

# Identificación de Diferencias y Similitudes Estructurales en las Redes Interindustriales de Empleo de Argentina

Sergio De Raco<sup>1</sup>[0000-0002-1240-5799] y Viktoriya Semeshenko<sup>2</sup>[0000-0003-0295-5946]

<sup>1</sup> Universidad de Buenos Aires. Facultad de Ciencias Económicas. Buenos Aires, Argentina. Universidad de Buenos Aires. Instituto Interdisciplinario de Economía Política de Buenos Aires. Buenos Aires, Argentina

<sup>2</sup> Universidad de Buenos Aires. Facultad de Ciencias Económicas. Buenos Aires, Argentina. CONICET-Universidad de Buenos Aires. Instituto Interdisciplinario de Economía Política de Buenos Aires. Buenos Aires, Argentina.  
<https://iiep.economicas.uba.ar/>

**Resumen** La movilidad del empleo entre industrias refleja relaciones entre actividades económicas que se representan naturalmente como redes que resaltan las propiedades de conectividad entre sectores económicos. Los flujos de empleo varían en el tiempo tanto por factores coyunturales como estructurales. Usando registros administrativos de Argentina para el período 1996-2020, exploramos las redes interindustriales y caracterizamos sus estructuras y dinámicas a cuatro dígitos de clasificación de actividades CIIU. Analizamos las transiciones interanuales de empleo entre casi 300 actividades económicas y encontramos redes de elevada conectividad con propiedades de mundo pequeño, cuya estructura cambia en el tiempo. ¿Cuán distintas son las estructuras subyacentes de intercambios interindustriales de empleo? ¿Es posible caracterizar conexiones estables y cambios significativos a lo largo del tiempo? Aplicando métricas y mediciones de similitud estructural para cuantificar las diferencias en las redes interanuales encontramos que es posible identificar distintos regímenes de conectividad que correlacionan con períodos de cierta estabilidad macroeconómica, a la vez que se detectan períodos transicionales en los que estas estructuras van cambiando entre regímenes.

**Keywords:** Redes Laborales · Distancias · Registros Administrativos.

## 1. Introducción

Las transiciones del mercado laboral, es decir, la movilidad de los trabajadores entre trabajos y dentro y fuera del empleo, son intrínsecas al funcionamiento de estos mercados y sus características particulares determinan la flexibilidad requerida para adaptarse a las crisis económicas y al cambio estructural. Como se ha notado en numerosos estudios, la movilidad laboral entendida a través de las transiciones de empleo presenta heterogeneidades marcadas entre países, tanto en volumen relativo como en la composición de los distintos tipos de transiciones (ver, por ejemplo [1]).

En los movimientos que ocurren dentro del empleo, como las transiciones entre distintos empleos entre períodos consecutivos (v.g. “*job-to-job transitions*”), resulta útil e informativo observar la movilidad laboral desde el punto de vista de la actividad económica de las firmas (empleadores), que provee una dimensión relevante para el análisis del entramado productivo. En particular, cuando estos cambios de empleo ocurren entre firmas de distintas industrias<sup>3</sup>, 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 este trabajo, nos enfocamos en los flujos **intersectoriales** de empleo, en particular aquellos que ocurren entre firmas dedicadas a actividades económicas distintas.

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 este trabajo se utilizan datos para períodos desde 1996 hasta 2020.

Usando un enfoque basado en datos y simultáneamente desagregado, representamos las transiciones de “job-to-job” *via* la red de flujo laboral, donde los sectores (nodos) están conectados entre sí cuando se ha observado un intercambio de empleo entre ellos. Estudiamos sus propiedades estructurales y encontramos redes densas, altamente conectadas con regularidades y estructuras interesantes ([3], [4]).

En la Fig. 1 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 *intra*sectoriales), y entre empresas de diferentes sectores de actividad (transiciones *inter*sectoriales), 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 representativos: 1) uno declinante asociado a la última parte de la vigencia del Plan de Convertibilidad y la crisis posterior (1996-2002, barra de color azul en Fig. 1); 2) uno de

<sup>3</sup> En este trabajo utilizamos indistintamente la referencia a “actividades económicas”, “sectores” productivos, e “industrias”, que se utiliza en sentido amplio, en referencia a las distintas clasificaciones incluidas en el sistema CIIU Rev. 3.

recuperación y crecimiento post-crisis (2003-2011, barra de color fucsia en Fig. 1); 3) otro de estancamiento (2012-2017, barra de color verde en Fig. 1); y 4) uno de fuerte caída asociado al cambio de gobierno y el comienzo de la pandemia COVID-19 (2018-2020, barra de color naranja en Fig. 1).

