que este mismo producto representa en las exportaciones totales de todos los países. Es decir, *RCA* muestra la relación en la importancia de un producto *i* en un país *c* respecto de la importancia promedio de ese producto en la economía mundial. De esta forma, un *RCA* alto implica que el país en cuestión tiene *ventaja relativa* para exportar un producto, mientras que un *RCA* bajo implica que el país tiene una *desventaja relativa* con el producto. De esta forma, se puede construir un grafo no ponderado estableciendo el punto de corte en $RCA > 1$.

Habiendo establecido nodos y aristas del grafo bipartito, el problema reside en que la mayor parte de las técnicas de análisis sobre grafos están definidas para grafos simples, no para grafos bipartitos. Por ello se continúa por realizar la proyección sobre uno de los tipos de nodos para luego analizar los resultados (Zhou et al., 2007). La proyección se realiza de forma ponderada y el peso asignado a una arista entre dos nodos es la cantidad de nodos del otro tipo con los cuales estaban conectados ambos nodos en el grafo original. Esto significa, en la proyección del grafo de países, que aquellos que compartían ventajas comparativas reveladas sobre un conjunto grande productos estarán fuertemente conectados, mientras que los países que poseen canastas exportadoras muy diferentes, estarán débilmente conectados.

Este procedimiento se realizó para los nodos-país. El resultado es un grafo de características similares al elaborado en el capítulo anterior, pero que en esta oportunidad los vínculos entre países no reflejan sus vínculos comerciales, sino su similitud respecto a la estructura productiva exportadora. A diferencia del grafo de relaciones comerciales, aquí las métricas de centralidad no resultan particularmente interesantes porque el centro del grafo está poblado por aquellos países que exportan las mercancías más comunes, mientras que en los márgenes se encuentran aquellos cuya canasta exportadora resulta muy diferente a la media. Lo que resulta particularmente interesante son las comunidades que se generan en el grafo, para ver si las mismas tienen alguna relación con lo dicho en la literatura respecto de la Nueva División Internacional del Trabajo (Fröbel et al., 1978). Para ello se utilizaron técnicas de detección de comunidades. En particular, se utilizó el *clustering de Louvain* (Blondel et al., 2008), donde se optimiza la modularidad de los clusters de forma *greedy*. También se utilizó el método de detección de comunidades *Walktrap* (Pons & Latapy, 2005), que se basa en la noción de que al realizar *random walks* en el grafo, estos caminos tienden a quedarse atrapados en las secciones más densas del grafo, que corresponden a las comunidades. En ambas técnicas en número de comunidades encontradas es definido de forma automática por el algoritmo.

Para el análisis de los productos, nos basamos en el concepto de *proximidad* de Hidalgo & Hausmann (2009); Hidalgo (2009); Hidalgo et al. (2007) definido como:

$$\phi_{ij} = \min(P(RCA_i > 1 / RCA_j > 1), P(RCA_j > 1 / RCA_i > 1))$$

dónde $P(RCA_i / RCA_j)$ es la probabilidad condicional de exportar el producto *i* dado que exporta el producto *j*. Es decir:

$$P(RCA_i > 1 / RCA_j > 1) = \frac{P(RCA_i > 1 \cap RCA_j > 1)}{P(RCA_j > 1)}$$

$$\text{con } P(RCA_j > 1) = \frac{\sum_c I(RCA_c > 1)}{N}$$

siendo N la cantidad de países y \sum_c la sumatoria en los países. Es decir, la proporción de países en los cuales el producto j tiene un $RCA > 1$. Por su parte:

$$P(RCA_i > 1 \cap RCA_j > 1) = \frac{\sum_c I(RCA_i > 1) \cap I(RCA_j > 1)}{N}$$

Esto es, la probabilidad de que un país tenga para ambos productos ventajas comparativas reveladas. Por lo tanto, ϕ_{ij} establece que dos productos son similares si son exportados por los mismos países. Dado que existen productos que son ubicuos, es decir que son exportados por la mayoría de los países, y son muchos quienes tienen un $RCA > 1$, y que una métrica de distancia debe cumplir con el principio de simetría ($d(x, y) = d(y, x)$), esta función toma el mínimo de ambos valores.

Φ es por lo tanto una matriz de distancias entre los productos. Esto resulta ideal para analizar los datos a través de la técnica de clustering *K-medoids* (Kaufman & Rousseeuw, 1987). Este método construye K clusters de forma iterativa modificando el centro del cluster para optimizar la bondad de ajuste. A diferencia de K-means (MacQueen & Others, 1967) el centro de cada cluster se define en una observación existente en el dataset, y por lo tanto se puede recuperar a las mismas para caracterizar el cluster. En el caso del espacio de productos esto resulta de utilidad para comprender los grupos que se conforman. Finalmente, dado que se presentan los datos en un *heatmap*, se decidió también utilizar la técnica de clustering jerárquico propuesta por Ward Jr (1963) para su ordenamiento.

3.4. Resultados

3.4.1. Latent Dirichlet Allocation Models

Análisis de Componentes

Los resultados obtenidos de los modelos de Topic Modelling para el análisis de textos normalmente se analizan en dos etapas: En primer lugar se etiquetan los tópicos obtenidos sobre la base de las palabras más salientes de cada tópico. Luego, se analiza su distribución en los textos. El etiquetado de los tópicos es una tarea subjetiva donde lo que se busca es un concepto generalizador de aquellas palabras que componen al tópico. Es posible, y frecuente, que esta tarea no se pueda realizar por falta de un concepto generalizador dentro del tópico. El hiperparámetro k , es decir la cantidad de tópicos, juega un rol fundamental en este punto, dado que con una cantidad baja de tópicos estos tenderán a reflejar conceptos amplios, mientras que si k es mayor que la cardinalidad del espacio latente que se busca, esto puede generar tópicos repetidos o sobre-específicos. Lo anterior no escapa al presente dominio sino que se refuerza, dado que la búsqueda, subjetiva, de un concepto abarcador entre productos puede resultar más compleja que la de un concepto generalizador de un grupo de palabras. Un problema que no existe en el presente dominio es el de la polisemia, dado que todos los significantes, índices del nomenclador, hacen referencia a un único significado no ambiguo. Sin embargo, nuevos problemas aparecen respecto a este punto, como qué nomenclador de base utilizar y en qué nivel de desagregación. Por

Código	Descripción
01	Primary products
02	Resource-based manufactures: agro-based
03	Resource-based manufactures: other
04	Low technology manufactures: textile, garment and footwear
05	Low technology manufactures: other products
06	Medium technology manufactures: automotive
07	Medium technology manufactures: process
08	Medium technology manufactures: engineering
09	High technology manufactures: electronic and electrical
10	High technology manufactures: other
99	Unclassified products

Tab. 3.1: Agrupamiento por complejidad tecnológica de Lall (2000)

motivos de comparabilidad con los resultados de Molinari & de Angelis (2016) e Hidalgo & Hausmann (2009) se decidió utilizar el nomenclador SITC a 4 dígitos (ONU, 2006). A su vez, se realizaron pruebas para varios valores de k : 2,4,6,8,10,20,30,40,50, 100 y 200, así como para diferentes valores de los hiperparámetros η y α ⁵. Para el etiquetado de los componentes se observó que la práctica usual de observar los primeros diez elementos de la distribución no bastaba para encontrar una etiqueta generalizadora, por lo que se elaboró un tablero dinámico⁶ para estudiar la distribución y su función acumulada, a la vez que se gráfico la distribución en función de un nomenclador de complejidad tecnológica (Lall, 2000). En la tabla 3.1 se puede observar dichas categorías.

A título ilustrativo, en la figura 3.10 se presentan una captura de la interfaz elaborada para la caracterización de los componentes en el caso de $k = 2$. Allí se presenta el código de SITC a 4 dígitos, la descripción del producto, su proporción y la proporción acumulada, ordenados de forma decreciente según su peso. En la figura 3.10(a) se puede observar que la distribución del primer componente asigna un gran peso al petróleo crudo, y que los siguientes productos, además de aquellos no especificados, son también derivados del petróleo, como el gasoil, gas propano (petroleum gases), etc. Una etiqueta plausible para este componente sería *Petróleo y derivados*. Nótese que también se encuentran otros productos como el carbón (coal) y metales como el hierro (iron) oro (gold) y cobre (copper). En la figura 3.10(b) se presenta la distribución del segundo componente. Allí no se observa una distribución tan asimétrica como en el caso del primer componente, sino que el primer producto pesa tan solo un 5 %. Los productos más destacados son los vehículos de pasajeros, microcircuitos electrónicos, partes y accesorios, etc. Podría decirse que este componente representa a los productos de origen industrial en general.

La figura 3.11 muestra la distribución de los componentes siguiendo la clasificación de Lall (2000). Allí se observa cómo el primer componente está constituido esencialmente por productos primarios y manufacturas derivadas de los mismos. El segundo componente presenta una distribución más uniforme, donde se destacan las manufacturas de tecnologías media y alta (ingenieriles y electrónicas)

Vale notar que los productos de origen agropecuario, ganadero, forestal, etc. no quedan

⁵ dado que los resultados son robustos a estos últimos parámetros, se decidió utilizar $\alpha = \eta = \frac{1}{k}$ para las corridas con distintos valores de k

⁶ ver https://treemaps.shinyapps.io/LDA_worldtrade/

	Componente	Code	Description	prop	cumprop
1	1	3330	Crude petroleum and oils obtained from bituminous materials	35%	35%
2	1	9310	Special transactions, commodity not classified according to class	6%	40%
3	1	3345	Lubricating petroleum oils, and preparations, nes	5%	46%
4	1	3414	Petroleum gases, nes, in gaseous state	4%	50%
5	1	3413	Petroleum gases and other gaseous hydrocarbons, nes, liquefied	4%	54%
6	1	3222	Other coal, not agglomerated	2%	56%
7	1	2815	Iron ore and concentrates, not agglomerated	2%	58%
8	1	9710	Gold, non-monetary (excluding gold ores and concentrates)	2%	60%
9	1	3344	Fuel oils, nes	1%	61%
10	1	6821	Copper and copper alloys, refined or not, unwrought	1%	63%

Showing 1 to 10 of 786 entries

Previous 1 2 3 4 5 ... 79 Next

(a) Componente 1

	Componente	Code	Description	prop	cumprop
1	2	7810	Passenger motor vehicles (excluding buses)	5%	5%
2	2	7764	Electronic microcircuits	3%	8%
3	2	9310	Special transactions, commodity not classified according to class	3%	11%
4	2	7849	Other parts and accessories, for vehicles of headings 722, 781-783	2%	14%
5	2	5417	Medicaments (including veterinary medicaments)	2%	16%
6	2	3345	Lubricating petroleum oils, and preparations, nes	2%	18%
7	2	7599	Parts, nes of and accessories for machines of headings 7512 and 752	2%	20%
8	2	7649	Parts, nes of and accessories for apparatus falling in heading 76	1%	21%
9	2	7643	Television, radio-broadcasting; transmitters, etc	1%	22%
10	2	7721	Switches, relays, fuses, etc; switchboards and control panels, nes	1%	23%

Showing 1 to 10 of 786 entries

Previous 1 2 3 4 5 ... 79 Next

(b) Componente 2

Fig. 3.10: Distribución de los componentes según Lall. K=2

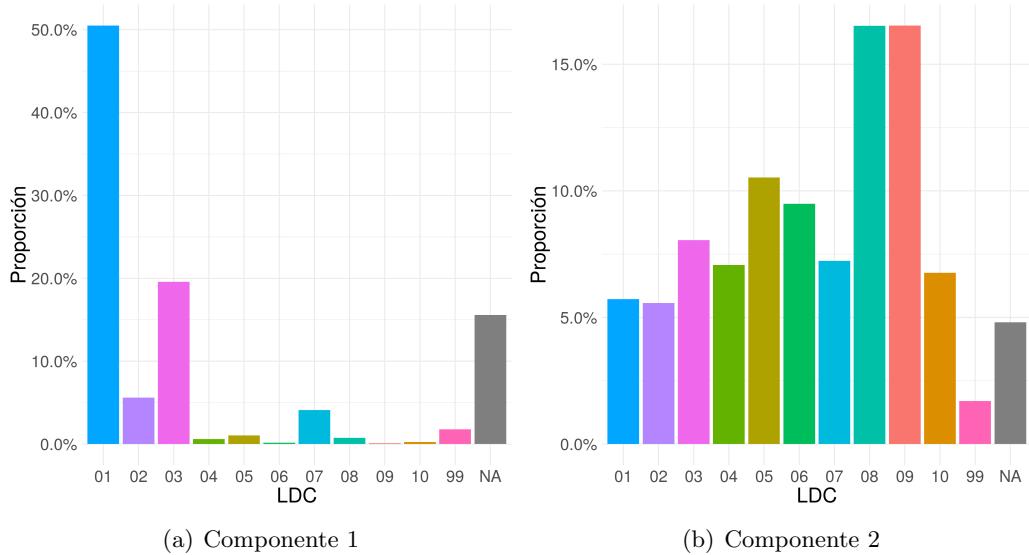


Fig. 3.11: Distribución de los componentes según Lall. K=2

bien definidos en $k = 2$. A su vez, resulta interesante como, bajo la consigna de dividir es espacio de productos en dos grupos, el modelo de LDA encuentra su óptimo en esta diferenciación. Por un lado, reproduce la dicotomía clásica de países industriales versus países productores de materias primas (Coffey, 1996; Hutchinson & Others, 2004). Por otro lado, dicha imagen tradicional pone a los productos agropecuarios como unos de los polos, y no exclusivamente al petróleo y sus derivados. Esto puede entenderse a partir del rol particular que la producción petrolera juega en la estructuración de las economías nacionales, y en particular de las exportaciones (Ross, 2012; Carrera, 2017; Starosta, 2016a). En ese sentido, las formas tan particulares de las exportaciones de los países petroleros llevan a que el modelo encuentre su óptimo construyendo uno de los dos componentes con estos productos.

una de las problemáticas observadas en el análisis de los diferentes ejercicios es definir la granularidad correcta del objeto de estudio. Como se mencionó arriba, el parámetro k define cuán específicos o genéricos son los componentes obtenidos. El problema surge en que en el análisis económico aquellos fenómenos que resulta de interés explorar pueden encontrarse a diferentes niveles de granularidad. Esto es, desde el punto de vista económico puede resultar de interés seguir el comportamiento de un componente que refiera a *derivados de la soja*, pero para poder dar con el mismo, sería necesario un nivel de granularidad que puede implicar que otro concepto interesante como *productos textiles* se multiplique a lo largo de varios componentes.

La figura 3.12 muestra la distribución de los componentes⁷. Allí se observa que la distribución es irregular, y que tiende a haber componentes que se destacan por sobre los demás, como el primer componente para k igual a 2,4 y 6 ; el componente octavo en k igual a 8,10,20, etc. Estos componentes están compuestos o bien por productos petroleros o por una mezcla de estos con otros productos primarios. Esto se puede interpretar como que, o bien son más los países productores primarios, o bien estos productos tienen un mayor peso en las canastas exportadoras de aquellos países que los producen, o una combinación

⁷ sin ponderar por el peso de las exportaciones, es decir de forma equiponderada entre países

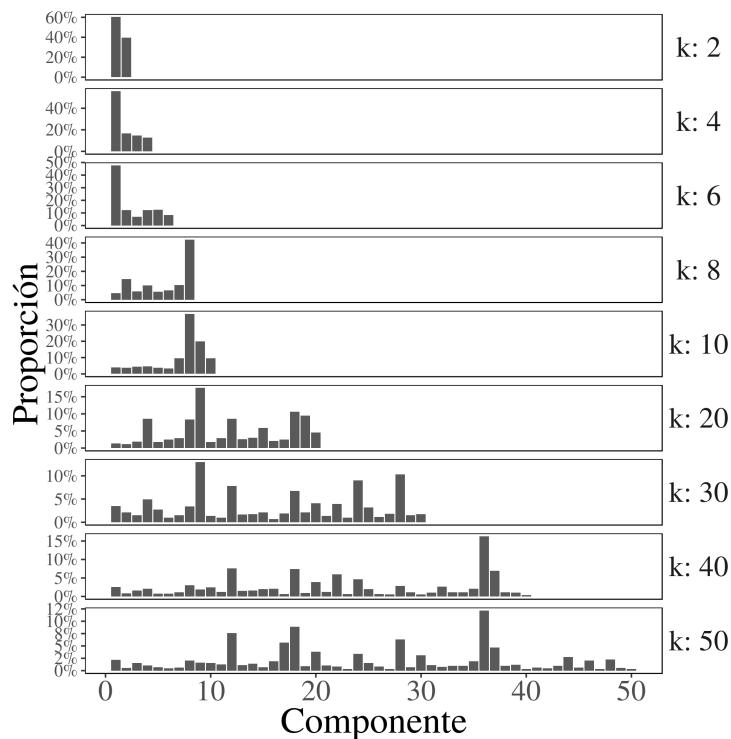


Fig. 3.12: Distribución promedio de los componentes en los países, varios valores de k

de ambos fenómenos.

Finalmente, para elaborar las etiquetas finales se consultó especialistas sectoriales. En la tabla 3.2 se pueden ver las etiquetas para el modelo con $k = 30$, que se utilizará para mostrar los resultados, dado que permite una buena granularidad para los objetivos del presente trabajo. Allí se presenta una descripción general por rama, excepto para el componente 19 donde esto no es posible, además se presenta como *subgrupo* algún conjunto característico de productos que permite una mejor especificación, y para el caso de los productos industriales, el nivel de complejidad, siguiendo la clasificación de Lall (2000). Por último, en la columna *país* se muestra aquel país para el que tiene una mayor participación del componente en el promedio de la serie. Un caso particular es el componente 5, que posee productos de alta complejidad para el comienzo de la serie, pero que luego cayeron en desuso o disminuyeron su participación en el comercio internacional. Por ejemplo, las cintas de grabación, líneas de teléfono o el papel fotográfico. Es natural que el país que caracteriza a dicho componente sea aquel cuya serie es de menor duración, como Checoslovaquia, que por disolverse en el año 92', su denominador es menor que en los demás países.

Análisis de países

Con los componentes definidos, es posible analizar la composición de la canasta exportadora de cada país. Dado que en la definición del problema nuestra unidad es el par país& año, eso significa que podemos comparar la evolución en la distribución de los componentes al interior de cada país.

