

Col·lecció d'estudis laborals
núm. 1

MONOGRÀFIC

Observatori del Treball de les Illes Balears

Anàlisi *de la* representativitat *de* l'enquesta de població activa *a les* Illes Balears

Versió en castellà a partir de la pàgina 57



**Govern
de les Illes Balears**



Direcció: Llorenç Pou Garcias
Director general de Planificació Estratègica

Coordinació: Delia Riera Tur

Autor: Joaquín Alegre Martín

ISBN 978-84-693-2893-4

Si desitjau més informació i/o us voleu subscriure als informes de l'Observatori del Treball de les Illes Balears, us podeu dirigir a otib@dgplaes.caib.es o bé:

Direcció General de Planificació Estratègica
Conselleria de Treball i Formació
Plaça de Son Castelló, 1
Polígon de Son Castelló. 07009 Palma
Tel.: 971 17 63 00
<http://www.caib.es>
<http://treballiformacio.caib.es>

ANÀLISI DE LA REPRESENTATIVITAT DE L'ENQUESTA DE POBLACIÓ ACTIVA A LES ILLES BALEARS

OBSERVATORI DEL TREBALL DE LES ILLES BALEARS

Joaquín Alegre Martín
Departament d'Economia Aplicada
Universitat de les Illes Balears

Col·lecció d'Estudis Laborals.
Núm.1

ÍNDEX

1. INTRODUCCIÓ	7
2. DISSENY MOSTRAL DE L'EPA.....	11
3. PROBLEMES DEL DISSENY MOSTRAL DE L'EPA	13
4. REPRESENTATIVITAT DE LA MOSTRA	17
4.1. Comparació amb fonts externes	18
4.2. Població.....	19
4.3. Nacionalitat espanyola i estrangera	21
4.4. Nacionalitats estrangeres	26
4.5. Edat.....	29
4.6. Sexe	34
4.7. Situació laboral	38
5. FACTORS D'ELEVACIÓ.....	43
6. CONCLUSIONS	51
7. REFERÈNCIES	55

1. INTRODUCCIÓ

En aquest treball s'analitza la capacitat que té l'Enquesta de Població Activa (EPA) de l'INE per proporcionar informació representativa de les Illes Balears. Les dificultats que ha tengut l'EPA per descriure correctament la situació del mercat laboral espanyol han estat la causa dels seus diversos canvis metodològics. Les millores introduïdes, especialment després de les modificacions de l'any 2005, han reafirmat l'EPA com una de les fonts d'informació més fiables sobre la situació laboral dels espanyols. En l'actualitat l'INE entrevista aproximadament 200.000 persones a 65.000 llars, oferint informació a l'àmbit nacional i provincial tant de la població activa i les seves diverses categories, com de la població inactiva. La validesa d'aquestes estimacions, amb tots els matisos o interpretacions que vulguin realitzar-se, ha estat confirmada en diversos estudis. No obstant això, encara hi ha pocs treballs que analitzin la validesa de les estimacions de l'EPA a l'àmbit de comunitat autònoma. Són diversos els serveis d'estudis i instituts estadístics regionals que reclamen un increment de la mida mostral corresponent a l'àmbit autonòmic i provincial; aquesta reclamació es justifica perquè en aquest nivell de desagregació és fàcil que en les anàlisis s'assoleixi la mida mínima a partir de la qual els resultats de l'Enquesta no són fiables¹. Com ja assenyalava l'INE el 2004, “és necessari prosseguir amb el contrast de les sèries de resultats amb els d'altres fonts estadístiques i registres administratius, a fi de conèixer i explicar les causes de certes inconsistències que s'han observat, sobretot quan es fa feina amb desagregacions espacials” (INE, 2004). No obstant això, encara falten estudis que valorin la capacitat de l'Enquesta per estimar correctament informació rellevant a l'àmbit de les comunitats autònomes.

El disseny mostral de l'EPA té per objectiu que les persones seleccionades a la mostra siguin representatives de tota la població nacional. Per a les comunitats autònomes (i

¹ L'INE adverteix que resultats inferiors a 5.000 persones no són fiables, en estar subjectes a fortes variacions conseqüència dels errors mostrals.

les províncies) l'INE indica que només és vàlida la informació amb un menor grau de desagregació, depenent dels coeficients de variació dels estimadors. El disseny mostral de l'EPA no garanteix, per tant, la representativitat estadística provincial en totes les variables. En conseqüència, l'objectiu d'aquest treball no és esbrinar la precisió de les estimacions a un nivell molt desagregat. Les variables emprades es defineixen, per tant, en un nivell d'agregació alt. No obstant això, s'ha de tenir en compte que, en el context de l'EPA, la precisió en el nivell agregat depèn del grau de precisió emprat en un nivell d'agregació menor. És possible que els processos de reponderació que emprava l'EPA puguin equilibrar la informació agregada que s'obté de fonts d'informació externa, però que això sigui a costa de desvirtuar la informació més desagregada que també es deriva de l'enquesta.

