

Trabajo Fin de Máster
Máster Universitario en Organización Industrial y
Gestión de Empresas

Un análisis de la movilidad laboral basado en la
Muestra Continua de Vidas Laborales

Autor: Eva M^a Rodríguez Guerrero

Tutor: Fernando Núñez Hernández

Dpto. Organización Industrial y Gestión de
Empresas I
Universidad de Sevilla

Sevilla, 2022



Trabajo Fin de Máster
Máster Universitario en Organización Industrial y Gestión de Empresas

Un análisis de la movilidad laboral basado en la Muestra Continua de Vidas Laborales

Autor:

Eva M^a Rodríguez Guerrero

Tutor:

Fernando Núñez Hernández

Profesor Titular de Universidad

Dpto. Organización Industrial y Gestión de Empresas I

Escuela Técnica Superior de Ingeniería

Universidad de Sevilla

Sevilla, 2022

Trabajo Fin de Máster: Un análisis de la movilidad laboral basado en la Muestra Continua de
Vidas Laborales

Autor: Eva M^a Rodríguez Guerrero

Tutor: Fernando Núñez Hernández

El tribunal nombrado para juzgar el Proyecto arriba indicado, compuesto por los siguientes miembros:

Presidente:

Vocales:

Secretario:

Acuerdan otorgarle la calificación de:

Sevilla, 2022

El Secretario del Tribunal

A mi familia y pareja
A mis profesores

Agradecimientos

Gracias Fernando Núñez Hernández, tutor de este Trabajo Fin de Máster por haber sido mi guía y mi estímulo tanto en la investigación como durante su redacción

A Pablo Álvarez de Toledo por su acompañamiento, su energía y su apoyo durante la elaboración de este Trabajo

Gracias a mis padres, a mi hermana y a mi pareja, sólo puedo expresar mi sincero agradecimiento por apoyarme durante la etapa académica que hoy culmina.

Resumen

El presente Trabajo de Fin de Máster propone un análisis de la movilidad laboral en el mercado de trabajo español. Empleando microdatos de la Muestra Continua de Vidas Laborales para los años 2019 y 2020, tratamos de identificar los patrones de movilidad de los trabajadores dados de alta en la Seguridad Social en dichos años. En concreto, estamos interesados en conocer las secuencias de cambios que se producen, a nivel de actividad económica, de grupo de ocupación y de provincia, cuando el trabajador cambia de empleo. La herramienta metodológica empleada ha sido la tabla de contingencia, la cual permite tabular el cruce de categorías dos variables nominales. En nuestro caso, dichas variables serán la actividad, la ocupación o la provincia del empleo actual (filas de la tabla) y del empleo anterior (columnas de la tabla); dichas movilizaciones serán analizadas por separado. A partir de dicha tabla, obtenemos unos mapas de calor (*heatmap*) que miden el grado de asociación entre las diferentes categorías de las variables analizadas. Las filas y columnas del *heatmap* se ordenan siguiendo una metodología de cluster jerárquico. Nuestro estudio ofrece resultados interesantes, ya que muestra aquellas movilizaciones laborales con un mayor grado de asociación y describe dichas movilizaciones en términos del individuo y del puesto que ocupa.

Abstract

This Master's thesis proposes an analysis of labor mobility in the Spanish labor market. Using microdata from the Continuous Sample of Working Lives for the years 2019 and 2020, we try to identify the mobility patterns of workers registered with the Social Security in those years. Specifically, we are interested in knowing the sequences of changes that occur occupation group and province at the level of economic activity when the worker changes employment. The methodological tool used was the contingency table, which allows us to tabulate the crossing of categories and two nominal variables. In our case, these variables will be the activity, occupation or province of the current job (rows of the table) and of the previous job (columns of the table); these mobilities will be analyzed separately. From this table, we obtain *heatmaps* that measure the degree of association between the different categories of the analyzed variables. The rows and columns of the *heatmap* are ordered following a hierarchical cluster methodology. Our study provides interesting results, as it shows those labor mobilities with a higher degree of association and describes these mobilities in terms of the individual and the position held.

Índice

<i>Agradecimientos</i>	9
<i>Resumen</i>	11
<i>Abstract</i>	13
<i>Índice</i>	15
<i>Índice de Tablas</i>	17
<i>Índice de Figuras</i>	19
1 INTRODUCCIÓN	21
2 REVISIÓN DE LA LITERATURA	25
3 MARCO TEÓRICO. MODELO DE CAPITAL HUMANO	29
EDUCACIÓN Y CAPITAL HUMANO	29
INTRODUCCIÓN	29
3.1.1 Algunos hechos	29
3.1.1.2 Tasas de graduación	29
3.1.1.3 Educación y resultados obtenidos en el mercado laboral	30
3.1.2 La teoría del capital humano	30
3.1.2.1 Relación entre las retribuciones y el capital humano	30
3.1.2.2 Duración de los estudios y retribuciones	31
3.1.2.3 Educación, formación y retribuciones a lo largo del ciclo de vida	32
3.1.3 La educación como señal	33
3.1.4 Identificación de la relación causal entre educación y retribuciones	34
3.1.4.1 La teoría del capital humano: del modelo a las estimaciones	34
3.1.4.2 El problema de la selección	34

3.1.5.1 Rendimientos privados de la educación.....	37
3.1.5.2 Rendimientos privados de la educación no pecuniarios	39
3.1.5.3 Rendimientos sociales de la educación	39
3.1.5.4 ¿Qué realmente es importante en la educación?	39
4 METODOLOGÍA ESTADÍSTICA. TABLAS DE CONTINGENCIA Y MAPAS DE CALOR	43
5 ORIGEN DE LOS DATOS. LA MUESTRA CONTINUA DE VIDAS LABORALES.....	49
6 APLICACIÓN	53
6.1 Secuencias de Actividades.....	54
6.2 Secuencias de Ocupaciones.....	63
6.3 Secuencias de Provincias.....	70
7 CONCLUSIONES.....	79
REFERENCIAS.....	81
ANEXOS	85

Índice de Tablas

Tabla 1. <i>Tabla de contingencia</i>	43
Tabla 2. <i>Ejemplo de mapa de calor y factor de asociación entre distintos sectores</i>	47
Tabla 3. <i>Tabla total colocaciones de secuencias de actividades</i>	55
Tabla 4. <i>Mapa de calor de actividades económicas</i>	56
Tabla 5. <i>Análisis celdas educación y sanidad</i>	58
Tabla 6. <i>Análisis celdas AAPP y agricultura</i>	60
Tabla 7. <i>Análisis celdas agricultura y construcción</i>	61
Tabla 8. <i>Análisis celdas construcción y suministros</i>	62
Tabla 9. <i>Tabla total colocaciones de secuencias de ocupaciones</i>	63
Tabla 10. <i>Mapa de calor de ocupaciones</i>	65
Tabla 11. <i>Análisis celdas ingenieros, licenciados, alta dirección y ingenieros técnicos, ayudantes titulados</i>	66
Tabla 12. <i>Análisis celdas ingenieros, licenciados, alta dirección y jefes administrativos y de taller</i>	68
Tabla 13. <i>Análisis celdas ingenieros licenciados, ingenieros técnicos y jefes administrativos y de taller</i>	69
Tabla 14. <i>Tabla total colocaciones de secuencias de provincias</i>	70
Tabla 15. <i>Mapa de calor de provincias</i>	72
Tabla 16. <i>Análisis celdas Madrid y Barcelona</i>	74
Tabla 17. <i>Análisis celdas Almería y Murcia</i>	75
Tabla 18. <i>Análisis celdas Murcia y Alicante</i>	76
Tabla 19. <i>Análisis celdas Cáceres y Badajoz</i>	77
Anexo 1. <i>Lista de agrupaciones de actividades económicas (CNAE-09)</i>	85
Anexo 2. <i>Mapa de calor de actividades económicas 2019</i>	86
Anexo 3. <i>Mapa de calor de ocupaciones 2019</i>	86
Anexo 4. <i>Mapa de calor de provincias 2019</i>	87

Índice de Figuras

Figura 1 <i>Aumentos de salarios medios en graduados de instituto o universidad en EE. UU. en 1996.</i>	32
Figura 2 <i>La ley del movimiento del tiempo dedicado (a la izquierda), stock de capital humano (derecha)</i>	33
Figura 3. <i>Trimestre de nacimiento y años de educación completa para los hombres nacidos en 1930-1939</i>	35
Figura 4. <i>Tasas interna de retorno</i>	38
Figura 5. <i>Tasa de rendimiento de la inversión en capital humano por edad</i>	40
Figura 6. <i>Ejemplo de obtención de un bicluster con estructura diagonal</i>	46
Figura 7. <i>Histograma de rendimientos de trabajo</i>	51
Figura 8. <i>Dendrograma de actividades</i>	55
Figura 9. <i>Dendrograma de ocupaciones</i>	64
Figura 10. <i>Dendrograma de provincias</i>	71

1 INTRODUCCIÓN

El presente Trabajo Fin de Máster busca aportar otra visión al análisis de las movilidades laborales que se producen en el mercado laboral español. El estudio se ha limitado a las movilidades entre actividades económicas, entre ocupaciones y entre provincias en España en el año 2020, aunque también analizaremos el año 2019 a efectos comparativos. Esta investigación laboral pretende dar información sobre la realidad del mercado laboral en la actualidad. El análisis tiene como objetivo identificar las movilidades laborales (regionales, de actividad y ocupacionales) más habituales y describirlas en términos del trabajador y de los puestos que este ocupa (tanto el de origen como el de destino) –contamos incluso con una *proxy* de la retribución del individuo–. Esperamos que esta fotografía de la movilidad laboral en España pueda orientar a los buscadores de empleo sobre las condiciones en que se produce de cada tipo de movilidad.

Los algoritmos propuestos se van a aplicar a los datos procedentes de la Muestra Continua de Vida Laboral (MCVL). La MCVL contiene un conjunto de datos individuales anonimados, procedentes de las bases de datos de la Seguridad Social, a los que se añaden otros datos procedentes del Padrón Continuo Municipal (INE) y del resumen anual de retenciones e ingresos a cuenta del IRPF (Modelo 190) de la Agencia Tributaria. Los datos corresponden a una muestra aleatoria simple de personas que fueron afiliados o pensionistas de la Seguridad Social durante el año de referencia. Para cada persona se incluyen tanto datos sobre su relación con la Seguridad Social en dicho año como datos históricos, en la medida en que se conserven en registros informatizados. Para los individuos que obtienen rendimientos del trabajo conocemos ciertas características suyas (como su género, edad, nivel de estudios o renta bruta) y del puesto que ocupan (como el grupo de ocupación y la provincia del puesto o el sector de actividad económica de la empresa). Conociendo estas características de los empleos, podemos obtener una visión global sobre las trayectorias laborales de los individuos de la Muestra que cambian de empleo dentro del año natural.

La metodología empleada se basa en la utilización de tablas de contingencia (TC), que son matrices que permiten tabular el cruce de categorías de dos o más variables nominales. En nuestro caso, emplearemos la TC para tratar de relacionar los sucesivos episodios de empleo de un trabajador en los que ha habido algún tipo de movilidad laboral. Es decir, nuestra TC sitúa en las filas la actividad o la ocupación o la provincia del empleo anterior y en las columnas la actividad o la ocupación o la provincia del empleo actual; estas tres TC (una para cada variable: actividad, ocupación, provincia) se convierten en matrices origen-destino que permiten contar cuántos tipos de movilidad (o cruce de categorías) se observan en la muestra –los tres tipos de movilidad se van a analizar por separado–. A partir de la TC, empleamos dos medidas que son el centro de nuestro análisis categórico: la similitud entre dos categorías-fila (o dos categorías-columna) y el factor de asociación entre una categoría-fila y una categoría-columna.

La primera medida, la de proximidad entre dos categorías de una variable, nos va a permitir generar clusters de categorías cuya representación gráfica recibe el nombre de dendrograma. La segunda medida, el factor de asociación, va a medir el grado de correspondencia o vinculación entre una determinada fila y una determinada columna de la tabla de contingencia, generando así un mapa de calor o *heatmap*. Con dicho factor, medimos el grado de asociación entre la actividad económica del empleo de origen (filas de la TC y del *heatmap*) y la actividad del empleo de destino (columnas de la TC y del *heatmap*) del individuo, y análogamente hacemos con las provincias y los grupos de ocupación de los puestos de origen y destino. Ofrecemos por tanto un análisis secuencial, que tiene un carácter dinámico, por cuanto analizamos sucesivos episodios de empleo de un mismo individuo a lo largo del tiempo, en concreto dentro de un año determinado. Asimismo, es importante tener presente que la base de cada dendrograma (actividades, ocupaciones y provincias) va a ser tomada como referencia para ordenar las filas y las columnas de la TC y del mapa de calor. Téngase en cuenta que dos elementos contiguos en la base del cluster o dendrograma se parecen más entre sí que dos elementos lejanos. Dicha similitud se basa en la forma en que esos dos elementos o categorías (provincias de origen, por ejemplo) se asocian o emparejan con las categorías de destino (provincias de destino). Por ejemplo, si las provincias de origen (provincias de la colocación anterior), Sevilla y Cádiz, hacen correspondencia fundamentalmente con las mismas provincias de destino (provincias de la colocación actual), pensemos en Sevilla, Cádiz, Madrid y Barcelona, dichas provincias de origen (Sevilla y Cádiz) se parecerán más entre sí en el cluster de provincias que Sevilla (o Cádiz) con otra provincia de origen (Álava, por ejemplo) que no muestra correspondencias fuertes con las cuatro provincias de destino indicadas.

Entre los resultados obtenidos en este Trabajo, destacamos el papel de Madrid como centro receptor de las movilidades regionales; esto es, la mayoría de las provincias tiene una mayor o menor tendencia a emparejarse con esta provincia a la hora de un traslado laboral. Por otro lado, en este estudio se dará a conocer qué tipo de movilidades de sectores se están produciendo, como es el caso de la educación a la sanidad. Además, en cuanto a la movilidad ocupacional, se verá la alta relación de movilidad entre los más altos cargos dentro del mercado laboral.

El resto del trabajo se estructura como sigue. En el segundo capítulo se lleva a cabo la revisión de la bibliografía, donde se sitúa este trabajo dentro de las investigaciones ya realizadas en este campo. En el tercer capítulo, se desarrolla el marco teórico que sirve de fundamento al trabajo. El modelo del capital humano parte de la premisa de que toda inversión educativa generará unas ganancias monetarias (privadas y sociales) en el futuro. Posteriormente en el capítulo 4, se describen las herramientas metodológicas que se van a utilizar para el estudio empírico propuesto, concretamente explicaremos las tablas de contingencias y los mapas de calor. En el capítulo quinto, se describe el origen de los datos empleados y se realiza una breve introducción a la Muestra Continua de Vidas Laborales. Posteriormente, en el capítulo 6 se lleva a cabo la aplicación del estudio: el análisis de las movilidades laborales. Se analizan (por separado) tres secuencias de movilidad distintas, secuencias de actividad, de ocupación y de provincia. Dentro de cada una se analizarán los casos más destacables, es decir, las combinaciones origen-destino

con un mayor grado de asociación. Finalmente, en el capítulo 7 se recopilan las conclusiones a las que los resultados de esta investigación han llegado. Además de las referencias bibliográficas, al final del Trabajo de Fin de Máster, en un Anexo, se han añadido los mapas de calor pertenecientes al año 2019 (a efectos comparativos con el año de estudio, 2020).

2 REVISIÓN DE LA LITERATURA

Para realizar la revisión bibliográfica del presente Trabajo Final de Máster sobre el estudio de movilidades en cuanto a sectores de actividad, ocupación o provincias, se analizarán investigaciones actuales sobre el tema para situar este trabajo dentro de los artículos más relevantes ya publicados. Se destacarán aquellos requisitos que potencian a los egresados a un mejor acceso a empleos más cualificados.

Se da la casualidad de que las economías española y andaluza poseen un mercado laboral que presenta un problema caracterizado por un nivel de desempleo relativamente alto y persistente en el tiempo (García-Cintado *et al.*, 2014, Usabiaga y Hernández-Salmerón, 2021). Como comentan autores como Usabiaga (2004) y López-Bazo y Motellón (2013), la economía andaluza presenta un dato de desempleo desfavorable frente a otras regiones del país y a su vez, muestra una estructura laboral parcialmente segmentada. Por su parte, se estima que existe un tipo de segmentación aplicable al mercado laboral descrito por el enfoque insider-outsider–mercado laboral dual– (Davía y Hernanz, 2004; Bentolila *et al.*, 2012), que permite distinguir entre trabajadores permanentes y temporales/desempleados, así como, permite evidenciar otros tipos de segmentación como es el caso de las aplicables en el caso andaluz nativos versus inmigrantes, segregación por género, etc.). El mercado andaluz se caracteriza por poseer diferentes tipos de desajustes, como es el caso de los términos de formación (y competencias) o de movilidad geográfica y ocupacional. Es en este punto, donde se debe destacar el problema que acontece a la sobre-cualificación/sobre-educación de los jóvenes de hoy en día, como comentan Nieto y Ramos (2017) y que será analizado en el presente estudio. En otro orden de cosas, deberá de ser analizada también la rigidez salarial y la baja movilidad geográfica existente y documentada en la economía andaluza y española, respectivamente (Jimeno y Bentolila, 1998; Liu, 2018).

Como se observa se desarrollan numerosas fricciones y desajustes que finalmente se acaban reflejando de manera negativa en la curva Beveridge (relación vacantes-desempleo), así como, en la función de emparejamiento laboral de la economía de la región andaluza. Estas relaciones macroeconómicas tienen su fundamento teórico en el *modelo de búsqueda y emparejamiento laboral* – sobre estos modelos, véanse, por ejemplo, los trabajos de Pissarides (2000, 2011), Petrongolo y Pissarides (2001) y Shimer (2007)–. Según dichos estudios, la presencia de fricciones en el mercado laboral, y más concretamente, la ausencia de información perfecta o la heterogeneidad de trabajadores y de puestos libres, provocan que el emparejamiento laboral no se realice de forma instantánea. Dicho emparejamiento, se desarrolla por tanto como un proceso de búsqueda con información errática y con bastante lentitud y costes para ambas partes, empresas e individuos trabajadores.

Al desarrollar los modelos microeconómicos que se desarrollan en el campo del emparejamiento laboral, el presente estudio ayuda a relacionar al menos dos tipos de modelos: los que parten de

microfundamentar la función macroeconómica de emparejamiento mencionada con anterioridad y los que se fundamentan en la asignación bilateral (*two-sided matching games*). Los primeros buscan hacer explícitas las fricciones y heterogeneidades que acontecen en la función de emparejamiento añadida. Así, algunos modelos como entre los que destacan el de islas, bolas en urnas, taxis, colas, *stock-flow* o el de desajuste o mismatch, han buscado la manera de conseguir diferentes tipos de fricciones que permiten microfundamentar a la función de emparejamiento, subdividiendo de alguna manera el mercado de trabajo en diferentes partes o segmentos. Con respecto a los segundos (los modelos de *two-sided matching games*), constituyen una base que fundamenta como marco teórico al presente trabajo y permite analizar los datos individuales de colocaciones detectando “quién se empareja con quién” (*who matches with whom*); dicha asignación viene determinada por la función o tecnología de asignación bilateral basada en las preferencias de los propios individuos.

El presente trabajo no tiene como objetivo el ampliar o evaluar los modelos teóricos ya creados, si no más bien proponer una serie de medidas empíricas que logren tener implicaciones en el marco de dichos modelos. Como ejemplo, se debe destacar que el “mapa” que se genera de propensiones al emparejamiento (entre segmentos de trabajadores y segmentos de puestos), y la opción de identificar los segmentos asociados (de trabajadores o de puestos) con una elevada semejanza con respecto a las maneras en la que participan en el emparejamiento laboral. Estos a su vez, pueden ser interpretados como procesos que ayudan a mejorar la información obtenida sobre el empleo de los propios individuos; estas mejoras pueden llegar a producir una reducción del *mismatch* que ya existe en la función de emparejamiento o en la tecnología relacionada con los modelos de asignación bilateral—provocando de hecho una revisión de las preferencias que muestran los individuos en dichos “juegos” de asignación—.

Este Trabajo, se va a centrar en analizar aquellas variables como la formación del individuo, su género y edad, la zona de pertenencia del puesto de trabajo, el sector de actividad y grupo de ocupación del puesto, etc. La mayor parte de variables de esta investigación tienen carácter categórico. Para el análisis de datos categóricos, se genera una tabla de contingencia con la frecuencia de cada celda correspondiente a la combinación de las diferentes categorías de m variables descriptivas. Como confirman Fienberg y Rinaldo (2007), la historia de las tablas de contingencia se remonta al menos hasta el siglo XIX, y actualmente sigue existiendo su área activa de investigación. Las tablas de contingencia se utilizan en diferentes campos como la genética, las encuestas sociológicas, la investigación de mercados, etc. Con los recientes avances en la tecnología de gestión de datos, las dimensiones de las tablas de contingencia crecen con el número de variables categóricas, dando lugar a problemas específicos de tablas de contingencia de elevada dimensión (Petitjean et al., 2013).

En cuanto a la rama de conocimiento económica, es muy común el uso de las tablas de contingencia. Seguramente los mercados de emparejamiento sean el tema económico más estrechamente relacionado con las tablas de contingencia. Desde el estudio de los enlaces matrimoniales, los modelos de emparejamiento han ampliado su campo de aplicación para incluir, por ejemplo, temas de movilidad ocupacional (Long y Ferrie, 2013), especialización sectorial y concentración regional (Haedo y Mouchart,

2018), rendimientos bursátiles (Rey et al., 2014), admisiones universitarias, asignación de préstamos bancarios, trasplantes de órganos e infinidad de casos más.

Con el objetivo de mejorar la información estadística disponible sobre el grado de desajuste laboral. Hoy en día existe un gasto relativamente bajo en Políticas Activas de Empleo, así como una evaluación insuficiente de su eficacia (Arranz et al., 2013). Por ejemplo, solo los programas específicos de formación han demostrado su utilidad. La mayoría de los recursos se destinan a incentivos a las empresas, con un alto efecto de “peso muerto”, mientras que existe un escaso papel de intermediación de las agencias públicas de empleo. En consecuencia, existe la necesidad de un mejor asesoramiento o tutorización de los desempleados en términos de formación útil, así como de movilidad ocupacional y geográfica. En este sentido, muchos investigadores solicitan datos más desagregados y estructurados y políticas de empleo específicas para cada región (Bande et al., 2019, Suárez et al., 2019). También hay propuestas para organizar las políticas laborales mediante unidades analíticas geográficas en lugar de administrativas. Por ejemplo, se sabe que vivir en grandes ciudades o cerca de áreas metropolitanas, o tener posibilidades de desplazamiento, aumenta la probabilidad de estar empleado (Viñuela et al., 2014). En este sentido, nuestra metodología puede ser muy útil, ya que somos capaces de detectar mercados laborales locales y relaciones laborales que no son fácilmente observables. Los “mapas laborales” a generar, con información revisada sobre el emparejamiento, pueden ser muy importantes en el diseño de una política específica y eficiente.