Comp	Grupo	Subgrupo	Complejidad	País
1	Minerales	Carbón, hierro	-	Australia
2	Industria	Textiles, ingenieril, otros	Baja y media	San Marino
3	Industria	Automóviles	Media	Bélgica
4	Industria	Textil, juguetes, etc	Baja	Macao
		Electrónicos no digitales.		
5	Industria	Cintas de grabación. Líneas de teléfono. Papel fotográfico	Alta (hasta 70')	Checoslovaquia
6	Industria	Autos, barcos, maquinaria	Media y alta	Japón
7	Petróleo	Estado Gaseoso	-	Turkmenistán
8	Minerales	Cobre	-	Chile
9	Agropecuario	Café, bananas, otros	-	Réunion
10	Industria	Autos y electrónicos	Media y alta	México
11	Industria	Autos, partes y componentes	Media y alta	Alemania
12	Petróleo	Crudo	-	Sudán del Sur
13	-	Oro, relojes, joyas	-	Suiza
14	Industria	Químicos	Alta	Curazao
15	Minerales	Diamantes	-	Botswana
16	Industria + Agro.	Aviones, autopartes, soja y maíz	Media y alta	USA
17	Industria + Agro.	Autopartes, madera y derivados	Media	Finlandia
18	Industria + Agro.	Productos primarios y textiles	Baja	Isla de Navidad
19	-	Transacciones especiales, no clasificadas	-	Isla San Martín
20	Combustibles	Fuel oil, gasolina, etc.	-	Rep. Dem. Pop. del Yemen
21	Industria	Medicamentos e instrumentos médicos.	Alta	Irlanda
22	Petróleo + Agro	Hidrocarburos, aceite de palma, cacao, etc.	-	Ghana
23	Industria	Procesadores, microcircuitos, juguetes y calzado.	Alta y baja	China
24	Agropecuario	Carnes, pescados, lácteos	-	Islandia
25	Minerales + Agro.	Soja y Hierro	-	Paraguay
26	Industria + Agro.	Aviones, perfumería, vino, queso	Alta	Francia
27	Industria	Microcircuitos	Alta	Filipinas
28	Industria + Agro.	Arroz, algodón, textiles, goma, etc.	Baja	Pakistán
29	Industria + Agro.	maquinaria, flores, quesos.	Alta	Países Bajos
30	Industria	Autos, motores, etc.	Media y alta	Reino Unido

Tab. 3.2: Descripción Componentes. K = 30

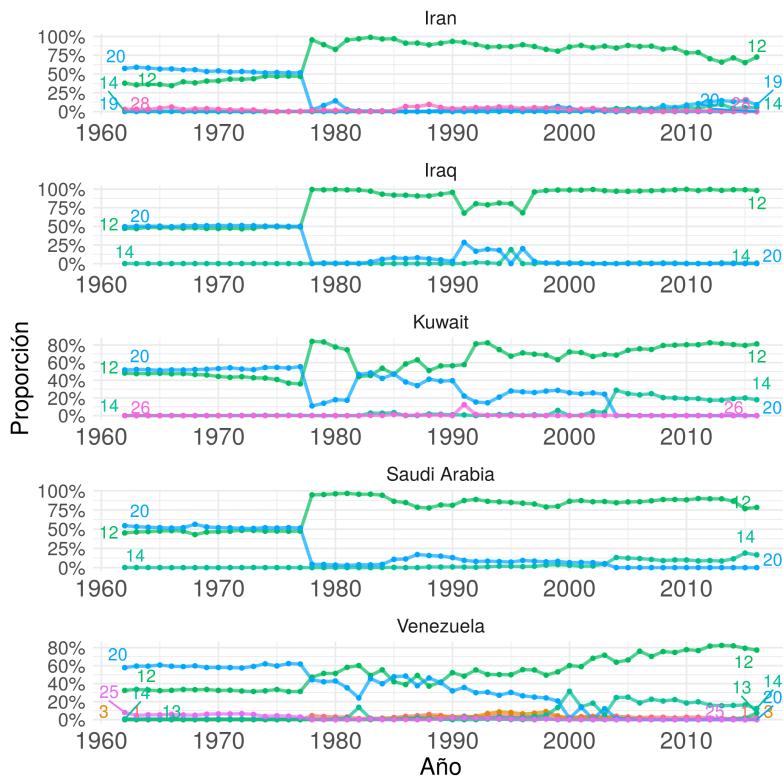


Fig. 3.13: Distribución componentes: Irán, Irak, Kuwait, Arabia Saudita y Venezuela.

Países petroleros En la figura 3.13 se observa la distribución de los componentes para los cinco países fundadores de la Organización de Países Exportadores de Petróleo (OPEP) en 1960. Tal como se refleja en el gráfico, las exportaciones de estos países se concentran en el petróleo y derivados. Es interesante notar cómo, previo a la crisis del petróleo de 1979 (Venn, 2016), las exportaciones se dividían simétricamente entre los componentes 12 (petróleo crudo) y 20 (combustibles), sin embargo, luego de este episodio, con la suba de los precios del barril de crudo, la participación del componente 12 aumenta fuertemente, y se mantiene de esta forma hasta la actualidad. El caso de Venezuela es un tanto particular, dado que las exportaciones de combustible previo a la crisis del 79' tenían un mayor peso que el crudo, y si bien esto se revierte luego de la crisis, durante las décadas de los 80' y 90', el componente 20 sigue teniendo un peso importante, aunque el petróleo crudo va tendencialmente ganando participación hasta la actualidad.

América En la figura 3.14 se presenta la evolución de los componentes de Canadá, México y Estados Unidos. Los tres países tienen un comportamiento diferente. En Canadá predomina el componente 17, una combinación de productos industriales de complejidad media y agropecuarios (madera). Con el comienzo del siglo XXI, este componente pareciera perder terreno frente al componente 12 (petróleo). México comienza la serie sin una especificidad clara, con algún peso mayor del componente 28 (Industrial de baja complejidad y agro) que rápidamente pierde terreno. Durante la década del 80', y parcialmente en los 90' cobra relevancia el componente 12 (petróleo crudo), pero al contrario de lo que sucede en Canadá, este componente pierde peso frente al componente 10 (industrial de

media y alta complejidad) que a fines de los 80' se convierte en el componente con más peso, y continúa creciendo en importancia hasta la actualidad. Este proceso es analizado en la literatura como la *Maquila Mexicana* (Sklair, 2012). En estos análisis se observa como las exportaciones de productos de alta complejidad por parte de México tiene como contrapartida la importación de los componentes desde Estados Unidos, siendo la *maquilas* fábricas de baja complejidad, donde se hace uso de la mano de obra barata en México para ensamblar los componentes importados, y volver a enviarlos a Estados Unidos. Es interesante notar como la metodología propuesta en esta sección lo logra captar este fenómeno en su totalidad, generando una apariencia engañosa respecto de la estructura productiva mexicana.

Estados Unidos comienza la serie con una predominancia de los componentes 5 (Electrónicos, analógicos) y 16 (Industria de alta complejidad y producción agropecuaria). Entre 1977 y 1978 cae la participación del componente 5 de manera abruptamente, para luego seguir su decaimiento hasta la actualidad. El componente 5 marca productos analógicos, considerados de alta complejidad hasta el quiebre producido por los cambios técnicos introducidos por las computadoras personales y las telecomunicaciones. Solo a modo de ejemplo, en junio de 1977 se crea la Apple II, las primeras computadoras personales de venta masiva. Es decir, hasta fines de los 70' la producción de Estados Unidos se dividía entre el agro y productos electrónicos no digitales. Sin embargo, en el componente 16, si bien aparece la industrial pesada (aviones, motores, maquinaria), no aparecen los microprocesadores, transistores y demás productos propios de la *era digital*. Es decir que no existe simplemente un reemplazo tecnológico. Independientemente del punto de quiebre, este proceso puede asociarse a lo que en la literatura es ampliamente conocido como *globalización*, donde a raíz de las posibilidades que brinda el mencionado cambio técnico, ciertos procesos productivos, como la producción de productos electrónicos, se trasladó hacia los países *periféricos* (Henderson, 2002).

En la figura 3.15 se puede observar la evolución de los componentes para Bolivia, Chile y Perú. En el caso de Bolivia, este país comienza la serie con una fuerte predominancia del componente 8 (cobre). Sin embargo, desde mediados de los 60' la importancia de este componente empieza a disminuir, y a crecer la importancia del componente 7 (gas), primero como un pico en la década del 80', y luego de forma sostenida a partir del cambio de siglo. Estos resultados son coincidentes con lo conocido por la bibliografía especializada para dicho país (Auty, 2002; Chávez-Rodríguez et al., 2016). Por su parte, tanto en Chile como en Perú predomina la producción de Cobre y derivados, aunque en el caso peruano existe una variedad de otros componentes con alguna importancia menor, mientras que para Chile las exportaciones son casi en su totalidad sobre el componente (Moran, 2014; Mikesell, 2013).

La figura 3.16 muestra la evolución de los componentes de Argentina, Brasil y Uruguay. En el caso argentino se observa una predominancia del componente 25 (productos primarios como soja y hierro). Al comienzo de la serie también tiene relevancia el componente 24 (ganadería, pesca y lácteos), pero su peso disminuye en el tiempo. Viendo el detalle de los componentes de menor peso, se puede observar que a partir de los 90' crecen los componentes 3 (automóviles) y 14 (químicos), probablemente debido a la creación del Mercado Común del Sur, y las nuevas exportaciones regionales que ello trajo (Bekerman & Rikap, 2010). En el caso Brasileño, se observa un cambio de componente dominante, desde el 9 (café, bananas) hacia el mencionado componente 25. Esto marca la misma tendencia que en el caso argentino a que predomine en este grupo de países la producción de com-

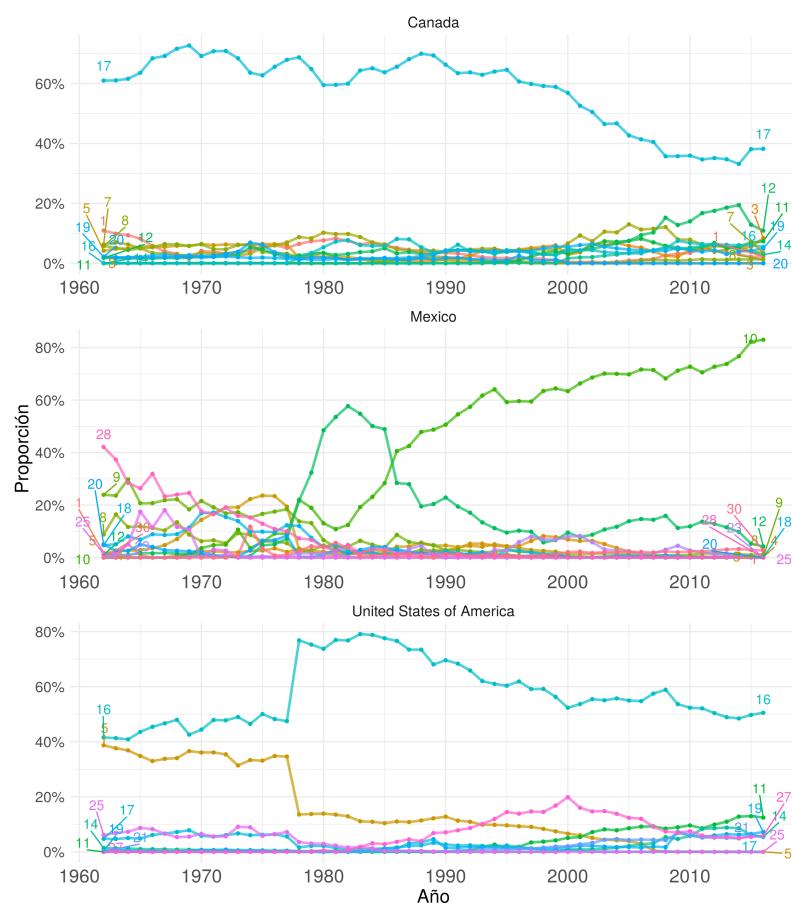


Fig. 3.14: Distribución componentes: Canadá, México y Estados Unidos.

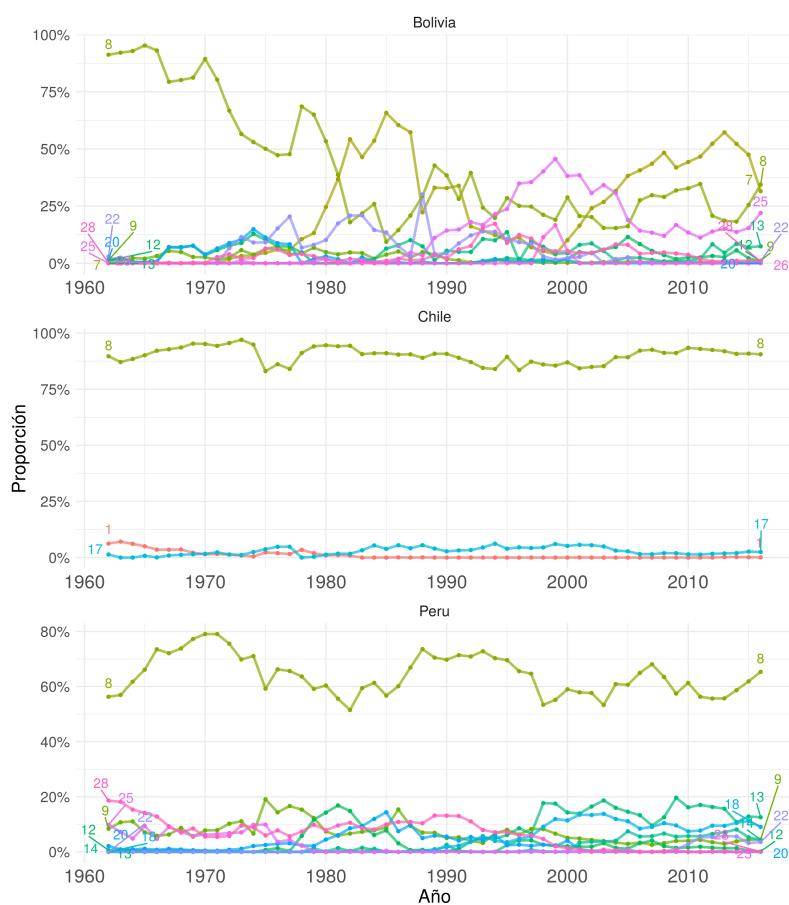


Fig. 3.15: Distribución componentes: Bolivia, Chile y Perú.

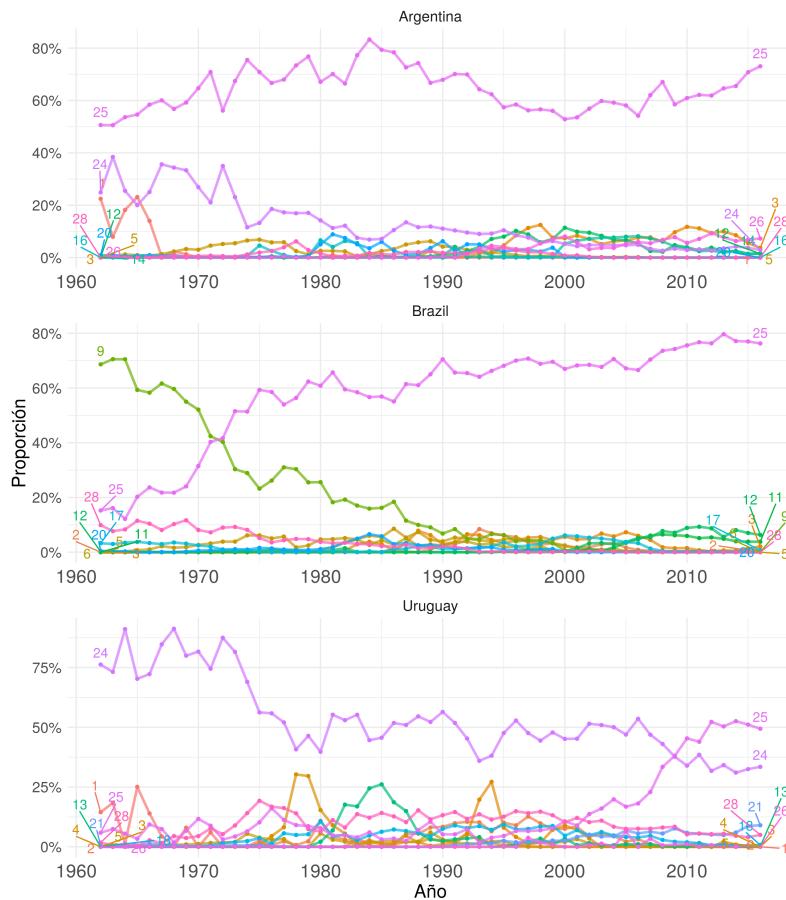


Fig. 3.16: Distribución componentes: Argentina, Brasil y Uruguay.

modities agrícolas. Por su parte, es posible que sea la influencia de Brasil la que determina el peso del hierro en el componente (Costantino, 2013). A su vez, al comienzo de la serie destaca el componente 28 (industria de baja complejidad), que pierde importancia en el tiempo, a la vez que se observa un crecimiento a partir de los 2000' de los componentes 11 (autos, partes y componentes) y 12 (petróleo), aunque siempre de importancia secundaria respecto del componente 25. En el caso de Uruguay, a diferencia de Argentina y Brasil, el componente más importante es el 24 (ganadería, pesca y lácteos). Sin embargo, el mismo pierde importancia, a la vez que desde los 2000' comienzan a crecer las exportaciones en el componente 25, que desde 2009 es el de mayor peso (Redo et al., 2012). En los tres países podemos observar una tendencia a la creciente exportación de commodities, en particular de la soja, y en el caso de Brasil del hierro.

Oceanía La figura 3.17 muestra la distribución de los componentes en Australia y Nueva Zelanda. En el caso australiano se marca una clara predominancia del componente 1 (Carbón y hierro), lo cual es esperable, dado que este país es el mayor exportador mundial de dichos minerales (Balat, 2009). A su vez, la literatura especializada en dicho país menciona un *boom* de los términos de intercambio para estos commodities, que permitiría explicar la tendencia ascendente a partir del año 2000 del componente 1 (McKissack et al., 2008). Nueva Zelanda, por su parte, refleja una predominancia del componente 24

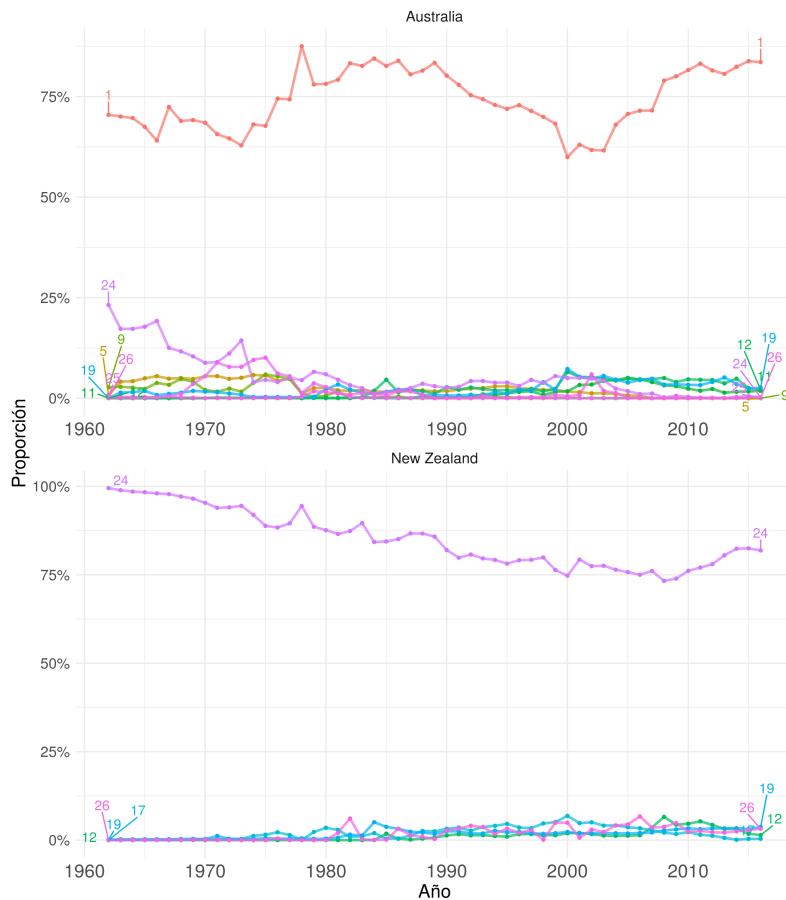


Fig. 3.17: Distribución componentes: Australia y Nueva Zelanda.

(Agropecuario, carnes y lácteos), lo cual se corresponde con la especialización productiva que señala la literatura (MacLeod & Moller, 2006; Ballingall & Lattimore, 2004). Es interesante notar como, a diferencia de lo observado para el caso de Argentina, Brasil y Uruguay, para Australia y Nueva Zelanda no se observa un cambio en la especialización productiva. El caso australiano, dado que se trata de extracción de minerales, es lógico compararlo con el caso de Bolivia, Chile y Perú, en donde estos dos últimos tampoco muestran signos de cambio en su especialización para el mercado mundial. Sin embargo, Nueva Zelanda exporta sobre el componente 24, que es el mismo que definía el perfil exportador de Uruguay y jugaba un rol importante en Argentina para el principio de la serie. No obstante esto último, no se observa para Nueva Zelanda un incremento en la participación del componente 25 sobre sus exportaciones.

