



RESUMEN AMPLIADO

Título: Metodología y desarrollo de variable(s) económica(s) municipal(es) para España desde 1950 hasta la actualidad

Autores y e-mail de todos ellos: Ignacio Cazcarro^{1,2*}, Guillermo Rodríguez³, Ana Serrano³, Miguel Martín-Retortillo⁴

- 1) ARAID (Aragonese Agency for Research and Development). Agrifood Institute of Aragon (IA2). Department of Economic Analysis, Faculty of Economics and Business Studies, University of Zaragoza, 50005 Zaragoza, Spain. icazcarr@unizar.es
- 2) Basque Centre for Climate Change, Parque Científico de UPV/EHU, 48940 Leioa, Spain.
- 3) Department of Economic Analysis, Faculty of Economics and Business Studies, Universidad de Zaragoza, 50005 Zaragoza, Spain. guillerodlopez96@gmail.com; asergon@unizar.es
- 4) Department of Economics, Faculty of Economics, Business and Tourism, Universidad de Alcalá, 28802 Alcalá de Henares, Spain. miguel.martin@uah.es

*** Autor de correspondencia**

Departamento: Department of Economic Analysis

Universidad: Universidad de Zaragoza

Área Temática: *S08 – La Riqueza de los Municipios – An Inquiry into the Nature and Causes of the Wealth of Municipalities*

Abstract:

El equipo investigador obtuvo en la Convocatoria 2021 una Beca de la Fundación Ramón Areces para el proyecto “[Una aproximación histórica al desarrollo económico de los municipios españoles: la dicotomía rural-urbano durante el siglo XX hasta la actualidad](#)”. El objetivo general del mismo consiste en analizar y evaluar los factores que han incidido en la evolución de la población en los municipios españoles en el siglo XX hasta la actualidad, así como en las implicaciones y consecuencias que de ello se pueden derivar, especialmente ante retos demográficos, como la despoblación de buena parte del territorio, principalmente rural. Ello implica cubrir una cierta carencia (*gap*) en cuanto a bases de datos coherentes y literatura, y objetivos específicos como los siguientes.

- 1) Elaboración y análisis por municipios de España de una base de datos históricos de la población, las superficies de regadío, secano y pastos, así como de variables climáticas y geográficas asociadas durante el siglo XX hasta la actualidad (1900-2020).
- 2) Elaboración y análisis de variable/s que cuantifique/n la actividad económica en términos municipales durante el siglo XX hasta la actualidad.

3) Análisis y evaluación de los factores explicativos de evolución de la población en los diferentes municipios en España, con especial énfasis en el papel del desarrollo del regadío, así como de la instalación de otras industrias y desarrollos en el largo plazo. Propuestas de desarrollo y opciones de futuro, para entender y proyectar los principales fenómenos y propuestas de interés (e.g. el reto demográfico).

El presente trabajo cubre el segundo de los objetivos. Este objetivo, el desarrollo de variables socioeconómicas, en especial las relacionadas con la población activa y el empleo, se suman a la base de datos desarrollada y proporcionada en (Rodríguez et al., 2022). The main motivation for this work is the noticeable lack of historical economic variables on a scale as specific as municipalities. It can be stated that there are no databases or studies that have been capable of systematically representing and analysing the economic evolution of Spanish municipalities, e.g. as regards the active population, employment, etc. (and hence neither combining these variables with other factors, such as biophysical, climatic, etc.)

Claves para la elaboración de las variables Socioeconómicas:

Preguntas que motivan la investigación

- ¿Existen variables económicas relevantes a nivel municipal, con las que poder realizar análisis relevantes sobre la evolución de los mismos en perspectiva histórica?
- Si la pregunta anterior es negativa, ¿existen informaciones con las que aproximarse, con las que unificar información dispersa o en formatos difícilmente utilizables (e.g. anuarios, en pdfs, papel, etc.) que combinar para obtener dichas variables?
- ¿cómo desarrollamos *ad hoc* una metodología con la que hacer el mejor uso posible de la información existente, para obtener variables socioeconómicas desde la mitad del siglo XX a la actualidad?
- ¿qué evolución de la población activa y el empleo resulta de la aplicación de la metodología desarrollada *ad hoc*?
- Para estudiar en el futuro con la base de datos: ¿De qué factores ha podido depender la evolución en la población, en su relación con el empleo, etc., frente a otros factores (e.g. biofísicos)?
- ¿El regadío condiciona la fijación de población y la generación de PIB en el medio rural español?