Existen numerosos autores que permiten conocer de primera mano como se ha de trabajar con las muestras continuas de vidas laborales (MCVL) para así poder obtener un mayor enfoque y manejo de la muestra de datos. (Lapuerta, 2010; Arranz y García-Serrano, 2011; Arranz et al., 2013). Otros temas que se tocan en la muestra son, por ejemplo, el análisis de la economía del conocimiento (Sánchez y Arellano, 2010) y la propuesta de calidad del empleo elaborada por (Arranz et al., 2018). A su vez, y tal y como comentan autores tras analizar las MCVL, el desempleo juvenil con anterioridad a los 30 años puede llegar a tener un impacto en la vida laboral de los trabajadores, ya que se tiende a vivenciar otro período de desempleo a lo largo de la vida laboral (Arrazola et al., 2018).

Como se analizará en el presente trabajo y a su vez comentan diversos autores, la movilidad geográfica laboral interior puede llegar a ser un mecanismo que ayude a combatir los elevados índices de desempleo, como ya se pudo observar tras algunas de las crisis económicas acontecidas en nuestro país, como la de 2007 (Prieto, 2017).

Otra propuesta estudiada por la literatura científica en la actualidad y que cobra sentido en términos de igualdad, es el análisis de la movilidad laboral de las mujeres inmigrantes en España, tal y como comentan diversos autores como Arranz et al. (2017) y Moreno-Colom y López-Roldán (2016). Por su parte, autores como Padilla y Planelles (2021) han enfocado su estudio de las MCVL desde el punto de vista salarial y de los trabajadores de tipología autónoma. En dicha investigación, analizan el salario medio de los andaluces y el nivel de ingresos que llegan a percibir los individuos andaluces dados de alta como autónomos en el régimen de la seguridad social por sexos y por franjas de edad. También existen otros

enfoques como puedan ser el elaborado por Cebrián y Toharia (2008), que analiza la entrada en el mercado de trabajo enfocando su análisis sobre la muestra (MCVL).

3 MARCO TEÓRICO. MODELO DE CAPITAL HUMANO

Como marco teórico para el presente Trabajo Final de Máster se ha utilizado el modelo de capital humano presentado por (P. Cahuc, 2004), en su capítulo 4. Education and Human Capital.

A continuación, se presenta un extracto de dicho capítulo, en el que se recogen aquellos puntos que nos han parecido de mayor interés.

EDUCACIÓN Y CAPITAL HUMANO

INTRODUCCIÓN

Según la teoría del capital humano la educación es una inversión que producirá ganancias en el futuro. El sistema educativo puede servir también para señalar las capacidades individuales a futuros empleadores. Si la educación sirve a la vez para adquirir conocimientos y para seleccionar a los individuos, debemos tratar de determinar el peso respectivo de cada una de estas dimensiones, no sólo para comprender el impacto de la educación en los ingresos y el crecimiento, sino también para evaluar la eficacia del gasto en educación. Las elecciones individuales son socialmente eficientes si el mercado laboral es competitivo y si la educación no produce externalidades.

3.1.1 Algunos hechos sobre la educación y capital humano

3.1.1.1 Gasto en educación

Por término medio, los países de la OCDE gastan el 6,3% de su PIB en instituciones educativas. El gasto oscila entre el 5% en Italia y el 8% en Dinamarca y Corea. En la mayoría de los países, la educación se financia principalmente con dinero público, por lo que el gasto en educación constituye hoy en día una importante partida presupuestaria.

3.1.1.2 Tasas de graduación

El porcentaje medio de la población en edad de trabajar que ha completado la educación secundaria en 2010 es del 74% para los países de la OCDE.

Los niveles educativos avanzan, ya que en todos los países la proporción de población con al menos estudios secundarios es mayor en la franja de edad de 25 a 35 años que en la de 55 a 64 años. En este

ámbito se observa un fenómeno de convergencia, ya que los países en los que las tasas de escolarización secundaria eran más bajas al principio han avanzado más rápidamente que los demás.

3.1.1.3 Educación y resultados obtenidos en el mercado laboral

En la OCDE en 2010 un trabajador con un nivel inferior a la secundaria superior recibe un salario equivalente al 77% del salario de un trabajador que ha alcanzado la educación secundaria superior. Los asalariados con un diploma de educación superior (nivel terciario) reciben un salario un 54% más alto que los que tienen un diploma de secundaria. Esto sugiere que adquirir una educación superior es una forma de elevar el salario, es decir, los niveles más altos de educación están relacionados con un mejor rendimiento del mercado laboral.

3.1.2 La teoría del capital humano

Adquirir competencias que el mercado laboral recompensará conlleva ‘costes de formación’ comparables a las inversiones que serán fuentes de ingresos futuros. Las diferencias salariales están influidas por las diferencias de productividad individual, que a su vez están influidas por las inversiones en educación o formación realizadas por los individuos a lo largo de su vida. Se supone que los agentes tienen la oportunidad de ampliar su formación a lo largo de toda su vida profesional.

3.1.2.1 Relación entre las retribuciones y el capital humano

La formación generalista, está asociada fundamentalmente al trabajador, que puede rentabilizarla en diferentes tipos de trabajo y hacer así que los empresarios compitan por sus servicios. Sin embargo, la formación específica está asociada a un tipo de trabajo concreto.

Se desarrolla un modelo del capital humano partiendo de una serie de supuestos: la población activa está formada por un continuo de trabajadores idénticos, cuyo tamaño se normaliza a 1; el primer período de vida se dedica únicamente a la educación y el segundo a producir; la tasa de descuento es igual a cero, es decir, no hay preferencia por el presente.

Equilibrio competitivo con formación generalista

Cada individuo tiene interés en invertir para mejorar su formación generalista siempre que el rendimiento marginal de esta inversión sea mayor que su coste marginal.

Los empresarios, por su parte, no tienen ningún incentivo para financiar este tipo de formación porque todo trabajador puede obtener un aumento de salario ofreciendo sus servicios a otros empresarios competidores tan pronto como aumente su productividad.

Equilibrio competitivo con formación específica

Cuando la formación es específica, los trabajadores sólo pueden rentabilizar su formación en un puesto de trabajo concreto. Una vez capacitados, no pueden exigir aumentos salariales a su empleador

haciéndolo competir contra otros empleadores. Por lo tanto, los empleadores pueden tener un incentivo para invertir en este tipo de formación.

Al igual que en el caso de la formación general, los trabajadores obtienen una renta igual a su productividad menos el coste de la inversión en formación

Las elecciones realizadas por los sujetos en el marco de la competencia perfecta conducen a la eficiencia social. Es decir, en una economía perfectamente competitiva, las elecciones individuales en materia de formación son socialmente eficientes.

La teoría del capital humano sugiere que los mecanismos de la competencia incentivan a los individuos a formarse con el fin de adquirir conocimientos o habilidades que el mercado prima y demuestra que las decisiones individuales en materia de educación son socialmente eficientes si el mercado laboral es perfectamente competitivo. Si los salarios son inferiores a la productividad porque las empresas disponen de un poder de monopsonio, la inversión en capital humano es inferior al óptimo social.

En resumen, la competencia en el mercado laboral permite a los trabajadores obtener valor, en forma de ingresos, de los conocimientos que mejoran su productividad. Por el contrario, en ausencia de competencia, desaparecen los incentivos a invertir para mejorar la productividad y es por ello por lo que generalmente, las empresas son las que invierten en formación específica.

3.1.2.2 Duración de los estudios y retribuciones

Se plantea la elección para un individuo entre si le conviene más seguir formándose o empezar a ganar dinero. En el modelo el sujeto puede adquirir educación a partir de la fecha $t=0$ y su vida laboral termina en la fecha $T > 0$. El factor de descuento es $r > 0$. En cada momento t es posible estudiar o trabajar, pero no hacer ambas cosas a la vez.

Se concluye que nunca hay interés en formarse si la tasa de descuento r es mayor que la eficiencia θ del esfuerzo educativo. Es por ello, que, si uno quiere adquirir educación extra, debe ser suficientemente paciente y los rendimientos de la educación deben ser suficientemente altos.

Una vez conocido el factor decisivo para que el individuo sepa si le conviene o no seguir con su formación educativa, es necesario saber cuál es la duración óptima de escolarización

La duración óptima de la escolarización aumenta con la duración de la vida T y con el parámetro de eficiencia. Por tanto, los individuos más eficientes son los que más tiempo dedican a la educación. La escolarización disminuye con la tasa de descuento r , lo que significa que los individuos más impacientes, o los que se enfrentan a mayores obstáculos financieros que aumentan el coste de los préstamos, deben estudiar durante períodos más cortos.

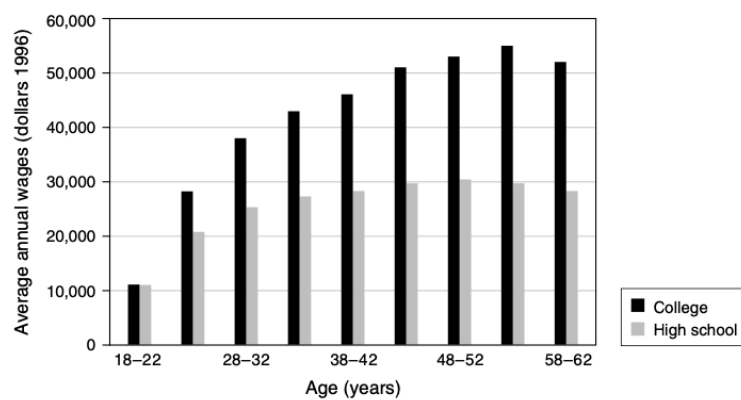
Podría ser óptimo no recibir ninguna formación o educación cuando el parámetro de eficiencia es demasiado pequeño, en cuyo caso el individuo conserva el mismo stock de conocimientos durante toda su vida.

El salario aumenta con la eficiencia de la inversión educativa. Cada período de educación aumenta el stock de capital humano en mayor medida, cuanto más eficiente es el individuo, y por otro, los individuos más eficientes estudian más tiempo. También el capital humano inicial influye en las ganancias del trabajo.

3.1.2.3 Educación, formación y retribuciones a lo largo del ciclo de vida

Por lo general, en todos los países desarrollados, para todas las profesiones, la relación entre la edad y los ingresos anuales del empleo a lo largo del ciclo vital presenta las mismas características, representado en el siguiente gráfico, ‘Life-Cycle-Earning’, el ciclo de vida de las ganancias.

Figura 1 Aumentos de salarios medios en graduados de instituto o universidad en EE. UU. en 1996



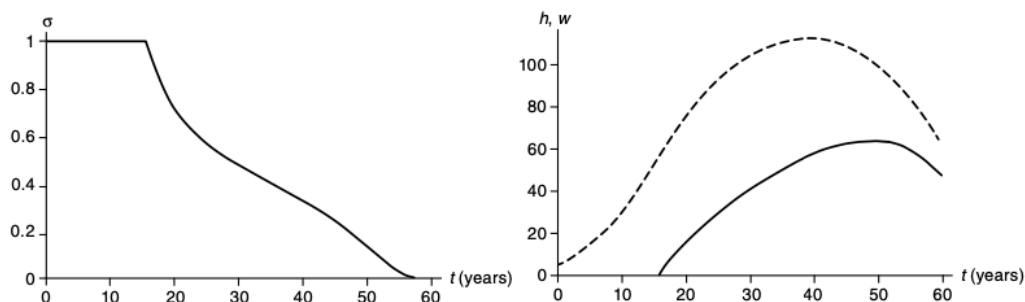
Fuente: Ashenfelter and Rouse (1999).

Tras un periodo inicial de formación en el que no se perciben ingresos salariales, esta curva es cóncava y alcanza un máximo entre los 45 y los 60 años, antes de reducirse gradualmente.

2.3.1 Modelo con formación a lo largo del ciclo vital

Este modelo permite replicar ingresos salariales a lo largo del ciclo de vida similares a los observados generalmente en la realidad. El modelo se calibra con datos anuales con un factor de descuento r que vale el 5%. El horizonte de 60 años de vida se justifica por la edad de jubilación, que es de 65 años en muchos países en el contexto temporal del modelo de Cahuc, y el inicio de la escolarización, que normalmente se produce a alrededor de los 5 años.

Figura 2 La ley del movimiento del tiempo dedicado (a la izquierda), stock de capital humano (derecha)



Fuente: P. Cahuc, 2004.

Estos gráficos reproducen la duración de la escolarización y la evolución de los ingresos salariales de los titulares de un título universitario en Estados Unidos. Estos resultados muestran que, en este modelo de capital humano, la heterogeneidad en las habilidades explica en gran medida tanto el comportamiento educativo como los ingresos laborales que se derivan de él. Cahuc demuestra que los individuos estudian a tiempo completo durante 16 años, pero después invierten cada vez menos en formación. El perfil de los ingresos salariales es creciente y cóncavo alcanzando un máximo de 60.000\$ unos 10 años antes de la jubilación.

Curiosamente, es posible representar la diferencia entre los comportamientos y los ingresos de los titulados universitarios y de los de secundaria modificando exclusivamente el valor del parámetro de eficiencia.

El modelo de capital humano en el que los individuos eligen el tiempo que desean dedicar a la formación reproduce muy bien la trayectoria temporal de los ingresos a lo largo del ciclo de vida. La teoría del capital humano se basa en la hipótesis de que las diferencias salariales reflejan las diferencias de productividad, que a su vez están influidas por la adquisición de competencias por parte de los trabajadores.

3.1.3 La educación como señal

Se puede proponer la idea de que la educación sirve también para seleccionar a los individuos, sin tener seguridad realmente en la eficiencia productiva que mostrarán en su futura vida profesional. La premisa de la teoría de señalización que toma Cahuc de Spence (1973) es que las personas que rinden más en la vida activa son también las que rinden mejor mientras estudian. Si la eficiencia productiva no es observable por los posibles empleadores, entonces el éxito como estudiante simplemente sirve para señalar la presencia de tales características productivas. Desde este punto de vista, una persona persigue la educación para señalar su eficiencia, sin que sus estudios modifiquen realmente esta eficiencia. Si la educación sólo sirve para señalar las cualidades intrínsecas del individuo, entonces el verdadero

significado de la correlación positiva entre la duración de los estudios y los ingresos es simplemente que los individuos más eficientes tienen mayores ingresos.

El punto de vista de la teoría del capital humano está en total desacuerdo con el de la teoría de la señalización, ya que para esta última la prolongación de los estudios no contribuye en absoluto a aumentar la capacidad productiva de la persona; lo único que hace es enviar una señal a los empresarios. La teoría de la señalización también llega a conclusiones muy diferentes en cuanto a la eficiencia de las inversiones en educación. Mientras que la teoría del capital humano indica que las decisiones individuales con respecto a la educación son socialmente eficientes en condiciones de competencia perfecta, Spence (1973) muestra que los trabajadores tienden a sobre-educarse con respecto al estándar de eficiencia social, si la educación sirve para señalar sus capacidades productivas a los empresarios.

Los numerosos estudios empíricos dedicados al problema sugieren que la educación sí mejora la eficiencia individual.

3.1.4 Identificación de la relación causal entre educación y retribuciones

3.1.4.1 La teoría del capital humano: del modelo a las estimaciones

La principal predicción de la teoría del capital humano es que la educación es la fuente de una acumulación de competencias que permite aumentar los ingresos. La evaluación de este resultado se realiza mediante la estimación de funciones de ingresos, que relacionan los salarios con la inversión en educación. A partir del modelo teórico de capital humano se llega a una estimación de la tasa interna de retorno de la inversión educativa. El tiempo de permanencia en la escuela tiene un efecto positivo significativo en los ingresos. La tasa de rendimiento de un año adicional de escolarización es del 7%. La precisión de estas estimaciones aumenta notablemente al tener en cuenta la experiencia profesional y la acumulación de capital humano que tiene lugar después de dejar atrás la educación formal.

3.1.4.2 El problema de la selección

En efecto, el modelo de capital humano muestra que las capacidades individuales influyen tanto en los salarios como en la duración de los estudios. Además, según la teoría de la señalización, anteriormente comentada, la educación desempeña un papel de filtro, ya que sirve para seleccionar a aquellos trabajadores que son innatamente eficientes y para señalar características productivas de los trabajadores que los empresarios no pueden observar directamente. Así pues, la correlación entre la duración de la escolaridad y los ingresos se derivaría del hecho de que los individuos más eficientes tienen mayores ingresos y permanecen más tiempo en la escuela.

Tanto la teoría del capital humano como la teoría de la señalización predicen que los individuos más productivos tienen interés en estudiar durante más tiempo. Esto conlleva la posibilidad del llamado sesgo de capacidad, es decir, el rendimiento atribuido a la educación puede provenir de las capacidades

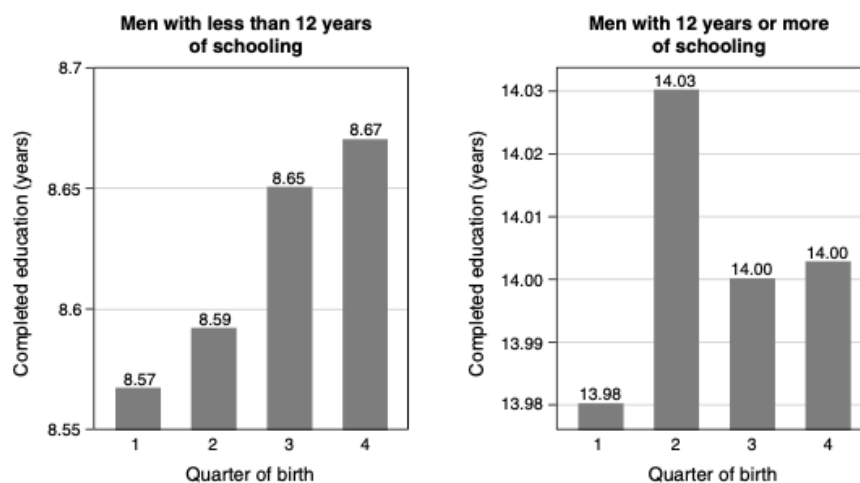
individuales. En estas circunstancias, los rendimientos de la educación estimados con el método de mínimos cuadrados ordinarios (MCO o OLS, Ordinary Least Squares) pueden estar sobreestimados.

La estimación se basa en la hipótesis de independencia entre el término de error y la duración de los estudios que es muy poco probable que se verifique, ya que, según los modelos de capital humano y de señales, las personas más capaces tienen un incentivo para cursar estudios más largos.

Este problema de selección, que tiene un carácter clásico en la econometría, ha sido abordado en numerosas contribuciones destinadas a evaluar el impacto causal de la educación sobre los ingresos.

A continuación, se presenta una estimación basada en el método común de variables instrumentales. En él se observa que los individuos nacidos a principios del año natural tienen una duración de la escolarización más corta que los nacidos más tarde. Este efecto se debe a que, en Estados Unidos, los distritos escolares suelen exigir que el alumno haya cumplido los 6 años el 1 de enero del año en que entra en la escuela. Además, las leyes de escolarización obligan a los alumnos a permanecer en la escuela hasta que cumplan 16 o 17 años. Por lo tanto, dos personas nacidas en el mismo año empiezan la escuela en la misma fecha, pero la que ha nacido antes está autorizada a dejar la escuela antes que la otra. Si suponemos que la fecha de nacimiento es independiente de los factores que influyen en las capacidades y las preferencias, este fenómeno puede suponer una variación exógena de la duración de la escolarización, que puede utilizarse como variable instrumental.

Figura 3. Trimestre de nacimiento y años de educación completa para los hombres nacidos en 1930-1939



Fuente: Angrist and Krueger (1991) conjunto de datos.

Este gráfico muestra que los individuos nacidos a principios de año y que estudian durante períodos más cortos, menos de 12 años, abandonan la escuela antes que los individuos nacidos a finales de año. En cambio, esta relación no se observa para los individuos que estudian más de 12 años. Esto sugiere que la

relación entre el trimestre de nacimiento y la duración de la escolarización está efectivamente inducida por la normativa que regula la asistencia obligatoria a la escuela y no por otros factores.

En promedio, si la fecha de nacimiento de un individuo cae en el último trimestre y no en los tres primeros, su probabilidad de completar 12 años de estudio aumenta. Las personas nacidas después del mes de septiembre deciden continuar sus estudios hasta el final del duodécimo año, correspondiente al final del bachillerato, sólo si el coste de hacerlo es suficientemente bajo. Las personas para las que este coste es demasiado elevado terminan sus estudios cuando alcanzan la edad legal, es decir, antes de terminar el 11º año. Los individuos cuyo comportamiento se ve modificado por el instrumento son los que prolongan sus estudios hasta el 12º año si han nacido después del mes de septiembre, y que los habrían interrumpido antes del 12º año si hubieran nacido antes del mes de septiembre.

Así, el hecho de nacer en el primer trimestre en lugar de en los tres últimos puede reducir la duración de los estudios para ciertas personas, o no tener ningún impacto, pero no debe aumentarla para nadie.

Otro método utilizado para evaluar los rendimientos de la educación consiste en utilizar datos sobre individuos cuyas capacidades son lo más parecidas posible. Desde esta perspectiva, varias contribuciones estiman los rendimientos de la educación para los hermanos, e incluso algunos estudios han utilizado poblaciones formadas por gemelos homocigóticos. Los gemelos perfectamente idénticos deberían, por definición, estudiar durante el mismo tiempo. Para que sus duraciones de estudio varíen, debe haber eventos que diferencien a estos gemelos de alguna manera. Debemos suponer que las razones por las que los gemelos no estudian durante el mismo tiempo no están correlacionadas con sus ganancias futuras.

Esta hipótesis es bastante dudosa, por lo que es poco probable que la utilización de datos relativos a los gemelos conduzca a estimaciones sólidas del impacto causal de la educación sobre los ingresos. Si se acepta la premisa de que los gemelos homocigóticos tienen capacidades idénticas y que las diferencias en su duración de la escolaridad se deben a eventos aleatorios que no modifican sus capacidades, estos resultados muestran que los sesgos de capacidad y selección tienen poco peso.