Els continguts d'aquest treball es presenten en l'ordre següent. En un primer apartat s'explica el disseny mostral de l'EPA, tal com és exposat per l'INE. Encara que es tracta d'un disseny conegut, s'ha preferit fer-ne una presentació breu, ja que la representativitat de l'enquesta depèn principalment del seu disseny. En un segon apartat es discuteixen els problemes relacionats amb el tipus de disseny mostral estratificat que emprava l'EPA. Es tracta d'una crítica específica al disseny, que es basa, essencialment, en la possibilitat que els estrats que defineix l'INE manquin del grau de precisió suficient per estimar algunes variables. En un tercer apartat s'analitza la representativitat de la mostra de l'EPA a les Illes Balears, entenent per representativitat la capacitat de l'Enquesta per proporcionar xifres coincidents amb fonts d'informació externes (per exemple, del Padró Municipal). Després d'exposar el procés de calibratge o reponderació que emprava l'INE, s'analitza la capacitat de l'EPA per equilibrar informació referida a població (total i per nacionalitats), distribució per edats, sexe i situació laboral. Aquestes anàlisis permeten detectar quins són els segments en els quals l'EPA subestima les variables poblacionals. En el darrer apartat s'analitzen els factors d'elevació poblacional emprats a l'INE. Els resultats que s'obtenen en aquest apartat suggereixen que l'excessiu valor dels factors d'elevació poblacional en alguns segments de l'Enquesta, així com la seva variabilitat en el temps, no solament ocasionen problemes d'interpretació dels resultats de l'Enquesta, sinó que poden generar estimacions errònies. En aquest apartat s'assenyalen a més les dificultats del procés de calibratge quan aquest és una alternativa de facto a l'ampliació de la mida mostral. Finalment, s'exposen les principals conclusions del treball.

Per dur a terme el treball s'han emprat els fitxers de microdades de l'EPA, del període comprès entre el primer trimestre de 2005 i el tercer trimestre de 2008. La majoria de la informació emprada es refereix a la identificació de les Illes Balears,

encara que també s'han fet algunes anàlisis referides a tot Espanya. Les comparacions externes s'han realitzat, essencialment, segons els padrons municipals.

2. DISSENY MOSTRAL DE L'EPA

El disseny de la mostra EPA (INE, 2005) utilitza un mostratge en dues etapes amb estratificació en les unitats de primera etapa. Les unitats de primera etapa estan constituïdes per les seccions censals. Les unitats de segona etapa són els habitatges familiars principals (ocupats permanentment) i els allotjaments fixos. Dins de les unitats secundàries no es fa cap submostratge, recollint-se informació de totes les persones que hi viuen.

Les unitats de primera etapa s'estratifiquen segons un doble criteri:

- Criteri geogràfic (d'estratificació), pel qual les seccions censals s'agrupen en estrats dins de cada província, d'acord amb la importància demogràfica del municipi a què pertanyen.
- Criteri socioeconòmic (de subestratificació), pel qual les seccions censals s'agrupen dins de cada un dels estrats, segons les seves característiques.

(1) Els estrats es formen distingint:

- Estrat 1: municipi capital de província.
- Estrats 2 i 3: municipis autorepresentats.
- Estrat 4: municipis entre 50.000 i 100.000 habitants.
- Estrat 5: municipis entre 20.000 i 50.000 habitants.
- Estrat 6: municipis entre 10.000 i 20.000 habitants.
- Estrat 7: municipis entre 5.000 i 10.000 habitants.
- Estrat 8: municipis entre 2.000 i 5.000 habitants.
- Estrat 9: municipis menors de 2.000 habitants.

(2) Dins de cada estrat s'han generat subestrats a partir de dos criteris. Per a les seccions dels estrats corresponents a municipis de menys de 10.000 habitants se'ls assigna com a subestrat la comarca del municipi a què pertanyen. En el cas de les Balears,

aquest criteri s'aplica al 20% de les seccions censals. Per a la resta de seccions els subestrats es generen mitjançant una anàlisi clúster que emprà informació demogràfica i socioeconòmica de les seccions. L'objectiu és formar grups homogenis de seccions. La informació utilitzada per generar els clústers procedeix del Cens 2001 i de l'Agència Estatal d'Administració Tributària. Segons l'INE (2005), s'empren les variables següents:

- els percentatges d'aturats, inactius i ocupats;
- el percentatge d'estrangers;
- el percentatge de població en diferents trams d'edat (entre 0 i 19 anys; entre 20 i 24; 65 o més anys);
- percentatge de persones per diferents graus d'estudis (analfabets o estudis primaris; ESO, EGB, batxillerat o FP; estudis universitaris);
- percentatges de la població ocupada segons la seva condició socioeconòmica (empresaris, treballadors, etc.);
- renda per habitatge;
- rendes de capital sobre la renda total;
- renda agrària sobre la renda total.

Com assenyala el document de l'INE, aquestes variables han estat seleccionades considerant que estan molt correlacionades amb l'objecte d'estudi de l'enquesta.

Dins de cada província l'afixació entre estrats és proporcional a la mida de cada un d'aquests, encara que s'han potenciat els municipis de dimensions més grans. Dins dels estrats, l'afixació entre subestrats és estrictament proporcional al seu nombre d'habitatges familiars. Inicialment, la mida de la mostra es va fixar mitjançant un mètode de mínima variància per a un cost fix, en un total de 3.000 seccions, amb una mitjana de 20 habitatges per secció. Per distribuir aquestes seccions entre províncies, estrats i subestrats, es van establir diversos criteris, que garantissin la fiabilitat de les estimacions a l'àmbit nacional i provincial. El criteri final d'afixació és un compromís entre l'afixació uniforme i la proporcional. Les Illes Balears distribueixen les seves seccions censals a cada estrat de la manera següent: estrat 1, 44; estrat 5, 27; estrat 6, 12; estrat 7, 12; estrat 8, 9.