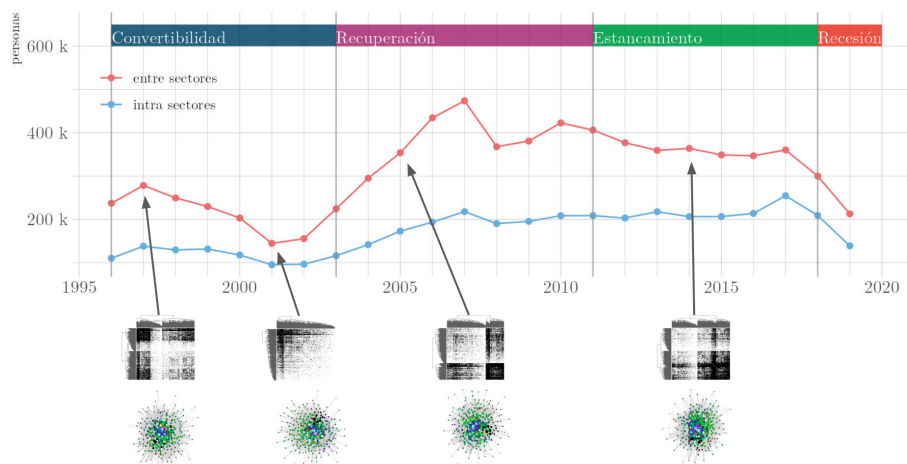


Figura 1: Transiciones de empleo formal privado, dentro y entre sectores de actividad económica a 4 dígitos CIU Rev. 3 del SIPA del MTEySS. Años organizados en periodos representativos según un criterio basado en la evolución de la actividad económica (representados por barra de colores). Debajo: Representaciones de grafos y respectivos mapas de calor para años seleccionados (1997, 2001, 2005 y 2014), de las redes binarias, es decir que indican con 1 si existe la conexión entre dos sectores cualesquiera y con 0 si no existe. Así, nos enfocamos exclusivamente en la estructura de conexiones intersectoriales, independientemente de la magnitud de flujos de empleo. Se aprecia la dificultad para comparar las estructuras de conectividad y caracterizar su evolución.

La estructura de las redes de empleo cambia durante el período de tiempo considerado. Sin embargo, dado que las redes son densas y altamente conectadas, no resulta trivial su análisis y comparación. Al observar directamente la estructura de la red (ver grafos y mapas de calor en Fig. 1), se evidencia la dificultad para detectar diferencias estructurales que plantean estas redes. Si seleccionamos momentos diferentes en la serie de tiempo y observamos la red de conexiones de flujos de empleo, la estructura es muy similar y es difícil concluir qué cambios topológicos se han realizado. Sin embargo, si miramos el mapa de calor (*heatmap*) correspondiente a cada red, podemos ver la configuración específica de la estructura y observar las diferencias, porque ellos brindan información topológica que funciona como una huella dactilar (*fingerprint*) de su estructura de conectividad.

En este artículo, el objetivo es identificar diferencias y/o similitudes en las estructuras de conexiones de las redes interindustriales de empleo para diferentes años bajo periodos de consideración, a fin de detectar los cambios ocurridos en el tiempo, cuantificar las diferencias entre las estructuras subyacentes de las redes, caracterizar e identificar conexiones estables y cambios significativos de la estructura a lo largo del tiempo. Por otro lado, queremos indagar si es posi-

ble relacionar los resultados con la evolución de la actividad macroeconómica. Para eso, proponemos utilizar distintas medidas de distancias y disimilitudes entre cada par de grafos interanuales y comparar los mapas de calor, utilizados comúnmente para visualizar matrices de adyacencias de estas redes.

La identificación de las diferencias estructurales de la red es un problema fundamental y desafiante ([2], [5], [7]). A partir de las matrices fundamentales asociadas a las redes de flujos, utilizaremos distintas métricas y disimilitudes para su comparación mediante el análisis a diferentes escalas de sus estructuras topológicas: local, intermedia (meso) y global [6]. Complementariamente, utilizaremos la distancia D-measure [9] como una métrica que promedia características locales y globales en la comparación de redes y permite capturar el papel diferente de los enlaces y nodos en las redes y, por lo tanto, identificar mejor las diferencias topológicas de las redes.

Como resultados principales, encontramos que es posible determinar distintos regímenes que se relacionan con períodos de ciertas estabilidad macroeconómica, y períodos transicionales en los que las estructuras de conectividad cambian entre estos regímenes. Además, al considerar distancias más globales que consideran la conectividad de la red en su conjunto, se logra identificar y cuantificar más detalladamente las estructuras en términos de regímenes, subregímenes, períodos transicionales y cambios de régimen. Eso se debe a que estas distancias consideran la conectividad integral de los nodos la red, condensando mayor información que las distancias que consideran vecindarios reducidos.