Cahuc muestra los resultados de Oreopoulos y Salvanes (2011) que utilizaron registros administrativos noruegos que proporcionan información sobre las trayectorias educativas y profesionales de todas las personas nacidas en ese país desde 1920. A partir de estos registros, construyeron una muestra muy amplia compuesta exclusivamente por hermanos y gemelos, lo que les permitió examinar las correlaciones entre la duración de los estudios y los ingresos futuros. En igualdad de condiciones, se constata que los hermanos (o gemelos, respectivamente) con un año más de estudios tienen, por término medio, un 5,2% (o un 4,8%, respectivamente) más de ingresos anuales que sus hermanos (o gemelos) con menos estudios.

3.1.5 Los rendimientos de la educación

Para conocer el rendimiento de la educación, Cahuc se basa en el método de Mincer que utiliza una simple relación entre los logaritmos de los salarios y la duración de los estudios para estimar la tasa de los rendimientos internos de la educación.

3.1.5.1 Rendimientos privados de la educación

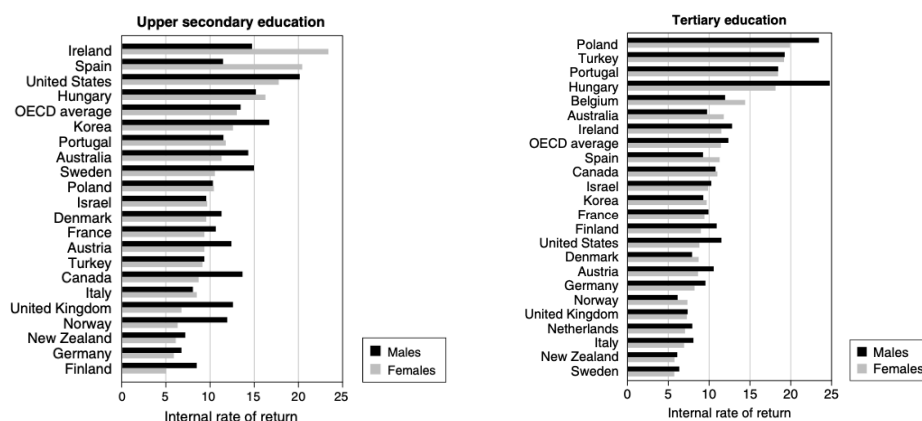
El modelo de Mincer se basa en un conjunto de hipótesis muy restrictivas, que lleva a estimar los rendimientos de la educación sobre la base de una regresión de los ingresos logarítmicos sobre los años de escolaridad, a los que a veces se añade la antigüedad.

Las hipótesis necesarias para derivar la ecuación de Mincer, son las siguientes. (1) La tasa de rendimiento de un año de escolaridad añadido es independiente de la duración de los estudios; (2) el coste de un año de escolaridad añadido es proporcional al salario; (3) la duración de la carrera es suficientemente larga (lo que corresponde a la hipótesis $T \rightarrow \infty$); (4) la duración de la carrera es independiente de la duración de la escolaridad. Además, cuando se tiene en cuenta la experiencia, la ecuación de ganancias de Mincer se basa en la hipótesis de que las funciones de ganancias son separables multiplicativamente en experiencia y escolaridad.

Es posible calcular los rendimientos de la educación en un marco de hipótesis menos restrictivo. En general, estos resultados indican que los supuestos de forma funcional que implican las estimaciones de las tasas de retorno basadas en Mincer conducen a sesgos muy grandes. Conduce a rendimientos notablemente superiores para los años de educación que cumplen los requisitos para obtener un diploma. Los resultados obtenidos indican que los impuestos y las tasas de matrícula reducen sustancialmente los rendimientos de la educación para los titulados universitarios, pero tienen un impacto menor en los rendimientos de la educación para los individuos con bajos niveles de escolaridad. Este resultado es coherente con la progresividad de los impuestos y las tasas de matrícula, que aumentan con el nivel de educación.

La magnitud de estos rendimientos estimados de la educación al final del duodécimo año plantea cuestiones importantes, en la medida en que más del 15% de las nuevas cohortes de jóvenes estadounidenses no obtienen un título de enseñanza secundaria, a pesar de su elevado rendimiento estimado. La estimación de un rendimiento medio de la educación oculta probablemente una fuerte heterogeneidad, cuya naturaleza exacta está aún por especificar.

Figura 4. Tasas interna de retorno



Fuente: P. Cahuc, 2004.

Estos gráficos muestran las estimaciones de las tasas internas de retorno privado de la educación secundaria y superior en 2005-2008 en 21 países de la OCDE (OCDE, 2012). Las tasas internas de retorno de la educación se calculan como los beneficios netos de obtener más educación. Estas cifras dan una idea de los incentivos para educarse a diferentes niveles. Los costes directos de la educación se basan en el gasto privado por año y la duración de la educación en cada país. Los ingresos medios se ajustan para tener en cuenta los impuestos y las transferencias sociales, así como el riesgo de desempleo para cada nivel educativo. Sin embargo, en los cálculos no se tienen en cuenta las becas y los préstamos subvencionados que a veces se reciben al principio de la educación, ni las diferencias de las pensiones después de la jubilación.

Por término medio, en la OCDE la tasa de retorno de la educación secundaria es bastante considerable, un 13%, un poco más alta que la de la educación universitaria (12%). Por supuesto, a medida que aumenta el desempleo, disminuye el coste de oportunidad de seguir formándose, mientras que aumentan las ganancias relativas de la educación universitaria. Las diferencias entre países pueden derivarse de una discrepancia entre la demanda y la oferta de mano de obra cualificada. Por ejemplo, en los países en los que hay una fuerte demanda de mano de obra altamente cualificada, pero todavía hay una gran parte de la población que carece de educación universitaria, esta última puede generar importantes beneficios. Por supuesto, esta situación tenderá a invertirse a medida que aumente progresivamente la proporción de la población con cualificaciones. Otro factor es la dispersión salarial general, una estructura salarial comprimida, típica de los países nórdicos, suele generar menores rendimientos de la educación superior. Esta es una de las razones por las que la educación superior es básicamente gratuita en estos países, con muchas subvenciones y becas para los estudiantes. Por el contrario, en los países con una desigualdad general de ingresos sustancialmente mayor, como Estados Unidos, las tasas de rendimiento tenderán a ser más altas.

3.1.5.2 Rendimientos privados de la educación no pecuniarios

Otros beneficios repercutidos de los años de estudio son influidos por la mejora en la toma de decisiones en materia de salud, elección de pareja y escolarización de los hijos. También puede modificar los estilos de consumo y el interés por el trabajo.

3.1.5.3 Rendimientos sociales de la educación

Los estudios empíricos dedicados a los rendimientos sociales de la educación tienden a mostrar que la educación ejerce efectivamente externalidades positivas y que los rendimientos sociales de la educación son superiores a los privados. En esencia, la educación mejora las capacidades de socialización de los individuos. La educación enseña a comunicarse con los demás y a comprenderlos, a aprender de ellos y a instruirlos. De ahí que la educación reduzca la delincuencia, mejore la participación en actividades cívicas como el voto y la participación en asociaciones, y promueva la difusión del conocimiento.

En general, los trabajos empíricos sugieren que los rendimientos sociales de la educación superan a los privados. Esta observación justifica hasta cierto punto el papel preponderante del Estado en el gasto en educación, un problema que debe investigarse más a fondo.

3.1.5.4 ¿Qué realmente es importante en la educación?

La duración de los estudios es uno de los ingredientes que permiten a los individuos acumular capital humano. Pero no es el único ingrediente y quizás ni siquiera sea el más importante. A continuación, Cahuc repasa los factores, además de la duración de los estudios, que más influyen en la acumulación de capital humano.

En primer lugar, los estudios empíricos suelen constatar que las personas que obtienen las mejores puntuaciones en las pruebas que miden los conocimientos obtienen mayores ingresos en el mercado laboral. Estas habilidades se miden con mayor precisión mediante pruebas que evalúan la capacidad cognitiva, como la prueba de cociente intelectual (CI), y los conocimientos generales, o mediante pruebas que evalúan las capacidades no cognitivas, como la perseverancia, la fiabilidad y la constancia. Estos resultados sugieren que es importante tener en cuenta la calidad de la escolarización, y no solo su duración, a la hora de evaluar los rendimientos de la educación.

Además, algunos estudios también constatan que la relación profesor/alumno, el gasto por alumno y los salarios de los profesores parecen tener un impacto positivo en los ingresos obtenidos por los estudiantes cuando abandonan la escuela. Estiman que reducir la ratio alumnos/profesor en 10 aumenta la tasa de rendimiento de la educación en unos 9 puntos porcentuales. Muestran que una reducción del tamaño de las clases durante la escuela primaria aumenta el nivel de éxito en las pruebas de rendimiento y la probabilidad de permanecer más tiempo en la escuela

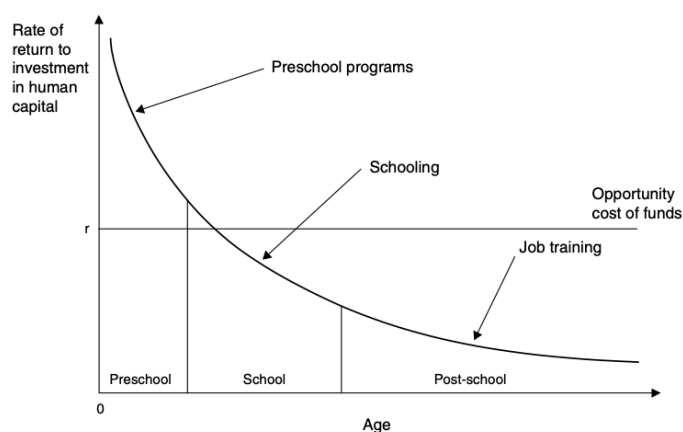
La adquisición de conocimientos, cuando es satisfactoria, se suele recompensar con un título capaz de influir en los beneficios derivados de un año más de estudio. En Francia, los trabajos de Goux y Maurin demuestran que los años de estudio no reconocidos por un título conllevan variaciones significativas en la remuneración (+3,2% al año), pero que, sin embargo, tienen un impacto de dos a tres veces más débil que los años que sí son reconocidos. Un título de ingeniero de una "grande école" (una institución postsecundaria de élite) da lugar a salarios un 25% superiores a los del título otorgado al finalizar el "deuxième cycle" en la universidad (el equivalente a un máster), aunque los períodos de estudio son de duración comparable.

La calidad del profesorado parece desempeñar un papel fundamental en el rendimiento de los alumnos. Se estima que, si pasa de un profesor situado en el cuartil inferior a otro situado en el cuartil superior, un alumno del percentil 50 (la mediana) en la distribución de los resultados obtenidos en matemáticas se encontrará al año siguiente en el percentil 58. Hanushek y Rivkin señalan que esta mejora de los resultados es mayor que la que se obtiene con las estimaciones de una reducción de 10 alumnos, en el tamaño de las clases.

También afecta a factores no cognitivos como la motivación, la personalidad, las preferencias temporales, la capacidad de cooperación, la concienciación, la extraversión, la estabilidad emocional y la sociabilidad. La consciencia del carácter es un predictor de las notas que obtendrá un estudiante en la universidad, tanto como el éxito en las pruebas de admisión a la universidad (Cunha et al., 2006). Se demuestra que los niveles altos de capacidad cognitiva y no cognitiva se asocian con menores tasas de abandono de la escuela secundaria.

En igualdad de condiciones, la tasa de retorno de un euro invertido en la educación de un joven supera la tasa de retorno de un euro invertido en la educación de una persona mayor. Para un coste de oportunidad dado (y constante) de las sumas invertidas igual a r , la curva óptima de inversión en educación dicta la mayor inversión cuando un individuo es joven, o incluso muy joven, ya que la tasa de rendimiento es visiblemente más alta para los programas preescolares.

Figura 5. Tasa de rendimiento de la inversión en capital humano por edad



Fuente: Cunha et al. (2006).

Hay muchos estudios que demuestran que la intervención dirigida desde una edad temprana tiene efectos fuertes y duraderos. Un análisis coste-beneficio de estos experimentos, con los beneficios descontados al 3%. Muestra que el Programa Preescolar Perry permitió a los padres reducir su gasto en atención infantil. Cada dólar invertido en el programa genera un rendimiento colectivo total de 9,11 dólares. Así que, además de su impacto positivo en el bienestar de los beneficiarios y la reducción de la desigualdad social, el dinero gastado en el Programa de Preescolar de Alta Calidad supone una contribución positiva sustancial al presupuesto del Estado.

3.1.6 Resumen y conclusiones

En conclusión, la teoría del capital humano justifica las opciones educativas asumiendo que la educación favorece la acumulación de competencias y aumenta los ingresos salariales. Este modelo predice que los individuos tienen interés, una vez completada su escolaridad, en reducir gradualmente el tiempo que dedican a la formación a lo largo del ciclo vital. El perfil de los salarios debería ser, pues, cóncavo con respecto a la edad. Si las características productivas de los individuos son inobservables, la educación puede considerarse como una actividad de señalización, que permite a los trabajadores más productivos darse a conocer a las empresas. La actividad de señalización puede conducir a un exceso de educación, que puede reducirse mediante subvenciones cruzadas destinadas a limitar el gasto en educación.

La estimación de los rendimientos de la educación debe hacer frente a la existencia de un sesgo de selección. Los estudios empíricos tienen muchas dificultades para detectar cualquier influencia sistemática del gasto en educación en el rendimiento de los estudiantes. La estimación de las funciones de ingresos que vinculan los ingresos con la duración de la escolaridad y la experiencia profesional permite evaluar el rendimiento de un año de educación adicional. En general, la investigación en este campo encuentra que este rendimiento se sitúa por término medio en la franja del 6-15%. A este respecto, también hay que señalar que los rendimientos de la educación son heterogéneos entre los individuos y entre los años de educación.

Los estudios empíricos disponibles indican que la educación mejora la implicación social, reduce la criminalidad, mejora la movilidad laboral y favorece la transmisión de conocimientos. Esto implica que los rendimientos sociales de la educación son mayores que los rendimientos privados de la misma. El tamaño de la clase y la calidad del profesorado influyen significativamente en el rendimiento de los alumnos y en su probabilidad de permanecer más tiempo en la escuela. Los rendimientos privados y sociales de la inversión en educación son mayores cuando se invierte en los jóvenes y en los que proceden de entornos desfavorecidos.

4 METODOLOGÍA ESTADÍSTICA. TABLAS DE CONTINGENCIA Y MAPAS DE CALOR

El presente Trabajo Fin de Máster realiza un análisis de la movilidad existente en el mercado laboral español utilizando para ello el análisis de correspondencias basado en tablas de contingencia – véanse en esta línea los trabajos de Álvarez de Toledo *et al.* (2013, 2014, 2017, 2018, 2020)–. Las Tablas de Contingencia (TC) se suelen utilizar para realizar una clasificación (tabulación) de forma cruzada de variables nominales o categóricas. Sea I el número de categorías de una variable X y J el número de categorías de una variable Y . Si las categorías de X se ordenan en filas y las categorías de Y se ordenan en columnas, obtenemos una tabla o matriz rectangular que tiene I filas y J columnas relacionadas entre sí. Esta tabla, que llamaremos TC_{IJ} , muestra el número de veces n_{ij} que se observa cada celda ij o combinación de categorías de X e Y , y se denomina tabla de contingencia o matriz de frecuencias. Con las frecuencias observadas n_{ij} , sus distribuciones de frecuencias marginales (totales de fila y columna), y el tamaño de muestra n , es posible cuantificar las similitudes que existen entre las categorías de una variable y el grado de asociación que existe entre las categorías de dos o más variables; estos dos conceptos, similitud de dos categorías-fila (o dos categorías-columna) y asociación entre categoría-fila y una categoría-columna son importantes en nuestro análisis categórico multivariante. La Tabla 1 muestra una tabla de contingencia de dimensión $I \times J$.

Tabla 1. *Tabla de contingencia*

		Categorías de Y					Total	
		1	2	...	j	...		J
Categorías de X	1	n_{11}	n_{12}	...	n_{1j}	...	n_{1J}	n_{1+}
	2	n_{21}	n_{22}	...	n_{2j}	...	n_{2J}	n_{2+}

	i	n_{i1}	n_{i2}	...	n_{ij}	...	n_{iJ}	n_{i+}

	I	n_{I1}	n_{I2}	...	n_{Ij}	...	n_{IJ}	n_{I+}
Total		n_{+1}	n_{+2}	...	n_{+j}	...	n_{+J}	n

Fuente: Álvarez de Toledo et al. (2018).

La noción de *propensión al emparejamiento* (o *factor de asociación*) entre cada categoría-fila y cada categoría-columna en la TC es importante en el presente estudio. Puede a su vez, asimilarse con la noción de alejamiento de la independencia (*departure from independence*) utilizada en el contexto metodológico de las TC, argumento que tiene su base en la diferencia existente entre las frecuencias analizadas de las celdas y las frecuencias esperadas bajo la hipótesis de independencia (o asignación aleatoria). Aunque el enfoque habitual en las TC está diseñado desde una perspectiva carácter global (se suele medir la inercia de la tabla completa, se llevan a cabo contrastes de independencia sobre la tabla, etc.), en este estudio se va a desarrollar un análisis más pormenorizado que consiste en observar la TC celda a celda, pretendiendo realizar una identificación de cuáles son las categorías-fila con mayor tendencia y probabilidad a asociarse con una determinada categoría-columna, y a la inversa; así se podrá observar *quién tiende a emparejarse con quién* dentro de la tabla de contingencia.

La TC desarrolla un conteo de observaciones en una clasificación cruzada de variables categóricas. A partir de la citada tabla podemos obtener una tabla de asociación (TA) $I \times J$ que permita mostrar los factores de asociación entre las diferentes categorías de las variables cruzadas. El factor de asociación (a_{ij}) entre la fila y la columna de la celda ij de la TC se obtiene como el cociente entre la probabilidad de la celda obtenida a partir de las frecuencias observadas p_{ij} y la probabilidad de la celda en el caso de que el reparto de sucesos entre filas y columnas fuera el resultado de un proceso aleatorio ($p_{i+}p_{+j}$):

$$a_{ij} = \frac{\text{Probabilidad observada de la celda } (i, j)}{\text{Probabilidad aleatoria de la celda } (i, j)} = \frac{p_{ij}}{p_{i+}p_{+j}} = \frac{n_{ij}/n}{\frac{n_{i+}}{n} \frac{n_{+j}}{n}} = \frac{n \cdot n_{ij}}{n_{i+}n_{+j}} \quad (4.1)$$

Los valores a_{ij} superiores a 1 significan que el grado de asociación entre la fila y la columna correspondiente a la celda analizada es mayor que en el caso de la asignación aleatoria, y viceversa.

Tanto en la TC como en la TA, cuando se da la circunstancia de que existen multitud de categorías en filas y columnas, el gran número ($I \times J$) de celdas que siguen un orden aleatorio, provoca que exista una dificultad para obtener como resultado una imagen genérica del objeto al que está referida la tabla; cuesta, por tanto, ver *quién se asocia con quién* si se parte de la base de seleccionar un número muy grande de celdas. La metodología de cluster o conglomerados consigue tratar este problema ordenando y agrupando las diferentes categorías (tanto de las filas como de las columnas) en un número menor de grupos o clusters de categorías y las diferentes celdas $I \times J$ en un número menor de biclusters, lo que permite obtener una visión más compacta y sintética de la TC y de la TA. En este trabajo no vamos a agrupar las frecuencias de categorías agrupadas, pero sí vamos a emplear la base de los clusters de filas y columnas para ordenar de forma coherente la TC y la TA.

La metodología de cluster tiene su base en la medida de similitud o proximidad entre los objetos (categorías) que se agrupan. Es preciso destacar que existen varias formas distintas de medir la semejanza o proximidad entre dos categorías de una variable; dado el concepto de asociación entre filas y columnas

(de la TC) que explora este trabajo, entenderemos que dos categorías-fila son más parecidas entre sí cuanto más se asemejen en la forma en que se asocien con las diferentes categorías-columna de la tabla, pudiéndose realizar el mismo razonamiento en la dirección contraria, desde las columnas hacia las filas. Consecuentemente, se medirá inicialmente la similitud entre cada par de filas (i_A e i_B) de la TC como la superposición o porcentaje de coincidencia de sus distribuciones (o perfiles) porcentuales respecto a las columnas de la tabla.

$$\widetilde{sm}_{i_A-i_B} = \sum_j \min \left(\frac{\tilde{p}_{i_Aj}}{\tilde{p}_{i_A+}}, \frac{\tilde{p}_{i_Bj}}{\tilde{p}_{i_B+}} \right) \quad (4.2)$$

El valor de la similitud se mueve entre 1 (si los perfiles porcentuales de las dos filas comparadas son idénticos) y cero (si la intersección entre ambas filas es nula). Podemos medir la similitud entre cada par de columnas (j_A y j_B) de la TC de forma análoga. Obsérvese que tendremos que transformar la TC en una tabla de perfiles-fila (*row profiles*) cuando queramos calcular las similitudes entre categorías-fila, y en una tabla de perfiles-columna (*column profiles*) cuando queramos calcular las similitudes entre categorías-columna.