Fuentes de datos, variables existentes, etc.

- Población: Censos reelaborados por Goerlich et al. (2021) (basados en los Censos del INE)
- GDP: Base de datos ISIMP extrapolado a nivel municipal a través de estadísticas zonales
- Dummies de tipo de sociedad según la estructura de la sociedad: Sociedad latifundista, mixta, o campesina. (No sé muy bien los criterios porque los datos me los pasaron Ana y Nacho)
- Población residente en núcleos de X población en un radio de Y km: PX_Y1_Y2_KM
- Variable de caracterización del municipio en base a la población (Criterio INE):
 - Urbano: >10000 habitantes
 - Intermedio: 2000 – 10000 habitantes

- Rural < 2000

Como se ha sintetiza, al comenzar la investigación obviamente se hizo una revisión de la literatura relevante, así como de las posibles fuentes de datos existentes para acometer la tarea. Mucha de la información disponible ya la habíamos trabajado a escala provincial.

En lo referente a bases de datos municipales, o espaciales, nos encontramos con algunas informaciones que potencialmente podían darnos información. Principalmente los censos, anuarios (e.g. industriales), y algunas otras informaciones dispersas.

Curiosamente encontramos también, de forma paralela a otras informaciones sobre los usos de suelo (cultivos, áreas de ganadería, etc.) de HYDE, información de bases espaciales (en formato raster) sobre Producto Interior Bruto (PIB o GDP). Para los no familiarizados con la terminología de la información espacial, Un archivo ráster es un modelo numérico compuesto por una cuadrícula georreferenciada de píxeles de un tamaño determinado. Cada píxel contiene el valor de la variable representada. El tamaño de píxel de esta base de datos es de 0,5 °. Para obtener los valores a nivel municipal se han utilizado estadísticas zonales. Partiendo de estos archivos ráster, y un archivo vectorial con los polígonos de los recintos municipales del Instituto Geográfico Nacional, y a través de diversas operaciones de análisis espacial, se ha logrado obtener los datos a escala municipal. Una de las operaciones de análisis espacial más utilizadas es, sin duda, las estadísticas zonales. La operación de estadística zonal calcula el valor medio de los píxeles incluidos en cada área provincial ponderando el valor del píxel por el porcentaje de píxel contenido en el área provincial. De esta forma, el valor de estas variables en cada municipio es la suma de los píxeles contenidos en cada municipio. Tras analizar el resultado, y comparándola con otras fuentes de datos, el resultado no se recomienda como primera opción para análisis municipal si se dispone de otros datos. Esto se debe a que la variabilidad espacial que presenta a escala municipal es muy reducida, por lo que no permite captar los contrastes entre municipios próximos. Sí que puede resultar útil para análisis a nivel nacional, autonómico (NUTS-2) o incluso provincial (NUTS-3).

La conclusión que obtenemos es que no es casualidad que no existan estudios ni bases de datos homogéneas y sistemáticas que capturen la evolución económica con cierto detalle regional o provincial (no digamos ya municipal), y con cierto detalle sectorial (al menos distinguiendo la agricultura de la industria y servicios) sea a través del PIB, del valor añadido, el empleo, etc. Como apuntábamos, existe información, pero muy dispersa, que hay que saber entender, combinar, procesar, interpretar, etc. Pero el motivo por el que nos atrevemos a plantear desarrollos en esta línea es porque las capacidades para sistematizar esta información, pese a no evitar determinado trabajo intensivo, son mayores actualmente, utilizando herramientas para escanear, leer, interpretar y procesar datos, combinando fuentes de información con bases de datos y herramientas históricas, estadísticas, geográficas y económicas.