El documento está organizado de la siguiente manera: los datos utilizados en el análisis se describen en la sección 2. En la sección 3 presentamos las características de las redes subyacentes y describimos la metodología aplicada en términos de las distancias utilizadas. En la sección 4 presentamos los resultados. La discusión se sigue en la sección 5, y al final se presenta el Material Complementario con la información detallada sobre el cálculo de distancias.

## 2. Datos

Para nuestro análisis utilizamos datos de empleo formal del sector privado de Sistema Previsional Argentino (SIPA) proporcionados por el Observatorio de Dinámica Empresarial y del Empleo (OEDE) del MTEySS. A partir de estos registros administrativos se arman las matrices de transición de empleo por actividad económica, cuya construcción se basa en la creación de paneles interanuales de individuos que aparecen en ambos períodos de registro para el mes de referencia<sup>4</sup>. Se contabiliza la cantidad 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. Los datos comprenden intercambios interanuales de empleo entre actividades productivas registrados entre los años 1996 y 2020. El conjunto de actividades incluye cerca de 300 sectores (ramas de actividad) a cuatro dígitos de detalle del clasificador CIIU Rev. 3.

<sup>4</sup> La metodología habitual de construcción de paneles interanuales de individuos utilizada por el MTEySS considera octubre como mes de referencia.

Después del preprocesamiento<sup>5</sup> de datos se cuenta con 24 matrices de transición anuales. Luego, a partir de estas matrices se procede a construir las redes de conexiones de flujos de empleo interanuales correspondientes.

### 3. Redes de Empleo Interanuales

Las matrices de transición representan redes dirigidas y ponderadas de interacciones. Analizamos redes interanuales para el período de tiempo de interés y encontramos estructuras densas, de elevada conectividad, con un único componente denso conexo, de diámetro reducido, con propiedades de mundo pequeño. Predominantemente, las redes presentan estructura de centro-periferia ([3], [4], [10]). Como se demuestra en la Fig. 1, si se analizan directamente redes de flujos es muy difícil detectar los cambios en la estructura y topología del período a período. Sin embargo, si miramos los mapas de calor de las redes correspondientes, podemos identificar diferencias entre los mapas de calor y detectar los cambios topológicos, se observa que la estructura cambia para diferentes períodos.

Por lo tanto, surgen varias preguntas para el análisis: ¿Cuán distintas son las estructuras subyacentes de intercambios interindustriales de empleo? ¿Es posible caracterizar y/o identificar conexiones estables y cambios significativos a lo largo del tiempo? ¿Cómo se relaciona la evolución temporal con los períodos macroeconómicos?

Para contestar estas preguntas utilizamos distintas medidas de distancias y de disimilitud para comparar los mapas de calor que brindan información estructural y topológica, para detectar los cambios ocurridos en el tiempo y analizar su estabilidad y evolución.

**Distancias.** Existen diversos métodos para comparar redes y, dependiendo del tipo de interacciones que se desea analizar, pueden utilizarse varios en forma complementaria para lograr una mejor comprensión de los fenómenos y casos atípicos que se desean estudiar, y para detectar mejor las características estructurales particulares. La selección de estas técnicas depende del tipo y el nivel de análisis de la comparación, que a su vez brindará distinta información de la estructura de las redes representadas.

En este trabajo, utilizamos el *framework* desarrollado por Donnat y Holmes [6], para comparar redes de nodos alineados en distintos períodos (es decir del

<sup>5</sup> Se excluyeron algunos sectores de modo que resultara posible obtener matrices de transición de igual dimensión para todos los períodos bajo análisis: a) Sector Público y Banca Central, dado que el análisis se enfoca en transiciones de empleo dentro del sector privado; b) 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[8]; c) sectores para los que no se publica el nivel de empleo por razones de secreto estadístico; d) sectores de bajo nivel de empleo sin transiciones de ingreso o de salida en alguno de los períodos utilizados ; y, e) 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.

mismo tamaño, que en nuestro caso representan a transiciones entre los mismos sectores de actividad a cuatro dígitos CIIU en todos los períodos de comparación), al que incorporamos la distancia D-measure [9]. A partir de las matrices fundamentales asociadas a estas redes, se utilizan distintas métricas y disimilitudes para su comparación mediante el análisis a diferentes escalas de sus estructuras topológicas: local, intermedia (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 que no ocurren. En cuanto a la estructura global, refiere a la red en su totalidad, en término de las propiedades del sistema que representan los nodos y los lazos que la componen, así como características de las dinámicas en base al análisis realizado sobre las matrices fundamentales (de adyacencias y/o Laplacianas). Por último, la existencia de meso-estructura a niveles intermedios, a veces referida como estructura “*glocal*” para destacar su composición mixta de estructura global y local, amplía el análisis local a entornos o vecindarios más extendidos, por ejemplo de segundo o tercer orden (v.g.: vecindarios de sus vecinos, etc.), se plantea como la posibilidad de identificar subconjuntos de nodos que compartan conectividad en grupos.