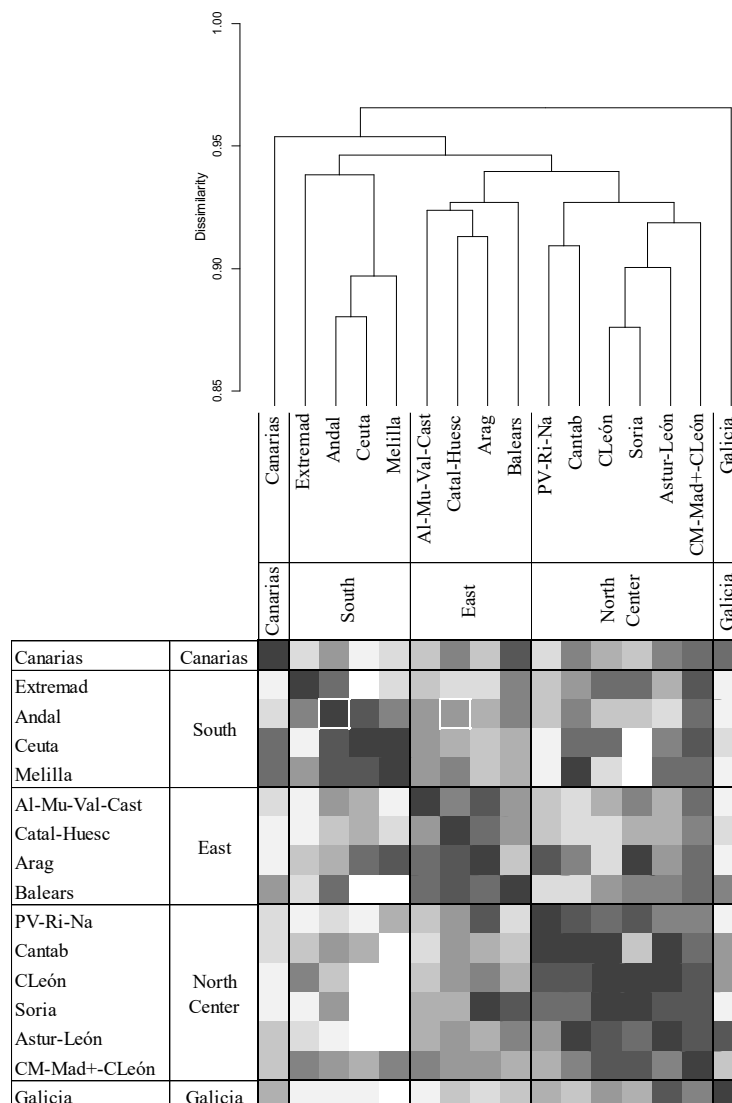
Obtenida la medida de semejanza entre categorías de una variable, se podrían utilizar diferentes métodos de agrupación de filas y de columnas. Por ejemplo, el método de cluster jerárquico, que nosotros aplicaremos en este trabajo, comienza uniendo las dos categorías (ya sean filas o columnas) de mayor similitud en un primer grupo o cluster de categorías, y se realiza dicho proceso de manera repetida en el tiempo de forma sucesiva con los nuevos grupos de categorías que se vayan generando; de esta forma, las categorías de partida se unen gradualmente para formar clusters de categorías cada vez mayores. Al finalizar dicho proceso, se obtiene un solo cluster que contiene a todas las categorías de partida. El proceso de agrupamientos se puede representar gráficamente mediante un dendrograma, que no es más que un diagrama en forma de árbol que muestra la generación sucesiva de clusters de categorías.

El agrupamiento obtenido de filas y de columnas en el proceso de cluster (esto es, la base de los respectivos dendrogramas) se pueda extrapolar al orden de filas y columnas tanto de la TC como de la TA. Se debe tener en cuenta que la base de un dendrograma permite observar todos los elementos de partida del cluster ordenados de acuerdo con su semejanza; o lo que es lo mismo, dos elementos contiguos en la base del dendrograma que se van a parecer más entre sí que dos elementos que no sean contiguos entre sí. Como se podrá observar en el capítulo de resultados, esto permite garantizar que el dendrograma se vaya generando visualmente de manera tal que no se desarrollen cruces entre las diferentes uniones que se van a ir generando, facilitando así su comprensión visual.

Si las filas y las columnas de la TA se ordenan según la base de sus respectivos dendrogramas, y los factores de asociación se muestran mediante una escala de color (de manera que colores más intensos se corresponden con mayores factores de asociación), se logra obtener una representación visual del

fenómeno analizado que recibe el nombre de “mapa de calor” (*heatmap*) –véase la Figura 6 a modo de ejemplo–.

Figura 6. Ejemplo de obtención de un bichuster con estructura diagonal



Fuente: Álvarez de Toledo et al. (2018).

La TC es un instrumento flexible que permite comprender la generación de empleo en el mercado laboral desde diferentes puntos de vista. Concretamente, en este estudio, se pretende obtener TC y TA o *heatmaps* sobre las secuencias de sectores de actividades, grupos de ocupación y provincias que siguen los trabajadores cuando van cambiando de empleo dentro del año. Se centrará en aquellos cambios de empleo que efectivamente implican un cambio en algunas de estas variables, cambios que se analizarán por separado, es decir, se van a generar un mapa de calor que relaciona las actividades del empleo anterior y del actual, otro que relaciona las ocupaciones del empleo anterior y del actual, y otro que hace lo propio

con la provincia del puesto de trabajo. De esta forma, se estará en condiciones de detectar las movilidades laborales de mayor asociación en el mercado de trabajo español durante el año 2020 –dada la particularidad de este año, primer año de la pandemia COVID-19, se comparará con el año 2019–.

A modo de ejemplo, se va a representar un mapa de calor de un bicluster de estructura diagonal (Tabla 2) en la que las filas recogen los sectores de origen donde trabajan los individuos y en las columnas las actividades económicas de destino. En cada celda viene representado el valor del factor de asociación entre las dos categorías que se cruzan en dicha fila-origen con la columna-destino de la celda. Si este valor es mayor que la unidad, la propensión de emparejamiento entre ese origen y ese destino es mayor a que si fuera un reparto de forma aleatoria. Sin embargo, si es menor que uno, se relacionan con más frecuencia de forma aleatoria que en el suceso a analizar. Estas celdas tienen un color más oscuro cuanto mayor sea el factor de asociación.

Tabla 2. Ejemplo de mapa de calor y factor de asociación entre distintos sectores

	Agricultura	IndManufacturera	Suministros	ServFros&Emp	OrganExtraterr	Educación	Sanidad	ComHostTte&Comunic	IndExtractivas	Construcción	AAPP	OtrosServicios
Agricultura		0,98	1,18	0,20	0,17	0,07	0,25	0,91	1,05	2,85	2,95	1,05
IndManufacturera	1,01		0,80	0,62	0,24	0,36	0,36	1,33	3,63	1,76	0,51	1,17
Suministros	1,31	0,87		0,35		0,41	0,44	1,04		1,53	0,77	1,13
ServFros&Emp	0,23	0,50	0,29			1,33	1,08	1,02		0,49	0,46	1,59
OrganExtraterr	0,31	0,37		0,64		0,68	0,76	1,16		0,21	2,62	1,32
Educación	0,09	0,26	0,34	0,95			2,24	0,92		0,13	1,67	1,64
Sanidad	0,25	0,30	0,32	0,84	0,92	3,06		1,04		0,20	2,03	1,18
ComHostTte&Comunic	0,85	1,21	1,05	0,96	1,34	0,92	1,20		0,88	0,99	0,62	1,92
IndExtractivas	1,01	2,93	1,21	0,84		0,15		0,91		2,40	1,15	0,61
Construcción	2,77	1,66	1,82	0,53	0,52	0,17	0,21	1,02	3,55		0,80	0,84
AAPP	3,25	0,41	1,09	0,35	2,44	1,60	1,99	0,62	0,19	0,83		0,71
OtrosServicios	1,06	1,34	1,09	1,75	1,36	1,43	1,20	1,90	0,49	0,80	0,72	

Fuente: Elaboración propia a partir de los datos.

5 ORIGEN DE LOS DATOS. LA MUESTRA CONTINUA DE VIDAS LABORALES

En este apartado se llevará a cabo la descripción de los datos necesarios para el análisis de las movilidades entre actividades, ocupaciones y provincias. Esta información ha sido extraída a partir de la Muestra Continua de Vidas Laborales (MCVL), donde aparece de forma muy detallada todos los episodios laborales e información personal sobre cada individuo de la muestra. Se usarán para el estudio y trabajo los datos provenientes del año 2020 pero se tomarán algunos datos de referencia del 2019 sólo a título comparativo. Además, se explicará el proceso de selección de la MCVL, su estructura y su contenido. La MCVL que se utiliza concretamente es la versión con los datos fiscales de los individuos de la muestra (MCVL-CDF) que añade a la información de la Muestra Sin Datos Fiscales, los rendimientos brutos anuales (dinerarios y en especie) de cada individuo. Esta última variable es importante en nuestro estudio porque, de alguna forma, nos va a permitir aproximar lo rentable que son las diferentes movilidades ocupacionales y geográficas detectadas.

La Muestra Continua de Vidas Laborales es una recopilación de datos individuales y anónimos. Estos son obtenidos de las bases de datos de la Seguridad Social, del Padrón Continuo Municipal (INE) y del resumen anual de retenciones e ingresos a cuenta del IRPF (Modelo 190) de la Agencia Estatal de Administración Tributaria de España (AEAT). Esta estadística ofrece anualmente un conjunto organizado de microdatos que están referidos a algo más de un millón de personas; se trata de una muestra representativa de todas las personas que tuvieron relación con la Seguridad Social en un determinado año.

El conjunto de los datos es recogido de una muestra de personas escogidas de forma azarosa de los que han sido afiliados o pensionistas de la Seguridad Social durante el año natural, ya sea 2020 o 2019, en nuestro caso. Los datos de cada individuo son los correspondientes a toda su vida laboral, incluyendo el cobro de pensiones y prestaciones en el caso de que los hubiere. En concreto, para el año de análisis, el 2020, la población de referencia eran 31.424.464 personas y para la muestra participan 1.255.522 individuos, es decir el porcentaje que la muestra representa sobre dicha población alcanza el 4%.

Toda esta información confidencial está sólo a disposición de investigadores que lo solicitan, quedando bajo su responsabilidad todo el protocolo de protección de datos. Los datos que se conceden no tienen modificaciones y se entregan con la misma estructura respecto a los registros informáticos de la Seguridad Social.

La información que aporta la MCVL se estructura en seis tablas que corresponden a sendas áreas de información. Para cada tabla, la primera columna de cada fila se describe el código identificador de cada afiliado o pensionista al que pertenecen los datos que se muestran en cada columna de esa fila, el llamado IPF (Identificador de la Persona Física). En primer lugar, la tabla de las PERSONAS contiene información de índole personal para cada individuo de la muestra; variables como la fecha de nacimiento,

el género, la provincia de afiliación, el domicilio residencial habitual y nivel de estudios, entre otras. Por otro lado, la tabla de AFILIACIÓN aporta datos sobre los episodios de afiliación a la Seguridad Social del individuo a lo largo de toda su vida laboral. La tabla contiene tantas filas como episodios de afiliación haya tenido para cada IPF (Identificador de la Persona Física). En el caso de tener varios empleos a la vez, se recoge cada alta a la Seguridad Social con su fecha, aunque sea con distintos empleadores. Los datos requeridos para este proyecto que se obtienen a partir de esta tabla son el tipo de contrato de trabajo, las fechas tanto de alta como de baja en la afiliación, el domicilio de la actividad de la cuenta de cotización y la Actividad Económica de la cuenta de cotización (CNAE 2009) de cada episodio por individuo.

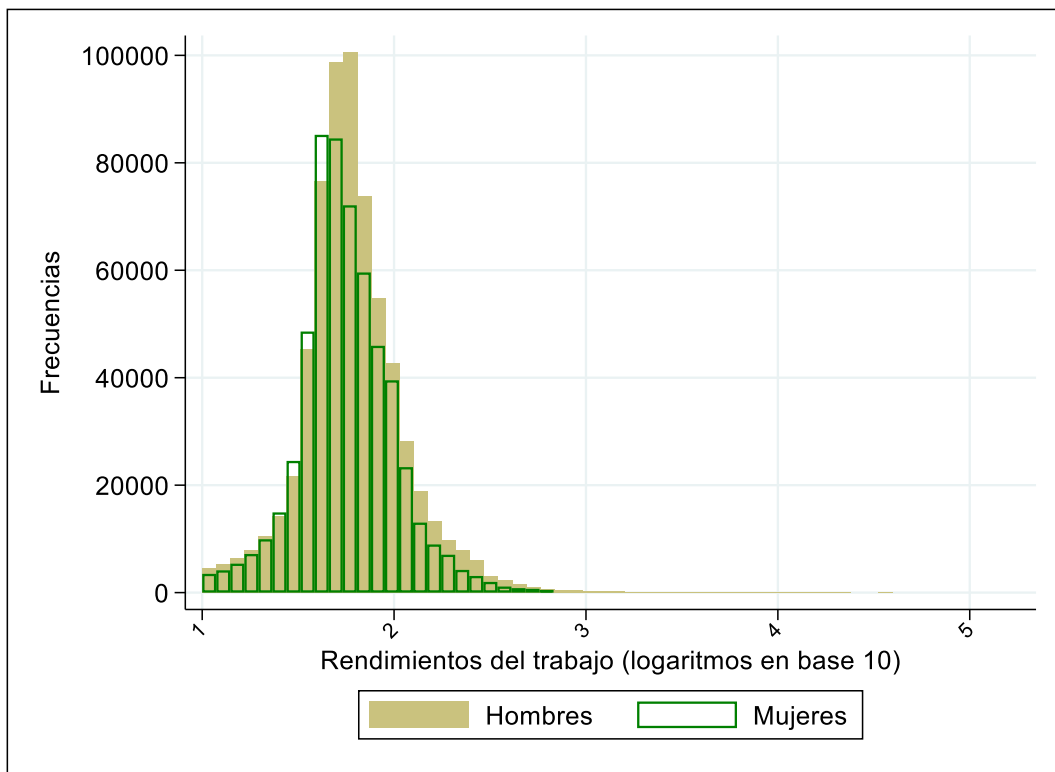
La tercera tabla que muestra la MCVL es sobre las BASES DE COTIZACIÓN, que no se usará para este trabajo. En esta tabla se registran los datos de IPF, año y Código de Cuenta de Cotización (CCC). El dato de la base de cotización mensual por contingencias comunes se refiere a la suma de todas las bases de cotización añadidas para un mismo IPF y CCC en cada mes.

Las siguientes dos tablas, sobre PENSIONES y CONVIVIENTES tampoco son utilizadas para la elaboración de este TFM. La tabla de pensiones muestra si el individuo tiene algún tipo de incapacidad, está jubilado o se encuentra en situación de viudedad u orfandad, y si por alguno de esos motivos recibe o lo ha recibido alguna retribución económica en forma de pensión. Por otro lado, la tabla de convivientes contiene tantas filas como registros del Padrón Municipal Continuo tenga cada persona física. Toda esta información no es necesaria para el análisis de las movilidades de nuestra muestra, por este motivo de estas tablas no se recogen datos para este trabajo.

Por último, la tabla 6 de la Muestra, incluye los DATOS FISCALES de los IPF perceptores de rentas. Esta información es de cada individuo y pagador, y muestra las retribuciones brutas producidas y las retenciones producidas por el IRPF en el año fiscal de referencia (2020 en nuestro caso). Esta última tabla sólo aparece en la versión de la MCVL con datos fiscales; los datos son obtenidos por la Agencia Estatal de la Administración Tributaria. La tabla del módulo fiscal muestra para cada individuo los rendimientos obtenidos (dinerarios o en especie) de cada pagador y las retenciones practicadas por cada pagador que haya tenido a lo largo del año fiscal. Para la investigación de este trabajo se considerará la variable clave de percepción que es una clave alfabética de un dígito la cuál indica el tipo de percepción ya sean rendimientos del trabajo por cuenta ajena, pensiones, prestaciones, etc. Estas claves se pueden consultar en las disposiciones que regulan el modelo 190 vigente para cada momento. Concretamente, la clave que consideramos para obtener los rendimientos del trabajo diarios es la clave A que corresponde a más del 90% de episodios fiscales. Nótese que para este análisis no se considerarán las claves B, C y D. Para relacionar estos datos fiscales con las anteriores tablas, en concreto con las tablas de datos personales y de datos de afiliación, se puede utilizar tanto el IPF como la letra del NIF del pagador y su identificador (llamado en la encuesta ‘identificador anonimado de la entidad pagadora’). Con esta información, en el trabajo se estudiarán si las movilidades geográficas, ocupacionales o por actividades de los trabajadores son más o menos beneficiosas en términos económicos.

Como en el presente trabajo se hablará de esta variable retributiva, se ha realizado un histograma con los datos medios que los individuos de retribuciones diarias, obtenidos de la MCVL. Estas retribuciones diarias en el eje de abscisas del gráfico es el valor de la variable de los rendimientos diarios de trabajo en logaritmo en base 10. En la Figura 7 se diferencian los datos entre el género masculino y femenino. Se puede apreciar que el histograma de las mujeres está levemente peor situado que el de los hombres ya que se mueve hacia cifras menores diarias, es decir, más cercano del origen que el histograma de los hombres. La parte derecha de la que podría ser una campana de Gauss es donde se encuentran los mayores rendimientos, más de 200 €/día. Como se ve reflejado en el histograma, la frecuencia con la que esos salarios medios diarios entre los hombres son considerablemente mayores a las colocaciones entre mujeres que lleven a esas retribuciones económicas.

Figura 7. *Histograma de rendimientos de trabajo*



Fuente: Elaboración propia a partir de los datos.

6 APLICACIÓN

En este capítulo se desarrollará el análisis empírico objetivo de este trabajo, la observación de las secuencias de empleo durante el año 2020 de los individuos de la muestra. En este año 2020, el mundo entero sufrió un parón económico (y vital) debido a la COVID-19; la pandemia ha afectado a las trayectorias de empleo de la población activa en el mercado de trabajo, por lo que, en nuestro estudio analizaremos también los datos de 2019 para tener una referencia pre-COVID. Este análisis se llevará a cabo mediante la elaboración de mapas de calor (basados en tablas de contingencia; ver sesión de metodología) que representan el factor de asociación entre cada fila y cada columna de la tabla o mapa analizada. Las filas vienen dadas por las categorías de una determinada variable descriptiva del puesto ocupado por el trabajador (como, por ejemplo, el sector de actividad o el grupo de ocupación del puesto), mientras que las columnas contienen las categorías de dicha variable descriptiva, pero en el siguiente episodio de empleo del individuo; obsérvese por tanto que analizamos trabajadores que necesariamente han cambiado de empleo en 2020 (o 2019). Las secuencias laborales serán observadas gracias a cambios de empleo de los individuos de la muestra dentro del año 2020 (el año 2019 también se valora, pero solo se usa a modo de referencia). La muestra empleada es una muestra de entrada (*inflow sample*), o lo que es lo mismo, se consideran todas las colocaciones que se producen dentro del año. Analizando los cambios de contrato de los trabajadores, podremos valorar tres tipos de movilidades: de sector de actividad, de grupos de ocupación y de provincia; movilidades que serán estudiadas por separado (no se analizarán dobles o triples movilidades).

En el caso de una movilidad de sector de actividad, el trabajador cambia de actividad entre dos empleos consecutivos. Estas actividades vienen determinadas por la clasificación CNAE (Clasificación Nacional de Actividades Económicas, 2009). Cada empresa en el momento que se constituye como sociedad debe definir la actividad económica que va a desarrollar. La CNAE consta de un código numérico seguido de su definición; su función es la de clasificar a las empresas según sectores y grupos de actividades homogéneos. Las movilidades de grupos de ocupación implican un cambio de ocupación entre dos empleos consecutivos. Dicha variable de ocupación viene codificada de acuerdo con la CNO (Código Nacional de Ocupaciones, 2011). Este código clasifica el tipo de puesto que ejerce el individuo dentro de la compañía. Finalmente, las movilidades de provincia son producidas por aquellas personas de la muestra que por un motivo u otro se han trasladado de provincia al cambiar de trabajo.

Estos tres tipos de movilidades se consideran de forma separada en nuestro estudio, esto es, no analizaremos cambios simultáneos en dos o más variables (aunque se hayan producido). Asimismo, resulta importante aclarar que aquellos cambios de contrato que no implican cambio de la variable analizada (actividad, ocupación o provincia) no son considerados en nuestro análisis de movilidad laboral. Nuestro análisis permite observar cómo ha ido evolucionando la movilidad laboral del individuo a lo

largo del año 2020 –a efectos comparativos, se han replicado los cálculos para el año 2019, como se muestra en los Anexos de este Trabajo.

El análisis de la movilidad laboral lo vamos a llevar a cabo usando la herramienta del *heatmap* o mapa de calor, que muestra el factor de asociación entre cada categoría fila y cada categoría columna de dicha tabla o mapa –dichas categorías pueden ser los sectores de actividad, los grupos de ocupación o las provincias de dos contratos consecutivos–. Además, el orden de las filas y columnas de los *heatmaps* de este capítulo no es exógeno o arbitrario, sino que viene dado por el proceso de clusterización descrito en el capítulo metodológico. De acuerdo con dicha metodología, dos categorías de una variable cualitativa (por ejemplo, dos filas de la tabla de contingencia) se agrupan formando un cluster porque se parecen bastante en la forma en que se emparejan con las categorías de otra variable cualitativa (categorías que se encontrarían en las columnas de la tabla de contingencia). Asimismo, la distancia entre dos clusters de categorías ya formados se calcula por el sistema de la vinculación promedio (*average linkage*) entre ellos, esto es a través de la media de las distancias entre los elementos de un clúster y los elementos del otro. A la representación visual de un clúster jerárquico como el descrito (ya sea de las filas de la tabla de contingencia o de las columnas) se le denomina dendrograma.

Como hemos comentado anteriormente, las filas y las columnas de nuestros mapas de calor están ordenadas igual que la base del dendrograma que clusteriza las categorías representadas en dichas filas y columnas. Por construcción matemática, dos elementos contiguos (o vecinos) en la base de un dendrograma se parecen más entre sí que dos elementos lejanos, por lo que se acabarán uniendo antes en el proceso de formación de clusters de elementos o categorías. Además, como nuestras tablas de contingencia y sus mapas de asociaciones (*heatmaps*) tienen las mismas categorías en las filas y las columnas, ya que analizamos movimientos entre actividades, ocupaciones o provincias por separado, hemos usado una matriz de distancias entre categorías común a filas y a columnas con el fin de generar un único dendrograma. En esta matriz de distancias promedio, cada celda viene dada por la media de la distancia entre las dos categorías-fila y las dos categorías-columna correspondientes. De esta forma, conseguimos que la diagonal del mapa de calor esté formada por una categoría fila y su simétrica por el lado de las columnas. Como veremos, dichas celdas implicarían colocaciones donde no se observa movilidad del trabajador (por ejemplo, dos contratos de un mismo trabajador ambos en la agricultura), colocaciones que no son tenidas en cuenta en nuestro estudio de movilidad.

6.1 Secuencias de Actividades

Los trabajadores que han tenido algún cambio de contrato en 2020 acumulan un total de 794.911 contrataciones en la MCVL –incluyendo aquellas colocaciones que empezaron antes de 2020 pero que no habían concluido en 2020–. De todos estos contratos, 106.210 han implicado movilidad de sector de actividad económica. Comparando con el año anterior a la pandemia, se observa que se han producido menos movilizaciones de sector de actividad en 2020 (106.210 vs. 139.120); nótese que también ha habido

un descenso de más de 200.000 colocaciones entre 2019 y 2020 dentro de nuestra muestra (esta muestra viene dada por los trabajadores que han tenido más de una colocación durante el año analizado).