Europa En la figura 3.18 se presentan los resultados para Hungría, Polonia y Rumanía, tres países que formaron parte de la zona de influencia de la Unión Soviética (Shama, 2000), y el *Council for Mutual Economic Assistance* (CMEA). En el caso de Hungría, se observa que entre mediados de los 60' hasta mediados de los 70' tiene un peso importante el componente 5(industrial, electrónicos analógicos), que luego cae abruptamente. Este proceso se debe a que los productos del componente 5 se vuelven obsoletos dado el cambio técnico del período. Sin embargo, si se tratara simplemente de un cambio técnico, sería

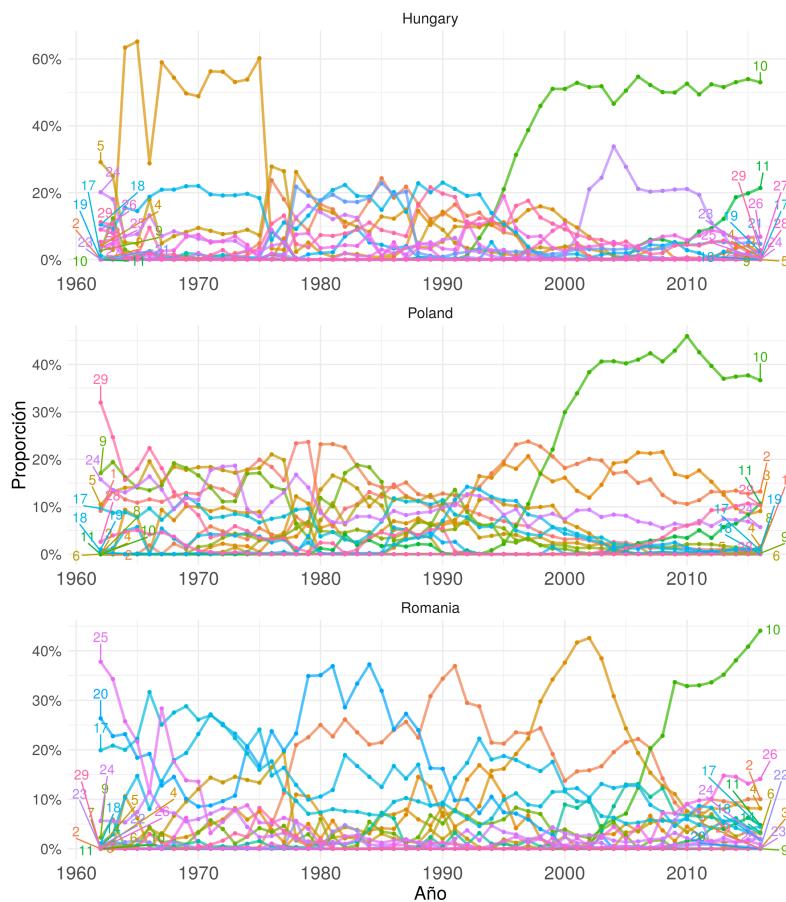


Fig. 3.18: Distribución componentes: Hungría. Polonia y Rumanía.

esperable que el componente 5 fuera reemplazado por otro que reflejase las nuevas tecnologías, lo cual no sucede. Esto puede estar relacionado con el deterioro de los términos de intercambio, tanto con el bloque socialista, como con los países de Europa Occidental, y la mayor apertura comercial con estos últimos (Köves & Oblath, 1983). Más allá de este punto, lo que se observa es una irregularidad que da cuenta de una falta de especialización productiva, en particular hasta mediados de los 90'. Con la caída de la Unión Soviética y la fuerte reestructuración económica dentro de estos países, se observa una creciente especialización sobre el componente 10 (industria, automotriz y electrónica)(Linden, 1998). Para el caso de Rumanía se observa además un pico del componente 4 (industria liviana, textiles y juguetes) que a partir de 2002 disminuye en importancia para dar lugar al componente 10.

La figura 3.19 presenta la evolución de Austria y Alemania desde 1962, y República Checa desde su fundación como estado independiente a principios de los 90'. Tanto en el caso de Austria como Alemania, se observa una fuerte caída a fines de los años 70' del mencionado componente 5. Este fenómeno ya fue comentado en el caso de Estados Unidos y Hungría, y a su vez también se puede observar para Checoslovaquia. Para el caso de Alemania y Austria, con un mayor o menor rezago, se observa que el componente que toma relevancia luego de la caída del componente 5, es el 11 (Industria de media tecnología, automóviles y maquinaria). Por su parte, para el caso de República Checa, junto con

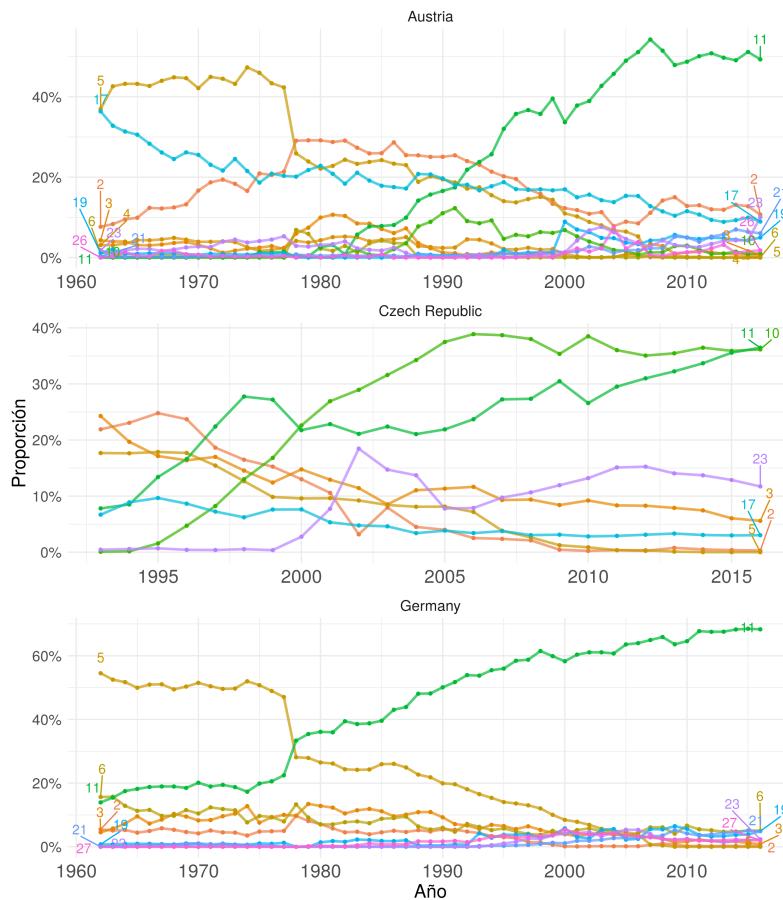


Fig. 3.19: Distribución componentes: Austria, República Checa y Alemania.

el componente 11, también toma igual importancia el 10, que también es industrial con un fuerte componente de automóviles y maquinaria, pero con mayor peso en productos electrónicos, como cables, paneles de control, televisores, etc.

En la figura 3.20 se presentan los resultados para algunos de los países nórdicos: Dinamarca, Finlandia, los Países Bajos y Suecia. Si bien para todos ellos, excepto Finlandia, se puede observar también la caída del componente 5, el mismo no ocupó en ningún momento de la serie el lugar del componente más importante. Al contrario, Todos estos países presentan una especialización bien definida, aunque no sin cambios en el tiempo. En los casos de los países escandinavos, Finlandia y Suecia, predomina el componente 17 (Industria y Agro. Motores, madera y derivados), aunque con una tendencia decreciente en el tiempo. Para el caso finlandés, crece desde mediados de los 90' el componente 23 (procesadores y microcircuitos) y el componente 11(automóviles y maquinaria) a partir de los 2000. El caso sueco es similar, aunque el componente 11 crece tendencialmente desde los 80' y su peso es en general mayor. Tanto el rol tradicional de estos países como productores de materias primas, y en particular madera y derivados, como los recientes procesos de complejización productiva son reconocidos en la literatura (Blomstrom & Kokko, 2003). En el caso de Dinamarca y los Países Bajos el componente característico es el 29 (Industria y Agro, maquinaria,productos químicos, flores, quesos, etc.). En el caso de Dinamarca, también se observa al principio de la serie un peso importante en el componente 24 (carnes

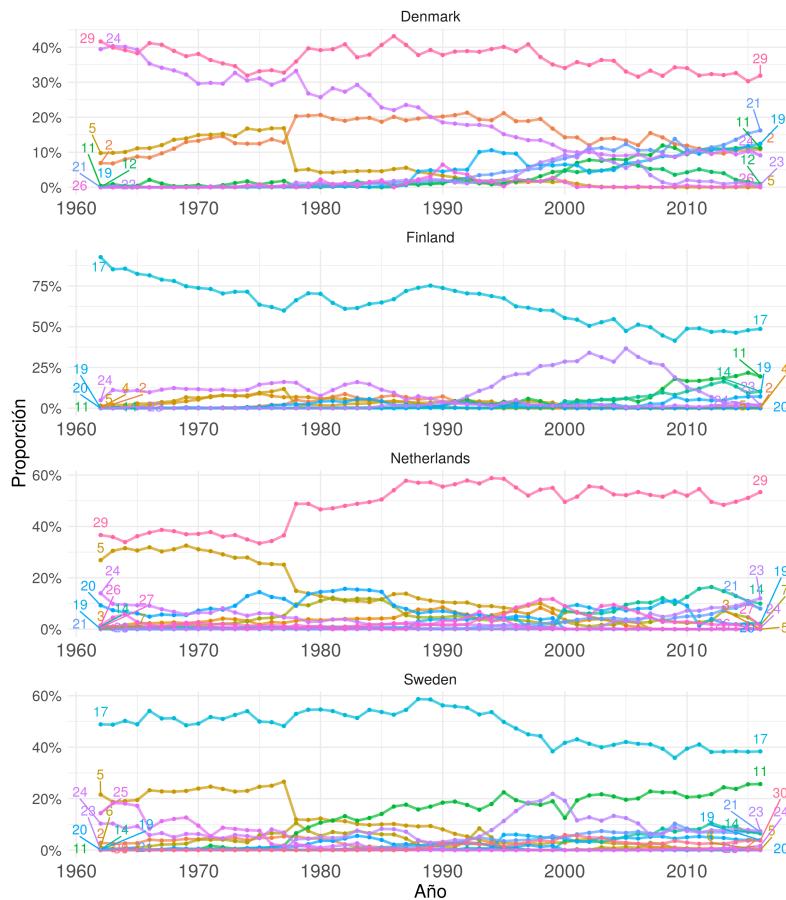


Fig. 3.20: Distribución componentes: Dinamarca, Finlandia, los Países bajos y Suecia.

y lácteos), que cae tendencialmente. Por su parte también se observa un peso importante del componente 2 (industria liviana), particularmente a partir de la mencionada caída del componente 5.

En la figura 3.21 se presenta la evolución de las exportaciones de Francia, Irlanda y el Reino Unido. Nuevamente se observa la caída del componente 5 entre 1977 y 1978, aunque en el caso de Irlanda este componente tenía un peso menor, aunque creciente. Las exportaciones de Irlanda al principio de la serie se caracterizaban por el componente 24 (carnes, lácteos, etc.). Si bien el componente 5 era predominante al comienzo de la serie en Francia y Reino Unido, tal como en los mencionados casos de Alemania y Austria, luego del punto de inflexión cada uno de estos países tiene un componente característico diferente: En el caso francés, es el 26 (aviones, motores, perfumería, vino y queso), todos ellos productos típicos, inclusive la producción de aviones, donde este país es uno de los líderes mundiales con la empresa Air Bus a la cabeza de dicho sector (Frenken, 2000). En el caso del Reino Unido el componente 30 (motores, medicamentos y bebidas alcohólicas) es el que toma la delantera. En ambos casos, estos componentes ya tenían un valor significativo al comienzo de la serie. Distinto es el caso de Irlanda, en donde el componente 31 (medicamentos, químicos y, en menor medida, maquinaria digital) no tenía relevancia alguna al comienzo de la serie. Este cambio desde un componente basado en productos primarios a uno de alta complejidad es reconocido en la literatura bajo el nombre de *celtic tiger*, en relación al

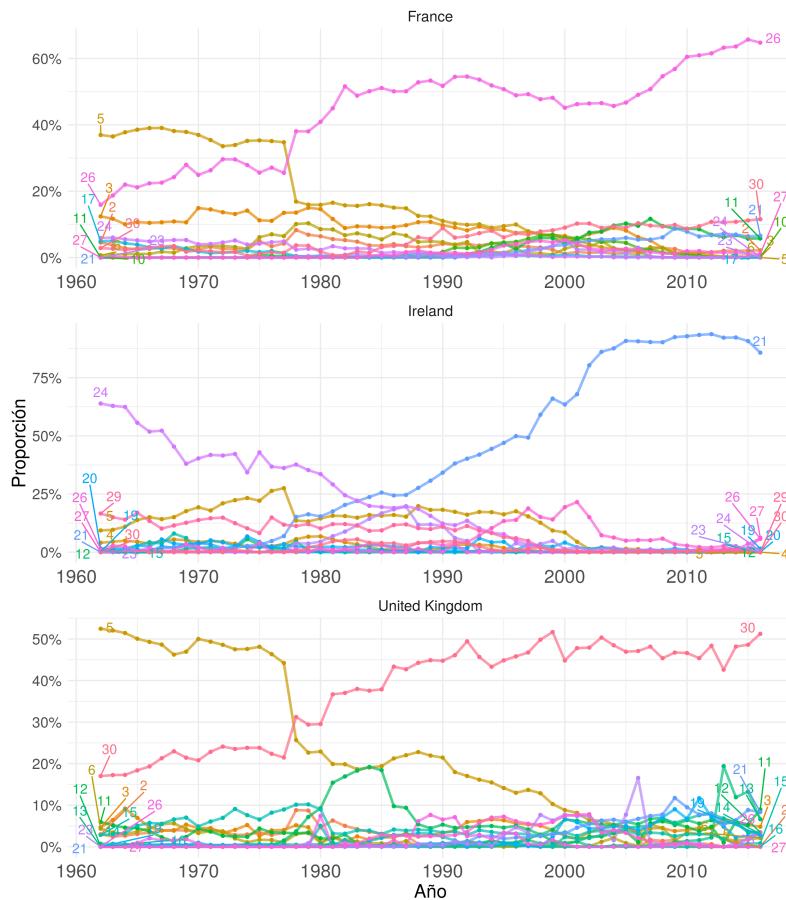


Fig. 3.21: Francia, Reino Unido e Irlanda.

proceso de industrialización reciente en el sudeste asiático (*asian tigers*) (Murphy, 2000).

La figura 3.22 muestra la evolución de Italia, Portugal y España. En el caso italiano, nuevamente se observa una caída del componente 5, que es reemplazado por el 2 (calzado, muebles y maquinaria), en términos de la clasificación de Lall, este cambio implica una simplificación productiva, dado que el peso de productos de alta tecnología en el componente 2 es menor que en el 5. El caso de España y Portugal es similar: desde el comienzo de la serie el componente dominante es el 18 (frutas y vegetales, y producción textil), aunque su peso disminuye a lo largo de la serie, lo cual se contrapone con el incremento de la importancia del componente 3 (automóviles, polietileno, hierro y acero). En particular para el caso español, a partir de la década del 90' este último componente es de mayor peso. Este proceso de industrialización tardía de España y Portugal es reconocido en la literatura (Holman, 2005).

La caracterización del comercio de los países europeos deja como conclusión por un lado la particularidad del comercio dentro del espacio de influencia de la unión soviética, así como la similitud de su estructura productiva en la actualidad de los países que lo componían. Por su parte, es marcada la regularidad temporal con la que el componente 5 cae en importancia en Austria, Alemania, Francia, Reino Unido e Italia. A su vez, parece notarse un proceso de especialización y complementariedad productiva al interior de lo que actualmente constituye la Unión Europea, dado que mientras al principio de la serie

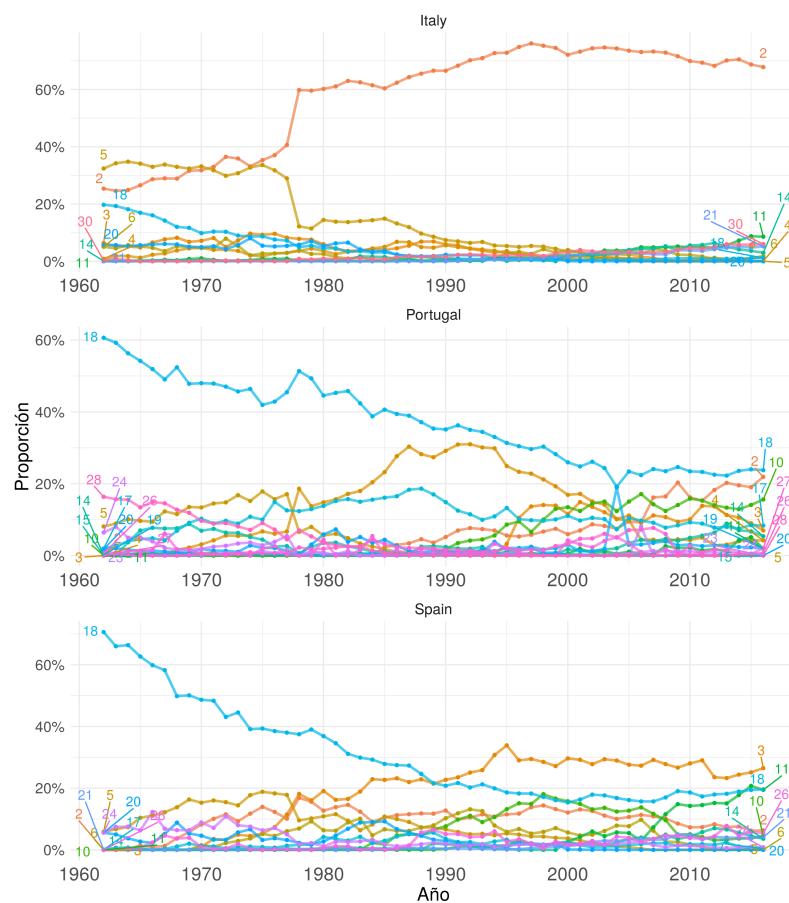


Fig. 3.22: Distribución componentes: España, Italia y Portugal.

muchos de estos países producían en un mismo componente, la caída del mismo da lugar a componentes más o menos bien diferenciados entre los países. También se observa la industrialización tardía de España, Portugal e Irlanda, y se logra apreciar la similitud de este proceso en los primeros dos países, y sus diferencias, tanto en magnitud como en especialización productiva, con Irlanda.

Asia en la figura 3.23 se puede observar la evolución de las exportaciones de China, Japón y Filipinas. En el caso de China, a comienzos de los 60' el componente más relevante es el 28, compuesto por arroz, algodón, té y algunos productos textiles. Este componente tiene una caída tendencial, acompañada de un incremento del componente 4 (vestimenta, juguetes, etc.) que a partir de la década del 80' es el de mayor peso. Este componente tiene su pico a comienzos de los 90', para comenzar a descender, en sintonía con el crecimiento del componente 23 (televisores, computadoras, microcircuitos y transistores), que constituye en la actualidad aproximadamente el 80% de las exportaciones de este país. Este cambio en la especificidad productiva da cuenta de la complejización de la industria China por etapas, primero desde una economía básicamente agropecuaria a una industrialización de baja complejidad, para luego convertirse en uno de los principales exportadores mundiales de productos de alta complejidad Chenery et al. (1986); Costantino (2013). En el caso de Japón, se observa al igual que en muchos países europeos la importancia del componente 5 al comienzo de la serie. Sin embargo, la evolución es distinta en este caso. En primer lugar, si bien hay una caída pronunciada en 1977, no tiene la misma magnitud que en los casos europeo o de Estados Unidos. A su vez, el componente 6 (motores, barcos y maquinaria eléctrica) cobra importancia desde el comienzo de la serie, y supera al 5 ya desde comienzos de la década del 70'. A su vez, desde mediados de los 80' aparece el componente 27 (microcircuitos, unidades periféricas, transistores, etc.). El caso de las Filipinas es paradigmáticamente distinto al de China. Desde comienzos de la serie hasta los 90', el componente principal es el 9 (Café, bananas, etc.). Este componente cae de forma irregular hasta 1990, y a partir de mediados de los 90' el componente 27 crece drásticamente (Chang & Zach, 2018).