Nos interesa investigar sobre los cambios estructurales y topológicos ocurridos en redes de empleo el tiempo. Para eso, las distancias utilizadas se calculan sobre las matrices de adyacencias binarias,  $A_{ij} \in \{0, 1\}$ , de las redes no dirigidas. Es decir que las matrices de la conectividad intersectorial de los flujos observados no tienen la magnitud de los flujos ni direccionalidad. O sea, nos interesa exclusivamente la existencia o no de la conexión entre los nodos. De aquí en adelante, seleccionamos las distancias estructural (local), polinomial (local extendida), espectral y D-measure (globales), para realizar el análisis comparativo. Los detalles sobre el desarrollo matemático y cómo se calcula cada distancia se tratan en Material Complementario (ver 5).

La *distancia estructural*. Utilizamos la distancia de Jaccard para identificar cambios en las redes a nivel de entornos locales de interconexión de cada sector y evaluar la tasa de reacomodamiento de lazos. 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.

La *distancia polinomial*. Esta distancia está basada en la utilización del polinomio de potencias de la matriz de adyacencias,  $A^k$ , que se relaciona con la topología local de la red a través de los coeficientes de  $A_{ij}^k$  que representan los caminos del nodo  $i$  al nodo  $j$  a través de  $k$  saltos o pasos. Esta distancia permite cuantificar diferencias entre dos redes 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. Dado que el diámetro de la red es reducido con esta distancia se captarían cambios en la estructura cercanos a la escala global, de toda la red.

La *distancia espectral*. Se utiliza para caracterizar el estado de la red mediante sus valores propios, y permite detectar cambios en estructuras globales e

intermedias. La distancia espectral se define como la distancia entre funciones del espectro de los valores propios de las matrices de dos redes bajo comparación. Al efectuar el análisis espectral (*eigen-análisis*) 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.

*D-measure*. Esta medida de disimilitud captura y cuantifica adecuadamente las diferencias topológicas, reconoce las diferentes funciones de los enlaces y nodos, teniendo en cuenta las desconexiones y otras condiciones estructurales. La idea principal para medir la disimilitud,  $d_D(G, \tilde{G})$ , de dos redes es asociar a cada estructura un conjunto de funciones de distribución de probabilidad (PDF de tres términos, ver ecuación 4 en 5), que representan las distancias de conectividad de todos los nodos, y compararlas, por métricas teóricas de la información estándar [9]. El primer término compara las redes, a través de las distribuciones de distancia de los grafos, capturando las diferencias topológicas globales. El segundo término compara la conectividad de cada nodo y cómo cada elemento está conectado a través de la red, observando las distribuciones de distancias de los nodos. El último término analiza las diferencias en la forma en que se produce esta conectividad, a través del análisis de la centralidad alfa de cada nodo.

Para evaluar los resultados de las distancias y medidas seleccionadas, utilizamos mapas de calor 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.

## 4. Resultados

En esta sección presentamos resultados del análisis de comparación de las redes interanuales de los  $T = 24$  períodos usando las cuatro distancias seleccionadas, que permiten distinguir diferencias a nivel local, intermedio (meso) y global de las estructuras de interacciones observadas, y relacionar importantes diferencias estructurales con los períodos macroeconómicos.

**Jaccard**. La distancia de Jaccard cuantifica la cantidad de cambios en las interacciones del entorno directo (local) de cada sector, es decir de los sectores con los que tienen intercambios de empleo en forma directa (medidos por el grado,  $k$ , del sector) y los resultados se presentan en la Fig.2.A1. 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) y entre 2004 y 2018 (etapas posteriores), que presentan relativa similitud hacia adentro de los mismos en los años consecutivos (regiones de color naranja más claro sobre la diagonal en la Fig.2.A1) y difieren entre períodos (regiones de color naranja más oscuro hacia los bordes, Fig.2.A1), separados por años de crisis 2001-2003 y 2019, en que aparecen distancias mayores con todos los

años (franjas de rojo oscuro, Fig.2.A1). 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.

En la Fig.2.A2 se mapea mediante escalamiento multidimensional métrico (MDS) la matriz de distancias calculada 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 horizontal 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. Es posible notar que la distancia de Jaccard ordena a las transiciones de 2003, incluidas en el período de recuperación Post-Convertibilidad, con el resto de las observadas en el período anterior y a 2004 y 2019 como años que se encuentran entre ambos regímenes, como años transicionales (*transient*).

