



# Working Paper

## Àrea de Societat

Mercè Farré  
Aureli Alabert  
Alejandra Cabaña  
Alba Tresens  
Sebastià Sarasa  
Sergio Porcel  
Lara Navarro-Varas  
Fernando Antón-Alonso  
Manel Pons  
Núria Ruiz  
Irene Cruz

### **Una estimació de la distribució de la renda familiar per àrees petites de l'àrea metropolitana de Barcelona**

IERMB Working Paper in Sociology, nº 18.01, September 2018



IERMB Working Paper in Sociology, n° 18.01, September 2018

**Title: Una estimació de la distribució de la renda familiar per àrees petites de l'àrea metropolitana de Barcelona**

Authors: Mercè Farré, Aureli Alabert, Alejandra Cabaña, Alba Tresens, Sebastià Sarasa, Sergio Porcel, Lara Navarro-Varas, Fernando Antón-Alonso, Manel Pons, Núria Ruiz, Irene Cruz

© Mercè Farré, Aureli Alabert, Alejandra Cabaña, Alba Tresens, Sebastià Sarasa, Sergio Porcel, Lara Navarro-Varas, Fernando Antón-Alonso, Manel Pons, Núria Ruiz, Irene Cruz, 2018

Published by:

Institut d'Estudis Regionals i Metropolitans de Barcelona  
Barcelona Institute of Regional and Metropolitan Studies

Parc de Recerca, Mòdul A  
Universitat Autònoma de Barcelona  
08193 Cerdanyola del Vallès, Barcelona, Spain.

<https://iermb.uab.cat>

## Una estimació de la distribució de la renda familiar per àrees petites de l'àrea metropolitana de Barcelona <sup>1</sup>

**Mercè Farré**  
**Aureli Alabert**  
**Alejandra Cabaña**  
**Alba Tresens**

Departament de matemàtiques, Universitat Autònoma de Barcelona  
08193 Cerdanyola del Vallès, Barcelona

**Sebastià Sarasa**

Departament de Ciències Polítiques i Socials, Universitat Pompeu Fabra  
Ramon Trias Fargas 25-27, 08005 Barcelona

**Sergio Porcel**  
**Lara Navarro-Varas**  
**Fernando Antón-Alonso**  
**Manel Pons**  
**Núria Ruiz**  
**Irene Cruz**

Institut d'Estudis Regionals i Metropolitans de Barcelona  
Parc de Recerca, Mòdul A, Universitat Autònoma de Barcelona  
08193 Cerdanyola del Vallès, Barcelona  
Email: sergio.porcel@uab.cat

**Abstract:** La manca d'informació sobre el nivell d'ingressos de la població per àrees territorials petites constitueix un dels principals dèficits d'informació del sistema estadístic espanyol. Aquest aspecte afecta especialment a línies de recerca molt importants dins dels estudis urbans, com per exemple, l'anàlisi de les dinàmiques de segregació residencial socioeconòmica que es produeixen a les grans ciutats. Aquest article presenta

---

<sup>1</sup> Aquest article forma part dels resultats obtinguts al projecte 'La segregació espacial de la pobresa a Catalunya: estructura i dinàmica de la desigualtat social' (Ref.PRO8113), finançat pel programa Recercaixa, convocatòria 2012 i dirigit pel Dr. Sebastià Sarasa.

una proposta metodològica d'estimació de renda en àrees petites aplicada a l'àmbit de l'àrea metropolitana de Barcelona pels anys 1991, 2001 i 2011. La proposta metodològica combina models de regressió logística i ajust proporcional iteratiu (IPF, en anglès). La informació resultant possibilita, per primer cop, l'estudi de les dinàmiques de segregació residencial socioeconòmica a partir del càlcul d'indicadors de segregació estàndards.

**Keywords:** Estimació en àrees petites, Iterative Proportional Fitting (IPF), segregació residencial, desigualtat urbana.

## 1. INTRODUCCIÓ

La manca d'informació sobre el nivell d'ingressos de la població per àrees territorials petites constitueix un dels principals dèficits d'informació del sistema estadístic espanyol. Aquest aspecte afecta especialment determinades línies de recerca molt importants dins dels estudis urbans, com per exemple, l'anàlisi de les dinàmiques de segregació residencial socioeconòmica que es produeixen a les grans ciutats. S'ha de dir que aquest problema no és exclusiu de les fonts estadístiques espanyoles, sinó que es tracta d'un problema força estès entre la resta de països europeus, llevat d'alguns països del nord que sí que disposen d'aquesta informació a partir de registres administratius. De fet, a Espanya s'estan produint darrerament avenços importants en aquesta línia. Cada vegada s'està tenint més accés a l'explotació estadística de registres fiscals a partir dels quals es pot obtenir informació fidedigna dels ingressos del conjunt de la població, però encara queda molt camí per recórrer<sup>2</sup>. Mentrestant, la comunitat científica ha desenvolupat diferents estratègies per superar aquesta manca d'informació, entre les quals destaca la utilització de *proxys* o les estimacions estadístiques per àrees petites.

En aquest article, precisament, es presenta una d'aquestes estratègies d'estimació de la distribució de les rendes de la població resident en àrees territorials petites. Es tracta d'un procediment que combina diverses metodologies, i que s'ha aplicat a l'àmbit de l'àrea metropolitana de Barcelona, basat en la combinació de dades d'enquesta i de dades censals, el qual està pensat i dissenyat per permetre l'anàlisi de la segregació residencial socioeconòmica. El treball s'emmarca en el projecte de recerca "La segregació espacial de la pobresa a Catalunya: estructura i dinàmica de la

---

<sup>2</sup> En el marc de l'Urban-Audit, l'INE està utilitzant informació de l'Agència Tributària per informar de la mitjana i de la distribució de la renda a escala municipal, però encara no es publica informació d'aquest tipus per àrees petites. Des de 2013, l'INE utilitza també aquesta informació fiscal en la construcció de la variable renda de l'*Encuesta de Condiciones de Vida* (ECV).

desigualtat social”<sup>3</sup>. L’objectiu general d’aquesta investigació és conèixer amb més profunditat la relació que s’estableix entre la desigualtat social i les dinàmiques d’estructuració socioresidencial a les principals metròpolis de Catalunya, per la qual cosa comptar amb la informació d’ingressos de la població per àrees petites ha esdevingut, des del primer moment, un aspecte cabdal del projecte.

El document s’estructura de la següent manera. En primer lloc, es fa una revisió dels exemples d’estimació de rendes que s’han realitzat fins ara, tant en el context espanyol com a nivell internacional. En segon lloc, s’introdueix una explicació de les principals fonts de dades que s’han fet servir per realitzar les estimacions. I, finalment, s’explica amb detall el mètode que s’ha implementat, tot avaluant els resultats obtinguts.

## 2. PRECEDENTS, LIMITACIONS I ALTERNATIVES

Evidentment, el mètode que es presenta en aquest article no és el primer intent d’estimació de la renda per àrees petites al context espanyol, sinó que ja hi han hagut d’altres exercicis, tot i que no del tot plausibles per l’anàlisi de la segregació residencial socioeconòmica. En els darrers anys, per exemple, tant a la comunitat autònoma de Madrid (Moreno, 2003) com al municipi de Barcelona (Calvo, 2007), s’han realitzat estimacions a partir de macromagnituds econòmiques. En els resultats d’aquestes propostes, s’han aconseguit estimacions de la renda mitjana *per càpita*, per cadascuna de les unitats territorials. Tanmateix, aquest tipus d’informació no és suficient per calcular els indicadors estàndards de segregació residencial, pels quals es precisa del nombre absolut de població que disposa d’un nivell de renda determinat. Observant aquesta mancança, Madariaga *et al.* (2012) van realitzar una estimació de la distribució dels salaris de la població ocupada a partir de la *Encuesta de Estructuras Salariales*, del *Padró municipal d’habitants* i del *Cens de Població i Habitatges* per l’àrea metropolitana de Barcelona. Els autors aconsegueixen disposar d’una bona desagregació territorial de la informació salarial que els permet analitzar la segregació residencial, però només disposen dels ingressos de la població assalariada. S’obvien altres tipus de fonts d’ingressos, així com la població aturada i inactiva, per la qual cosa els resultats només es poden considerar com una aproximació parcial.

A l’esfera internacional, diversos autors han proposat i utilitzat d’altres metodologies per abordar l’estimació de la renda per àrees petites, les quals s’ajusten millor als requisits de la informació que precisen computar els

---

<sup>3</sup> Projecte finançat pel Programa RecerCaixa 2012 (ref. PRO8113).

indicadors de segregació residencial més utilitzats. Entre aquests mètodes es troben la modelització multinivell (Elbers i van der Weide, 2014) o les metodologies IPF (de les seves sigles en anglès, Iterative Proportional Fitting) (Simpson i Tranmer, 2005; Lomax i Norman, 2016). Aquests són, doncs, les principals metodologies que s'han pres de referència a l'hora de dissenyar el mètode que es presenta en aquest article, tot per tal de superar les limitacions de les experiències que s'han citat anteriorment en relació a l'estudi de la segregació residencial.