3. PROBLEMES DEL DISSENY MOSTRAL DE L'EPA

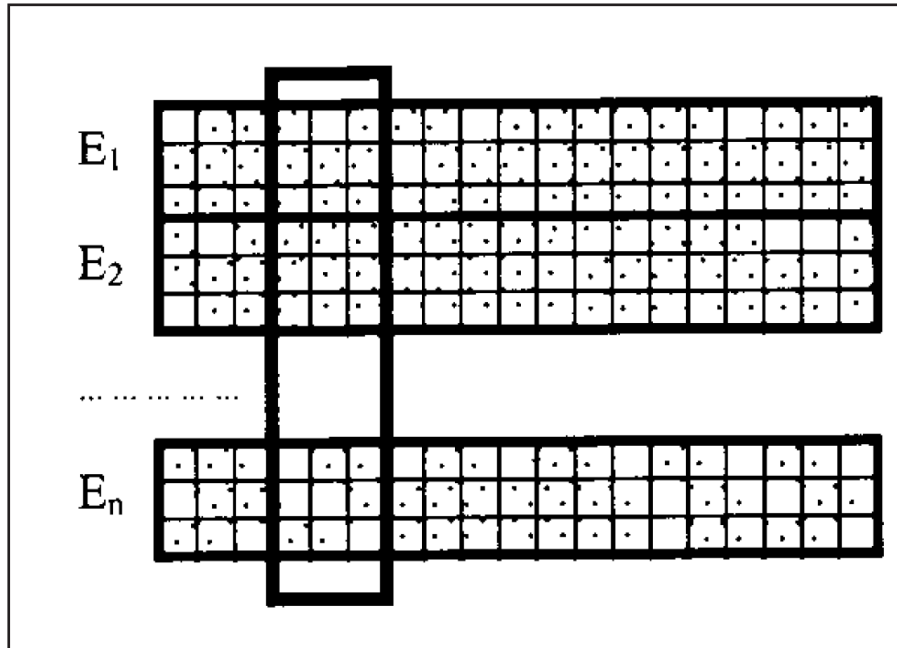
En el mostatge estratificat, els estrats es defineixen de manera que, respecte a les variables d'interès, cada un dels estrats presenti unitats homogènies entre si i, al seu torn, siguin heterogènies quan es comparen entre ells. L'eficàcia del mostatge estratificat depèn que els grups o estrats presentin aquestes característiques. Tanmateix, aquest no sembla ser el cas d'algunes característiques d'interès a l'EPA. Per exemple, Martí i Ródenas (2004) assenyalen que els individus amb nacionalitat estrangera constitueixen una categoria poblacional que es presenta per igual a tots els estrats de la mostra i es distribueix de manera homogènia dins de l'habitatge. Això pot tenir com a conseqüència que les estimacions lligades a aquest tipus de característica siguin imprecises.

Martí i Ródenas (2004) exposen gràficament les dificultats d'estimar determinades característiques mitjançant un disseny mostral per estrats. En les gràfiques (que reproduïm del seu article), cada estrat està constituït per un conjunt d'habitatges, representats mitjançant caselles. Dins de cada habitatge els individus que tenen la característica estudiada se simbolitzen amb un punt. La mostra d'habitatges seleccionats en el mostatge s'assenyala amb la barra vertical que travessa tots els estrats. Martí i Ródenas (2004) distingeixen tres possibles situacions:

En la primera, es vol estimar una característica que és present per igual a tots els estrats i a més és una unitat heterogènia en la darrera unitat de mostatge (en el cas de l'EPA, l'habitatge). Exemples d'aquest tipus de característica són el sexe, l'edat o l'estat civil. Si a més l'habitatge està compost per persones de diferent sexe,

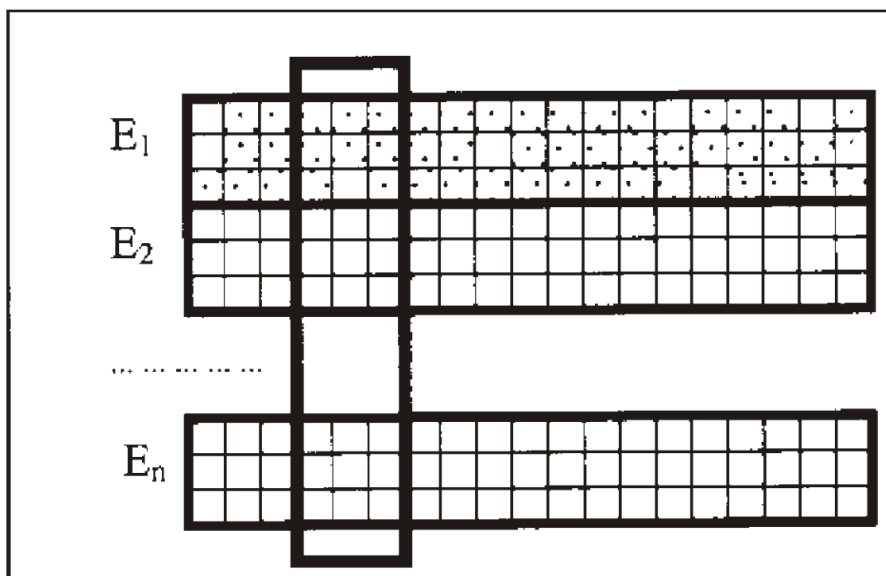
edat i estat civil, la mostra que s'extreu a qualsevol estrat serà representativa de la població. La Gràfica 1 il·lustra aquesta situació.