La idea para lograr esta/s variable/s económicas es utilizar diferentes fuentes de datos e instrumentos para desarrollar y analizar una base completa de variable/s económica/s de los municipios en todo el siglo XX y XXI. En particular esto comprende: 1) bases de datos geográficas nacionales e internacionales, históricas y recientes, sobre población, usos de suelo, Producto Interior Bruto (PIB), PIB espacial, instalaciones (minas, utilidades), etc.; 2) Atlas sobre la industrialización¹ (e.g. J Nadal, 2003) y Anuarios

¹ En este tipo de atlas se pone también de manifiesto cómo existe información localizada, pero dispersa en los años y sectores disponibles. En este en concreto la información sobre la ubicación y otros datos de empresas para Cataluña y el País Vasco es muy detallada en determinados periodos.

estadísticos, en los que de forma muy dispersa espacialmente y en el tiempo, y poco sistemática, se ha ido recogiendo información localizada de producciones e instalaciones de diferentes actividades económicas, pero que de forma interesante es más completa en las 3 primeras décadas del siglo XX que posteriormente hasta los años 70; 3) censos de población: donde además de la población, en algunos de ellos se recogen algunas informaciones sobre grupos profesionales y actividad de la población activa por provincias y para determinados municipios, y distinguiendo entre zonas rurales, intermedias y urbanas; 4) bases de datos de empresas (principalmente recientes), como pueda ser (SABI-Amadeus, 2021), donde se da información de entorno a los últimos 25 años de buena parte de las empresas, del empleo, ingresos y otras variables económicas, así como de su año de fundación, además de su localización (frecuentemente exacta, con coordenadas)²; 5) para combinarse con la información anterior, de cara a “proyectar hacia atrás” (especialmente en esas décadas de 2ª mitad del siglo XX con poca información) información reciente de empresas, conociendo los momentos de fundación, las evoluciones provinciales que sirvan como comprobaciones y restricciones (*checks*) de los agregados, utilizar datos (OECD, UNECE, Banco de España, CEOE, CEPYME, etc.) y marcos teóricos sobre “demografía industrial” (*business demography*, e.g. Eurostat e INE); 6) Toda la literatura (e.g. Carreras & Tafunell, 2005) e informaciones regionales y especialmente provinciales, más agregadas, nos deben servir para realizar las comprobaciones, repartos, restricciones, y otras asignaciones.

Más en concreto, en esta elaboración nos centramos en las variables de población activa y empleo desde 1950 hasta la actualidad. La información más relevante desde mediados de siglo XX a nivel sectorial, o por grupos de actividad económica, nos resulta para 1950 el “Cuadro VIII – Población económicamente activa, clasificada según las características de los establecimientos donde trabaja” (INE, 2022a) y para 1960 (INE, 2022b). Análogamente, utilizamos información relevante del mismo tipo para otras décadas, e.g. 1970, de los respectivos censos. Para los municipios mayores de 10.000 habitantes en el caso del censo de 1950 y 1960, y mayores de 50.000 en el caso del censo de 1970, se dispone del detalle para 9 grupos de actividad³ económica (y mayor desagregación en el caso del censo de 1970, donde se desagregan los servicios de “establecimientos financieros, seguros, bienes inmuebles y servicios prestados a las empresas” y “Servicios comunales, sociales y personales”, ambos grupos que estaban agregados para 1950). Todo ello permite un relevante detalle sectorial de la población activa y empleo a nivel municipal.

Además, al disponer de diferentes estructuras para municipios rurales (menores de 2000 habitantes de acuerdo con la clasificación del INE⁴) e intermedios (entre 2000 y 10.000 habitantes de acuerdo con la clasificación del INE), podemos aplicar distribuciones sectoriales de población activa de acuerdo con las mismas para todos los municipios, clasificados en una u otra categoría, para los que no se dispone información exacta (i.e. iguales o menores de 10.000 habitantes en 1950 y 1960, e iguales o menores de 50.000 habitantes en 1970).

² Hasta donde hemos podido observar, serán necesarias varias reclasificaciones para aunar la información de las coordenadas, códigos postales (en ambos casos no presentes para un número elevado de empresas), ciudad/pueblo en el que se instala (donde incluso para la asignación puede ser necesaria la información y metodología para el análisis intra-municipal (Natera Rivas et al., 2021)), para su clasificación en los municipios y sus códigos, con los que principalmente trabajaremos.

³ Las 9 categorías en realidad son 8 categorías concretas más una llamada “actividades mal especificadas”. Las 8 categorías son: “Agricultura, selvicultura, caza y pesca”, “Explotación de minas y canteras”, “Industrias fabriles”, “Construcción”, “Electricidad, gas, agua y servicios sanitarios”, “Comercio, Transporte, almacenaje y comunicaciones” y “Servicios oficiales, públicos y personales”.