**Polinomial 2.** Esta distancia considera las diferencias de la conectividad ampliada a los sectores en sus entornos de segundo orden (v.g.: sectores vinculados con aquellos que tiene lazos directos) y los resultados se presentan en la Fig. 2.B1. Con esta distancia también se identifican dos grandes “regímenes” 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. 2.B1. Así, se agrupan las transiciones de 1996-2003 y de 2004-2019 en dos grandes regímenes 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 del “segundo régimen” de estabilidad 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. 1, 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. 2.B2). 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 de conectividad ampliadas en la proyección MDS de la Fig. 2.B2, tienden a estabilizarse en el período 2011-2017 identificado en la sección 1 (Fig. 1, período en verde) como de “estancamiento” (Fig. 2.B2, en rojo).

**Espectral.** La distancia espectral cuantifica las diferencias en todo el espectro de frecuencias de los valores propios de los grafos asociados a cada red de flujos interanuales. 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. 2.C1 identifica los mismos grandes grupos de períodos (1996-2000 y 2004-2018) 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 (que computa el más largo de los caminos promedio entre cada par de nodos) resulta del orden de 3 pasos (ver [3]), 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 la polinomial. No obstante, dado que para calcular la distancia espectral se utilizan todos los autovalores de los grafos, aparece información adicional que proviene de las interacciones de todos los sectores en conjunto.

En la proyección MDS (Fig. 2.C2) 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), prácticamente la totalidad de la variabilidad y el doble que la explicada en el caso de la distancia polinomial. Allí se observa que el eje horizontal (Comp 1) ordena hacia la izquierda a los años del primer régimen de movilidad (Convertibilidad y colapso, 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, atípicos y con máxima distancia  $l_2$  entre sí, y hacia abajo a los años de transición entre régimenes y los extremos 2018 y 2019.

**D-measure.** La distancia D-measure,  $d_D(G, \tilde{G})$ , identifica y cuantifica adecuadamente las diferencias topológicas estructurales, que afectan el flujo de información a través de las redes. Esto se aprecia en la Fig. 2.D, en que las diferencias topológicas crecientes corresponden a valores  $d_D$  más altos.