GRÀFICA 1. SISTEMA D'ESTRATIFICACIÓ D'UNA MOSTRA



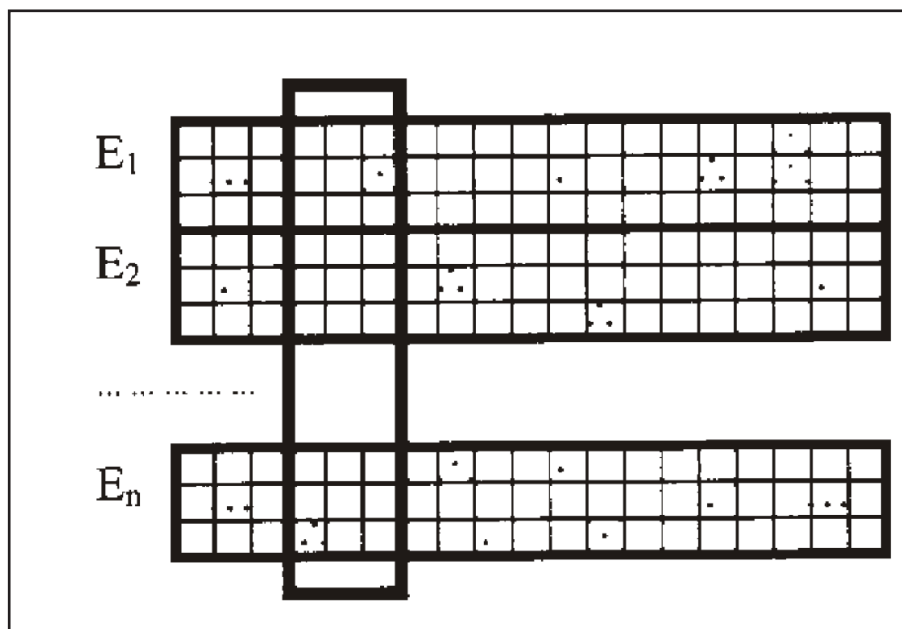
En la segona situació, la característica a estimar la posseeixen un nombre reduït d'habitatges, distribuint-s'hi de manera homogènia (vegi's la Gràfica 2). Si els estrats es construeixen tenint en compte aquesta característica, creant un estrat que contengui els habitatges amb individus que la posseeixen, la mostra resultarà representativa.

GRÀFICA 2. SISTEMA D'ESTRATIFICACIÓ D'UNA MOSTRA



El problema pot sorgir quan la característica d'interès no es té en compte quan es construeixen els estrats. Aquesta situació quedaria reflectida en la Gràfica 3. Si cap estrat no es dissenya per recollir aquella característica, la mostra seleccionada inclourà pocs habitatges amb individus que tinguin la característica. Una possible conseqüència d'aquest problema és que la presència d'aquesta característica a la mostra pot presentar gran variabilitat. Però fins i tot si una mida mostral adequada aconseguix que a la mostra la característica poblacional estigui ben representada, la representativitat d'altres variables associades a aquesta característica serà extremadament baixa.

GRÀFICA 3. SISTEMA D'ESTRATIFICACIÓ D'UNA MOSTRA



Encara que s'ha apuntat que la falta de resposta pot ser una de les raons per les quals les persones de nacionalitat estrangera no estan ben representades a la mostra de l'EPA, les dades disponibles per a l'any 2007 no semblen confirmar-ho. Segons l'informe de l'INE d'avaluació de la qualitat de l'EPA-2007, les negatives a respondre a l'enquesta per primera vegada corresponents a persones estrangeres és de 4,9 % (per a tot Espanya, tercer trimestre), una mica inferior al percentatge de persones estrangeres que s'enquesta a l'EPA (5,7 % en el tercer trimestre).

Els possibles problemes d'utilització de l'EPA poden resumir-se, per tant, que la precisió de les estimacions depèn de la mida de la mostra i l'heterogeneïtat o homogeneïtat de la variable estudiada:

-
- (1) Respecte a la mida mostral, Martí i Ródenas (2004) assenyalen que “mostres dimensionades per extreure conclusions generals o mesurar determinats paràmetres no es poden utilitzar per a l'anàlisi de caràcters no gaire freqüents, ja que el nombre d'entrevistes realitzades a persones que disposin de l'esmentat tret no serà suficient per donar suport a l'estimació, ja que es pot superar àmpliament el que es podria anomenar l'error permès”. I, per tant, només s'haurien d'estimar característiques poblacionals la presència de les quals a la mostra és prou àmplia.
 - (2) En el disseny de l'EPA s'inclouen un nombre limitat de característiques per definir els estrats de població. Aquesta definició pot perjudicar la precisió de característiques que, en primer lloc, es distribueixen de manera uniforme entre els estrats de la població i, en segon lloc, ho fan homogèniament dins la darrera unitat de mostratge (l'habitatge).

Com s'ha explicat en exposar el disseny mostral de l'EPA, els criteris de definició dels estrats han intentat incloure no només criteris geogràfics, sinó socioeconòmics, emprant en aquest darrer cas informació censal correlacionada amb l'objecte d'estudi de l'enquesta. Entre les variables emprades per generar aquests estrats es troba el percentatge d'estrangers. Així i tot, el procediment exposat per l'INE deixa algunes qüestions imprecises. En primer lloc, no es publica quina és la importància que cada una de les variables socioeconòmiques té en la determinació dels subestrats. Donat el conjunt de variables que s'utilitza no és descartable que siguin aquelles variables amb una major variabilitat les que més influeixin en el procés d'agrupació. En segon lloc, aquestes variables només juguen un paper secundari en la definició dels subestrats, havent-hi els estrats definits per un criteri geogràfic.