⁴ Véase e.g. (Naranjo & Pérez, 2003). Más información y consideraciones sobre delimitaciones de áreas rurales y urbanas puede encontrarse en Reig Martínez et al. (2016) y Lewis Dijkstra & Poelman (2014).

Además, estas estructuras se ven algo modificadas, ligeramente, haciendo uso de la base de datos SABI (2021). Esta base de datos sobre empresas, que es bastante actual, recogiendo con detalle la actividad y empleo de los últimos 30 años nos ofrece información muy relevante distribuida en el espacio para la actualidad (para nuestras 4 cohortes por décadas de 1990-2020 aproximadamente). Para un periodo tan como mediados de siglo XX la información útil es mucho más escasa obviamente. En todo caso existe una información que puede servir como proxy del nivel de actividad en una determinada década. En la base SABI aparece el año de fundación de las empresas que han seguido vigentes hasta la recogida de datos de SABI (si han desaparecido en el período 1990-2020, suelen aparecer como “extintas”). Es decir, la base de datos no captura todas aquellas empresas que se fundaron, pudieron tener vigencia en las diferentes cohortes de 1950, y después desaparecieron antes de tener ningún dato en la base SABI. Pese a esa gran limitación, el dato que sacamos es, para cada cohorte por década (1900, 1910, 1920... hasta 2020) es el número de las empresas existentes en cada municipio, y de los empleados “actuales” de las empresas existentes en cada municipio, por cada sector de actividad ii (más desagregado que i , pero fácilmente agregable a dicho nivel), desde 1850⁵ hasta la cohorte, la década estudiada en cuestión. Este empleo lo denominamos para la década de 1950, $e_{i,m}^{P,1850-1950}$. Pese a lo extraño aparentemente de utilizar información de “empleados actuales”, esta información sólo la usamos de forma relativa en la relación entre cada municipio y la media de la provincia (es decir, que va a re-ponderar la distribución entre actividades según su importancia mayor o menor que la media de la provincia). El empleo provincial por actividad se define como $e_i^{P,1850-1950} = \sum_i e_{i,m}^{P,1850-1950}$. También precisamente por el hecho de que, como hemos dicho, se pierde mucha información con la base SABI, en presencia actividades con valor 0, no queda ponderada por 0, sino que mantiene la ponderación media provincial. Además, el hecho de combinar población activa con empleo, pese a que proporcionamos resultados de ambos tipos, conceptual y numéricamente no se diferencian demasiado en este período (ni habría forma de afinarlos adecuadamente a nivel municipal), puesto que el desempleo se estima en torno a un 1-2% de la población activa.

En todos los pasos anteriores realizamos varias comprobaciones. Por ejemplo, estudiamos cuáles son los niveles porcentajes de población activa e inactiva que resultan, tanto a nivel provincial, como municipal. Estudiamos también la máxima “corrección” o cambio en la distribución por actividad que se produce desde la distribución inicial (para los municipios rurales o intermedios, i.e. que no disponen de datos específicos) con los ajustes con la base de datos SABI, y con el ajuste final que describimos a continuación. Porque finalmente, obviamente de forma agregada se reajustan los resultados para el conjunto de los municipios rurales e intermedios para cumplir con los datos por sector de los agregados provinciales (i.e., se puede comprobar que los totales de población activa por provincia y sector se cumplen).

Para las décadas más recientes, especialmente las 2 más actuales (2010 y 2020), la información actual es utilizada, a partir de los datos de afiliados proporcionados por la Seguridad Social. En concreto de todos los municipios de España para los 87 códigos CNAE. Una primera comprobación que se realiza es la relación (los ratios, etc.) existentes entre estos datos y los datos de SABI (2021) de empleados actuales,

⁵ La elección de 1850 como año inicial es relativamente arbitraria, para poder tener algún inicio anterior a 1900, de forma que no se obtengan prácticamente todos 0s por municipios. A la vez ese inicio parece necesario limitarse, ya que -probablemente por algún error en la recogida de información de la base SABI- existen una decena de casos en los que aparecen años de fundación poco realistas.

detectando los lugares de mayor infrarrepresentación en esta última base de datos. Con esta base de datos que también recoge empleos para unos 25 años atrás (datos que se recogieron hace unos años, con lo que se tendría recogido el empleo -con la representación y salvedades indicadas- hasta finales de los 90), con informaciones de los censos anteriores (e.g. 1991, 2001, 2011), se busca la reconstrucción y convergencia con los datos citados elaborados para el período 1950-1970.