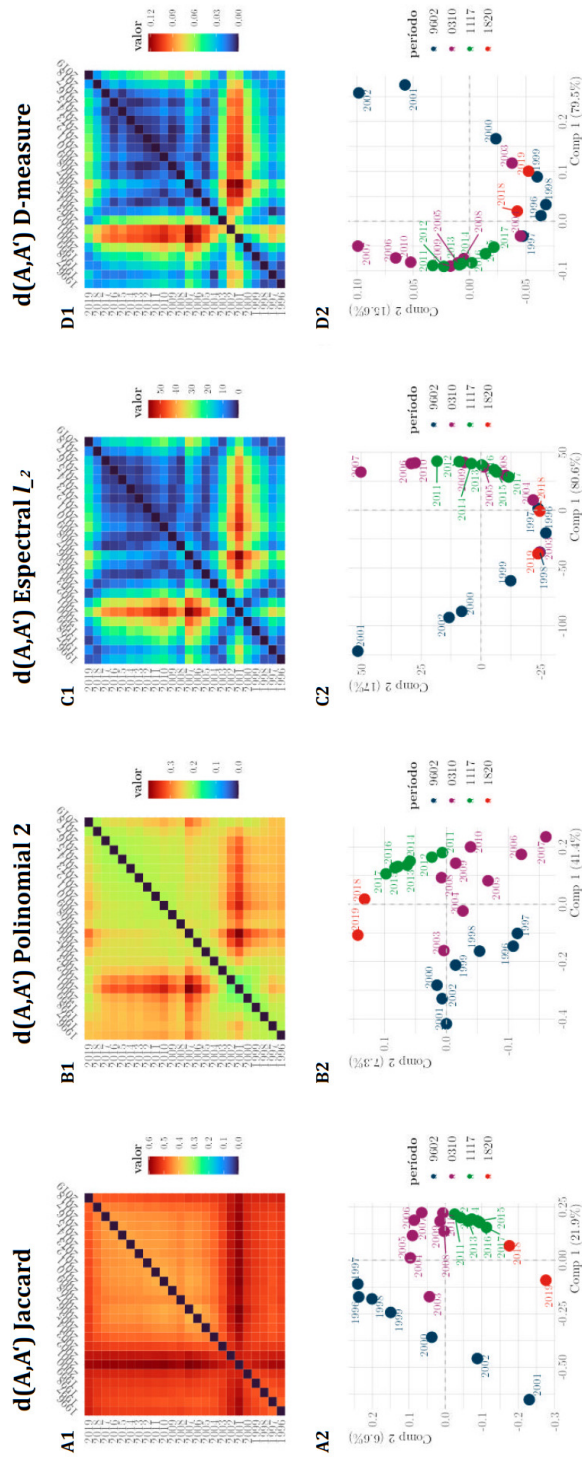


Figura 2: Distancias de grafos binarios no dirigidos de redes interanuales 1996-2019. En columnas A-D se representan las matrices de las cuatro distancias utilizadas para comparar las matrices de adyacencia correspondientes. De izquierda a derecha: Jaccard, Polinomial de segundo grado, Espectral, y D-measure, que se utilizan para comparar la estructura de conectividad local (de primero y de segundo orden), global y una combinación de métricas locales y de mayor generalidad (D-measure). En la fila superior se presentan (A1-D1) mapas de calor en los que cada celda representa el valor de la distancia seleccionada entre las redes (en orden cronológico) y en la fila inferior se presentan (A2-D2) gráficos de escalamiento multidimensional métrico (MDS) sobre los primeros dos componentes, utilizando colores para remarcar los periodos agrupados por clustering jerárquico (completo) calculado en la matriz de D-measure, con  $k = 5$  (corte en dendrogramas). En los mapas de calor se puede observar que Jaccard permite identificar el periodo de crisis 2001-2002 como diferente a la estructura de las redes de años anteriores y posteriores. La distancia Polinomial, que compara estructuras de conectividad directa e indirecta de segundo orden, permite identificar dos subperiodos (antes y después de 2007) en el régimen posterior a la crisis de 2001. Las distancias espectral  $l_2$  de valores propios y D-measure presentan mayor capacidad para identificar diferencias en la estructura de conectividad, con mayores niveles de detalle y sensibilidad que las otras métricas en la identificación de subregímenes en periodos de estabilidad y de contrastes en los periodos transicionales. En las representaciones MDS se destaca el paralelismo entre la mayor amplitud de la conectividad capturada por cada distancia (creciente en amplitud en sentido A-D) y la proporción de variabilidad de las matrices de distancias capturada por los primeros dos componentes de MDS, que van desde 28,5% con Jaccard hasta 97,6% y 95,1% con la espectral y D-measure, respectivamente.

Aunque esta disimilitud promedia el aporte de dos métricas locales y una global, tanto en el mapa de calor asociado a esta distancia (Fig. 2.D1) como en la proyección MDS (Fig. 2.D2) los resultados de la D-measure se asemejan a los obtenidos con la distancia espectral, que presenta un ordenamiento muy similar de las redes interanuales en ambas representaciones. Esta similitud posiblemente se explique en buena medida por las características de mundo pequeño y diámetro reducido de las redes analizadas. En el mapa de calor (Fig. 2.D1) se observa una distinción en la identificación del año 2002 como el año más diferente a todos, dentro del período transicional 2000-2002 que ya se detectaba con la distancia espectral. La proyección MDS (Fig. 2.D2) acumula una variabilidad explicada de 95,1 %, ligeramente inferior respecto de la espectral. Se observa 2002 como el período más atípico hacia la izquierda y arriba y un ordenamiento similar en los ejes horizontal (Comp 1, 79,5 %) y vertical (Comp 2, 15,6 %).

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, intermedia y global) que los períodos seleccionados aparecen representando distintos “régimenes” de movilidad del empleo, separados entre sí por períodos transicionales y cambios de régimen.

## 5. Discusión

Las redes interindustriales de empleo representan interacciones entre sectores de actividad mediados por cambios en el empleo. Esta representación pone énfasis en la relevancia de la estructura de conectividad intersectorial para el análisis de la movilidad laboral. En particular, la movilidad de los individuos que cambian de empleo (*job-to-job transitions*) hacia una industria distinta a la de origen refleja información acerca de la valorización de habilidades adquiridas por una trabajadora en una industria diferente.

Con el objeto de cuantificar las diferencias estructurales de conectividad intersectorial y de identificar el potencial de esta representación para extraer información útil del mercado laboral, en este trabajo se abordó el problema de comparar dichas redes para la Argentina en un período que contiene al menos un cambio estructural sustancial.

En particular, aquí nos interesa examinar sólo los cambios de la estructura que conecta a las industrias, donde las transiciones de empleo tienen lugar entre diferentes actividades económicas. En nuestro análisis dejamos de lado los efectos del tamaño y la representatividad de los flujos de trabajadores que se intercambian entre estas actividades dado que la magnitud de ellos no afecta la estructura de intercambios. El análisis de los flujos de empleo, es decir de las redes interindustriales de empleo *ponderadas*, queda para trabajos futuros donde sería interesante en este sentido abordar el análisis de la naturaleza y las características particulares de las transiciones laborales en cada año que permita atender preguntas como: el tipo de actividades involucradas en transiciones con mayor (menor) frecuencia, las actividades “atractoras” (“expulsoras”) de empleo, así

como los agrupamientos de actividades que habitualmente intercambian empleo entre sí.