La figura 3.24 muestra la evolución de los denominados *tigres asiáticos*: Hong-Kong, Singapur, Corea del sur y Taiwan. En el caso de Hong-Kong, el componente cuatro (calzado, ropa y juguetes) es el de mayor relevancia a lo largo de la serie, hasta el cambio de siglo. A partir de allí, y de forma irregular, disminuye su importancia respecto del componente 27 (microcircuitos, transistores, etc.). Un comportamiento similar tienen Taiwan y Corea del Sur, donde al comienzo de la serie tiene mayor peso el componente 28 (arroz, algodón, etc.), que pierde importancia a mediados de los 60' para ser remplazado por el componente 4, y por el 27 a mediados de los 90'. Vale mencionar que el componente 28 también tiene un peso significativo y decreciente al comienzo de la serie de Hong-Kong, lo que da la pauta de que esta región podría haber atravesado un proceso similar al de los demás países, con anterioridad en el tiempo. Estos procesos son similares a su vez con lo analizado para el caso Chino, donde también se observa el proceso de sofisticación productiva de los componentes 28-4-27. En el caso de Corea del sur, junto con el componente 27, también crece en importancia el 6 (motores, barcos y maquinaria eléctrica), visto en el caso japonés. Singapur tiene un recorrido diferente al de los demás *tigres*. Hasta fines de los 80' el componente más importante es el 20 (fuel oil, combustibles, etc.), para ser reemplazado por el 27, que se encontraba en rápido crecimiento desde principios de la década. Siguiendo a Ng (2013) podría decirse que la industrialización de Singapur se basa

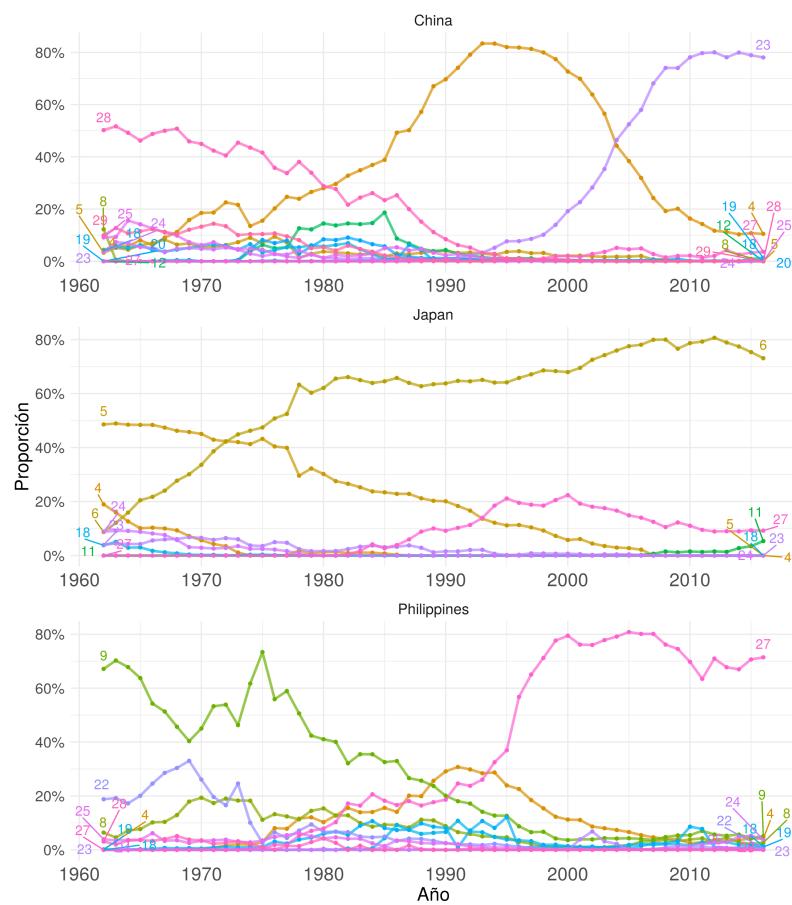


Fig. 3.23: Distribución componentes: China, Japón y Filipinas.

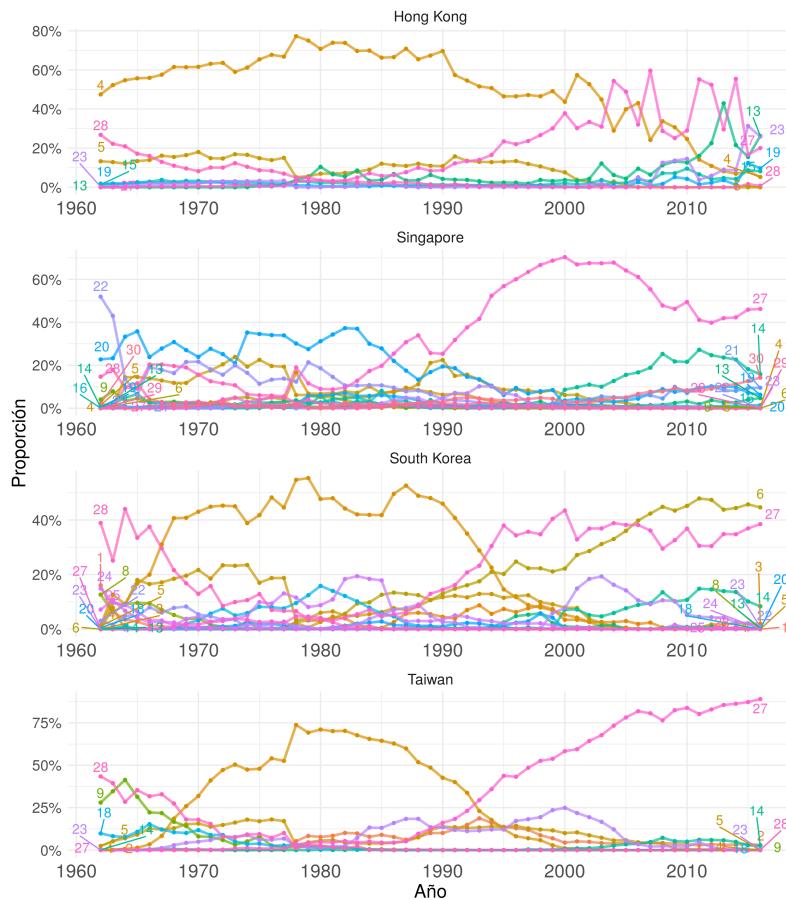


Fig. 3.24: Distribución componentes: Hong Kong, Corea del Sur, Singapur y Taiwán.

en su producción de energía. Sin embargo, a diferencia de los demás países, también se observa en una caída del componente 27 a partir de los 2000, y un crecimiento del 14 (derivados del petróleo)

Finalmente, en el gráfico 3.25 se presenta el caso de Bangladesh, India y Pakistán. Estos tres países tienen un recorrido muy diferente al visto en el resto de Asia. En el caso de Pakistán, el componente 28 representa la absoluta mayoría de las exportaciones, con un decrecimiento tendencial muy leve a partir de los 90' y un incremento muy pequeño de los componentes 4 (ropa juguetes, etc.) y 18 (productos primarios y textiles). En el caso de la India, el componente 28 es predominante, pero su tendencia descendente es marcada, y se comporta en espejo con el crecimiento del componente 15 (diamantes). Este último se explica no por las escasas minas de diamantes que existen en la actualidad en la India, sino por la industria de la talla de diamantes en el oeste de dicho país (Henn, 2013; Sevdermish et al., 1998). Por su parte, Bangladesh muestra un recorrido similar al de China, pasando del componente 28 al 4.

La evolución de las exportaciones en Asia deja un imagen muy diferente a la de Europa. Mientras en aquel continente se pasa de una concentración de los distintos países en un único componente a una diferenciación, en el caso asiático sucede lo contrario: ocurre un proceso de indiferenciación en las exportaciones, concentrándose en la producción de industria liviana de baja complejidad primero, y la producción de micro-componentes

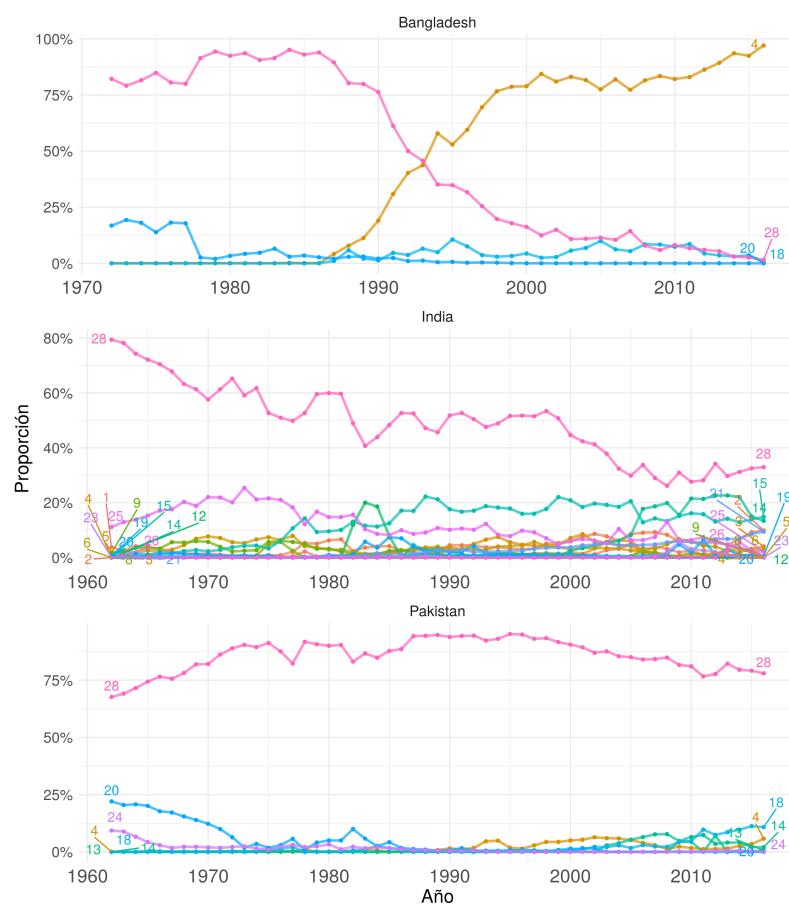


Fig. 3.25: Distribución componentes: Bangladesh, India y Pakistán.

luego. A su vez, lo que se observa es que la producción del hardware asociado a las nuevas tecnologías que se producen a partir del cambio técnico de la *era digital*, y que generan la obsolescencia de los productos comprendidos en el componente 5, se producen en el sudeste asiático.

3.4.2. Grafo Bipartito

Proyección a nivel países

Como se mencionó en la metodología, se realizó una proyección del grafo bipartito para reconstruir un grafo simple ponderado de la similitud de la estructura productiva de los países. En esta red, por lo tanto, la cercanía entre dos países esta determinada por cuán similares son sus canastas exportadoras. Sobre la base de esto, se realizó un análisis de comunidades. El objetivo de este ejercicio es encontrar nueva evidencia empírica respecto de un debate difundido en la literatura económica sobre la Nueva División Internacional del trabajo Starosta (2016b). Por ello, en la figura 3.26 se representan las comunidades resultantes del clustering de Louvain⁸, en mapas de diferentes épocas. La figura 3.26(a) muestra el resultado para la proyección en 1966. en este período el mundo se encuentra fundamentalmente dividido en dos clusters:

- La primera comunidad esta constituida por Estados Unidos, Europa occidental y algunso países del sudeste asiático, como Japón, China y Corea del Sur.
- La segunda comunidad esta compuesta por prácticamente la totalidad del resto del planeta, con excepción de unos pocos países de Centroamérica.

Esta división se podría caracterizar como la diferenciación entre países productores de mercancías industriales de diverso tipo, respecto de países productores de otras mercancías, de origen agropecuario o minero. Esta imagen de un mundo polarizado es coincidente con la bibliografía respecto de la División Internacional del Trabajo clásica, en la cual el mundo se caracterizaba como dividido entre *centro* y *periferia*, donde los primeros son productores de mercancías industriales, mientras que los segundo son productores de materias primas (Coffey, 1996; Hutchinson & Others, 2004).

En la figura 3.26(b) esta imagen empieza a cambiar. Si bien existe aún un dos clusters que coincidentes con los caracterizados para el decenio anterior, surge una nueva comunidad, de países del continente asiático como China e India y Pakistán, entre otros, junto con países africanos y de medio oriente con un cierto grado de desarrollo como Turquía Marruecos y Egipto, y algunos países americanos como México y Colombia. Este cluster, disperso geográficamente, puede caracterizarse como la comunidad de países cuya producción es industrial, de baja complejidad. Esta comunidad es coincidente en tiempo y espacio con lo propuesto por la teoría de la Nueva División Internacional del trabajo, según la cual el fenómeno que ocurre durante este decenio es la fragmentación del proceso productivo y la transferencia de las etapas simples de la producción industrial a estos países, en busca de la mano de obra disponible (Fröbel et al., 1978; Haggard, 1986). A partir de este punto, lo que se observa para los siguientes decenios es que se conserva la estructura de países productores de mercancías industriales de alta complejidad, baja complejidad y de materias primas. En la figura 3.26(c) se puede observar la distribución de las comunidades

⁸ En el anexo, sección E se presentan los resultados utilizando el método de walktrap.

en el año 1986. Los cambios en este punto no son de tipo *estructural*, sino que implican movimientos entre las comunidades anteriormente descriptas. En la década de los 80' el principal cambio se observa en América, en dónde tanto Canadá como Brasil se incorporan en la primera comunidad, mientras que México parecería complejizar su producción industrial (Cardoso, 2009; Feenstra & Hanson, 1997). En la figura 3.26(d) se incorpora la información respecto del comercio de la ex Unión Soviética, que se incorpora al mercado mundial como productora de materias primas en el caso de Rusia y los países de Asia central y como exportadora de productos industriales en el caso de los países de Europa oriental. A su vez, se observa como nuevos países del Sudeste asiático se incorporan a la comunidad de países productores de mercancías de baja complejidad, cómo en el caso de Indonesia, Laos, Vietnam, mientras que países como Corea del Sur, Taiwan y Malasia complejizan su producción, este proceso es reconocido en la literatura como la *segunda ola* de los tigres asiáticos (Rodrik, 1995). Finalmente, en la figura 3.26(e) se puede observar la configuración de las comunidades en el año 2016. Allí se observa un proceso de reprimarización de Latinoamérica, dado que por primera vez desde 1966 en subcontinente en pleno se encuentra en la comunidad de productores de materias primas (Bolinaga & Slipak, 2015; Svampa, 2015). Por su parte, también se observa como continúa la complejización de la producción en el sudeste asiático, con la incorporación de China, Tailandia y Filipinas a la primera comunidad (Suehiro, 2008).

Por comparación y de forma complementaria a los resultados anteriores, en la figura 3.27 se presentan los resultados de la misma metodología, aplicada sobre los datos de comercio agregado entre países del capítulo dos⁹. En primer lugar lo que destaca es la mayor cantidad de comunidades, así como el número cambiante en el tiempo. Esto puede interpretarse como que las relaciones comerciales, bilaterales o multilaterales, son más inestables en el tiempo que el rol que los países juegan en el mercado mundial en tanto productores de cierto tipo específico de productos. En otros términos, con quién se comercia es una forma concreta circunstancial del verdadero contenido puesto en juego en dicho comercio, que es el rol de los países en el mercado mundial (Starosta, 2016b). En segundo lugar, lo que se observa es que las distancias geográficas parecieran tener un rol más relevante en el caso de las comunidades comerciales, respecto de las comunidades según estructura productiva. Esto es esperable, dado los costos involucrados en el transporte de mercancías (Head & Mayer, 2014). En todas las figuras, con excepción de 3.27(b), existe una comunidad de América. A su vez, en todas las figuras hasta 2016 se puede observar el efecto de la Unión Soviética sobre los vínculos comerciales¹⁰ en la actualidad, representada en la figura 3.27(e), lo que se observa es una comunidad comercial de Europa y Asia Central, con excepción del Reino Unido. También se observa una comunidad en África austral, central y oriental, mientras que África occidental constituye una comunidad propia, a la cual pertenecen también parte del sudeste asiático y Oceanía.

Espacio de productos

Como se mencionó en la metodología, para analizar las relaciones entre los productos se toma como base a Hidalgo & Hausmann (2009) y su concepto de *Espacio de productos*. El mismo se define a partir de un concepto de distancia, definido como *proximidad* analizado

⁹ En el anexo, sección E se presentan los resultados utilizando el método de Walktrap

¹⁰ Vale recordar en este punto que, mientras los datos hasta el 2000 provienen de Gleditsch (2002), los datos de 2016 fueron elaborados a partir de la base de datos de *comtrade*.