En definitiva, hasta ahora no ha sido posible evaluar cuantitativamente el cambio de población de forma conjunta con una variable económica bien delimitada a nivel municipal, ni con otros factores potencialmente explicativos, como se está haciendo en este trabajo, que lo inicia a través de empleo a nivel sectorial detallado. En el artículo hemos tenido como objetivo explicar paso a paso el proceso seguido para obtener algunos de ellos. El trabajo también pretende proporcionar un ejemplo para otros trabajos metodológicos y empíricos en la Unión Europea a gran escala, p. en la elaboración de variables económicas y geográficas y análisis conjunto con la demografía.

Referencias

- Carreras, A., & Tafunell, X. (2005). *Estadísticas históricas de España, siglos XIX-XX* (2ª). Fundación BBVA.
- Goerlich, F. J., Maudos, J., & Mollá, S. (2021). *Distribución de la población y accesibilidad a los servicios en España*. Instituto Valenciano de Investigaciones Económicas (Ivie) - Fundación Ramón Areces.
- INE. (2022a). *Censo de 1950. Tomo II. Clasificaciones- Clasificaciones por provincias y municipios de más de 10.000 habitantes. Cuadro VIII – Población económicamente activa, clasificada según las características de los establecimientos donde trabaja*. Fondo documental, Instituto Nacional de Estadística (INE). <https://www.ine.es/inebaseweb/71807.do?language=0#>
- INE. (2022b). *Censo de 1960 / Tomo III. Volúmenes provinciales. Icono que indica el tipo de nodo Clasificación de los habitantes por actividad económica. Población económicamente activa*. Fondo documental, Instituto Nacional de Estadística (INE). <https://www.ine.es/inebaseweb/treeNavigation.do?tn=92676&tns=126961#126961>
- Lewis Dijkstra, & Poelman, H. (2014). *A harmonised definition of cities and rural areas: the new degree of urbanisation* (p. 28). Regional Working Paper 2014. European Commission (EC).
- Nadal, J. (2003). *Atlas de la industrialización de España 1750-2000*. Grupo Planeta. https://books.google.es/books?id=A3IQSqV%5C_5wsC
- Naranjo, F. Z., & Pérez, A. A. (2003). *Tendencias demográficas durante el siglo XX en España. La población de España* (pp. 17–76). Instituto Nacional de Estadística (INE).
- Natera Rivas, J. J., Batista Zamora, A. E., & Larrubia Vargas, R. (2021). Sobre el empleo del Nomenclátor para el estudio de la distribución espacial de la población en municipios pequeños frente al seccionado censal. *Empiria. Revista de Metodología de Ciencias Sociales*, 51, 63–88. <https://doi.org/10.5944/empiria.51.2021.30808>
- Reig Martínez, E., Goerlich Gisbert, F. J., & Cantarino Martí, I. (2016). Delimitación de áreas rurales y urbanas a nivel local. In *FBBVA - Informe Técnico* (Issue enero). <http://www.fbbva.es/TLFU/tlfu/esp/publicaciones/informes/fichainforme/index.jsp?codigo=818>
- Rodríguez, G., Serrano, A., Martín-Retortillo, M., & Cazcarro, I. (2022). The

construction of a climatic, demographic and land use database for the Spanish municipalities. *XLVII Reunión de Estudios Regionales (RER), Asociación Española de Ciencia Regional*.

SABI-Amadeus. (2021). *SABI: Sistema de Análisis de Balances Ibéricos: base de datos*. Bureau van Dijk. Computer File.

SABI. (2021). *Sistema de análisis de balances ibéricos: base de datos*. Bureau van Dijk.

Palabras Clave: Variables económicas, población activa, empleo, base de datos histórica, municipios españoles.

Clasificación JEL: J1, N94, O18, Q15, Q56