Mediante el uso de un conjunto de distancias que difieren entre sí en el alcance de la estructura de conectividad entre nodos considerada, para las redes analizadas se observó que: a) al considerar únicamente vecindarios directos (Jaccard) es posible distinguir la presencia de estructuras de largo plazo (regímenes) detrás de los intercambios de empleo intersectoriales, así como períodos de cambio estructural; b) al incorporar vecindarios de conectividad indirecta (polinomial de segundo orden) se logran identificar subregímenes y períodos transicionales con mayor detalle; y c) al considerar distancias más globales, que evalúan la conectividad de la red en su conjunto, se identifican detalladamente los parecidos entre las estructuras interanuales en términos de regímenes, subregímenes, períodos transicionales y cambios de régimen, a la vez que se explica prácticamente la totalidad de la variabilidad informada por las métricas utilizadas.

Por otro lado, se encontró que la información obtenida al analizar las distancias permite interpretar los resultados a la luz de la evolución histórica de la actividad macroeconómica, tanto en términos de las fases cíclicas que caracterizan “regímenes” estructurales como de los períodos de transición y cambio de régimen asociados a crisis económicas que imponen redireccionamientos en la estructura de relaciones intersectoriales.

Estos hallazgos resultan interesantes en tanto que aportan un elemento adicional, la *conectividad interindustrial*, para el análisis de la estructura y la evolución del mercado laboral e incluso para el análisis de la estructura productiva. Las distancias utilizadas aportan una manera resumida de comparar estructuras complejas de interacción como las que surgen de las redes de empleo interindustrial.

El ejercicio realizado sugiere que la información de conectividad entre actividades económicas al nivel analizado (cuatro dígitos del clasificador CIU Rev. 3) resulta relevante para caracterizar la evolución de la estructura del mercado laboral y, posiblemente, de la estructura productiva subyacente. En particular, dado que los datos provienen de registros administrativos parece útil su aplicación para el monitoreo de la situación comparativa y la evolución del mercado laboral.

## Material Complementario

En este Material Complementario describimos brevemente cómo se calculan las distancias utilizadas en el análisis de la comparación de redes interanuales realizado. Las distancias se calculan sobre las matrices de adyacencias binarias,  $A_{i,j} \in \{0, 1\}$  de los grafos no dirigidos asociados a las redes correspondientes.

*Jaccard*. La distancia de similitud de Jaccard identifica cambios en las redes a nivel de entornos locales de interconexión de cada sector, y se calcula en la siguiente manera:

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

donde  $G$ , y  $\tilde{G}$  son dos redes que se comparan,  $A_{ij}$ , y  $\tilde{A}_{ij}$  las matrices fundamentales de las redes correspondientes. La distancia varía de 0 a 1, si su valor es cercano a 1 indica una reestructuración completa de la red entre dos periodos.

*Distancia Polinomial.* La distancia polinomial utiliza polinomio de potencias de la matriz de adyacencias,  $A^k$ . Dicho polinomio contiene una gran cantidad de información sobre la matriz, los más significativos son los valores propios, y que se relaciona con la topología local del grafo a través de los coeficientes de  $A_{ij}^k$  que representan los caminos del nodo  $i$  al nodo  $j$  a través de  $k$  saltos. 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.

Debido a que el camino promedio más largo (diámetro) de las redes de flujos de empleo es en promedio de 3 pasos (ver [3]), utilizamos la distancia polinomial con  $k = 2$ . Para calcular esta distancia se parte de la descomposición de valores propios de la matriz de adyacencias  $A$  de una red,  $A = Q\Lambda_A Q^T$ , donde  $\Lambda_A$  es la matriz diagonal con los valores propios de la matriz  $A$ , mientras que  $Q$  es la matriz cuadrada con los vectores propios asociados.

$P(A) = QWQ^T$ , donde

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

y

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

$V$  son nodos, y que para el caso de uso con  $k = 2$  se expresa:

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

Esta distancia permite cuantificar diferencias entre dos redes en términos de la estructura de conectividad que presenta cada sector con el vecindario de sectores con los que tiene intercambios directos, ampliado con las conexiones directas de estos con el resto de la red.

*Distancia espectral.* Esta distancia se utiliza para caracterizar el estado de la red mediante sus valores propios, y detectar cambios en estructuras globales e intermedias. Al efectuar el análisis espectral (*eigen-análisis*) 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 redes y se expresa de la siguiente forma:

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

Se utiliza la distancia  $l_p$  con  $p = 2$ , la distancia Euclídea, sobre el espectro de las matrices  $A$  de adyacencias de las redes 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.