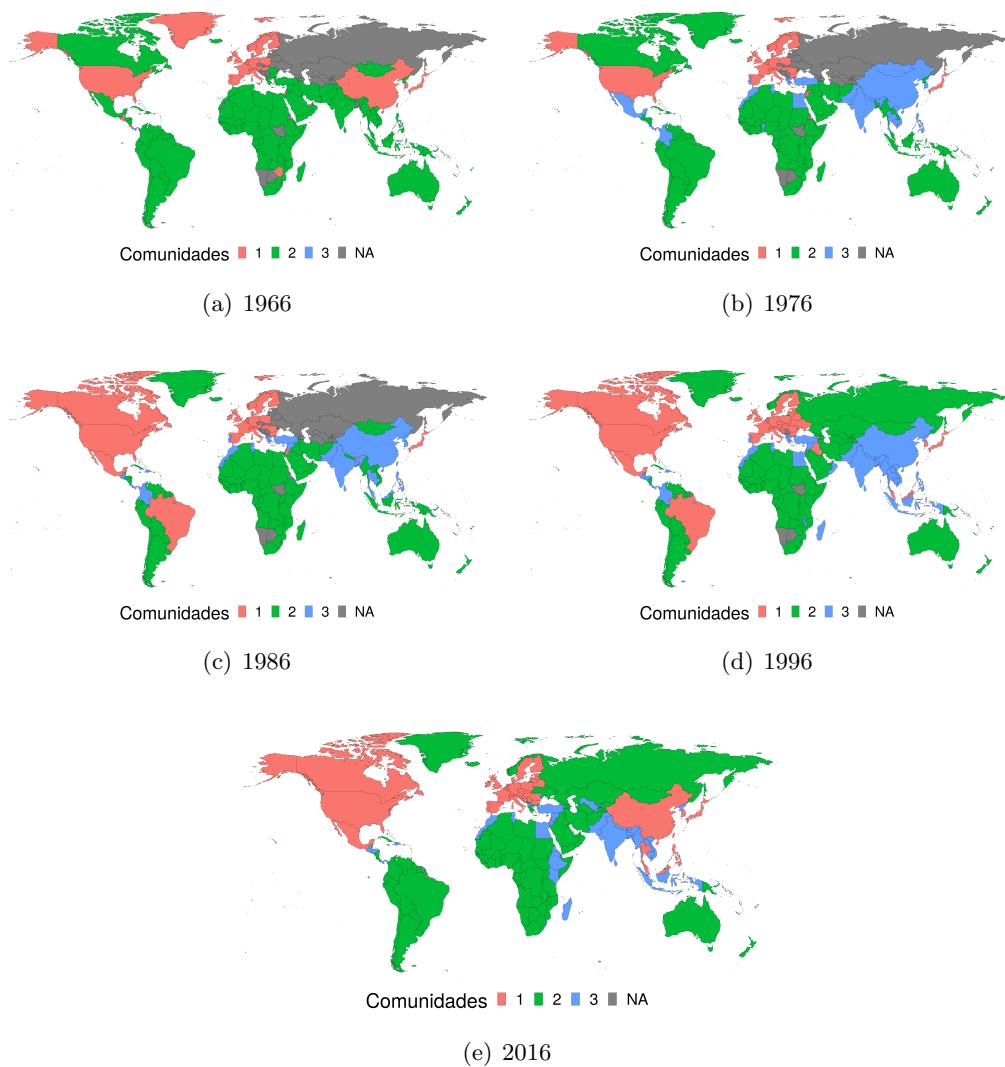


Fig. 3.26: Proyección a países del grafo bipartito. Clustering Louvain. Exportaciones

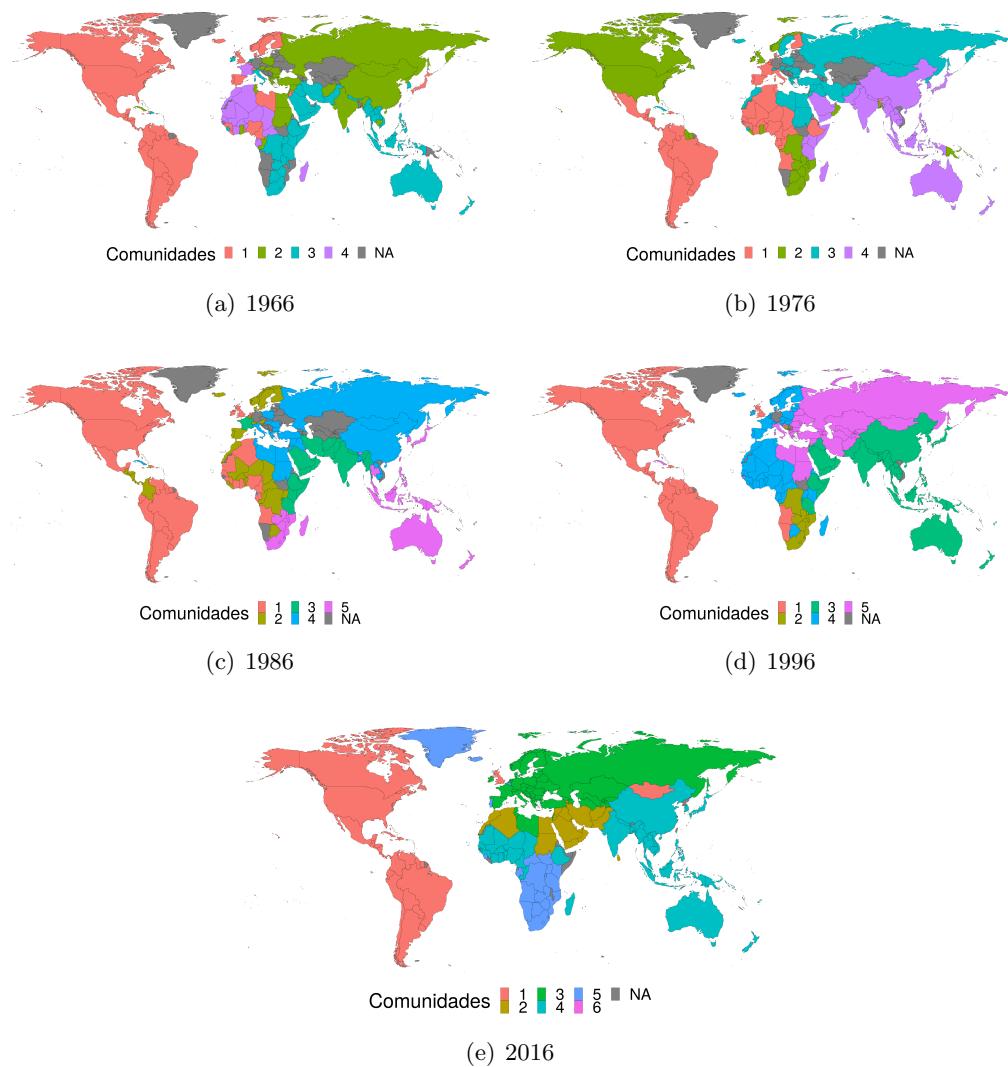


Fig. 3.27: Grafo a nivel países. Clustering Louvain. Exportaciones

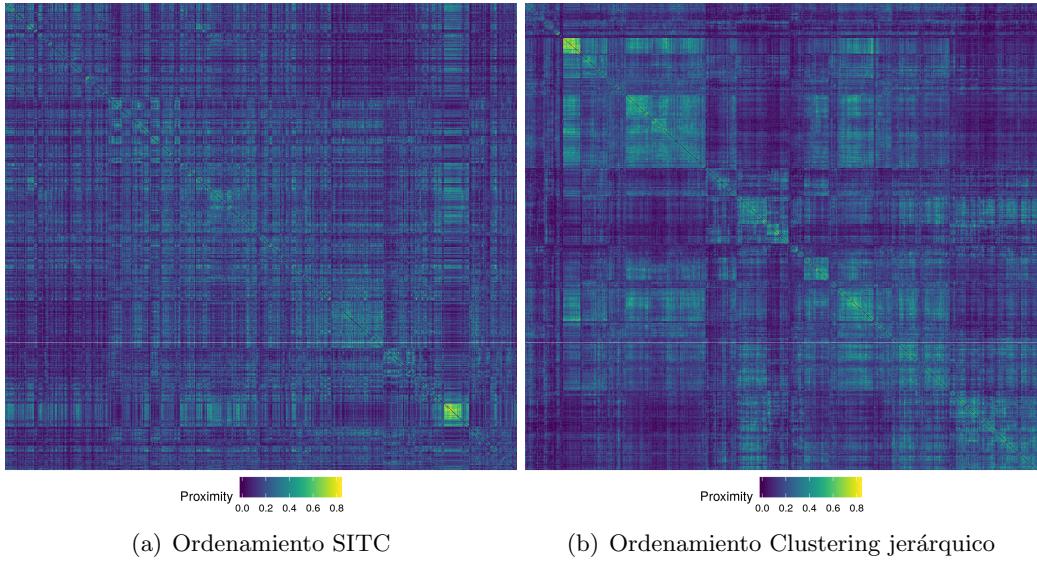


Fig. 3.28: Heatmap proximidad. Espacio de productos. Exportaciones. Promedio 1996-2017

en la sección correspondiente de la metodología. Esto se aplicó sobre el RCA promedio del período 1996-2017, para SITC desagregado a 4 dígitos. Para realizar una primera caracterización del espacio de productos, en 3.28 se muestra el mapa de calor. En 3.28(a) los productos se encuentran ordenados alfabéticamente según su nomenclador SITC. En 3.28(b) la matriz se ordena a partir de realizar un clustering jerárquico. Dado que algunos pares de productos tienen una proximidad igual a cero, al calcular la matriz de distancias estos pares tienen distancia infinita, por lo que la misma se reemplaza por un valor alto fuera de rango.

Más allá de este ajuste para realizar el clustering, en 3.28(b) se mantienen los valores originales para todos los ϕ_{ij} . Las diferencias que se observan entre ambas figuras dan cuenta que el nomenclador SITC no mantiene el ordenamiento propio del espacio de productos dada la similitud en términos de proximidad. Este resultado refuerza la motivación del nomenclador construido a partir de *Latent Dirichlet Allocation Models* propuesto en la sección precedente.

En ambos mapas de calor se puede observar que existe un pequeño conjunto de productos de muy alta proximidad. Para observar dicho fenómeno, en la tabla 3.3 se destacan los diez pares productos de mayor proximidad, y su descripción a cuatro dígitos. Como se observa, todos ellos son productos textiles, algunos de ellos incluso compartiendo descripción a cuatro dígitos. Los procesos productivos para realizar estas mercancías son muy similares y requieren básicamente de los mismos insumos, maquinaria y mano de obra, lo cual explica su similitud.

Teniendo presente lo anterior, se utilizó la inversa de la matriz de similaridad como una matriz de distancias para calcular clusters con mediante el método *Partition around medoids*. A su vez, con estos datos se reconstruyó el grafo de productos utilizando una versión modificada de la matriz de proximidad como matriz de adyacencia del grafo. Específicamente, para aquellos productos con una proximidad menor a 0.5 se anuló la métrica:

Sim	SITC₁	Descripción	SITC₂	Descripción
0.82	8438	underwear,nightwear etc.	8448	underwear, nightwear etc
0.82	8425	skirts and divided skirts, women's or girls', of textile materials, not k	8427	blouses, shirts and shirt-blouses, women's or girls', of textile material
0.82	8416	underwear,nightwear etc.	8428	underwear,nightwear etc.
0.81	8424	dresses, women's or girls', of textile materials, notted or crochete	8425	skirts and divided skirts, women's or girls', of textile materials, not k
0.80	8427	blouses, shirts and shirt-blouses, women's or girls', of textile material	8424	dresses, women's or girls', of textile materials, notted or crochete
0.80	8425	shirts and shirt-blouses,	8426	trousers, bib and brace overalls, breeches and shorts, women's or girls
0.80	8428	women's or girls',	8448	underwear, nightwear etc
0.79	8411	of textile material	8421	overcoats,oth.coats etc.
0.79	8415	shirts	8426	trousers, bib and brace overalls, breeches and shorts, women's or girls
0.79	8426	trousers, bib and brace overalls, breeches and shorts, women's or girls	8427	blouses, shirts and shirt-blouses, women's or girls', of textile material

Tab. 3.3: Top 10 pares de productos más similares

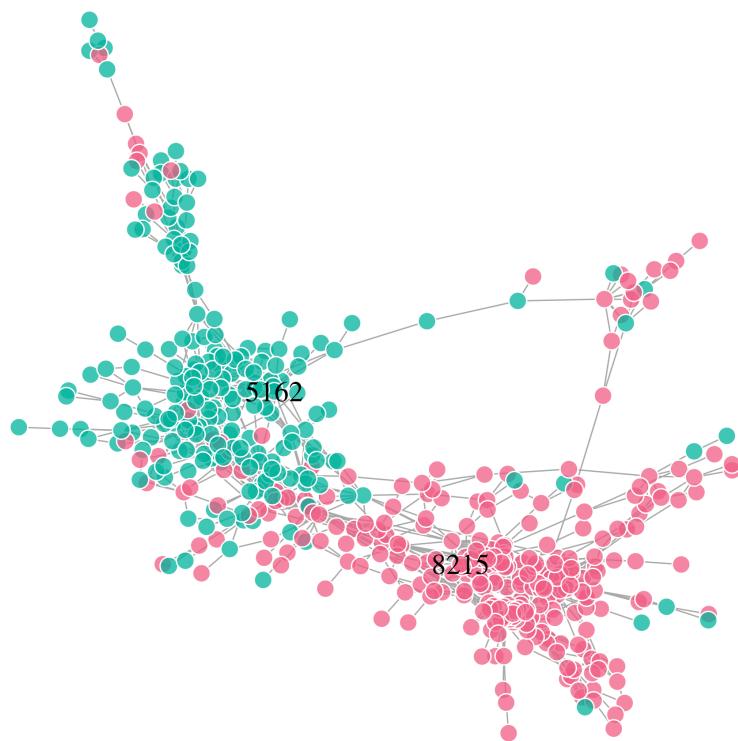


Fig. 3.29: Componente gigante. Grafo de proximidad y clustering por K-medoides. Exportaciones. Promedio 1996-2017. 2 medoides

mediodide	Description
8215	Furniture,nes,of wood
5162	Aldehyde,etc.fnct.cmpnds

Tab. 3.4: Medoides. PAM. k=2

$$\phi_{ij} = \begin{cases} \phi_{ij} & \text{si } \phi_{ij} > 0,5 \\ 0 & \text{si } \phi_{ij} < 0,5 \end{cases}$$

Con esta matriz de proximidad modificada se reconstruyó el grafo. En la figura 3.29 se observa el componente gigante, identificando la pertenencia de cada nodo y su mediodide.

En la figura 3.29 se puede observar una buena separación entre ambos clusters dentro del componente gigante espacio de productos. En la tabla 3.4 se describen los medoides. Dado que el *Aldehyde* es un químico para sintetizar otros compuestos, mientras que el otro mediodide esta caracterizado por muebles de madera, podemos decir que el cluster en verde representa productos de alta complejidad, mientras que el cluster rojo representa a los productos de baja complejidad.

En la figura 3.30 podemos observar qué sucede con 10 medoides. Aquí, si bien se mantiene una separación de los grupos en muchos casos, podemos observar que existe un solapamiento en los puntos más densos del componente gigante. Estos clusters representan a productos de alta complejidad, como componentes de maquinarias (7285), componentes químicos (5416), o productos electrónicos. Con mejor separación aparece la industria pe-

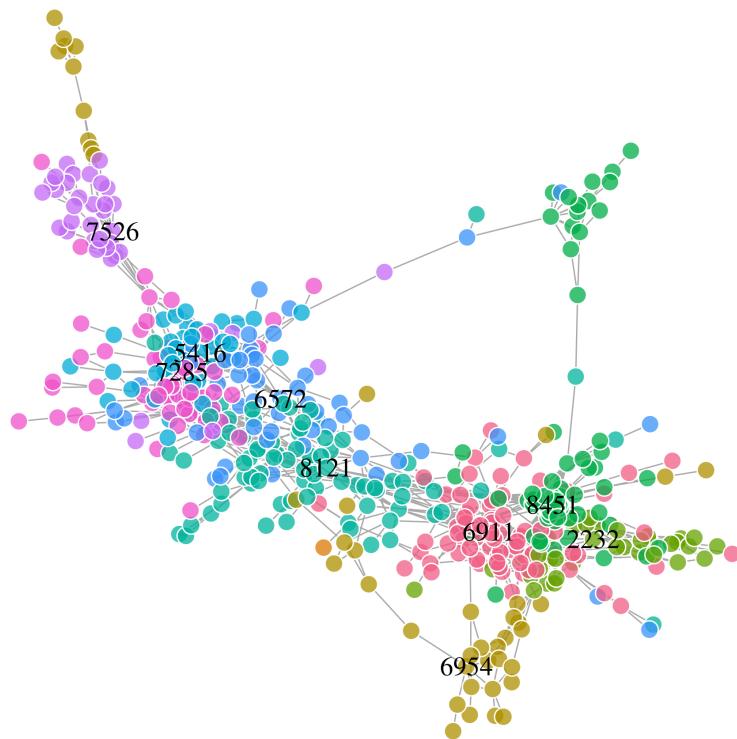


Fig. 3.30: Componente gigante. Grafo de proximidad y clustering por K-medoides. Exportaciones. Promedio 1996-2017. 10 medoides

sada (6911) y en especial equipos para el procesamiento automático de datos (7526). En la figura 3.5 se describen los demás medoides.

En la figura 3.31 se analiza el caso de 50 medoides. La descripción de cada uno de ellos se encuentra en el anexo, en la tabla 4.4. Allí se fuerza lo mencionado anteriormente sobre el solapamiento en la parte más densa del grafo, siendo casi todos los medoides productos del capítulo 7, de maquinaria y equipos industriales. Es decir, es un sector del grafo de alta complejidad. Se mantiene como un cluster más o menos bien diferenciado los procesadores automáticos de datos y microcomponentes (7526, 7722), el cual se podría caracterizar como un sector de mayor complejidad que el centro del grafo. También se configura separadamente el sector de productos textiles (6562) y el de herramientas manuales (6954).

3.4.3. Conclusiones

En el presente capítulo se realizó una segunda aproximación a la cuantificación del comercio internacional, considerando la dimensión *producto*. Este enfoque busca no considerar el lugar que ocupan los países en el mercado mundial, no ya desde sus centralidad en el comercio agregado, sino en función del rol de los mismos en el sistema productivo mundial. Para ello, y dado que parte del presente capítulo fue elaborada en un marco de trabajo multidisciplinario¹¹, se desarrollaron diversas herramientas de análisis. En primer

¹¹ Parte del trabajo del presente capítulo se realizó en el marco del proyecto de Investigación Científica y Tecnológica (PICT) 1085-2016. 'Abordando la restricción externa en América Latina a partir de la

medioide	Description
6911	Metal structures,parts
2831	Copper ores and concentrates
6954	Hand tools,etc. nes
2232	Palm nuts and kernels
8451	Babies'garmnts,clths acc
8121	Boilrs.radiatrs,etc.n.el
5416	Glycosides; glands etc.
6572	Non-wovens, whether or not impregnated, coated, covered or laminated, n.e
7526	Input or output units for automatic data processing machines, whether or
7285	Parts public wrk mach etc

Tab. 3.5: Medioides. PAM. k=10

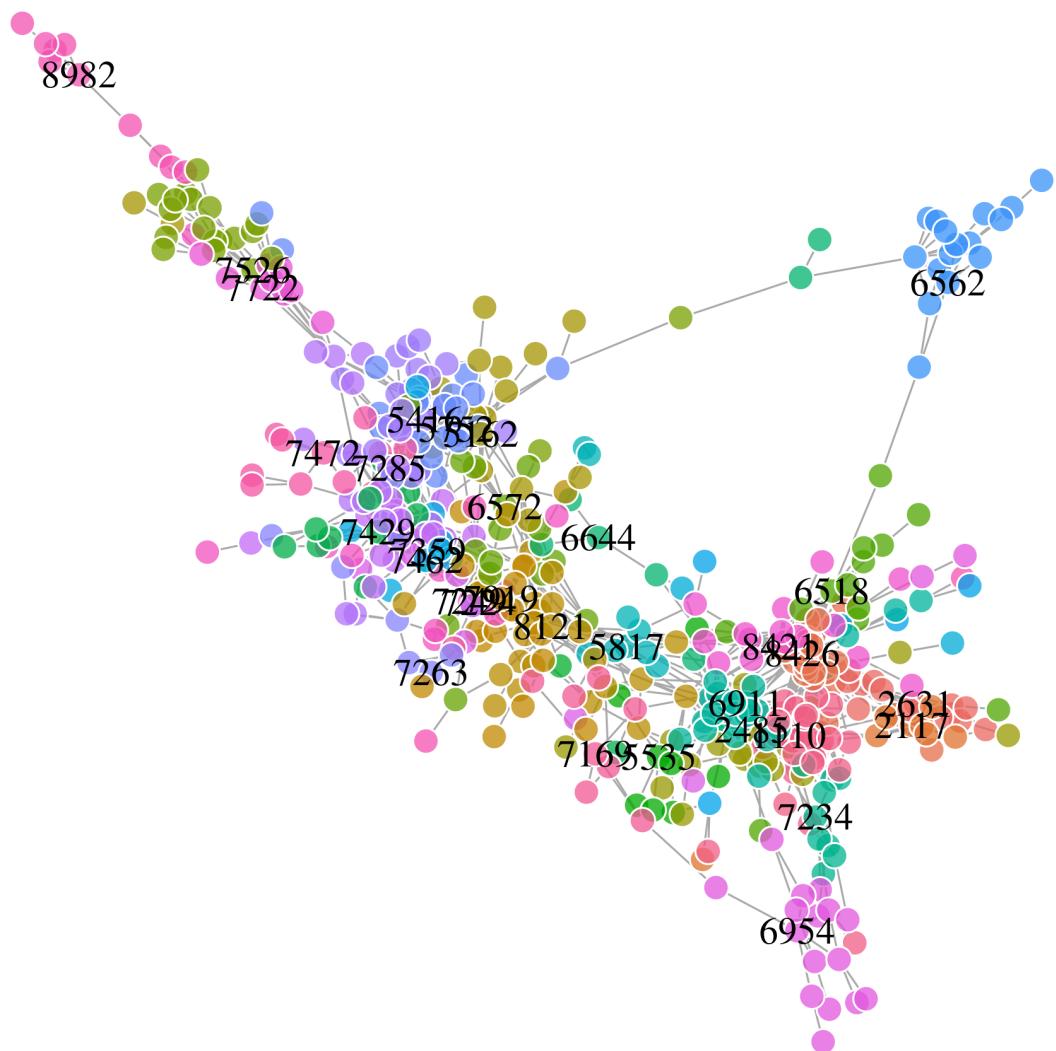


Fig. 3.31: Componente gigante. Grafo de proximidad y clustering por K-medoides. Exportaciones. Promedio 1996-2017. 50 medoides

lugar, para poder dar a conocer los resultados a investigadores temáticos, se elaboraron diversas aplicaciones web en *shiny*, tanto para el análisis exploratorio de datos¹², como para la caracterización de los componentes del modelo de *LDA*, y sus resultados¹³. La propuesta metodológica del capítulo consta de dos partes. En primer lugar, se propone una nueva utilización del modelo propuesto por Blei et al. (2003), proveniente del campo del análisis de textos, para la detección de las dimensiones latentes del comercio mundial, nombradas *componentes*. Allí se exploró los parámetros óptimos del modelo, se caracterizaron los componentes, y con esto, se analizaron los resultados desde el punto de vista de la evolución de la estructura productiva exportadora de los países a lo largo del período 1962-2016. Esta nueva herramienta se presenta como promisoria para la elaboración de una caracterización sintética del comercio de los países, sin necesidad de elaborar nomencladores ad-hoc. Por su parte, tomando el trabajo de Balassa (1965) se elaboró un grafo bipartito entre países y productos, con el que luego se elaboró una proyección al grafo de países y de allí se caracterizaron las comunidades presentes en el mercado mundial a lo largo del período. Este análisis permitió agrupar a los países según el rol que los mismos juegan en la producción mundial de valor. De esta forma, se produjo nueva evidencia empírica respecto de la Nueva División Internacional del trabajo. Finalmente, sobre la base del trabajo de Hidalgo & Hausmann (2009) se caracterizó el espacio de productos, en función de su similitud. Se encontró de esta forma que los nomencladores oficiales no son una representación uniforme respecto a la distancia entre los distintos productos, y que por lo tanto resultan de interés los ejercicios como el propuesto en utilizando *LDA*. A su vez, mediante técnicas de clustering, se caracterizó a los productos procurando construir tipologías basadas en la información empírica disponible. En síntesis, la propuesta del presente capítulo es elaborar metodologías complementarias a las tradicionalmente utilizadas en el análisis estadístico del comercio internacional. Las diferentes propuestas tienen en común buscar tipologías que se basen en los datos generados por el comercio internacional.

integración regional: integración productiva, cooperación Sur-Sur y financiamiento para el desarrollo'.

