



Facultad de Ciencias Exactas y Naturales



Universidad de Buenos Aires

Maestría en Explotación de Datos y Descubrimiento del Conocimiento

Tesis de Maestría

Movilidad interindustrial de empleo en la Argentina: Análisis usando redes complejas

Autor

Lic. Sergio Andrés De Raco

Directora

Dra. Viktoriya Semeshenko

Diciembre 2021

Agradecimientos

Agradezco a mi directora de tesis, Viktoriya Semeshenko, por introducirme al mundo de las redes complejas y por las numerosas conversaciones y sugerencias que ayudaron a organizar y mejorar el trabajo plasmado en esta tesis.

Agradezco a Daniel Heymann por invitarme a la aventura de la investigación y a formar parte del Instituto Interdisciplinario de Economía Política, sin lo cual esta tesis no habría sido escrita.

Agradezco a la Universidad de Buenos Aires por la formación y por brindarme la posibilidad de devolver con la docencia y la investigación parte de lo recibido en mi formación.

Agradezco a mi familia por el apoyo y la contención a lo largo de todo este proyecto.

A Verónica, Goran y Andrik

Resumen

En esta tesis se analiza el fenómeno de la movilidad intersectorial del empleo a partir de registros administrativos de gran nivel de detalle del período 1996-2020 mediante enfoques de redes complejas.

La movilidad laboral resume aspectos dinámicos y estructurales del funcionamiento de los mercados de trabajo. En particular, la movilidad laboral de las personas que transitan desde un empleo hacia otro conllevan intercambios de capacidades y habilidades entre sectores de actividad. Al observar estos movimientos es posible extraer información valiosa que permite caracterizar aspectos subyacentes de la estructura productiva nacional.

En economía usualmente se analiza información de actividad económica sectorial agregada en un conjunto reducido de categorías, familiares a analistas y tomadores de decisiones de políticas públicas, que proveen información comprensible en términos de conceptos tradicionales, analizables mediante métodos también tradicionales. Los datos de empleo que surgen de registros administrativos identifican a gran nivel de detalle la actividad de los empleadores, lo cual presenta un desafío en cuanto a la interpretación de los sectores productivos y a su abordaje analítico. Estos datos originan naturalmente una red temporal de interacciones bipartita entre trabajadores y empresas. A partir de estas observaciones surge interés en indagar acerca de las características estructurales que revelan los intercambios de empleo entre industrias a tales niveles de detalle, aún poco explorados, así como la medida en que son estables en el tiempo. Adicionalmente, interesa identificar agrupamientos relevantes de los flujos entre industrias y la información que aquellos permitan extraer respecto al entramado productivo.

En esta tesis se pone el foco en extraer información de estos flujos intersectoriales de empleo en base a datos de registro con identificación de la actividad económica de los empleadores a un nivel de detalle prácticamente inexplorado. Se encuentra que las relaciones intersectoriales que surgen de las transiciones de empleo interindustriales revelan una historia de elevada y persistente conectividad que evidencia una estructura de interacciones densa, organizada en un solo componente conexo con caminos cortos que permiten verificar la propiedad de "mundo pequeño". Por otro lado, se identifican estructuras de organización sectorial de centro y periferia, que aportan información de aspectos funcionales de las interacciones entre sectores mediadas por el empleo en consonancia con la organización productiva, que permite detectar dos grandes núcleos de alta interacción y otros subgrupos de sectores periféricos relativamente estables en el tiempo. En términos de la evolución del sistema, se evidencian cambios en la organización de la estructura de interconexiones de mediano plazo en "regímenes" de conectividad de la movilidad laboral, en correlación con la actividad económica agregada.

En suma, se encuentra que el abordaje del análisis de flujos de empleo entre industrias a partir de la conectividad provista por metodologías de redes complejas enriquece el estudio del fenómeno de la movilidad laboral aportando nuevas analogías y métricas para su mejor comprensión. Desde el punto de vista de la minería de datos se logra extraer conocimiento original y útil, y aportar métodos accesibles y replicables para reproducir, mejorar y extender el análisis aquí propuesto.

Palabras clave: Movilidad laboral, Redes complejas, Detección de comunidades.

Summary

This thesis analyzes the phenomenon of intersectoral mobility of employment based on highly detailed administrative records for the period 1996-2020 using complex network approaches.

Labor mobility summarizes dynamic and structural aspects of the functioning of labor markets. In particular, the labor mobility of people who move from one job to another entails exchanges of skills and abilities between sectors of activity. By observing these movements it is possible to extract valuable information that allows characterizing underlying aspects of the national productive structure.

In economics, information on aggregate sectoral economic activity is usually analyzed in a reduced set of categories, familiar to analysts and public policy decision-makers, which provide information that is understandable in terms of traditional concepts, which can be analyzed using traditional methods. The employment data arising from administrative records identifies the activity of employers at a high level of detail, which presents a challenge in terms of the interpretation of the productive sectors and their analytical approach. This data naturally can be interpreted as a temporary network of bipartite interactions between workers and companies. From these observations, interest arises in inquiring about the structural characteristics that reveal the exchanges of employment between industries at such levels of detail, still little explored, as well as the extent to which they are stable over time. Additionally, it is interesting to identify relevant groupings of flows between industries and the information that they allow to extract regarding the productive network.

This thesis focuses on extracting information from these intersectoral employment flows based on registration data identifying the economic activity of employers at a virtually unexplored level of detail. It is found that the intersectoral relationships that arise from the interindustrial employment transitions reveal a history of high and persistent connectivity that evidences a dense structure of interactions, organized in a single connected component with short paths that allow verification of "small world" ownership. On the other hand, center and periphery sectoral organization structures are identified, which provide information on functional aspects of the interactions between sectors mediated by employment in consonance with the productive organization, which allows detecting two large nuclei of high interaction and other subgroups of peripheral sectors relatively stable over time. In terms of the evolution of the system, there are changes in the organization of the structure of interconnections in the medium term in "regimes" of connectivity of labor mobility, in correlation with aggregate economic activity.

In sum, it is found that the approach to the analysis of employment flows between industries based on the connectivity provided by complex network methodologies enriches the study of the phenomenon of labor mobility, providing new analogies and metrics for better understanding. From the point of view of data mining, it is possible to extract original and useful knowledge, and provide accessible and replicable methods to reproduce, improve and extend the analysis proposed here.

Keywords: Labor mobility, Complex networks, Community detection.

Movilidad interindustrial de empleo en la Argentina: Análisis usando redes complejas

Sergio A. De Raco

Diciembre 2021

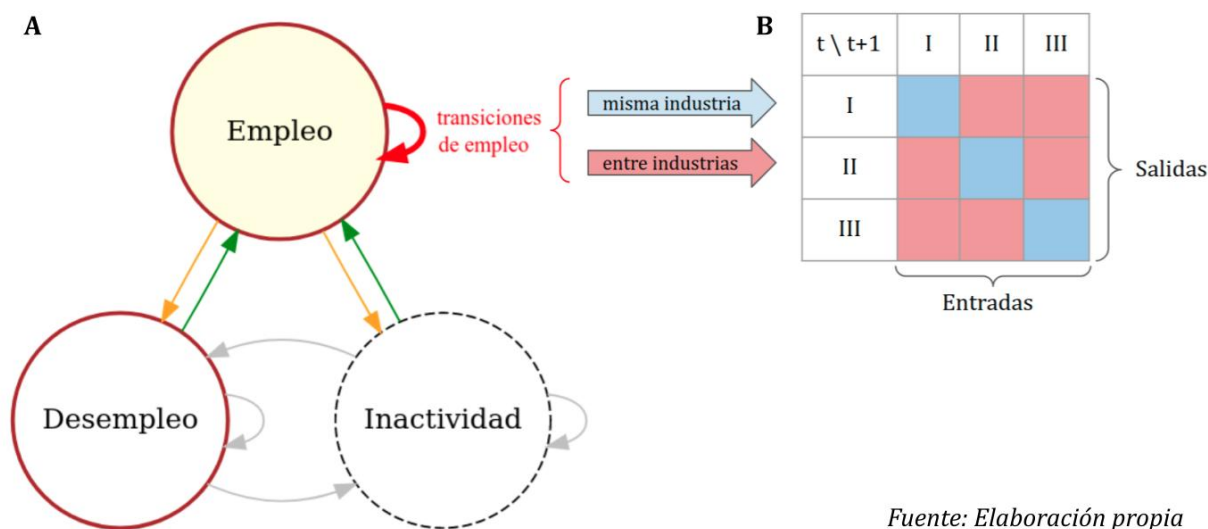
Índice

1. Introducción	1
1.1. Objetivos	7
2. Materiales y métodos	8
2.1. Datos	8
2.2. Métodos	10
3. Redes de flujos de empleo	16
3.1. Perfil	16
3.2. Evolución temporal y agregación	17
3.3. Análisis descriptivo	24
4. Análisis de agrupamientos	31
4.1. Coreness	32
4.2. Comunidades	33
5. Discusión final	42
Epílogo	44
Referencias	45
Anexos	47
5.1. Material suplementario	47
5.2. Anexo metodológico	56
5.3. Clasificaciones sectoriales y nodos	57
5.4. Reproducibilidad	67

1. Introducción

Motivación. La movilidad laboral resume aspectos dinámicos y estructurales del funcionamiento de los mercados de trabajo. Refiere a un tema primordial y recurrente en la literatura económica moderna por su aporte a la explicación de la evolución de la actividad económica en función del rol del trabajo como factor productivo y su relevancia en la determinación del desempleo (Keynes 1936).

A lo largo de la vida, las personas adquieren y desarrollan habilidades y trabajan en diversas actividades económicas a intervalos de distintas duraciones. La movilidad laboral es un fenómeno que surge de analizar en términos agregados estos procesos individuales en términos de transiciones entre los estados de empleo, desempleo e inactividad de las personas (Fig. 1), que se vinculan entre sí mediante flujos dirigidos entre cada uno de ellos y respecto a sí mismos.



Fuente: Elaboración propia

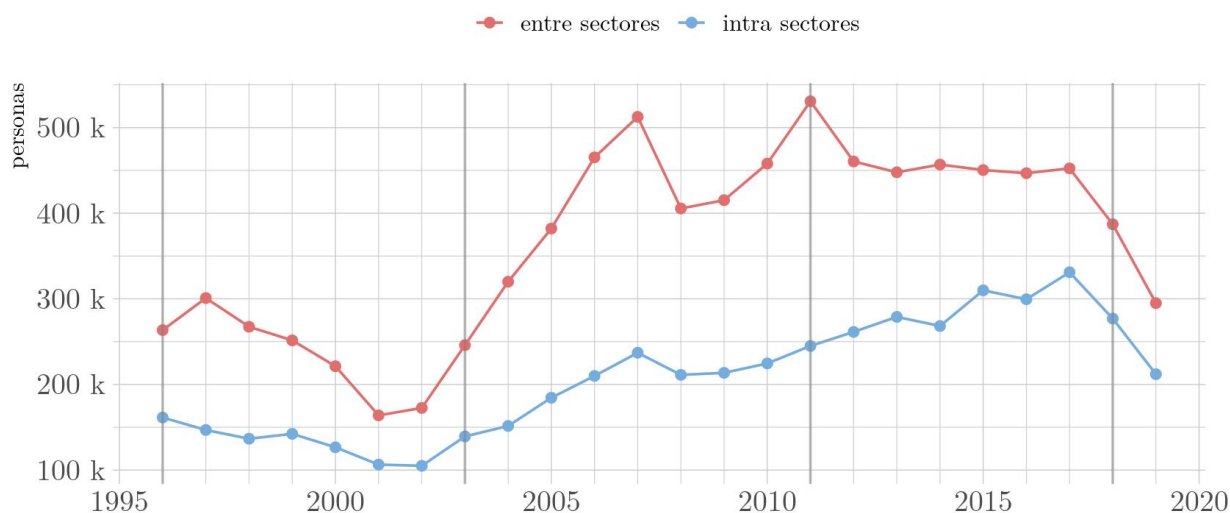
Figura 1: Flujos y estados de movilidad laboral, transiciones de empleo entre y dentro de sectores. Esquema general que representa la dinámica de flujos de empleo (A). Ejemplo de matriz de transiciones interanuales de empleo dentro y entre sectores de actividad I, II y III (B).

Las historias de ingreso al empleo (flechas verdes en Fig. 1) suelen revestirse de un carácter económico y social positivo, tanto para las personas como para la economía en general, dado que aparecen naturalmente como eventos en los que ‘todos ganan.’ En cambio, no resulta tan clara la valoración de los flujos de salida del empleo (flechas naranjas y roja en Fig. 1), que pueden considerarse individual y socialmente positivos si resultan en un cambio de empleo y negativos si su destino es el desempleo o la inactividad.

En particular, la movilidad laboral de las personas que transitan desde un empleo hacia otro (flecha roja en Fig. 1) pone en evidencia vinculaciones entre actividades productivas que conllevan información acerca de la estructura productiva de la economía. Estas transiciones entre empleos (*job-to-job transitions*) resumen coincidencias entre personas que parten de un empleo ofreciendo sus habilidades y empresas que demandan un puesto de trabajo para una actividad productiva específica. Al observar los movimientos dentro del empleo a la luz de la actividad económica de los empleadores, es posible extraer información valiosa que permite caracterizar aspectos subyacentes de la estructura productiva nacional asociados a los intercambios de capacidades y habilidades entre sectores de actividad. En la Fig. 1 (derecha) se presenta un esquema de la matriz típica que resume las transiciones de empleo entre sectores productivos. En esta tesis se pone el foco en extraer información de estos flujos **intersectoriales** de empleo, en particular aquellos que ocurren entre firmas dedicadas a actividades económicas distintas (celdas rojas en la matriz de transición, en Fig. 1.B).

En Argentina, el Ministerio de Trabajo, Empleo y Seguridad Social (MTEySS) de la Nación cuenta con datos de registros administrativos del empleo formal privado que es informado periódicamente al fisco por

las empresas. Mediante el armado de paneles dinámicos de individuos, estos datos se utilizan para analizar transiciones interanuales de empleo. En la Fig. 2 se observa la evolución de las transiciones de empleo formal privado en las últimas décadas, registradas entre empresas que declaraban el mismo sector de actividad (transiciones *intrasectoriales*, en celdas azules de la diagonal principal de la matriz en la Fig. 1.B) y entre empresas de diferentes sectores de actividad (transiciones *intersectoriales*, en celdas rojas de dicha matriz) a cuatro dígitos del clasificador CIIU Rev 3 utilizado en el Sistema Integrado Previsional Argentino (SIPA). Según se observa, es posible caracterizar a grandes rasgos el comportamiento de la serie en términos de cuatro períodos característicos: 1) uno declinante asociado a la última parte de la vigencia del Plan de Convertibilidad y la crisis posterior (1996-2002); 2) uno de recuperación y crecimiento post-crisis (2003-2011); 3) otro de estancamiento (2012-2017); y 4) uno de fuerte caída asociado al cambio de gobierno y el comienzo de la pandemia COVID-19 (2018-2020).



Fuente: Elaboración propia en base a SIPA

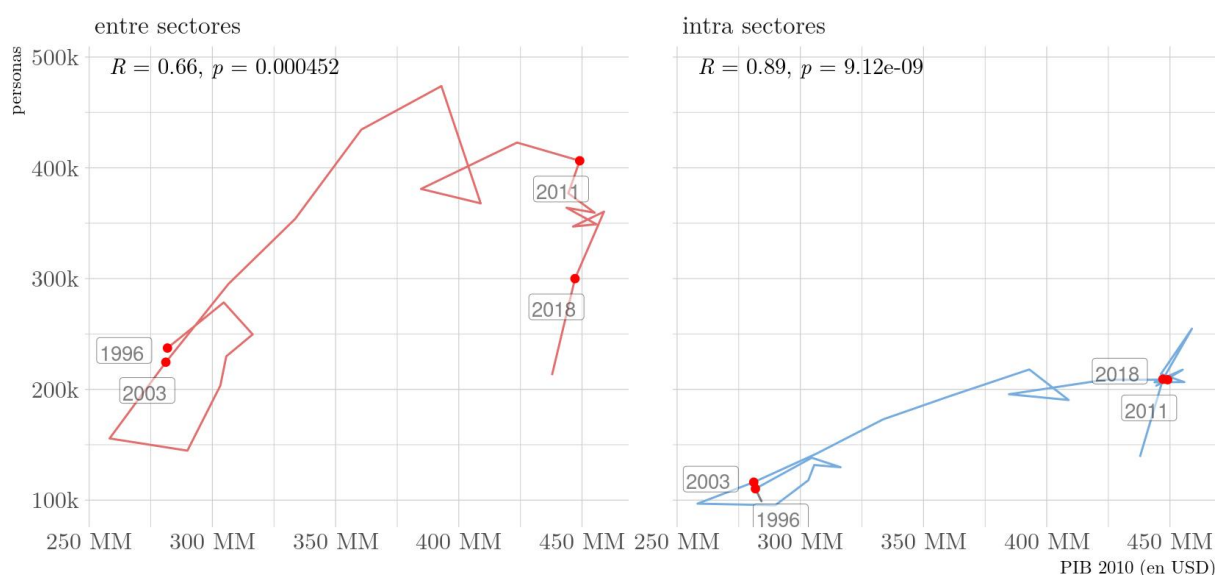
Figura 2: Transiciones de empleo formal privado dentro y entre sectores de actividad económica. Actividades a 4 dígitos CIIU Rev 3.

En particular, al observar las transiciones de empleo formal privado, en la Fig. 3 se destaca una correlación elevada con el producto interno (PIB) real, que resume el comportamiento de la actividad económica en términos de volumen. En el caso de las transiciones *intrasectoriales*, que representan cambios de empleo ocurridos entre firmas de la misma actividad económica, la correlación es esperablemente mayor en razón de que se trata de movimientos que posiblemente resultan de valorar la acumulación de capital humano dentro de la industria particular y potencialmente impliquen mejoras en la productividad. En cuanto a las transiciones *intersectoriales*, la correlación con el nivel general de actividad resulta sensiblemente elevada, sin embargo no resulta tan clara la intuición de las vías en las que estos cambios de empleo afectan el desempeño productivo.

Los datos de registro incluyen la identificación detallada de la actividad económica de los empleadores. Sin embargo, la complejidad de trabajar con datos muy desagregados usando métodos tradicionales, tanto por el volumen como por la heterogeneidad, lleva usualmente a plantear el análisis en términos agregados con lo que buena parte de la información contenida en ellos permanece inexplorada desde hace décadas. En este sentido, las estadísticas de trabajo registrado publicadas periódicamente por el MTEySS¹ y los datos de actividad económica sectorial publicados por el Instituto Nacional de Estadísticas y Censos (INDEC)², se

¹Ver <https://www.trabajo.gob.ar/estadisticas/index.asp>.

²Agregados macroeconómicos de las cuentas nacionales, <https://www.indec.gob.ar/indec/web/Nivel3-Tema-3-9>, y Esti-



Fuente: Elaboración propia en base a SIPA, World Bank WDI

Figura 3: Transiciones de empleo formal privado entre y dentro de industrias a 4 dígitos CIIU. Correlación (Pearson) con PIB real. En ambos casos la correlación es fuertemente positiva, mayor en las transiciones intrasectoriales aunque más voluminosas en el caso de las intersectoriales.

proveen usualmente a niveles de letra (v.g.: máximo nivel de agregación) y dos dígitos de desagregación que corresponde al detalle utilizado tradicionalmente para analizar la actividad económica sectorial. Por ejemplo, desde el punto de vista de las políticas públicas de empleo se pueden mencionar los informes mensuales de “Panorama del Trabajo Registrado³” y “Situación y Evolución del Trabajo Registrado⁴” que presentan información del empleo registrado a nivel de letra. Por otro lado, en cuanto a las publicaciones de investigación del empleo sectorial y su dinámica realizadas por el MTEySS, usualmente también refieren a estos niveles de agregación (Castillo et al. 2002, 2005). En particular, en (Castillo et al. 2002) los autores analizan la dinámica del empleo industrial para 1995-2000 mediante flujos agregados de creación y destrucción de empleo, combinando datos de dinámica de empleo y dinámica de las empresas a nivel sectorial de dos dígitos. La elección de publicar datos de empleo y su análisis a niveles agregados de la actividad económica resuelve el aspecto comunicacional, dado que se trata de conceptos conocidos por analistas económicos y políticos, como también en cierta medida por el público en general. No obstante, surge la necesidad de explorar enfoques y metodologías alternativos para explotar información más desagregada de los intercambios de empleo entre industrias y abordar con mayor riqueza preguntas de interés para el diseño de políticas públicas. En este sentido, la teoría de redes complejas y las metodologías asociadas a ella proponen enfoques alternativos que permiten analizar y resumir la información contenida en datos de gran complejidad como estos registros administrativos.

Para analizar los flujos de empleo interindustriales⁵ se utilizó como base información proveniente de las declaraciones juradas que las empresas presentan periódicamente al fisco⁶. Estos datos originan naturalmente una red temporal de interacciones bipartita entre trabajadores y empresas. No obstante, a fines de

mador mensual de actividad económica (EMAE), <https://www.indec.gob.ar/indec/web/Nivel4-Tema-3-9-48>.

³Ver https://www.trabajo.gob.ar/downloads/estadisticas/panorama/panorama_del_trabajo_registrado_2110.pdf el correspondiente a Octubre de 2021.

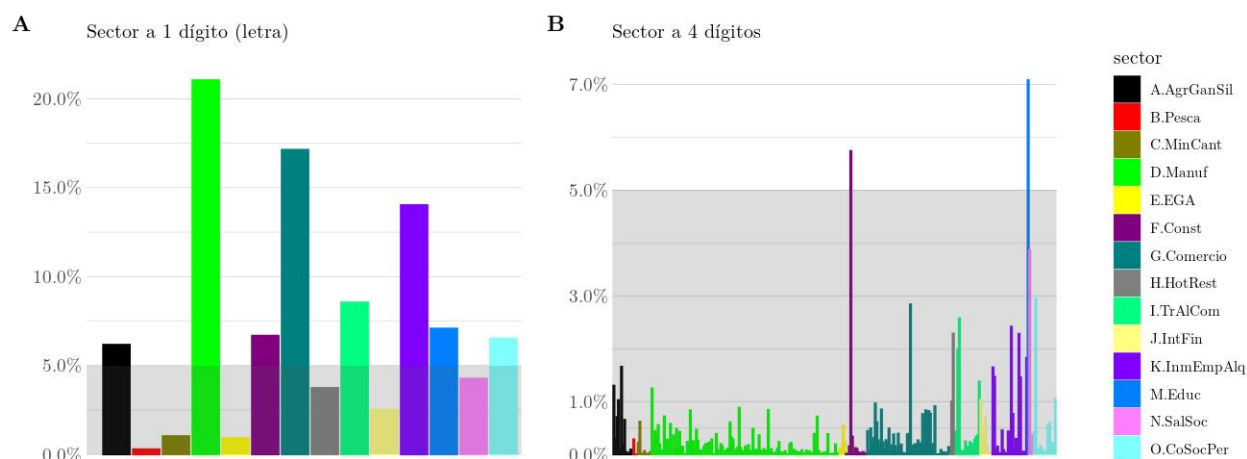
⁴Ver https://www.trabajo.gob.ar/downloads/estadisticas/trabajoregistrado/trabajoregistrado_2107_informe.pdf el correspondiente a Julio de 2021.

⁵En esta tesis se utiliza el término “interindustrial” en sentido amplio, como sinónimo de “intersectorial,” al nivel de detalle que corresponda.

⁶Nóminas salariales presentadas mensualmente a la Administración Federal de Ingresos Públicos a través del Sistema Integrado de Pensiones Argentino (SIPA), formularios 931/921/941.

preservar la privacidad de los datos, se utilizó como unidad de medida la cantidad de personas que cambiaron interanualmente de empleo en términos de las actividades productivas declaradas por sus empleadores. De esta forma, es posible extraer la proyección (unimodal) de intercambios de empleo entre las empresas. Así, estos datos pueden analizarse mediante la teoría grafos y el formalismo de redes complejas de interacciones a fin de extraer información acerca de la estructura topológica y las características particulares de los procesos de transiciones de empleo que las sustentan.

Como se mencionó, en economía se analiza usualmente información de actividad económica sectorial a nivel de letra y de dos dígitos, que proveen información comprensible en términos de conceptos tradicionales, que resultan familiares tanto a analistas y tomadores de decisiones de políticas públicas. No obstante, en ocasiones estos niveles de detalle limitan la riqueza del análisis por lo que para profundizar se recurre a un análisis restringido de algún subconjunto de sectores productivos de interés con datos más desagregados. Analizar los datos sectoriales con máximo nivel de detalle, en este caso a cuatro dígitos, presenta un desafío en cuanto a la interpretación de los sectores productivos, así como de la relevancia económica y política de la información a este nivel. Por ejemplo, una variable clave como el nivel de empleo que absorbe cada sector de actividad a nivel de letra presenta una imagen muy sintética y fácilmente comprensible, mientras que la que surge cuando se consideran sectores a cuatro dígitos presenta inmediatas dificultades para su comprensión a cualquier analista. Como se observa en la Fig. 4, la letra D que representa a la industria manufacturera concentra el 21 % del empleo promedio en el período 2003-2010 (panel A), mientras que al observar la distribución a 4 dígitos (panel B) la misma letra se subdivide en 122 sectores que acumulan entre 0.002 % y 1.24 % del empleo promedio del mismo período.



Fuente: Elaboración propia en base a SIPA

Figura 4: Empleo promedio 2003-2010 por sectores de actividad a nivel de letra y a cuatro dígitos CIIU. Participación relativa de sectores agregados a nivel de letra (A) y de sectores a cuatro dígitos (B). En gris se destaca la participación de hasta el 5 % del total. En la clasificación de industria manufacturera, letra D, se observa el mayor contraste: a nivel agregado capta el 21 %, mientras que a cuatro dígitos se compone de 122 sectores que captan entre 0.002 % y 1.24 % del empleo formal privado total.

Preguntas de investigación. A partir de estas observaciones se plantean las siguientes preguntas de investigación. ¿Qué características estructurales revelan los intercambios de empleo entre industrias? ¿Cuán estables son estas características en el tiempo? ¿Es posible identificar agrupamientos relevantes de estos flujos entre industrias? ¿Qué información se puede extraer de ellos respecto al entramado productivo? Dado que el fenómeno de los flujos de empleo analizado desde las relaciones entre sectores puede conceptualizarse naturalmente como un grafo, ¿pueden las metodologías de redes complejas aportar marcos de análisis para extraer nuevo conocimiento de los datos de registro?

Expectativa de resultados. Desde el punto de vista del análisis económico, se espera que los resultados de este estudio aporten a la mejor comprensión de aspectos económicos y productivos de los flujos

interindustriales de empleo. En particular, dado que las transiciones observadas entre empleos surgen bajo una configuración particular del sistema productivo, se esperaría encontrar aspectos estructurales relevantes al analizar estos flujos. Otra parte de estos podría explicarse por factores coyunturales, atribuibles a la fase corriente del ciclo económico. Por otro lado, se esperaría encontrar agrupaciones naturales entre las actividades productivas mediadas por los intercambios de empleo. Sin embargo, al ampliar el horizonte temporal se esperaría que las estructuras de interrelaciones obtenidas en períodos más lejanos evidencien diferencias con las más actuales en virtud de los efectos de cambios más permanentes, tanto en las relaciones de trabajo como en la organización de la estructura productiva. Es decir que se esperaría cierta evolución de estas estructuras de interrelaciones.

Por otro lado, desde el punto de vista de la minería de datos, se espera que las técnicas aquí utilizadas para la extracción de conocimiento de los datos de registro de empleo a máximo nivel de detalle permitan sentar las bases para un marco de referencia orientado al análisis de la conectividad interindustrial.

Por último, desde el punto de vista de la aplicación de las metodologías utilizadas, se espera aportar una herramienta útil para el diseño y el soporte de políticas públicas. Las metodologías para analizar redes complejas de interacciones, de bases interdisciplinarias y en continuo desarrollo, resultan apropiadas para la descripción y modelización de fenómenos complejos desde diversos puntos de vista y bajo distintas hipótesis de análisis. Las interacciones resumidas en los cambios de empleo presentan características de redes complejas, con topologías no triviales que representan el resultado de procesos dinámicos no lineales de búsqueda y matching entre trabajadores y empleadores, de relaciones de reciprocidad y de agrupamientos de sectores con mayor densidad de lazos entre sí. En base a estas características es que se considera que el formalismo de redes complejas resulta aplicable para abordar las preguntas de investigación planteadas.

Revisión de literatura. En las últimas décadas se desarrollaron enfoques interdisciplinarios que explotaron la aplicación de metodologías de redes complejas para abordar el análisis de diversos fenómenos económicos, con énfasis en la extracción de conocimiento a partir de datos de distintas fuentes, de gran volumen y complejidad. (Hidalgo and Hausmann 2009) examinaron las economías nacionales interpretando los datos de comercio exterior como una red bipartita en la que los países están conectados a los productos exportados y cuantificaron la complejidad de esa economía caracterizando la estructura de la red resultante. Este “espacio de productos” formaliza la idea de asociación entre productos intercambiados en la economía global y cuantifica la relación de productos con una medida llamada proximidad. Las distancias entre los productos que surgen en este espacio sirven como un indicador relevante para predecir el potencial de diversificación comercial de los países, también utilizado para cuantificar y mapear similitudes del nivel de desarrollo de los países.

En cuanto al análisis de la movilidad laboral, en lo que se refiere como “enfoque de flujos del mercado de trabajo,” (Gianelle 2010, 2014)) analizó la movilidad del empleo entre firmas en base a datos longitudinales de registros administrativos en regiones de Italia a nivel de empleadores (firmas) y trabajadores con métodos de redes complejas. En su análisis el autor encontró que estas redes de del empleo entre firmas presentaban propiedades de **mundo pequeño**, dominadas esencialmente por clustering local y conformando un sistema interconectado, en que algunas pocas firmas (grandes y de renombre) funcionaban como importantes redireccionadores de empleo (v.g.: *hubs*) acortando caminos entre distintas regiones de la red de firmas.

También en términos del análisis de las relaciones de intercambios de empleo entre firmas, (Guerrero and Axtell 2013) y (López, Guerrero, and Axtell 2015) desarrollaron el enfoque de *labor flow network*, que consiste en la modelización de la movilidad laboral (en sentido amplio, como el proceso descrito en el diagrama de flujos de la Fig. 1.A) a nivel de firmas y trabajadores considerando matrices de transición basadas en un modelo de procesos aleatorios sobre grafos en los que los nodos de la red son firmas y los lazos representan las transiciones de empleo observadas. En su análisis calcularon equilibrios teóricos de estado estacionario y compararon sus resultados con evidencia estadística de datos administrativos de México y Finlandia, para los que corroboraron la consistencia de los supuestos con la realidad observada. Los autores obtienen grafos estáticos (estacionarios, que reflejan la persistencia de la movilidad entre firmas observada) de transiciones de empleo para aportar un enfoque alternativo al modelo tradicional (agregado) de búsqueda y coincidencia de empleo (*job search and matching*) basado en datos de alta resolución (López, Guerrero, and Axtell 2015) y para la identificación de empresas con alto potencial de crecimiento (Guerrero and Axtell 2013).

Recientemente, (Neffke, Otto, and Weyh 2017) desarrollaron un marco de referencia para analizar flujos de empleo basados en datos administrativos e introdujeron el indicador de parentesco de habilidades (*skill-relatedness*), que proporciona información sobre las habilidades intercambiadas entre las industrias a través de los flujos de empleo. Este indicador cuantifica las desviaciones de los flujos observados con respecto a los flujos esperados según un criterio de aleatoriedad (modelo nulo) y, a partir del mismo, es posible extraer una red de interconexiones entre las distintas industrias denominada “espacio industrial.” Previamente, (Neffke and Henning 2013) construyeron el espacio industrial y encontraron que este indicador permite extraer una estructura relevante de conexiones interindustriales para el mercado laboral de Alemania, que difiere de la estructura jerárquica de los sistemas estándar de clasificación de actividades industriales y resultó significativa en la predicción de la diversificación de actividades de las empresas.