*D-measure.* La disimilitud  $d_D(G, \tilde{G})$  entre  $G$  y  $\tilde{G}$  propuesta por Schieber et al. [9] viene dada por:

$$d_D(G, \tilde{G}) = \omega_1 \sqrt{\frac{\mathcal{J}_{\mathcal{H}}(\mu_G, \mu_{\tilde{G}})}{\log_2}} + \omega_2 \left| \sqrt{NND(G)} - \sqrt{NND(\tilde{G})} \right| + \frac{\omega_3}{2} \left( \sqrt{\frac{\mathcal{J}_{\mathcal{H}}(P_{\alpha G}, P_{\alpha \tilde{G}})}{\log_2}} + \sqrt{\frac{\mathcal{J}_{\mathcal{H}}(P_{\alpha G^c}, P_{\alpha \tilde{G}^c})}{\log_2}} \right), \quad (4)$$

donde  $G^c$  indica el complemento del grafo,  $\mathcal{J}_{\mathcal{H}}(\cdot, \cdot)$  es la divergencia Jensen-Shannon, una medida de disimilitud entre las distribuciones de probabilidad de distancias de conectividad de cada nodo  $i$  de red,  $\mathcal{P}_i$ ,  $i \in (0 \dots V)$ . Los pesos  $\omega_1$ ,  $\omega_2$  y  $\omega_3$  son arbitrarios, con  $\omega_1 + \omega_2 + \omega_3 = 1$ , y  $NND$  es una medida de la heterogeneidad de la red en términos de distancias de conectividad:

$$NND(G) = \frac{\mathcal{J}(\mathcal{P}_1 \dots \mathcal{P}_N)}{\log(d+1)},$$

siendo  $d$  el diámetro de la red.

El primer término, a través de la divergencia de Jensen-Shannon ( $\mu_G$  y  $\mu_{\tilde{G}}$ ), compara la distribución de distancias entre las dos redes. El segundo término analiza la heterogeneidad de los nodos comparando los valores de NND de las redes. Las redes que presentan el mismo NND son redes que tienen los mismos perfiles de distancia de conectividad. El tercer término compara medidas de centralidad alfa de ambas redes. Se considera la centralidad de cada nodo, teniendo en cuenta las conexiones directas e indirectas de cada nodo. Al considerar el complemento del grafo, la medida captura el efecto de los nodos desconectados. Así definida, la métrica  $d_D(G, \tilde{G})$  captura las disimilitudes globales y locales de las redes, y se espera que esta medida identifique adecuadamente las diferencias topológicas estructurales, que afectan el flujo de información a través de las redes, en caso de estudio, el flujo de empleo.

## Referencias

1. Causa, O., Luu, N., Abendschein, M.: Labour market transitions across oecd countries: Stylised facts. Tech. rep., OECD (2021). <https://doi.org/10.1787/62c85872-en>
2. Conte, D., Foggia, P., Sansone, C., Vento, M.: Thirty years of graph matching in pattern recognition. *Int. J. Pattern Recogn.* **18**, 265–298 (2004)
3. De Raco, S.A., Semeshenko, V.: Labor mobility and industrial space in Argentina. *Journal of Dynamics & Games* **6**(2), 107 (2019). <https://doi.org/10.3934/jdg.2019008>
4. De Raco, S.A., Semeshenko, V.: The network structure of inter-industry labor mobility in Argentina. In: 6th Regulating for Decent Work Conference. ILO, Geneva (2019)
5. Dehmer, M., Emmert-Streib, F., Kilian, J.: A similarity measure for graphs with low computational complexity. *Appl. Math. Comput.* **182**, 447–459 (2006)
6. Donnat, C., Holmes, S.: Tracking network dynamics: A survey using graph distances. *The Annals of Applied Statistics* **12**(2), 971–1012 (2018). <https://doi.org/10.1214/18-AOAS1176>
7. Gao, X., Xiao, B., Tao, D., Li, X.: A survey of graph edit distance. *Pattern Anal. Appl.* **13**, 113–129 (2010)
8. Gianelle, C.: Discovering the regional small world of labour mobility. Evidence from linked employer–employee data. *Regional Studies* **48**(7), 1263–1278 (2014). <https://doi.org/10.1080/00343404.2012.697993>
9. Schieber, Tiago A. and Carpi, L., Díaz-Guilera, A., Pardalos, P.M., Masoller, C., Ravett, M.G.: Quantification of network structural dissimilarities. *Nat Commun* **8**(13928) (2017). <https://doi.org/10.1038/ncomms13928>
10. Semeshenko, V., De Raco, S.A.: Analysis of the evolution of labor market flows in Argentina. *Proceedings 50 JAIIO-AGRANDA* pp. 20–24 (2021)