¹² <https://treemaps.shinyapps.io/treemaps/>

¹³ https://treemaps.shinyapps.io/LDA_worldtrade/

4. CONCLUSIONES

El presente trabajo buscó, mediante diferentes técnicas, aportar nueva evidencia empírica respecto de la organización del comercio internacional de bienes. Dado el objeto de estudio, la presentación del problema y los resultados se organizó en dos grandes bloques temáticos: En primer lugar el análisis de las relaciones bilaterales entre países, y en segundo lugar el rol de los países dados los productos que exportan. En ambos casos, el trabajo se estructuró a partir de un análisis exploratorio de datos, una definición metodológica del problema, y un análisis de los resultados obtenidos. En el capítulo dos, en donde se estudiaron las relaciones bilaterales, se observó el distinto rol que juegan los países en tanto productores y consumidores del mercado mundial. Por su parte, en el análisis de la serie 1948-2000, se pudo observar los fuertes cambios en las características topológicas de la red durante las décadas del 50' y 60', y el cambio en la centralidad del continente asiático. A su vez, en el análisis de la importancia de los países en forma individual, se observó un cambio en el rol de Japón, China, Corea de Sur y México, en dicho orden cronológico. En el capítulo tres, donde se estudiaron las relaciones comerciales a nivel producto, se observó en primer lugar las diferencias en la forma de comerciar de los países latinoamericanos en tanto se trata del comercio regional, con el resto del mundo o con China específicamente. Luego, sobre la base de la propuesta metodológica de utilizar el modelo de *Latent Dirichlet Allocation* en el presente objeto de estudio, se definió treinta dimensiones subyacentes del comercio internacional, que se utilizaron luego para caracterizar la evolución de la estructura productivo-exportadora de un conjunto amplio y diverso de países. Esta metodología nos permitió analizar la especificidad, y los cambios en la misma, del conjunto de países analizados, y dar con un panorama de la división internacional del trabajo. A su vez, se observaron los cambios técnicos ocurridos en la década del 70', y el fenómeno de la industrialización tardía en varios países de Asia, así como en México, Irlanda, España y Portugal. Además, la herramienta resultó lo suficientemente sensible como para poder caracterizar las diferencias en los distintos procesos de industrialización, y la diversificación productiva en Europa luego de los 70'. Respecto de América Latina, se observó la creciente importancia de la producción sojera en Argentina, Brasil y Uruguay, así como el cambio en Bolivia desde la producción de cobre a la producción gasífera, así como también la relevancia que las exportaciones de cobre tienen en Chile y Perú, y el petróleo en Venezuela. El análisis de la proyección de los nodos de países en el grafo bipartito, y su clustering posterior, confirma los resultados obtenidos mediante LDA, al mostrar una división, a partir de los años 70', entre países productores de materias primas, países industriales de baja complejidad y países industriales de alta complejidad. Se observa también allí, los cambios ocurridos en el continente asiático. Por su parte, el análisis de los clusters de la matriz de proximidad en el espacio de productos da cuenta de que los agrupamientos no solo son por ramas de la actividad, sino también por el grado de complejidad de los productos. En este análisis, sin embargo, no se logra la granularidad obtenida mediante LDA. El presente trabajo no se propuso indagar sobre las causas de la división internacional del trabajo, la industrialización tardía de ciertos países, ni el porqué del rol central que algunos países juegan en el grafo del comercio bilateral. El objetivo es realizar una serie de propuestas metodológicas para el análisis empírico del comercio mundial, en sus diferentes facetas, que puedan ser de utilidad para otros trabajos que indaguen respecto a dichas causas. Dado

que se trata de generar nuevas representaciones de la información empírica disponible, de forma tal de echar luz sobre fenómenos largamente estudiados en la literatura, este trabajo hizo un uso intensivo de la visualización de la información como una herramienta propicia para objetivos propuestos. Por su parte, si la representación del capítulo tercero generaba la impresión de que se trata de un fenómeno en el cual elementos aislados, países, se interrelacionan; el capítulo cuarto resuelve dicha impresión dando la pauta de que el sistema económico funciona a nivel de sistema-mundo, y en donde los países son partes de una organización más general, que en la literatura se describe como la División Internacional del Trabajo. El presente trabajo deja como conclusión que la caracterización clásica de un mundo bipolar entre países industrializados, del *centro* y países productores de materias primas, la *periferia*, es insuficiente para dar cuenta de la complejidad que presenta la división geográfica actual de la producción. Por un lado países del *centro* dan cuenta de un componente agrario importante en sus exportaciones, como los países nórdicos en Europa, Canadá, Australia y Nueva Zelanda. Por otro lado, países de la *periferia* son los grandes productores industriales de la actualidad, en particular en el sudeste asiático, e incluso de productos de alta complejidad técnica, como los microprocesadores y transistores, entre otros. Este punto incluso se opone al análisis clásico de la Nueva División Internacional del trabajo, en dónde se planteaba que la relocalización productiva sólo sería para productos de baja complejidad técnica. Es importante destacar, sin embargo, que el presente análisis es también restringido respecto de su objeto de estudio. En particular, no se consideró en el presente trabajo la producción de servicios, lo cual incluye la multiplicidad de trabajos relacionados con la producción de software, contenidos digitales, tareas de diseño, investigación, etc. Por su parte, al considerar las exportaciones de un país, no se tiene en cuenta el origen de los capitales que producen dichos bienes. Estos dos elementos son indispensables para una caracterización de conjunto de la economía mundial, y para poder dar con las causas de los fenómenos que este trabajo simplemente se propuso graficar en sus expresiones empíricas. Las futuras líneas de investigación que este trabajo plantea, por lo tanto, es la utilización del herramiental propuesto aquí como un insumo más en una explicación que busque dar cuenta de las causas de aquello que aquí se busca reflejar.

ANEXO

La nomenclatura de Cadenas y Subcadenas utilizada en el presente trabajo proviene de Molinari & de Angelis (2016). A su vez la clasificación en Usos proviene de Flóres (2008)

A. Tabla de referencia: Descripción cadenas

Cadena	Descripción
Agroind	Agroindustria
Animal	Manufacturas De Origen Animal
Automotriz	Automotriz
Bebidas	Bebidas
Bk	Bienes De Capital
Combyene	Combustibles Y Energía
Cuero	Marroquinería, Calzado Y Otras Manufacturas De Cuero Y Pieles
Foreind	Foresto-Industria
Insdif	Insumos Difundidos
Medic	Medicamentos
Oeqtransp	Otro Equipo De Transporte
Texteind	Textil E Indumentaria
Omanuf	Otras Manufacturas
Oquim	Otros Químicos

Tab. 4.1: Descripción Cadenas

B. Tabla de referencia: Descripción subcadenas

Cadena	Subcadena	Descripción
Agroindustria	Agroquim	Agroquímicos
Agroindustria	Alimanim	Alimentos Para Animales
Agroindustria	Ceryoleag	Cereales Y Oleaginosas
Agroindustria	Flh	Frutas, Legumbres Y Hortalizas
Agroindustria	Oagroind	Otros Agroindustria
Animal	Aviar	Aviar
Animal	Lácteos	Lácteos
Animal	Oanimal	Otros Cadena Animal
Animal	Pescado	Pescado
Animal	Porcina	Porcina
Animal	Vacuna	Vacuna
Automotriz	Autopartes	Autopartes
Automotriz	Autos	Autos
Automotriz	Camyomnib	Camiones Y Ómnibus
Automotriz	Oautom	Otros Automotriz
Bebidas	Alcohol	Alcohólicas

Cadena	Subcadena	Descripción
Bebidas	Noalco	No Alcohólicas
Bebidas	Vino	Vino
Bienes De Capital	Eqelec	Equipos Eléctricos
Bienes De Capital	Eqmedico	Equipamiento Médico
Bienes De Capital	Maqagr	Maquinaria Agrícola
Bienes De Capital	Maqalim	Maquinaria Para Alimentos
Bienes De Capital	Maqcomb	Maquinaria Para Combustibles
Bienes De Capital	Maqfor	Maquinaria Forestal
Bienes De Capital	Maqmet	Maquinaria Para Metales
Bienes De Capital	Maqtextycuero	Maquinaria Para Textiles Y Cuero
Bienes De Capital	Obk	Otros Bienes De Capital
Bienes De Capital	Oeqtrans	Otro Equipo De Transporte
Bienes De Capital	Vehyautop	Vehículos Y Autopartes
Combustibles Y Energía	Biocomb	Biocombustibles
Combustibles Y Energía	Electricidad	Electricidad
Combustibles Y Energía	Gas	Gas
Combustibles Y Energía	Nuclear	Nuclear
Combustibles Y Energía	Ocombyene	Otros Combustibles Y Energía
Combustibles Y Energía	Petroleo	Petróleo
Combustibles Y Energía	Quimicos	Químicos
Marroquinería, Calzado Y Otras Manufacturas De Cuero Y Pieles	Calzado	Calzado
Marroquinería, Calzado Y Otras Manufacturas De Cuero Y Pieles	Cueros	Cueros
Marroquinería, Calzado Y Otras Manufacturas De Cuero Y Pieles	Ocuero	Otros Cueros
Marroquinería, Calzado Y Otras Manufacturas De Cuero Y Pieles	Quimcuero	Químicos Para Cueros
Foresto-Industria	Grafica	Gráfica
Foresto-Industria	Madymueb	Madera Y Muebles
Foresto-Industria	Oforeind	Otros Foresto-Industria
Foresto-Industria	Papel	Papel
Insumos Difundidos	Metales	Metales
Insumos Difundidos	Minerales	Minerales
Insumos Difundidos	Oinsdif	Otros Insumos Difundidos
Insumos Difundidos	Plasticos	Químicos
Medicamentos	Medic	Medicamentos
Medicamentos	Omedic	Otros Medicamentos
Medicamentos	Pactivos	Principios Activos
Otro Equipo De Transporte	Aeronautico	Aeronáutico
Otro Equipo De Transporte	Ferroviario	Ferroviario
Otro Equipo De Transporte	Motos	Motos
Otro Equipo De Transporte	Naval	Naval
Otro Equipo De Transporte	Oeqtrans	Otro Equipo De Transporte
Otro Equipo De Transporte	Op partes	Partes Para Otro Equipo De Transporte
Textil E Indumentaria	Fibras	Fibras

Cadena	Subcadena	Descripción
Textil E Indumentaria	Hilados	Hilados
Textil E Indumentaria	Indyacces	Indumentaria Y Accesorios
Textil E Indumentaria	Otexteind	Otros Textiles E Indumentaria
Textil E Indumentaria	Primtext	Primarios Textiles
Textil E Indumentaria	Quimtext	Químicos Para Textiles
Textil E Indumentaria	Tejidos	Tejidos
Otras Manufacturas	Animal	Otras Manufacturas De Origen Animal
Otras Manufacturas	Armas	Armamento
Otras Manufacturas	Cauyplast	Manufacturas De Caucho Y Plástico
Otras Manufacturas	Cineyfoto	Manufacturas Para Cine Y Fotografía (no Bk)
Otras Manufacturas	Corcho	Corcho Y Sus Manufacturas
Otras Manufacturas	Electrodom	Electrodomésticos
Otras Manufacturas	Eqelec	Otro Equipo De Eléctrico (no Bk)
Otras Manufacturas	Herramientas	Herramientas
Otras Manufacturas	Instrumentos	Aparatos E Instrumentos Varios (no Bk)
Otras Manufacturas	Juguetes	Juguetes
Otras Manufacturas	Mateqconst	Materiales Y Equipos De Construcción
Otras Manufacturas	Metypieprec	Metales Y Piedras Preciosas
Otras Manufacturas	Muebles	Otros Muebles
Otras Manufacturas	Oarte	Obras De Arte
Otras Manufacturas	Oartmet	Otros Artículos De Metal
Otras Manufacturas	Oficina	Artículos De Oficina
Otras Manufacturas	Omateriales	Otros Materiales
Otras Manufacturas	Oomanuf	Otras Manufacturas
Otras Manufacturas	Quimicos	Otros Químicos Y Sus Manufacturas
Otras Manufacturas	Tabaco	Tabaco Y Sus Manufacturas
Otras Manufacturas	Vegetal	Otras Manufacturas De Origen Vegetal
Otros Químicos	Caucho	Caucho
Otros Químicos	Cosmeticos	Cosméticos
Otros Químicos	Ooquim	Otros Químicos
Otros Químicos	Quimicos	Químicos

Tab. 4.2: Descripción subcadenas

C. Tabla de referencia: Descripción Usos

Uso	Descripción
PP	Productos Primarios
P&C	Partes Y Componentes
ST	Bienes Semiterminados
BK	Bienes De Capital
BC	Bienes De Consumo

Tab. 4.3: Descripción Usos

D. Medioides PAM clustering. K=50

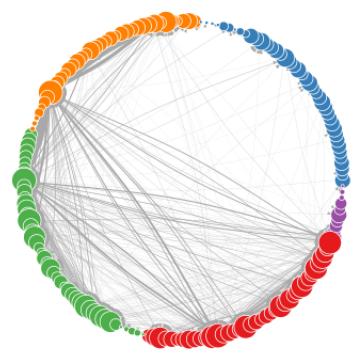
SITC	Description
1110	Non-alcohol.beverage,nes
6511	Yarn,of wool,animal hair
2631	Cotton (other than linters), not carded or combed
8426	Trousers, bib and brace overalls, breeches and shorts, women's or girls',
2117	Sheepskins and lambskins without wool on, raw (fresh, or salted, dried, l
3223	Peat (including peat litter), whether or not agglomerated
2891	Prec.mtl.ore,concentrats
4225	Castor oil and its fractions
7919	Rail.track fixtures,prts
8121	Boilrs.radiatrs,etc.n.el
2658	Veg. textile fibres, nes
5162	Aldehyde,etc.fnct.cmpndns
2485	Wood of non-coniferous species (including strips and friezes for parquet
5983	Artificial,prepared waxes
7526	Input or output units for automatic data processing machines, whether or
6572	Non-wovens, whether or not impregnated, coated, covered or laminated, n.e
6518	Yarn,staple fibres, etc.
2440	Cork,natural,raw;waste
5535	Toiletries, etc.
2513	Chemical wood pulp, dissolving grades
7429	Parts,pumps,liq.elevator
6966	Knives,forks,spoons,etc.
6644	Float,ground,plshd.glass
2683	Fine animal hair, not carded or combed
7234	Const.& mining machs.nes
6911	Metal structures,parts
3421	Propane, liquefied
5817	Fittings for tubes, pipes and hoses (e.g., joints, elbows, flanges), of p
6672	Diamonds.excl.industrial
6861	Zinc,zinc alloy,unwrght.
6726	Semi-finish.iron,steel
7462	Tapered roller bearings (including cone and tapered roller assemblies)
5233	Hypochlorites, etc.
4222	Palm oil, fractions
3345	Lubricating petroleum oils and oils obtained from bituminous minerals, ot
6562	Labels,badge etc.not emb
5752	Acrylic polymers
7263	Typesetting,mch,type etc
5416	Glycosides; glands etc.
7285	Parts public wrk mach etc
7359	Parts,nes,mch-tool w.mtl
8854	Watches,other than p.mtl
7224	Wheeled tractors nes
6954	Hand tools,etc. nes

SITC	Description
7722	Printed circuits
8421	Overcoats, oth. coats etc.
7249	Pts, txtle, dom washng mch
8982	Musical instruments nes
7472	Valves for oleohydraulic or pneumatic transmissions
7169	Parts, n.e.s., suitable for use with the machines of group 716

Tab. 4.4: Medioides. PAM. k=50

E. Figuras Complementarias

● Asia ● Europe ● Africa ● Americas ● Oceania



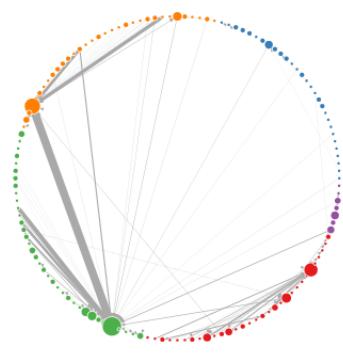
(a) 1 %



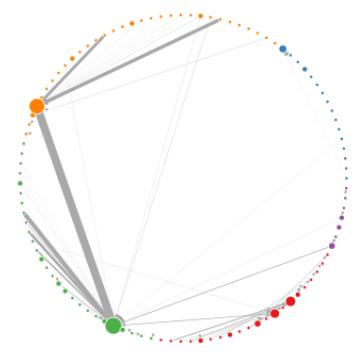
(b) 5 %



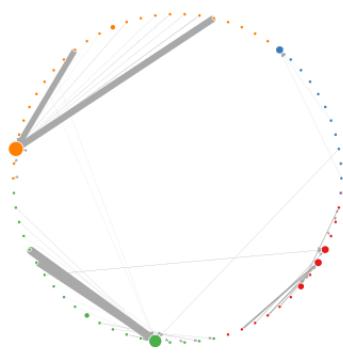
(c) 10 %



(d) 15 %



(e) 20 %



(f) 25 %

Fig. 4.1: Grafos. 2016, según punto de corte. Ponderado por Valor comercial. Importaciones