La metodologia que es presenta integra informació de l'*Enquesta de condicions de vida i hàbits de la població* (ECVHP) i informació censal. Es parteix de la premissa que a totes dues fonts es disposa de variables coincidents capaces de predir el nivells de renda de la població (categoria professional, nivell d'estudis, relació amb l'activitat dels membres de la llar, lloc de naixement, superfície de l'habitatge, etc.) i que la informació de renda només es disposa a l'ECVHP. L'objectiu final d'aquest exercici d'estimació és obtenir la proporció de població resident en cadascuna de les seccions censals que integren l'àmbit territorial objecte d'estudi —en aquest cas, l'àrea metropolitana de Barcelona—, segons diferents nivells de renda. Concretament, s'ha decidit treballar amb tres estrats de renda definits de forma relativa: *població amb rendes altes* (>150% de la mediana de la distribució de renda metropolitana), *població amb rendes intermèdies* ( $\geq 50\%$  i  $\leq 150\%$  de la mediana de la distribució de renda metropolitana) i *població amb rendes baixes* (<50% de la mediana de la distribució de la renda metropolitana). El mètode d'estimació, però, només s'aplica sobre la població amb rendes altes i la població amb rendes baixes, ja que són les que millor es prediuen. La proporció de població amb rendes intermèdies, en canvi, es dedueix posteriorment restant les anteriors, un cop aquestes han sigut estimades.

### 3. FONTS DE DADES

Com s'ha esmentat anteriorment, el procediment d'estimació de rendes per àrees petites que es presenta es basa en dos tipus de dades. Unes procedents l'*Enquesta de condicions de vida i hàbits de la població* (ECVHP) i unes altres procedents del *Cens de Població i Habitatges*.

L'ECVHP<sup>4</sup> és una enquesta de periodicitat quinquennal impulsada, dissenyada i executada per l'Institut d'Estudis Regionals i Metropolitans de

---

<sup>4</sup> Per a més informació sobre aquesta enquesta vegeu el següent enllaç: <https://iermb.uab.cat/ca/enquestes/cohesio-social-i-urbana/>

Barcelona (IERMB)<sup>5</sup>. Aquesta operació estadística es realitza sobre l'àmbit metropolità de Barcelona des de l'any 1985, encara que en les últimes edicions —concretament les del 2006 i 2011— l'enquesta va abastar el conjunt del territori català, al mateix temps que va passar a ser estadística oficial. No obstant això, tot i les successives ampliacions del territori objecte d'estudi, sempre s'ha mantingut un nivell de mostra estadísticament significatiu per als diferents àmbits locals de Barcelona (municipi, àrea, regió i província). Es tracta, per tant, d'una de les sèries històriques d'informació sobre estructura social i condicions de vida a escala metropolitana més importants d'Europa.

Respecte els continguts, l'ECVHP ofereix una àmplia informació sobre l'estructura de les llars, l'activitat i l'ocupació, la formació, els ingressos i la renda, la salut i l'autonomia personal, les relacions personals i la xarxa social, la habitatge i les condicions d'habitabilitat, la mobilitat residencial i l'ús i la percepció del territori. Es tracta, per tant, d'una enquesta dissenyada específicament per a l'estudi de l'estratificació social i de la cohesió social, amb una clara vocació socioterritorial, que es reflecteix tant en algunes de les temàtiques que tracta com en el seu disseny mostral. A més, un altre aspecte important d'aquesta enquesta que afavoreix l'anàlisi socioterritorial és que disposa de la variable sobre la secció censal, la qual cosa permet una major versatilitat a l'hora de construir variables territorials. Precisament, disposar de la variable de secció censal ha facilitat el desenvolupament d'aquest mètode d'estimació de la distribució de la renda en àrees petites.

L'última edició, la del 2011, va suposar un punt d'inflexió metodològic i de continguts d'aquesta enquesta, en la mesura que es van adoptar una sèrie de criteris que van permetre harmonitzar-la amb l'enquesta de condicions de vida europea, la *Statistics on Income and Living Conditions* (EU-SILC). D'una banda, es va dotar a l'ECVHP de representativitat individu-llar. En les edicions anteriors la mostra només oferia informació sobre individus majors de 16 anys. A més, el qüestionari passa a recollir la informació de la renda de la mateixa manera que ho fa l'*Encuesta de Condiciones de Vida* (ECV) que realitza l'Institut Nacional d'Estadística (INE) i que forma part de l'EU-SILC, així com altres informacions que permeten calcular indicadors sobre desigualtat social, pobresa i exclusió de forma estandarditzada a nivell europeu. Tota aquesta transformació es fa respectant al màxim la sèrie històrica de dades i reforçant també el caràcter socioterritorial de part dels continguts.

---

<sup>5</sup> Històricament aquesta operació estadística ha comptat també amb el suport d'administracions com l'Àrea Metropolitana de Barcelona, la Diputació de Barcelona i, en les darreres edicions, també per part de l'Institut d'Estadística de Catalunya (Idescat).

Per la implementació d'aquest mètode d'estimació s'ha utilitzat la mostra referida al conjunt dels 36 municipis que conformen l'Àrea Metropolitana de Barcelona de les edicions 1990, 2000 i 2011. Les característiques tècniques de la mostra d'individus de les diferents edicions de l'ECVHP per a aquest àmbit territorial són les següents:

*Taula 1. Característiques de la mostra de la Enquesta de condicions de vida i hàbits de la població, 1990-2000-2011*

	1990	2000	2011 <sup>(1)</sup>
Àmbit territorial	Àrea metropolitana de Barcelona	Àrea metropolitana de Barcelona	Àrea metropolitana de Barcelona
Municipis	36	36	36
Població total	3.048.479	2.921.563	3.226.944
Univers estadístic	Individus de 18 anys i més	Individus de 18 anys i més	Individus de 16 anys i més
Mida de la mostra	3.579	3.789	3.701
Error mostral	±1,6%	±1,6%	±1,6%

<sup>(1)</sup> En aquesta edició es disposa també de mostra representativa de llars.  
Font: elaboració pròpia.

Per altra banda, també s'han utilitzat les dades de diferents edicions del *Cens de Població i Habitatges*, desenvolupat per l'Institut Nacional d'Estadística (INE) cada 10 anys. En concret, s'han utilitzat els censos de 1991, de 2001 i de 2011. Els censos proveeixen informació del total de la població (sexe, edat, lloc de naixement, nivell d'estudis, relació amb l'activitat, categoria ocupacional, etc.), de les llars (composició demogràfica), els edificis (estat, nombre de plantes, etc.) i els habitatges (superfície, règim de tinença, etc.), és per això que es considera la font d'informació de referència a petita escala. La unitat territorial més petita sobre la qual ofereixen informació dels censos és la secció censal<sup>6</sup>, de manera que s'ha treballat a aquesta escala a través dels percentatges totals marginals de cadascuna de les categories de les variables utilitzades. De manera que si una variable censal té 3 categories, com per exemple, el nivell d'estudis, la informació que ha proporcionat l'INE per a cada secció censal són els percentatges (respecte la població total de cada secció censal) de població amb estudis primaris o inferiors, amb estudis secundaris i amb estudis superiors. Aquesta informació ha estat anomenada en el marc d'aquest exercici com a *categories-variable*.

<sup>6</sup> La secció censal és una unitat territorial infra-municipal que està delimitada preferentment per límits fàcilment identificables (accidents naturals, vials, etc.) i que aglutina una població d'entre 1.000 i 2.000 persones.



S'ha de dir, però, que els canvis metodològics que va introduir l'INE en l'última edició del cens han empitjorat el detall territorial de la informació censal. El 2011, el cens va passar de ser un recompte universal a ser una macro-enquesta, de manera que no es cobreix el total de la població, sinó una mostra d'aproximadament el 13% del total de la població. Posteriorment, la informació recollida és calibrada amb un fitxer precensal, construït a partir de tota aquella informació útil procedent d'altres registres administratius disponibles i que té el Padró d'habitants com a element bàsic de la seva estructura. Aquesta metodologia censal ha permès obtenir una bona informació sociodemogràfica en termes agregats amb menys recursos econòmics, però en canvi no ha permès oferir informació estadísticament significativa en la totalitat de seccions censals per a totes les categories de les variables del cens, la qual cosa ha suposat un handicap important per a aquesta investigació. Per superar aquest obstacle i evitar perdre detall territorial de la informació s'ha optat per fer un tractament de correcció de la manca d'informació a escala de secció censal. Els detalls d'aquest tractament s'expliquen més endavant.

En definitiva, pel que fa a la informació original a partir de la qual es realitza aquest mètode d'estimació dels nivells de renda per àrees petites es poden distingir dos tipus de dades:

- a) **Dades d'enquesta (tipus 1):** Dades desagregades a nivell individual, representant el conjunt d'un àmbit territorial. Inclouen informació sociodemogràfica detallada a títol individual. Constitueixen una mostra representativa a una escala territorial gran. A partir de l'enquesta es disposa de les variables resposta 'Y' (indicadors binaris dels estrats de renda) per a cada individu i de diverses variables explicatives 'X' que classifiquen els individus segons diferents atributs, com puguin ser la situació d'activitat, el nivell d'estudis, el sexe, o el grup d'edat.
- b) **Dades censals (tipus 2):** Dades agregades a una escala territorial petita (seccions censals), que representen les unitats d'informació. La informació dels *Censos de Població i Habitatges* s'expressen com a freqüències absolutes (recomptes) i relatives (percentatges) d'individus per cada categoria de les variables explicatives 'X'. Es disposa de les distribucions marginals de les variables explicatives, però no les distribucions conjuntes.

La qüestió fonamental és que els censos (tipus 2) no inclouen informació sobre renda familiar. Aquesta informació existeix de manera desagregada a

l'ECVHP (tipus 1), però amb aquesta font de dades no es pot fer un càlcul directe dels valors a escala de secció censal, ja que la mostra de l'enquesta no és estadísticament representativa a aquest nivell territorial. L'objectiu per tant és combinar les dues informacions de manera que es pugui obtenir una estimació de la població segons els seu nivell d'ingressos per cada secció censal.