Más recientemente, (O’Clery, Flaherty, and Kinsella 2021) utilizando datos de transiciones de empleo de registros administrativos para Irlanda analizaron la estructura de relaciones interindustriales de las redes de flujos de empleo en términos de redes de parentesco de habilidades (*skill-relatedness networks*) y encontraron una estructura modular en la que identificaron agrupamientos de industrias en diferentes escalas de aglomeración. A partir de ellas construyeron una variable de “empleo agrupado” (*cluster employment*) que captura la fuerza de trabajo disponible para las industrias del agrupamiento a las que pertenecen, en términos de “cuencas de habilidades” (*skill basins*). Sin embargo, aunque para el análisis utilizaron datos de los intercambios entre industrias a cuatro dígitos (CIIU, o ISIC por sus siglas en inglés), el análisis de agrupamientos se realizó a dos dígitos.

Por otro lado, en la misma línea de investigación, (Straulino, Landman, and O’Clery 2021) desarrollaron un método de comparación de estructuras modulares entre redes de nodos alineados. Los autores comparan la estructura modular de las redes de parentesco de habilidades derivadas de registros administrativos a 4 dígitos ISIC de transiciones de empleo entre industrias para Irlanda, Alemania, Holanda y Suecia, junto con Reino Unido, para el cual utilizan como soporte encuestas longitudinales. En su trabajo concluyen que las estructuras modulares de estas redes presentan similitudes entre los países nórdicos, lo que indicaría que la organización de industrias en términos de *skill-relatedness* es parecida en estos países, aunque se diferencian de las del Reino Unido e Irlanda, que presentan a su vez cada uno meso-estructuras disímiles.

Para el caso de Argentina, (De Racó and Semeshenko 2019a) exploraron las redes de flujos de empleo y del indicador de parentesco de habilidades utilizando datos de transiciones de empleo formal privado 2009-2014 (SIPA, CIIU Rev4 a 4 dígitos), y luego construyeron el espacio industrial asociado. En su trabajo encontraron que las redes de flujos de empleo presentaban un único componente conexo denso, de diámetro reducido, con estructura de centro-periferia (*core-periphery*). En los grafos asociados a las redes de flujos se identificaron actividades productivas de la misma clasificación de letra (sección o gran división) organizadas en regiones de la red (i.e.: actividades primarias, servicios públicos, construcción, información y telecomunicaciones, entre otras), mientras otras se expandían a través de la misma (i.e.: industria, transporte, actividades profesionales, científicas y técnicas, actividades administrativas y otras actividades de servicios). Por otro lado, las redes de parentesco de habilidades y el espacio industrial presentaron una estructura organizada en encadenamientos productivos parciales.

En base al mismo conjunto de datos, (De Racó, Tumini, and Ernst 2019) calcularon el indicador de parentesco de habilidades y analizaron agrupamientos jerárquicos de las actividades productivas. En su análisis encontraron agrupamientos caracterizados por encadenamientos productivos parciales conectados entre sí e interpretables en términos de covariantes como niveles de formación, participación femenina y otros atributos económicos y sociales asociados a la movilidad laboral.

En (De Racó and Semeshenko 2019b), los autores comparan los espacios industriales de Argentina y Alemania, construido a partir de matrices de parentesco de habilidades publicadas por Neffke correspondientes a un período (2007-2014) y clasificación industrial (NACE Rev2 a 4 dígitos) similares. En sus resultados reportaron similitudes y diferencias en términos de las redes de parentesco de habilidades entre los países. Ambos países presentaron un único gran componente conexo, densidad y coeficiente de clustering elevados, largo promedio de caminos muy corto y diámetro pequeño. En términos de conectividad ambas presentaron estructuras de mundo pequeño y de centro-periferia. En el caso de Alemania la red contenía más nodos y lazos y resultaba más densa que la obtenida para la Argentina. En términos del espacio industrial, Alemania presentaba un grupo de mayor conectividad interna (máximo *k*-coreness) compuesto por sectores de

manufacturas de alto valor agregado y maquinaria industrial que concentraban una pequeña proporción del empleo total, mientras que en Argentina el grupo de sectores productivos de mayor conectividad interna resultaba muy heterogéneo y captaba más del 70 % del empleo total.

Dado lo expuesto, esta tesis se orienta a profundizar la temática y aportar en la identificación y evolución de las estructuras de los flujos de empleo entre industrias, en particular de aquellas que surgen de registros administrativos a gran nivel de detalle. A continuación se exponen los objetivos generales y específicos de trabajo.

1.1. Objetivos

El objetivo general del trabajo es la extracción de información de datos de flujos de empleo provenientes de registros administrativos, examinados a partir de un nivel de detalle sustancialmente mayor al utilizado habitualmente en estudios académicos y evaluaciones de políticas públicas. Se propone analizar la estructura de interacciones que surge de datos de transiciones de empleo formal privado de la Argentina en períodos seleccionados entre 1996 y 2020 en términos de las matrices de flujos interindustriales interanuales y promedio, y se articula en dos objetivos específicos (OE):

- **OE 1:** Se plantea analizar las características estadísticas y topológicas de las redes generadas para luego evaluar la estabilidad de las mismas a lo largo del tiempo.
- **OE 2:** Se busca identificar agrupamientos sectoriales que permitan caracterizar intercambios por criterios de similitud.

La tesis se compone de distintos Capítulos que abordan el análisis de las redes de flujos de empleo interindustrial y la detección de comunidades en un subconjunto de ellas, que en función de ello se organiza de la siguiente manera. En el Capítulo 2 se describen los materiales y los métodos utilizados. En el Capítulo 3 se analizan las redes de flujos de empleo y en el Capítulo 4 se presentan resultados de detección de comunidades y agrupamientos en las distintas redes. Finalmente, en el Capítulo 5 se presentan las conclusiones y perspectivas de trabajo futuro.

2. Materiales y métodos

En este Capítulo se presentan las acciones requeridas para atender los objetivos propuestos. Como primer paso se procesaron los datos y se construyeron las redes asociadas. Se utilizaron datos de flujos de empleo interindustriales para construir matrices de transición interanuales que resumen los intercambios entre sectores productivos. Por otro lado, para el análisis comparativo de su estructura en un horizonte temporal de mediano plazo, se agruparon flujos de empleo interanuales en períodos representativos según criterios que se definen en el Capítulo 3. Las matrices de transición constituyen los *lazos* (o *aristas*) y representan redes dirigidas y ponderadas de interacciones. A partir de estas redes se calculan medidas estadísticas y se extraen un conjunto de características topológicas a fin de compararlas temporalmente. Adicionalmente, se utilizan datos de dotación de empleo como atributos que permitan dimensionar los sectores (*nodos* ó *vértices*, en terminología de grafos) de las redes construidas. De forma complementaria se utilizan datos de la actividad económica agregada para cotejar la correlación existente entre ésta y los flujos de empleo analizados.

En segundo lugar, para identificar agrupamientos sectoriales se utiliza el método multinivel Leiden (Traag, Waltman, and Eck 2019), de detección de comunidades en redes complejas, basado en la optimización de la modularidad en etapas sucesivas.

2.1. Datos

El Ministerio de Trabajo, Empleo y Seguridad Social (MTEySS), como administrador del Sistema Previsional Argentino (SIPA), dispone de datos de empleo asalariado del sector privado provenientes de declaraciones juradas presentadas por los empleadores mensualmente a la AFIP con la nómina salarial. Estos registros administrativos, adaptados para su uso estadístico, expresan la relación entre el empleador y su nómina de empleados y contienen información de características económicas y sociales de ambos actores. La metodología usual de construcción de las matrices de transición de empleo utilizada por el MTEySS parte de la creación de paneles interanuales de individuos que aparecen en ambos o alguno de los períodos de registro para un mes de referencia. Así, se contabiliza al conjunto unión de trabajadores declarados en los períodos t y $t+1$ para el cual se identifican cambios de estado en su situación respecto del empleo formal, es decir: ingresos/egresos a/de la formalidad (v.g.: entre empleo y desempleo o inactividad), permanencia en el mismo empleo y transición a otro empleo. En este trabajo se utilizan sólo los flujos de transiciones interindustriales del empleo privado, que representan individuos que aparecen empleados en t y que en $t+1$ aparecen empleados en otra empresa perteneciente a un sector de actividad distinto⁷.

Los datos utilizados fueron proporcionados por el Observatorio de Dinámica Empresarial y del Empleo (OEDE) del MTEySS⁸ y comprenden intercambios interanuales de empleo entre actividades productivas registrados entre 1996 y 2020, es decir $T = 24$ períodos con la última matriz de transición observada entre 2019 y 2020, año en que sobreviene la pandemia COVID-19.⁹ El conjunto de actividades productivas incluye cerca de 300 sectores o ramas de actividad (ver Anexo) que se registran a cuatro dígitos de detalle, organizadas en 14 secciones (letras) del clasificador CIIU Rev 3. La cantidad de sectores captados anualmente a este nivel de detalle no es necesariamente fija. Puede variar dado que la baja cantidad de empleo que concentran algunos sectores en cabeza de pocas empresas que en algunos períodos desaparecen del registro. Adicionalmente, con el objeto de dimensionar el tamaño de cada sector de actividad, se utilizaron datos de la dotación sectorial media anual a cuatro dígitos del clasificador CIIU Rev 3 disponibles al público en el sitio de la OEDE¹⁰. En el Cuadro 1 se presenta un resumen de la clasificación de actividades utilizada según surgió de los datos de flujos de empleo disponibles, cuya nómina completa se presenta en el Anexo metodológico.

A los fines del análisis de cambios estructurales, se seleccionaron períodos relevantes a comparar para los cuales se construyeron matrices de transiciones promedio según se mencionó anteriormente. Se utilizaron datos de PBI real del repositorio de Indicadores Mundiales de Desarrollo del Banco Mundial¹¹ para evaluar la correlación de los flujos de empleo privado con la actividad económica (Fig. 3).

Exclusiones y filtrado de datos. Para el análisis comparativo de matrices de transiciones se utilizan métodos de comparación de redes con nodos alineados, es decir aplicables a redes compuestas con los mismos nodos, dado que éstos surgen de un sistema de clasificación estándar de las actividades económicas mediante el que es posible asociar a cada nodo una identidad particular. En el preprocesamiento de los datos se excluyeron algunos sectores de modo que resultara posible obtener matrices de transición de igual dimensión para todos los períodos analizados. Así, se excluyeron sectores a cuatro dígitos en base a los siguientes criterios de relevancia y pertinencia: 1) sector público y Banca Central, dado que el análisis se enfoca en transiciones de empleo dentro del sector privado; 2) agencias de empleo transitorio, debido a que su rol de intermediario en relaciones formales de empleo no aporta habilidades adquiridas en las personas que temporariamente se declaran empleadas en dicho sector, siguiendo a Giannelle (2010); 3) sectores para los que no se publica el nivel de empleo por razones de secreto estadístico; 4) sectores de bajo nivel de empleo

⁷Adicionalmente el MTEySS realiza controles adicionales en la creación de estos paneles, por ejemplo utilizando el empleo de mayor remuneración en el caso de pluriempleo y controlando por migraciones masivas a empresas nuevas.

⁸Datos producidos en el marco de una cooperación OIT-MTEySS y miembros del Instituto Interdisciplinario de Economía Política (IIEP) para el estudio de transiciones de empleo.

⁹Respecto de los efectos del último año incluido en el período de análisis, cabe la aclaración de que los datos corresponden a empleo formal privado, que en términos de transiciones dentro del empleo 2020 resultó un año condicionado por las políticas de protección del empleo llevadas a cabo ante el desplome de la actividad devenido a su vez por las medidas de aislamiento obligatorio, con su efecto directo en la fuerte reducción de los cambios en el empleo. La evaluación completa del cambio estructural en la movilidad laboral posterior a la pandemia COVID-19 requerirá de nuevas observaciones para cotejar la adaptación del sistema al funcionamiento lejos del período de crisis.

¹⁰<https://www.trabajo.gob.ar/estadisticas/oede/estadisticasnacionales.asp>.

¹¹<https://datos.bancomundial.org/indicador/NY.GDP.MKTP.KN?locations=AR>.

Cuadro 1: Actividades a cuatro dígitos CIU Rev 3 incluidas en el análisis, por letra y empleo promedio por período.

Letra	Descripción	# Activ.	Empleo promedio			
			1996-2002	2003-2010	2011-2017	2018-2020
A	Agricultura, ganadería, caza y silvicultura	13	247.718	310.613	346.464	351.984
B	Pesca	3	10.010	14.918	13.223	12.596
C	Explotación de minas y canteras	9	33.668	52.014	71.362	75.354
D	Industrias manufactureras	122	874.791	1.059.175	1.275.338	1.167.951
E	Suministro de electricidad, gas y agua	4	42.202	46.450	60.489	67.644
F	Construcción	14	219.815	336.302	441.707	426.335
G	Comercio al por mayor y al por menor; reparación de vehículos automotores, motocicletas, efectos personales y enseres domésticos	52	601.037	862.241	1.148.282	1.156.582
H	Hoteles y restaurantes	4	117.234	188.455	271.521	281.727
I	Transporte, almacenamiento y comunicaciones	16	338.864	430.264	564.229	565.944
J	Intermediación financiera	8	132.412	126.721	148.193	145.958
K	Actividades inmobiliarias, empresariales y de alquiler	23	498.275	705.987	855.625	841.852
M	Enseñanza	1	248.973	356.464	460.485	505.934
N	Servicios sociales y de salud	3	175.128	214.818	306.208	347.091
O	Otras actividades de servicios comunitarios, sociales y personales	15	254.192	328.283	399.185	401.811
Total		287	3.794.317	5.032.703	6.362.310	6.348.761

sin transiciones de ingreso o de salida en alguno de los períodos utilizados, en razón de que el indicador de parentesco de habilidades utilizado requiere subtotales positivos para su cálculo; y, 5) sector no especificado por el declarante (v.g.: código 0), que consiste en una categoría residual poco significativa y marcadamente decreciente en volumen durante el período de análisis. En el Cuadro 23 en Anexo se detallan los sectores excluidos del análisis con datos del nivel medio de empleo en los casos que se encontró disponible.

Matrices de transición. A partir de los paneles interanuales de transiciones de empleo entre sectores se construyen $t = 24$ matrices de transición anuales, F_t , con $t = \{1996, \dots, 2020\}$, y matrices promedio para los $p = 4$ períodos seleccionados, F_p , con $p = \{9602, 0310, 1117, 1820\}$. Luego, se construyen las redes correspondientes de flujos de empleo, F_p . Sobre estas redes se calculan estadísticas descriptivas y se analizan los grafos y mapas de calor de las matrices de adyacencias binarias y/o ponderadas, según el caso.

Para el análisis temporal se comparan las t matrices de adyacencias interanuales y las matrices de adyacencias de los p períodos mediante métodos para nodos alineados, es decir matrices con las mismas entidades e igual dimensión.

Por último, se aplica el algoritmo de Leiden de detección de comunidades para el análisis de agrupamientos de sectores.

Software utilizado. El procesamiento y el análisis de datos de esta tesis se realizó íntegramente en el lenguaje R (R Core Team, 2021) con la familia de paquetes de **tidyverse** (Wickham et al. 2019) en el entorno **RStudio** (RStudio Team 2020). Para la construcción, procesamiento, análisis y visualización de las redes en los Capítulos 3 y 4 se utilizó **igraph** (Csardi and Nepusz 2006) y **skynet** para el algoritmo de disparity filter, y extensiones como **ggraph**, **tidygraph** y **graphlayouts**. Complementariamente, para las visualizaciones se usaron los paquetes **RColorBrewer**, **viridis**, **hrbrthemes** y **cowplot**. Para la detección de comunidades en el Capítulo 4 se utilizó el paquete **leidenAlg**, que implementa el algoritmo desarrollado por los autores Traag et al. (2017) en **python**, junto con **ggalluvial** para la visualización de la evolución de los grupos. Para la escritura y procesamiento del documento, realizado íntegramente en **Rmarkdown** con el paquete **knitr**, se utilizó el paquete **bookdown** con plantillas y macros adaptadas del paquete **thesisdown** que facilitaron la exportación y post-procesamiento del documento final en **L^AT_EX**. Por último, en la producción y composición

de imágenes resultaron de utilidad los paquetes **DiagrammeR** y **FactoMineR**.

Tanto el documento final como los datos¹² y el código fuente se publicarán en el repositorio <https://github.com/sderaco/tesism> con la versión final posterior a su defensa.

2.2. Métodos

Las redes de transiciones interindustriales de empleo que se analizan en este trabajo excluyen los cambios de empleo dentro de la misma actividad industrial, es decir que se consideran todos los flujos fuera de la diagonal principal de las matrices de transiciones de referencia, que representan movimientos interanuales de salida de e ingreso a la misma actividad productiva. En términos de teoría de grafos equivale a decir que se estudian los grafos sin bucles (*loops*), que se analizan como grafos dirigidos o no dirigidos, binarios o pesados, según se indique en el análisis, asociados a las matrices de flujos de empleo que presenta en esta sección.

Representaciones de redes. Las redes pueden representarse de diversas formas, como se presenta en la Fig. 5, según los aspectos que se desea analizar o resaltar. En su formalización más convencional mediante teoría de grafos, una red se representa mediante un grafo $G = (V, E)$ (Fig. 5.D), que relaciona entidades designadas en sus vértices o nodos (i.e.: sectores económicos en este caso), V , mediante aristas o lazos, E (Fig. 5.B, que bien pueden tratarse de la salida de una consulta a una base de datos como la descrita en la sección anterior).

En este estudio, los grafos representan las redes de interacciones entre distintas actividades productivas (vértices o nodos) tomadas de a pares mediadas por cambios de empleo formal privado (aristas o lazos). La magnitud de los flujos representa los pesos o ponderaciones en cada lazo. En este sentido, los grafos asociados a estas redes son simples (i.e.: sin bucles, se excluyen las relaciones dentro del mismo sector), dirigidos y ponderados.

Asociado a G hay un conjunto de **matrices fundamentales** que sirven de representaciones alternativas. La **matriz de adyacencias**, A , indica en los valores de cada celda las interacciones entre las entidades representadas en sus filas y columnas (Fig. 5.A). Si sólo se requiere representar la existencia (binaria) de conexiones entre nodos, A representa al grafo *no dirigido* y *binario* (con $A_{ij} \in \{0, 1\}, \forall_{i,j}$) de la red original y resulta una matriz cuadrada, no negativa y simétrica tal que esta matriz es igual a su transpuesta, $A = A^T$, en la cual la suma de una fila (o columna) i representa el *grado*, k_i , de conectividad del nodo i -ésimo, mientras que la suma de sus elementos $\sum_{i,j} A_{ij}$ es el doble de las conexiones presentes en la red. Esta matriz mostrará distintas propiedades según el tipo de interacciones que represente. Si la red que representa fuera *ponderada*, como es el caso de las matrices de transiciones observadas que se analizan, la suma de sus elementos puede interpretarse como el peso (*weight*) de las interacciones observadas, a la que es posible asociar una interpretación física en relación al fenómeno que describe. En el caso de un grafo *dirigido*, las interacciones tienen una dirección particular de las relaciones entre dos nodos que puede resultar no correspondida (no recíproca), con lo que A puede resultar asimétrica¹³.

Utilizando matrices fundamentales es posible obtener una visualización bidimensional mediante **mapas de calor** (*heatmaps*, Fig. 5.C), que permite distinguir rápidamente la organización de la densidad de interacciones de la red. Comúnmente estos mapas se utilizan en conjunto con algoritmos de agrupamientos jerárquicos representados por dendrogramas (Fig. 5.C2), que reorganizan las matrices en grupos de nodos que interactúan con mayor intensidad entre sí (en bloques), grupos de nodos que interactúan entre sí y con toda la red (centro, *core* o núcleo), o grupos que sólo interactúan con subconjuntos de nodos con mayor conectividad (periferia), por mencionar algunas estructuras comunes.

Así, las matrices de transiciones de empleo observadas (Fig. 5.A) se corresponden directamente con la representación de la matriz de adyacencias de un grafo sin bucles dirigido y ponderado, para el cual es posible

¹²Sujeto al acuerdo de aprobación para la publicación de los mismos por parte del OEDE.

¹³Existen otras matrices fundamentales asociadas al grafo que resultan particularmente útiles para el análisis de propiedades globales de la red, como el Laplaciano, $L = D - A$ en que D es la matriz cuya diagonal contiene el grado k_i correspondiente a cada nodo i , y el Laplaciano normalizado por la matriz de grados, $L_n = I - D^{-1/2} A D^{-1/2}$. No obstante, en este trabajo sólo se analizan propiedades extraídas de las matrices de adyacencias.

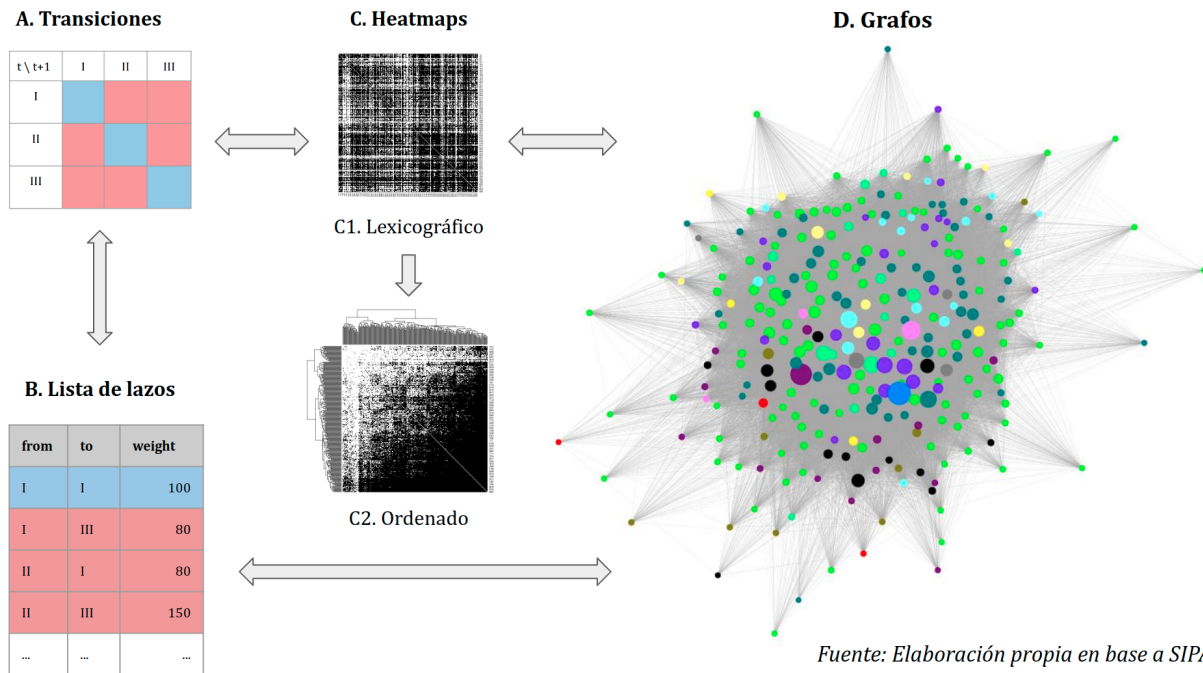


Figura 5: Representaciones alternativas de redes de transiciones de empleo. Matrices de transiciones (A), lista de lazos (B), mapas de calor (C1, lexicográfico, por orden alfabético, y C2, con ordenamiento jerárquico), y grafos (D). A y B son esquemas didácticos de las representaciones correspondientes. C y D corresponden a transiciones promedio de empleo formal privado, para las que se utilizaron como ejemplo datos de 2003-2010 (SIPA). La visualización D está realizada con layout de fuerzas dirigidas LGL y presenta la red como un gran componente conexo denso con forma de "bola de pelos" (hairball), donde los nodos son sectores a 4 dígitos CIIU Rev 3, su tamaño representa la dotación media del período, los colores representan la letra o gran división del sector y los lazos representan flujos a escala.

alternar su representación como una lista de lazos (Fig. 5.B), mapas de calor (Fig. 5.C) o mediante la visualización de un grafo (Fig. 5.D).

Estadística descriptiva. Para las redes generadas se calculan estadísticas descriptivas *globales*, que consideran a los grafos en su totalidad, y métricas *locales*, relativas a cada nodo y su entorno o relación con el resto de la red, que se detallan en el Anexo metodológico.

Dentro de las métricas globales se obtiene la cantidad de *componentes conexas*, es decir los subconjuntos de nodos interconectados por transiciones de empleo, que como se verá más adelante consiste usualmente en un único componente conexo. Un aspecto relevante de la red refiere a su *densidad*, es decir la cantidad de lazos observados respecto de los posibles para el tamaño de la red. Una propiedad básica ya mencionada de cada nodo i -ésimo en una red es su *grado* de conectividad (*degree*) k_i , que mide la cantidad de lazos adyacentes al mismo. En una red dirigida cada nodo tendrá dos tipos de grados, uno de entrada k_i^{in} y otro de salida k_i^{out} . La *distribución de grados* es una medida global de la estructura de conectividad de una red. El *coeficiente de asortatividad* mide la correlación entre pares de nodos conectados según el grado de conectividad de cada uno. Esta métrica permite distinguir si en la red los nodos se relacionan entre sí con nodos de grado similar, en cuyo caso la red resulta asortativa evidenciando la propiedad de homofilia típica en redes sociales, o los nodos de mayor conectividad (grado) se relacionan con nodos de baja conectividad, en cuyo caso la red es disasortativa, y resulta útil para evaluar la tipología de la red en términos globales. En términos de las relaciones entre pares de nodos, en una red direccionada (o dirigida) como las de flujos de empleo observadas resulta relevante la información acerca de la *reciprocidad* de los lazos entre dos nodos (v.g.: díadas), es decir en qué medida ocurre que observando transiciones de V_i hacia V_j también ocurren transiciones de V_j hacia V_i , y el *coeficiente de clustering*, que resume información de los lazos entre tres nodos (v.g.: tríadas). Cuanto mayor la reciprocidad observada, resultará mejor la aproximación del grafo dirigido mediante un grafo no dirigido, lo cual simplifica la aproximación analítica a esa red y permite la aplicación de un mayor conjunto de técnicas a su estudio. Respecto al coeficiente de clustering, cuanto mayor resulte conformará más triángulos en las tríadas observadas, fortaleciendo la estructura de conectividad necesaria para el agrupamiento de los nodos. El coeficiente *local* de clustering se calcula para cada nodo y mientras que el coeficiente *global* de clustering representa el cociente de las tríadas cerradas respecto de las tríadas posibles, que se utiliza como medida resumen para todo el grafo. Este coeficiente resulta relevante para identificar estructuras de mundo pequeño y detectar meso-estructuras, que se verán en el Capítulo 3.

En cuanto a las estadísticas asociadas a los nodos, resultan de interés las medidas de *centralidad* y *conectividad* de cada nodo para evaluar su posición y rol en la red. En este trabajo interesa caracterizar a los nodos (i.e.: sectores) en función de estas métricas para observar la estructura de relaciones entre ellos. Las métricas utilizadas se detallan en Anexo metodológico.

Para ambas clases de métricas se utiliza el método estadístico de *análisis de componentes principales* (PCA), de reducción de dimensionalidad, para extraer información comparativa de las distintas redes y nodos y mapearlas en los componentes que capturan la mayor variabilidad del conjunto de datos.

Análisis temporal. Dado que se dispone de la serie de tiempo de transiciones de empleo, el análisis temporal se puede realizar en distintas maneras según la pregunta que se quiera responder. En esta tesis la comparación de las matrices de transición de distintos períodos (anuales y de períodos agregados) se realiza en dos etapas.

En primer lugar, para evaluar la persistencia de las estructuras de flujos de empleo interindustriales se calcula el coeficiente de correlación de Spearman sobre las matrices de pesos vectorizadas. Un método utilizado para comparar matrices de igual dimensión consiste en crear una nueva matriz cuyas columnas son las matrices a comparar vectorizadas, es decir la concatenación de las filas en un vector columna, para luego calcular la correlación de Pearson entre estas columnas (Hidalgo and Hausmann 2009; Neffke and Henning 2013). Este método se utiliza más adelante en el Capítulo 3 para la comparación entre matrices interanuales y entre matrices de períodos agregados.

En segundo lugar, para evaluar diferencias estructurales a distintos niveles de generalidad de los grafos, se utilizan las distancias presentadas a continuación.

Comparación de redes. La comparación de redes es una tarea cuyo objetivo es analizar las diferencias y similitudes entre las estructuras de interacción que representan, y atraviesa múltiples áreas de aplicación

en un amplio espectro de las ciencias. Existen diversos métodos para comparar redes y, dependiendo del tipo de interacciones que se desea analizar, pueden utilizarse varios de ellos en forma complementaria para lograr una mejor comprensión de los fenómenos que se desean estudiar. En la teoría de grafos este problema se refiere al isomorfismo de grafos, es decir a la correspondencia (biyectiva, uno-a-uno) entre dos grafos. Para ello, se utilizan medidas de distancia ó se calculan (dis)similitudes. En aplicaciones, a veces se requiere comparar redes de interacciones que provienen de distintos ámbitos (v.g.: tecnológicas, biológicas, financieras, etc.), mientras que en otras se busca comparar la evolución temporal de una misma red. En el caso de las interacciones manifestadas en los flujos de empleo interindustriales, interesa comparar los cambios ocurridos en el tiempo y analizar su estabilidad y evolución.

Estas técnicas pueden organizarse en términos del tipo y el nivel de generalidad de las características topológicas que enfatizan como base de comparación, que brindan información distinta de la estructura de las redes representadas. Algunas utilizan información de aspectos globales, considerando propiedades estructurales y topológicas del grafo completo, otras comparan características de la estructura local, es decir los entornos de conexión respecto de cada nodo, y otras usan estructuras intermedias, de agrupamientos de subconjuntos de nodos ó meso-estructura.

(Donnat and Holmes 2018) desarrollan un enfoque para comparar redes de nodos alineados en distintos momentos del tiempo. A partir de las matrices fundamentales de los grafos asociados a estas redes, utilizan distintas métricas y disimilitudes para su comparación mediante el análisis a diferentes escalas de sus estructuras topológicas: local, intermedia (meso) y global.¹⁴ Desde el punto de vista de la teoría de grafos, es posible analizar la estructura de las relaciones entre (pares de) nodos desde estas tres escalas o niveles de resolución local, meso-y global. La estructura local refiere a los entornos o vecindarios de primer orden de cada nodo y enfatiza el análisis del grado de conectividad y densidad de lazos que presenta cada uno, así como de los lazos que no ocurren (huecos estructurales). En cuanto a la estructura global, refiere al grafo en su totalidad, en término de las propiedades del sistema que representan los nodos y los lazos que lo componen, así como características de las dinámicas en base al análisis realizado sobre las matrices fundamentales (de adyacencias y Laplacianas). Por último, la existencia de meso-estructura a niveles intermedios, a veces referida como estructura “*local*”, amplía el análisis local a entornos o vecindario más extendidos, por ejemplo de segundo o tercer orden (i.e.: vecindarios de sus vecinos), se plantea como la posibilidad de identificar subconjuntos de nodos que compartan conectividad en grupos.

En esta tesis se utiliza parcialmente el enfoque propuesto por (Donnat and Holmes 2018) para comparar redes de flujos de empleo en el Capítulo 3. Las distancias utilizadas se calculan sobre las matrices de adyacencias binarias, $A_{ij} \in \{0, 1\}$, de los grafos no dirigidos asociados a las redes correspondientes, es decir de las matrices que indican la conectividad intersectorial de los flujos observados sin la magnitud de los pesos ni direccionalidad. Se seleccionan las distancias estructurales (locales), polinomiales (locales extendidas) y espectrales (globales), que se describen a continuación, para el análisis comparativo entre dos grafos G y $\tilde{G}$, mediante sus respectivas matrices de adyacencias binarias, A_{ij} y $\tilde{A}_{ij}$.