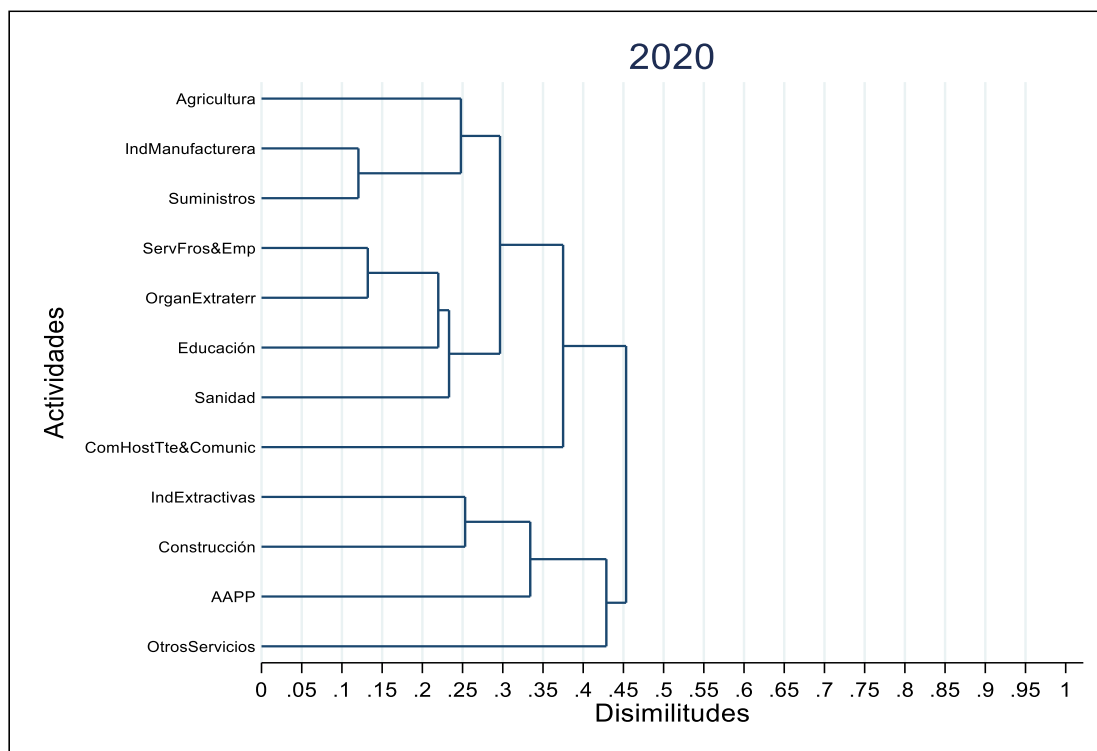
Tabla 3. *Tabla total colocaciones de secuencias de actividades*

PERIODO	CASO	COLOCACIONES
Año 2020	Individuos con más de una colocación en 2020	794.911
	Colocaciones de las que conocemos origen y destino	584.659
	Sin movilidad de sector de actividad	478.449
	Con movilidad de sector de actividad	106.210
Año 2019	Individuos con más de una colocación en 2019	1.006.961
	Colocaciones de las que conocemos origen y destino	782.361
	Sin movilidad de sector de actividad	643.241
	Con movilidad de sector de actividad	139.120

Fuente: Elaboración propia a partir de los datos.

Para una mejor visualización de nuestros resultados, hemos agrupado los 99 grupos de la CNAE 2009 a d dos dígitos en 12 grupos de sectores homogéneos, tal y como se puede ver en el Anexo 1 para generar el dendrograma de la Figura 8.

Figura 8. *Dendrograma de actividades*



Fuente: Elaboración propia a partir de los datos.

Analizando el dendograma, podemos identificar tres clusters y un sector poco próximo a los demás. En primer lugar, la pareja *agricultura, industria manufacturera* junto con *suministros* forman un primer cluster de actividades –esto significaría, aproximadamente, que estos sectores de actividad se parecen entre sí actuando en los dos sentidos: como sectores de origen (actividades del empleo anterior) y como sectores de destino (actividades del empleo actual)–. Seguido de *servicios financieros* junto con *organismos extraterritoriales, educación y sanidad*, que forman así un segundo cluster. El tercer grupo está compuesto por los tres sectores de *industrias extractivas con construcción y administración pública*. Y, por último, queda *comercio, hostelería, transportes y comunicaciones* como sector menos homogéneo al resto.

La Tabla 4 contiene el mapa de calor que refleja la propensión al emparejamiento o factor de asociación entre cada actividad-fila y cada actividad-columna. La diagonal de la tabla implica no movilidad entre actividades cuando se produce un nuevo contrato de un trabajador, y esas colocaciones (que son la mayoría) no las consideramos porque el trabajo implica movilidad –nótese que esas celdas no contienen factores de asociación–. Los mapas de calor propuestos en este trabajo se pueden ver como matrices origen-destino, ya sea, de provincias, ocupaciones o actividades. Vamos a optar por interpretar el *heatmap* desde el origen hacia el destino, es decir, interpretando cómo las filas se reparten en las diferentes columnas –aunque también la interpretación de columnas a filas podría tener su interés–.

Tabla 4. Mapa de calor de actividades económicas

	Agricultura	IndManufacturera	Suministros	ServFros&Emp	OrganExtraterr	Educación	Sanidad	ComHostTte&Comunic	IndExtractivas	Construcción	AAAPP	OtrosServicios
Agricultura		0,98	0,81	0,17	0,35	0,09	0,24	0,94	1,91	2,57	2,65	1,21
IndManufacturera	1,02		0,81	0,66		0,37	0,44	1,37	2,55	1,60	0,46	1,30
Suministros	1,01	0,83		0,31		0,48	0,47	1,15	3,03	1,81	0,78	1,11
ServFros&Emp	0,19	0,59	0,45		1,04	1,45	1,41	1,06		0,57	0,79	1,48
OrganExtraterr	0,17	0,26		2,17		1,45	1,17	1,59		0,28		1,36
Educación	0,12	0,33	0,37	1,34	2,46		2,15	1,04		0,13	1,68	1,62
Sanidad	0,24	0,34	0,50	1,30	1,32	3,01		1,01		0,22	1,99	1,26
ComHostTte&Comunic	0,89	1,19	1,18	0,93	0,91	0,97	1,23		0,71	1,02	0,62	1,78
IndExtractivas	1,74	2,98	1,21	0,60		0,16	0,16	0,93		1,53	0,56	0,54
Construcción	2,57	1,56	2,02	0,54	0,30	0,20	0,24	1,06	2,78		0,75	0,81
AAAPP	2,78	0,37	0,74	0,49	0,39	1,73	1,86	0,56	0,44	0,77		0,63
OtrosServicios	1,21	1,43	1,08	1,69	1,58	1,40	1,27	1,89	0,56	0,81	0,70	

Fuente: Elaboración propia a partir de los datos.

- Justificación de la formación de clusters de actividades

Como se ha comentado anteriormente, en el dendrograma, dos o más categorías (actividades en este caso) forman un clúster porque tienden a asociarse de forma similar con las actividades del otro lado de la tabla de contingencia. Por ejemplo, *servicios financieros y a empresas, organismos extraterritoriales, educación y sanidad* se encuentran cercanos en el dendrograma porque tienden a asociarse con las actividades del otro lado de la tabla de forma parecida, y eso sucede tanto cuando actúan como filas (actividades del empleo anterior) que se asocian con ciertas columnas como cuando lo hacen como columnas (actividades del empleo actual) que se asocian con ciertas filas; aunque nosotros atenderemos al emparejamiento de filas hacia columnas. Así, los cuatro sectores mencionados tienden a asociarse entre ellos y con sectores como el de *otros servicios*; es decir, que trabajadores que se colocan en el sector de actividad de *otros servicios* tienen una probabilidad elevada de proceder de las actividades de este cluster. Por ejemplo, el de *industria manufacturera y suministros*, que se parecen porque se asocian, fundamentalmente, con *industrias extractivas y construcción*

- Análisis descriptivo de las celdas con mayor factor de asociación (FA)

Observando cada celda de la Tabla 4, se identifican cuáles son las categorías-fila con mayor propensión a asociarse con una determinada categoría-columna, cuyas celdas son de un color azul más oscuro. Este es el caso, por ejemplo, de los factores de asociación 3,03 y 3,01. El primero correspondiente a los individuos que trabajando en la actividad económica de *suministros* se mueven hacia la actividad de *industrias extractivas*; sin embargo, a pesar de tener un factor de asociación elevado, la frecuencia de colocaciones de este caso es muy escasa, por lo que no es un caso particularmente relevante. Es por ello por lo que los criterios para describir celdas de interés del mapa de calor serán que las frecuencias de colocaciones sean mayores a 100 y que el factor de asociación sea mayor a 2. Existen siete celdas que cumplen los criterios anteriormente descritos. A continuación, las describimos en términos de una serie de variables de interés laboral.

El primer caso más relevante, correspondiente al $FA = 3,01$ en la Tabla 5, representa la propensión de asociarse entre los grupos de sanidad y educación, concretamente, muestra las movilidades de trabajadores en el sector sanitario que han cambiado su trabajo yéndose al sector de la educación. Esta fuerte asociación, que también se observa (aunque en menor medida) a la inversa ($FA = 2,15$), se observa también en el año 2019, por lo que no puede ser atribuible a la pandemia; más bien se debe a que son sectores afines que experimentan trasvases importantes de trabajadores en materia de dirección, gestión, conserjería, mantenimiento o incluso en materia de docencia e investigación. La movilidad de sanidad a educación se ha observado en 1444 casos. Nuestros microdatos permiten describir los contratos de origen y de destino de estos trabajadores. En concreto, las variables de interés van a ser la provincia y ocupación del puesto, y el nivel educativo, la renta bruta diaria, el género, la edad y la duración del empleo del trabajador. Dicha descripción se muestra en la Tabla 5.

Tabla 5. Análisis celdas educación y sanidad

ACTIVIDAD ORIGEN	ACTIVIDAD DESTINO	COLOCACIONES	FA	CARACTERÍSTICA	SUBPARTADO	TOTAL	PORCENTAJE
Sanidad	Educación	1444	3,01	PROVINCIA	Barcelona	330	22,85%
					Madrid	274	18,98%
					Valencia	99	6,86%
					Zaragoza	52	3,60%
					Sevilla	46	3,19%
				OCUPACIÓN	Ingenieros, licenciados, alta dirección	727	50,35%
					Ingenieros técnicos, ayudantes titulados	240	16,62%
					Oficiales administrativos	90	6,23%
					Ayudantes no titulados	84	5,82%
				N. ESTUDIO	Máster Universitario, Doctorado y postgrado	548	37,95%
					Licenciado o Grado Universitario	343	23,75%
					Diplomados en Escuelas Universitarias	157	10,87%
					Bachiller o más	100	6,93%
OTROS	RENTA AL DÍA (€/día)	81,7					
	GÉNERO (Hombre:0 - Mujer:1)	0,3					
	EDAD MEDIA (años)	40,5					
	DURACIÓN (días)	47,1					
Educación	Sanidad	972	2,15	PROVINCIA	Barcelona	248	25,51%
					Madrid	138	14,20%
					Valencia	77	7,92%
					Málaga	40	4,12%
					Zaragoza	30	3,09%
				OCUPACIÓN	Ingenieros técnicos, ayudantes titulados	244	25,10%
					Ingenieros, licenciados, alta dirección	171	17,59%
					Ayudantes no titulados	144	14,81%
					Auxiliares administrativos	117	12,04%
				N. ESTUDIO	Licenciado o Grado Universitario	281	28,91%
					Máster Universitario, Doctorado y postgrado	226	23,25%
					Diplomados en Escuelas Universitarias	96	9,88%
					Bachiller	85	8,74%
OTROS	RENTA AL DÍA (€/día)	86,1					
	GÉNERO (Hombre:0 - Mujer:1)	0,2					
	EDAD MEDIA (años)	36,1					
	DURACIÓN (días)	84,4					

Fuente: Elaboración propia a partir de los datos.

Dentro de estas nuevas colocaciones la mayoría se han producido en Barcelona y Madrid en un 22,85% y 18,98% de los casos respectivamente; después en menor proporción, en Valencia, Zaragoza y Sevilla. Con respecto a la ocupación de esta migración de personal del ámbito de la *sanidad* al de *educación* más de la mitad de los casos son licenciados de alta dirección o ingenieros técnicos ya que seguramente sean en su mayoría médicos o enfermeros. En relación con el nivel de estudios, existe cierta correspondencia con los grupos de ocupación, ya que observamos mayoritariamente titulados universitarios con o sin postgrado. Gracias a los datos reflejados en la MCVL se sabe cuál es la renta bruta diaria para cada individuo, esto es, el salario bruto diario ganado por el individuo en el año 2020 en todas sus colocaciones, no sólo las recogidas en esta celda del mapa de calor. Se observa que, en promedio, han declarado los individuos pertenecientes a esta movilidad unos 81,7€/día en todo el 2020 –el salario mínimo interprofesional fue de 36 € (con pagas extra prorrateadas al día) en 2020–. Es destacable que el género predominante entre las profesiones relacionadas con la enfermería sea el femenino con una edad media de 40,5 años. Por último, hay que señalar que, por lo general, estos individuos cambian a un contrato de 47,1 días de media de duración. Aunque son períodos de corta duración, los contratos de cambio en promedio son mayores a un mes y se puede decir que está muy bien remunerado, posiblemente debido a la alta cualificación de los trabajadores con este tipo de movilidad.

En cuanto al movimiento en la otra dirección, de educadores hacía sanitarios, tiene unas características muy parecidas en todos los aspectos, a pesar de producirse en menor frecuencia, 972 colocaciones durante el año 2020. En lo que concierne a en qué provincias se han producido estas movi­lidades son casi idénticas a las anteriores a excepción del cambio de las provincias andaluzas, de Málaga por Sevilla entre las más repetidas. Tanto la ocupación como el nivel de estudios es también parecido, son mayormente, titulados universitarios, pero de menor rango que en el caso anterior. A pesar de que el nivel educativo y los años de cada individuo de media bajen levemente, las condiciones de estos contratos suben considerablemente en su duración casi a los tres meses y la renta es algo mayor también, por lo que podríamos decir que, aunque a la vista está que es más complicado que los educadores pasen a ser sanitarios, ya que el número de cambios es menor, sí que es buscado porque estos individuos tienen una remuneración mayor. El perfil de estos individuos sigue siendo más pronunciado en mujeres que hombres con una edad media de 36 años.

Otra celda destacable del mapa de calor es la relación entre el sector de administraciones públicas y la agricultura en ambos sentidos; con un factor de asociación del movimiento de AAPP hacía agricultura de un 2,78 y viceversa de un 2,65. Para trabajar en la administración pública se requiere, por lo general aprobar, unas oposiciones, por tanto, podemos predecir que es fácil pasar de un sector a otro, existiendo para cada caso en torno a 2000 colocaciones. En lo que atañe al movimiento geográfico de esta movilidad, en la Tabla 6 se afirma que, tanto para un sentido como para otro, la mayoría se producen en provincias andaluzas, sobretodo en Jaén; junto con Badajoz, que es otra ciudad con una gran extensión dedicada al cultivo. Este tipo de sectores se caracterizan por requerir un bajo nivel de estudios, como se refleja en la Tabla 6, y asimismo las ocupaciones que ejercen los trabajadores son de escasa cualificación. Hay que resaltar que, aunque parezca en un primer momento que este tipo de puestos los ocupan generalmente hombres, en estos dos casos se cumple la paridad de género. Por otro lado, como estos individuos no poseen títulos universitarios en su mayoría, las rentas que obtienen son menores, sobre todo en la movilidad hacía el grupo de agricultura. Aunque los contratos que pasan a trabajar para la administración pública son mejor remunerados, la duración media hasta que cambian de trabajo suele ser en promedio más reducida, llegando al mes aproximado en ambos casos.

Tabla 6. Análisis celdas AAPP y agricultura

ACTIVIDAD ORIGEN	ACTIVIDAD DESTINO	COLOCACIONES	FA	CARACTERÍSTICA	SUBPARTADO	TOTAL	PORCENTAJE
AAPP	Agricultura	1984	2,78	PROVINCIA	Jaén	396	19,96%
					Sevilla	303	15,27%
					Badajoz	238	12,00%
					Granada	154	7,76%
					Huelva	145	7,31%
				OCUPACIÓN	Mayores 18 no cualif	1827	92,09%
					Oficiales 1ª y 2ª	93	4,69%
					Oficiales 3ª y especialistas	43	2,17%
					Ingenieros, licenciados, alta dirección	4	0,20%
				N. ESTUDIO	Primaria y bachiller elemental; FP básica	708	35,69%
					Sin graduado escolar	443	22,33%
					Primaria incompleta	226	11,39%
					Graduado escolar	217	10,94%
OTROS	RENDA AL DÍA (€/día)	33,5					
	GÉNERO (Hombre:0 - Mujer:1)	0,5					
	EDAD MEDIA (años)	42,4					
	DURACIÓN (días)	35,8					
Agricultura	AAPP	2046	2,65	PROVINCIA	Jaén	412	20,14%
					Sevilla	263	12,85%
					Badajoz	219	10,70%
					Granada	182	8,90%
					Huelva	159	7,77%
				OCUPACIÓN	Mayores 18 no cualif	1805	88,22%
					Oficiales 1ª y 2ª	114	5,57%
					Subalternos	33	1,61%
					Auxiliares administrativos	21	1,03%
				N. ESTUDIO	Primaria y bachiller elemental; FP básica	778	38,03%
					Sin graduado escolar	451	22,04%
					Primaria incompleta	256	12,51%
					Graduado escolar	199	9,73%
OTROS	RENDA AL DÍA (€/día)	53,6					
	GÉNERO (Hombre:0 - Mujer:1)	0,5					
	EDAD MEDIA (años)	42,7					
	DURACIÓN (días)	28					

Fuente: Elaboración propia a partir de los datos.

La relación entre la probabilidad observada y la probabilidad aleatoria de la celda de *construcción* con *agricultura* es en ambos sentidos fila-columna de 2,57, con casi 2500 colocaciones por movilidad en el 2020. Ambos sectores son muy manuales y técnicos en los que la persona que se desenvuelve bien en un grupo es muy probable que también lo haga en el otro. Estos sectores, como se puede ver en la Tabla 7 no requieren gran titulación a la hora de contratar personal para un nuevo puesto, además son sectores muy volátiles porque cada proyecto suele tener una corta duración. Es por ello por lo que desencadenan continuamente un puesto con otro mediante contratos temporales, es por ello el gran número de colocaciones en un solo año, a pesar de la pandemia. Los individuos de la muestra que pasan de trabajar en la *agricultura* para ocuparse de la *construcción*, concretamente, 2420 personas; y en sentido contrario 2339.

En estos movimientos vuelven a destacar las provincias andaluzas junto con Badajoz como las mayoritarias donde se han producido las movilidades de sector, condicionado seguramente por el sector de la *agricultura*. En ambos casos con casi la misma proporción y destacando en primer lugar, la capital de Andalucía, Sevilla y en segundo lugar Córdoba. En cuanto a la ocupación de estos perfiles se refiere, más de la mitad son mayores de edad sin cualificación alguna y es por ello por lo que en ambos casos de

movilidad, el nivel de estudios es únicamente el certificado escolar. A pesar de la bajísima cualificación que poseen, la renta de la migración de individuos hacia el sector de *construcción* es casi el doble del otro sentido, esto puede deberse a la grandísima disminución de personal que hay en la actualidad de albañiles y técnicos para obras. Cabe destacar que el perfil mayoritario que cambia de sector en ambos sentidos es de hombres de en torno a 41 años de edad. Estos cambian de actividad económica en busca de una mejora de calidad. Y por último señalar la similitud de los períodos de contratos medios en ambos casos, por ello, que se pueda intercalar trabajos de un sector u otro.

Tabla 7. Análisis celdas agricultura y construcción

ACTIVIDAD ORIGEN	ACTIVIDAD DESTINO	COLOCACIONES	FA	CARACTERÍSTICA	SUBPARTADO	TOTAL	PORCENTAJE
Agricultura	Construcción	2420	2,57	PROVINCIA	Córdoba	362	14,96%
					Sevilla	337	13,93%
					Jaén	258	10,66%
					Granada	218	9,01%
					Badajoz	215	8,88%
				OCUPACIÓN	Mayores 18 no cualif	1644	67,93%
					Oficiales 1ª y 2ª	524	21,65%
					Oficiales 3ª y especialistas	230	9,50%
					Auxiliares administrativos	7	0,29%
				N. ESTUDIO	Primaria y bachiller elemental; FP básica	921	38,06%
					Sin graduado escolar	558	23,06%
					Primaria incompleta	330	13,64%
					Graduado escolar	221	9,13%
OTROS	RENTA AL DÍA (€/día)	60,5					
	GÉNERO (Hombre:0 - Mujer:1)	0,7					
	EDAD MEDIA (años)	41					
	DURACIÓN (días)	38,9					
Construcción	Agricultura	2339	2,57	PROVINCIA	Córdoba	348	14,88%
					Sevilla	333	14,24%
					Jaén	317	13,55%
					Badajoz	218	9,32%
					Granada	188	8,04%
				OCUPACIÓN	Mayores 18 no cualif	2084	89,10%
					Oficiales 1ª y 2ª	164	7,01%
					Oficiales 3ª y especialistas	70	2,99%
					Ingenieros, licenciados, alta dirección	5	0,21%
				N. ESTUDIO	Primaria y bachiller elemental; FP básica	899	38,44%
					Sin graduado escolar	538	23,00%
					Primaria incompleta	328	14,02%
					Graduado escolar	195	8,34%
OTROS	RENTA AL DÍA (€/día)	33,6					
	GÉNERO (Hombre:0 - Mujer:1)	0,7					
	EDAD MEDIA (años)	41,1					
	DURACIÓN (días)	34,9					

Fuente: Elaboración propia a partir de los datos.

Finalmente, acentuar la relación de asociación de ratio 2,02 entre los individuos que trabajan en el grupo de *construcción* y se trasladan a trabajar en el de *suministros* durante el año 2020. El número de colocaciones en las que se produce esta movilidad es bastante más bajo que las anteriores, sólo 139 personas son las que se mueven de un sector a otro. Entre esta migración, la mayoría se producen en Madrid, capital del país, seguido de Barcelona, Las Palmas, Sevilla y Málaga; todas forman parte de las principales provincias donde se realizan la mayoría de las actividades económicas de España, además, de la ampliación estructural tanto en viviendas como en nuevas oficinas. Estos individuos ocupan puestos no cualificados o como oficiales y por tanto su nivel de estudio es escaso; escolar o de la primaria

incompleta. A pesar de que este fragmento de la muestra no posee titulación universitaria, la renta es elevada, aunque no al mismo nivel que los empleos con titulación universitaria. Existe un 5% de ingenieros o de alta dirección en esta movilidad de actividad y seguramente este segmento de individuos sea el que suba la media del salario bruto diario. La duración promedio de estos nuevos contratos supera los tres meses que en comparación con otras movilizaciones es sustancialmente superior. Por otro lado, el perfil que realiza este cambio suele ser en su gran mayoría hombres de 41 años de nuevo que tratan de mejorar su situación.