4. REPRESENTATIVITAT DE LA MOSTRA

El nombre d'individus de les Illes Balears sobre qui l'EPA ofereix informació efectiva varia en cada ona de l'enquesta. En el període comprès entre el primer trimestre de 2005 i el tercer trimestre de 2008 els enquestats a les Balears oscil·len entre un mínim de 3.752 i un màxim de 4.235. La informació de la mostra s'eleva a la població emprant els factors d'elevació calculats per l'INE. En principi, aquests es deriven del disseny mostral i eleven la informació recollida en un estrat a allò que aquest representa en la població. A més, l'EPA corregeix els pesos o factors d'elevació originals (els que es dedueixen del disseny mostral), mitjançant un procés de calibratge o reponderació, que emprà fonts estadístiques externes. El seu objectiu és millorar la qualitat de les estimacions, de manera que l'estimació de les variables per a les quals es disposa d'informació externa fiable coincideixi amb les dades d'aquella font.

La base de microdades de l'enquesta ofereix els factors d'elevació de cada individu. Aquests valors indiquen el nombre de persones de la població a la qual representen cada una de les persones enquestades. A partir de l'any 2005 aquests factors s'han modificat per reflectir millor el pes de la població estrangera. Els factors d'elevació s'obtenen en dues etapes. En primer lloc, s'obtenen els factors d'elevació a partir del disseny mostral de l'enquesta. Així i tot, les dades mostrals elevades no concorden amb les dades de població. La diferència entre les estimacions derivades de l'enquesta i la d'altres fonts externes les justifica l'INE per diverses causes (INE, Mètode de reponderació emprat a l'EPA):

- Falta de resposta, que produeix biaixos en els estimadors finals en afectar de manera relativament diferent llars o grups de població de determinada composició.

- Possible tendència de la mostra a sobrerrepresentar llars d'una determinada estructura en detriment d'altres.

Per això, en una segona etapa, denominada de calibrat o reponderació, els factors d'elevació es modifiquen perquè els resultats poblacionals coincideixin amb els d'una o diverses fonts externes de referència. En el cas de l'EPA s'ajusta a la piràmide de població per sexe i edat de cada comunitat autònoma, el total de població de cada província i el total de població per nacionalitat en cada comunitat autònoma (distingint en aquest cas entre espanyols i estrangers). El procés de calibratge permet equilibrar la mostra, en el sentit que les estimacions d'aquestes variables a la mostra coincidiran aproximadament amb el valor observat en les fonts externes.

A l'hora de valorar l'impacte del calibrat sobre les xifres de l'EPA es poden tenir en compte dues qüestions. La primera, relacionada amb el grau en el qual s'aconsegueix l'objectiu perseguit en el procés de calibratge. La segona, sobre si l'estructura de pesos que es genera en el procés de calibratge ocasiona que alguns segments de la mostra estiguin hiperrepresentats. La primera qüestió pot valorar-se comprovant si efectivament la informació elevada a població està equilibrada respecte a les diferents fonts externes emprades. Addicionalment, aquesta comprovació du aparellada a més l'anàlisi de quins són els segments sobre els quals té major incidència la reponderació, la qual cosa enllaça amb la segona qüestió. Per contestar a aquesta darrera, és necessari realitzar una anàlisi de la importància relativa dels factors d'elevació en alguns segments de la mostra, valorant quin és l'impacte relatiu dels elements mostrals quan s'eleva la seva informació a nivell poblacional.

4. 1. Comparació amb fonts externes

El procés de calibrat empra com a fonts externes les següents variables a cada comunitat autònoma:

- Projeccions de població per grups d'edat i sexe (22 grups) quinquennals.
- Projeccions de població per nacionalitat (espanyols i estrangers).
- Totals de població per província.

Un dels objectius del calibrat és garantir que la suma de les observacions ponderades coincideixi amb els valors de les fonts externes. El procés garanteix, per tant, que quan s'eleva a la població els resultats de l'EPA, coincidiran amb les fonts externes la informació agregada dels diferents segments. S'ha de tenir en compte que, en principi,

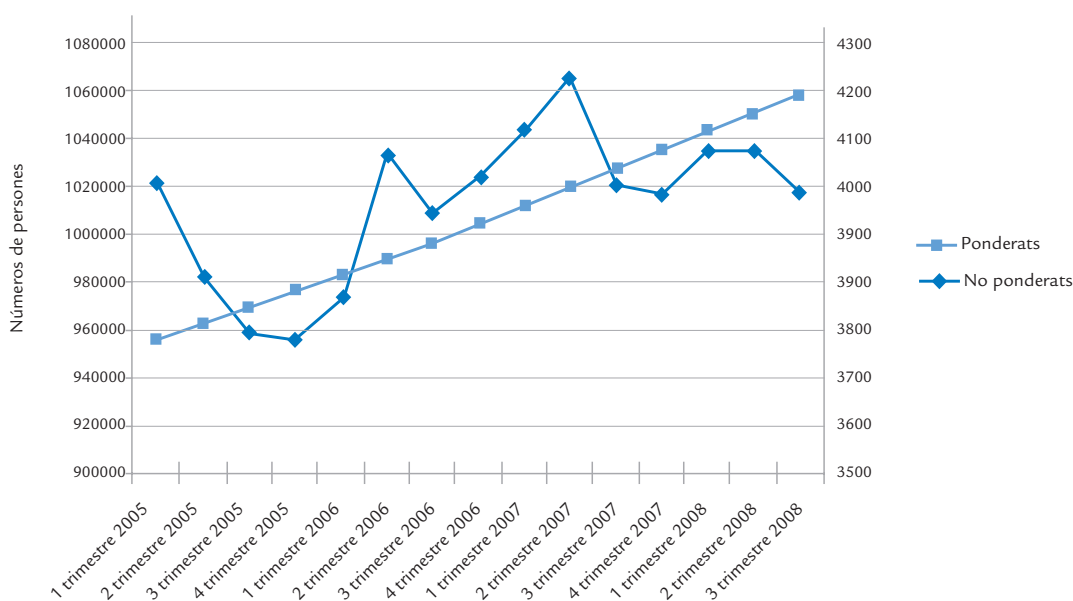
el procés de calibrat hauria de garantir el mateix equilibri en els segments creuats de la variable. És a dir, els resultats per edat i sexe (22 grups) s'han d'equilibrar amb la font externa, igual que el total de població o la quantitat de població espanyola i estrangera.