Dentro de las *distancias estructurales*, de la familia de las “distancias de edición,” se utiliza la distancia de Jaccard que permite identificar cambios en las redes a nivel de entornos locales de interconexión de cada sector y resulta apropiada para evaluar la tasa de reacomodamiento de lazos. Según (Donnat and Holmes 2018), el coeficiente de Jaccard puede interpretarse como la proporción de lazos eliminados o incorporados con respecto al total de lazos que aparecen en ambas redes, y si su valor es cercano a 1 indica un remodelamiento completo de la estructura del grafo entre dos períodos. En el Capítulo 3 se utiliza la distancia de Jaccard:

$$d_{Jacc}(G, \tilde{G}) = \frac{|G \cup \tilde{G}| - |G \cap \tilde{G}|}{|G \cup \tilde{G}|} = \frac{\sum_{i,j} |A_{ij} - \tilde{A}_{ij}|}{\sum_{i,j} \max(A_{ij}, \tilde{A}_{ij})}. \quad (1)$$

En cuanto a las *distancias polinomiales*, basadas en la utilización del polinomio de potencias de la matriz de adyacencias, A^k , que se relaciona con la topología local del grafo a través de los coeficientes A_{ij}^k que representan los caminos del nodo i al nodo j a través de k saltos o pasos. En este sentido, es posible utilizar

¹⁴(Wills and Meyer 2020) utilizan una aproximación similar en la comparación a distintas escalas de las estructuras topológicas de distintos grafos explorando métodos para grafos alineados y no alineados.

dicho polinomio para cuantificar las diferencias entre dos grafos considerando los efectos en vecindarios extendidos k -pasos. En términos de transiciones de empleo este permite incorporar al análisis otros eslabones de las cadenas de conectividad que cada sector tiene con otros mediados por los sectores con los que tiene intercambios directos en órdenes sucesivos. En este trabajo se utiliza la distancia polinomial con $k = 2$ debido a que el camino promedio más largo (diámetro) de las redes de flujos de empleo es en promedio de 3 pasos. Para calcular esta distancia se parte de la descomposición de valores propios de la matriz de adyacencias A de un grafo, $A = Q\Lambda_A Q^T$, donde Λ_A es la matriz diagonal con los valores propios de la matriz A y Q es la matriz cuadrada con los vectores propios asociados. Se propone una aproximación polinomial para la matriz de adyacencias, $P(A) = QWQ^T$, donde

$$P(x) = x + \frac{1}{(V-1)}x^2 + \dots + \frac{1}{(V-1)^{k-1}}x^k,$$

y

$$W = \Lambda_A + \frac{1}{(V-1)}\Lambda_A^2 + \dots + \frac{1}{(V-1)^{k-1}}\Lambda_A^k,$$

que para el caso de uso con $k = 2$ se expresa:

$$d_{Pol2}(G, \tilde{G}) = \frac{1}{V^2} \|P(A) - P(\tilde{A})\|_{2,2}. \quad (2)$$

Esta distancia permite cuantificar diferencias entre dos grafos de la estructura de conectividad que presenta cada sector con el vecindario de sectores con los que tiene intercambios directos y de la conectividad de estos con el resto de la red. Como se dijo, dado que el diámetro de la red es de tres pasos, con esta distancia se captarían cambios en la estructura cercanos a la escala global, de toda la red.

Respecto de las *distancias espectrales*, utilizadas para caracterizar el estado del grafo mediante sus valores propios, permiten detectar cambios en estructuras globales e intermedias. Por ejemplo, al efectuar el análisis espectral (*eigen-analysis*) sobre las matrices fundamentales se extrae información de la estabilidad de la dinámica del sistema complejo que representa y acerca de la suavidad de su evolución mediante el seguimiento de los cambios en los valores propios (autovalores o *eigenvalues*) de estas matrices, de las que a su vez es posible extraer una interpretación en términos físicos (matriz de adyacencias, A). La distancia espectral se define como la distancia l_p entre funciones del espectro de los valores propios $\lambda_0 \leq \lambda_1 \leq \dots \leq \lambda_{V-1}$ de las matrices de dos grafos y se expresa de la siguiente forma:

$$d_{Eig}(G, \tilde{G})^p = \|f(\lambda_i) - f(\tilde{\lambda}_i)\|_p = \left(\sum_{i=0}^{V-1} |f(\lambda_i) - f(\tilde{\lambda}_i)|^p \right)^{(1/p)}. \quad (3)$$

En este trabajo se utiliza la distancia l_p con $p = 2$, la distancia Euclídea, sobre el espectro de las matrices A de adyacencias de los grafos correspondientes. En particular, dado que en este caso interesa cuantificar el cambio en el espectro completo de frecuencias representado en los autovalores de las matrices de adyacencias de las redes a comparar, se utiliza la función identidad $f(\lambda) = \lambda$ para enfatizar la amplitud del cambio en los autovalores.

Para evaluar los resultados de las distancias introducidas en este acápite, en el Capítulo 3 se utilizan mapas de calor (Fig. 5.C) y el método estadístico de *escalamiento multidimensional métrico* (MDS), una técnica clásica de reducción de dimensionalidad no lineal similar a PCA, basada en la descomposición de valores propios de una matriz de distancias, que permite mapear la proximidad entre las distintas redes en los componentes que capturan la mayor variabilidad de la matriz de distancias analizada.

Agrupamientos y comunidades. Un aspecto relevante al analizar flujos de empleo es la necesidad de indagar acerca de grupos de industrias que intercambian empleo más asiduamente entre sí que con el resto de las industrias. En términos de redes esta tarea refiere a la exploración de la meso-estructura en la red bajo análisis mediante la identificación de agrupamientos y la detección de comunidades. En términos de la identificación de agrupamientos, algunas técnicas de inspección utilizadas hacen uso de los mapas de calor

ordenados por algún algoritmo de clustering, que permite identificar la existencia de estructuras modulares y de centro periferia, entre otras, en las matrices fundamentales que en su ordenamiento natural no son distinguibles (i.e.: orden lexicográfico, que en este caso refiere a la clasificación CIIU Rev 3 de los datos, ver Fig. 5.C1). Dentro de las familias de algoritmos de detección de comunidades, es posible clasificar los métodos de detección de comunidades en dos categorías, según se asignen los nodos a grupos disjuntos (particiones) o se permita asignar a los nodos a múltiples grupos de pertenencia (comunidades superpuestas). Por otro lado, es relevante distinguir si el análisis de agrupamientos se realiza sobre grafos dirigidos o no dirigidos, es decir si posee información de la direccionalidad de los lazos, dado que las técnicas correspondientes difieren como se reporta en (Fortunato 2010). En esta tesis se utiliza un algoritmo sobre las redes agrupadas de flujos dirigidos, F_p .

En primer lugar, se parte de las representaciones de mapas de calor de las matrices fundamentales ordenadas por dendrogramas de agrupamientos de los nodos mediante algoritmos jerárquicos aglomerativos, basados en grafos con el método de vinculación completo (máximo), que permite equiparar el concepto de agrupamiento con un subconjunto de nodos completamente conectados entre sí (subgrafo completo, o *clique*). Este análisis se realiza en el Capítulo 4 para los períodos seleccionados sobre las matrices de adyacencias A_p binarias y ponderadas, según el caso, para identificar meso-estructura.

En segundo lugar, para la detección de comunidades de nodos en grafos se utiliza el algoritmo multinivel Leiden (Traag, Waltman, and Eck 2019), un método de *partición* de los nodos, aplicado sobre grafos *dirigidos* de las redes originales de flujos. Este algoritmo mejora, extiende y refina el método de Louvain (Blondel et al. 2008) basado en la optimización en tres etapas de la modularidad, una medida de calidad que calcula la densidad de lazos dentro de las comunidades en relación con la densidad de lazos entre las comunidades. En particular, este algoritmo mejora a Louvain garantizando la conectividad de las comunidades detectadas. En el Capítulo 4 de esta tesis se utiliza este algoritmo para identificar comunidades en las redes originales de flujos, F_p .

Luego, en el mismo Capítulo 4 se analiza su persistencia en el tiempo, para lo cual se comparan las pertenencias de los sectores a los agrupamientos obtenidos en cada período mediante los métodos de información mutua normalizada (NMI) e índice de Rand ajustado (ARI), que varían entre 0 (ninguna coincidencia de asignación entre pares de asignaciones para el mismo nodo) y 1 (coincidencia total entre pares).

3. Redes de flujos de empleo

En este Capítulo se analizan y caracterizan las redes observadas de flujos de empleo entre sectores industriales. En la Fig. 5 se resumen las principales representaciones alternativas de una red de interacciones que se utilizan para abordar el análisis de los datos. A partir del registro de la movilidad laboral dentro del empleo que reflejan entradas y salidas para cada sector (matrices de transiciones en Fig. 1.B), estas interacciones pueden expresarse en forma de una lista de lazos, de un mapa de calor (*heatmap*) que puede ordenarse por algún criterio de agrupamiento (por ejemplo, jerárquico aglomerativo como en la Fig. 5.C2), o visualizarse en forma de grafo. A partir de la representación de grafos es posible utilizar las matrices fundamentales asociadas a los grafos (ver sección 2.2) para extraer estadísticas descriptivas de su estructura de interconexiones. Cada representación resulta útil para obtener distinto tipo de información acerca de la estructura de las interacciones que representa.

En primer lugar, se visualizan y describen las principales características observables de las redes de flujos de empleo. En segundo lugar, se analiza la evolución de las estructuras interanuales de transiciones interindustriales observadas y de aquellas que surgen de sus promedios para los períodos seleccionados, que se utilizan para el resto del análisis. Por último, se presenta un análisis descriptivo de estadísticas de las redes extraídas y una caracterización de la tipología de las redes obtenidas en base a las propiedades observadas.

3.1. Perfil

En la introducción se mencionó que la movilidad laboral dentro del empleo formal correlaciona positivamente con la actividad económica (Fig. 3) y que, al nivel de detalle utilizado (4 dígitos CIIU), las transiciones de empleo entre industrias cobran importancia respecto de las que se conservan dentro de la industria así definida (Fig. 2). Esto equivale a decir que a mayor nivel de detalle (o granularidad) de la actividad económica, la definición de industria es más restrictiva (Fig. 4) dado que es cada vez más específica por lo que, a medida que se incrementa la granularidad, resulta cada vez menos probable que los cambios de empleo se realicen dentro del mismo sector. En términos de las matrices de transición (Fig. 5.A) se espera que presenten cada vez menos flujos de empleo entre empresas de la misma industria (en azul, dentro de la diagonal principal) que entre empresas de otra industria (en rojo, fuera de la diagonal), lo cual implica que la carga de flujos se desplazará hacia afuera de la diagonal principal a medida que se incrementa la granularidad. Esta observación constituye en sí misma una línea de investigación, actualmente en curso, que indaga acerca de la determinación de una escala de resolución adecuada para el análisis de los flujos de empleo (Semeshenko and De Raco 2021).

En base a estas observaciones, ¿qué forma característica es esperable encontrar en las redes de flujos de empleo? ¿Qué información es posible extraer de los flujos de empleo a partir de matrices fundamentales asociadas a estas interacciones? En la Fig. 6 se muestran representaciones alternativas de los flujos de empleo formal privado promedio durante 2003-2010. A partir de la lista de lazos que es posible extraer del sistema de registro (SIPA), que permite construir la matriz de transiciones de empleo, es posible extraer información valiosa de la estructura de interacciones de las matrices de adyacencias binaria y ponderada (original) que la representan (ver, Capítulo 2, matrices fundamentales).

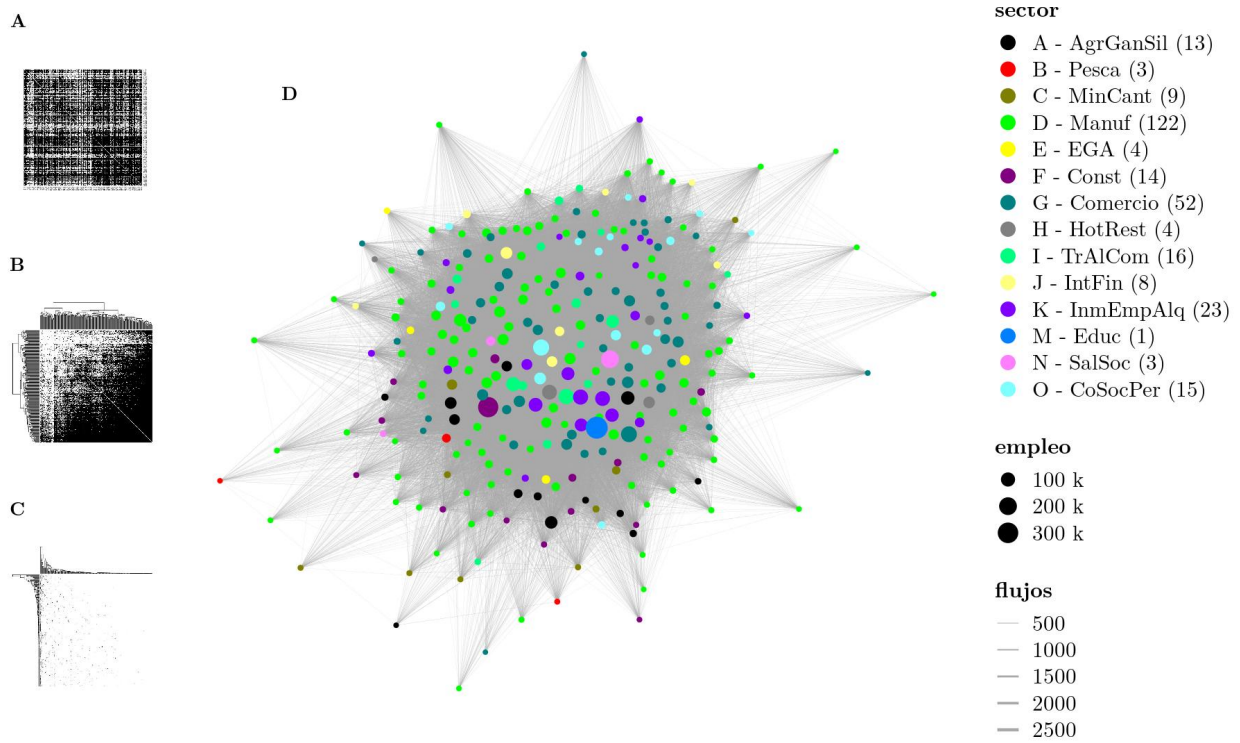
Utilizando la matriz de adyacencias binaria, que indica la presencia o ausencia de lazos entre sectores, es posible extraer información de la conectividad entre sectores. En la Fig. 6 se observan dos representaciones posibles de mapas de calor de esta matriz indicador. En (A) se presentan las interacciones tal cual se ordenan por el clasificador estadístico CIIU Rev 3 y en (B) su representación luego de ordenar la matriz con un algoritmo jerárquico aglomerativo de vinculación completa, que devuelve grupos de sectores con mayor interacción entre todos los integrantes favoreciendo “cliques” hacia su interior, es decir subgrafos de interacción completa. Comparativamente, en (A) es difícil distinguir una estructura de relaciones interindustriales mientras que en (B) se evidencia una estructura de grupos con baja conectividad entre sí (en bloques más claros, ausentes de interacción, desde arriba a la izquierda a través de la diagonal principal) y otros con mayor conectividad entre sí (bloques más oscuros, hacia abajo a la derecha a través de la diagonal principal). Los últimos constituyen distintas clases de *centros*, que mediante los cambios de empleo se relacionan dentro de los sectores que los componen y con el resto de la red, y son sectores con alto nivel de conectividad. Los primeros en cambio constituyen la “periferia,” que se caracteriza porque los sectores que la componen poseen notablemente menos conectividad y se relacionan casi únicamente con sectores centrales.

Al observar la matriz (C) de adyacencias ponderada (de flujos de empleo original) ordenada jerárquicamente y normalizada por el total de flujos, es posible observar que pocos sectores captan la mayor intensidad de carga (flujos) de la matriz mientras que la mayoría presenta flujos de baja intensidad. La atomización de los flujos guarda relación con las implicancias de la elección del nivel de granularidad en la definición de sector o “industria.” Por último, en (D) se presenta el grafo de la matriz de adyacencias ponderada en el que se representa una red cuyos nodos son sectores a cuatro dígitos y cuyos lazos muestran los flujos de empleo. Se observa un gran componente único y conexo, denso y con forma de “bola de pelos” (*hairball*), en que los sectores en el centro interactúan con mayor intensidad entre sí y con toda la red, mientras que los sectores más distantes corresponden a sectores de la periferia, con baja conectividad e intensidad de interacción.

3.2. Evolución temporal y agregación

En esta sección se analizan los cambios en la estructura de los flujos interindustriales de empleo a lo largo del tiempo para el período 1996-2020. Dado que los datos utilizados provienen del registro formal de empleo de las empresas y reflejan decisiones de cambios (incorporaciones y desvinculaciones¹⁵) en su planta de personal

¹⁵El tipo de desvinculaciones, voluntarias (renuncia) o involuntarias (despido), quedan fuera del análisis dado que no resulta posible distinguirlo a partir de los datos de registro.



Fuente: Elaboración propia en base a SIPA

Figura 6: Representaciones de redes de flujos de empleo interindustrial 2003-2010. Mapas de calor de la matriz binaria de transiciones: (A) ordenada según la clasificación de registro CIU Rev 3, y (B) agrupada por clustering jerárquico de vinculación completa. Ambos muestran una red densa, poblada de interconexiones. La red ordenada por agrupamientos jerárquicos permite observar la organización en torno a la diagonal de grupos de baja conectividad (periferia, bloques de mayor proporción en blanco) y grupos de distintos niveles de conectividad (centros, bloques diagonales con distintos niveles de densidad en negro). (C) Mapa de calor de la matriz ponderada (normalizada) de flujos de empleo ordenada por agrupamiento jerárquico: destaca que pocos sectores captan la mayor intensidad de la matriz. (D) Grafo ponderado: realizado con layout de fuerzas dirigidas LGL, presenta una red con un gran componente único y conexo, denso con forma de "bola de pelos", en que los nodos de la periferia corresponden a sectores con baja conectividad e intensidad de interacción y hacia el centro se encuentran los que interactúan con mayor intensidad. El color de los nodos representa la letra del sector industrial y su tamaño la dotación relativa.

en función de las condiciones coyunturales—que pueden variar levemente en períodos de estabilidad ó abruptamente frente a eventos de crisis macroeconómicas—sujetos al contexto institucional relativamente estable del mercado laboral, se esperaría encontrar poca variabilidad de corto plazo y alta correlación cronológica entre las estructuras observadas interanualmente. En la evolución de estos cambios de corto plazo desde el punto de vista de las redes de interacciones interesa distinguir el grado en que es posible identificar diferencias y similitudes en las características globales de estos flujos de empleo que permitan sustentar la definición de períodos de estructura interanual con similitud relativa que a su vez se diferencien con otros períodos a fin de permitir una evaluación del cambio estructural de mediano plazo. Se utilizan los cuatro períodos identificados en base a la correlación y evolución de métricas globales de las redes interanuales y se construyen las redes promedio correspondientes, sobre las que se realiza el análisis posterior¹⁶.

3.2.1. Interanual

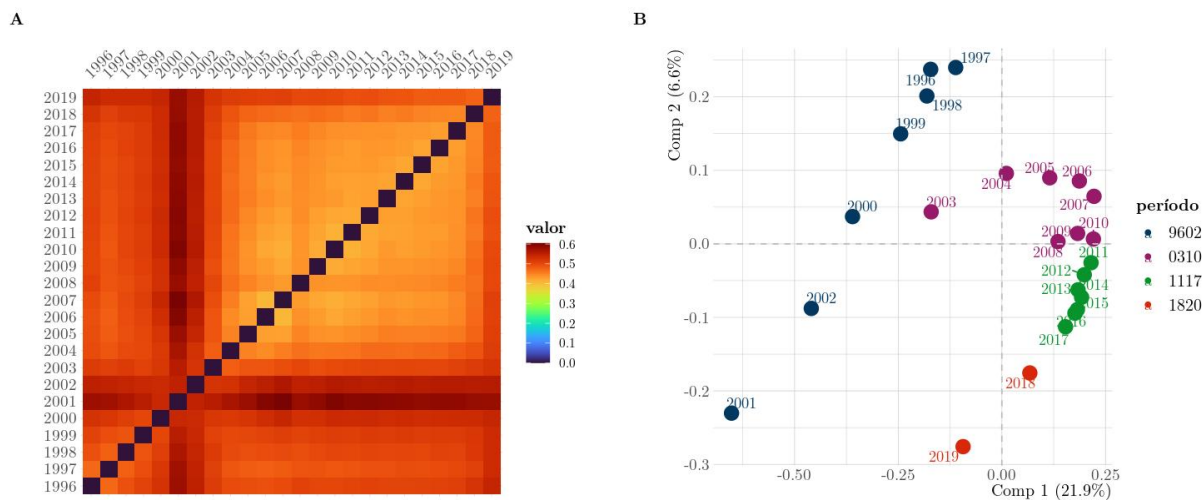
Correlación. A fin de analizar la evolución temporal de las matrices de flujos se evalúa la similitud de las mismas mediante el cálculo del coeficiente de Pearson entre cada par de matrices de adyacencias, A_t , vectorizadas (i.e.: filas concatenadas en un solo vector columna, a fin de comparar la correlación de a pares de celdas, a_{ij}). La comparación de las matrices de transiciones interanuales entre 1996 y 2020 presentan correlación positiva en un rango de 0.24 y 0.93. En la Fig. 9 del Anexo se presenta el Cuadro completa donde se destaca una correlación positiva entre todas las matrices interanuales, que resulta más débil en años de crisis o cambios macroeconómicos de alto impacto. En particular, las transiciones observadas en 2017 lo muestran como un año notablemente atípico (correlaciones con el resto entre 0.24 y 0.43), mientras que en la misma dirección 2003 y 2006 se presentan como años con menor correlación con el resto de todos los años (entre 0.24 y 0.93). Por otro lado, se observa una fuerte correlación entre los años comprendidos en los períodos 1996-2000 (0.65-0.79), 2007-2016 (0.80-0.93), y 2018-2019 (0.88-0.90). Estos resultados sugieren que las redes que surgen de las transiciones interanuales de empleo interindustrial presentan habitualmente una estructura de conectividad persistente, con mayor estabilidad (en el sentido de la correlación analizada) en años consecutivos de un mismo “período.” En particular, en el período de Convertibilidad 1996-2000 se presenta menor correlación entre las estructuras interanuales dentro del período a la vez que ocurren una menor cantidad de transiciones (ver Fig. 2) sugiriendo que se detecta menor cantidad y variedad de lazos intersectoriales. Las transiciones de los períodos 2007-2016 y 2018-2020 aparecen como más similares hacia adentro de los mismos y ligeramente distintas entre ellos. Por otro lado, los resultados también permiten distinguir años en los que las transiciones aparecen como “anómalas,” con baja correlación con todos los restantes en mayor (2017) o menor (2001-2003) intensidad. Estos resultados servirán de base para identificar más adelante “régimenes” de transiciones de empleo basados en las diferencias de la estructura de conectividad y fundamentar la segmentación de períodos de mediano plazo en línea con ellos.

Distancias. Las redes interanuales de los $t = 24$ períodos se comparan mediante las tres distancias introducidas en la sección 2.2, que permiten distinguir diferencias a nivel local, intermedio (meso) y global de las estructuras de interacciones observadas. Este ejercicio resulta en cierta forma análogo a utilizar un microscopio con distintas lentes de aumento para observar un mismo conjunto de objetos, en que con cada lente es posible observar estructuras que destacan características observables a la resolución que éste lo permite. Así, se espera detectar diferencias particulares aportadas por cada distancia como también características comunes de las redes comparadas que son captadas por todas ellas. En el primer caso se utiliza la distancia de Jaccard, $d_{Jacc}(G, \bar{G})$, que cuantifica la cantidad de cambios en las interacciones del entorno directo (local) de cada sector, es decir de los sectores con los que tiene conexión directa (medidos por el grado, k , del sector). En segundo lugar se utiliza la distancia de la distancia polinomial de segundo grado, $d_{Poly2}(G, \bar{G})$, para cuantificar las diferencias en un entorno ampliado que abarca al vecindario de segundo orden de cada sector, es decir las conexiones de los sectores del entorno de intercambios directos de un sector (i.e.: amigos de los amigos). En tercer lugar se utiliza la distancia l_2 de valores propios, $d_{Eig}(G, \bar{G})^p$, que permite cuantificar las diferencias en todo el espectro de frecuencias de los valores propios de los grafos asociados a cada red de flujos interanuales. Este análisis de comparación de las redes intertemporales se realiza sobre las *matrices de adyacencias binarias* con lo que se pone el foco en comprender las diferencias en la conectividad presente en cada período más que en el volumen y la dirección de las transiciones observadas.

¹⁶La definición de los períodos se realizó en conjunto con el MTEySS en el transcurso de la cooperación mencionada.

Jaccard. En la comparación de las matrices de adyacencias de las redes interanuales, A con $d_{Jacc}(A, A')$ es posible observar dos grandes subperíodos, entre 1996 y 2000 (etapa de Convertibilidad, en términos de la descripción temporal propuesta en el Capítulo 1) y entre 2004 y 2018 (etapas posteriores), que presentan relativa similitud dentro de los mismos en los años consecutivos (regiones de color naranja más claro sobre la diagonal en la Fig.7.A) y difieren entre períodos (regiones de color naranja más oscuro hacia los bordes, Fig.7.A), separados por años de crisis 2001-2003 y 2019, en que aparecen distancias mayores con todos los años (franjas de rojo oscuro, Fig.7.A). Ello evidencia diferencias en la estructura local de transiciones de empleo que se sostienen en el mediano plazo, es decir algún tipo de cambio estructural en la manera en que se organizaron y la cuantía en que registraron las transiciones interindustriales de empleo respecto de los intercambios directos entre sectores, que es lo que permite distinguir la medida de Jaccard.

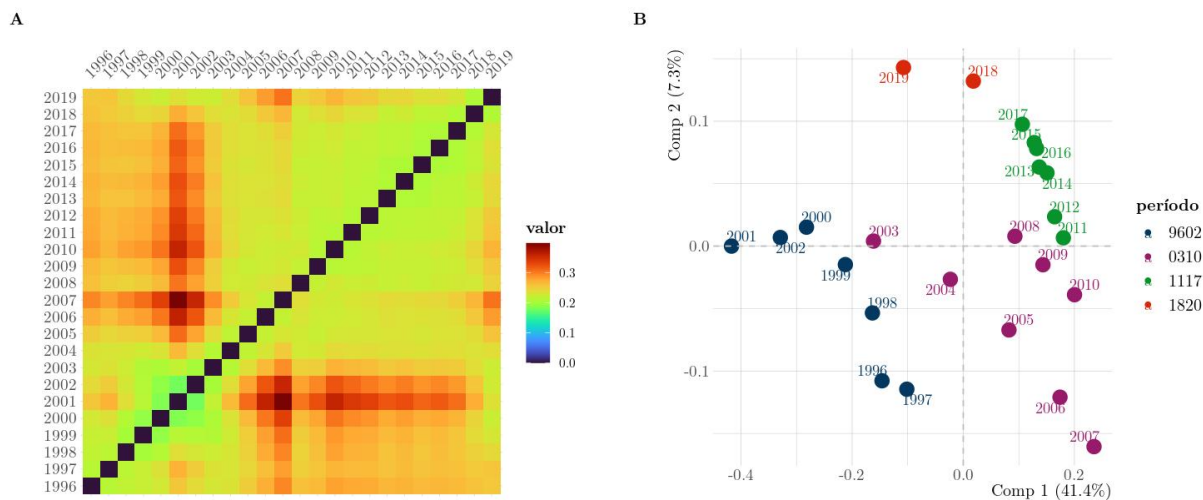
En la Fig.7.B se mapea la matriz de distancias calculada mediante escalamiento multidimensional métrico (MDS) en los dos primeros componentes que capturan la mayor variabilidad de esta matriz Comp 1 (21,9 %) y Comp 2 (6,6 %) acumulando 28,5 % de la variabilidad total. En esta proyección se observa que el eje vertical ordena hacia la izquierda, en los valores negativos del Comp 1 que captura la mayor parte de la varianza de las distancias, al primer grupo 1996-2000 junto con los años de crisis 2001-2003 y 2019, mientras que ordena en los valores positivos de Comp 1 a las transiciones del segundo grupo 2005-2018. Hacia los valores positivos se observan los años Post-Convertibilidad que el componente vertical ordena desde arriba hacia abajo a los años 2004 a 2018, bastante cercanos entre sí en línea con lo observado en el mapa de calor. En particular, el tercer cuadrante pareciera captar mermas en la conectividad y muestra hacia valores más negativos los años de crisis, con 2019 alejándose de la estructura del “régimen” de mayor movilidad y acercándose a 2001, un año de transición del régimen anterior de movilidad laboral. En línea con la descripción temporal propuesta en el Capítulo 1 y en términos de los cuatro períodos definidos en la sección 2.1, en la Fig.7.B y en las siguientes comparaciones se utilizan colores para identificar las transiciones anuales con estos períodos de referencia. Allí es posible notar que en esta proyección la distancia de Jaccard ordena a las transiciones de 2003, incluidas en el período de recuperación Post-Convertibilidad (0310, en violeta), con el resto de las observadas en el período anterior (9602, en azul) y a 2004 y 2019 como años que se encuentran entre ambos regímenes.



Fuente: Elaboración propia en base a SIPA

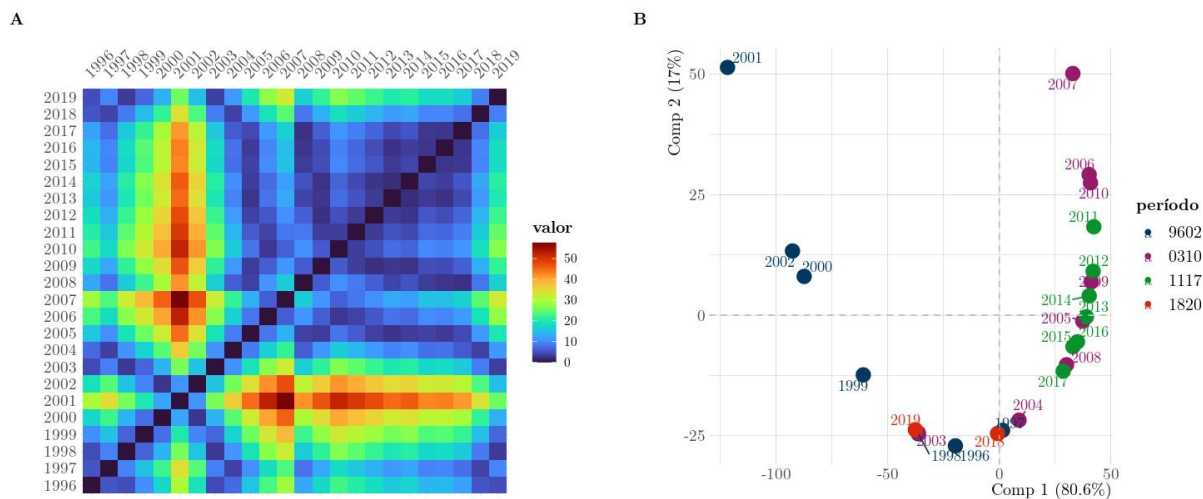
Figura 7: Distancias de Jaccard entre matrices de adyacencias binarias de flujos de empleo interanuales. (A) Mapa de calor. (B) Escalamiento multidimensional. Se distinguen dos regímenes de transiciones de empleo, 1996-2000 y 2004-2018, en términos de los cambios en los entornos directos de conexión de los sectores, y se diferencian los años de transición y crisis macroeconómicas 2001-2003 y 2019. El agrupamiento de períodos propuesto (en colores) se orienta en regiones características del mapa de los componentes que captan la mayor variabilidad de la matriz de distancias (Comp 1 y 2).