Tabla 8. Análisis celdas construcción y suministros

ACTIVIDAD ORIGEN	ACTIVIDAD DESTINO	COLOCACIONES	FA	CARACTERÍSTICA	SUBPARTADO	TOTAL	PORCENTAJE
Construcción	Suministros	139	2,02	PROVINCIA	Madrid	18	12,95%
					Barcelona	11	7,91%
					Las Palmas	10	7,19%
					Sevilla	10	7,19%
					Málaga	9	6,47%
				OCUPACIÓN	Mayores 18 no cualif	58	41,73%
					Oficiales 1ª y 2ª	43	30,94%
					Oficiales 3ª y especialistas	18	12,95%
					Ingenieros, licenciados, alta dirección	7	5,04%
				N. ESTUDIO	Primaria y bachiller elemental; FP básica	49	35,25%
					Primaria incompleta	18	12,95%
					FP segundo grado y superior	15	10,79%
					Sin graduado escolar	14	10,07%
				OTROS	RENDA AL DÍA (€/día)	67,3	
					GÉNERO (Hombre:0 - Mujer:1)	0,9	
					EDAD MEDIA (años)	41,3	
DURACIÓN (días)	93,5						

Fuente: Elaboración propia a partir de los datos.

Con este análisis podemos asegurar que siempre los individuos que poseen más formación suelen ocupar los puestos más elevados y con un mejor salario medio. Aunque hay excepciones como en el sector de la construcción, que en la actualidad se está sufriendo una gran demanda de personal, pero al no haber suficientes profesionales su remuneración se eleva en cierta forma para encontrar trabajadores. La paridad en estas movilizaciones de actividad económica suele ser predominante en hombres para trabajos manuales relacionados con el manejo de herramientas o máquinas y en mujeres para trabajos en relación con las personas.

- Comparativa temporal: 2019 vs. 2020

A pesar de que en el 2020 hay menos movilidad entre actividades que en el 2019, las asociaciones que se han creado entre los sectores de actividad son algo más fuertes en 2020 que en el 2019 (el mapa de calor de 2019 se muestra en el Anexo 1 de este trabajo). En concreto, destacaríamos la aparición de nuevas conexiones en 2020, como las observadas en el bicluster formado por los sectores de *servicios financieros y a empresas, organismos extraterritoriales, educación y sanidad*.

6.2 Secuencias de Ocupaciones

En este caso se analizan las movilidades de ocupación, es decir, el cambio de puesto que ejercen los individuos de la muestra dentro de una compañía, con independencia de si ha cambiado o no de empresa. Dentro de las 794.783 contrataciones en la MCVL –incluyendo aquellas colocaciones que empezaron antes de 2020 pero que no habían concluido en 2020–, 129.135 han implicado movilidad de ocupación. Teniendo en cuenta el año anterior a la pandemia, se establece que se han producido menos movilidades de ocupación en 2020 (129.135 respecto a 188.561). Hay que tener en cuenta que los datos totales del número de colocaciones en ambos años difirieren un poco de los contabilizados en la Tabla 9 de actividades porque hay movilidades desconocidas distintas en cada tipo que no se contabilizan.

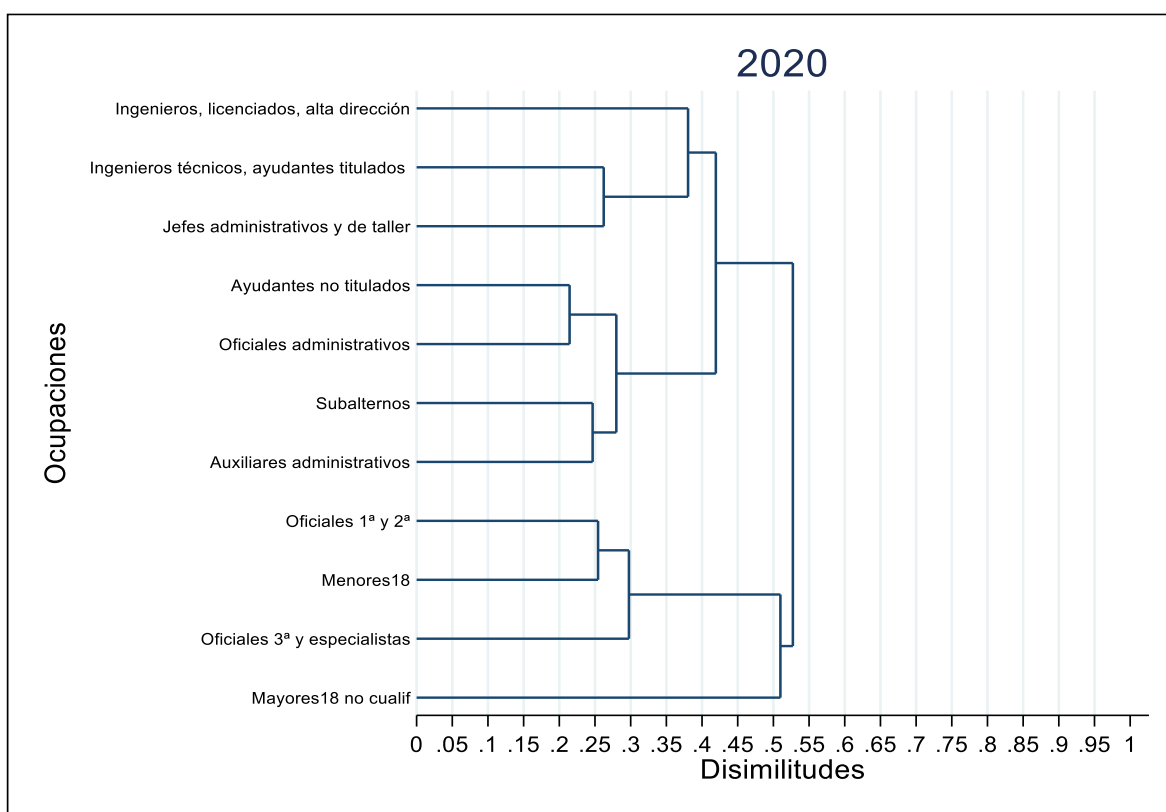
Tabla 9. *Tabla total colocaciones de secuencias de ocupaciones*

PERIODO	CASO	COLOCACIONES
Año 2020	Individuos con más de una colocación en 2020	794.783
	Colocaciones de las que conocemos origen y destino	584.581
	Sin movilidad de ocupación	455.446
	Con movilidad de ocupación	129.135
Año 2019	Individuos con más de una colocación en 2019	1.006.847
	Colocaciones de las que conocemos origen y destino	782.291
	Sin movilidad de ocupación	593.730
	Con movilidad de ocupación	188.561

Fuente: Elaboración propia a partir de los datos.

Para la secuencia de ocupaciones se ha configurado 11 posiciones principales diferentes teniendo como referencia la CON 2011. Con estas ocupaciones se representa el dendrograma de la Figura 9 que es la representación visual donde las ocupaciones se agrupan según la forma en que se emparejan con las ocupaciones del siguiente puesto.

Figura 9. Dendrograma de ocupaciones



Fuente: Elaboración propia a partir de los datos.

Dos ocupaciones se emparejan entre sí relativamente pronto en el dendrograma de ocupaciones porque la forma en la que se asocian con el resto de las ocupaciones lo hacen de una forma parecida. Las relaciones entre las ocupaciones de la Figura 9 se clasifican en tres grandes grupos. Los titulados y de alta categoría como los *ingenieros o licenciados*, los *ingenieros técnicos* y los *jefes administrativos*; Seguido de la categoría media como *ayudantes no titulados*, *oficiales administrativos*, *subalternos* y *auxiliares administrativos*; Y por último las ocupaciones con menos cualificación, los *Oficiales de 1ª y 2ª*, *menores de 18*, *oficiales de 3ª y especialistas* y los *mayores de 18 no cualificados*.

Observando el dendrograma es necesario resaltar que la ocupación de *menores de 18* se asocia de manera más similar a los *oficiales de 1ª y 2ª* que a los *mayores de 18 no cualificados*. Esto puede ser debido a que por lo general cada vez las nuevas generaciones tienen más nivel educativo y por ello se emparejan con las demás ocupaciones de forma parecida a los *oficiales de 1ª y 2ª*. Este orden de posiciones se aprovechará también para la construcción del mapa de calor de ocupaciones de la Tabla 10.

Tabla 10. Mapa de calor de ocupaciones

	Ingenieros, licenciados, alta dirección	Ingenieros técnicos, ayudantes titulados	Jefes administrativos y de taller	Ayudantes no titulados	Oficiales administrativos	Subalternos	Auxiliares administrativos	Oficiales 1ª y 2ª	Menores 18	Oficiales 3ª y especialistas	Mayores 18 no cualif
Ingenieros, licenciados, alta dirección		7,3	4,5	1,4	1,9	0,4	0,8	0,4	0,2	0,2	
Ingenieros técnicos, ayudantes titulados	7,9		3,1	1,8	1,7	0,8	1,0	0,4	0,2	0,3	
Jefes administrativos y de taller	4,0	2,8		1,7	2,3	0,6	1,3	0,6	0,6	0,4	0,4
Ayudantes no titulados	1,4	2,0	1,7		1,4	1,3	1,3	0,7	0,5	0,7	0,8
Oficiales administrativos	2,0	1,9	2,3	1,6		1,1	2,0	0,7	0,3	0,5	0,6
Subalternos	0,6	1,0	0,7	1,3	1,2		1,9	0,9	0,8	0,8	1,0
Auxiliares administrativos	1,0	1,2	1,4	1,3	2,3	1,8		0,8	0,4	0,8	1,0
Oficiales 1ª y 2ª	0,4	0,5	0,7	0,8	0,8	0,9	0,9		0,5	1,5	1,9
Menores 18	0,2			0,8	0,3	0,5	0,6	0,5		0,8	2,7
Oficiales 3ª y especialistas	0,2	0,3	0,4	0,7	0,6	0,8	0,9	1,5	1,0		2,0
Mayores 18 no cualif	0,2	0,3	0,3	0,7	0,6	1,1	1,1	1,9	2,6	2,2	

Fuente: Elaboración propia a partir de los datos.

- Justificación de la formación de clusters de actividades

Un mapa de calor es una matriz origen destino, en este caso de ocupaciones, que refleja cómo las ocupaciones origen situadas en las filas se emparejan con las ocupaciones destino de las columnas. Para trabajar con la similitud promedio hay que moverse en dos direcciones, tanto en cómo se emparejan las filas, ocupaciones del empleo anterior, como las columnas, ocupaciones del empleo actual. Si en ambas direcciones se parecen entre sí, la proximidad promedio es alta, en estos casos, se puede hablar de que forman un cluster. Las asociaciones más fuertes han creado un cluster entre las posiciones más altas, con más nivel de estudios. Por otro lado, también se han creado clúster entre las posiciones bajas con gran número de colocaciones, aunque el factor de asociación no es tan elevado como en las ocupaciones altas.

- Análisis descriptivo de las celdas con mayor factor de asociación (FA)

Analizando cada celda del mapa de calor, se identifican cuáles son las categorías-fila con mayor propensión a asociarse con una determinada categoría-columna. De todas estas celdas se pasarán a analizar las más destacables (FA= 7,9 7,3 4,5 3,9 3,1 y 2,8). Estos seis casos están conectados dos a dos ya que, en este mapa de calor, las asociaciones fuertes entre dos ocupaciones se presentan tanto en un sentido como en otro. Aunque no es lo mismo moverse de un titulado medio a titulado superior, en el que el empleado asciende de forma natural su posición que moverse de titulado superior a medio, ya que este caso, sería un escenario de sobre-educación o sobre-cualificación.

Si analizamos la categoría superior de *ingenieros, licenciados, alta dirección*, siendo el puesto de origen (como fila), estos individuos se reparten en su mayoría a *ingenieros técnicos, ayudantes titulados*

y seguidos de *jefes administrativos y de taller*. En ambas movilidades se produce una reducción a nivel jerárquico en la que se infravalora la ocupación de los individuos altamente cualificados. Sin embargo, si observamos la posición más elevada como destino (su columna), es el destino más propenso tanto de *ingenieros técnicos, ayudantes titulados* como de *jefes administrativos y de taller*. Aquí se consigue una visión más optimista en el ámbito laboral, donde si un individuo esta en los dos estratos más alto por debajo del de *ingenieros, licenciados, alta dirección*; puede pensar que tiene bastante opciones de que su próxima posición sea la más alta, obteniendo un ascenso en su carrera profesional.

El clúster con las relaciones de asociación más fuertes corresponde a la pareja de *ingenieros, licenciados, alta dirección* con los *ingenieros técnicos, ayudantes titulados*. El mayor factor de asociación es de 7,9 y representa la movilidad de los *ingenieros técnicos, ayudantes titulados* hacia los *ingenieros, licenciados, alta dirección*, con 1695 colocaciones, este movimiento es el natural de un trabajador que asciende. Sin embargo, en el sentido contrario con 1367 colocaciones y un factor de asociación de 7,3 representa la propensión de asociarse siendo la ocupación origen la categoría más alta, *ingenieros, licenciados, alta dirección*.

Tabla 11. Análisis celdas *ingenieros, licenciados, alta dirección y ingenieros técnicos, ayudantes titulados*

OCUPACIÓN ORIGEN	OCUPACIÓN DESTINO	COLOCACIONES	FA	CARACTERÍSTICA	SUBPARTADO	TOTAL	PORCENTAJE
Ingenieros técnicos, ayudantes titulados	Ingenieros, licenciados, alta dirección	1695	7,90	PROVINCIA	Madrid	362	21,36%
					Barcelona	333	19,65%
					Valencia	133	7,85%
					Sevilla	73	4,31%
					Murcia	59	3,48%
				ACTIVIDAD	Educación	849	50,09%
					OtrosServicios	258	15,22%
					Sanidad	195	11,50%
					AAPP	169	9,97%
				N. ESTUDIO	Licenciado o Grado Universitario	621	36,64%
					Máster Universitario, Doctorado y postgrado	551	32,51%
					Diplomados en Escuelas Universitarias	153	9,03%
					Bachiller o más	96	5,66%
OTROS	RENTA AL DÍA (€/día)	102,8					
	GÉNERO (Hombre:0 - Mujer:1)	0,5					
	EDAD MEDIA (años)	38,7					
	DURACIÓN (días)	88,9					
Ingenieros, licenciados, alta dirección	Ingenieros técnicos, ayudantes titulados	1367	7,30	PROVINCIA	Madrid	292	21,36%
					Barcelona	230	16,83%
					Valencia	106	7,75%
					Sevilla	54	3,95%
					Alicante	53	3,88%
				ACTIVIDAD	Educación	499	36,50%
					OtrosServicios	272	19,90%
					Sanidad	266	19,46%
					AAPP	153	11,19%
				N. ESTUDIO	Licenciado o Grado Universitario	565	41,33%
					Máster Universitario, Doctorado y postgrado	422	30,87%
					Diplomados en Escuelas Universitarias	103	7,53%
					Bachiller o más	71	5,19%
OTROS	RENTA AL DÍA (€/día)	148,3					
	GÉNERO (Hombre:0 - Mujer:1)	0,4					
	EDAD MEDIA (años)	37,7					
	DURACIÓN (días)	92,3					

Fuente: Elaboración propia a partir de los datos.

Estas colocaciones en ambos sentidos se han producido en las cuatro principales provincias españolas a nivel industrial como son Madrid y Barcelona, seguidas de Valencia y Sevilla como se representa en la Tabla 11. Lo característico de este movimiento es que en ambas direcciones la mayor parte se producen hacia el sector de la educación, entendemos que siendo licenciados estos individuos pueden impartir clase tanto en institutos como en la Universidad además de otros contratos intrínsecamente ligados con la educación. El perfil de estos individuos en su mayoría titulados con incluso una parte importante con titulación de postgrado, jóvenes y con una paridad entre hombres y mujeres casi exacta. Gracias a los datos reflejados en la MCVL los individuos que hacen estas dos movilizaciones en promedio han declarado entre 102,8 y 148,3€/día en todo el 2020. Es destacable que, el que posee mayor remuneración sea el escenario de sobrecualificación en el que la categoría destino sea más baja que la de origen, por esto se puede deducir que, aunque estas personas hayan bajado de categoría el sueldo posiblemente haya permanecido estable. Por último, hay que señalar que, por lo general, los contratos en este escenario también son de media más duraderos que los de ascenso.

Otro clúster de pareja de ocupaciones relacionadas entre sí es la de *ingenieros, licenciados, alta dirección* con *jefes administrativos y de taller*, representado en la Tabla 12. Estas celdas poseen una propensión a emparejarse mayor el escenario de sobreeducación, por el cuál un *ingeniero o licenciado de alta dirección* se convierte en *jefe* descendiendo así su categoría profesional independientemente del nivel de estudios que tengan, que en el escenario de ascenso. Ambas movilizaciones se vuelven a producir hacia las mismas provincias lideradas por Madrid y Barcelona, aunque Sevilla sigue estando presente entre las cuatro principales provincias donde se produce el escenario de ascenso cuando la movilidad es hacia *ingeniero o licenciado de alta dirección* es sustituida por Girona. En cuanto a la actividad económica, más de la mitad de los casos de sobrecualificación se producen en otros servicios, donde se encuentra toda la parte industrial mientras que en el escenario de ascenso esta más repartido los casos entre las actividades de educación y otros servicios. En este clúster sí se confirma que en la movilidad hacia la categoría más alta los individuos declaran un salario mayor que en el sentido contrario, además las condiciones de contrato, en cuanto a duración son bastante superiores en promedio. Es muy característico que el perfil de individuos de *alta dirección* que bajen a *jefes administrativos y de taller* sean de edad media de casi 70 años, ya entrados en la jubilación. Como podría ser predecible, existe mayoría de hombres respecto a mujeres en ambas movilizaciones.

Tabla 12. Análisis celdas ingenieros, licenciados, alta dirección y jefes administrativos y de taller

OCUPACIÓN ORIGEN	OCUPACIÓN DESTINO	COLOCACIONES	FA	CARACTERÍSTICA	SUBPARTADO	TOTAL	PORCENTAJE
Ingenieros, licenciados, alta dirección	Jefes administrativos y de taller	725	4,50	PROVINCIA	Madrid	227	31,31%
					Barcelona	171	23,59%
					Valencia	40	5,52%
					Girona	23	3,17%
					Pontevedra	21	2,90%
				ACTIVIDAD	OtrosServicios	479	66,07%
					Educación	78	10,76%
					ComHostTte&Comunic	59	8,14%
					IndManufacturera	35	4,83%
				N. ESTUDIO	Licenciado o Grado Universitario	256	35,31%
					Máster Universitario, Doctorado y postgrado	164	22,62%
					Bachiller	79	10,90%
				OTROS	Bachiller o más	60	8,28%
RENTA AL DÍA (€/día)	174,6						
GÉNERO (Hombre:0 - Mujer:1)	0,6						
EDAD MEDIA (años)	39,3						
Jefes administrativos y de taller	Ingenieros, licenciados, alta dirección	747	3,90	PROVINCIA	Madrid	207	27,71%
					Barcelona	158	21,15%
					Valencia	50	6,69%
					Sevilla	25	3,35%
					Pontevedra	23	3,08%
				ACTIVIDAD	Educación	249	33,33%
					OtrosServicios	235	31,46%
					ComHostTte&Comunic	90	12,05%
					AAPP	46	6,16%
				N. ESTUDIO	Licenciado o Grado Universitario	256	34,27%
					Máster Universitario, Doctorado y postgrado	176	23,56%
					Bachiller	84	11,24%
				OTROS	Bachiller o más	58	7,76%
RENTA AL DÍA (€/día)	286,7						
GÉNERO (Hombre:0 - Mujer:1)	0,6						
EDAD MEDIA (años)	39						
				DURACIÓN (días)	101		

Fuente: Elaboración propia a partir de los datos.

Por último, hay que destacar las celdas correspondientes a los factores de asociación 3,1 y 2,8, como la relación de *ingenieros técnicos, ayudantes titulados* hacia *jefes administrativos y de taller* y viceversa, respectivamente (Tabla 13). En ambos casos de movilidad se producen con menos frecuencia y menos propensión a emparejarse que los anteriores. Por lo que afecta a la actividad económica, en el escenario de sobrecualificación (de *ingenieros técnicos, ayudantes titulados* hacia *jefes administrativos y de taller*), sucede en más de la mitad de los casos en el sector otros servicios, sin embargo, en el escenario de ascenso, sobresale el sector educativo por encima del de otros servicios. En otro orden de cosas, los individuos que sufren la movilidad de ascenso suelen tener un contrato promedio más largo, aunque en la situación de sobreeducación, estas personas han declarado un salario bruto medio superior durante todo el 2020. Respecto a la igualdad de género, destaca que en la movilidad hacia los *jefes administrativos y de taller* la mayoría en hombres, mientras que, en el sentido contrario sí que existe paridad.

Tabla 13. *Análisis celdas ingenieros licenciados, ingenieros técnicos y jefes administrativos y de taller*

OCUPACIÓN ORIGEN	OCUPACIÓN DESTINO	COLOCACIONES	FA	CARACTERÍSTICA	SUBPARTADO	TOTAL	PORCENTAJE
Ingenieros técnicos, ayudantes titulados	Jefes administrativos y de taller	663	3,10	PROVINCIA	Madrid	152	22,93%
					Barcelona	141	21,27%
					Valencia	32	4,83%
					Zaragoza	24	3,62%
					Sevilla	23	3,47%
				ACTIVIDAD	OtrosServicios	375	56,56%
					Educación	107	16,14%
					Sanidad	79	11,92%
					ComHostTte&Comunic	38	5,73%
				N. ESTUDIO	Licenciado o Grado Universitario	222	33,48%
					Máster Universitario, Doctorado y postgrado	89	13,42%
					Diplomados en Escuelas Universitarias	75	11,31%
					Bachiller	74	11,16%
OTROS	RENTA AL DÍA (€/día)	140,4					
	GÉNERO (Hombre:0 - Mujer:1)	0,6					
	EDAD MEDIA (años)	37,3					
	DURACIÓN (días)	58,7					
Jefes administrativos y de taller	Ingenieros técnicos, ayudantes titulados	622	2,80	PROVINCIA	Madrid	159	25,56%
					Barcelona	109	17,52%
					Sevilla	24	3,86%
					Granada	23	3,70%
					Valencia	22	3,54%
				ACTIVIDAD	Educación	192	30,87%
					OtrosServicios	186	29,90%
					AAPP	94	15,11%
					Sanidad	73	11,74%
				N. ESTUDIO	Licenciado o Grado Universitario	218	35,05%
					Máster Universitario, Doctorado y postgrado	87	13,99%
					Diplomados en Escuelas Universitarias	81	13,02%
					Bachiller	68	10,93%
OTROS	RENTA AL DÍA (€/día)	117,8					
	GÉNERO (Hombre:0 - Mujer:1)	0,5					
	EDAD MEDIA (años)	37,6					
	DURACIÓN (días)	77,6					

Fuente: Elaboración propia a partir de los datos.