*Taula 2. Tipus de dades, nivell d'agregació i tipus d'informacions*

	<b>Dades tipus 1</b>	<b>Dades tipus 2</b>
Agregació/unitat	Desagregades/individus	Agregades/seccions censals
Territori	Grans àmbits territorials	Àrees petites
Disponibilitat de dades	Objectiu i predictors	Predictors

Font: elaboració pròpia.

#### **4. DESCRIPCIÓ DEL PROCEDIMENT D'ESTIMACIÓ DE RENDA PER ÀREES PETITES A L'ÀREA METROPOLITANA DE BARCELONA**

El mètode d'estimació dels nivells de renda de la població resident a cada secció censal de l'àrea metropolitana de Barcelona es desplega en 6 etapes consecutives, que queden reflectides en el gràfic 3.1. i que s'enumeren a continuació:

1. S'ha aplicat un mètode d'estimació i imputació de valors mancants a les dades del cens 2011 per aquelles *categories-variable* on no hi ha dades.
2. A partir de les dades censals completes i sense dades mancants, s'ha fet una anàlisi de clústers de les seccions per obtenir sub-àmbits segons una caracterització econòmica i social que permet afinar millor les estimacions dels indicadors finals.
3. S'han aplicat models de regressió logística a les dades de l'ECVHP per a determinar els factors explicatius (X) que conjuntament millor expliquen les variables a estimar.
4. S'ha aplicat el mètode iteratiu d'ajust proporcional, IIPF, sobre les dades del cens per a obtenir la llei conjunta dels factors explicatius a totes les seccions censals.
5. Donada la llei conjunta estimada a cada secció, i mitjançant les dades de l'ECVHP, s'ha obtingut finalment l'estimació de màxima versemblança o projecció de la distribució de l'indicador a cada secció censal.
6. Finalment s'ha fet una estimació dels marges d'error.

Seguidament s'explica cada etapa en major detall.

#### 4.1. Mètode d'imputació de dades mancants al Cens de Població i Habitatges, 2011

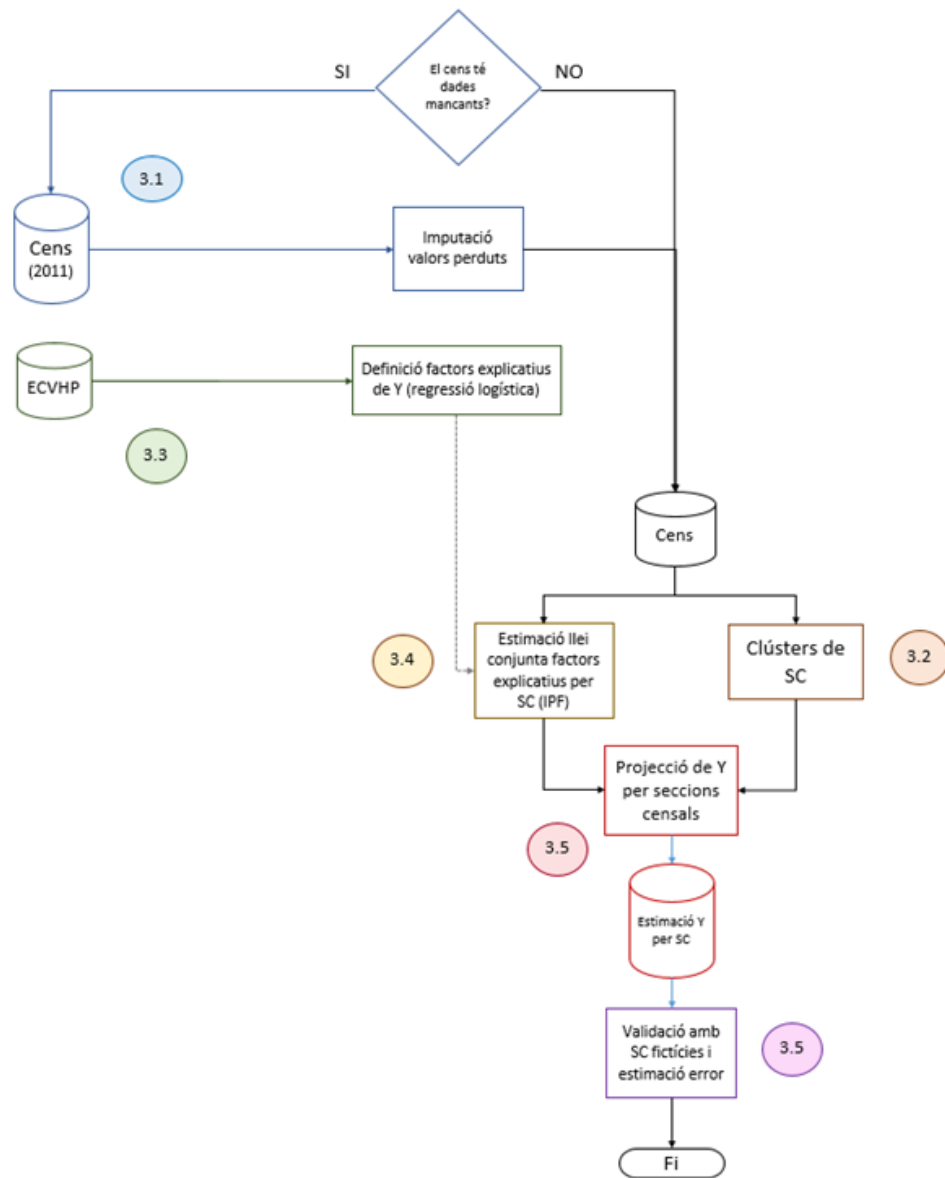
El procediment d'imputació de valors a les dades mancants del cens es pot resumir en els següents quatre passos:

1. Quantificar el nombre de dades mancants que hi ha en cada *categoria-variable* censal, és a dir, per a cada secció censal i variable censal.
2. Quantificar el nombre de *categories-variable* que no té cap dada.
3. Com es disposa de les freqüències relatives dels totals marginals, la suma total de les diferents *categories-variable* ha de ser 100 per cada secció censal. Així doncs, en els casos que només hi ha 1 categoria faltant, es pot fer una imputació directa mitjançant un simple càlcul, on la categoria faltant ( $C_j$ ) equival a 100 menys la suma de la resta de categories:

4. 
$$C_j = 100 - \sum_{i \neq j} C_i \quad (1)$$

5. En els casos que tenen més d'una categoria faltant es realitza la imputació mitjançant el mètode de *regressió seqüencial multivariant*.

Figura 1. Diagrama de procés del tractament de dades per l'estimació de rendes per àrees petites.



Font: elaboració pròpia.

El mètode de *regressió seqüencial multivariant* s'ha implementat a través de la macro *IVEware*, desenvolupada al Survey Research Center de la Universitat de Michigan (Raghunathan, Lepwoski, Van Hoewyk, i Solenberger, 2000; Raghunathan, Solenberger, i Van Hoewyk, 2002), la qual executa un procediment d'imputació general multivariable que pot tractar dades amb una estructura complexa. Aquest procediment crea imputacions de dades mancants a través d'una sèrie de regressions. El model de regressió depèn de la mètrica de la variable imputada (si és categòrica o numèrica).

Les variables explicatives o predictives que s'han utilitzat han estat les categories d'altres paràmetres de la base de dades precensal, la majoria d'elles procedents del Padró d'habitants. Aquestes variables explicatives s'han seleccionat sota la condició de no tenir cap valor perdut, és a dir, que eren variables completes, ja sigui perquè tenien la informació completa d'entrada, ja sigui perquè en el pre-processat se'ls ha imputat els valors mancants.

En el moment de la imputació s'imposa una restricció construïda a partir del percentil 10 i 90 de la distribució de la variable a imputar, per eliminar els valors extrems. L'aleatorietat en l'estimació es fa de dues formes: A) el terme residual és aleatori, i B) s'integra una pertorbació en els coeficients de regressió estimats de forma que reculli la variabilitat dels coeficients<sup>7</sup>.

La seqüència d'imputacions dels valors buits es realitza cíclicament fins arribar a un criteri de parada (un nombre màxim d'iteracions) especificat per l'equip investigador, actualitzant en cada iteració les imputacions prèvies per recollir millor l'estructura de correlacions del conjunt de les variables<sup>8</sup>.

#### **4.2. Anàlisi de clúster de les seccions censals**

Les seccions censals s'han agrupat en clústers d'acord amb les seves categories-variable amb l'objectiu d'afinar les estimacions posteriors. Cada secció s'ubica en un sub-àmbit més afí pel que fa a la situació econòmica i social, de manera que dins dels clústers es produeix un millor encaix de la llei conjunta, que s'explica a la secció 4.4.

Els clústers s'han definit a partir de totes les variables de les seccions censals, tant les variables explicatives de Y, descrites a l'apartat anterior, com la resta de variables que no s'havien utilitzat en aquell pas. Això ha permès construir

---

<sup>7</sup> Veieu Gelman et al. (1995) per més detalls sobre aquest aspecte.

<sup>8</sup> A la Taula A.1. de l'annex es poden observar el percentatge de seccions censals en les que s'ha imputat informació mancant segons aquest mètode.

agrupacions de seccions censals similars en una diversitat d'aspectes socials i demogràfics.

El mètode d'agrupació emparat ha estat el clúster jeràrquic sobre les observacions (seccions censals), amb el mètode de conglomeració de Ward i la mesura de distàncies euclidianes. La solució final s'ha establert en tres grups.

A posteriori s'han avaluat els resultats per clústers i sense clústers, i els primers donen millors estimacions. Per a cada àrea petita, s'ha considerat que el clúster de pertinença és l'àrea global de referència.