Polinomial. Esta distancia considera las diferencias de la conectividad ampliada a los sectores en sus entornos



Fuente: Elaboración propia en base a SIPA

Figura 8: Distancias polinomiales de segundo grado (vecindarios de segundo orden) entre matrices de adyacencias binarias de flujos de empleo interanuales. (A) Mapa de calor. (B) Escalamiento multidimensional. Se distinguen también dos regímenes de transiciones de empleo aunque los años de transición macroeconómica distinguidos por Jaccard quedan integrados a los regímenes. Aparece un subrégimen de con cierta variabilidad entre 2004-2007 y otro más estable entre 2008-2019 y la distancia máxima se observa entre 2001 y 2007.



Fuente: Elaboración propia en base a SIPA

Figura 9: Distancias l_2 de valores propios entre matrices de adyacencias binarias de flujos de empleo interanuales. (A) Mapa de calor. (B) Escalamiento multidimensional. Se identifican dos grandes grupos de períodos (1996-2005 y 2007-2018) o regímenes de movilidad laboral, a la vez que 2006 se identifica como un período distinto a todos los demás y 2019 como distinto a los años del régimen post-Convertibilidad y más semejante a régimen de movilidad laboral de la Convertibilidad.

de segundo orden (v.g.: sectores vinculados con aquellos que tiene lazos directos) y los resultados se presentan en la Fig. 8. Con esta distancia también se identifican dos grandes “régimenes” aunque no aparecen tan claramente los años de transición entre ellos como aparecían al analizar los cambios en las conexiones directas con Jaccard, como se observa en el mapa de calor Fig. 8.A. Así, se agrupan las transiciones de 1996-2003 y de 2004-2019 en dos grandes régimenes de conectividad extendida. Sin embargo, se observa un mayor contraste entre los años 2001-2003 y el resto de los años del segundo régimen Post-Convertibilidad en las estructuras comparadas. Esto sugiere que la mayor conectividad observada en general en dichos años se logró en la persistencia de las transiciones de segundo orden que no estaban presentes durante el régimen de movilidad laboral vigente en la Convertibilidad. En particular, se destaca el contraste entre las estructuras de 2001 y 2007, año a partir del cual se observan mayores similitudes interanuales entre 2008-2018. Esto podría estar relacionado con cambios en las vinculaciones intersectoriales a partir de 2008, año de crisis interna en Argentina con el conflicto con el campo y las restricciones a la movilidad interior. Por otro lado, como se observa en la serie temporal de transiciones en la Fig. 2, las transiciones de 2001 ocurren hacia 2002 que marca el primer año en que se revierte la tendencia negativa post-crisis en la serie, mientras que las transiciones de 2007 ocurren hacia 2008 que marca una fuerte caída en la serie, por lo que su comparación captura información de fuerte contraste que esta distancia logra captar.

En la proyección MDS, los dos primeros componentes Comp 1 (41,4%) y Comp 2 (7,3%) acumulan 48,7% del total de la variabilidad de esta matriz de distancias (Fig. 8.B). En esta proyección, el Comp 1 en el eje horizontal ordena nuevamente en sus valores negativos al período de Convertibilidad incluyendo a los años de transición 2003-2004 y a 2019 y hacia la derecha en los valores positivos a los restantes años. Sin embargo, al ampliar el orden de conectividad 2019 ya no aparece tan cercano a 2001 como cuando se consideraban sólo las conexiones directas. En particular, se observa que al comparar estas estructuras ampliadas en la proyección MDS de la Fig. 8.B, tienden a estabilizarse en el período 1117 identificado en la sección 2.1 como de “estancamiento” (Fig. 8.B, en verde).

Spectral. La distancia espectral l_2 utilizada recurre a encontrar la amplitud en la diferencia de los autovalores de las matrices comparadas en la totalidad del espectro de frecuencias, $\lambda_i, i = 0, 1, \dots, V - 1$. En este sentido resulta una medida amplia de las diferencias en la estructura global de las redes analizadas por lo que sugiere una interpretación más acorde al “estado del sistema” de transiciones interindustriales dentro del empleo formal privado a lo largo del tiempo. El mapa de calor de la Fig. 9.A identifica los mismos grandes grupos de períodos (1996-2005 y 2007-2018) distinguibles con la distancia polinomial, que consideraba diferencias en la estructura de segundo orden de conectividad, descriptos como dos “régimenes” de movilidad laboral. En este sentido, dado que el diámetro de las redes de flujos de empleo (más largo de los caminos cortos que conecta a los dos sectores más lejanos) resulta del orden de 3 pasos (De Raco and Semeshenko 2019a, y 2019b, y análisis más adelante) y que la distancia polinomial analizada captura información de entornos de hasta 2 pasos, se espera que la información adicional proveniente de una métrica relacionada con el espectro del grafo completo no resulte demasiado alejada de la obtenida por $d_{Pol2}(G, \bar{G})$. No obstante, dado que para calcular la distancia $d_{Eig}(G, \bar{G})$ se utilizan todos los autovalores de los grafos, aparece información adicional. En la proyección MDS (Fig. 9.B) los primeros dos componentes acumulan 97,6% de la variabilidad total de la matriz de distancias (80,6% Comp 1 y 17,0% Comp 2). Allí se observa que el eje horizontal (Comp 1) ordena hacia la izquierda a los años del primer régimen de movilidad (Convertibilidad, 1996-2003) y hacia la derecha el segundo régimen (2004-2018), con 2018 y 2019 más cercanos al primer régimen. El eje vertical (Comp 2) separa hacia arriba a los años 2001 y 2007, con máxima distancia l_2 entre sí, y hacia abajo a los años de transición entre régimenes y los extremos 2018 y 2019.

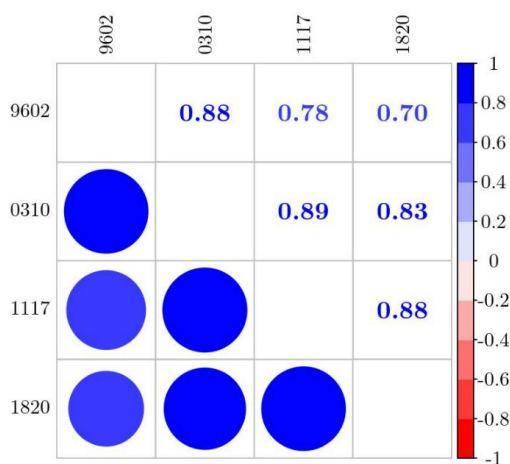
Hasta aquí se logra confirmar mediante las diferencias de las estructuras de las redes de flujos de empleo a distintos niveles de análisis de su conectividad (local, intermedio y global) que los períodos seleccionados aparecen representando distintos “régimenes” de movilidad del empleo. A continuación se definen y analizan las redes de flujos promedio de estos períodos agregados, que se utilizarán para el análisis de comunidades.

3.2.2. Mediano plazo

En términos cualitativos las correlaciones interanuales observadas (Fig. 9) y los resultados del análisis de las matrices interanuales implican que los períodos identificados en las Figs. 2 y 3 seleccionados para el análisis

de redes conservan cierta estabilidad en el sentido que presentan alta correlación interanual dentro de los años que los componen. En base a ello, se construyen matrices que promedian las transiciones interanuales de los períodos 1996-2002 (de aquí en más 9602), 2003-2010 (ídem, 0310), 2011-2017 (ídem, 1117) y 2018-2020 (ídem, 1820), que se utilizan y comparan en el análisis posterior de redes. Las matrices obtenidas resultan más densas que las anuales, ya que incluyen algunas transiciones que ocurrieron sólo en algún año, y contienen pesos medios no necesariamente enteros. Nótese que el período 1820 promedia transiciones de sólo dos años que a su vez son de caída de la actividad, mientras que los tres períodos anteriores agrupan entre 7 y 8 años de movilidad laboral presumiblemente similar, por lo que esperaríamos que este período contraste con los anteriores por las características de su definición.

Correlación. En la Fig. 10 se presentan las correlaciones de Pearson entre las matrices promedio vectorizadas—es decir con sus filas concatenadas en un solo vector columna a fin de comparar la correlación de a pares de celdas, a_{ij} —que varían entre 0.71 y 0.91 con lo que se consideran matrices diferentes aunque relativamente estables. En particular, se observa mayor correlación entre los períodos contiguos que entre los distantes.



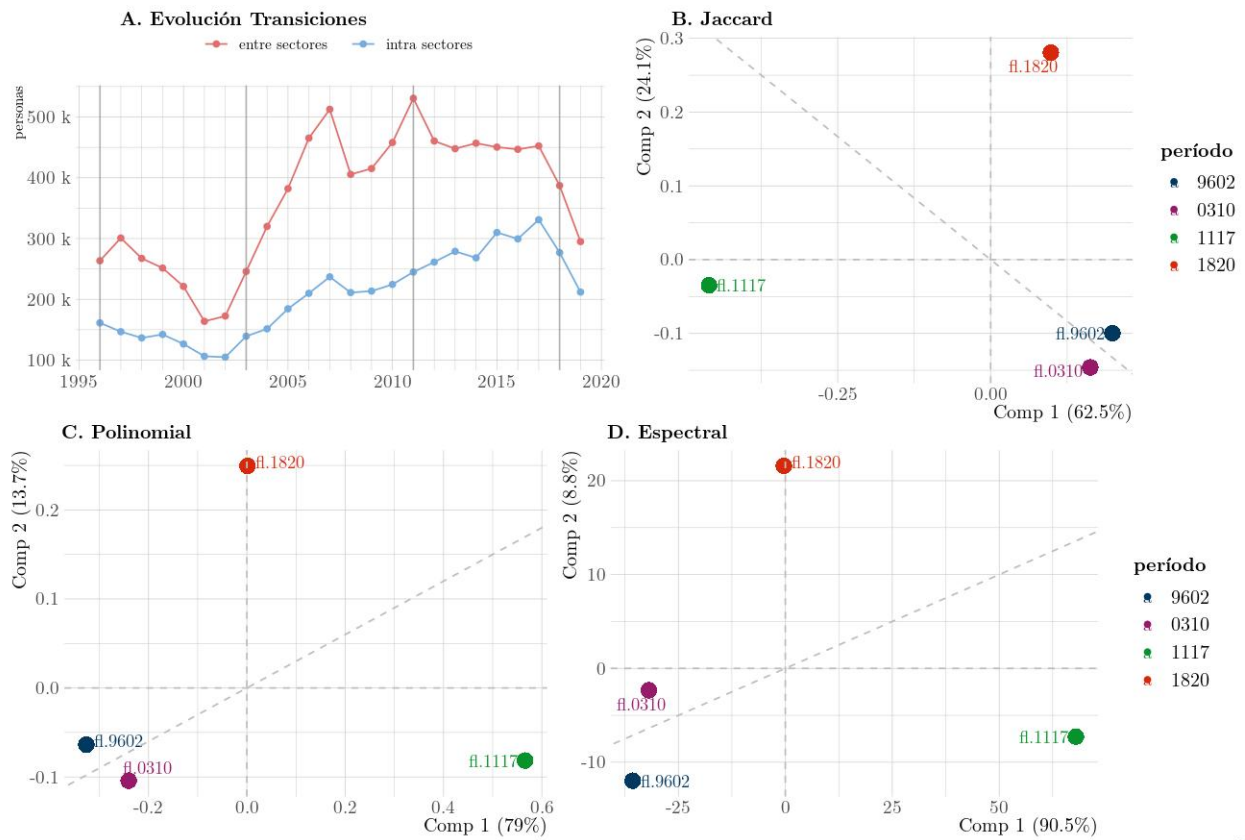
Fuente: Elaboración propia en base a SIPA

Figura 10: Transiciones de empleo interindustrial de períodos seleccionados. Correlación, de Pearson entre matrices vectorizadas.

Distancias. En la Fig. 11 se comparan mediante MDS las matrices de adyacencias binarias de los cuatro períodos agregados mediante las distancias de Jaccard, polinomial y espectral previamente utilizadas.

En la comparación de períodos agregados mediante las métricas propuestas resulta más complejo determinar el aporte de cada una de ellas. En la Fig. 11 se presentan las proyecciones MDS sobre los componentes que capturan la mayor variabilidad de cada matriz de distancias. Allí se observa que con las tres distancias utilizadas se ordena de forma cualitativamente similar a las matrices de las redes 1117 y 1820, que aparecen separadas entre sí y de 9602 y 0310, que en todos los casos aparecen más cercanas entre sí. En términos del comportamiento de mediano plazo de las transiciones de empleo entre sectores (Fig. 2 en Capítulo 1, incluida aquí como Fig. 11.A), se espera que 1820 se distinga del resto por su característica de menor densidad y por su retroceso a estructuras de interacción con algunas características posiblemente similares a las de algunos años de la Convertibilidad (ver orientación relativa en Fig. 7.B y 8.B). El período agregado 1117 condensa períodos de estancamiento con matrices interanuales más similares entre sí que en cualquier otro período. No obstante, al trazar un eje de simetría entre las dos componentes (diagonal punteada) que separe las redes más distantes y las más cercanas se observan algunas diferencias.

En resumen, la construcción de los períodos agregados en base a los períodos según la correlación con el PIB parece razonable desde el punto de vista del análisis de correlación de las matrices interanuales y de



Fuente: Elaboración propia en base a SIPA

Figura 11: Distancias entre matrices de adyacencias de flujos de empleo de períodos seleccionados: Evolución de las transiciones de empleo y Escalamiento multidimensional (MDS) de las matrices de distancias, proyección en primeros componentes (Comp 1 y 2). A. Evolución de las transiciones de empleo. B. Jaccard (local), C. Polinomial (vecindarios de segundo orden), y D. Espectral l_2 (global, autovalores). El ordenamiento cualitativo de los períodos de mediano plazo se mantiene con todas las distancias. En relación con la construcción de las redes agregadas, que promedian transiciones de empleo interanuales, las redes contienen mayor cantidad de lazos por lo que son más densas y similares entre sí y no se identifican períodos particularmente anómalos como sucede en las redes interanuales.

las distancias analizadas, que consideran similitudes y diferencias en la conectividad de entornos directos y entornos de segundo orden de los nodos, así como la evaluación de la conectividad global de los grafos completos.

3.3. Análisis descriptivo

Las estadísticas de las redes permiten identificar propiedades globales del grafo y de los nodos que la componen. Las medidas de densidad y de cohesión caracterizan la red en términos globales, mientras que las de centralidad y de grado de conectividad permiten caracterizar la estructura local de una red. En esta sección se evalúan estadísticas descriptivas de las matrices de transición interanuales y agregadas de flujos de empleo en términos de su caracterización y comparación en el tiempo.

Una medida global de conectividad utiliza el *largo de los caminos promedio* (*average path length*, APL) entre pares de nodos, que en las redes sociales se refiere al grado de separación o saltos (*hops*) entre ellos. El *diámetro*, d , de la red se mide como el camino promedio más largo, la distancia lineal entre los nodos más alejados. En las redes de flujos de empleo estas métricas se pueden interpretar como los caminos que surgen de conectar los distintos sectores mediante las incorporaciones y desvinculaciones de empleo (formal privado en este caso) de trabajadoras que se desvinculan de un sector distinto al que se incorporan. El diámetro es el camino más largo de todos los caminos cortos que conectan todos los pares de sectores. Otras métricas relevantes refieren a: la cantidad de *componentes conexos* que presenta el grafo, que cuenta la cantidad de grupos de nodos que tienen caminos de conexión entre ellos; la *densidad* de la red, medida como el cociente de lazos observados sobre lazos posibles entre los nodos de la red; la *reciprocidad*, que en redes dirigidas mide el cociente de lazos observados en ambas direcciones entre dos nodos respecto de la cantidad de lazos posibles; el *coeficiente global de agrupamiento o clustering* (CCoef), que se basa en las relaciones de nodos tomadas de a tres y se calcula como el ratio entre los triángulos transitivos de conexión observados respecto de las posibles combinaciones de tres nodos; y la *distribución de grados* de los nodos, que presenta la cantidad de conexiones de a pares en las que participa cada nodo.

Otras estadísticas relevantes refieren a la centralidad de los nodos. Existe una variedad de métricas para caracterizar la centralidad (véase Anexo metodológico). La centralidad por *intermediación* (*betweenness centrality*) mide el grado en que un nodo participa en los caminos más cortos entre los distintos nodos de la red. La centralidad por *cercanía* (*closeness centrality*) cuantifica los pasos que se requieren para acceder desde un nodo dado a cada uno de los demás nodos.

Cuadro 2: Estadísticas globales en redes interanuales y períodos agregados.

Redes	Orden	Tamaño	$\langle CCon \rangle$	Flujos	Dens	$\langle k \rangle$	$\langle k^{in} \rangle$	$\langle k^{out} \rangle$	$\langle knn \rangle$	Asort	$\langle Recip \rangle$	$\langle CCoef \rangle$	LCP	Diám	$\langle Betwn \rangle$	$\langle Core \rangle$
Promedio																
$\langle F_L \rangle$	287	26.961	1,1	278.117	0,328	188	94	94	285	-0,26	0,69	0,65	1,68	3,4	0,0024	116
$\langle F_p \rangle$	287	46.190	1,0	269.257	0,563	322	161	161	391	-0,24	0,82	0,80	1,44	2,8	0,0015	222
Períodos seleccionados																
9602	287	45.613	1,0	193.559	0,556	318	159	159	390	-0,25	0,82	0,79	1,44	3,0	0,0016	215
0310	287	54.592	1,0	318.602	0,665	380	190	190	432	-0,22	0,87	0,85	1,33	2,0	0,0012	272
1117	287	52.169	1,0	328.269	0,636	364	182	182	422	-0,23	0,86	0,84	1,36	3,0	0,0013	259
1820	287	32.388	1,0	236.598	0,395	226	113	113	319	-0,26	0,74	0,71	1,61	3,0	0,0022	143

Cuadro 3: Distribución de pesos por tramos, períodos agregados.

[illegible]

Como se mencionó en la sección anterior, dado que las redes promedio de los períodos seleccionados abarcan un conjunto de años, incorporan una mayor variedad de lazos por lo que las métricas globales presentan estadísticas de mayores niveles de densidad y cohesión (ver Cuadro 2 más adelante). Las diferencias de las métricas promedio de transiciones interanuales y de períodos agregados surgen debido a que las últimas agrupan flujos de empleo poco frecuentes, que ocurren sólo en algunos años y en baja magnitud, por lo que en las redes interanuales se capta menor heterogeneidad y menor cantidad de interacciones. El beneficio de la agregación de los flujos de empleo para un conjunto de años relativamente asimilables reside en que es posible captar mayor heterogeneidad de interacciones y permite generar un grafo con flujos potencialmente más representativos del estado del sistema productivo en un horizonte temporal de mediano plazo (análisis estructural). El costo se manifiesta en redes más grandes (densas) que las observadas interanualmente que presentan una mayor proporción de flujos de baja intensidad.

En el Cuadro 2 se presenta un resumen de estadísticas globales promedio de las redes de flujos de empleo interanuales y de los períodos seleccionados, así como las estadísticas individuales de cada período agregado. Se observa que el diámetro resulta $d_a \approx 3,4$ grados en las redes interanuales y $d_p \approx 2,5$ en las redes promedio de los períodos analizados. Esto significa que es posible encontrar relaciones entre sectores que permitan llegar de un sector a cualquier otro en poco más de tres cambios de empleo entre distintos sectores, lo cual en términos de la variedad de sectores de actividad (definidos a 4 dígitos CIIU) resulta un camino muy corto. La densidad promedio resulta de 0.33 y 0.58, respectivamente, que en términos de flujos de empleo implica que la variedad de intercambios observados en las redes interanuales representan 1/3 de todos los posibles intercambios, mientras que para las redes agregadas esta proporción casi se duplica. La reciprocidad promedio de los lazos de 0.69 y 0.83, respectivamente, lo que significa que en las redes interanuales en promedio poco más de 2/3 de los intercambios de empleo entre dos sectores cualesquiera ocurre en ambas direcciones mientras que en las redes agregadas esta proporción ocurre en poco más de 4/5 de las veces. El coeficiente global de clustering medio es de 0.65 y 0.80, respectivamente, lo que implica que en las redes interanuales tomando 3 sectores cualesquiera 2/3 de las veces ocurre que hay relación entre estos tres sectores mientras que en el caso de las redes agregadas esta relación ocurre en promedio 4/5 de las veces. En cuanto a la cohesión, las redes construidas presentan prácticamente en su totalidad un único componente conexo, con la excepción de la red de transiciones de 2002 con 3 componentes en la que dos sectores quedan aislados del resto del grafo. Las redes de transiciones interanuales presentan relativa estabilidad en sus estadísticas globales y también es posible identificar el impacto de años particulares, como 2002 que presenta una red de más de un componente conexo y los períodos 1998-2003 y 2015-2019 en que se reduce la centralidad por cercanía.

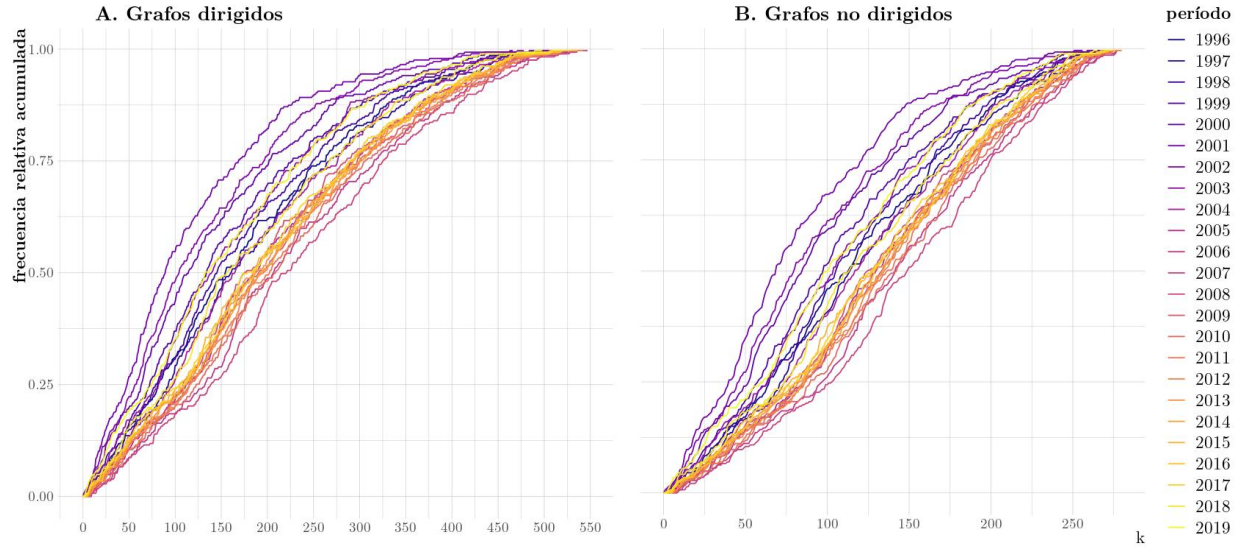
Distribución de pesos. En las redes de flujos de empleo la cantidad de transiciones intersectoriales observadas (registradas) representan los pesos o ponderaciones de los lazos que las componen. En términos económicos la cuantía de estos flujos resulta muy relevante, aunque una característica de las redes de flujos de empleo construidas interrelacionando sectores al nivel de detalle que se analiza aquí (4 dígitos CIIU), que consiste en un conjunto de $V = 287$ sectores, es la distribución muy concentrada de estos en valores pequeños. Así, al ordenar los flujos en orden creciente de tamaño y la cantidad de lazos que se acumulan en cada período se observa que los tramos inferiores, de transiciones de empleo intersectorial de hasta 10 personas, acumulan una proporción sustantiva de lazos a la vez que en términos de flujos acumulan una pequeña proporción del total (Cuadro 3).

Nótese que esta distribución de los pesos, característica de los flujos de empleo entre sectores definidos a gran nivel de detalle, plantea el problema de la significatividad de los flujos pequeños y su distinción entre estructura y ruido, en particular para la definición de corredores de empleo en esta escala. Este problema comúnmente se aborda con criterios de filtrado de los flujos para la reducción de las redes, que se exploró parcialmente en una etapa previa al desarrollo de esta tesis en (De Raco and Semeshenko 2019a) mediante el método tradicional de umbrales globales y constituye una línea de investigación activa.

Distribución de grados. En cuanto a la distribución de grados totales, $k_i, i \in V$, en las Figs. 12 y 13 se observa su evolución para las redes interanuales y de períodos seleccionados, respectivamente, en las que se presentan las versiones observadas (flujos dirigidos, A) y su aproximación de promedio de flujos en ambas direcciones (flujos no dirigidos, B). Allí se distingue la mayor conectividad en los años de recuperación post-Convertibilidad asociados a la mayor movilidad laboral en las curvas más cóncavas hacia el eje horizontal.

Tanto en las redes interanuales como en las agregadas en sus dos versiones (dirigidas y no dirigidas) se observa que las distribuciones de grados acumulan frecuencias de nodos con valores de k relativamente bajos en los períodos de baja correlación con la actividad económica. Este patrón pone en evidencia que se trata de redes densas y de alta conectividad media, con menor densidad (mayor proporción acumulada de k bajos) enfatizada en las redes que agregadas (13). Esta característica implica que los flujos de empleo ocurren muy frecuentemente entre diversos sectores económicos, muchas veces aunque no siempre en ambas direcciones.

Como se mencionó en el Capítulo 2, la elevada reciprocidad de las redes disminuye la distorsión de aproximar las redes originalmente dirigidas (con flujos direccionados) mediante la simplificación en redes no dirigidas, de flujos simétricos en ambas direcciones. En términos de grados de conectividad, el grado total k_i de un nodo en una red dirigida es $k_i = k_i^{in} + k_i^{out}$, que en el caso de un nodo conectado con todo el resto de los nodos de la red será $k_i = 2(V - 1)$, mientras que en una red no dirigida es sólo k_i , cuyo valor máximo será $k_i = V - 1$.

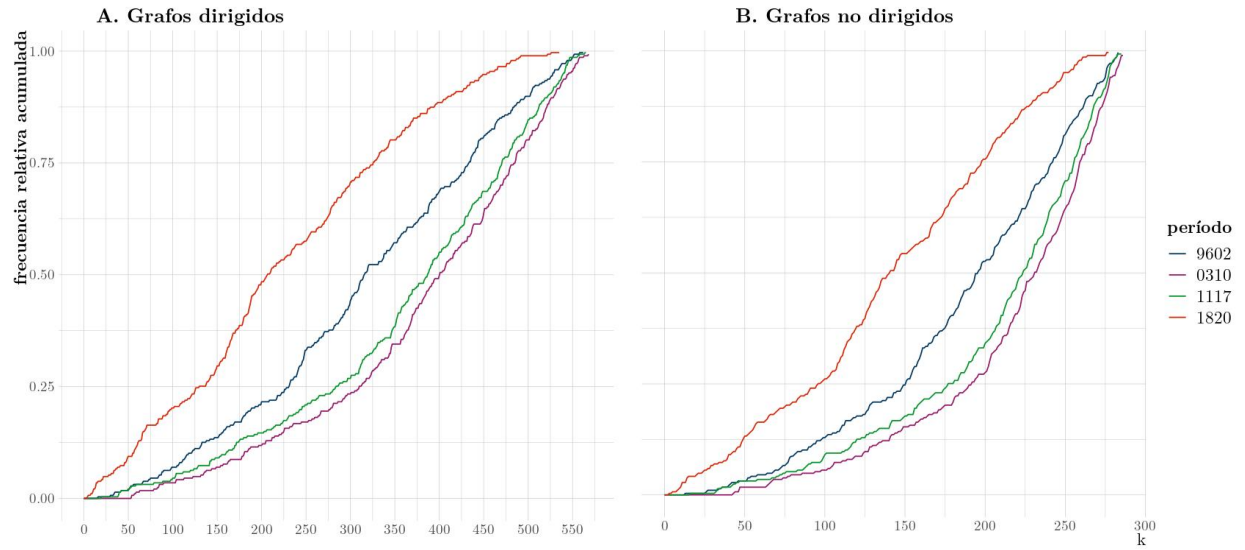


Fuente: Elaboración propia en base a SIPA

Figura 12: Distribución acumulada de grados de las redes interanuales. (A) Grafos dirigidos: Escala de color para los años de violeta, naranja y amarillo, desde 1996 (violeta) al 2019 (amarillo).

Las estadísticas reportadas ponen en evidencia que las transiciones de empleo (formal privado) interanuales entre sectores de actividad definidos a 4 dígitos observadas en la Argentina las últimas dos décadas componen una red densa e interconectada, aunque en menor medida que las transiciones agregadas previamente estudiadas, que a su vez también se concentran en un único gran componente conexo. Esta evidencia de conectividad intersectorial prevalente en la movilidad laboral del empleo le otorga relevancia a los métodos de análisis introducidos en esta tesis.

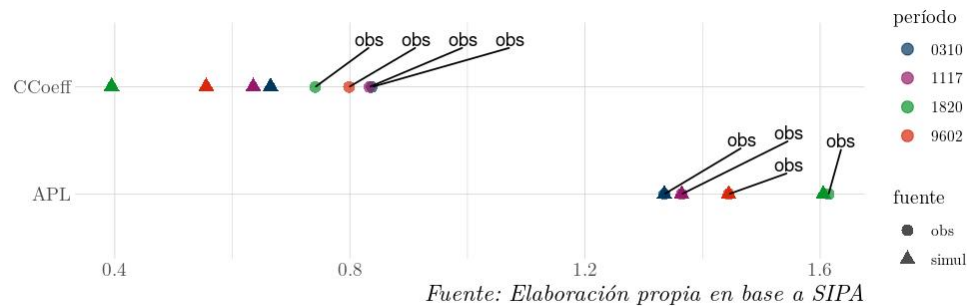
Propiedad de mundo pequeño. En líneas generales y de forma similar a los resultados reportados en (De Raco and Semeshenko 2019a, 2019b), las redes de flujos de empleo analizadas poseen una estructura de interconexión densa de diámetro pequeño y caminos muy cortos, con reciprocidad y coeficiente de clustering elevados que se adecua a una estructura de red de la tipología denominada “mundo pequeño.” En este aspecto, en simulaciones realizadas mediante la generación de redes aleatorias usando los parámetros de orden (cantidad de nodos, V) y tamaño (cantidad de lazos, E) de las redes promedio de cada período, con $n = 1000$ simulaciones, se encontró que los coeficientes de clustering global observados presentaron valores muy distantes del rango de las redes aleatorias generadas (ver Fig.14). Por el contrario, las distancias medias observadas (largo promedio de caminos, se encontraron dentro de las generadas aleatoriamente en las simulaciones (ver Fig.14), con lo que se confirma que las redes observadas se comportan estadísticamente como redes de mundo pequeño. Este tipo de estructuras son comunes a las redes sociales e implican que los caminos para llegar de un nodo (sector) a otro (en este caso a través de cambios intersectoriales de empleo)



Fuente: Elaboración propia en base a SIPA

Figura 13: Distribución de grados de las redes de períodos agregados. Cuanto mayor conectividad, más cóncava es la curva hacia el eje horizontal indicando que se acumula menos nodos con grados k de conectividad bajos. El período 1820 (rojo) presenta niveles de conectividad sustancialmente inferiores al resto de los períodos.

son cortos y que la red presenta una conectividad elevada.