A modo de conclusión, es necesario afirmar que, con respecto a la secuencia de ocupaciones mostrada anteriormente, la mayoría de las movilidades se desarrollan en las tres categorías más altas (*ingenieros licenciados, ingenieros técnicos y jefes administrativos*). Si se observa cada tabla, los casos de ascenso se producen mayoritariamente en la actividad económica de la educación, sector donde a simple vista hay bastante movilidad entre las altas ocupaciones. No obstante, en algunas ocasiones se producen movilidades hacia una categoría inferior o incluso hacia dos categorías por debajo. Este hecho permite atisbar que en estos casos se produce una sobre-cualificación en el individuo de su formación académica con respecto al rango laboral que ocupa. En definitiva, la carga salarial que perciben los trabajadores es mayor cuanto mayor es el grado de desempeño de la persona, o lo que es lo mismo, cuanto mayor nivel de estudio posean, con independencia en muchos casos del cargo.

- Comparativa temporal: 2019 vs. 2020

Además de que en el 2020 hay al menos 50.000 movilidades menos que en el año anterior, además las asociaciones que se han creado entre las ocupaciones no son tan fuertes como en el 2019 donde llegaban a FA de 9,2 (mapa de calor en anexo 3). Posiblemente no se han podido producir tantas movilidades entre estos dos grupos durante el año 2020 como si pasó en el año pre-COVID.

6.3 Secuencias de Provincias

En este apartado se analizarán las movilidades de provincias, los cambios de trabajo de los individuos de la muestra que han necesitado trasladarse a una provincia distinta con independencia de si ha cambiado o no de empresa. Los trabajadores que han tenido algún cambio de contrato en 2020 acumulan un total de 794.821 contrataciones en la MCVL. De todos estos contratos, 55.832 han implicado movilidad de provincia. Comparando con el año anterior a la pandemia, se observa que se han producido menos movilidades de provincias en 2020 (55.832 vs. 73.764); nótese que también ha habido un descenso de más de 200.000 colocaciones entre 2019 y 2020 dentro de nuestra muestra (esta muestra viene dada por los trabajadores que han tenido más de una colocación durante el año analizado).

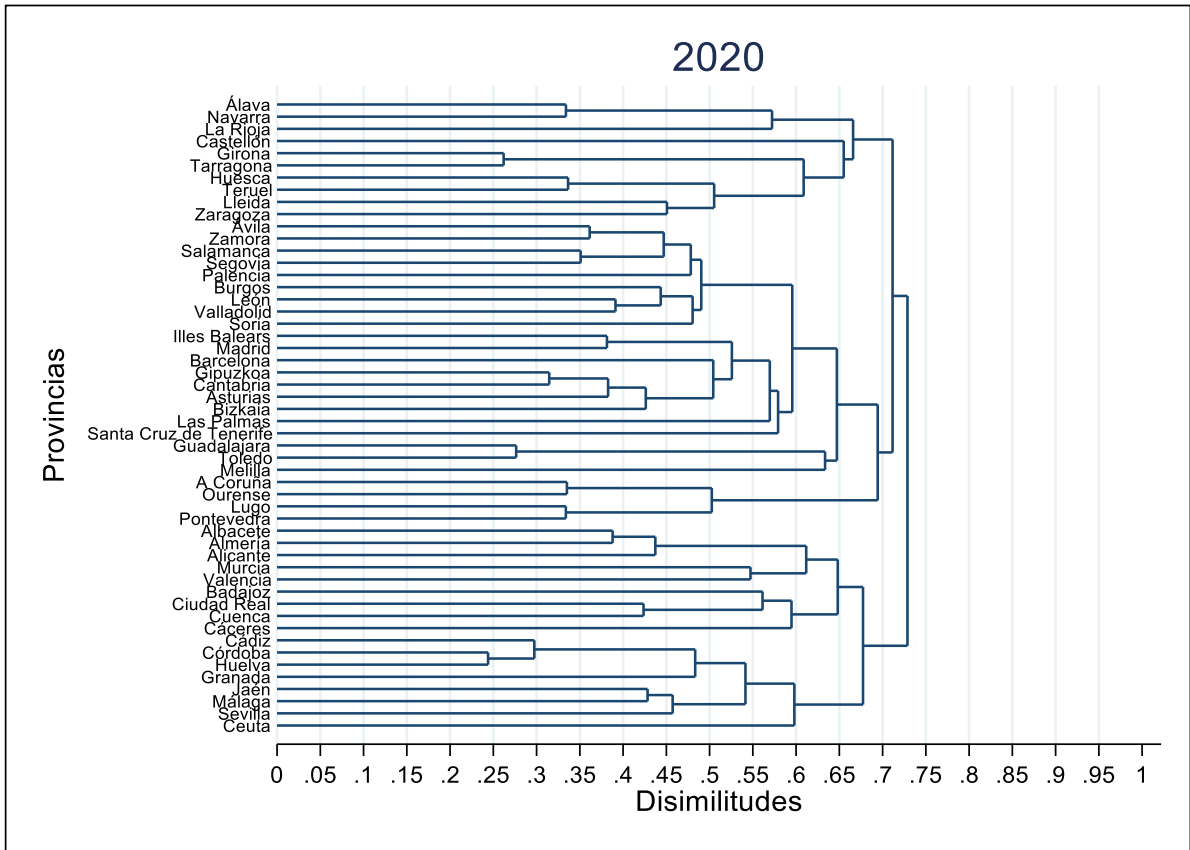
Tabla 14. *Tabla total colocaciones de secuencias de provincias*

PERIODO	CASO	COLOCACIONES
Año 2020	Individuos con más de una colocación en 2020	794.821
	Colocaciones de las que conocemos origen y destino	584.584
	Sin movilidad de provincias	528.752
	Con movilidad de provincias	55.832
Año 2019	Individuos con más de una colocación en 2019	1.006.855
	Colocaciones de las que conocemos origen y destino	782.272
	Sin movilidad de provincias	708.508
	Con movilidad de provincias	73.764

Fuente: Elaboración propia a partir de los datos.

A continuación, se puede visualizar el dendrograma (Figura 10) de provincias como resultado de la tabla de contingencias.

Figura 10. Dendrograma de provincias



Fuente: Elaboración propia a partir de los datos.

Analizando el dendrograma, podemos identificar tres grandes clusters. En primer lugar, desde si se observa el orden vertical jerárquico del gráfico lo formarían desde Álava hasta Zaragoza; el segundo de Ávila a Pontevedra y por último las provincias desde Albacete hasta Ceuta. Estos tres grupos de provincias se relacionan con el resto de las provincias con un tipo de semejanza teniendo en cuenta los dos sentidos: como filas, provincia del empleo anterior y como las columnas, provincia del empleo actual.

Tabla 15. Mapa de calor de provincias

Fuente: Elaboración propia a partir de los datos.

- Justificación de la formación de clusters de actividades

El mapa de calor de la Tabla 15 que refleja la propensión al emparejamiento o factor de asociación entre cada provincia origen-fila y cada provincia destino-columna. Estamos ante un *heatmap* con clusters de filas y columnas, bicluster, en el que se describe cómo las filas se reparten en las diferentes columnas –aunque también la interpretación de columnas a filas podría tener su interés–. Los clusters se agrupan entre sí porque se asocian de la misma forma en las diferentes provincias destino incluso con sus simétrica y sus no simétricas. En este caso, el mapa de calor puede darnos muchísima información de movimientos entre provincias por motivos laborales, tanto entre comunidades autónomas como entre provincias más o menos alejadas entre sí. En este trabajo se comentará a grandes rasgos las asociaciones de dos direcciones, se centrará en una de las dos movibilidades dejando la asociación inversa a título informativo.

Dentro de los clusters, observando el mapa de calor se pueden sacar dos ideas principales en cuanto a las movibilidades entre provincias. Este movimiento se suele producir en primer lugar en función de su cercanía física, por regiones; y la otra opción es que la movilidad se produzca hacia o desde la capital de España, *Madrid*. La movilidad regional se aprecia en la Tabla 15 porque forman cuadrados de color azul más oscuro y los suelen formar provincias cercanas entre ellas. Entre estas comunidades autónomas resaltan por su movilidad entre sus provincias, Castilla y León, Andalucía y Galicia. Por otro lado, es bastante destacable el caso de *Madrid* ya que como se aprecia en el *heatmap* forma una cruz, es decir es

una capital que sobresale tanto como provincia de origen, del empleo anterior, como de destino, del empleo actual. Madrid cuenta con la mayor actividad económica del país, además es la sede de la mayoría de las empresas españolas, es por ello por lo que al ser la provincia con mayor generación de empleo es previsible que este relacionada con la mayoría de las provincias ya sea de origen o de destino. En un primer momento se podía pensar que *Barcelona* también cumpliera un gran desempeño entre la movilidad laboral de España en el 2020, pero, aunque existe una relación con muchas provincias no se puede comparar con las de *Madrid*.

- *Análisis descriptivo de las celdas con mayor factor de asociación (FA)*

Las celdas de las que vamos a entrar más en detalle serán entre las restricciones anteriormente comentadas, las movilidades que se produzcan entre provincias de distintas comunidades autónomas y entre las regionales una a destacar. Aunque en un primer momento puede parecer importante los factores de asociación entre *Ceuta* y *Melilla*, el número de colocaciones que existen en el año 2020 son sólo dos para cada dirección. El valor tan elevado es provocado porque ambas provincias se emparejan casi de la misma forma con el resto de las provincias.

El primer caso más relevante son las movilidades entre *Barcelona* y *Madrid*, correspondiente a los factores de asociación, $FA = 2,58$ y $2,66$ y con un número de colocaciones entorno a las 1200 en cada sentido, como se puede apreciar en la Tabla 16. La actividad económica principal es la de otros servicios y en menor medida el sector de comercio, hostelería, transportes y comunicaciones. Por otro lado, la ocupación que en mayor medida tienen los individuos que realizan esta movilidad es la de auxiliares administrativos con un nivel de estudios repartido entre bachiller o titulados universitarios. Por último, la media de la renta declarada por estas personas es bastante elevada, entre 170 €/día, seguramente debido a que gran parte de estos individuos tienen la ocupación de jefes administrativos y de taller.

Tabla 16. Análisis celdas Madrid y Barcelona

PROVINCIA ORIGEN	PROVINCIA DESTINO	COLOCACIONES	FA	CARACTERÍSTICA	SUBPARTADO	TOTAL	PORCENTAJE
Madrid	Barcelona	1187	2,58	ACTIVIDAD	OtrosServicios	848	71,44%
					ComHostTe&Comunic	177	14,91%
					Sanidad	44	3,71%
					Educación	33	2,78%
					IndManufacturera	21	1,77%
				OCUPACIÓN	Auxiliares administrativos	328	27,63%
					Jefes administrativos y de taller	173	14,57%
					Oficiales administrativos	163	13,73%
					Mayores18 no cualif	156	13,14%
				N. ESTUDIO	Bachiller	298	25,11%
					Licenciado o Grado Universitario	243	20,47%
					Primaria y bachiller elemental; FP básica	175	14,74%
					Bachiller o más	75	6,32%
OTROS	RENTA AL DÍA (€/día)	178,9					
	GÉNERO (Hombre:0 - Mujer:1)	0,6					
	EDAD MEDIA (años)	36,9					
	DURACIÓN (días)	45,2					
Barcelona	Madrid	1274	2,66	ACTIVIDAD	OtrosServicios	967	75,90%
					ComHostTe&Comunic	149	11,70%
					Educación	40	3,14%
					Construcción	31	2,43%
					ServFros&Emp	31	2,43%
				OCUPACIÓN	Auxiliares administrativos	366	28,73%
					Oficiales 1ª y 2ª	213	16,72%
					Jefes administrativos y de taller	157	12,32%
					Oficiales administrativos	153	12,01%
				N. ESTUDIO	Bachiller	291	22,84%
					Licenciado o Grado Universitario	265	20,80%
					Primaria y bachiller elemental; FP básica	205	16,09%
					FP segundo grado y superior	83	6,51%
OTROS	RENTA AL DÍA (€/día)	160					
	GÉNERO (Hombre:0 - Mujer:1)	0,6					
	EDAD MEDIA (años)	36,6					
	DURACIÓN (días)	40,1					

Fuente: Elaboración propia a partir de los datos.

Otra celda de movilidad interregional destacable es entre las provincias de *Almería* y *Murcia*, con un factor de asociación de 6,74 y 6,41 y entorno a 400 colocaciones en cada caso. En este escenario, los sectores predominantes son el de otros servicios y la agricultura con unos perfiles de bajo nivel de estudios que, por ello, declaran un menor salario bruto de 41 €/día de media.

Tabla 17. Análisis celdas Almería y Murcia

PROVINCIA ORIGEN	PROVINCIA DESTINO	COLOCACIONES	FA	CARACTERÍSTICA	SUBPARTADO	TOTAL	PORCENTAJE
Almería	Murcia	413	6,74	ACTIVIDAD	OtrosServicios	237	57,38%
					Agricultura	87	21,07%
					ComHostTte&Comunic	49	11,86%
					Sanidad	14	3,39%
					Construcción	12	2,91%
				OCUPACIÓN	Mayores18 no cualif	318	77,00%
					Oficiales 1ª y 2ª	45	10,90%
					Ingenieros técnicos, ayudantes titulados	12	2,91%
					Oficiales 3ª y especialistas	10	2,42%
				N. ESTUDIO	Sin graduado escolar	153	37,05%
					Desconocido	80	19,37%
					Primaria y bachiller elemental; FP básica	54	13,08%
					Primaria incompleta	41	9,93%
OTROS	RENDA AL DÍA (€/día)	40,5					
	GÉNERO (Hombre:0 - Mujer:1)	0,8					
	EDAD MEDIA (años)	38,2					
	DURACIÓN (días)	49,3					
Murcia	Almería	401	6,41	ACTIVIDAD	OtrosServicios	190	47,38%
					Agricultura	109	27,18%
					ComHostTte&Comunic	58	14,46%
					Construcción	18	4,49%
					Educación	10	2,49%
				OCUPACIÓN	Mayores18 no cualif	300	74,81%
					Oficiales 1ª y 2ª	58	14,46%
					Oficiales 3ª y especialistas	13	3,24%
					Ingenieros técnicos, ayudantes titulados	8	2,00%
				N. ESTUDIO	Sin graduado escolar	140	34,91%
					Desconocido	83	20,70%
					Primaria y bachiller elemental; FP básica	56	13,97%
					Primaria incompleta	44	10,97%
OTROS	RENDA AL DÍA (€/día)	42,3					
	GÉNERO (Hombre:0 - Mujer:1)	0,8					
	EDAD MEDIA (años)	38,9					
	DURACIÓN (días)	45,8					

Fuente: Elaboración propia a partir de los datos.

Para las celdas de movilidad entre *Murcia* y *Alicante* de la Tabla 18, las principales actividades económicas donde se realizan estos cambios vuelven a ser otros servicios y la agricultura, además del en menor medida el sector de comercios, hostelerías, transportes y comunicaciones. Ambas provincias son destinos muy recurridos para vacaciones, por lo que es posible que el turismo provoque también una posibilidad de empleo para trabajadores de otros destinos. Los perfiles mayoritarios no tienen mucha cualificación, por tanto, el salario bruto medio también es más bajo.

Tabla 18. Análisis celdas Murcia y Alicante

PROVINCIA ORIGEN	PROVINCIA DESTINO	COLOCACIONES	FA	CARACTERÍSTICA	SUBAPARTADO	TOTAL	PORCENTAJE
Murcia	Alicante	941	8,46	ACTIVIDAD	OtrosServicios	407	43,25%
					Agricultura	256	27,21%
					ComHostTte&Comunic	143	15,20%
					Construcción	66	7,01%
					Sanidad	26	2,76%
				OCUPACIÓN	Mayores18 no cualif	652	69,29%
					Oficiales 1ª y 2ª	121	12,86%
					Oficiales 3ª y especialistas	43	4,57%
					Auxiliares administrativos	29	3,08%
				N. ESTUDIO	Sin graduado escolar	271	28,80%
					Primaria y bachiller elemental; FP básica	178	18,92%
					Desconocido	120	12,75%
					Primaria incompleta	116	12,33%
OTROS	RENDA AL DÍA (€/día)	45,1					
	GÉNERO (Hombre:0 - Mujer:1)	0,8					
	EDAD MEDIA (años)	38					
	DURACIÓN (días)	41,3					
Alicante	Murcia	937	8,08	ACTIVIDAD	OtrosServicios	494	52,72%
					Agricultura	185	19,74%
					ComHostTte&Comunic	95	10,14%
					Construcción	65	6,94%
					Sanidad	43	4,59%
				OCUPACIÓN	Mayores18 no cualif	652	69,58%
					Oficiales 1ª y 2ª	101	10,78%
					Oficiales 3ª y especialistas	55	5,87%
					Ingenieros técnicos, ayudantes titulados	31	3,31%
				N. ESTUDIO	Sin graduado escolar	247	26,36%
					Primaria y bachiller elemental; FP básica	191	20,38%
					Desconocido	123	13,13%
					Primaria incompleta	105	11,21%
OTROS	RENDA AL DÍA (€/día)	48,8					
	GÉNERO (Hombre:0 - Mujer:1)	0,8					
	EDAD MEDIA (años)	38					
	DURACIÓN (días)	46,6					

Fuente: Elaboración propia a partir de los datos.

Una vez vistas las movilidades que se desarrollan dentro de las comunidades autónomas, destacaría las celdas referidas a *Cáceres* y *Badajoz* con casi 300 colocaciones durante el 2020 cada sentido, el factor de asociación de cada movilidad es muy fuerte, entorno a 30 (Tabla 19). Sus dos filas se reparten de la misma forma entre las columnas de destino. Ambas provincias pertenecientes a Extremadura tienen bastante movilidad de una a otra por su cercanía y similitud de cultura. Los sectores que provocan esta movilidad están más repartidos que en casos anteriores, la mayor parte de las movilidades se producen en la agricultura por detrás, en la sanidad. Por otro lado, estarían el resto de las actividades económicas como otros servicios, comercios, hostelerías, transportes y comunicaciones o construcción, y por último la educación. La declaración de la renta media en estas movilidades es media, algo más elevada que en casos donde hay más proporción en los sectores de la agricultura y comercios, hostelerías, transportes y comunicaciones.

Tabla 19. Análisis celdas Cáceres y Badajoz

PROVINCIA ORIGEN	PROVINCIA DESTINO	COLOCACIONES	FA	CARACTERÍSTICA	SUBPARTADO	TOTAL	PORCENTAJE
Cáceres	Badajoz	290	31,21	ACTIVIDAD	Agricultura	80	27,59%
					Sanidad	42	14,48%
					OtrosServicios	39	13,45%
					ComHostTte&Comunic	34	11,72%
					Educación	28	9,66%
				OCUPACIÓN	Mayores18 no cualif	102	35,17%
					Oficiales 1ª y 2ª	42	14,48%
					Ingenieros técnicos, ayudantes titulados	32	11,03%
					Oficiales 3ª y especialistas	26	8,97%
				N. ESTUDIO	Primaria y bachiller elemental; FP básica	79	27,24%
					Licenciado o Grado Universitario	35	12,07%
					Máster Universitario, Doctorado y postgrado	34	11,72%
				OTROS	FP segundo grado y superior	25	8,62%
RENTA AL DÍA (€/día)	68,7						
GÉNERO (Hombre:0 - Mujer:1)	0,6						
Badajoz	Cáceres	300	30,11	ACTIVIDAD	EDAD MEDIA (años)	37,5	
					DURACIÓN (días)	58,4	
					Agricultura	85	28,33%
					Sanidad	62	20,67%
					OtrosServicios	37	12,33%
				OCUPACIÓN	Construcción	30	10,00%
					Educación	26	8,67%
					Mayores18 no cualif	109	36,33%
					Oficiales 1ª y 2ª	61	20,33%
				N. ESTUDIO	Ingenieros técnicos, ayudantes titulados	37	12,33%
					Ingenieros, licenciados, alta dirección	21	7,00%
					Primaria y bachiller elemental; FP básica	69	23,00%
				OTROS	Licenciado o Grado Universitario	33	11,00%
Máster Universitario, Doctorado y postgrado	33	11,00%					
Sin graduado escolar	30	10,00%					
OTROS	RENTA AL DÍA (€/día)	60,7					
	GÉNERO (Hombre:0 - Mujer:1)	0,6					
	EDAD MEDIA (años)	38,8					
				DURACIÓN (días)	62,4		

Fuente: Elaboración propia a partir de los datos.

En conclusión, España es un país en el que una importante parte del PIB que se produce es gracias al turismo, y por ello, esta presente en la mayoría de los cambios de movilidad. Aunque si destacamos un sector donde se realizan la mayoría de las movilidades entre las diferentes provincias sería el de otros servicios seguido de la agricultura. Según la movilidad sea hacia otra provincia de la comunidad autónoma o sea una movilidad interregional con *Madrid*, la actividad económica principal puede ser una u otra. Para el caso de una movilidad interregional, es decir hacia otra provincia suele ser en el sector de otros servicios como es el caso entre *Barcelona* y *Madrid*. En otro orden de las cosas, las movilidades dentro de una región o entre la región contigua parece ser que el sector que destaca es el de la agricultura. En cuanto a la igualdad entre hombres y mujeres, suele ser mayoritario que el genero masculino sea el que predomine entre estas movilidades, aunque no con mucha diferencia.

- *Comparativa temporal: 2019 vs. 2020*

Durante el 2020 hay menos movilidad entre provincias que en el 2019, y las asociaciones que se han creado entre las provincias son distribuidas casi de la misma forma que en el 2020 (el mapa de calor de 2019 se muestra en el Anexo 5 de este trabajo). La provincia de *Madrid* en el 2019 seguía siendo un punto de encuentro entre los trabajos de origen y de destino. En concreto, se destacaría la aparición de

nuevas conexiones en 2020, como las observadas en el bicluster formado por los sectores de *servicios financieros y a empresas, organismos extraterritoriales, educación y sanidad*.