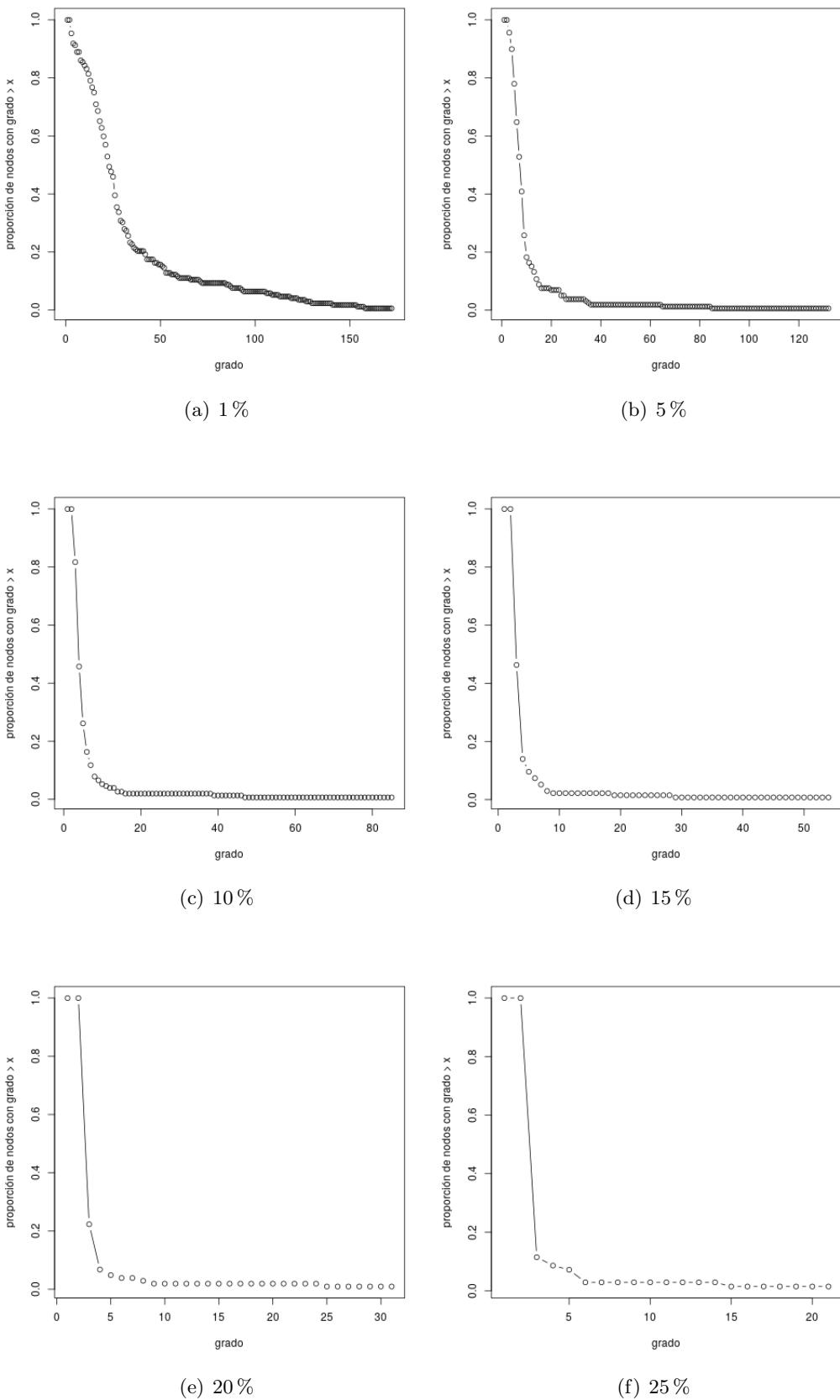


Fig. 4.2: Distribución acumulada del grado. 2016, según punto de corte. Importaciones

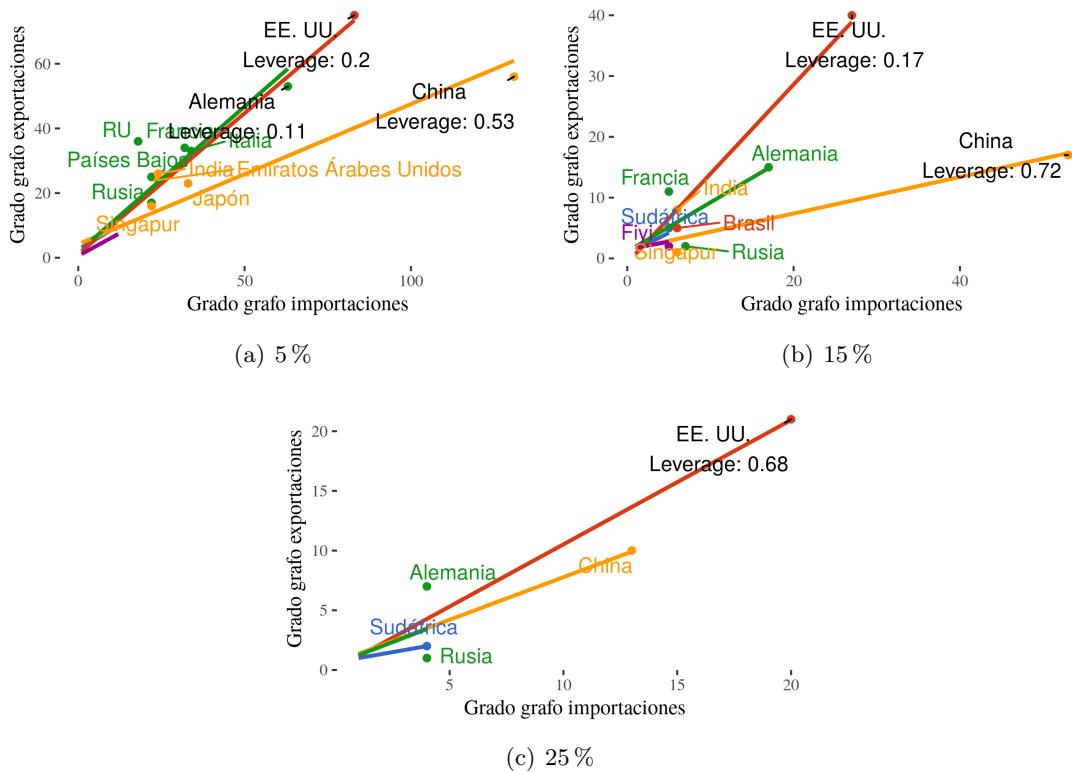


Fig. 4.3: Grado total. Grafo exportaciones respecto grafo importaciones. Año 2016

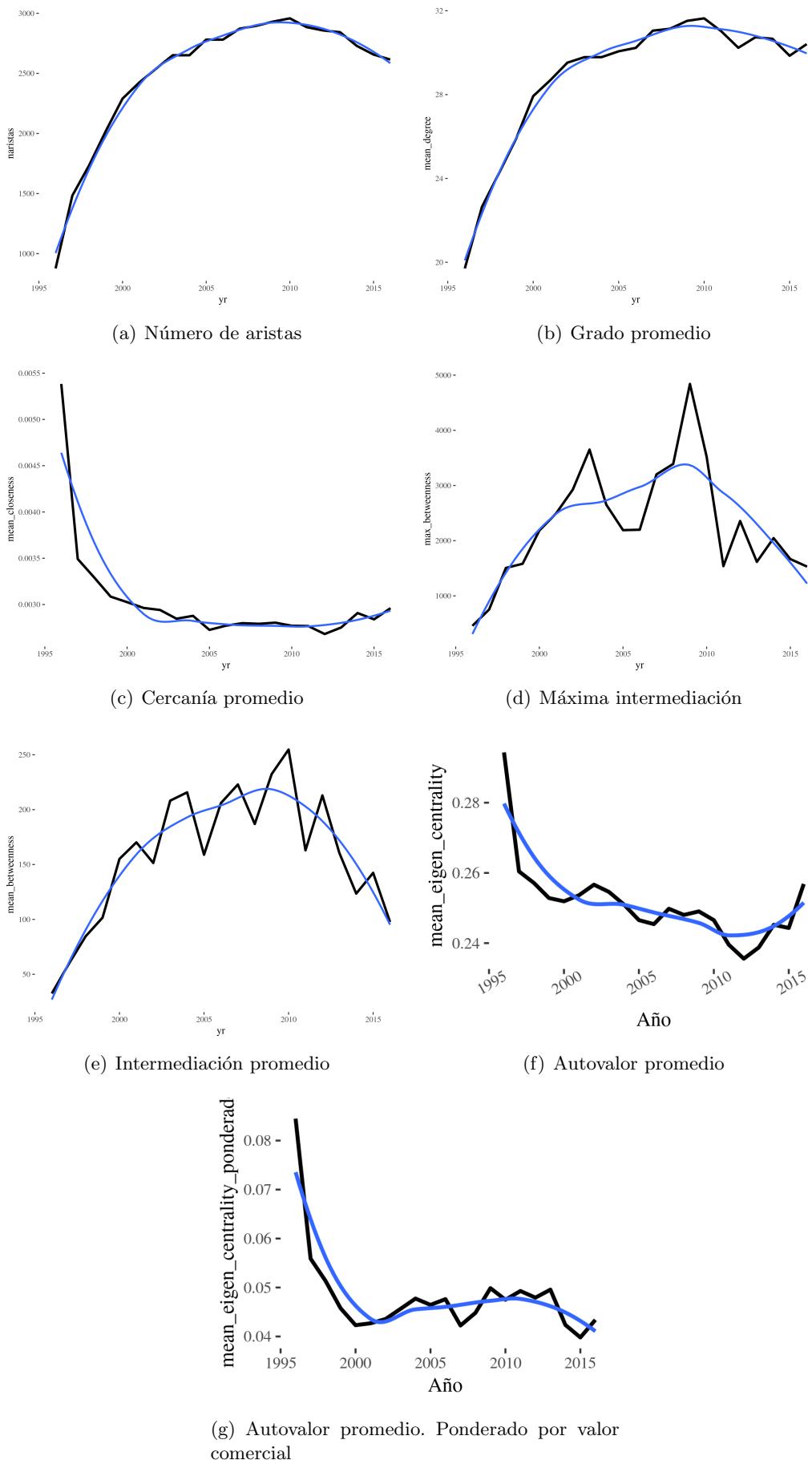


Fig. 4.4: Importaciones. Umbral 1 %. 1996-2016

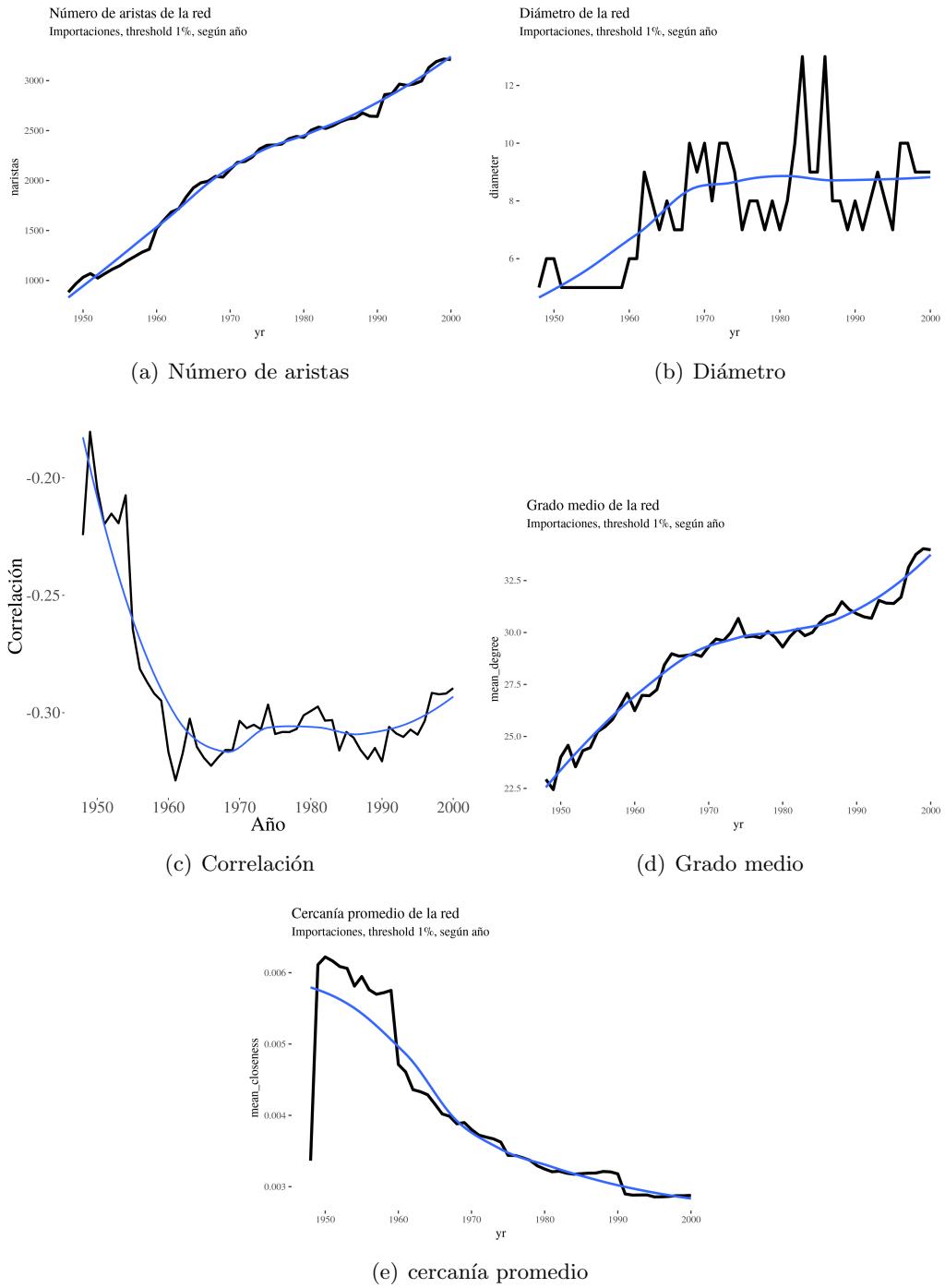


Fig. 4.5: Evolución de la estructura de la red. Importaciones. Umbral 1 %.1948-2000

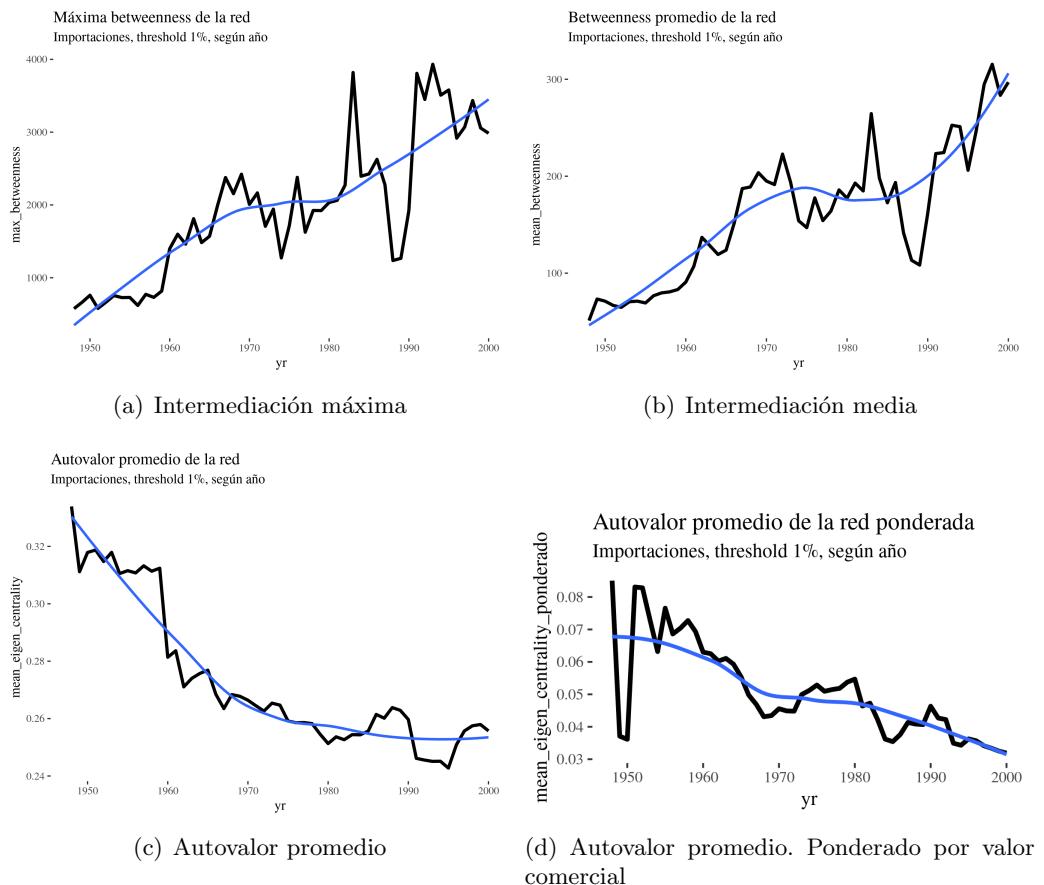


Fig. 4.6: Evolución de la estructura de la red. Importaciones. Umbral 1 %. 1948-2000

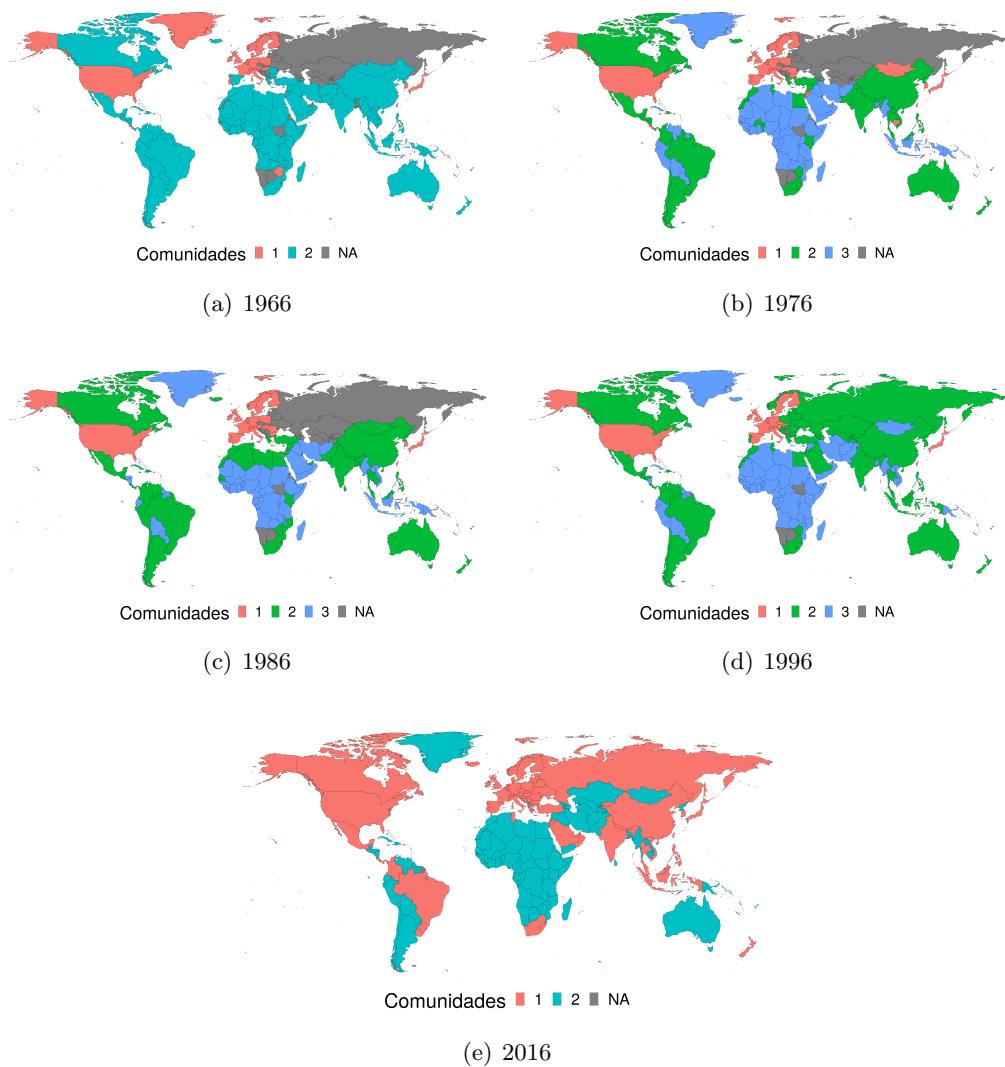


Fig. 4.7: Proyección a países del grafo bipartito. Clustering Walktrap. Exportaciones