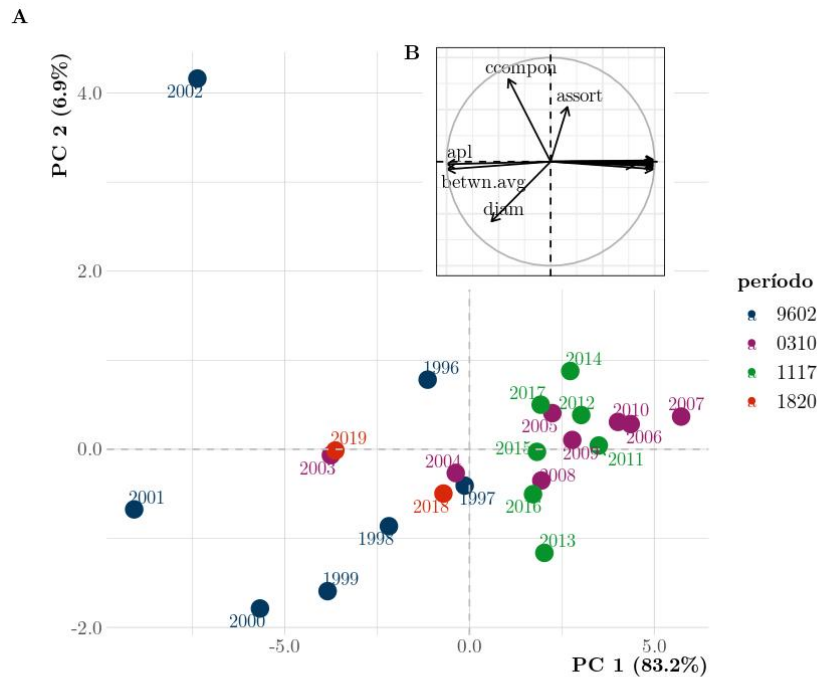


Fuente: Elaboración propia en base a SIPA

Figura 14: Coeficientes de clustering globales y Largo promedio de caminos. Valores simulados vs. observados en los períodos de referencia. Las redes de flujos de empleo observadas presentan coeficiente de clustering sensiblemente superiores a las generadas en redes aleatorias simuladas ($n=1000$) con la misma cantidad de nodos y lazos, mientras que en todos los casos el largo promedio de los caminos entre nodos observado no fue significativamente distinto al generado en redes aleatorias. Estos resultados indican la presencia de la propiedad de mundo pequeño en las redes observadas.

Comparación temporal. En la Fig. 15 se observan similitudes entre las redes interanuales según surge de la aplicación del método de análisis de componentes principales (PCA) sobre la matriz de métricas globales de las redes interanuales, proyectadas en las primeras dos componentes extraídas que captan la mayor proporción de la variabilidad de este conjunto de datos y acumulan 90,1 % del total. Allí (Fig. 15.A) es posible identificar en los cuadrantes de la izquierda a los años correspondientes al primer período (1996-2002), a la vez que separa en el cuadrante superior a los primeros años de este período y en el cuadrante superior a los cercanos al final de la Convertibilidad con 2002 (única red que presenta más de un componente conexo). A su vez, en estos cuadrantes se ubican las transiciones de 2018 y 2019 que aparecen similares a 2003 y 2004, emparentados por este conjunto de métricas con aquellos años de restricción de la movilidad del empleo formal. Como se observa en la Fig. 15.B, estos cuadrantes captan hacia la izquierda la influencia de las métricas de cohesión (componentes conexos, hacia abajo) y distancias (largo de caminos promedio y diámetro, hacia la izquierda

y hacia abajo). Por otro lado, en los cuadrantes de la derecha (dominados por la influencia de métricas de conectividad como el grado, la reciprocidad y el coeficiente de clustering, y de tamaño), se observa mayor cercanía entre las redes de los años 2005 a 2017, con 2007 como el año más “interconectado” dentro de los datos analizados.

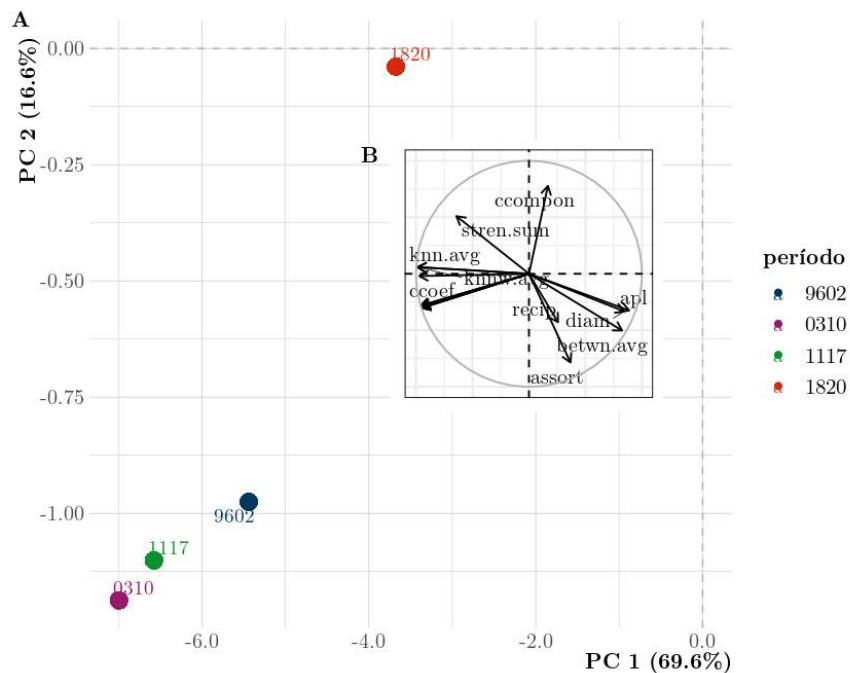


Fuente: Elaboración propia en base a SIPA

Figura 15: Análisis de componentes principales (PCA) sobre el conjunto de métricas seleccionadas de las redes de transiciones interanuales. En color el período de agrupamiento. (A) Similitud de las redes interanuales en el espacio de las primeras dos componentes principales, que explican el 90,1% de la variabilidad acumulada de los datos. (B) Orientación de las variables representadas (métricas) en este espacio: hacia la derecha variables de conectividad (grados, reciprocidad, transitividad) y tamaño de las redes; a la izquierda APL y betweenness promedio; a la izquierda arriba el diámetro; a la izquierda abajo cantidad de componentes conexos; a la derecha abajo asortatividad. El eje vertical separa los años con mayor conectividad hacia la derecha que corresponden al régimen de recuperación post-Convertibilidad y estancamiento (0310 y 1117). Los años del régimen de Convertibilidad (9602) se ubican a la izquierda, con 2002 hacia abajo que es el único año con más de un componente. Se distingue la secuencia 1996-2001 como creciente en la longitud media de los caminos y del diámetro. Las métricas de 2018 y 2019 (1820, períodos de declinación) se asimilan a 2003 y 2004.

En la Fig. 16 se mapean las redes agregadas según se proyectan en las primeras dos componentes resultantes de la aplicación PCA sobre la matriz de métricas globales de dichas redes, que captan la mayor proporción de la variabilidad de este conjunto de datos y acumulan 86,2% del total. Las redes se ubican en el cuadrante negativo de ambos componentes, abajo a la izquierda. Dentro de las variables representadas en este espacio, hacia la izquierda se ubican métricas de conectividad (grados, reciprocidad, transitividad) y tamaño de las redes; a la izquierda arriba peso de las transiciones; a la derecha arriba cantidad de componentes conexos; a la derecha abajo asortatividad, reciprocidad, diámetro, betweenness y APL. Las transiciones agregadas 1820 aparecen más a la derecha indicando menor tamaño (cantidad de lazos), fuerza (pesos) y conectividad media de los sectores, mientras que 9602, 1117 y 0310 se ubican hacia abajo a la izquierda en orden creciente de aquellos factores. A diferencia de la comparación entre las redes interanuales, que en la representación de los primeros componentes de PCA presentan mayor variabilidad, las redes agregadas de mediano plazo se ordenan más distintivamente por métricas de tamaño. Nuevamente, al comparar las redes interanuales en función del período de mediano plazo que las agrega, se encuentra que 2018 y 2019 en la Fig. 15.A no aparecen separados del resto como 1820 en la Fig. 16.A sino superpuestos a redes del período 9602. Ello

posiblemente ocurre en parte por la diferencia en la cantidad de períodos promediados que contiene 1820 respecto de los demás.



Fuente: Elaboración propia en base a SIPA

Figura 16: Análisis de componentes principales (PCA) sobre el conjunto de métricas seleccionadas de las redes de transiciones agregadas. En color el período de agrupamiento. (A) Similitud de las redes agregadas en el espacio de las primeras dos componentes principales, que explican el 86,2% de la variabilidad acumulada de los datos. (B) Orientación de las variables representadas (métricas) en este espacio: hacia la izquierda variables de conectividad (grados, reciprocidad, transitividad) y tamaño de las redes; a la izquierda arriba peso de las transiciones; a la derecha cantidad de componentes conexos; a la derecha abajo asortatividad, reciprocidad, diámetro, betweenness y APL. Las transiciones agregadas 1820 aparecen más a la derecha indicando menor tamaño (cantidad de lazos), fuerza (pesos) y conectividad media de los sectores, mientras que 9602, 1117 y 0310 se ubican hacia abajo a la izquierda en orden creciente de aquellos factores.

Acerca de la unidad temporal de análisis. En base a estas primeras observaciones de las redes interanuales y agregadas obtenidas, surge la pregunta acerca de las ventajas de trabajar con unas u otras. Las primeras son más esparsas y menos heterogéneas, aunque son simples de construir y resulta posible y de bajo costo su comparación con redes de períodos anteriores a partir de métricas globales como las presentadas aquí. En cuanto a las redes agregadas, plantean el problema de la selección y evaluación de la representatividad del período de agregación e introducen cierto ruido al incorporar transiciones de bajo peso, que luego será preciso evaluar al utilizar métodos de filtrado de las redes. A priori, en términos de la utilidad práctica de estos enfoques parece razonable utilizar el análisis interanual para evaluar la similitud de las redes observadas año tras año. A la vez, los resultados de esta comparación pueden resultar útiles para, en conjunto con la evaluación macroeconómica de las transiciones de empleo en el tiempo, determinar los períodos a incluir en un análisis agregado de mediano plazo acorde a la pregunta de investigación a satisfacer.

Conclusiones. En el análisis realizado de las redes de flujos de empleo en este capítulo, se encontró que se trata de redes con alta conectividad, persistentes, habitualmente densas con coeficiente de clustering y reciprocidad elevadas, que se organizan en un solo componente conexo donde los sectores se interrelacionan mediante caminos cortos que permiten verificar la propiedad de “mundo pequeño.” La distribución de grados de conectividad de los sectores acumula presenta una relación positiva con la actividad económica. Por otro lado, las correlaciones de matrices interanuales permiten identificar similitudes y persistencia de estructuras de conectividad intersectorial, así como brindar soporte a la selección de años para la agregación de períodos.

Del análisis de las redes interanuales resulta que proveen información útil para distinguir anomalías y hacer seguimiento de evolución comparativa período a período que no es posible realizar con redes agregadas. Por otro lado, las redes agregadas manifiestan su utilidad en que: a) permiten captar mayor heterogeneidad de transiciones, lo cual es una línea de trabajo futuro en si misma; y, b) resultan aptas para el análisis de mediano plazo, aunque se debe prestar especial atención en la selección de la ventana temporal de agregación en términos de la cantidad de transiciones anuales a incluir y las características de los años seleccionados respecto del ciclo económico, considerando criterios que transparenten la elección.

En base a estos hallazgos y en función del objetivo de analizar la evolución a mediano plazo de los flujos de empleo intersectorial, se continúa el análisis sobre las redes de los períodos agregados.

4. Análisis de agrupamientos

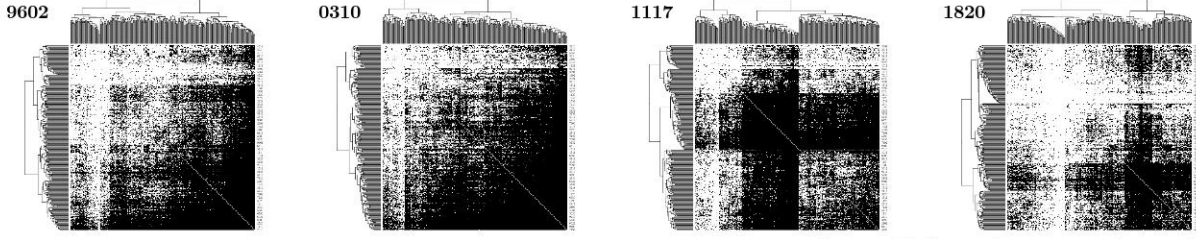
La conectividad entre sectores no sólo interesa a nivel de global. La identificación de grupos o comunidades de sectores más densamente conectados entre sí que con otros grupos y el resto de la red reviste especial importancia. En las redes analizadas de relaciones intersectoriales de flujos de empleo al nivel de detalle analizado (4 dígitos, CIIU Rev 3) se encontró que de los cambios de empleo observados entre sectores emergen estructuras persistentes de interacciones entre ellos. Sin embargo, dada la cantidad de sectores analizados ($V = 287$) interesa indagar acerca de la posibilidad de resumir estas interacciones en términos de agrupamientos y acerca de su estabilidad temporal.

Al introducir las representaciones alternativas de las matrices de transición en el Capítulo 2, en los mapas de calor ordenados por clustering jerárquico (ver Fig. 5.C2) se presentó como una posible forma de agrupar los sectores. Sin embargo, en redes ponderadas como las de flujos de empleo estos agrupamientos resultan diferentes si se analiza el grado de conectividad de los sectores en base a las matrices A binarias, como en la Fig. 6.B, que si se considera el volumen de intercambios que resumen los pesos de las matrices A ponderadas, como en la Fig. 6.C.

En trabajos previos se reportó la presencia de estructuras de centro-periferia en redes de flujos de empleo para el período 2009-2014 (De Raco y Semeshenko, 2019a y 2019b). Esta clase de meso-estructura se caracteriza por la existencia de grupos de sectores que interactúan entre sí y con el resto de la red (centro) y otros grupos de sectores que presentan interacciones con sectores del centro a la vez que prácticamente no interactúan entre sí. En la Fig. 17 se presentan los mapas de calor de las matrices binarias de adyacencias de los F_p períodos agregados ordenados por clustering jerárquico, en los que se observa la existencia de grupos de conectividad identificables en este sentido. Dado que estas matrices son cuadradas con los mismos nodos en filas y columnas, al reordenarlas se observan bloques más blancos sobre la diagonal que representan agrupamientos de sectores de baja o nula conectividad entre sí (periferia), que a su vez intercambian (resto de las franjas, con partes en blanco y en negro) con sectores de alta interconectividad (bloques en negro sobre la diagonal) que funcionan como centros que proveen conexión a la periferia.

Una forma convencional de caracterizar la organización de una red en términos de conectividad es a través de calcular la descomposición en k -cores del grafo correspondiente, que consiste en determinar los subgrafos que resulten maximales de orden k , es decir para los que cada nodo tenga al menos un grado k de conectividad. Así, el *coreness* de un sector es k si pertenece al k -core y no al $(k + 1)$ -core. Los sectores con el mismo coreness pueden compartir lazos entre sí y también con sectores con otro nivel de coreness. De esta forma, al referirse a esta métrica, cuanto más elevado el grado de coreness mayor resulta el nivel de conectividad de estos mismos. En la próxima sección se elabora sobre la utilización de esta medida en las redes bajo análisis.

Por otro lado, la búsqueda de estructuras modulares es un tema muy activo en la ciencia de redes dado que los hallazgos en esta área facilitan la comprensión del funcionamiento de sistemas de interacciones de diversos grados de complejidad, por ejemplo mediante la reducción del análisis a la interacción y evolución de los grupos identificados. Como se mencionó en el Capítulo 2, existe una gran variedad de métodos para la *detección de comunidades*. Para realizar este tipo de análisis sobre las redes de flujos de empleo se utiliza



Fuente: Elaboración propia en base a SIPA

Figura 17: Mapas de calor de la matrices binarias de adyacencias ordenados por clustering jerárquico aglomerativo con vinculación completa, períodos agregados. En negro (blanco) la presenencia (ausencia) de un lazo intersectorial. En todos los períodos se observa la presencia de un grupo de sectores que prácticamente no intercambian empleo entre sí (periferia, bloques en blanco sobre la diagonal) y grupos con distintos niveles de interconexión (centros, bloques en blanco sobre la diagonal con corredores en negro que conectan grupos de la periferia con los centros). En 0310 se observa la mayor conectividad. En contraste en 1820 la periferia es más extendida y el centro más reducido. En 1117 aparecen dos niveles de periferia y dos centros, uno bastante menos conectado que en otros períodos.

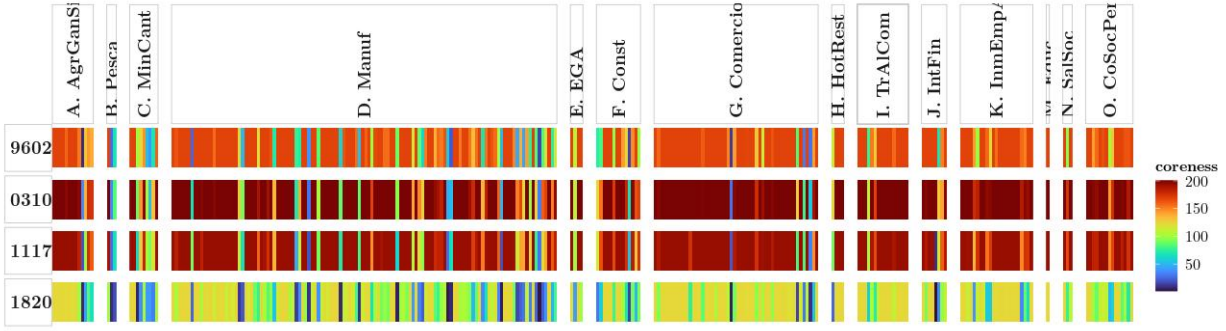
el algoritmo de Leiden (Traag, Waltman, and Eck 2019), un algoritmo aglomerativo en dos etapas que mejora el defecto en la generación de grupos internamente mal conectados del conocido algoritmo de Louvain (Blondel et al. 2008). Se seleccionó el algoritmo de Leiden sobre Louvain dado que garantiza la conectividad en la detección de comunidades mediante un paso de refinamiento de particiones y porque está diseñado originalmente para trabajar con grafos dirigidos y ponderados.

4.1. Coreness

La descomposición en k -cores de la red de flujos de empleo permite observar en forma resumida los sectores agrupados por niveles crecientes de conectividad. En las redes F_p bajo análisis, al ordenar los sectores según la clasificación de registro (CIU Rev 3) se observa una distribución de coreness muy heterogénea hacia el interior de las grandes divisiones por letra (Fig. 18). No obstante, como se mencionó en el Capítulo 1 y se observa en la Fig. 4, la distribución de la dotación de empleo a 4 dígitos también resulta heterogénea, aunque en este caso la correlación (Pearson) entre dotación y coreness se encuentra en el rango 0.22-0.29 por lo que no pareciera condicionar la conectividad de los sectores medidos por el nivel de coreness. En la misma Fig. 4 se observa un efecto de nivel entre los distintos períodos. En los períodos 0310 y 1117 los mismos sectores aparecen con niveles de coreness más elevados respecto de 9602, mientras que en 1820 los sectores presentan niveles sensiblemente inferiores a los demás períodos. Esta observación se corresponde con la evolución de la conectividad entre los períodos analizados, como se observa en las estadísticas globales de Dens, $\langle k \rangle$, $\langle k^{in} \rangle$, $\langle k^{out} \rangle$ y $\langle \text{Core} \rangle$ en el Cuadro 2. Por otro lado, la correlación del coreness de cada sector entre períodos (Cuadro 4) sugiere que los sectores conservan relativamente su nivel entre períodos contiguos mientras que para los períodos más distantes el parecido decae. Esta característica se observa en la Fig. 18 como la persistencia entre períodos del patrón de *cambio de colores* (v.g.: la alternancia horizontal).

Cuadro 4: Correlación (Pearson) de coreness sectorial por períodos.

	9602	0310	1117
0310	0,93	...	...
1117	0,88	0,94	...
1820	0,87	0,88	0,93



Fuente: Elaboración propia en base a SIPA

Figura 18: k -coreness en redes dirigidas de flujos, F_p , períodos agregados. Sectores ordenados según la clasificación de actividades económicas CIIU Rev 3 agrupados por letra. Se observa un incremento del grado de conectividad medido por el k -coreness en los períodos 0310 y 1117 respecto de 9602, en línea con la evolución de la actividad económica, y niveles sensiblemente inferiores en 1820, en parte posiblemente relacionados con la menor cantidad de transiciones que capta este período. Escala de intensidad de colores azul (baja) a rojo (alta).

4.2. Comunidades

Mediante otro tipo de análisis más enfocado en la identificación de grupos con mayor interacción entre sus integrantes, los algoritmos de detección de comunidades como Leiden identifican estructuras basadas en el indicador de *modularidad* como criterio de calidad. Este indicador mide la fuerza de una partición de la red en módulos o grupos mediante la densidad de lazos dentro de los grupos y la baja densidad de lazos entre grupos.

En las redes de flujos de empleo esto equivale a identificar grupos de sectores entre los cuales se observan mayor cantidad de transiciones de empleo que entre miembros de un mismo grupo y el resto de la red. Esta característica resulta útil en sí misma para identificar sectores que pueden estar compartiendo empleo en función de las habilidades de los trabajadores que cambian de empleo. Por otro lado, también resulta útil conocer su persistencia y su evolución en el tiempo a fin de identificar aquellos sectores que determinan los agrupamientos principales y sectores que pueden asociarse a distintas comunidades a lo largo del tiempo.

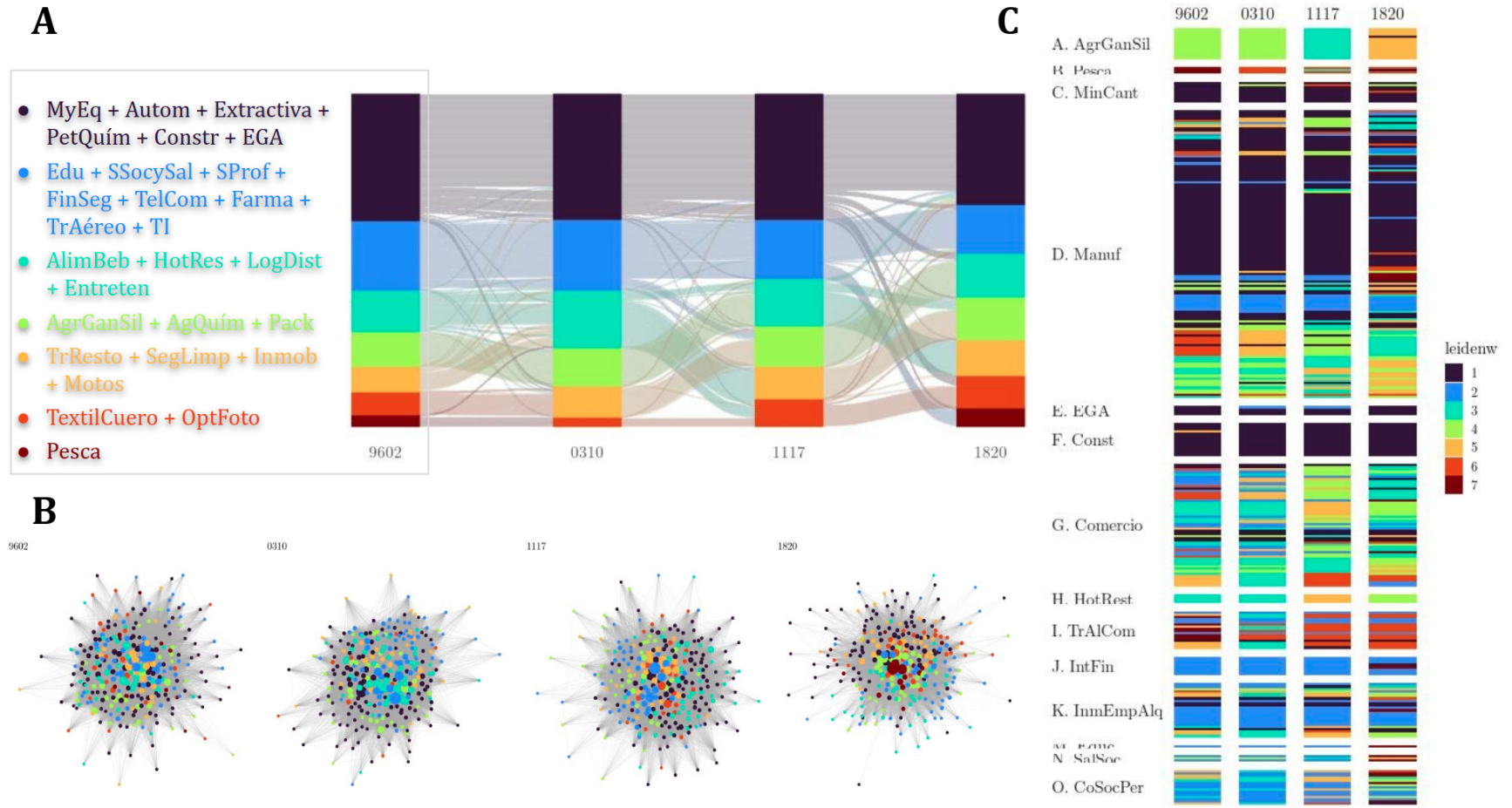


Figura 19: Comunidades detectadas con Leiden en redes de flujos de empleo, F_p , períodos agregados. Escala de colores de las 7 comunidades detectadas. (A) Evolución temporal con principales grupos de actividades identificadas. (B) Comunidades en grafos (layout LGL, tamaño de los nodos proporcional a dotación sectorial de empleo) e histograma de pesos. (C) Comunidades en el orden lexicográfico CIHU Rev 3.

Las comunidades detectadas se enumeran en orden decreciente de la cantidad de sectores que agrupan. En el Cuadro 5 se resume la cantidad de sectores por su clasificación de letra agrupados en cada comunidad y período. En el Cuadro 7 se resume la proporción de flujos de empleo dentro y entre comunidades detectadas en cada período, mientras que en el Cuadro 6 se presenta la misma desagregación en términos de la dotación relativa de empleo asociada a los sectores agrupados. Por último, en el Cuadro 8 se presenta la comparación de las particiones entre períodos según los criterios de información mutua normalizada (NMI) e índice de Rand ajustado (ARI). En la sección 5.1.2 del Material Suplementario se encuentran tablas resumidas de los 10 sectores con mayor y con menor empleo en las comunidades detectadas en cada período.

A fin de ayudar al análisis y la comprensión de los agrupamientos, un aspecto relevante a considerar refiere a la distribución no uniforme de sectores según su agrupamiento por letra de clasificación. En la Fig. 4 del Capítulo 2 se observa que la variabilidad del tamaño de la dotación de empleo a nivel de 4 dígitos del clasificador se caracteriza porque la mayoría de las clasificaciones sectoriales por letra (ver Cuadro 1 en la sección 2.1 del Capítulo 2) se desagregan en pocos sectores, siendo las más numerosas D-Manufacturas (122 sectores), G-Comercio (52) y K-Actividades inmobiliarias, empresariales y de alquiler (23 sectores). Como resultaba esperable debido a esta característica, las letras que se desagregan en una mayor cantidad de sectores a 4 dígitos se encuentran representadas en más comunidades mientras que las menos desagregadas tienden a concentrarse en una comunidad, según se observa en el Cuadro 5. En particular, las actividades agrupadas en las letras A (agricultura, ganadería y silvicultura), B (pesca), C (minería), E (energía, gas y agua), F (construcción), H (hoteles y restaurantes), J (finanzas y seguros), M (enseñanza) y N (servicios sociales y de salud), se desagregan en pocos sectores y tienden a concentrarse en bloque en una o dos comunidades en los distintos períodos.

Como se mencionó en la sección 2.2, el algoritmo Leiden utiliza para detectar comunidades el criterio de calidad denominado “modularidad” que evalúa de forma aglomerativa y en etapas los agrupamientos en función de los lazos que ocurren *dentro* de los miembros de cada grupo y los que relacionan a miembros del grupo *entre* grupos. Desde el punto de vista de los lazos que interconectan sectores y los flujos de empleo que los recorren (pesos), en el Cuadro 7 se presenta un resumen de los flujos sumariados por comunidad, $c \in \mathcal{C}$, en cada período. Esta información puede servir para indagar acerca de la identificación de grupos de sectores en términos de la estructura mencionada de centro-periferia mediante la observación del valor del ratio de las transiciones ocurridas dentro de la comunidad (diagonal principal de cada matriz de transición, $a_{ii}^p \in A^p$, en este Cuadro) respecto de las ocurridas entre este grupo y los demás (resto de fila, $a_{i\bullet}^p - a_{ii}^p$, y resto de columna, $a_{\bullet j}^p - a_{ii}^p$, de interacciones con las demás comunidades). Se calculan los indicadores CP_c^+ y CP_c^- como:

$$CP_c^+ = \frac{a_{ii}}{a_{\bullet j} - a_{ii}} = \frac{\text{flujos dentro de } c}{\text{flujos netos hacia } c},$$

para el caso de los flujos de ingreso (suma horizontal) y

$$CP_c^- = \frac{a_{ii}}{a_{i\bullet} - a_{ii}} = \frac{\text{flujos dentro de } c}{\text{flujos netos desde } c},$$

para los flujos de egreso (suma vertical) en esta Cuadro. Estos indicadores permiten asociar al grupo de sectores dentro de una comunidad como formando parte de un centro cuando $CP_c^-, CP_c^+ > 1$ o de la periferia cuando presenta valores inferiores a la unidad.

En la Fig. 19.A se observa que en los períodos analizados se identificaron entre 6 (0310 y 1117) y 7 comunidades (9602 y 1820). Esta visualización permite apreciar que entre cada período algunos sectores mantuvieron la asignación de comunidad del período inicial a lo largo del tiempo, en particular algunos pertenecientes a las dos primeras comunidades (1, en azul, y 2, en celeste) que se comentan más adelante, mientras que una proporción no menor de sectores migraron hacia distintas comunidades, incluso más de una vez en los períodos analizados. En particular 1820 presenta una estructura de comunidades diferente a los demás períodos como se observa en los Cuadros 5 y 6, y se visualiza en la Fig. 19.

En los grafos de la Fig. 19.B, realizados con layout de fuerzas dirigidas LGL, se dificulta la identificación visual de las comunidades debido a la densidad del único componente conexo de las redes como se analizó en el Capítulo 3, que le otorga la forma de “bola de pelos,” y a que las comunidades detectadas contienen sectores tanto del centro como de la periferia.

En la Fig. 19.C se presenta la asignación de comunidades para cada sector según el orden natural de la clasificación de registro CIU Rev 3, y resulta a la vista evidente que una buena parte de estos no conserva la comunidad detectada inicialmente en 9602 aunque en este caso permite distinguir que ciertos grupos de sectores cercanos en la misma rama de actividad se mueven en bloque a lo largo del tiempo. Esto implica que se mantienen como miembros de la misma comunidad aunque cambien de comunidad de referencia. Por ejemplo, las actividades de las letras A (Agricultura, Ganadería y Silvicultura), B (Pesca), F (Construcción) y H (Hoteles y Restaurantes) se conservan prácticamente en bloque durante todos los períodos analizados.

Las primeras dos comunidades revisten particular importancia. En el Cuadro 5 se observa que capturan más de la mitad de los sectores en prácticamente todos los períodos excepto 1820, que presenta una estructura diferente debido a las características particulares del mismo según se analizó en el Capítulo 3, en relación con la (menor) cantidad de años que agrupa y repercute en la menor carga de flujos de empleo y densidad de lazos. A la vez, resultan las comunidades más estables y que, como se verá, representan en conjunto al *core* productivo de la economía. Juntas incluyen algo más de la mitad de los sectores y del total de dotación de empleo (Cuadro 6).