#### 4.3. Definició dels factors explicatius de Y

Mitjançant una regressió logística, s'han obtingut models explicatius de les variables binàries descrites a l'apartat anterior respecte algunes de les variables explicatives com a regressores. Els models ajustats són del tipus següent:

$$\log \frac{P_i}{1 - P_i} = \beta X_i \quad (2)$$

on  $i$  varia per a cada individu,  $P_i$  expressa la probabilitat que l'individu pertanyi a la categoria binària codificada amb el valor 1 ("èxit") i  $X_i$  és el vector de regressores (*dummies*) de les diverses categories de classificació de les variables explicatives descrites a les variables inicials de l'individu. Els models logístics com (2) ajusten el log-odds a partir d'una combinació lineal de les variables regressores.

L'ECVHP permet determinar els factors explicatius més rellevants o significatius per als nivells de renda mitjançant una regressió logística. S'estimen els models logístics per als individus que tenen uns ingressos per sota del 50% de la mediana i els que tenen més del 150% de la mediana. Així, en aquest procés només s'estimaran aquests dos grups d'individus i no s'estimarà el del tram central (entre 50% i el 150% de la mediana), ja que aquest es pot obtenir directament a partir dels altres dos trams. S'ha tingut en compte que els individus no són tots independents, atès que alguns d'ells comparteixen la mateixa llar.

S'han buscat models amb un nombre reduït de variables (2 a cada model) per diversos motius. El motiu principal és que a les seccions censals només es disposava de distribucions marginals de cada variable per separat, per tant, si s'utilitzen tres variables, els errors d'estimació podrien ser molt superiors

perquè primer s'haurien d'estimar les conjuntes de les tres possibles parelles per acabar estimant la conjunta global de les 3 variables. Un altre motiu és que la mostra repartida en moltes tipologies ( $4 \times 3 \times 3$ ) = 36 seria insuficient, fins i tot en l'àmbit global on s'ha realitzat l'enquesta.

#### 4.4. Llei conjunta dels factors explicatius

El mètode principal per a realitzar el càlcul d'estimació de la renda a nivell de secció censal ha sigut el mètode iteratiu d'ajust (IPF) introduït i estudiat, des dels anys 40 (Deming i Stephan, 1940; Stephan (1942), Fienberg (1968, 1970). L'ajust iteratiu proporcional i les estimacions posteriors dels indicadors de pobresa s'expliquen de manera esquemàtica a l'apartat 4.5. El *Mathematical Consulting Service* (UAB) ha col·laborat en definir el procés metodològic i ha desenvolupat el codi amb R (2017), que permet obtenir aquestes estimacions de forma massiva per a tot el conjunt de seccions censals. Una aplicació del mètode IPF en un context similar al d'aquest projecte es troba a Simpson i Tranmer (2005).

Per a cada àrea petita (secció censal), el mètode IPF estima o ajusta la distribució conjunta de dues o més variables categòriques respectant les lleis marginals d'aquestes variables en l'àrea petita i ho fa projectant la llei conjunta coneguda d'aquestes variables en un àmbit global que conté a l'àrea petita (àmbit metropolitana, Catalunya o clústers grans). La projecció de la llei conjunta conserva la mateixa estructura d'interacció que hi ha entre les variables en l'àrea global (Fienberg, 1970).

Lewis (1959) i Brown (1959) proven que, partint d'una llei conjunta uniforme a l'àrea global, i d'unes marginals per àrees petites, la llei conjunta obtinguda aplicant IPF afegeix la informació addicional mínima (màxima entropia) respectant les marginals de l'àrea petita.

IPF és un procediment iteratiu que s'inicia amb una distribució conjunta proporcional a la de l'àrea global, que no respecta les marginals i, en els passos successius, ajusta primer una marginal i després l'altra fins arribar a la convergència. Convé remarcar que IPF no és un procediment inferencial clàssic, sinó un mètode determinista basat en uns principis d'aproximació.

S'ha implementat el mètode IPF seguint el procediment desenvolupat a Tresens (2014). Concretament, per a cada secció censal, aplicant IPF s'estimen les freqüències absolutes  $\{n_{ij}\}_{ij}$  és a dir, els recomptes de totes les caselles de la taula de contingència de les dues variables explicatives

seleccionades al model logístic, les quals determinen la seva distribució conjunta:

$$P\{X_1 = x_1, X_2 = x_j\} \approx \frac{n_{ij}}{n} \quad (3)$$

Recordem que per obtenir aquestes estimacions  $\{n_{ij}\}_{ij}$  s'ha aplicat la IPF a les freqüències absolutes  $\{N_{ij}\}_{ij}$  de l'àrea global (enquesta) i les freqüències marginals l'àrea petita (cens).

#### 4.5. Projectió de dades

Per a l'estimació final, es tornaran a utilitzar les dades de l'àrea global (enquesta) per estimar les probabilitats de l'indicador d'interès, condicionades per la combinació de valors dels factors:

$$P\{Y = 1 | X_1 = x_1, X_2 = x_j\} = p_{ij} \quad (4)$$

on  $Y = 1$  indica l'èxit de la variable binària (indicador de nivell de renda). Més explícitament, si suposem el cas per a les rendes superiors al 150% de la mediana, es denota  $O_{ij}$  el nombre observat d'individus de la mostra global que tenen aquests nivells superiors de renda, per a cada tipologia definida per (*ocup*)  $X_1 = x_i$ , i (*educ*)  $X_2 = x_j$ , i es denota  $N_{ij}$  el nombre total d'individus d'aquesta tipologia a la mostra global. Aleshores, s'estimen les probabilitats condicionades:

$$\hat{p}_{ij} = \frac{O_{ij}}{N_{ij}} \quad (5)$$

Finalment, aplicant la fórmula de la probabilitat total s'obtindrà l'estimació de la probabilitat d'aquest indicador de renda:

$$P\{Y = 1\} = \sum_{ij} P\{Y = 1 | X_1 = x_i, X_2 = x_j\} P\{X_1 = x_i, X_2 = x_j\} \quad (6)$$

De manera equivalent, s'estimen les freqüències absolutes de renda a les àrees petites:



$$\begin{aligned} \#\{\text{renda} > 150\% \mid M_d = 1\} &\approx \frac{O_{11}}{N_{11}} \cdot n_{11} + \frac{O_{21}}{N_{21}} \cdot n_{21} + \dots + \frac{O_{44}}{N_{44}} \cdot n_{44} \quad (7) \\ &\approx \sum_{i,j} n_{ij} \hat{p}_{ij} \end{aligned}$$

on  $\hat{p}_{ij}$  són les estimacions de les probabilitats condicionades i  $n_{ij}$  el nombre d'individus de la tipologia definida per  $X_1$  i  $X_2$  a l'àrea petita.

La justificació d'aquest pas del procediment es basa en que les freqüències relatives són els estimadors de màxima versemblança de les probabilitats. Condicionant a cada casella, pot ser que resulti una mostra petita i per això s'utilitza l'interval exacte de Fisher quan es calcula el marge d'error.

#### 4.6. Marges d'error

Per als marges d'error s'han fet aproximacions teòriques i s'han aplicat tècniques de simulació (generació de seccions hipotètiques) i de remostreig (bootstrap).

Per als resultats teòrics, cal tenir en compte les consideracions següents:

- A cada secció censal, els valors  $n_{ij}$  obtinguts aplicant IPF són estimacions del nombre d'individus.
- Les probabilitats  $p_{ij}$  per una categoria-variable concreta s'han estimat a partir de les freqüències observades a la mostra, però corresponen a individus no-independents degut a que alguns pertanyen a la mateixa llar.
- El nombre esperat, obtingut aplicant

$$ne := \sum_{i,j} n_{ij} p_{ij}, \quad (8)$$

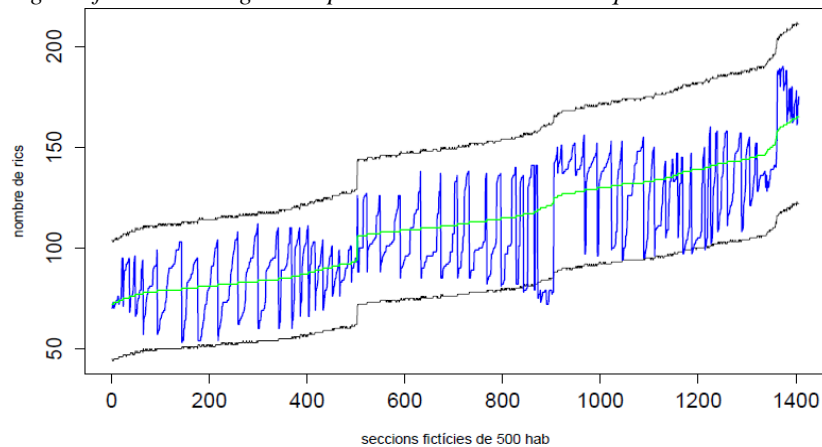
es pot entendre com el valor esperat d'una suma de variables Bernoulli ( $Y \in \{0, 1\}$ ) que no son ni independents ni idènticament distribuïdes, perquè a cada casella els individus tenen una probabilitats  $p_{ij}$  diferent i hi ha individus correlacionats pel factor llar.

- La distribució Poisson-binomial modela la suma de variables de Bernoulli independents però no idènticament distribuïdes (probabilitats  $p_{ij}$  diferents) i és aproximable (sota certes condicions sobre  $n$  i les probabilitats  $p_{ij}$ ) per una distribució de Poisson amb paràmetre  $\lambda = ne$  (veure Chen and Liu, 1997).