4.2. Població

Un dels principals criteris d'equilibri definit per l'INE és el de població. Per al conjunt de comunitats autònomes el procés de calibrat té en compte el total de població per província. En el cas de comunitats amb una única província no s'estableix un agregat de menor entitat. En el cas de les Illes Balears, es considera únicament el total de la comunitat, ignorant la divisió territorial definida per la insularitat.

Les projeccions de població que realitza l'INE a les Balears per al període 1-2005 a 3-2008 suposen un increment de la població des de 956.673 a 1.057.259 persones. Això implica una taxa trimestral de creixement mitjà acumulatiu del 0,62%, amb un increment final de la població del 10,51% en el conjunt del període. Malgrat el creixement de la població, la mida de la mostra corresponent a les Balears s'ha mantingut la majoria dels trimestres entorn d'unes 4.000 persones. En la Gràfica 4 pot observar-se l'evolució de la mida de la mostra (dades no ponderades) i de la projecció de població (dades ponderades a partir dels factors d'elevació) durant el període.

GRÀFICA 4. EVOLUCIÓ DE LA MIDA MOSTRAL (DADES NO PONDERADES) I DE LA POBLACIÓ REPRESENTADA (DADES PONDERADES) A L'EPA PER A LES ILLES BALEARS

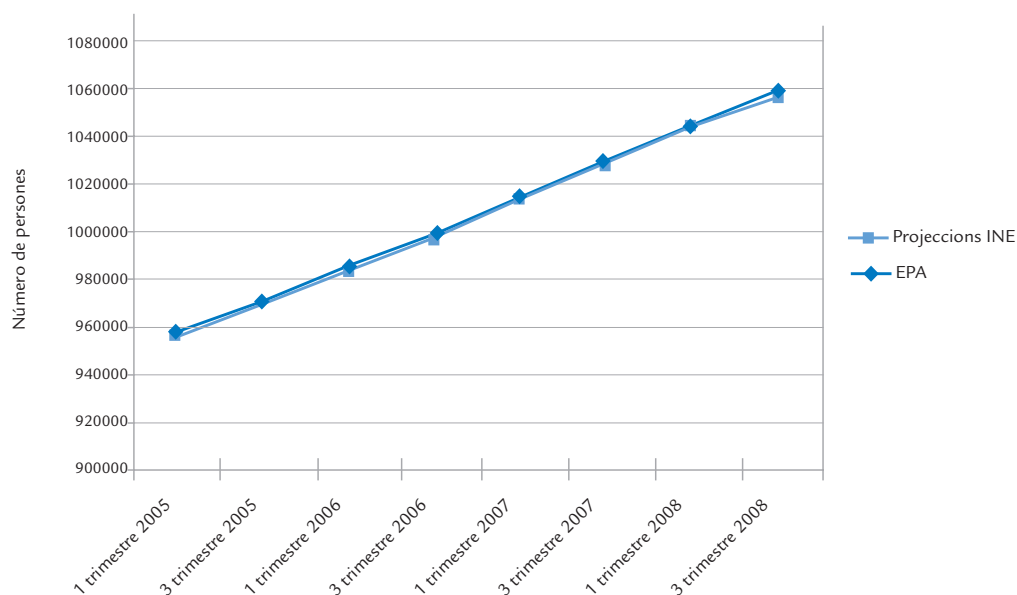


El Quadre 1 mostra l'evolució de la població de les Illes (per a l'1 de gener i l'1 de juliol del període 2005-2008) segons és estimada a partir de l'EPA i les projeccions de l'INE de població, calculades a partir del Cens de 2001 (font: INE, Estimacions de la població actual, ePOBa). Les dades es representen a la Gràfica 5, resultant evident que el procés de calibratge compleix l'objectiu d'equilibri en el total de població de la comunitat autònoma, quan es compara amb l'ePOBa.

QUADRE 1. EVOLUCIÓ DE LA POBLACIÓ DE LES ILLES BALEARS, 2005-2008

Població Illes Balears		
	EPA - Ponderats	Projeccions INE
1 trimestre 2005	956.673	957.953
3 trimestre 2005	970.188	971.779
1 trimestre 2006	984.223	985.620
3 trimestre 2006	997.000	998.054
1 trimestre 2007	1.014.250	1.014.405
3 trimestre 2007	1.027.886	1.028.635
1 trimestre 2008	1.043.845	1.045.008
3 trimestre 2008	1.057.259	1.058.668