La primera comunidad (1, azul en la Fig. 19), la más numerosa y estable de todas, agrupa poco más de un tercio de los sectores en todos los períodos excepto en 1820 y se podría considerar el *núcleo industrial* de la red, con una gran cantidad de intercambios de empleo dentro de su comunidad (entre 11.9 y 12.7 % del total de flujos) así como con las demás comunidades, como se observa en el Cuadro 7. En promedio la dotación de los sectores que la componen suma 20 % del total del empleo privado. En 9602 agrupa a 109 sectores (un tercio de los sectores) que incluyen a: F-Construcción, (13 sectores) prácticamente en su totalidad; G-Comercio (7 sectores), relacionados a la venta de materiales e insumos para la construcción, y servicios de plomería e instalaciones de gas; D-Manufacturas (71 sectores), que incluyen fabricación de vehículos y autopartes, de maquinaria agropecuaria, de productos metálicos para uso de infraestructura, industrias básicas de hierro y acero, de productos de plástico y metal, de muebles, entre otros; E-Electricidad, Gas y Agua (EGA, 4 sectores) en su totalidad; y, C-Explotación de minas y canteras (EMyC, 9 sectores) en su totalidad. En Material Suplementario se presenta una selección de los sectores de la comunidad por orden decreciente en la dotación de empleo (ver sección 5.1.2 en Material Suplementario). Esta comunidad captura consistentemente en el tiempo a los sectores de la construcción (F), explotación de minas y canteras (C), EGA (E) y las industrias (D) y sectores de comercio (G) mencionados.

La segunda comunidad más relevante y estable (2, celeste en la Fig. 19), agrupa aproximadamente un quinto del total de sectores y acumula en promedio el 30 % del total del empleo privado, con excepción de 1820 que parte de los sectores con mayor cantidad de empleo (enseñanza, servicios sociales y de salud, y asociaciones) se reagrupa en la comunidad 7, como surge de los Cuadros 5 y 6. Los sectores de esta comunidad constituyen lo que podría denominarse el *núcleo de servicios* de la red, con alto dinamismo del empleo, captando la mayor proporción de flujos de empleo del sistema dentro de la propia comunidad y participando en intercambios de empleo con sectores de las restantes comunidades también en una elevada proporción. En 9602 agrupa a 60 sectores (un quinto del total) que incluyen a: M-Enseñanza (1 único sector); N-Servicios de salud y sociales (2 sectores); K-Actividades inmobiliarias, empresariales y de alquiler (11 sectores), servicios profesionales, informáticos, y otros; O-Otras actividades de servicios comunitarios, sociales y personales (8 sectores), servicios de entretenimiento y medios de comunicación; I-Transporte, almacenamiento y comunicaciones (4 sectores), servicios de telecomunicaciones, publicidad, turismo y transporte aéreo; J-Intermediación financiera (8 sectores), en su totalidad; G-Comercio al por mayor y al por menor, y otros (13 sectores), comercio minorista de farmacia, cosmética y perfumería, de muebles y otros; D-Manufacturas (13 sectores), industria farmacéutica, fabricación de productos de limpieza y perfumería, maquinaria de oficina e informática, edición e impresión, entre otros. Dentro de esta comunidad aparecen consistentemente en el tiempo la mayor parte de los sectores de finanzas y seguros (J), servicios de telecomunicaciones, de transporte aéreo y de turismo (I), servicios profesionales, empresariales, informáticos, y publicidad (K); y fabricación de maquinaria de oficina, de aeronaves, de edición, impresión y grabaciones (D). Otros sectores que participan en todos los períodos excepto en 1820, son: industria farmacéutica, fabricación de productos de limpieza y perfumería (D); y sectores de comercio (G) mencionados.

En el Cuadro 7 se observa que mientras que las dos comunidades descritas registran transiciones de empleo interindustriales tanto hacia *adentro* de la comunidad respectiva (v.g.: sobre la diagonal principal de las

Cuadro 5: Cantidad de sectores por comunidad (Leiden) y período.

Per./Com.	Sector (#)															Total (%)
	A	B	C	D	E	F	G	H	I	J	K	M	N	O	Total	
9602																
1			9	71	4	13	7		1		4				109	38,0
2				13			13		4	8	11	1	2	8	60	20,9
3				11			16	4			2			3	36	12,5
4	13			12			3				1		1		30	10,5
5				2		1	6		5		4			4	22	7,7
6				12			7				1				20	7,0
7		3		1					6						10	3,5
Total	13	3	9	122	4	14	52	4	16	8	23	1	3	15	287	100,0
0310																
1			8	72	3	14	6		1		4				108	37,6
2				12	1		11		4	8	12	1	2	10	61	21,3
3				7			22	4	7		5			5	50	17,4
4	13		1	14			3				1		1		33	11,5
5				16			10				1				27	9,4
6		3		1					4						8	2,8
Total	13	3	9	122	4	14	52	4	16	8	23	1	3	15	287	100,0
1117																
1			8	71	3	14	5		2		5				108	37,6
2				11	1		5		4	8	10	1	2	9	51	17,8
3	13	1		17			8				1		1		41	14,3
4				16			18				1				35	12,2
5				5			9	4			5			5	28	9,8
6		2	1	2			7		10		1			1	24	8,4
Total	13	3	9	122	4	14	52	4	16	8	23	1	3	15	287	100,0
1820																
1	1	1	7	54	4	14	7		1	1	4			1	95	33,1
2				13			5		4	6	10			4	42	14,6
3				19			17				1			1	38	13,2
4			1	9			14	4			5			4	37	12,9
5	12			14			3				1		1		31	10,8
6		2	1	8			4		11		1			1	28	9,8
7				5			2			1	1	1	2	4	16	5,6
Total	13	3	9	122	4	14	52	4	16	8	23	1	3	15	287	100,0

Cuadro 6: Participación del empleo sectorial por comunidad (Leiden) y período.

Per./Com.	Sector (%)														
	A	B	C	D	E	F	G	H	I	J	K	M	N	O	Total
9602															
1			0,9	10,5	1,1	5,8	1,5		0,0		1,2				21,0
2				2,2			3,7		2,1	3,5	4,3	6,6	4,6	4,1	31,1
3				5,0			6,9	3,1			2,5			0,5	18,0
4	6,5			2,3			1,0				0,0		0,0		9,9
5				0,1		0,0	1,4		6,1		5,0			2,1	14,7
6				2,8			1,3				0,1				4,2
7		0,3		0,2					0,7						1,2
Total	6,5	0,3	0,9	23,1	1,1	5,8	15,8	3,1	8,9	3,5	13,1	6,6	4,6	6,7	100,0
0310															
1			1,0	9,4	0,7	6,7	0,9		0,0		1,2				19,9
2				2,1	0,2		3,3		2,0	2,5	5,8	7,1	4,2	4,3	31,5
3				2,9			9,0	3,7	6,1		6,9			2,3	30,9
4	6,2		0,0	3,6			1,0				0,1		0,0		10,8
5				2,9			3,0				0,0				5,9
6		0,3		0,2					0,5						0,9
Total	6,2	0,3	1,0	21,0	0,9	6,7	17,1	3,7	8,5	2,5	14,0	7,1	4,3	6,5	100,0
1117															
1			1,1	9,4	0,7	6,9	0,9		0,6		1,2				20,8
2				1,7	0,2		1,9		1,9	2,3	5,8	7,2	4,8	4,3	30,2
3	5,4	0,0		4,1			2,8				0,1		0,0		12,5
4				2,4			4,9				0,0				7,4
5				2,3			4,7	4,3			6,3			1,6	19,0
6		0,2	0,0	0,2			2,9		6,4		0,1			0,4	10,1
Total	5,4	0,2	1,1	20,0	1,0	6,9	18,0	4,3	8,9	2,3	13,4	7,2	4,8	6,3	100,0
1820															
1	0,0	0,0	1,1	5,9	1,1	6,7	1,8		0,0	0,0	0,9			0,5	18,0
2				1,1			1,2		1,9	1,6	5,9			0,4	12,1
3				2,5			4,6				0,0			0,1	7,2
4			0,0	3,9			6,9	4,4			6,0			0,7	22,0
5	5,5			2,6			0,9				0,0		0,0		9,1
6		0,2	0,1	1,2			1,2		6,9		0,2			0,6	10,4
7				1,3			1,6			0,7	0,1	8,0	5,4	4,1	21,2
Total	5,5	0,2	1,2	18,4	1,1	6,7	18,2	4,4	8,9	2,3	13,3	8,0	5,5	6,3	100,0

matrices de cada período) como *entre* sectores de otras comunidades, las restantes comunidades presentan menor proporción de transiciones dentro de las comunidades en relación con la proporción de cambios de empleo con las demás comunidades.

Como se observa en el diagrama aluvial de la Fig. 19.A, las restantes comunidades resultan menos estables en el tiempo en el sentido que los sectores incluidos se reorganizan hacia otra comunidad o hacia distintas comunidades con el correr del tiempo, a veces en bloque, por lo que pierde sentido el seguimiento de éstas. Sin embargo, es posible identificar algunos subgrupos de sectores que persisten apareciendo juntos en comunidades distintas en el tiempo. A continuación se mencionan los grupos más relevantes en orden decreciente de relevancia respecto de la dotación relativa de empleo (aproximada) que acumulan los sectores involucrados.

- El primero de estos grupos, que podría denominarse “Alimentos y bebidas, alojamiento y entretenimientos” participa en las comunidades 3 en 9602 y 0310, 5 en 1117 y 4 1820 en el Cuadro 5. Dentro de este subgrupo se incluyen la totalidad de los sectores de hotelería y restaurantes, que incluyen servicios de alojamiento (H); comercio de alimentos y bebidas (G); elaboración de alimentos y bebidas, fabricación de pastas, lácteos, panadería, confitería (D); servicios de esparcimiento, espectáculos y diversión (O).
- Un segundo subgrupo, que podría denominarse “Agro,” relaciona a las actividades vinculadas a la explotación agrícola-ganadera, silvicultura y servicios vinculados, que se acompañan en las comunidades 4 en 9602 y 0310, 3 en 1117 y 5 1820 en el Cuadro 5. En este grupo se incluyen prácticamente la totalidad de los sectores de producción agrícola-ganadera y silvicultura (A); elaboración de aceites, de vinos, de productos de molinería, de alimentos para animales, fabricación de agroquímicos y plaguicidas, de abonos, aserrado y cepillado de madera (D); y servicios veterinarios (N).
- Un tercer subgrupo, que podría denominarse “Transporte automotor y correos,” relaciona actividades relativas al transporte automotor, la eliminación de desperdicios, y correos, que se acompañan en las comunidades 5 en 9602, 3 en 0310, 6 en 1117 y 1820 en el Cuadro 5, que aparece vinculado a veces con actividades de logística y distribución de cargas.
- Otro subgrupo bien identificado son las actividades vinculadas a la “Pesca” (letra B completa, sector de industria pesquera en letra D y sectores de transportes marítimo y fluvial en letra I) y parte de la logística de cargas, que se encuentran en distintas comunidades a lo largo de los períodos aunque siempre aparecen juntas (en Cuadro 5, comunidad 7 en 9602, 6 en 0310, 1117 y 1820), lo cual resulta esperable en función de que la vinculación proviene de intercambios de empleo entre sectores que, aún siendo de clases (letra) distintas según la clasificación, resultan actividades conexas o complementarias a la producción pesquera.
- Un caso particular que relaciona actividades aparentemente dispares, que podría denominarse “Textil y cuero, óptica y fotografía,” es la combinación de las actividades de la industria textil y de cueros, que aparecen juntas con actividades de óptica y fotografía (en Cuadro 5, comunidad 6 en 9602, 5 en 0310, 4 en 1117 y 3 en 1820)
- Finalmente, los servicios de “Seguridad, limpieza e inmobiliarios” (K) son otro grupo que aparecen acompañándose en tándem en distintas comunidades a lo largo del tiempo (en Cuadro 5, comunidad 5 en 9602, 3 en 0310, 4 en 1117 y 5 en 1820).

Mediante la inspección realizada de las comunidades y grupos identificados, se concluye que la organización de los sectores productivos lograda por las comunidades detectadas sobre las redes agregadas de flujos de empleo con el algoritmo de Leiden resulta comprensible y razonable. En este sentido, los hallazgos de organización sectorial emergentes de las estructuras de transiciones interindustriales de empleo encontrados identifican dos grandes estructuras de elevada conectividad intersectorial hacia dentro y afuera de estas comunidades que sugieren su condición de centros: un núcleo industrial y un núcleo de servicios. El resto de las comunidades presenta mayor interacción relativa con aquellas comunidades centrales que con el resto por lo que se consideran distintas clases de periferia. Adicionalmente se identifican subgrupos de sectores que se mantienen cercanos y migran entre comunidades con el correr del tiempo.

Cuadro 7: Asignaciones de comunidades (Leiden) y proporción de flujos dentro y entre comunidades.

Com./Per.	Comunidad								Indicadores	
	1	2	3	4	5	6	7	8	CP_c^{++}	CP_c^{-}
9602										
1	12,7	3,9	2,5	1,3	3,1	0,4	0,4	24,4	1,1	1,1
2	3,9	15,1	3,5	0,9	3,2	0,7	0,3	27,7	1,1	1,2
3	2,9	4,4	7,2	1,4	2,7	0,6	0,2	19,5	0,7	0,6
4	1,3	0,9	1,4	3,7	0,7	0,1	0,1	8,3	0,8	0,8
5	3,0	3,4	2,6	0,6	4,6	0,4	0,3	14,9	0,4	0,4
6	0,5	0,9	0,5	0,1	0,4	1,3	0,0	3,8	0,6	0,5
7	0,3	0,2	0,2	0,1	0,2	0,0	0,6	1,6	0,5	0,7
Total	24,7	28,8	17,9	8,2	14,9	3,6	1,9	100,0		
0310										
1	11,3	3,6	4,8	1,5	0,7	0,2		22,0	0,9	1,1
2	4,2	15,7	6,3	1,2	1,6	0,2		29,1	1,1	1,2
3	5,8	7,4	14,2	2,3	1,8	0,3		31,8	0,9	0,8
4	1,8	1,0	2,2	4,2	0,2	0,0		9,5	0,8	0,8
5	0,9	1,9	1,6	0,2	1,7	0,0		6,5	0,4	0,4
6	0,2	0,1	0,3	0,1	0,0	0,4		1,2	0,6	0,6
Total	24,2	29,8	29,3	9,5	6,1	1,2	0,0	100,0		
1117										
1	12,6	3,6	2,0	0,9	2,8	2,3		24,1	1,1	1,1
2	3,5	13,7	1,6	1,7	3,3	1,7		25,6	1,0	1,2
3	2,1	1,5	6,1	0,5	1,5	1,2		12,9	0,9	0,9
4	1,2	2,2	0,5	2,6	1,2	0,7		8,4	0,5	0,4
5	3,2	4,4	1,8	1,5	7,3	1,8		20,0	0,7	0,6
6	2,0	1,6	1,1	0,4	1,2	2,8		9,1	0,4	0,5
Total	24,5	27,0	13,0	7,7	17,2	10,6	0,0	100,0		
1820										
1	11,9	2,2	0,7	2,4	1,6	2,2	1,2	22,2	1,2	1,2
2	2,2	9,1	1,1	2,3	0,9	1,6	2,3	19,5	0,9	0,9
3	1,0	1,3	2,1	1,3	0,3	0,6	0,9	7,4	0,5	0,4
4	2,5	2,4	1,2	7,6	1,6	1,8	2,2	19,2	0,7	0,6
5	1,3	0,8	0,2	1,5	6,6	0,7	0,5	11,6	1,1	1,3
6	2,2	1,3	0,4	1,7	1,0	3,2	0,8	10,5	0,4	0,4
7	0,8	1,7	0,4	1,2	0,4	0,6	4,2	9,5	0,5	0,8
Total	22,0	18,8	6,1	18,0	12,4	10,6	12,1	100,0		

Comparación de comunidades. Para la comparación de la similitud en las membresías de los sectores en las comunidades detectadas en las redes F_p se utilizan los indicadores de información mutua normalizada (NMI) y Rand ajustado (ARI). Ambos indicadores comparan los grupos identificados en un rango de variación entre 0 (completamente distintos) y 1 (agrupamientos iguales). En el Cuadro 8 se observa que con ambos indicadores se encuentran resultados concordantes y, en línea con las correlaciones de las adyacencias binarias entre períodos en la Fig. 10, muestran que las similitudes se mantienen relativamente entre períodos contiguos y decaen hacia los más alejados, aunque en particular 1820 presenta una estructura de comunidades diferente a los demás períodos. Sin embargo, como se analizó en esta sección, la asignación de sectores cambió en el tiempo aunque varios sectores siguieron siendo identificados en subgrupos dentro de distintas comunidades en los cuatro períodos analizados. Esto sugiere que un análisis de conectividad a una escala a nivel más local, como por ejemplo de vecindarios de primer orden o vecindarios ampliados de segundo orden, podría aportar a explicar aspectos estructurales más permanentes que los identificados con el algoritmo Leiden. Posiblemente, parte de la degradación de las particiones identificadas con estos indicadores se relacione con la elevada densidad característica de las redes de flujos de empleo, para las que podría resultar útil explorar métodos de detección de comunidades superpuestas en las que resulte posible asignar sectores a más de una comunidad.

Conclusiones. En el análisis de las redes de flujos de empleo mediante la descomposición de k -core y la detección de comunidades con el algoritmo de Leiden se encuentra la presencia de meso-estructuras en las que son detectables grupos de sectores en base a características estructurales de conectividad (lazos dirigidos) en el caso del coreness, o en base a los flujos de empleo observados (lazos ponderados dirigidos) con las comunidades detectadas usando el algoritmo de Leiden. Mediante ambos métodos se identifican estructuras con cierto grado de permanencia en el tiempo y también diferencias que evidencian su evolución. En ambos casos se detecta el período 1820 como particularmente distinto al resto.

En primer lugar, las estructuras identificadas con coreness indican la presencia de conectividad relativa persistente en el tiempo, con elevadas correlaciones entre períodos. Esta característica de conectividad de los sectores no presenta una relación lineal con la dotación de empleo, lo cual sugiere que puede resultar una métrica que aporta información útil de conectividad no asociada a estadísticas comúnmente utilizadas como el nivel de empleo.

En segundo lugar, el análisis de las comunidades detectadas en los distintos períodos permite distinguir agrupamientos que asocian sectores lejanos según la clasificación estadística (v.g.: actividades con letra distinta) aunque cercanos en términos de la actividad productiva en el sentido más amplio de la organización productiva. En los agrupamientos obtenidos se distinguen claramente dos comunidades estables que conforman cada una un grupo de sectores de alta interacción interna y con el resto de la red que funcionan como centros, un núcleo industrial y un núcleo de servicios, que en conjunto participan en más de la mitad de los flujos de empleo interindustrial registrados y acumulan más de la mitad de la dotación total de empleo privado. Por otro lado, se detectan entre cuatro y cinco comunidades más pequeñas e inestables que se reorganizan entre períodos y funcionan como periferia, con mayor cantidad de intercambios de empleo con los dos núcleos identificados que con el resto de la red. Adicionalmente, dentro de estas comunidades se detectan subgrupos más estables que se asocian en distintas comunidades con el correr del tiempo.

Cuadro 8: Comparación de comunidades (Leiden) por períodos. Información Mutua Normalizada (NMI) e Índice de Rand Ajustado (ARI).

Período	NMI			ARI		
	9602	0310	1117	9602	0310	1117
0310	0,73	...	...	0,76	...	...
1117	0,66	0,70	...	0,72	0,76	...
1820	0,55	0,52	0,57	0,54	0,51	0,55

5. Discusión final

En esta tesis se explora la aplicación de enfoques de redes complejas al análisis de flujos de empleo a un nivel de detalle poco explorado que revela una historia de conectividad interindustrial con estructuras persistentes, que correlacionan positivamente con la actividad económica agregada y en el mediano plazo evolucionan reacomodando el entramado productivo. Aunque no es la primera vez que se utilizan metodologías interdisciplinarias como las aquí utilizadas para analizar flujos de empleo, la aplicación a datos del mercado de trabajo formal argentino en un período tan abarcativo y desagregado es novedosa. Recientemente se analizaron redes de flujos de empleo con metodologías similares para períodos agregados (De Raco and Semeshenko 2019a, 2019b; De Raco, Tumini, and Ernst 2019), la disponibilidad de datos de transiciones de empleo interanuales permitió analizar redes de interacciones a una escala de agregación temporal inexplorada. Esto permitió comprender los efectos de la agregación de estas matrices transiciones interanuales y encontrar el valor complementario de la observación a distintas escalas temporales para indagar acerca de la evolución de las estructuras latentes de relaciones entre industrias. En este punto, los hallazgos presentados en esta tesis respecto de la persistencia, propiedades y características de las redes de flujos de empleo intersectorial proveen información novedosa y complementaria a la metodología tradicional de análisis de flujos de empleo, más enfocada en resumir el comportamiento de creación y destrucción de empleo y las tasas de ingreso y salida de empleo por sectores.

Resultados, respuestas y expectativas. Al comienzo de esta tesis, en el Capítulo 1 se plantearon preguntas que trazaron la línea conductora de la investigación aquí realizada y se deseaba responder en su desarrollo.

La primera pregunta indagaba acerca de las características estructurales que podían extraerse de los intercambios de empleo entre industrias, mientras que la segunda indagaba acerca de la estabilidad en el tiempo de las características encontradas. Los resultados y el análisis presentados en el Capítulo 3 aportan elementos para responder estas preguntas. Se encontró que las relaciones que de allí surgen presentan una estructura de elevada conectividad, densas, con coeficiente de clustering y reciprocidad elevadas, que se organizan en un solo componente conexo donde los sectores se interrelacionan mediante caminos cortos que permiten verificar la propiedad de “mundo pequeño.” Más adelante, en el Capítulo 4 se lograron identificar estructuras de organización sectorial de centro y periferia, que aportan además información de aspectos funcionales de las interacciones entre sectores mediadas por el empleo.

Respecto de la estabilidad de las estructuras encontradas, los resultados presentados en ambos Capítulos permiten brindar una respuesta parte afirmativa en términos generales. Se encontró que a pesar de fluctuaciones en la densidad de interconexión, la estructura de un gran componente conexo con la propiedad de mundo pequeño persiste en las redes agregadas y está presente en la mayor parte de las redes interanuales, que por otro lado presentan correlaciones elevadas de las estructuras de conectividad, con la excepción de años de crisis macroeconómicas. En este sentido, en el análisis también se evidenciaron que con el correr del tiempo también surgen cambios en la organización de la estructura de interconexiones que caracterizan un período, aquí referidos como “regímenes” de movilidad laboral, en correlación con la actividad económica agregada. Así, en el análisis realizado la historia que cuentan los flujos de empleo interindustriales aparece hilvanada con el hilo de la conectividad, que teje y desteje el entramado productivo al ritmo y en la dirección de la actividad económica.

A partir de estos hallazgos de la organización general que se identifica en los flujos de empleo analizados se avanzó sobre aspectos más particulares, continuando con las preguntas planteadas, hacia la búsqueda de estructuras intermedias en la organización de las actividades productivas emergentes de estas interacciones. ¿Es posible identificar agrupamientos relevantes de los flujos de empleo entre industrias? ¿Qué información se puede extraer de estas interacciones respecto al entramado productivo? El análisis realizado en el Capítulo 4 avanzó en dirección de brindar parcialmente respuesta a estas preguntas. Allí se encontró que es posible identificar sectores productivos agrupados en comunidades a partir de la optimización del criterio de modularidad (ver sección 2.2) que resultan interpretables en términos de la articulación de todas las actividades productivas bajo una estructura de centro y periferia. Aunque no todas las comunidades detectadas presentan estabilidad en el tiempo, resulta posible distinguir dos grandes comunidades persistentes que capturan la mitad de las actividades productivas y del empleo privado que funcionan como núcleos (uno con predom-

inancia de actividades industriales y otro con predominancia de actividades de servicios) de alto dinamismo del empleo dentro y entre comunidades.

Por último, al inicio se planteó la pregunta acerca del potencial aporte de las metodologías de redes complejas como marco de análisis para extraer conocimientos nuevos a partir de los datos de registros administrativos, dado que los flujos de empleo intersectoriales se pueden conceptualizar naturalmente como un grafo. En esta tesis desde principio a fin se encuentra una respuesta afirmativa a esta pregunta. El abordaje del análisis de flujos de empleo entre industrias a partir de la conectividad enriquece el estudio del fenómeno de la movilidad laboral aportando nuevas analogías y métricas para su mejor comprensión.

Por otro lado, también se plantearon expectativas sobre los resultados a obtener, que ahora es menester retomar para cotejar la medida en que se satisficieron. Desde el punto de vista del análisis económico, se esperaba que los resultados aportaran a mejorar la comprensión de aspectos económicos y productivos de los flujos interindustriales de empleo. Se esperaba encontrar a partir del análisis de estos flujos de empleo algunos aspectos de carácter estructural y otros de carácter coyuntural, como en una descomposición de factores de largo y de corto plazo. En este sentido, el análisis llevado a cabo permitió distinguir claramente aspectos estructurales de organización de las redes de intercambio de empleo entre sectores productivos, como también características de estas estructuras asociadas a determinados períodos de tiempo como se reportó en especial en la sección 3.2 del Capítulo 3. En base a ello se podría decir que aquellas expectativas se lograron cumplir, aunque con un detalle que no permitió discriminar plenamente los factores de largo y de corto plazo lo cual requeriría posiblemente la adecuación de las aproximaciones metodológicas. En particular, respecto a la componente de corto plazo, que en esta tesis se relegó en función de atender a la evolución de más largo plazo planteada como objetivo de investigación, quedó sólo parcialmente atendida y para su estudio posiblemente requiera de mayor precisión acerca de los factores coyunturales que se desee incluir en el análisis.

Respecto de las expectativas planteadas desde el punto de vista de la minería de datos, se esperaba que la utilización de las técnicas de redes complejas aquí utilizadas sirvieran para sentar las bases de un marco de referencia orientado al análisis de la conectividad interindustrial, para la extracción de conocimiento de datos de registro de empleo a máximo nivel de detalle. Estas expectativas se consideran en principio cumplidas dado que con los resultados obtenidos se logró extraer conocimiento original y útil, y aportar métodos accesibles y replicables para reproducir, mejorar y extender el análisis aquí propuesto.

Por último, se plantearon expectativas desde el punto de vista de la aplicación de las metodologías utilizadas como herramienta útil para el diseño y el soporte de políticas públicas. Estas expectativas podrían considerarse en proceso de cumplimiento dado que los desarrollos realizados en la aplicación de metodologías y el análisis sobre las transiciones de empleo forman parte de una colaboración activa con funcionarios y analistas del MTEySS. En este sentido, la evaluación del nivel de cumplimiento de estas expectativas se efectuará en conjunto con los responsables de la gestión pública y forma parte de la agenda de investigación y trabajo por delante.

Futuros pasos. En cuanto a los caminos que se abren a partir del análisis y la aplicación de herramientas de redes complejas al estudio de flujos de empleo, en el transcurso de la elaboración de esta tesis se identificaron numerosas áreas de interés y preguntas de investigación que pueden abordarse mediante los métodos aquí explorados y una variedad de métodos adicionales. Un paso futuro inmediato, que constituye una línea de investigación en curso, está orientado a explorar distintos métodos de reducción de las redes de flujos de empleo para caracterizar los corredores de empleo y estudiar su evolución y las implicancias de la granularidad de los datos en las características informacionales de las redes de flujos de empleo (Semeshenko and De Raco 2021). Asociado a ello, otros temas de interés en los que también se está avanzando se orientan a profundizar el análisis de las comunidades y las interrelaciones de cada sector con el resto y a realizar el análisis presentado sobre las redes de parentesco de habilidades, en una línea emparentada con (O’Clery, Flaherty, and Kinsella 2021) y (Straulino, Landman, and O’Clery 2021). Otro aspecto de interés se relaciona con el análisis más detallado de la conectividad local de sectores o comunidades específicas, para el cual los métodos de análisis redes sociales focalizados en el estudio de vecindarios y estructuras recurrentes y significativas de relaciones básicas presentes en subgrafos (*network motifs*) resultan apropiados.

Extensiones. Utilizando los mismos datos analizados en esta tesis, de empleo privado registrado, existen

algunas inmediatas que surgen de ampliarlos con metadatos relevantes que surgen del SIPA asociados a nodos sectoriales, como por ejemplo tamaño de las empresas, salarios medios, género y rangos de edad, a fin de analizar covariantes que permiten enriquecer la información que surge de los flujos de empleo analizados. En esta línea, algunos de estos aspectos se exploraron en una etapa previa al desarrollo de esta tesis en (De Racó, Tumini, and Ernst 2019) aplicados a nuevos sectores tecnológicos mediante la exploración del parentesco de habilidades para Argentina 20019-2014. Otra extensión posible en base a datos de registro se orienta a la comparación internacional de redes de parentesco de habilidades de la Argentina con redes de parentesco de habilidades recientemente publicadas para Alemania, Suecia, Holanda e Irlanda. En términos de otra clase de datos de registro, Argentina publicó recientemente la Muestra longitudinal de empleo registrado (MLER, 2018) que consiste en una muestra de historias laborales de empleo registrado 1996-2015 que permite el análisis de trayectorias laborales y redes sectoriales con memoria temporal.

Por otro lado, en esta tesis se analizaron transiciones de empleo privado registrado como un sistema cerrado. Sin embargo, la movilidad laboral en términos más generales como se describe en la Fig. 1, el universo de las transiciones de empleo incluye al sector público y al sector privado informal para los cuales se requieren otro tipo de datos que pueden presentar distinta granularidad, calidad y disponibilidad. En este sentido ampliado, una extensión natural del análisis aquí realizado refiere a la utilización de la Encuesta Permanente de Hogares (EPH) publicada por el Instituto Nacional de Estadística y Censos (INDEC) como fuente de datos para el análisis de las transiciones de empleo. Esta fuente permite identificar las relaciones entre las actividades productivas y también las ocupaciones de las personas, tanto del empleo formal como del informal. En particular, en el Instituto Interdisciplinario de Economía Política (IIEP) de la Universidad de Buenos Aires y CONICET, donde me desempeño como investigador, hay actualmente una línea de investigación en curso dedicada al análisis de las transiciones de empleo en base a la EPH y la MLER.

Epílogo

El objetivo planteado originalmente para este trabajo de tesis se orientaba a extraer información relevante de un conjunto de datos de alta definición como son los registros administrativos de empleo del SIPA, el cual se encontraba poco explorado debido al desafío que plantea la complejidad del nivel de detalle para su comprensión y análisis. El plan de tesis aprobado en 2019 tuvo como principal impronta extraer información de datos históricos que nunca se habían analizado con los enfoques desarrollados en esta tesis para poder evaluar su aporte a futuro, tanto para el análisis como para el diseño de políticas públicas. Entonces estaba muy activa la discusión del “Futuro del Trabajo” y los desafíos de la automatización como las difíciles transiciones y aprendizajes que deberían enfrentar, tarde o temprano, tanto las empresas en su rol dentro del sector productivo como las personas en su rol de trabajadoras. Y el futuro no tuvo contemplaciones.

Lejos de imaginar la disrupción, los reacomodamientos sistémicos generalizados y la aceleración de profundos cambios avizorados en las relaciones laborales producto de la pandemia desencadenada en 2020, esta tesis fue planeada con el espíritu de detectar patrones estables como también evolutivos de la dinámica de los cambios de empleo y del entramado productivo emergente utilizando métodos de análisis de redes complejas. A partir de estos acontecimientos los hechos indican que el trabajo y la organización productiva evolucionarán en los próximos años en formas difíciles de predecir a partir de la historia reciente. Sin embargo este trabajo se hace para aportar, más que una predicción, una forma de observar y analizar el mismo fenómeno con otros ojos, con herramientas que permitan identificar estructuras emergentes de interacciones que seguirán existiendo aunque posiblemente organizarán las actividades productivas de otras maneras que las conocidas.

A lo largo del camino recorrido en la elaboración de esta tesis se lograron cooperaciones y colaboraciones que enriquecieron notablemente la comprensión y la discusión acerca de los flujos de empleo desde enfoques de redes complejas. Tanto funcionarios públicos como investigadores aportaron a enfocar las preguntas y los métodos aquí explorados, como también a depurar la comprensión de problemáticas asociadas al funcionamiento y la estructura del entramado productivo emergente de las interacciones entre sectores industriales mediadas por cambios de empleo.