- Si les condicions no se satisfan perquè les probabilitats no són prou petites i els individus de la mostra no són independents, aproximant per la distribució Poisson podem establir una aproximació de l'interval de confiança que després hem contrastat mitjançant validació creuada.
- Per fer la validació, s'han creat seccions fictícies mitjançant simulació. Imposant un nivell de confiança del 99% a la distribució de Poisson (quantils 0.005 i 0.995), com a fita superior del nivell de confiança de l'estimació final. Cal tenir en compte que en l'estimació final, les probabilitats  $\hat{p}_{ij}$  també tenen marge d'error i les estimacions per l'IPF de les freqüències també. Per això, del 99% de confiança prefixat es té un nivell de confiança final desconegut però menor (segons hem vist amb les simulacions efectuades, depenent de la mida de la secció el nivell de confiança final se situava al voltant del 95%). El marge d'error (precisió) de les bandes de confiança està al voltant d'un 10% en les simulacions efectuades.

La imatge de la Figura 2. mostra que la majoria de seccions fictícies tenen valors observats dins de la banda de confiança. La banda de confiança té uns marges d'error que poden superar  $\pm 50$  en seccions de 500 habitants, és a dir, ser superiors 10% en les estimacions finals. Les línies de color blau recorren els valors reals, la línia verda són les estimacions i les línies negres les bandes de confiança.

*Figura 2. Bandes de confiança per a l'estimació final de la categoria-variable "nivell de renda per sobre del 150% de la mediana" a un seguit de regions fictícies obtingudes a partir de la mostra de l'enquesta*

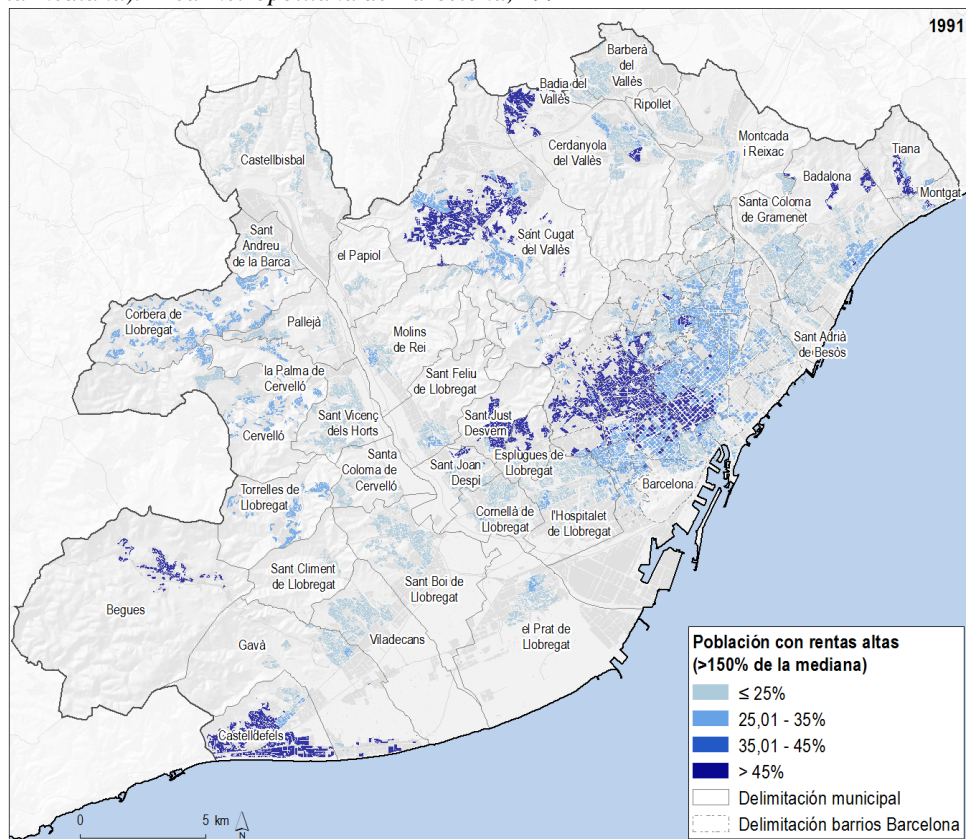


Font: elaboració pròpia.

## 5. RESULTATS

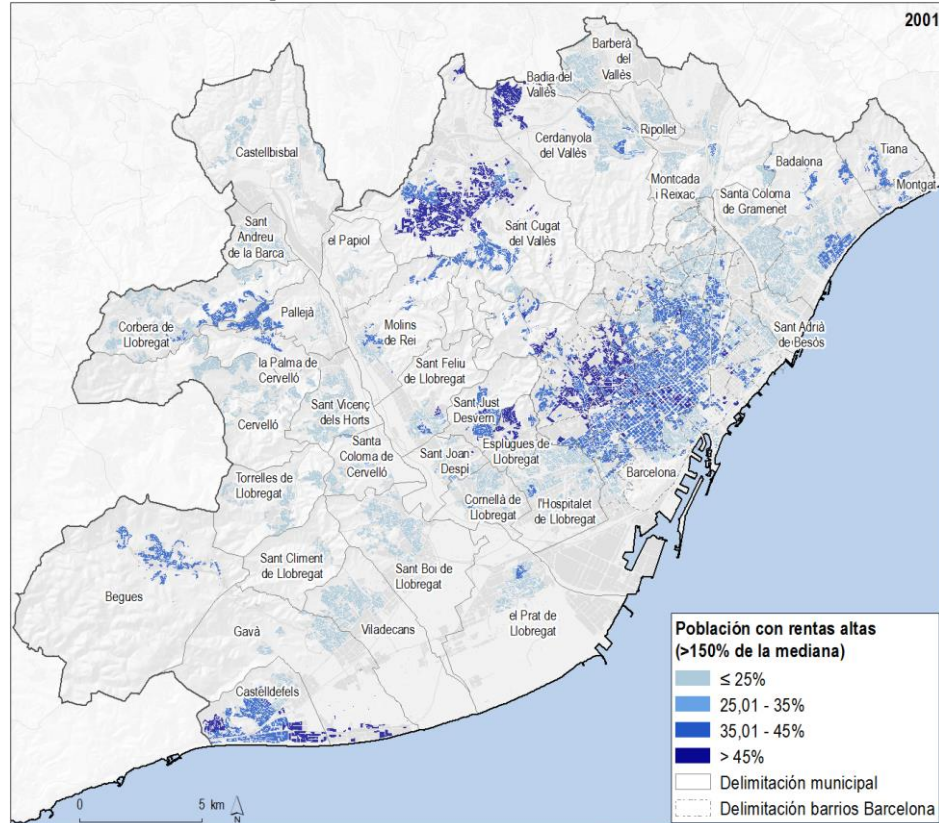
En els següents mapes es recull la cartografia dels principals resultats d'aquest mètode d'estimació de renda per àrees petites. Es tracta dels resultats descriptius de la distribució territorial de la població amb rendes altes (>150% de la mediana) i de la població amb rendes baixes (<50% de la mediana) a l'àmbit de l'àrea metropolitana de Barcelona pels anys 1991, 2001 i 2011.

*Figura 3. Distribució territorial de la població amb rendes altes (>150% de la mediana). Àrea metropolitana de Barcelona, 1991*



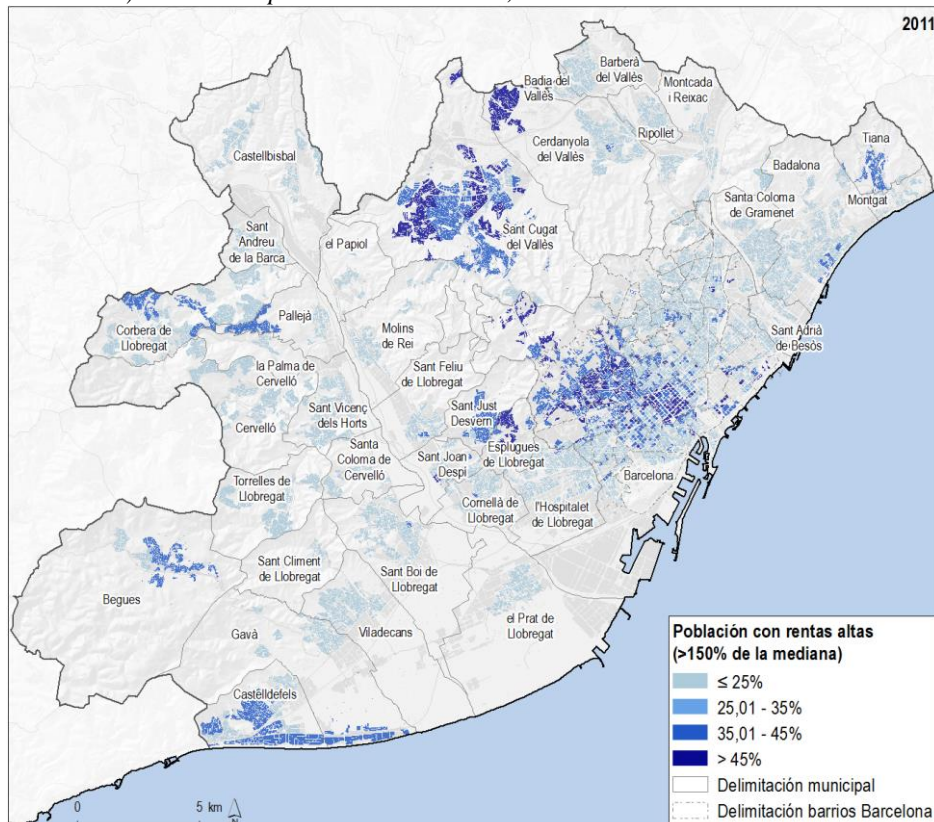
Font: IERMB i CRM-UAB, grups socials segons nivell de renda estimats a escala de secció censal a partir del Cens de Població i Habitatges, 1991 (INE) i l'Enquesta de condicions de vida i hàbits de la població, 1990 (IERMB).