GRÀFICA 5. POBLACIÓ I PROJECCIONS DE POBLACIÓ PER A LES ILLES BALEARS, 2005-2008

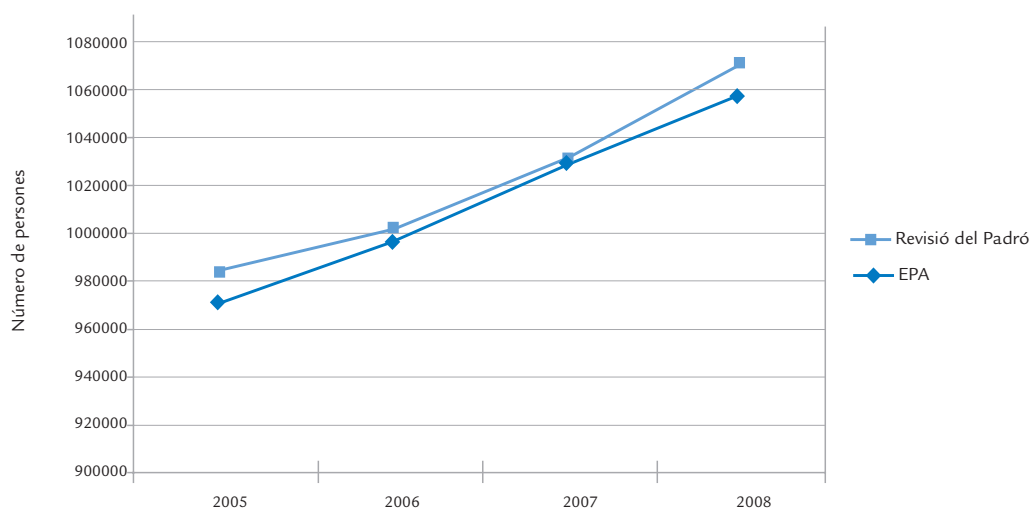


Malgrat tot, l'objectiu d'equilibri no s'assoleix quan es compara la població de l'EPA amb les revisions del padró municipal amb referència a 1 de gener de cada any (INE. Revisió del padró municipal). Les dades de l'INE corresponents a les Projeccions de l'INE i a la Revisió del padró municipal no són totalment coincidents. Com pot comprovar-se en el Quadre 2, en el període 2005-2008, les Projeccions de l'INE semblen infravalorar la població de les Balears. De manera equivalent, la població estimada que s'estima en l'EPA és una mica menor (vegi's la Gràfica 6).

CUADRO 2. POBLACIÓ DE LES ILLES BALEARS SEGONS FONTS, 2005-2008

Població Illes Balears		
	Projeccions INE	Revisió Padró Municipal
01-gener-05	957.953	983.131
01-gener-06	985.620	1.001.062
01-gener-07	1.014.405	1.030.650
01-gener-08	1.045.008	1.071.221

GRÁFICA 6. POBLACIÓ DE LES ILLES BALEARS SEGONS EL PADRÓ MUNICIPAL I L'EPA, 2005-2008



4.3. Nacionalitat espanyola i estrangera

Respecte a les projeccions de la població derivades del Cens de 1991, una de les principals característiques de les noves projeccions és la revisió de les hipòtesis sobre immigració. Aquesta revisió s'introdueix el 2005 i afecta no només al nombre

absolut d'immigrants (que s'havien infraestimat anteriorment), sinó que modifica els factors d'elevació amb l'objectiu de compensar les dificultats d'entrevistar la població immigrant. Segons l'INE, els factors d'elevació tenen en compte les dificultats d'entrevistar la població immigrant com a conseqüència de diverses causes: una part no resideix a llars familiars, la seva situació és en ocasions irregular, o les dificultats de l'idioma que dificulten l'entrevista. Un dels objectius de la revisió l'any 2005 de la metodologia de l'EPA era reflectir millor el pes específic dels estrangers a determinats sectors o professions. Així i tot, l'INE no publica les projeccions de població corresponents a les diferents nacionalitats. Per això, per conèixer la distribució per nacionalitats s'ha de recórrer a les revisions del padró municipal, respecte al total del qual, com s'ha indicat, l'EPA manté una infravaloració.

per l'INE per al període 2005 a 2008. Les dades corresponents a l'EPA es mostren en el Quadre 4. Pot observar-se que entre 2005 i 2008 la població de nacionalitat espanyola ha augmentat en 22.029 persones (una taxa anual mitjana acumulativa del 0,88 %), mentre que la de nacionalitat estrangera ha crescut en 66.061 persones (amb una taxa anual mitjana acumulativa del 12,47 %), la qual cosa suposa un increment de població de 88.090 persones (taxa mitjana del 2,9 %). Les dades de l'EPA mostren increments de 6.320 persones de nacionalitat espanyola, 80.852 de nacionalitat estrangera, amb un total de 87.172. Aquests increments suposen taxes mitjanes acumulatives del 0,26 %, 16,09 % i 2,95 %, respectivament (vegi's Quadre 5).

QUADRE 3. REVISIÓ DEL PADRÓ MUNICIPAL DE L'INE PER A LES ILLES BALEARS, 2005-2008

	Nombre de persones				Percentatge			
	2005	2006	2007	2008	2005	2006	2007	2008
Espanyols	826.861	833.311	840.480	848.890	84,1	83,2	81,6	79,3
Estrangers	156.270	167.751	190.170	222.331	15,9	16,8	18,5	20,8
Total	983.131	1.001.062	1.030.650	1.071.221	100	100	100	100

Font: Revisió del padró municipal (a 1 de gener de cada any).