Considero que con esta tesis se logra un aporte original para la extracción de conocimiento de las transiciones de empleo a gran nivel de detalle y un avance en la comprensión de la conectividad interindustrial argentina. Es mi deseo que resulte también de utilidad para la mejor comprensión de las transiciones de empleo, así como para el diseño de políticas públicas.

Referencias

- Blondel, Vincent D, Jean-Loup Guillaume, Renaud Lambiotte, and Etienne Lefebvre. 2008. “Fast Unfolding of Communities in Large Networks.” *Journal of Statistical Mechanics: Theory and Experiment* 2008 (10): P10008. <https://doi.org/10.1088/1742-5468/2008/10/P10008>.
- Castillo, Victoria, V Cesa, A Filippo, S Rojo Brizuela, D Schleser, and G Yoguel. 2002. “Dinámica Del Empleo y Rotación de Empresas: La Experiencia En El Sector Industrial de Argentina Desde Medios de Los años Noventa.” CEPAL. <https://www.cepal.org/es/publicaciones/4833-dinamica-empleo-rotacion-empresas-la-experiencia-sector-industrial-argentina>.
- Castillo, Victoria, Marta Novick, Sofía Rojo, and Gabriel Yoguel. 2005. “Trayectorias Laborales y Rotación Del Empleo: Restricciones Para El Desarrollo de Competencias técnicas.” In *Actas Del 7mo Congreso de Especialistas En Estudios Del Trabajo: “Nuevos Escenarios En El Mundo Del Trabajo: Rupturas y Continuidades*.
- Csardi, Gabor, and Tamas Nepusz. 2006. “The Igraph Software Package for Complex Network Research.” *InterJournal Complex Systems*: 1695. <https://igraph.org>.
- De Racó, Sergio A., and Viktoriya Semeshenko. 2019a. “Labor Mobility and Industrial Space in Argentina.” *Journal of Dynamics & Games* 6 (2): 107. <https://doi.org/10.3934/jdg.2019008>.
- . 2019b. “The Network Structure of Inter-Industry Labor Mobility in Argentina.” In. Geneva: ILO.
- De Racó, Sergio A., Lucía Tumini, and Christophe Ernst. 2019. “Parentesco de Habilidades y Nuevos Sectores Económicos. Un análisis Exploratorio Para La Argentina.” In *Cambio Tecnológico y Futuro Del Trabajo*. Organización Internacional del Trabajo. https://www.ilo.org/wcmsp5/groups/public/---americas/---rolima/---ilo-buenos_aires/documents/publication/wcms_734829.pdf.
- Donnat, Claire, and Susan Holmes. 2018. “Tracking Network Dynamics: A Survey Using Graph Distances.” *The Annals of Applied Statistics* 12 (2): 971–1012. <https://doi.org/10.1214/18-AOAS1176>.
- Fortunato, Santo. 2010. “Community Detection in Graphs.” *Physics Reports* 486 (3-5): 75–174. <https://doi.org/10.1016/j.physrep.2009.11.002>.
- Gianelle, Carlo. 2010. “Essays on Labor Mobility and Network Analysis.” {PhD} dissertation. Sienna University. https://phdeconomics-colne.wp.unisi.it/wp-content/uploads/sites/127/2015/01/Gianelle_PhD_2010.pdf.
- . 2014. “Discovering the Regional Small World of Labour Mobility. Evidence from Linked Employer–Employee Data.” *Regional Studies* 48 (7): 1263–78. <https://doi.org/10.1080/00343404.2012.697993>.
- Guerrero, Omar A., and Robert L. Axtell. 2013. “Employment Growth Through Labor Flow Networks.” *PLOS ONE* 8 (5): e60808. <https://doi.org/10.1371/journal.pone.0060808>.
- Hidalgo, César A., and Ricardo Hausmann. 2009. “The Building Blocks of Economic Complexity.” *Proceedings of the National Academy of Sciences* 106 (26): 10570–75. <https://doi.org/10.1073/pnas.0900943106>.
- Keynes, John Maynard. 1936. *The General Theory of Employment, Interest, and Money*.
- López, Eduardo, Omar Guerrero, and Robert L. Axtell. 2015. “The Network Picture of Labor Flow.” *arXiv:1507.00248 [Cond-Mat, Physics:physics]*, July. <http://arxiv.org/abs/1507.00248>.
- Neffke, Frank, and Martin Henning. 2013. “Skill Relatedness and Firm Diversification.” *Strategic Management Journal* 34 (3): 297–316. <https://doi.org/10.1002/smj.2014>.

- Neffke, Frank, Anne Otto, and Antje Weyh. 2017. “Inter-Industry Labor Flows.” *Journal of Economic Behavior & Organization* 142 (October): 275–92. <https://doi.org/10.1016/j.jebo.2017.07.003>.
- O’Clery, Neave, Eoin Flaherty, and Stephen Kinsella. 2021. “Modular Structure in Labour Networks Reveals Skill Basins.” *arXiv:1909.03379 [Econ, q-Fin]*, April. <http://arxiv.org/abs/1909.03379>.
- RStudio Team. 2020. *RStudio: Integrated Development Environment for r*. Boston, MA: RStudio, PBC. <http://www.rstudio.com/>.
- Semeshenko, Viktoriya, and Sergio A. De Raco. 2021. “Analysis of the Evolution of Labor Market Flows in Argentina.” *Proceedings 50 JAIIO-AGRANDA, Simposio Argentino de Ciencia de Datos y GRANdes DATos*, en imprenta.
- Straulino, Daniel, Mattie Landman, and Neave O’Clery. 2021. “A Bi-Directional Approach to Comparing the Modular Structure of Networks.” *EPJ Data Science* 10 (1): 13. <https://doi.org/10.1140/epjds/s13688-021-00269-8>.
- Traag, V. A., L. Waltman, and N. J. van Eck. 2019. “From Louvain to Leiden: Guaranteeing Well-Connected Communities.” *Scientific Reports* 9 (1): 5233. <https://doi.org/10.1038/s41598-019-41695-z>.
- Wickham, Hadley, Mara Averick, Jennifer Bryan, Winston Chang, Lucy McGowan, Romain François, Garrett Grolemund, et al. 2019. “Welcome to the Tidyverse.” *Journal of Open Source Software* 4 (43): 1686. <https://doi.org/10.21105/joss.01686>.
- Wills, Peter, and François G. Meyer. 2020. “Metrics for Graph Comparison: A Practitioner’s Guide.” *PLOS ONE* 15 (2): e0228728. <https://doi.org/10.1371/journal.pone.0228728>.

Anexos

5.1. Material suplementario

5.1.1. Correlaciones interanuales

Cuadro 9: Correlaciones (Pearson) de matrices de adyacencias ponderadas.

	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018
1997	0,79	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
1998	0,83	0,76	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
1999	0,74	0,73	0,79	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
2000	0,66	0,65	0,69	0,70	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
2001	0,60	0,53	0,63	0,57	0,58	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
2002	0,67	0,67	0,71	0,71	0,71	0,62	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
2003	0,55	0,54	0,58	0,58	0,54	0,46	0,60	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
2004	0,73	0,72	0,76	0,75	0,70	0,60	0,77	0,63	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
2005	0,76	0,73	0,80	0,78	0,74	0,59	0,79	0,64	0,87	...	...	...	...	...	...	...	...	...	...	...	...	...	...
2006	0,66	0,69	0,66	0,67	0,61	0,47	0,66	0,53	0,70	0,74	...	...	...	...	...	...	...	...	...	...	...	...	...
2007	0,74	0,72	0,77	0,77	0,71	0,57	0,76	0,63	0,86	0,92	0,75	...	...	...	...	...	...	...	...	...	...	...	...
2008	0,70	0,71	0,75	0,75	0,70	0,57	0,75	0,62	0,83	0,89	0,72	0,93	...	...	...	...	...	...	...	...	...	...	...
2009	0,67	0,66	0,71	0,71	0,66	0,54	0,72	0,59	0,80	0,86	0,70	0,90	0,89	...	...	...	...	...	...	...	...	...	...
2010	0,68	0,67	0,71	0,72	0,67	0,55	0,74	0,60	0,81	0,87	0,72	0,90	0,89	0,90	...	...	...	...	...	...	...	...	...
2011	0,66	0,63	0,69	0,69	0,64	0,52	0,69	0,57	0,77	0,82	0,66	0,85	0,84	0,82	0,84	...	...	...	...	...	...	...	...
2012	0,69	0,67	0,72	0,73	0,69	0,56	0,74	0,59	0,81	0,86	0,70	0,89	0,89	0,87	0,90	0,86	...	...	...	...	...	...	...
2013	0,69	0,66	0,72	0,72	0,73	0,56	0,72	0,59	0,79	0,85	0,68	0,87	0,88	0,85	0,88	0,84	0,91	...	...	...	...	...	...
2014	0,67	0,67	0,70	0,71	0,66	0,53	0,71	0,59	0,79	0,83	0,73	0,87	0,87	0,85	0,89	0,83	0,92	0,91	...	...	...	...	...
2015	0,66	0,63	0,71	0,69	0,65	0,55	0,71	0,57	0,77	0,82	0,67	0,86	0,86	0,86	0,88	0,84	0,90	0,89	0,90	...	...	...	...
2016	0,63	0,61	0,67	0,66	0,62	0,51	0,67	0,54	0,74	0,78	0,64	0,82	0,83	0,81	0,84	0,80	0,86	0,86	0,86	0,87	...	...	...
2017	0,30	0,30	0,31	0,34	0,30	0,24	0,32	0,26	0,35	0,37	0,30	0,39	0,38	0,38	0,39	0,37	0,40	0,40	0,43	0,40	0,39	...	...
2018	0,60	0,60	0,65	0,67	0,62	0,53	0,67	0,55	0,74	0,79	0,64	0,83	0,84	0,83	0,85	0,82	0,88	0,88	0,89	0,90	0,87	0,40	...
2019	0,53	0,53	0,56	0,57	0,54	0,46	0,60	0,48	0,64	0,69	0,56	0,75	0,77	0,76	0,78	0,75	0,80	0,81	0,81	0,83	0,80	0,38	0,88

5.1.2. Comunidades por período

5.1.2.1. Cuadros resumen por período Los Cuadros 10, 11, 12 y 13 presentan información por período para cada comunidad detectada con un resumen de la cantidad de sectores, la suma y la participación en el empleo total que agrupan, junto con el grados de conectividad promedio, $\langle k \rangle$, y el grado de conectividad promedio de los vecinos de primer orden, $\langle k_{nn} \rangle$.

Cuadro 10: Comunidades Leiden. Período 9602. Resumen por comunidad.

Com.	# Sec.	Empleo	% Emp.	$\langle k \rangle$	$\langle k_{nn} \rangle$
1	109	795.381	20.5	292	177
2	60	1.180.182	31	340	197
3	36	682.585	18	368	210
4	30	375.892	9.6	296	177
5	22	558.028	14.7	376	210
6	20	158.198	4.1	292	178
7	10	44.052	1	277	168
	287	3.794.318			

Cuadro 11: Comunidades Leiden. Período 0310. Resumen por comunidad.

Com.	# Sec.	Empleo	% Emp.	$\langle k \rangle$	$\langle knn \rangle$
1	108	1.002.085	19.4	357	208
2	61	1.585.725	31.3	394	222
3	50	1.555.961	31	452	246
4	33	545.874	10.9	364	210
5	27	295.963	5.8	352	203
6	8	47.096	0.9	311	183
	287	5.032.704			

Cuadro 12: Comunidades Leiden. Período 1117. Resumen por comunidad.

Com.	# Sec.	Empleo	% Emp.	$\langle k \rangle$	$\langle knn \rangle$
1	108	1.323.473	20.6	349	204
2	51	1.923.362	29.8	367	208
3	41	792.616	12.5	360	205
4	35	468.243	7.2	330	190
5	28	1.211.218	19.1	427	235
6	24	643.397	9.7	404	226
	287	6.362.309			

Cuadro 13: Comunidades Leiden. Período 1820. Resumen por comunidad.

Com.	# Sec.	Empleo	% Emp.	$\langle k \rangle$	$\langle knn \rangle$
1	95	1.140.217	17.8	192	126
2	42	768.107	11.7	207	131
3	38	455.911	7	201	131
4	37	1.396.547	22.2	283	171
5	31	577.950	9.1	230	146
6	28	661.300	10.4	271	168
7	16	1.348.729	21.3	314	188
	287	6.348.761			

5.1.2.2. Detalle de comunidades núcleo En los Cuadros 17 y 21 se detallan por período los 10 sectores más relevantes en términos de empleo de las dos comunidades más estables, identificadas como *núcleo industrial* y *núcleo de servicios*. Para cada una de estas dos comunidades se presenta información del código de sector, la clasificación de letra de pertenencia y la descripción de la actividad, junto con datos de el grado de conectividad del sector, k , el grado de conectividad de los vecinos de primer orden, knn y el empleo promedio del período.

Cuadro 14: Comunidad 1. Período 9602. Sectores por comunidad, top-10 según dotación de empleo.

Código	Sec.	Sec. Abrev.	Sector Descripción	Indicador		
				<i>k</i>	<i>knn</i>	Empleo
4520	F	Const	Construcción de edificios y sus partes y obras de ingeniería civil	562	283	189.894
2520	D	Manuf	Fabricación de productos de plástico	525	276	34.068
2899	D	Manuf	Fabricación de productos elaborados de metal n.c.p.	518	275	31.743
3430	D	Manuf	Fabricación de partes; piezas y accesorios para vehículos automotores y sus motores	490	262	27.885
4013	E	EGA	Distribución de energía eléctrica	391	230	25.798
5236	G	Comercio	Venta al por menor de materiales de construcción, artículos de ferretería, pinturas, cristales y espejos, y artículos para la decoración	515	270	25.393
7421	K	InmEmpAlq	Servicios de arquitectura e ingeniería y servicios conexos de asesoramiento técnico	543	281	25.173
2710	D	Manuf	Industrias básicas de hierro y acero	436	245	22.612
3410	D	Manuf	Fabricación de vehículos automotores	422	238	19.981
1120	C	MinCant	Actividades de servicios relacionadas con la extracción de petróleo y gas, excepto las actividades de prospección	462	255	17.215

Cuadro 15: Comunidad 1. Período 0310. Sectores por comunidad, top-10 según dotación de empleo.

Código	Sec.	Sec. Abrev.	Sector Descripción	Indicador		
				<i>k</i>	<i>knn</i>	Empleo
4520	F	Const	Construcción de edificios y sus partes y obras de ingeniería civil	568	287	289.073
2520	D	Manuf	Fabricación de productos de plástico	537	277	44.554
2899	D	Manuf	Fabricación de productos elaborados de metal n.c.p.	543	278	42.508
7421	K	InmEmpAlq	Servicios de arquitectura e ingeniería y servicios conexos de asesoramiento técnico	555	285	38.716
3430	D	Manuf	Fabricación de partes; piezas y accesorios para vehículos automotores y sus motores	530	277	36.199
1120	C	MinCant	Actividades de servicios relacionadas con la extracción de petróleo y gas, excepto las actividades de prospección	480	259	31.320
4013	E	EGA	Distribución de energía eléctrica	438	248	27.598
2710	D	Manuf	Industrias básicas de hierro y acero	508	270	24.593
5143	G	Comercio	Venta al por mayor de madera, materiales de construcción, artículos de ferretería y materiales para plomería e instalaciones de gas	525	274	19.868
3410	D	Manuf	Fabricación de vehículos automotores	479	263	19.626

Cuadro 16: Comunidad 1. Período 1117. Sectores por comunidad, top-10 según dotación de empleo.

Sector				Indicador		
Código	Sec.	Sec. Abrev.	Descripción	<i>k</i>	<i>knn</i>	Empleo
4520	F	Const	Construcción de edificios y sus partes y obras de ingeniería civil	565	286	377.821
2520	D	Manuf	Fabricación de productos de plástico	528	277	54.454
2899	D	Manuf	Fabricación de productos elaborados de metal n.c.p.	534	277	49.912
7421	K	InmEmpAlq	Servicios de arquitectura e ingeniería y servicios conexos de asesoramiento técnico	538	274	44.127
3430	D	Manuf	Fabricación de partes; piezas y accesorios para vehículos automotores y sus motores	505	268	43.056
1120	C	MinCant	Actividades de servicios relacionadas con la extracción de petróleo y gas, excepto las actividades de prospección	467	256	42.976
6010	I	TrAlCom	Servicio de transporte ferroviario	402	249	35.318
4013	E	EGA	Distribución de energía eléctrica	411	246	33.919
3410	D	Manuf	Fabricación de vehículos automotores	447	244	31.430
5143	G	Comercio	Venta al por mayor de madera, materiales de construcción, artículos de ferretería y materiales para plomería e instalaciones de gas	509	270	26.363

Cuadro 17: Comunidad 1. Período 1820. Sectores por comunidad, top-10 según dotación de empleo.

Sector				Indicador		
Código	Sec.	Sec. Abrev.	Descripción	<i>k</i>	<i>knn</i>	Empleo
4520	F	Const	Construcción de edificios y sus partes y obras de ingeniería civil	535	278	359.053
5236	G	Comercio	Venta al por menor de materiales de construcción, artículos de ferretería, pinturas, cristales y espejos, y artículos para la decoración	448	248	54.352
2520	D	Manuf	Fabricación de productos de plástico	430	240	50.292
2899	D	Manuf	Fabricación de productos elaborados de metal n.c.p.	453	253	45.485
1120	C	MinCant	Actividades de servicios relacionadas con la extracción de petróleo y gas, excepto las actividades de prospección	344	202	41.002
7421	K	InmEmpAlq	Servicios de arquitectura e ingeniería y servicios conexos de asesoramiento técnico	475	264	40.738
4013	E	EGA	Distribución de energía eléctrica	279	187	37.356
9000	O	CoSocPer	Eliminación de desperdicios y aguas residuales, saneamiento y servicios similares	332	198	30.774
5143	G	Comercio	Venta al por mayor de madera, materiales de construcción, artículos de ferretería y materiales para plomería e instalaciones de gas	397	228	26.746
4531	F	Const	Ejecución y mantenimiento de instalaciones eléctricas, electromecánicas y electrónicas	426	236	26.077

Cuadro 18: Comunidad 2. Período 9602. Sectores por comunidad, top-10 según dotación de empleo.

Sector				Indicador		
Código	Sec.	Sec. Abrev.	Descripción	<i>k</i>	<i>knn</i>	Empleo
8000	M	Educ	Enseñanza	534	276	248.973
8510	N	SalSoc	Servicios relacionados con la salud humana	527	276	159.833
9100	O	CoSocPer	Servicios de asociaciones n.c.p.	537	277	118.700
7410	K	InmEmpAlq	Servicios jurídicos y de contabilidad, teneduría de libros y auditoría; asesoramiento en materia de impuestos; estudios de mercados y realización de encuestas de opinión pública; asesoramiento empresarial y en materia de gestión	557	284	91.969
6420	I	TrAlCom	Servicios de telecomunicaciones	507	268	57.886
6521	J	IntFin	Servicios de las entidades financieras bancarias	462	251	53.971
7499	K	InmEmpAlq	Servicios empresariales n.c.p.	551	282	32.506
6610	J	IntFin	Servicios de seguros	503	270	30.304
5239	G	Comercio	Venta al por menor en comercios especializados n.c.p.	529	276	28.531
5235	G	Comercio	Venta al por menor de muebles, artículos de mimbre y corcho, colchones y somieres, artículos de iluminación y artefactos para el hogar	502	268	24.185

Cuadro 19: Comunidad 2. Período 0310. Sectores por comunidad, top-10 según dotación de empleo.

Sector				Indicador		
Código	Sec.	Sec. Abrev.	Descripción	<i>k</i>	<i>knn</i>	Empleo
8000	M	Educ	Enseñanza	552	285	356.464
8510	N	SalSoc	Servicios relacionados con la salud humana	543	280	194.430
9100	O	CoSocPer	Servicios de asociaciones n.c.p.	550	282	148.420
7410	K	InmEmpAlq	Servicios jurídicos y de contabilidad, teneduría de libros y auditoría; asesoramiento en materia de impuestos; estudios de mercados y realización de encuestas de opinión pública; asesoramiento empresarial y en materia de gestión	567	287	121.779
7499	K	InmEmpAlq	Servicios empresariales n.c.p.	568	287	92.040
6420	I	TrAlCom	Servicios de telecomunicaciones	531	278	69.527
6521	J	IntFin	Servicios de las entidades financieras bancarias	472	259	52.212
5235	G	Comercio	Venta al por menor de muebles, artículos de mimbre y corcho, colchones y somieres, artículos de iluminación y artefactos para el hogar	539	276	41.460
5231	G	Comercio	Venta al por menor de productos farmacéuticos, cosméticos, de perfumería, instrumental médico y odontológico y artículos ortopédicos	497	263	38.504
6610	J	IntFin	Servicios de seguros	532	278	35.179

Cuadro 20: Comunidad 2. Período 1117. Sectores por comunidad, top-10 según dotación de empleo.

Sector				Indicador		
Código	Sec.	Sec. Abrev.	Descripción	<i>k</i>	<i>knn</i>	Empleo
8000	M	Educ	Enseñanza	545	282	460.485
8510	N	SalSoc	Servicios relacionados con la salud humana	541	279	277.326
9100	O	CoSocPer	Servicios de asociaciones n.c.p.	544	278	191.627
7499	K	InmEmpAlq	Servicios empresariales n.c.p.	557	283	128.323
7410	K	InmEmpAlq	Servicios jurídicos y de contabilidad, teneduría de libros y auditoría; asesoramiento en materia de impuestos; estudios de mercados y realización de encuestas de opinión pública; asesoramiento empresarial y en materia de gestión	563	286	122.200
6420	I	TrAlCom	Servicios de telecomunicaciones	510	266	76.598
6521	J	IntFin	Servicios de las entidades financieras bancarias	457	256	64.824
5231	G	Comercio	Venta al por menor de productos farmacéuticos, cosméticos, de perfumería, instrumental médico y odontológico y artículos ortopédicos	499	267	56.850
6610	J	IntFin	Servicios de seguros	546	278	47.302
7220	K	InmEmpAlq	Servicios de consultores en informática y suministros de programas de informática	517	268	46.591

Cuadro 21: Comunidad 2. Período 1820. Sectores por comunidad, top-10 según dotación de empleo.

Sector				Indicador		
Código	Sec.	Sec. Abrev.	Descripción	<i>k</i>	<i>knn</i>	Empleo
7499	K	InmEmpAlq	Servicios empresariales n.c.p.	525	276	137.167
7410	K	InmEmpAlq	Servicios jurídicos y de contabilidad, teneduría de libros y auditoría; asesoramiento en materia de impuestos; estudios de mercados y realización de encuestas de opinión pública; asesoramiento empresarial y en materia de gestión	521	276	106.272
6420	I	TrAlCom	Servicios de telecomunicaciones	386	226	74.541
6521	J	IntFin	Servicios de las entidades financieras bancarias	310	184	66.348
7220	K	InmEmpAlq	Servicios de consultores en informática y suministros de programas de informática	444	249	59.823
5011	G	Comercio	Venta de vehículos automotores nuevos, excepto motocicletas	416	241	48.146
7290	K	InmEmpAlq	Actividades de informática n.c.p.	404	230	41.195
3410	D	Manuf	Fabricación de vehículos automotores	274	174	26.837
5159	G	Comercio	Venta al por mayor de máquinas, equipo y materiales conexos n.c.p.	434	244	24.836
6590	J	IntFin	Servicios financieros excepto los de la banca central y las entidades financieras	352	208	22.567

5.2. Anexo metodológico

5.2.1. Estadísticas de red

Aquí se provee una breve reseña de las estadísticas descriptivas utilizadas para el análisis de las redes analizadas de flujos de empleo.

Métricas globales. Dentro de las métricas básicas que refieren al grafo completo asociado a una red se encuentran el *orden* y el *tamaño*, que refieren respectivamente a la cantidad de nodos, V , y la cantidad de lazos, E , de la red, que informan acerca de la escala de complejidad que puede adquirir la misma. Una medida relevante respecto del total de lazos es la *densidad* del grafo, calculada como el cociente de lazos observados respecto de la cantidad de lazos posibles dada la cantidad de nodos.

Al analizar la conectividad del grafo completo interesa conocer características de la distribución de los lazos, por ejemplo en términos de los *caminos más cortos* que existan para llegar de un nodo, v_i , a otro cualquiera, v_j , que permiten cuantificar distancias dentro de la red. Una medida representativa considera el *largo promedio de los caminos más cortos* (LPC, o *average path length*, APL), que puede calcularse sobre grafos dirigidos y no dirigidos. Otra medida de distancia utilizada es el *diámetro* de la red, que es el máximo camino más corto y se expresa comúnmente en términos de la cantidad de “pasos” requeridos para atravesar la red entre los nodos más alejados en el sentido de la conectividad.

Otras métricas globales que brindan información relevante del grafo refieren a la *reciprocidad* de los lazos en un grafo dirigido, calculada como la cantidad de lazos que ocurren en ambas direcciones respecto de las posibles entre cada par de nodos. Esta métrica permite estimar el error de aproximar un grafo dirigido mediante un grafo no dirigido, sobre el cual existen una variedad de resultados de aplicación disponibles.

Una métrica fundamental para la organización modular dentro de una red refiere a las características que presentan los lazos al considerar relaciones entre tres nodos, o tríadas. En particular, el *coeficiente global de clustering* se calcula como el ratio de la cantidad de tríadas cerradas (v.g.: triángulos)¹⁷ observadas en la red respecto de la cantidad total de tríadas del grafo. Este coeficiente brinda información respecto del grado en que los nodos se agrupan entre sí e incide en la densidad del grafo. Este coeficiente tiene su versión local, calculada respecto de las tríadas cerradas en que participa cada nodo.

Por último, otra métrica global utilizada en esta tesis es el *coeficiente de asortatividad*, calculado como el coeficiente de correlación de Pearson del grado de conectividad k (definido debajo) entre pares de nodos. Esta métrica puede tomar valores entre -1 y 1 e indica en qué medida los nodos se conectan con nodos de grado similar, en cuyo caso será positiva y cercana a 1 cuanto mayor sea esta asociación¹⁸ o, viceversa, será negativa y cercana a -1 cuanto mayor la diferencia de conectividad entre nodos, en este caso.

Métricas locales. Dentro de las estadísticas relativas a los nodos y sus vecinos, el *grado*, k , de un nodo es la medida más básica y representa la cantidad de lazos en los que éste participa. En términos de las redes analizadas representa la cantidad de sectores con los que intercambia empleo cada sector de actividad. Dado que estas redes poseen una dirección, los lazos pueden ir en una o ambas direcciones con lo que el grado total se compone de la suma de las conexiones por flujos de entrada y por los flujos de salida, es decir $k_i = k_i^{in} + k_i^{out}$.

Otras métricas de conectividad local utilizadas en esta tesis relacionadas con el grado refieren a los entornos de primero (vecinos directos) y segundo orden (vecinos directos de sus vecinos directos) de cada nodo.

El *vecindario de primer orden* de un nodo es igual al número de nodos con que está conectado directamente, independientemente de la dirección, por lo que para grafos no dirigidos brinda información provista por el grado.

¹⁷En rigor la cantidad de tríadas cerradas es equivalente a tres veces la cantidad de triángulos observados, ya que éstos se computan a partir de cada nodo de la tríada.

¹⁸En la literatura de análisis de redes sociales este comportamiento se conoce como homofilia, preferencia por asociación con otros con atributos similares.

El *grado promedio de vecinos cercanos*, *knn*, se calcula como el promedio de los *grados* de los nodos que conforman el vecindario de primer orden, con lo que se considera la dirección de sus conexiones. Esta métrica cuantifica la densidad de conexiones de los vecindarios inmediatos de un nodo.

El grado de *coreness* o *k-coreness* de un nodo se calcula a partir de la descomposición de *k-core* del grafo que consiste en encontrar para cada nodo el subgrafo maximal en el cual posee como mínimo un grado de conectividad de orden *k*. Así, el *coreness* de un nodo es *k* si pertenece al *k-core* pero no pertenece al $(k + 1)$ -core. Esta métrica, utilizada en el Capítulo 4 para mostrar agrupamientos por conectividad, es usada comúnmente para organizar una red en grados crecientes de conectividad de modo que los subconjuntos de nodos con mayor nivel de *coreness* resultan asociados a la noción de “centro” mientras que los conjuntos de nodos con menor nivel se asocian con la “periferia.”

El grado pertenece a la familia de medidas de **centralidad**, que miden la importancia de cada nodo $V_i \in V$ dentro de la red. Otras medidas utilizadas de esta familia son la *centralidad de intermediación* (*betweenness*, Freeman, 1977), que cuantifica la cantidad de caminos cortos que pasan por un nodo v_i ; la *centralidad de cercanía* (*closeness*, Bavelas, 1950), que representa la distancia media entre un nodo cualquiera y los restantes dentro de la red analizada.

5.3. Clasificaciones sectoriales y nodos

5.3.1. Sectores CIIU Rev 3

En el Cuadro 22 se detallan los sectores CIIU Rev 3 incluidos con datos del empleo promedio en los períodos agregados analizados.

Cuadro 22: Actividades a cuatro dígitos CIIU Rev 3. Descripción y empleo promedio.