Figura 4. Distribució territorial de la població amb rendes altes (>150% de la mediana). Àrea metropolitana de Barcelona, 2001



Font: IERMB i CRM-UAB, grups socials segons nivell de renda estimats a escala de secció censal a partir del Cens de Població i Habitatges, 2001 (INE) i l'Enquesta de condicions de vida i hàbits de la població, 2000 (IERMB).

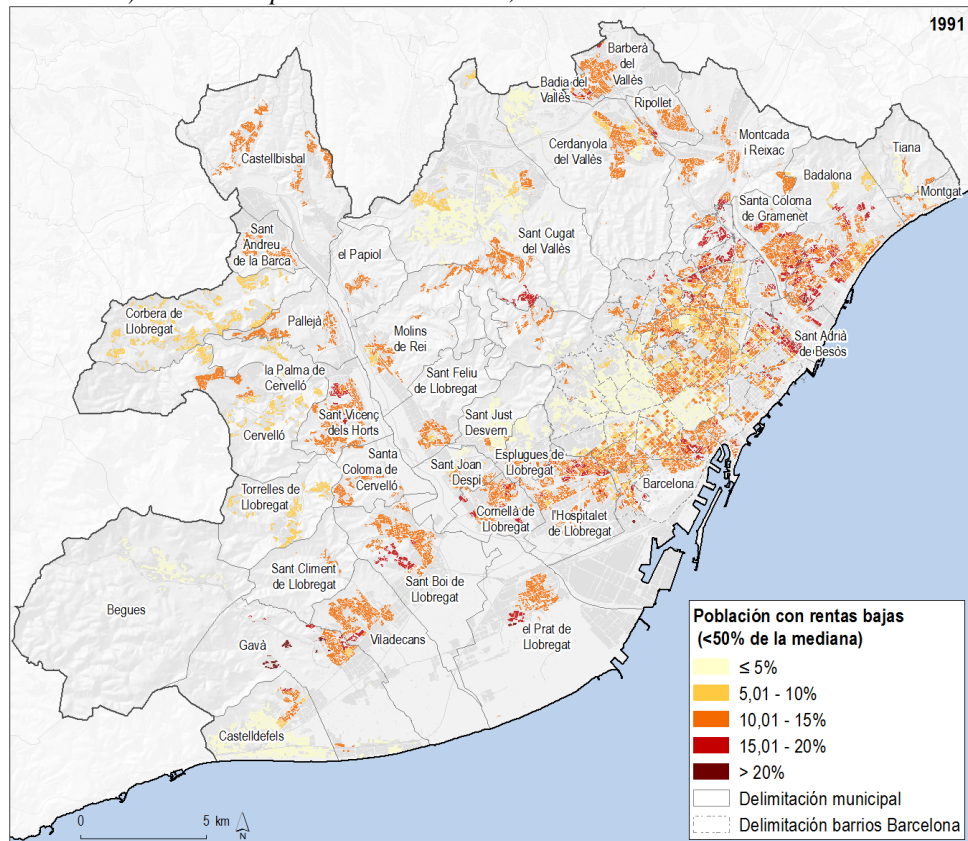
Figura 5. Distribució territorial de la població amb rendes altes (>150% de la mediana). Àrea metropolitana de Barcelona, 2011



Font: IERMB i CRM-UAB, grups socials segons nivell de renda estimats a escala de secció censal a partir del Cens de Població i Habitatges, 2011 (INE) i l'Enquesta de condicions de vida i hàbits de la població, 2011 (Idescat i IERMB).

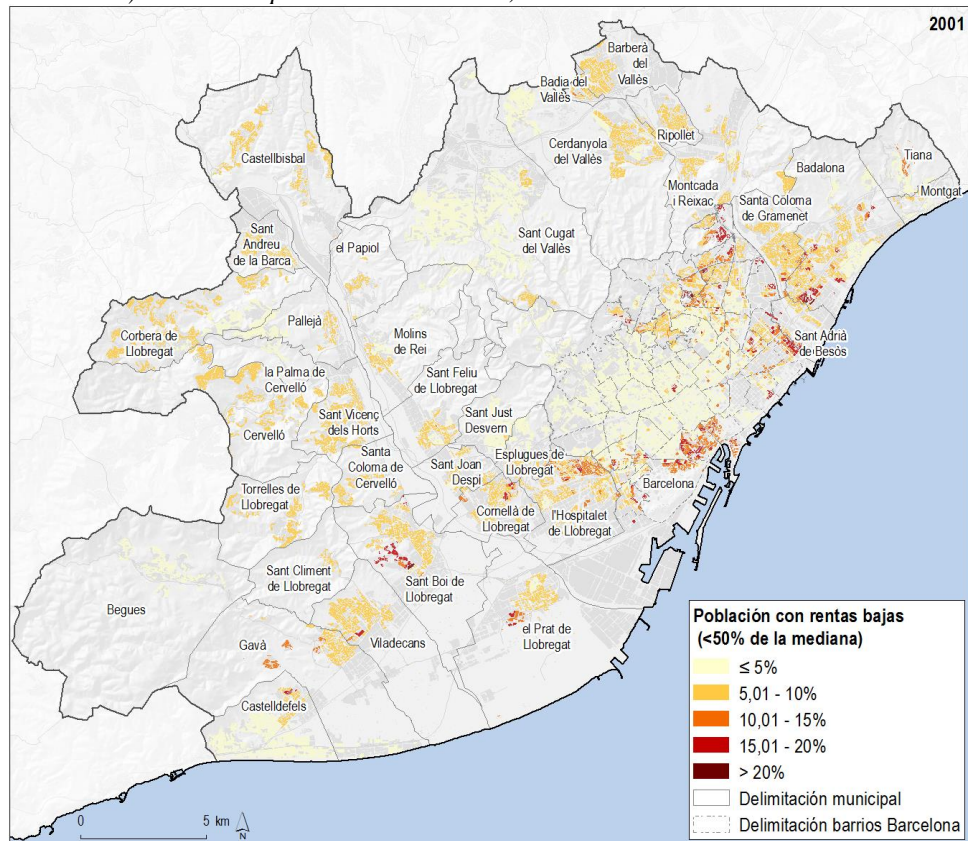


Figura 6. Distribució territorial de la població amb rendes baixes (<50% de la mediana). Àrea metropolitana de Barcelona, 1991



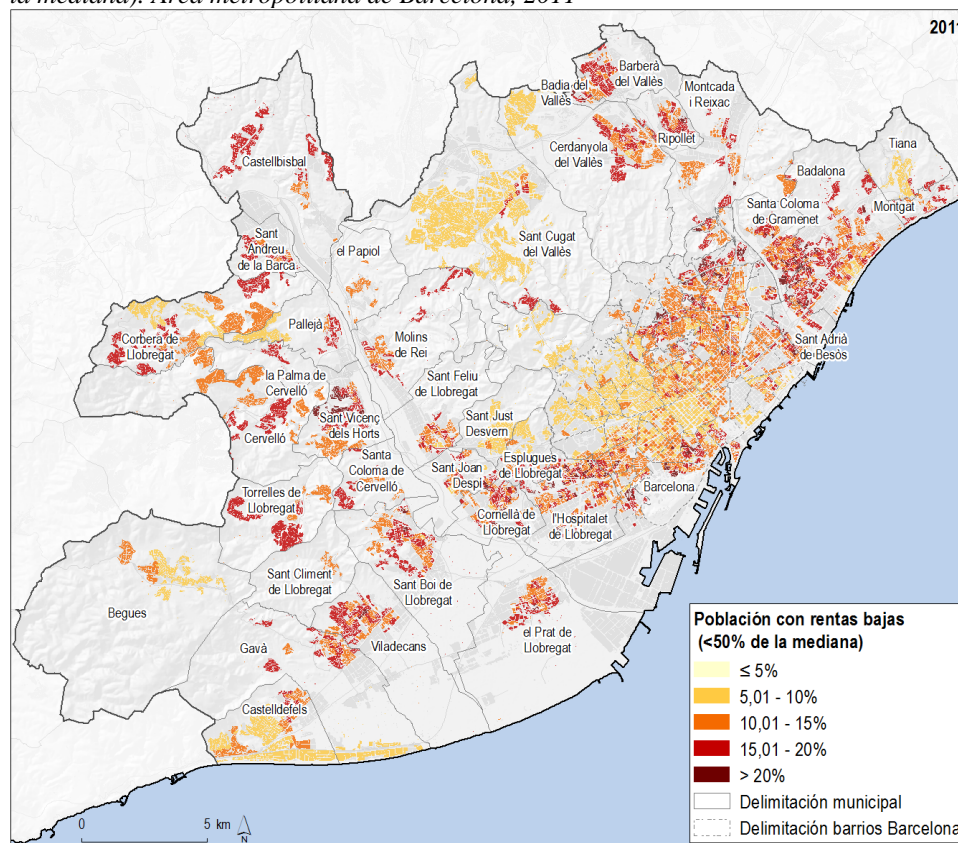
Font: IERMB i CRM-UAB, grups socials segons nivell de renda estimats a escala de secció censal a partir del Cens de Població i Habitatges, 1991 (INE) i l'Enquesta de condicions de vida i hàbits de la població, 1990 (IERMB).