7 CONCLUSIONES

Este Trabajo Fin de Máster pretende identificar las movilidades laborales (regionales y ocupacionales) más recurrentes entre los trabajadores dados de alta en la Seguridad Social que cambian de contrato dentro del año. Los microdatos proceden de la Muestra Continua de Vidas Laborales, años 2019 y 2020, que publica la Seguridad Social. A nivel metodológico, se encuadra este estudio en el marco del análisis categórico de las tablas de contingencia y los mapas de calor (o *heatmaps*). Estas tablas permiten medir la frecuencia y el grado de asociación de los cruces de las distintas categorías de dos variables nominales. En nuestro caso, dichas variables van a ser el sector de actividad del empleo anterior y el sector de actividad del empleo actual (cuando un trabajador cambia de empleo), el grupo de ocupación del puesto en el empleo anterior y en el actual, y la provincia en el empleo anterior y en el actual. Estos tres análisis se desarrollan por separado (no se consideran dobles o triples movilidades) y solo tienen en cuenta aquellos cambios de empleo donde, efectivamente, el trabajador ha cambiado de actividad, ocupación o provincia (según corresponda).

La tabla de contingencia cuenta el número de colocaciones observadas en cada cruce de categorías origen-destino (de actividad, ocupación o provincia). Por su parte, los mapas de calor generados ofrecen una representación visual de las asociaciones entre las filas y las columnas de la tabla de contingencia. Cada celda muestra el factor de asociación entre la categoría-fila (por ejemplo, la actividad en el empleo de origen) y la categoría-columna (la actividad en el empleo de destino). Este factor de asociación muestra la tendencia a emparejarse entre dos categorías concretas (fila y columna de la tabla de contingencia). Si este factor de asociación es mayor que la unidad, es porque existe una asociación entre las dos categorías correspondientes a la celda más fuerte que la que se observaría en un escenario donde la asignación de movilidades a filas y columnas fuera el resultado de un proceso aleatorio. Para identificar patrones de correspondencia de forma visual, en los *heatmaps* las celdas se colorean con mayor intensidad cuando el factor de asociación de la celda aumenta.

Con este análisis de movilidades entre actividades económicas se puede concluir que siempre los individuos que poseen más nivel educativo suelen ocupar los puestos más elevados y con renta diaria media más elevada. Sin embargo, en la construcción, su remuneración no depende del nivel de estudios, sino de la oferta de trabajo que haya en el sector. Así pues, es común encontrar más hombres para trabajos de manipulación de herramientas o máquinas y, por el contrario, hay mayor cantidad de mujeres para trabajos en relación con las personas.

En las movilidades de ocupaciones analizadas hay casos en los que se produce sobre-educación, por ejemplo, trabajadores del grupo de *ingenieros, licenciados, y alta dirección* obtienen nuevos empleos en los grupos de *jefes administrativos y de taller o ingenieros licenciados, ingenieros técnicos*, todo ello

con un factor de asociación no despreciable. En este escenario, la ocupación a desempeñar por el trabajador baja de nivel, por eso hablamos de sobre-educación.

En el análisis de este Trabajo se han observado algunas asociaciones fuertes en cuanto a la movilidad provincial, que en un primer momento podrían ser esperadas, como por ejemplo la propensión a la asociación entre los individuos que tienen el empleo de origen en *Cáceres* y el de destino en *Badajoz* (su factor de asociación es de 31,21) o la tendencia a relacionarse de los *ingenieros técnicos, ayudantes titulados que pasan a ser ingenieros, licenciados, alta dirección* en su nuevo empleo (su factor de asociación es de 7,9). Lo importante de nuestra metodología no es tanto que identifique asociaciones que puedan resultar esperables, sino que cuantifica con una medida estadística el grado o fortaleza de dicha asociación.

Las políticas activas de empleo deben tener en cuenta la existencia de movilidad a nivel ocupacional, a nivel de sectores de actividad y a nivel geográfico en el mercado de trabajo español. Dichas políticas deben favorecer estas movilidades. Las medidas de fomento de la movilidad geográfica (como mejorar el acceso a la vivienda y a los alquileres o mejorar las infraestructuras y los costes del transporte) deben ser compatibles con los fenómenos migratorios y demográficos del territorio español. Pensamos que resulta necesario incentivar el establecimiento de empresas en zonas deprimidas o que tienden a la despoblación. A nivel ocupacional, se debe favorecer el reciclaje de ciertos trabajadores para mejorar sus opciones de movilidad (por ejemplo, mediante la Formación Profesional Dual o el emprendimiento). Finalmente, a nivel de sectores de actividad, resulta necesario facilitar la transición del empleo desde sectores en declive hacia sectores emergentes.

El campo de estudio abierto en este Trabajo permite plantear nuevas líneas de investigación futuras. Por ejemplo, se pueden agregar al análisis nuevas dimensiones de la movilidad laboral, como las que hacen referencia al tipo de contrato (fijo/temporal, jornada completa/parcial, etc.), a la duración de cada colocación o a la existencia de episodios de cobro de prestaciones por desempleo que haya podido tener el trabajador entre empleo y empleo. Asimismo, se podrían analizar varias movilidades al mismo tiempo, es decir, en cuántos episodios de empleo el individuo ha cambiado simultáneamente de ocupación y provincia o de provincia y actividad, etc.

REFERENCIAS

- Álvarez de Toledo, P., Núñez, F. y Usabiaga, C. (2013). Análisis “Cluster” de los Flujos Laborales Andaluces”. *Revista de Estudios Regionales*, 97, 195-221.
- Álvarez de Toledo, P., Núñez, F. y Usabiaga, C. (2014): “An Empirical Approach on Labour Segmentation. Applications with Individual Duration Data”. *Economic Modelling*, 36, 252-267.
- Álvarez de Toledo, P., Núñez, F. y Usabiaga, C. (2017). ¿Quién se Empareja con Quién en el Mercado Laboral Español? Un Análisis Cluster basado en la Muestra Continua de Vidas Laborales. *Investigación Económica*, 76 (299), 87-124.
- Álvarez de Toledo, P., Núñez, F. y Usabiaga, C. (2018). Matching and Clustering in Square Contingency Tables. Who Matches with Whom in the Spanish Labour Market. *Computational Statistics and Data Analysis*, 127, 135-159.
- Álvarez de Toledo, P., Núñez, F. y Usabiaga, C. (2020). Matching in Segmented Labor Markets: An Analytical Proposal based on High-Dimensional Contingency Tables. *Economic Modelling*, 93, 175-186.
- Arazola, M., Galán, S., y de Hevia, J. (2018). Desempleo juvenil en España: situación, consecuencias e impacto sobre la vida laboral de los adultos. *Papeles de economía española*, 3 (156), 62-75.
- Arranz, J. M., Carrasco, C., y Massó, M. (2017). La movilidad laboral de las mujeres inmigrantes en España (2007-2013). *Revista Española de Sociología*, 26 (3), 329-344.
- Arranz, J. M., García-Serrano, C., y Hernanz, V. (2018). Calidad del empleo: una propuesta de índice y su medición para el periodo 2005-2013. *Hacienda Pública Española*, 1 (225), 133-164.
- Arranz, J. M., García-Serrano, C., y Hernanz, V. (2013). How do we pursue “labormetrics”? An application using the MCVL. *Estadística Española*, 55 (181), 231-254.
- Arranz, J. M., y Serrano, C. G. (2011). Los datos fiscales de la Muestra Continua de Vidas Laborales: Algunas ideas para su explotación. *Instituto de estudios fiscales*, 5 (2011), 1-80.
- Arranz, J.M., García-Serrano, C. y Hernanz, V. (2013). Active Labour Market Policies in Spain: A Macroeconomic Evaluation. *International Labour Review*, 152 (2), 327-348.
- Bande, R., Karanassou, M. y Sala, H. (2019). Employment in Spanish Regions: Cost-Control or Growth-Enhancing Policies?. *The Annals of Regional Science*, 62 (3), 601-635.
- Bentolila, S., Dolado, J.J. y Jimeno, J.F. (2012). Reforming an Insider-Outsider Labor Market: The Spanish Experience. *IZA Journal of European Labor Studies*, 1, n. 4.
- Cebrián, I., y Toharia, L. (2008). La entrada en el mercado de trabajo. Un análisis basado en la MCVL. *Revista de economía aplicada*, 16 (1), 137-172.
- Cahuc P. y Zylberberg A. (2014). *Labor Economics*, The MIT Press, Cambridge-Massachusetts y Londres-Inglaterra.

- Cunha, F., Heckman, J. J., Lochner, L., y Masterov, D. V. (2006). Interpreting the evidence on life cycle skill formation. *Handbook of the Economics of Education*, 1, 697-812.
- Davia, M.A. y Hernanz, V. (2004). Temporary Employment and Segmentation in the Spanish Labour Market: An Empirical Analysis through the Study of Wage Differentials. *Spanish Economic Review*, 6 (4), 291-318.
- Fienberg, S. y Rinaldo, A. (2007). Three Centuries of Categorical Data Analysis: Log-Linear Models and Maximum Likelihood Estimation. *Journal of Statistical Planning and Inference*, 137 (11), 3430-3445.
- García-Cintado, A., Romero-Ávila, D. y Usabiaga, C. (2014). *Spanish Regional Unemployment. Disentangling the Sources of Hysteresis*, Springer, Berlin.
- Haedo, C. y Mouchart, M. (2018). A Stochastic Independence Approach for Measuring Regional Specialization and Concentration. *Papers in Regional Science*, 97 (4), 1151-1168.
- Jimeno, J.F. y Bentolila, S. (1998). Regional Unemployment Persistence (Spain, 1976-1994). *Labour Economics*, 5 (1), 25-51.
- Lapuerta, I. (2010). Claves para el trabajo con la Muestra Continua de Vidas Laborales. DemoSoc Working Paper, 37 (2010), 1-32.
- Liu, L. (2018). Regional Labor Mobility in Spain. IMF (International Monetary Fund), Working Paper. *IMF Working Paper*, 18 (282), 1-28.
- Long, J. y Ferrie, J. (2013). Intergenerational Occupational Mobility in Great Britain and the United States since 1850. *American Economic Review*, 103 (4), 1109-1137.
- López-Bazo, E. y Motellón, E. (2013). The Regional Distribution of Unemployment: What Do Micro-Data Tell Us?. *Papers in Regional Science*, 92 (2), 383-406.
- Moreno-Colom, S., & López-Roldán, P. (2018). El impacto de la crisis en las trayectorias laborales de las mujeres inmigrantes en España. *Cuadernos de Relaciones Laborales*, 36 (1), 65.
- Nieto, S. y Ramos, R. (2017). Overeducation, Skills and Wage Penalty: Evidence for Spain using PIAAC Data. *Social Indicators Research*, 134 (1), 219-236.
- Núñez, F. y Usabiaga, C. (2007). *La Curva de Beveridge y la Función de Emparejamiento: Revisión de sus Fundamentos Teóricos y de la Literatura Empírica, con Especial Énfasis en el Caso Español*, Factoría de Ideas, Centro de Estudios Andaluces, Sevilla.
- Padilla, I. y Planelles, J. (2021). Estudios monográficos del mercado de trabajo con la MCVL. Salarios y autónomos.
- Petitjean, F., Webb, G. y Nicholson, A. (7-10 de Diciembre de 2013). *Scaling Log-Linear Analysis to High-Dimensional Data*. Proceedings of the 13th IEEE International Conference on Data Mining, Dallas, Estados Unidos.
- Petrongolo, B. y Pissarides, C. A. (2001). Looking into the Black Box: A Survey of the Matching Function. *Journal of Economic Literature*, 39 (2), 390-431.

- Pissarides, C. (2000). *Equilibrium Unemployment Theory, 2nd Ed.*, The MIT Press, Cambridge, Mass., US.
- Pissarides, C. (2011). Equilibrium in the Labor Market with Search Frictions. *American Economic Review*, 101 (4), 1092-1105.
- Prieto, M. (2017). La movilidad geográfica laboral interior: un mecanismo para combatir el desempleo. *Revista de Economía Mundial*, 47, 179-202.
- Rey, C., Rey, S. y Viala, J. (2014). Detection of High and Low States in Stock Market Returns with MCMC Method in a Markov Switching Model. *Economic Modelling*, 41 (C), 145-155.
- Sanchez, S., & Arellano, A. Aproximación a la economía del conocimiento en las ciudades españolas a través de la Muestra Continua de Vidas Laborales: aspectos metodológicos y avance de resultados1. *Fundación de estudios de Economía Aplicada. Colección de Estudios Económicos*, 18-2010.
- Shimer, R. (2007). Mismatch. *American Economic Review*, 97 (4), 1074-1101.
- Suárez, P., Mayor, M., Cueto, B. y Salas-Olmedo, M.H. (2019). The Spatial Structure of the Labour Market Across Public Jobcentres. Does their Accessibility Matter?. *Papers in Regional Science*, 98 (3), 1359-1372.
- Usabiaga, C. (2004). *El Diferencial de Desempleo Andaluz. Análisis Macroeconómico del Mercado de Trabajo Andaluz en Comparación con el del Resto de España (1980-2000)*, Aconcagua, Sevilla.
- Usabiaga, C. y Hernández-Salmerón, M. (2021). Reflections on Idiosyncratic Labour Markets: The Spanish and Andalusian Cases. *Progress in Economics Research*, Nova Science Publishers, New York, 1-96.
- Viñuela, A., Rubiera-Morollón, F. y Fernández-Vázquez, E. (2014). Applying Economic-Based Analytical Regions: A Study of the Spatial Distribution of Employment in Spain. *The Annals of Regional Science*, 52 (1), 87-102.

ANEXOS

A continuación, se adjuntan para la secuencia de actividades se muestra una tabla con las agrupaciones de las actividades económicas según la CNAE-09 de dos dígitos. Además, se muestran las tablas relacionadas con los datos del 2019 a uso comparativo con las del 2020.

Anexo 1. Lista de agrupaciones de actividades económicas (CNAE-09)

CNAE-09	Grupos de sectores
01.- Agricultura, ganadería, caza y servicios relacionados con las mismas. 02.- Silvicultura y explotación forestal 03.- Pesca y acuicultura	Agricultura
05.- Extracción de antracita, hulla y lignito 06.- Extracción de crudo de petróleo y gas natural 07.- Extracción de minerales metálicos 08.- Otras industrias extractivas 09.- Actividades de apoyo a las industrias extractivas	Industrias Extractivas (IndExtractivas)
10.- Industria de la alimentación y 11.- Fabricación de bebidas 12.- Industria del tabaco 13.- Industria textil 14.- Confección de prendas de vestir 15.- Industria del cuero y del calzado 16.- Industria de la madera y del corcho, excepto muebles; cestería y espartería 17.- Industria del papel 18.- Artes gráficas y reproducción de soportes grabados y 58.- Edición 19.- Coquerías y refino del petróleo 20.- Industria química y 21.- Fabricación de productos farmacéuticos 22.- Fabricación de productos de caucho y plásticos 23.- Fabricación de otros productos minerales no metálicos 24.- Metalurgia: fabricación de productos de hierro, acero y ferroaleaciones 25.- Fabricación de productos metálicos excepto maquinaria y equipo 26.- Fabricación de productos informáticos, electrónicos y ópticos y 32.- Otras industrias manufactureras 27.- Fabricación de material y equipo eléctrico 28.- Fabricación de maquinaria y equipo n.c.o.p. (No clasificado en otra parte) y 33.- Reparación e instalación de maquinaria y equipo 29.- Fabricación de vehículos de motor, remolques y semirremolques 30.- Fabricación de otro material de transporte 31.- Fabricación de muebles	Industrias Manufactureras (IndManufacturera)
35.- Suministro de energía eléctrica, gas, vapor y aire acondicionado 36.- Captación, depuración y distribución de agua 37.- Recogida y tratamiento de aguas residuales, 38.- Recogida, tratamiento y eliminación de residuos: valorización y 39.- Actividades de descontaminación y otros	Suministros
41.- Construcción de edificios, 42.- Ingeniería civil y 43.- Actividades de construcción especializada	Construcción
45.- Venta y reparación de vehículos de motor y motocicletas 46.- Comercio al por mayor e intermediarios del comercio, excepto de vehículos de motor y motocicletas 47.- Comercio al por menor, excepto de vehículos de motor y motocicletas, 95.- Reparación de ordenadores, efectos personales y artículos de uso doméstico 55.- Servicios de alojamiento y 56.- Servicios de comidas y bebidas 49.- Transporte terrestre y por tubería 50.- Transporte marítimo y por vías navegables interiores 51.- Transporte aéreo 52.- Almacenamiento y actividades anexas al transporte y 79.- Actividades de agencias de viajes, operadores turísticos, servicios de reservas y actividades relacionadas 61.- Telecomunicaciones y 53.- Actividades postales y de correos	Comercio, hostelería, transporte y comunicaciones (ComHostTte&Comunic)
64.- Servicios financieros, excepto seguros y fondos de pensiones 65.- Seguros, reaseguros y fondos de pensiones, excepto Seguridad Social obligatoria 66.- Actividades auxiliares a los servicios financieros y a los seguros 77.- Actividades de alquiler 72.- Investigación y desarrollo	Servicios financieros y a empresas (ServFros&Emp)
84.- Administración Pública y defensa; Seguridad Social obligatoria	AAPP
85.- Educación	Educación
86.- Actividades sanitarias, 87.- Asistencia en establecimientos residenciales, 88.- Actividades de servicios sociales sin alojamiento y 75.- Actividades veterinarias	Sanidad
59.- Actividades cinematográficas, de vídeo y de programas de televisión, grabación de sonido y edición musical, 60.- Actividades de programación y emisión de radio y televisión, 90.- Actividades de creación, artísticas y espectáculos, 91.- Actividades de bibliotecas, archivos, museos y otras actividades culturales, 92.- Actividades de juegos de azar y apuestas, 93.- Actividades deportivas, recreativas y de entretenimiento, 96.- Otros servicios personales. 68.- Actividades inmobiliarias 62.- Programación, consultoría y otras actividades relacionadas con la informática, 63.- Servicios de información 69.- Actividades jurídicas y de contabilidad, 70.- Actividades de las sedes centrales; actividades de consultoría de gestión empresarial, 71.- Servicios técnicos de arquitectura e ingeniería; ensayos y análisis técnicos, 73.- Publicidad y estudios de mercado, 74.- Otras actividades profesionales, científicas y técnicas, 78.- Actividades relacionadas con el empleo, 80.- Actividades de seguridad e investigación, 81.- Servicios a edificios y actividades de jardinería, 82.- Actividades administrativas de oficina y otras actividades auxiliares a las empresas. 94.- Actividades asociativas 97.- Actividades de los hogares como empleadores de personal doméstico 98.- Actividades de los hogares como productores de bienes y servicios para uso propio 99.- Actividades de organizaciones y organismos extraterritoriales	Otros servicios (OtrosServicios)
	(OrganExtraterr)

Fuente: Elaboración propia a partir de la CNAE 2009.

Anexo 2. Mapa de calor de actividades económicas 2019

	Agricultura	IndManufacturera	Suministros	ServFros&Emp	OrganExtraterr	Educación	Sanidad	ComHostTe&Comunic	IndExtractivas	Construcción	AAAPP	OtrosServicios
Agricultura		0,98	1,18	0,20	0,17	0,07	0,25	0,91	1,05	2,85	2,95	1,05
IndManufacturera	1,01		0,80	0,62	0,24	0,36	0,36	1,33	3,63	1,76	0,51	1,17
Suministros	1,31	0,87		0,35		0,41	0,44	1,04		1,53	0,77	1,13
ServFros&Emp	0,23	0,50	0,29			1,33	1,08	1,02		0,49	0,46	1,39
OrganExtraterr	0,31	0,37		0,64		0,68	0,76	1,16		0,21	2,62	1,32
Educación	0,09	0,26	0,34	0,95			2,24	0,92		0,13	1,67	1,64
Sanidad	0,25	0,30	0,32	0,84	0,92	3,06		1,04		0,20	2,03	1,18
ComHostTe&Comunic	0,85	1,21	1,05	0,96	1,34	0,92	1,20		0,88	0,99	0,62	1,92
IndExtractivas	1,01	2,93	1,21	0,84		0,15		0,91		2,40	1,15	0,61
Construcción	2,77	1,66	1,82	0,53	0,52	0,17	0,21	1,02	3,55		0,80	0,84
AAAPP	3,25	0,41	1,09	0,35	2,44	1,60	1,99	0,62	0,19	0,83		0,71
OtrosServicios	1,06	1,34	1,09	1,75	1,36	1,43	1,20	1,90	0,49	0,80	0,72	

Fuente: Elaboración propia a partir de los datos.

Anexo 3. Mapa de calor de ocupaciones 2019

	Ingenieros, licenciados, alta dirección	Ingenieros técnicos, ayudantes titulados	Jefes administrativos y de taller	Ayudantes no titulados	Oficiales administrativos	Subalternos	Auxiliares administrativos	Oficiales 1ª y 2ª	Menores 18	Oficiales 3ª y especialistas	Mayores 18 no cualif
Ingenieros, licenciados, alta dirección		8,4	4,3	1,3	1,8	0,5	1,0	0,4		0,2	0,2
Ingenieros técnicos, ayudantes titulados	9,2		3,2	1,8	1,8	0,8	1,0	0,4		0,3	0,3
Jefes administrativos y de taller	4,3	3,3		1,8	2,3	0,7	1,2	0,8		0,4	0,4
Ayudantes no titulados	1,2	1,9	2,0		1,4	1,2	1,2	0,8	1,8	0,8	0,8
Oficiales administrativos	2,0	2,1	2,3	1,5		1,1	2,0	0,8	0,5	0,6	0,6
Subalternos	0,5	0,9	0,7	1,2	1,1		1,6	1,0	0,3	0,9	1,1
Auxiliares administrativos	1,1	1,1	1,3	1,2	2,2	1,6		1,0	0,3	0,9	1,0
Oficiales 1ª y 2ª	0,4	0,5	0,8	0,9	0,8	1,0	1,0		0,3	1,5	1,8
Menores 18			0,4	0,9	0,6	0,1	0,8	0,4		1,0	2,7
Oficiales 3ª y especialistas	0,2	0,3	0,5	0,8	0,6	0,9	0,9	1,6	1,3		2,0
Mayores 18 no cualif	0,2	0,3	0,3	0,8	0,6	1,1	1,0	1,7	2,5	2,0	

Fuente: Elaboración propia a partir de los datos.