Codigo	Descripcion	Empleo promedio			
		1996-2002	2003-2010	2011-2017	2018-2020
111	Cultivo de cereales, oleaginosas y forrajeras	49.083	65.608	68.946	64.829
112	Cultivo de hortalizas, legumbres, flores y plantas ornamentales	10.930	14.848	19.283	18.921
113	Cultivo de frutas -excepto vid para vinificar- y nueces	28.252	35.765	42.169	41.802
114	Cultivos industriales, de especias y de plantas aromáticas y medicinales	47.011	51.820	53.409	49.862
115	Producción de semillas y otras formas de propagación de cultivos agrícolas	4.165	4.058	7.391	9.534
121	Cría de ganado y producción de leche , lana y pelos	75.188	83.563	84.914	80.803
122	Producción de granja y cría de animales, excepto ganado	6.858	8.602	10.848	10.885
141	Servicios agrícolas	15.240	33.092	47.064	64.457
142	Servicios pecuarios, excepto los veterinarios	1.925	2.470	2.807	2.684
150	Caza y captura de animales vivos, repoblación de animales de caza y servicios conexos	40	95	108	101
201	Silvicultura	1.825	1.744	1.406	1.099
202	Extracción de productos forestales	3.820	4.656	4.451	4.218
203	Servicios forestales	3.382	4.292	3.669	2.788
501	Pesca y recolección de productos marinos	9.440	14.452	12.767	12.166
502	Explotación de criaderos de peces, granjas piscícolas y otros frutos acuáticos (acuicultura)	71	86	75	60
503	Servicios para la pesca	499	381	380	370
1110	Extracción de petróleo crudo y gas natural	8.939	11.089	16.648	21.317

Cuadro 22: Actividades a cuatro dígitos CIIU Rev 3. Descripción y empleo promedio. *(continued)*

Codigo	Descripcion	1996-2002	2003-2010	2011-2017	2018-2020
1120	Actividades de servicios relacionadas con la extracción de petróleo y gas, excepto las actividades de prospección	17.215	31.320	42.976	41.002
1310	Extracción de minerales de hierro	...	...	...	...
1320	Extracción de minerales metalíferos no ferrosos, excepto minerales de uranio y torio	...	...	...	...
1411	Extracción de rocas ornamentales	573	564	454	383
1412	Extracción de piedra caliza y yeso	833	886	1.053	1.042
1413	Extracción de arenas, canto rodado y triturados pétreos	2.223	3.384	4.211	4.549
1414	Extracción de arcilla y caolín	416	627	825	883
1421	Extracción de minerales para la fabricación de abonos y productos químicos, excepto turba	525	570	730	1.010
1422	Extracción de sal en salinas y de roca	808	920	967	907
1429	Explotación de minas y canteras n.c.p.	2.136	2.654	3.497	4.262
1511	Producción y procesamiento de carne y productos cárnicos	48.678	62.996	73.148	78.061
1512	Elaboración de pescado y productos de pescado	6.241	9.033	9.392	10.569
1513	Preparación de frutas, hortalizas y legumbres	18.576	22.493	28.855	29.190
1514	Elaboración de aceites y grasas de origen vegetal	13.151	17.417	24.095	22.804
1520	Elaboración de productos lácteos	26.793	28.167	33.274	32.008
1531	Elaboración de productos de molinería	8.446	9.616	12.184	14.010
1532	Elaboración de almidones y productos derivados del almidón	443	741	1.268	1.210
1533	Elaboración de alimentos preparados para animales	2.474	3.971	6.187	6.936
1541	Elaboración de productos de panadería	34.330	36.369	50.468	47.346
1542	Elaboración de azúcar	12.999	14.146	16.179	15.934
1543	Elaboración de cacao y chocolate y de productos de confitería	11.697	14.037	14.310	14.114
1544	Elaboración de pastas alimenticias	6.378	7.796	10.031	9.830
1549	Elaboración de productos alimenticios n.c.p.	27.958	29.214	36.361	38.543
1551	Destilación, rectificación y mezcla de bebidas alcohólicas; producción de alcohol etílico	1.804	2.140	3.120	3.333
1552	Elaboración de vinos y otras bebidas fermentadas a partir de frutas	15.541	18.222	22.444	21.599
1553	Elaboración de cerveza, bebidas malteadas y de malta	4.073	6.034	7.933	8.878
1554	Elaboración de bebidas no alcohólicas; producción de aguas minerales	24.247	24.350	27.265	25.735
1600	Elaboración de productos de tabaco	5.325	6.962	6.985	5.801
1711	Preparación e hilandería de fibras textiles; tejeduría de productos textiles	21.600	21.659	21.093	16.553
1712	Acabado de productos textiles	4.963	7.221	7.266	6.141
1721	Fabricación de artículos confeccionados de materiales textiles, excepto prendas de vestir	5.157	7.454	9.453	9.110
1722	Fabricación de tapices y alfombras	824	641	609	352
1723	Fabricación de cuerdas, cordeles, bramantes y redes	450	694	788	739
1729	Fabricación de productos textiles n.c.p.	7.738	12.094	19.623	16.944
1730	Fabricación de tejidos de punto y artículos de punto y ganchillo	7.450	9.181	9.772	7.713
1810	Fabricación de prendas de vestir, inclusive de cuero	32.620	42.035	48.112	37.328
1820	Terminación y teñido de pieles; fabricación de artículos de piel	276	224	105	44

Cuadro 22: Actividades a cuatro dígitos CIIU Rev 3. Descripción y empleo promedio. *(continued)*

Codigo	Descripcion	1996-2002	2003-2010	2011-2017	2018-2020
1911	Curtido y terminación de cueros	11.223	12.741	9.672	8.235
1912	Fabricación de maletas, bolsos de mano y similares, artículos de talabartería y artículos de cuero n.c.p.	1.485	2.746	2.761	2.010
1920	Fabricación de calzado y de sus partes	21.554	23.652	30.397	17.961
2010	Aserrado y cepillado de madera	8.957	13.379	14.086	12.728
2021	Fabricación de hojas de madera para enchapado; fabricación de tableros contrachapados, tableros laminados, tableros de partículas, y tableros y paneles n.c.p.	2.768	3.097	2.781	2.677
2022	Fabricación de partes y piezas de carpintería para edificios y construcciones	3.794	4.175	4.402	3.806
2023	Fabricación de recipientes de madera	710	900	950	1.060
2029	Fabricación de productos de madera n.c.p.; fabricación de artículos de corcho, paja y materiales trenzables	5.730	7.558	7.418	5.902
2101	Fabricación de pasta de madera, papel y cartón	8.067	9.997	9.926	9.363
2102	Fabricación de papel y cartón ondulado y envases de papel y cartón	8.594	9.628	11.411	11.012
2109	Fabricación de artículos de papel y cartón	9.549	11.504	13.339	11.983
2211	Edición de libros, folletos, partituras y otras publicaciones	6.650	6.511	6.084	4.666
2212	Edición de periódicos, revistas y publicaciones periódicas	13.046	14.367	14.685	10.377
2213	Edición de grabaciones	436	264	211	222
2219	Edición n.c.p.	309	521	522	530
2221	Impresión	18.637	19.687	19.730	14.984
2222	Servicios relacionados con la impresión	3.458	4.504	4.929	4.179
2230	Reproducción de grabaciones	169	90	72	100
2310	Fabricación de productos de hornos de coque	...	...	...	...
2320	Fabricación de productos de la refinación del petróleo	8.875	9.931	8.865	4.097
2330	Elaboración de combustible nuclear	...	...	...	...
2411	Fabricación de sustancias químicas básicas, excepto abonos y compuestos de nitrógeno	7.297	7.754	7.651	7.052
2412	Fabricación de abonos y compuestos de nitrógeno	301	618	770	724
2413	Fabricación de plásticos en formas primarias y de caucho sintético	5.267	3.504	3.824	3.902
2421	Fabricación de plaguicidas y productos químicos de uso agropecuario	4.194	4.916	12.184	9.437
2422	Fabricación de pinturas; barnices y productos de revestimiento similares; tintas de imprenta y masillas	4.947	5.990	7.346	7.033
2423	Fabricación de productos farmacéuticos, sustancias químicas medicinales y productos botánicos	22.877	30.691	38.204	39.150
2424	Fabricación de jabones y detergentes, preparados para limpiar y pulir, perfumes y preparados de tocador	13.330	16.300	19.892	17.869
2429	Fabricación de productos químicos n.c.p.	11.073	14.163	17.166	16.261
2430	Fabricación de fibras manufacturadas	408	344	298	203
2511	Fabricación de cubiertas y cámaras de caucho; recauchutado y renovación de cubiertas de caucho	3.708	4.762	6.446	6.195
2519	Fabricación de productos de caucho n.c.p.	5.306	6.780	7.609	6.169

Cuadro 22: Actividades a cuatro dígitos CIIU Rev 3. Descripción y empleo promedio. (continued)

Codigo	Descripcion	1996-2002	2003-2010	2011-2017	2018-2020
2520	Fabricación de productos de plástico	34.068	44.554	54.454	50.292
2610	Fabricación de vidrio y productos de vidrio	5.533	6.256	7.830	7.074
2691	Fabricación de productos de cerámica no refractaria para uso no estructural	3.019	3.254	4.020	3.495
2692	Fabricación de productos de cerámica refractaria	614	688	675	620
2693	Fabricación de productos de arcilla y cerámica no refractaria para uso estructural	7.292	7.776	8.232	6.615
2694	Elaboración de cemento, cal y yeso	5.337	5.144	7.071	7.021
2695	Fabricación de artículos de hormigón, cemento y yeso	7.005	9.069	12.504	13.216
2696	Corte, tallado y acabado de la piedra	1.360	2.031	2.227	2.211
2699	Fabricación de productos minerales no metálicos n.c.p.	2.746	3.749	4.190	3.721
2710	Industrias básicas de hierro y acero	22.612	24.593	25.553	23.771
2720	Fabricación de productos primarios de metales preciosos y metales no ferrosos	3.817	3.890	4.367	3.955
2731	Fundición de hierro y acero	3.269	4.493	4.693	3.955
2732	Fundición de metales no ferrosos	2.711	4.052	4.396	3.874
2811	Fabricación de productos metálicos para uso estructural y montaje estructural	12.543	18.416	25.633	25.421
2812	Fabricación de tanques, depósitos y recipientes de metal	2.727	3.744	4.737	4.355
2813	Fabricación de generadores de vapor	348	215	180	137
2891	Forjado, prensado, estampado y laminado de metales; pulvimetalurgia	3.458	5.261	5.964	4.884
2892	Tratamiento y revestimiento de metales; obras de ingeniería mecánica en general realizadas a cambio de una retribución o por contrata	2.832	3.886	4.643	4.174
2893	Fabricación de artículos de cuchillería, herramientas de mano y artículos de ferretería	3.979	4.282	4.518	3.994
2899	Fabricación de productos elaborados de metal n.c.p.	31.743	42.508	49.912	45.485
2911	Fabricación de motores y turbinas, excepto motores para aeronaves, vehículos automotores y motocicletas	800	1.381	1.011	914
2912	Fabricación de bombas; compresores; grifos y válvulas	3.511	5.340	6.896	6.696
2913	Fabricación de cojinetes; engranajes; trenes de engranaje y piezas de transmisión	887	1.317	1.320	862
2914	Fabricación de hornos; hogares y quemadores	284	424	542	566
2915	Fabricación de equipo de elevación y manipulación	1.559	2.299	3.651	3.890
2919	Fabricación de maquinaria de uso general n.c.p.	4.014	5.601	6.807	7.257
2921	Fabricación de maquinaria agropecuaria	7.960	11.975	13.243	12.238
2922	Fabricación de máquinas herramienta	1.147	1.622	1.940	1.945
2923	Fabricación de maquinaria metalúrgica	1.691	1.828	1.833	1.516
2924	Fabricación de maquinaria para la explotación de minas y canteras y para obras de construcción	1.220	1.486	2.313	2.673
2925	Fabricación de maquinaria para la elaboración de alimentos, bebidas y tabaco	2.913	3.341	3.867	3.676
2926	Fabricación de maquinaria para la elaboración de productos textiles, prendas de vestir y cueros	196	168	136	105
2927	Fabricación de armas y municiones	182	264	332	269

Cuadro 22: Actividades a cuatro dígitos CIIU Rev 3. Descripción y empleo promedio. *(continued)*

Codigo	Descripcion	1996-2002	2003-2010	2011-2017	2018-2020
2929	Fabricación de maquinaria de uso especial n.c.p.	5.999	9.827	11.841	10.396
2930	Fabricación de aparatos de uso doméstico n.c.p.	10.794	11.103	15.594	14.047
3000	Fabricación de maquinaria de oficina, contabilidad e informática	1.198	2.153	3.934	3.075
3110	Fabricación de motores, generadores y transformadores eléctricos	3.461	4.577	5.773	4.928
3120	Fabricación de aparatos de distribución y control de la energía eléctrica	2.163	2.499	3.259	3.543
3130	Fabricación de hilos y cables aislados	2.455	2.489	3.266	3.264
3140	Fabricación de acumuladores, pilas y baterías primarias	1.150	1.663	2.141	2.003
3150	Fabricación de lámparas eléctricas y equipo de iluminación	1.713	2.090	2.309	2.125
3190	Fabricación de equipo eléctrico n.c.p.	3.319	4.131	5.389	5.252
3210	Fabricación de tubos, válvulas y otros componentes electrónicos	383	602	855	676
3220	Fabricación de transmisores de radio y televisión y de aparatos para telefonía y telegrafía con hilos	1.271	1.086	3.248	3.010
3230	Fabricación de receptores de radio y televisión, aparatos de grabación y reproducción de sonido y video, y productos conexos	2.699	3.847	9.053	6.253
3310	Fabricación de aparatos e instrumentos médicos y de aparatos para medir, verificar, ensayar, navegar y otros fines, excepto instrumentos de óptica	1.926	2.630	2.960	3.051
3311	Fabricación de equipo médico y quirúrgico y de aparatos ortopédicos	2.032	3.154	4.206	4.461
3320	Fabricación de instrumentos de óptica y equipo fotográfico	1.276	1.434	1.423	1.337
3330	Fabricación de relojes	158	126	85	67
3410	Fabricación de vehículos automotores	19.981	19.626	31.430	26.837
3420	Fabricación de carrocerías para vehículos automotores; fabricación de remolques y semirremolques	4.060	5.841	7.803	6.782
3430	Fabricación de partes; piezas y accesorios para vehículos automotores y sus motores	27.885	36.199	43.056	36.898
3511	Construcción y reparación de buques	1.464	3.314	3.788	3.236
3512	Construcción y reparación de embarcaciones de recreo y deporte	406	761	939	624
3520	Fabricación de locomotoras y de material rodante para ferrocarriles y tranvías	1.266	1.384	1.563	805
3530	Fabricación y reparación de aeronaves	1.149	1.261	1.977	1.700
3590	Fabricación de motocicletas, bicicletas y sillones ortopédicos	1.682	2.161	3.141	2.810
3599	Fabricación de equipo de transporte n.c.p.	407	402	522	491
3610	Fabricación de muebles y colchones	16.290	19.502	24.371	22.055
3691	Fabricación de joyas y artículos conexos	857	718	628	537
3692	Fabricación de instrumentos de música	130	204	244	185
3693	Fabricación de artículos de deporte	451	672	729	606
3694	Fabricación de juegos y juguetes	723	917	1.307	1.164
3699	Industrias manufactureras n.c.p.	8.900	9.760	11.531	11.111
3710	Reciclamiento de desperdicios y desechos metálicos	631	1.013	1.726	1.409

Cuadro 22: Actividades a cuatro dígitos CIIU Rev 3. Descripción y empleo promedio. *(continued)*

Codigo	Descripcion	1996-2002	2003-2010	2011-2017	2018-2020
3720	Reciclamiento de desperdicios y desechos no metálicos	460	1.362	1.709	1.665
4011	Generación de energía eléctrica	4.580	5.560	9.446	10.286
4012	Transporte de energía eléctrica	1.400	1.454	1.904	2.288
4013	Distribución de energía eléctrica	25.798	27.598	33.919	37.356
4020	Fabricación de gas y distribución de combustibles gaseosos por tuberías	...	...	...	...
4030	Suministro de vapor y agua caliente	...	...	...	...
4100	Captación, depuración y distribución de agua	10.423	11.838	15.221	17.714
4511	Demolición y voladura de edificios y de sus partes	304	636	787	711
4512	Perforación y sondeo, excepto: perforación de pozos de petróleo, de gas, de minas e hidráulicos y prospección de yacimientos de petróleo	331	671	878	1.052
4519	Movimiento de suelos y preparación de terrenos para obras n.c.p.	4.250	7.767	12.270	12.917
4520	Construcción de edificios y sus partes y obras de ingeniería civil	189.894	289.073	377.821	359.053
4531	Ejecución y mantenimiento de instalaciones eléctricas, electromecánicas y electrónicas	10.734	17.099	23.382	26.077
4532	Aislamiento térmico, acústico, hídrico y antivibratorio	424	938	1.340	1.512
4533	Instalaciones de gas, agua, sanitarios y de climatización, con sus artefactos conexos	4.014	6.179	7.503	7.465
4539	Instalaciones para edificios y obras de ingeniería civil n.c.p.	3.278	5.461	7.419	7.956
4541	Instalaciones de carpintería, herrería de obra y artística	1.008	1.924	2.369	2.014
4542	Terminación y revestimiento de paredes y pisos	904	1.800	2.137	1.738
4543	Colocación de cristales en obra	83	186	300	270
4544	Pintura y trabajos de decoración	1.411	2.465	2.887	2.530
4549	Terminación de edificios y obras de ingeniería civil n.c.p.	2.504	808	648	740
4550	Alquiler de equipo de construcción o demolición dotado de operarios	676	1.293	1.966	2.300
5011	Venta de vehículos automotores nuevos, excepto motocicletas	18.883	22.590	43.534	48.146
5012	Venta de vehículos automotores usados, excepto motocicletas	1.920	2.083	2.712	2.733
5020	Mantenimiento y reparación de vehículos automotores, excepto motocicletas	15.358	24.973	33.778	31.694
5031	Venta al por mayor de partes, piezas y accesorios de vehículos automotores	6.942	10.912	14.059	14.757
5032	Venta al por menor de partes, piezas y accesorios de vehículos automotores	10.890	15.705	20.913	22.047
5050	Venta al por menor de combustible para vehículos automotores y motocicletas	40.537	48.733	57.122	60.873
5111	Venta al por mayor en comisión o consignación de productos agrícolas y pecuarios	22.983	30.797	35.491	35.332
5119	Venta al por mayor en comisión o consignación de mercancías n.c.p.	9.172	13.514	15.974	16.305
5121	Venta al por mayor de materias primas agropecuarias y de animales vivos	8.958	9.751	11.304	11.452
5122	Venta al por mayor de alimentos	28.554	42.907	60.293	63.403
5123	Venta al por mayor de bebidas	7.632	11.771	16.672	17.427

Cuadro 22: Actividades a cuatro dígitos CIIU Rev 3. Descripción y empleo promedio. (continued)

Codigo	Descripcion	1996-2002	2003-2010	2011-2017	2018-2020
5124	Venta al por mayor de cigarrillos y productos de tabaco	2.993	3.987	3.824	3.562
5131	Venta al por mayor de productos textiles, prendas de vestir, calzado excepto el ortopédico, cueros, pieles, artículos de marroquinería, paraguas y similares	10.654	13.666	17.682	16.026
5132	Venta al por mayor de libros, revistas, diarios, papel, cartón, materiales de embalaje y artículos de librería	8.542	9.715	10.488	9.289
5133	Venta al por mayor de productos farmacéuticos, veterinarios, cosméticos y de perfumería, instrumental médico y odontológico y artículos ortopédicos	20.472	24.848	32.505	36.503
5134	Venta al por mayor de artículos de óptica, fotografía, relojería, joyería y fantasías	1.843	2.799	3.464	3.185
5135	Venta al por mayor de muebles, artículos de iluminación y demás artefactos para el hogar	7.958	8.840	9.799	9.379
5139	Venta al por mayor de artículos de uso doméstico y/o personal n.c.p.	11.776	19.286	25.451	26.477
5141	Venta al por mayor de combustibles, incluso gaseosos y productos conexos	6.608	8.008	9.280	9.984
5142	Venta al por mayor de metales y minerales metalíferos	2.627	3.101	3.552	3.444
5143	Venta al por mayor de madera, materiales de construcción, artículos de ferretería y materiales para plomería e instalaciones de gas	14.787	19.868	26.363	26.746
5149	Venta al por mayor de productos intermedios n.c.p., desperdicios y desechos	6.424	8.602	11.847	11.859
5151	Venta al por mayor de máquinas, equipos e implementos de uso especial	6.452	9.587	13.744	15.312
5152	Venta al por mayor de máquinas-herramienta de uso general	1.006	1.606	2.334	2.540
5153	Venta al por mayor de vehículos, equipos y máquinas para el transporte ferroviario, aéreo y de navegación	46	71	65	66
5154	Venta al por mayor de muebles e instalaciones para la industria, el comercio y los servicios	1.018	1.287	1.691	1.642
5159	Venta al por mayor de máquinas, equipo y materiales conexos n.c.p.	14.059	19.379	27.714	24.836
5190	Venta al por mayor de mercancías n.c.p.	14.524	19.718	26.301	26.440
5211	Venta al por menor en comercios no especializados con predominio de productos alimenticios y bebidas	108.376	143.029	167.954	161.907
5212	Venta al por menor excepto la especializada, sin predominio de productos alimenticios y bebidas	2.123	4.707	8.861	8.323
5221	Venta al por menor de productos de almacén, fiambrería y dietética	6.826	8.661	12.764	14.439
5222	Venta al por menor de carnes rojas y productos de granja y de la caza	5.868	9.941	16.228	18.258
5223	Venta al por menor de frutas, legumbres y hortalizas frescas	1.716	2.321	3.583	4.200
5224	Venta al por menor de pan y productos de panadería y confitería	9.642	13.538	22.253	24.391
5225	Venta al por menor de bebidas	834	1.781	3.146	3.231
5229	Venta al por menor de productos alimenticios n.c.p. y tabaco, en comercios especializados	5.839	10.823	17.616	20.598

Cuadro 22: Actividades a cuatro dígitos CIIU Rev 3. Descripción y empleo promedio. *(continued)*

Codigo	Descripcion	1996-2002	2003-2010	2011-2017	2018-2020
5231	Venta al por menor de productos farmacéuticos, cosméticos, de perfumería, instrumental médico y odontológico y artículos ortopédicos	22.820	38.504	56.850	65.122
5232	Venta al por menor de productos textiles, excepto prendas de vestir	3.509	4.782	5.853	5.434
5233	Venta al por menor de prendas y accesorios de vestir excepto calzado, artículos de marroquinería, paraguas y similares	22.078	42.038	59.369	52.738
5234	Venta al por menor de calzado excepto el ortopédico, artículos de marroquinería, paraguas y similares	7.494	13.046	20.497	18.259
5235	Venta al por menor de muebles, artículos de mimbre y corcho, colchones y somieres, artículos de iluminación y artefactos para el hogar	24.185	41.460	59.593	53.725
5236	Venta al por menor de materiales de construcción, artículos de ferretería, pinturas, cristales y espejos, y artículos para la decoración	25.393	38.995	52.210	54.352
5237	Venta al por menor de artículos de óptica, fotografía, relojería, joyería y fantasía	5.204	7.632	9.824	9.855
5238	Venta al por menor de libros, revistas, diarios, papel, cartón, materiales de embalaje y artículos de librería	6.559	10.093	12.108	11.353
5239	Venta al por menor en comercios especializados n.c.p.	28.531	46.029	63.200	62.180
5241	Venta al por menor de muebles usados	78	150	145	87
5242	Venta al por menor de libros, revistas y similares usados	89	38	38	26
5249	Venta al por menor de artículos usados n.c.p.	451	503	424	396
5251	Venta al por menor por correo, televisión, internet y otros medios de comunicación	1.321	4.448	4.463	2.916
5252	Venta al por menor en puestos móviles	89	233	220	273
5259	Venta al por menor no realizada en establecimientos n.c.p.	1.793	2.192	2.633	2.379
5261	Reparación de calzado y artículos de marroquinería	229	308	315	200
5262	Reparación de artículos eléctricos de uso doméstico	1.318	1.948	1.595	1.159
5269	Reparación de efectos personales y enseres domésticos n.c.p.	5.220	7.125	8.768	9.434
5511	Servicios de alojamiento en camping	764	735	868	784
5512	Servicios de alojamiento excepto en camping	32.666	50.714	65.410	63.422
5521	Servicios de expendio de comidas y bebidas en restaurantes, bares y otros establecimientos con servicio de mesa y/o en mostrador excepto en heladerías	72.020	115.243	170.520	180.732
5522	Preparación y venta de comidas para llevar	11.784	21.763	34.723	36.789
6010	Servicio de transporte ferroviario	17.178	21.855	35.318	39.540
6021	Servicio de transporte automotor de cargas	61.465	99.974	142.074	140.009
6022	Servicio de transporte automotor de pasajeros	115.804	129.796	159.347	149.940
6030	Servicio de transporte por tuberías	1.463	1.715	1.838	2.109
6110	Servicio de transporte marítimo	1.895	3.491	4.809	4.325
6120	Servicio de transporte fluvial	1.415	2.046	3.005	2.435
6200	Servicio de transporte aéreo de carga y de pasajeros	11.451	12.668	16.436	18.350
6310	Servicios de manipulación de carga	3.803	7.531	9.489	10.129

Cuadro 22: Actividades a cuatro dígitos CIIU Rev 3. Descripción y empleo promedio. (continued)

Codigo	Descripcion	1996-2002	2003-2010	2011-2017	2018-2020
6320	Servicios de almacenamiento y depósito	4.857	5.495	7.903	9.422
6331	Servicios complementarios para el transporte terrestre	9.295	11.903	15.005	13.636
6332	Servicios complementarios para el transporte por agua	8.114	10.077	13.150	14.357
6333	Servicios complementarios para el transporte aéreo	2.334	4.599	7.460	8.387
6340	Servicios de agencias de viaje y otras actividades complementarias de apoyo turístico	8.120	14.088	21.860	22.344
6350	Servicios de gestión y logística para el transporte de mercaderías	7.717	17.002	30.129	35.696
6410	Servicios de correos	26.067	18.494	19.808	20.723
6420	Servicios de telecomunicaciones	57.886	69.527	76.598	74.541
6521	Servicios de las entidades financieras bancarias	53.971	52.212	64.824	66.348
6522	Servicios de las entidades financieras no bancarias	6.930	4.134	3.302	2.657
6590	Servicios financieros excepto los de la banca central y las entidades financieras	20.949	21.556	26.649	22.567
6610	Servicios de seguros	30.304	35.179	47.302	47.369
6620	Administración de fondos de jubilaciones y pensiones (AFJP)	16.037	8.340	164	73
6711	Servicios de administración de mercados financieros	941	1.149	1.198	1.122
6712	Servicios bursátiles de mediación o por cuenta de terceros	1.325	1.431	2.024	3.052
6719	Servicios auxiliares a la actividad financiera n.c.p., excepto a los servicios de seguros y de administración de fondos de jubilaciones y pensiones	1.953	2.719	2.731	2.769
6721	Servicios auxiliares a los servicios de seguros	...	...	...	...
6722	Servicios auxiliares a la administración de fondos de jubilaciones y pensiones	...	...	...	...
7010	Servicios inmobiliarios realizados por cuenta propia, con bienes propios o arrendados	94.191	82.996	56.213	33.613
7020	Servicios inmobiliarios realizados a cambio de un retribución o por contrata	64.747	74.002	81.728	81.992
7111	Alquiler de equipo de transporte para vía terrestre, sin operarios	3.222	3.694	3.707	2.641
7112	Alquiler de equipo de transporte para vía acuática, sin operarios ni tripulación	...	...	...	...
7113	Alquiler de equipo de transporte para vía aérea, sin operarios ni tripulación	...	...	...	...
7120	Alquiler de maquinaria y equipo n.c.p.	6.180	7.553	7.671	7.103
7130	Alquiler de efectos personales y enseres domésticos n.c.p.	1.051	1.444	926	753
7210	Servicios de consultores en equipo de informática	368	931	1.545	2.483
7220	Servicios de consultores en informática y suministros de programas de informática	6.777	23.153	46.591	59.823
7230	Procesamiento de datos	3.128	4.197	6.495	5.863
7240	Servicios relacionados con bases de datos	816	1.336	1.290	1.137
7250	Mantenimiento y reparación de maquinaria de oficina, contabilidad e informática	586	1.371	1.434	1.244
7290	Actividades de informática n.c.p.	10.594	22.041	35.000	41.195
7300	Investigación y desarrollo experimental	2.699	4.598	6.911	8.037

Cuadro 22: Actividades a cuatro dígitos CIIU Rev 3. Descripción y empleo promedio. *(continued)*

Codigo	Descripcion	1996-2002	2003-2010	2011-2017	2018-2020
7410	Servicios jurídicos y de contabilidad, teneduría de libros y auditoría; asesoramiento en materia de impuestos; estudios de mercados y realización de encuestas de opinión pública; asesoramiento empresarial y en materia de gestión	91.969	121.779	122.200	106.272
7421	Servicios de arquitectura e ingeniería y servicios conexos de asesoramiento técnico	25.173	38.716	44.127	40.738
7422	Ensayos y análisis técnicos	1.436	3.388	5.533	5.749
7430	Servicios de publicidad	10.511	14.947	19.208	18.875
7491	Obtención y dotación de personal	11.020	12.601	19.735	14.955
7492	Servicios de investigación y seguridad	72.947	114.986	150.729	149.275
7493	Servicios de limpieza de edificios	50.359	73.506	109.559	117.557
7494	Servicios de fotografía	1.910	1.807	1.112	663
7495	Servicios de envase y empaque	1.445	2.765	3.451	3.039
7496	Servicios de impresión heliográfica, fotocopia y otras formas de reproducciones	4.643	2.136	2.136	1.675
7499	Servicios empresariales n.c.p.	32.506	92.040	128.323	137.167
7500	Agencias de empleo temporario	34.913	84.333	67.576	39.731
8000	Enseñanza	248.973	356.464	460.485	505.934
8510	Servicios relacionados con la salud humana	159.833	194.430	277.326	314.980
8520	Servicios veterinarios	769	1.554	2.273	2.428
8530	Servicios sociales	14.526	18.834	26.608	29.683
9000	Eliminación de desperdicios y aguas residuales, saneamiento y servicios similares	23.373	21.558	26.927	30.774
9100	Servicios de asociaciones n.c.p.	118.700	148.420	191.627	205.738
9211	Producción y distribución de filmes y videocintas	1.750	4.918	7.321	7.712
9212	Exhibición de filmes y videocintas	1.775	2.772	3.836	3.436
9213	Servicios de radio y televisión	6.676	8.046	9.271	8.451
9214	Servicios teatrales y musicales y servicios artísticos n.c.p.	2.001	3.419	5.946	6.081
9219	Servicios de espectáculos artísticos y de diversión n.c.p.	3.483	4.193	5.847	5.374
9220	Servicios de agencias de noticias	913	833	914	668
9230	Servicios de bibliotecas, archivos y museos y servicios culturales n.c.p.	1.141	2.148	2.290	1.962
9241	Servicios para prácticas deportivas	19.811	27.598	37.668	39.849
9249	Servicios de esparcimiento n.c.p.	12.992	30.377	41.535	36.738
9301	Lavado y limpieza de artículos de tela, cuero y/o de piel, incluso la limpieza en seco	3.464	4.747	5.181	4.116
9302	Servicios de peluquería y tratamientos de belleza	6.872	11.127	11.844	10.230
9303	Pompas fúnebres y servicios conexos	4.667	5.032	5.458	5.542
9309	Servicios n.c.p.	46.574	53.093	43.520	35.140
		3.829.673	5.117.448	6.430.175	6.388.649

5.3.2. Sectores excluidos

En el Cuadro 23 se detallan los sectores CIIU Rev 3 excluidos con datos disponibles del empleo promedio en los períodos agregados analizados.

Cuadro 23: Actividades a cuatro dígitos excluidas del análisis. Flujos totales y empleo promedio en períodos.

Código	Descripción	Empleo promedio			
		1996-2002	2003-2010	2011-2017	2018-2020
1310	Extracción de minerales de hierro	...	...	...	...
1320	Extracción de minerales metalíferos no ferrosos, excepto minerales de uranio y torio	...	...	...	...
1820	Terminación y teñido de pieles; fabricación de artículos de piel	276	224	105	44
2310	Fabricación de productos de hornos de coque	...	...	...	...
2330	Elaboración de combustible nuclear	...	...	...	...
4020	Fabricación de gas y distribución de combustibles gaseosos por tuberías	...	...	...	...
4030	Suministro de vapor y agua caliente	...	...	...	...
5241	Venta al por menor de muebles usados	78	150	145	87
5242	Venta al por menor de libros, revistas y similares usados	89	38	38	26
6721	Servicios auxiliares a los servicios de seguros	...	...	...	...
6722	Servicios auxiliares a la administración de fondos de jubilaciones y pensiones	...	...	...	...
7112	Alquiler de equipo de transporte para vía acuática, sin operarios ni tripulación	...	...	...	...
7113	Alquiler de equipo de transporte para vía aérea, sin operarios ni tripulación	...	...	...	...
7500	Agencias de empleo temporario	34.913	84.333	67.576	39.731
	Subtotal Sectores Excluidos	35.356	84.745	67.864	39.888

5.4. Reproducibilidad

Datos. Los datos utilizados en esta tesis fueron provistos por el Observatorio de Empresas y Dinámica del Empleo (OEDE) del MTEySS de la Nación Argentina, provenientes del Sistema Integrado Previsional Argentino (SIPA) que registra el empleo formal. Cada registro presenta la cantidad de individuos que cambiaron interanualmente de empleador agrupadas por la actividad económica declarada por éste. De esta forma, dado que no se muestra la identidad de empleador ni empleado, se garantiza la anonimidad de los datos casi en su totalidad, excepto por aquellas actividades en que la cantidad de empresas y/o empleados resulte muy pequeña. Por lo tanto se excluyeron algunos sectores a fin de completar los requisitos oficiales de anonimización. En cuanto a la publicidad de los datos utilizados, actualmente se encuentran en proceso de revisión con OEDE para su eventual publicación.

Código. El código fuente se realizó en su totalidad en R y se encontrará disponible en <https://github.com/sderaco/tesis> con los datos utilizados luego que se apruebe su publicación por parte del OEDE.