Figura 7. Distribució territorial de la població amb rendes baixes (<50% de la mediana). Àrea metropolitana de Barcelona, 2001



Font: IERMB i CRM-UAB, grups socials segons nivell de renda estimats a escala de secció censal a partir del Cens de Població i Habitatges, 2001 (INE) i l'Enquesta de condicions de vida i hàbits de la població, 2000 (IERMB).

Figura 8. Distribució territorial de la població amb rendes baixes (<50% de la mediana). Àrea metropolitana de Barcelona, 2011



Font: IERMB i CRM-UAB, grups socials segons nivell de renda estimats a escala de secció censal a partir del Cens de Població i Habitatges, 2011 (INE) i l'Enquesta de condicions de vida i hàbits de la població, 2011 (Idescat i IERMB).

## 6. REFLEXIONS FINALS

Tot i que malauradament no es disposa d'una font de dades alternativa per poder realitzar un contrast rigorós dels resultats obtinguts de l'estimació realitzada, els mapes descriptius són, en general, força coherents amb el coneixement del territori. Per tant, els resultats estimats es poden considerar acceptables com a aproximació als nivells de renda de la població resident a



cada secció censal. Si bé no són del tot precisos per determinar valors per cada unitat territorial (atenent al nivell d'error que s'ha d'assumir en l'estimació), sí que mostren un gran potencial per ser analitzats de manera conjunta en termes estructurals. En aquest sentit, per exemple, aquesta informació estimada permet el càlcul d'indicadors de segregació residencial estàndards, oferint coneixement rellevant sobre quina han estat l'evolució de les pautes d'estructuració socioresidencial a l'àrea metropolitana de Barcelona durant les darreres dècades (Porcel i Navarro, 2016; Porcel, 2016). D'aquesta manera, s'ha assolit un dels principals objectius pels quals es va dissenyar aquest mètode d'estimació.

Evidentment, la metodologia compta també amb limitacions importants que s'han d'assenyalar. Una d'elles és la qualitat dels models explicatius (explicat a l'apartat 4.3), ja que les variables predictives que es poden introduir venen condicionades per la seva presència també en les dades censals. A més, per evitar fer més complexa l'estimació de la llei conjunta de les variables predictives a escala de secció censal (explicat a l'apartat 4.4), aquestes s'han reduït únicament a dues variables per model, essent la categoria professional i el nivell d'estudis finalment les variables predictives seleccionades en tots els casos.

Per acabar, cal reconèixer que la utilització de dades estimades a escala de secció censal és una opció no desitjada *a priori*, sinó que es tracta d'una necessitat per avançar en determinades línies de recerca socioterritorials davant dels dèficits que presenta sobre aquest aspecte el sistema estadístic espanyol. En aquest sentit i per millorar la qualitat de la informació disponible per àrees petites, resulta absolutament necessari que les oficines estadístiques avancin en el tractament procedent de registres administratius. No obstant això, també és necessari continuar avançant en la millora de les tècniques d'estimació d'informació per àrees petites, sobretot en el camp dels estudis urbans.

## BIBLIOGRAFÍA

- Brown, D. T. (1959). A note on Approximations to Discrete Probability Distributions. *Information and control*, 2(4), 386-392.
- Calvo, M.J. (2007). *Distribució territorial de la Renda Familiar a Barcelona*. Barcelona: Ajuntament de Barcelona.
- Chen, S.X., i Liu, J.S. (1997). Statistical applications of the Poisson-binomial and conditional Bernoulli distributions. *Statistica Sinica*, 7, 875-892. Disponible a <https://pdfs.semanticscholar.org/6dae/7b9a1238f94f6a5e11c3686f5962f407c7de.pdf>
- Deming, W.E., i Stephan, F.F. (1940). On a least squares adjustment of a sampled frequency table when the expected marginal totals are known. *The Annals of Mathematical Statistics*, 11(4), 427-444. Disponible a <https://projecteuclid.org/euclid.aoms/1177731829>
- Elbers, C., i van der Weide, R. (2014). *Estimation of Normal Mixtures in a Nested Error Model with an Application to Small Area Estimation of Poverty and Inequality* (Policy Research Working Paper, N° 6962). Washington, DC. World Bank Group. Disponible a <https://openknowledge.worldbank.org/handle/10986/19362>
- Fienberg, S.E. (1968). The Geometry of an  $r \times c$  Contingency Table. *The Annals of Mathematical Statistics*, 39(4), 1186-1190. Disponible a <https://www.jstor.org/stable/pdf/2239685.pdf>
- Fienberg, S.E. (1970). An iterative procedure for estimation in contingency tables. *The Annals of Mathematical Statistics*, 41(3), 907-917. Disponible a <https://projecteuclid.org/euclid.aoms/1177696968>
- Gelman, A., Carlin, J.B., Stern, H.S., i Rubin, D.B. (1995). *Bayesian Data Analysis*. New York: Chapman and Hall.
- Lewis, P.M. (1959). Approximating Probability Distributions to Reduce Storage Requirements. *Information and control*, 2(3), 214-225.
- Lomax, N., i Norman, P. (2016). Estimating Population Attribute Values in a Table: “Get Me Started in” Iterative Proportional Fitting. *The Professional Geographer*, 68(3), 451-461. doi: [10.1080/00330124.2015.1099449](https://doi.org/10.1080/00330124.2015.1099449)

- Madariaga, R., Martori, J. C., i Oller, R. (2012). Distribución espacial y desigualdad de la renta salarial en el Área Metropolitana de Barcelona. *Scripta Nova: revista electrònica de geografia y ciencias sociales*, 16.
- Moreno, A. (2003). La distribución especial de la renta en la Comunidad de Madrid: análisis y aplicaciones. Madrid: Instituto de Estadística de la Comunidad de Madrid.
- Porcel, S. (2016). *Dinàmiques de estructuració socioresidencial en la Barcelona metropolitana postindustrial. ¿Hacia una ciudad dual o cuarteada?* Tesi doctoral, Departament de Sociologia, Universitat Autònoma de Barcelona.
- Porcel, S., i Navarro-Varas, L. (2016). Desigualtat social i segregació residencial a l'AMB. A J. Trullén i V. Galetto (dirs.) *Anuari Metropolità de Barcelona 2015* (pp. 56-87). Bellaterra: Institut d'Estudis Regionals i Metropolitans de Barcelona i Àrea Metropolitana de Barcelona.
- R Core Team (2017). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. <http://www.R-project.org/>.
- Raghunathan, T.E., Lepkowski, J., Van Hoewyk, J.H., i Solenberger, P.W. (2000). A Multivariate Technique for Multiply Imputing Missing Values Using a Sequence of Regression Models. *Survey Methodology*, 27(1). Disponible a <https://www.researchgate.net/publication/244959137>
- Raghunathan, T.E., Solenberger, P.W., i Van Hoewyk, J.H. (2002). *IVEware: Imputation and Variance Estimation Software user guide*. Ann Arbor (MI). Survey Research Center, Institute for Social Research. University of Michigan. Disponible a [ftp://ftp.isr.umich.edu/pub/src/smp/ive/ive\\_user.pdf](ftp://ftp.isr.umich.edu/pub/src/smp/ive/ive_user.pdf)
- Simpson, L., i Tranmer, M. (2005). Combining sample and census data in small area estimates: Iterative Proportional Fitting with standard software. *The Professional Geographer*, 57(2), 222-234. Disponible a <https://onlinelibrary.wiley.com/doi/abs/10.1111/j.0033-0124.2005.00474.x>
- Stephan, F.F. (1942). An iterative method of adjusting sample frequency tables when expected marginal totals are known. *The Annals of Mathematical Statistics*, 13(2), 166-178. Disponible a <https://projecteuclid.org/euclid.aoms/117731598>

Tresens, A. (2014). *Métode IPF: procediment iteratiu per estimar lleis conjuntes amb restriccions*. Treball de fi de grau de les titulacions de Matemàtiques i Estadística Aplicada (2014). Disponible a <https://ddd.uab.cat/record/126227>

**ANNEX**
*Taula A.1. Proporció de seccions censals imputades segons categoria-variable i procediment utilitzat. Cens de Població i Habitatges, 2011*

<b>Categoria-variable</b>	<b>Descripció</b>	<b>Original</b>	<b>1 missing result</b>	<b>Més d'1 missing result</b>
GRADO1	Estudios. Primer grado	64,20%	20,40%	15,40%
GRADO2	Estudios. Segundo grado	98,90%	0,00%	1,10%
GRADO3	Estudios. Tercer grado	66,00%	18,90%	15,10%
PARO	Tasa de paro	63,70%	34,80%	1,50%
OCUP	Tasa de ocupación	98,30%	0,20%	1,50%
ACT	Tasa de actividad	99,30%	0,00%	0,70%
ACHOGAR1	Hogares con todos sus miembros inactivos	65,00%	2,20%	32,80%
ACHOGAR2	Hogares con todos sus miembros activos ocupados	36,70%	11,40%	51,90%
ACHOGAR3	Hogares con sus miembros ocupados y parados	93,40%	0,10%	6,50%
ACHOGAR4	Hogares con todos sus miembros activos parados	33,00%	23,20%	43,80%
SECTOR1	Agricultura, ganadería y pesca	66,10%	3,80%	30,10%
SECTOR2	Industria	33,40%	7,60%	59,00%
SECTOR3	Construcción	19,10%	15,30%	65,60%
SECTOR4	Servicios	95,50%	0,30%	4,20%
SITPROF1	Empresario o profesional que emplea personal	24,40%	2,00%	73,60%
SITPROF2	Empresario o profesional que no emplea personal	22,00%	1,30%	76,70%
SITPROF3	Trabajador por cuenta ajena con carácter fijo o indefinido	94,50%	0,20%	5,30%
SITPROF4	Trabajador por cuenta ajena con carácter eventual, temporal.	24,60%	2,50%	72,90%