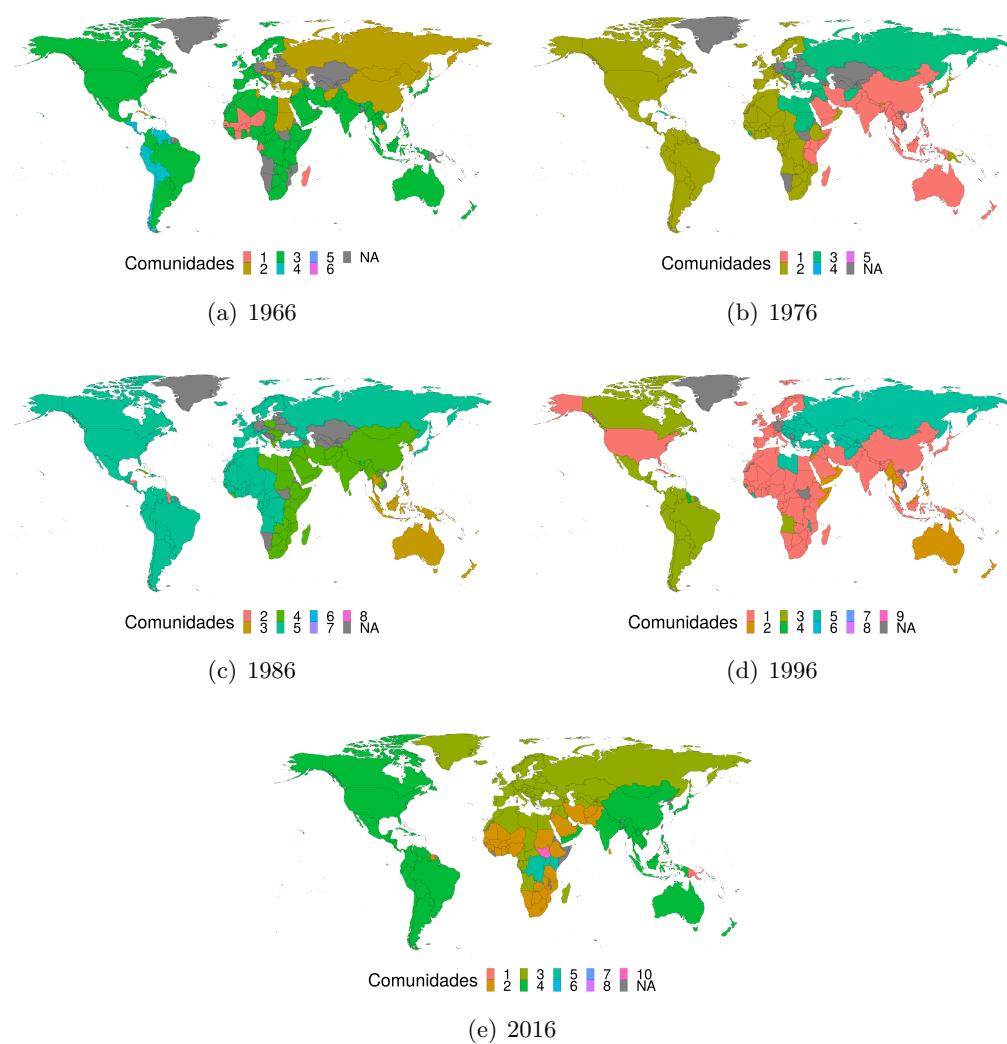


Fig. 4.8: Grafo a nivel países. Clustering walktrap. Exportaciones

Bibliografía

- Charu C Aggarwal, Alexander Hinneburg, and Daniel A Keim. On the surprising behavior of distance metrics in high dimensional space. In *International conference on database theory*, pages 420–434. Springer, 2001.
- J J Allaire. *rsconnect: Deployment Interface for R Markdown Documents and Shiny Applications*, 2019. URL <https://cran.r-project.org/package=rsconnect>.
- Vincent Arel-Bundock. countrycode: Convert Country Names and Country Codes, 2017.
- Jeffrey B Arnold. ggthemes: Extra Themes, Scales and Geoms for 'ggplot2', 2017.
- Richard Auty. *Sustaining development in mineral economies: the resource curse thesis*. Routledge, 2002.
- Bela Balassa. Trade liberalisation and “revealed” comparative advantage. *The manchester school*, 33(2):99–123, 1965.
- M Balat. Coal in the global energy scene. *Energy Sources, Part B: Economics, Planning, and Policy*, 5(1):50–62, 2009.
- John Ballingall and Ralph Lattimore. Farming in New Zealand: The state of play and key issues for the backbone of the New Zealand economy. Technical report, New Zealand Trade Consortium Working Paper, 2004.
- Marta Bekerman and Cecilia Rikap. Integración regional y diversificación de exportaciones en el Mercosur: el caso de Argentina y Brasil. *Revista Cepal*, 2010.
- K. Bhattacharya, G. Mukherjee, J. Saramäki, K. Kaski, and S. S. Manna. The international trade network: Weighted network analysis and modelling. *Journal of Statistical Mechanics: Theory and Experiment*, 2008(2), 2008. ISSN 17425468. doi: 10.1088/1742-5468/2008/02/P02002.
- David M Blei, Andrew Y Ng, and Michael I Jordan. Latent dirichlet allocation. *Journal of machine Learning research*, 3(Jan):993–1022, 2003.
- M. I. Blei, D. M.; Ng, A. Y.; Jordan. Latent dirichlet allocation. *Journal of Machine Learning Research*, 3(4-5):993–1022, 2003. ISSN 15324435. doi: 10.1162/jmlr.2003.3.4-5.993.
- Magnus Blomstrom and Ari Kokko. From natural resources to high-tech production: the evolution of industrial competitiveness in Sweden and Finland. 2003.
- Vincent D Blondel, Jean-Loup Guillaume, Renaud Lambiotte, and Etienne Lefebvre. Fast unfolding of communities in large networks. *Journal of statistical mechanics: theory and experiment*, 2008(10):P10008, 2008.
- Luciano Bolinaga and Ariel Slipak. El Consenso de Beijing y la reprimarización productiva de América Latina: el caso argentino. *Problemas del desarrollo*, 46(183):33–58, 2015.

- Guido Caldarelli, Matthieu Cristelli, Andrea Gabrielli, Luciano Pietronero, Antonio Scala, and Andrea Tacchella. A network analysis of countries' export flows: firm grounds for the building blocks of the economy. *PloS one*, 7(10):e47278, 2012.
- Eliana Cardoso. A brief history of trade policies in Brazil: From ISI, export promotion and import liberalization to multilateral and regional agreements. In *conference on “The Political Economy and Trade Policy in the BRICS”*, March, pages 27–28. Citeseer, 2009.
- Juan Iñigo Carrera. *La renta de la tierra: formas, fuentes y apropiación*. Ediciones Imago Mundi, 2017.
- Jorge Carrillo and Alfredo Hualde. Third generation maquiladoras? The Delphi-General Motors case. *Journal of Borderlands Studies*, 13(1):79–97, 1998.
- Ha-joon Chang and Kiryl Zach. Industrial development in Asia: Trends in industrialization and industrial policy experiences of developing Asia. Technical report, WIDER Working Paper, 2018.
- Winston Chang, Joe Cheng, J J Allaire, Yihui Xie, and Jonathan McPherson. *shiny: Web Application Framework for R*, 2018. URL <https://cran.r-project.org/package=shiny>.
- Mauro F Chávez-Rodríguez, Rafael Garaffa, Gisela Andrade, Gonzalo Cárdenas, Alexandre Szklo, and André F P Lucena. Can Bolivia keep its role as a major natural gas exporter in South America? *Journal of Natural Gas Science and Engineering*, 33:717–730, 2016.
- Hollis Burnley Chenery, Sherman Robinson, Moshe Syrquin, and Syrquin Feder. *Industrialization and growth*. Oxford University Press New York, 1986.
- William Chow. An Anatomy of the World Trade Network. pages 1–20, 2013.
- Neil M Coe, Martin Hess, Henry Wai-chung Yeung, Peter Dicken, and Jeffrey Henderson. ‘Globalizing’regional development: a global production networks perspective. *Transactions of the Institute of British geographers*, 29(4):468–484, 2004.
- William J Coffey. The ‘newer’international division of labour. *The global economy in transition*, pages 40–61, 1996.
- Agostina Costantino. ¿Gatopardismo sojero?: Los efectos de la bonanza sojera sobre el cambio estructural en Argentina y Brasil. *Nueva Sociedad*, (244):84–96, 2013.
- Gabor Csardi and Tamas Nepusz. The igraph software package for complex network research, 2006.
- Dilip K Das. *Regionalism in global trade*. Edward Elgar Publishing, 2004.
- Giorgio Fagiolo, Javier Reyes, and Stefano Schiavo. The Evolution of the World Trade Web Giorgio. 2007.
- Giorgio Fagiolo, Javier Reyes, and Stefano Schiavo. The evolution of the world trade web: A weighted-network analysis. *Journal of Evolutionary Economics*, 20(4):479–514, 2010. ISSN 09369937. doi: 10.1007/s00191-009-0160-x.

- Ying Fan, Suting Ren, Hongbo Cai, and Xuefeng Cui. The state's role and position in international trade: A complex network perspective. *Economic Modelling*, 39:71–81, 2014. ISSN 02649993. doi: 10.1016/j.econmod.2014.02.027. URL <http://dx.doi.org/10.1016/j.econmod.2014.02.027>.
- Robert C Feenstra and Gordon H Hanson. Foreign direct investment and relative wages: Evidence from Mexico's maquiladoras. *Journal of international economics*, 42(3-4):371–393, 1997.
- M Ennes Ferreira and Others. The Topology of African Exports: emerging patterns on spanning trees. Technical report, arXiv. org, 2016.
- Renato G Jr. Flôres. The World Fragmentation of Production and Trade Concepts and Basic Issues. In *International Workshop Integração Produtiva-Lições da Ásia e Europa para o MERCOSUL*, CEPAL/Brasil, page 7, 2008.
- Koen Frenken. A complexity approach to innovation networks. The case of the aircraft industry (1909–1997). *Research Policy*, 29(2):257–272, 2000.
- Folker Fröbel, Jürgen Heinrichs, and Otto Kreye. The new international division of labour. *Information (International Social Science Council)*, 17(1):123–142, 1978.
- Gary Gereffi. The organization of buyer-driven global commodity chains: how US retailers shape overseas production networks. *Contributions in economics and economic history*, page 95, 1994.
- Gary Gereffi, John Humphrey, and Timothy Sturgeon. The governance of global value chains. *Review of international political economy*, 12(1):78–104, 2005.
- Kristian S Gleditsch. Expanded Trade and GDP Data. *Journal of Conflict Resolution*, 46:712–24, 2002.
- Jun Guan, Xiaoyu Xu, Shan Wu, and Lizhi Xing. Measurement and simulation of the relatively competitive advantages and weaknesses between economies based on bipartite graph theory. *PLOS ONE*, 13(5):1–28, 2018. doi: 10.1371/journal.pone.0197575. URL <https://doi.org/10.1371/journal.pone.0197575>.
- Stephan Haggard. The newly industrializing countries in the international system. *World Politics*, 38(2):343–370, 1986.
- Jon Haveman and David Hummels. Alternative hypotheses and the volume of trade: the gravity equation and the extent of specialization. *Canadian Journal of Economics/Revue canadienne d'économique*, 37(1):199–218, 2004.
- Keith Head and Thierry Mayer. Gravity Equations: Workhorse, Toolkit, and Cookbook. *Handbook of International Economics*, 4:131–195, jan 2014. ISSN 1573-4404. doi: 10.1016/B978-0-444-54314-1.00003-3. URL <https://www.sciencedirect.com/science/article/pii/B9780444543141000033?via%23Dihub>.
- Jeffrey Henderson. *Globalisation of high technology production*. Routledge, 2002.

- Sebastian Henn. Transnational entrepreneurs and the emergence of clusters in peripheral regions. The case of the diamond cutting cluster in Gujarat (India). *European Planning Studies*, 21(11):1779–1795, 2013.
- C. A. Hidalgo, B. Winger, A. L. Barabási, and R. Hausmann. The product space conditions the development of nations. *Science*, 317(5837):482–487, 2007. ISSN 00368075. doi: 10.1126/science.1144581.
- César Hidalgo and Ricardo Hausmann. The building blocks of economic complexity. *Proceedings of the National Academy of the Sciences of the United States of America*, 106(26):10570–10575, 2009. ISSN 0027-8424. doi: 10.1073/pnas.0900943106.
- César A. Hidalgo. The Dynamics of Economic Complexity and the Product Space over a 42 year period. *CID Working Paper*, 189(189):20, 2009. ISSN 6507247197.
- Matthew D Hoffman, David M Blei, Chong Wang, and John Paisley. Stochastic variational inference. *The Journal of Machine Learning Research*, 14(1):1303–1347, 2013.
- Otto Holman. *Integrating Southern Europe: EC expansion and the transnationalization of Spain*. Routledge, 2005.
- John Humphrey and Hubert Schmitz. *Governance and upgrading: linking industrial cluster and global value chain research*, volume 120. Institute of Development Studies Brighton, 2000.
- Francis Hutchinson and Others. Globalisation and the 'Newer' international division of labour. 2004.
- Leonard Kaufman and Peter J Rousseeuw. Clustering by means of medoids. Statistical Data Analysis based on the L1 Norm. *Y. Dodge, Ed*, pages 405–416, 1987.
- Andras Köves and Gabor Oblath. Hungarian foreign trade in the 1970's. *Acta Oeconomica*, pages 89–109, 1983.
- Sanjaya Lall. The Technological structure and performance of developing country manufactured exports, 1985-98. *Oxford development studies*, 28(3):337–369, 2000.
- Sanjaya Lall, John Weiss, and Jinkang Zhang. The "sophistication" of exports: a new trade measure. *World development*, 34(2):222–237, 2006.
- Greg Linden. Building production networks in Central Europe: the case of the electronics industry. 1998.
- Catriona J MacLeod and Henrik Moller. Intensification and diversification of New Zealand agriculture since 1960: An evaluation of current indicators of land use change. *Agriculture, ecosystems & environment*, 115(1-4):201–218, 2006.
- James MacQueen and Others. Some methods for classification and analysis of multivariate observations. In *Proceedings of the fifth Berkeley symposium on mathematical statistics and probability*, volume 1, pages 281–297. Oakland, CA, USA, 1967.
- Adam McKissack, Jennifer Chang, Robert Ewing, Jyoti Rahman, and Others. Structural effects of a sustained rise in the terms of trade. *Australian Treasury*, 2008.

- Raymond F Mikesell. *The world copper industry: structure and economic analysis.* RFF Press, 2013.
- Andrea Molinari and Jesica Yamila de Angelis. Especialización y complementación productiva en el MERCOSUR: un Enfoque de Cadenas Productivas de Valor. 2016.
- Theodore H Moran. *Multinational corporations and the politics of dependence: Copper in Chile.* Princeton University Press, 2014.
- Antoin E Murphy. The' Celtic Tiger', An Analysis of Ireland's Economic Growth Performance. 2000.
- Erich Neuwirth. RColorBrewer: ColorBrewer Palettes, 2014.
- Weng Hoong Ng. *Singapore, the Energy Economy: From the First Refinery to the End of Cheap Oil, 1960-2010.* Routledge, 2013.
- ONU. Standard International Trade Classification Revision 4. *Statistical Papers Series M No,* 34, 2006.
- F Pedregosa, G Varoquaux, A Gramfort, V Michel, B Thirion, O Grisel, M Blondel, P Prettenhofer, R Weiss, V Dubourg, J Vanderplas, A Passos, D Cournapeau, M Brucher, M Perrot, and E Duchesnay. Scikit-learn: Machine Learning in Python . *Journal of Machine Learning Research,* 12:2825–2830, 2011.
- Pascal Pons and Matthieu Latapy. Computing communities in large networks using random walks. In *International symposium on computer and information sciences*, pages 284–293. Springer, 2005.
- David M W Powers. Applications and Explanations of Zipf's Law. In *Proceedings of the Joint Conferences on New Methods in Language Processing and Computational Natural Language Learning, NeMLaP3/CoNLL '98*, pages 151–160, Stroudsburg, PA, USA, 1998. Association for Computational Linguistics. ISBN 0-7258-0634-6. URL <http://dl.acm.org/citation.cfm?id=1603899.1603924>.
- R Core Team. R: A Language and Environment for Statistical Computing, 2017.
- Daniel J Redo, T Mitchell AIDE, Matthew L Clark, and María José Andrade-Núñez. Impacts of internal and external policies on land change in Uruguay, 2001–2009. *Environmental Conservation,* 39(2):122–131, 2012. doi: 10.1017/S0376892911000658.
- David Ricardo. *Principios de economía política y tributación.* Claridad, Buenos Aires, 2007.
- Dani Rodrik. Getting interventions right: how South Korea and Taiwan grew rich. *Economic policy,* 10(20):53–107, 1995.
- Michael Ross. *The oil curse: How petroleum wealth shapes the development of nations.* Princeton University Press, 2012.
- Menahem Sevdermish, Alan R Miciak, and Alfred A Levinson. The rise to prominence of the modern diamond cutting industry in India. *Gems & Gemology,* 34(1):4–23, 1998.

- Avraham Shama. Determinants of entry strategies of US companies into Russia, the Czech Republic, Hungary, Poland, and Romania. *Thunderbird International Business Review*, 42(6):651–676, 2000.
- Leslie Sklair. *Assembling for development: The maquila industry in Mexico and the United States*. Routledge, 2012.
- Kamil Slowikowski. ggrepel: Repulsive Text and Label Geoms for 'ggplot2', 2017.
- Guido Starosta. *The new international division of labour: global transformation and uneven development*. Springer, 2016a.
- Guido Starosta. Revisiting the new international division of labour thesis. In *The new international division of labour*, pages 79–103. Springer, 2016b.
- Mika J Straka, Guido Caldarelli, and Fabio Saracco. Grand canonical validation of the bipartite international trade network. *Physical Review E*, 96(2):22306, 2017.
- Akira Suehiro. *Catch-up industrialization: The trajectory and prospects of East Asian economies*. NUS Press, 2008.
- Maristella Svampa. Commodities consensus: Neoextractivism and enclosure of the commons in Latin America. *South Atlantic Quarterly*, 114(1):65–82, 2015.
- G van Rossum. Python tutorial. Technical Report CS-R9526, Centrum voor Wiskunde en Informatica (CWI), Amsterdam, may 1995.
- Fiona Venn. *The oil crisis*. Routledge, 2016.
- Vincent Q Vu. ggbiplot: A ggplot2 based biplot, 2011.
- Joe H Ward Jr. Hierarchical grouping to optimize an objective function. *Journal of the American statistical association*, 58(301):236–244, 1963.
- Carsten Weerth. Basic Principles of Customs Classifications under the Harmonized System. *Global Trade & Cust. J.*, 3:61, 2008.
- Hadley Wickham. tidyverse: Easily Install and Load the 'Tidyverse', 2017.
- Claus O Wilke. ggridges: Ridgeline Plots in 'ggplot2', 2017.
- WTO. World Trade Statistics Review 2017. Technical report, 2017.
- Tao Zhou, Jie Ren, Matúš Medo, and Yi-Cheng Zhang. Bipartite network projection and personal recommendation. *Physical Review E*, 76(4):46115, 2007.