SITPROF5	Otra situación (Ayuda familiar)	84,20%	0,80%	15,00%
SITPROF6	Otra situación (Miembro de cooperativas)	81,50%	1,40%	17,10%
OCUPACION1	Dirección de las empresas y de las administraciones públicas	27,30%	0,60%	72,10%
OCUPACION2	Técnicos y profesionales científicos e intelectuales	39,60%	0,00%	60,40%
OCUPACION3	Técnicos y profesionales de apoyo	26,10%	0,00%	73,90%
OCUPACION4	Empleados de tipo administrativo	29,60%	0,00%	70,40%
OCUPACION5	Trabajadores de servicios de restauración, personales, protección	36,90%	0,10%	63,00%
OCUPACION6	Trabajadores cualificados en la agricultura y la pesca	66,60%	0,30%	33,10%
OCUPACION7	Artesanos y trabajadores cualificados de las industrias manufactureras	24,80%	0,10%	75,10%
OCUPACION8	Operadores de instalaciones y maquinaria, y montadores	19,50%	0,10%	80,40%
OCUPACION9	Trabajadores no cualificados	16,00%	0,10%	83,90%
VIVIEDIF	Relación viviendas/edificios	100,00%	0,00%	0,00%
CVVIVIF1	Tipo vivienda. Principales	99,40%	0,40%	0,20%
CVVIVIF2	Tipo vivienda. Secundarias	57,20%	11,10%	31,70%
CVVIVIF3	Tipo vivienda. Vacías	35,50%	32,70%	31,80%
ESTADO1	Edificios en estado ruinoso	100,00%	0,00%	0,00%
ESTADO2	Edificios en mal estado	100,00%	0,00%	0,00%
ESTADO3	Edificios en estado deficiente	100,00%	0,00%	0,00%
ESTADO4	Edificios en buen estado	100,00%	0,00%	0,00%
PLANTAS1	Edificios con 1 planta	100,00%	0,00%	0,00%
PLANTAS2	Edificios con 2 plantas	100,00%	0,00%	0,00%

PLANTAS3	Edificios con 3 plantas	100,00%	0,00%	0,00%
PLANTAS4	Edificios con 4 plantas	100,00%	0,00%	0,00%
PLANTAS5	Edificios con 5 plantas	100,00%	0,00%	0,00%
PLANTAS6	Edificios con 6 plantas o más	100,00%	0,00%	0,00%
ANYOCONS1	Año de construcción. Antes de 1951	100,00%	0,00%	0,00%
ANYOCONS2	Año de construcción. 1951-1960	100,00%	0,00%	0,00%
ANYOCONS3	Año de construcción. 1961-1970	100,00%	0,00%	0,00%
ANYOCONS4	Año de construcción. 1971-1980	100,00%	0,00%	0,00%
ANYOCONS5	Año de construcción. 1981-1990	100,00%	0,00%	0,00%
ANYOCONS6	Año de construcción. 1991-2000	100,00%	0,00%	0,00%
ANYOCONS7	Año de construcción. Después de 2000	100,00%	0,00%	0,00%
SUT1	Superficie de la vivienda. De 1 a 50m <sup>2</sup>	28,10%	8,10%	63,80%
SUT2	Superficie de la vivienda. De 51 a 70m <sup>2</sup>	54,20%	3,60%	42,20%
SUT3	Superficie de la vivienda. De 71 a 100m <sup>2</sup>	82,30%	2,00%	15,70%
SUT4	Superficie de la vivienda. De 101 a 150m <sup>2</sup>	47,00%	3,10%	49,90%
SUT5	Superficie de la vivienda. Más de 150m <sup>2</sup>	53,90%	1,20%	44,90%
TENEN1	Propiedad por compra totalmente pagada	84,00%	0,20%	15,80%
TENEN2	Propiedad por compra con pagos pendientes	66,50%	0,30%	33,20%
TENEN3	Propiedad por herencia o donación	33,20%	0,80%	66,00%
TENEN4	Alquiler	31,60%	1,90%	66,50%
TENEN5	Cedida gratuitamente o a bajo precio	47,80%	1,00%	51,20%

TENEN6	Otro régimen de tenencia	21,90%	1,20%	76,90%
LN1	Nacidos en Cataluña	99,50%	0,10%	0,40%
LN2	Nacidos en el resto de España	85,80%	1,30%	12,90%
LN3	Nacidos en la UE-15	41,00%	2,70%	56,30%
LN4	Nacidos en el resto de Europa	56,30%	1,40%	42,30%
LN5	Nacidos en América del Norte	86,30%	0,20%	13,50%
LN6	Nacidos en América Central y del Sur	31,30%	5,60%	63,10%
LN7	Nacidos en África	54,40%	2,60%	43,00%
LN8	Nacidos en Asia y Oceanía	76,00%	0,70%	23,30%
ANYOC1	Residencia en Catalunya hace menos de 11 años	35,40%	2,20%	62,40%
ANYOC2	Residencia en Catalunya entre 11 y 20 años	5,10%	30,80%	64,10%
ANYOC3	Residencia en Catalunya hace más de 20 años	90,80%	0,10%	9,10%
ANYOC4	Siempre ha residido en Cataluña	99,50%	0,00%	0,50%
ANYOE1	Residencia en España hace menos de 2 años	65,20%	0,20%	34,60%
ANYOE2	Residencia en España entre 2 y 5 años	31,40%	0,60%	68,00%
ANYOE3	Residencia en España entre 6 y 10 años	26,70%	1,00%	72,30%
ANYOE4	Residencia en España entre 11 y 20 años	31,10%	0,70%	68,20%
ANYOE5	Residencia en España hace más de 20 años	17,60%	3,70%	78,70%
ANYOE6	Siempre han residido en España	99,90%	0,00%	0,10%
ANYOM1	Residencia en el municipio hace menos de 2 años	25,90%	14,90%	59,20%
ANYOM2	Residencia en el municipio entre 2 y 5 años	42,40%	1,10%	56,50%



ANYOM3	Residencia en el municipio entre 6 y 10 años	45,30%	1,10%	53,60%
ANYOM4	Residencia en el municipio entre 11 y 20 años	46,00%	0,60%	53,40%
ANYOM5	Residencia en el municipio hace más de 20 años	94,50%	0,10%	5,40%
ANYOM6	Siempre han residido en el municipio	98,40%	0,00%	1,60%
CID1	Conocimiento del catalán. No lo entiende	99,90%	0,10%	0,00%
CID2	Conocimiento del catalán. Lo entiende	99,90%	0,10%	0,00%
CID3	Conocimiento del catalán. Lo sabe hablar	99,60%	0,00%	0,40%
CID4	Conocimiento del catalán. Lo sabe leer	99,70%	0,00%	0,30%
CID5	Conocimiento del catalán. Lo sabe escribir	98,80%	0,00%	1,20%
NMIEM1	Miembros del hogar. 1 persona	55,60%	0,30%	44,10%
NMIEM2	Miembros del hogar. 2 personas	78,40%	0,10%	21,50%
NMIEM3	Miembros del hogar. 3 personas	48,40%	0,30%	51,30%
NMIEM4	Miembros del hogar. 4 personas	33,70%	1,40%	64,90%
NMIEM5	Miembros del hogar. 5 personas	20,30%	7,00%	72,70%
NMIEM6	Miembros del hogar. 6 o más personas	56,70%	0,90%	42,40%
TIPOH1	Hogares unipersonales	55,60%	44,40%	0,00%
TIPOH2	Hogares de 2 o más personas sin núcleo	31,30%	0,10%	68,60%
TIPOH3	Hogares de pareja sin hijos (sin otras personas)	56,70%	0,00%	43,30%
TIPOH4	Hogares de pareja sin hijos (con otras personas)	44,30%	0,00%	55,70%

TIPOH5	Hogares de pareja con hijos (sin otras personas)	73,60%	0,10%	26,30%
TIPOH6	Hogares de pareja con hijos (con otras personas)	36,60%	0,00%	63,40%
TIPOH7	Hogares con padre o madre con hijos(sin otras personas)	9,80%	0,40%	89,80%
TIPOH8	Hogares con padre o madre con hijos (con otras persona)	52,30%	0,00%	47,70%
TIPOH9	Hogares de 2 o más núcleos	43,70%	0,00%	56,30%
ESTH1	Hogares unipersonales de 16 a 64 personas	21,00%	9,80%	69,20%
ESTH2	Hogares unipersonales de 65 y más años	15,70%	15,10%	69,20%
ESTH3	Hogares 1 adulto con menores	31,90%	0,80%	67,30%
ESTH4	Hogares 2 o más adultos con 1 menor	20,80%	0,30%	78,90%
ESTH5	Hogares 2 o más adultos con 2 menores	16,10%	0,70%	83,20%
ESTH6	Hogares 2 o más adultos con 3 o más menores	54,50%	0,40%	45,10%
ESTH7	Hogares 2 o más adultos de 16 a 64 años sin menores	66,20%	0,00%	33,80%
ESTH8	Hogares 2 uno al menos de 65 años o más sin menores	53,10%	0,10%	46,80%
CALEF_1	No dispone Calefacción	100,00%	0,00%	0,00%
CALEF_2	Dispone de Calefacción	100,00%	0,00%	0,00%