QUADRE 4. EPA. DADES PONDERADES DE POBLACIÓ PER A LES ILLES BALEARS, 2005-2008 (PRIMER TRIMESTRE)

	Número de persones				Percentatge			
	gener-2005	gener-2006	gener-2007	gener-2008	gener-2005	gener-2006	gener-2007	gener-2008
Espanyols	813.463	813.908	818.155	819.783	85,0	82,7	80,7	78,5
Estrangers	143.210	170.315	196.095	224.062	15,0	17,3	19,3	21,5
Total	956.673	984.223	1.014.250	1.043.845	100	100	100	100

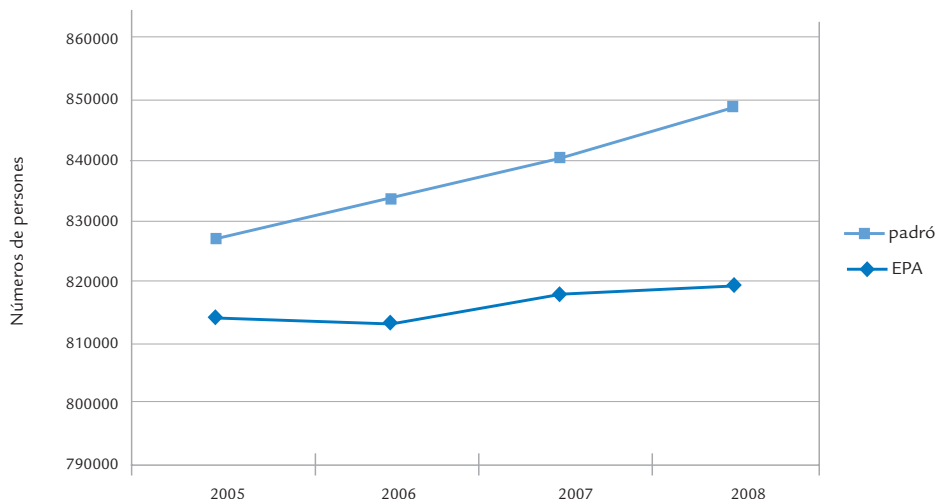
QUADRE 5. VARIACIONS DE POBLACIÓ PER A LES ILLES BALEARS, 2005-2008

	Variació 2008-2005		Taxa mitja acumulativa	
	Padró	EPA	Padró	EPA
Espanyols	22.029	6.320	0,88%	0,26%
Estrangers	66.061	80.852	12,47%	16,09%
Total	88.090	87.172	2,90%	2,95%

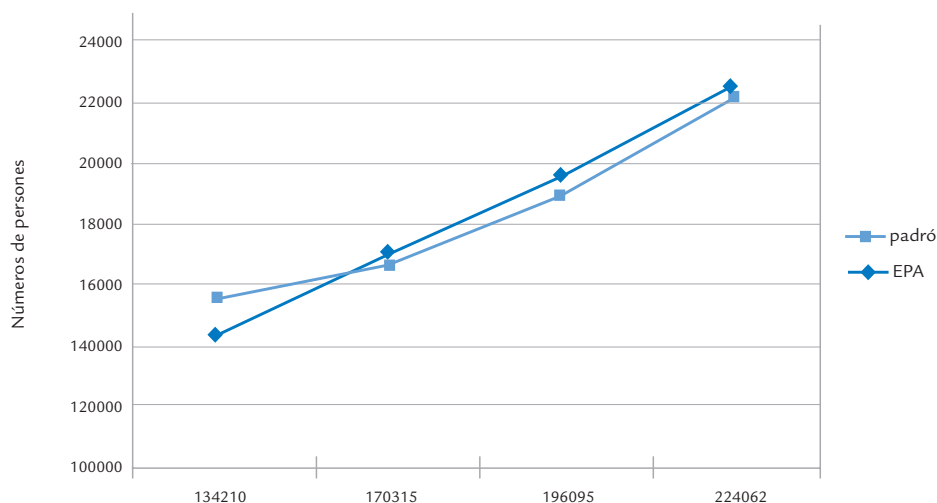
Font: INE, Revisió del padró municipal

Per tant, les xifres estimades per l'EPA (Quadre 4) difereixen de les obtingudes de la revisió del padró no solament en termes absoluts, sinó també en la seva evolució temporal. La comparació d'ambdues fonts mostra que l'EPA (vegin-se les Gràfiques 7 i 8) subestima de manera sistemàtica la població de nacionalitat espanyola, no corregint-se aquest biaix en el període (vegi's el Quadre 5). No ocorre el mateix amb la població estrangera, que sí es presenta equilibrada respecte al padró. Aquest resultat no sembla tenir una conseqüència important en la distribució estimada de la població segons nacionalitat, encara que l'EPA presenta un percentatge lleugerament major d'estrangers que el padró.

GRÀFICA 7. EVOLUCIÓ DE LA POBLACIÓ ESPANYOLA SEGONS EL PADRÓ MUNICIPAL I L'EPA, 2005-2008



GRÀFICA 8. EVOLUCIÓ DE LA POBLACIÓ ESTRANGERA SEGONS EL PADRÓ MUNICIPAL I L'EPA, 2005-2008



El pes que presenta la nacionalitat estrangera en les dades poblacionals de l'EPA són, en una part important, resultat del calibratge. L'anàlisi de les dades no ponderades mostra (vegin-se els Quadres 6 i 7 i la Gràfica 9) que el percentatge d'espanyols enquestats presenta una tendència clara a la seva disminució, passant d'un 91,37 % en el primer trimestre de 2005 a un 85,76 % en el tercer trimestre de 2008. No obstant això, la mostra sobreestima excessivament el pes de la nacionalitat espanyola. Quan les dades són elevades a nivell poblacional, la tendència de la sèrie s'aguditza a més de disminuir

de manera important el percentatge, que passen de 85,03 % (primer trimestre de 2005) a 77,64 % (tercer trimestre de 2008). Els resultats indiquen, per tant, que sembla existir un fort biaix en la representació de la població de nacionalitat estrangera a la mostra. Encara que l'ús dels factors d'elevació poblacional corregeix aquest biaix, la important diferència en els percentatges de les dades no ponderades davant els ponderats assenyala que hi ha una molt deficient representació de la població estrangera a la mostra.

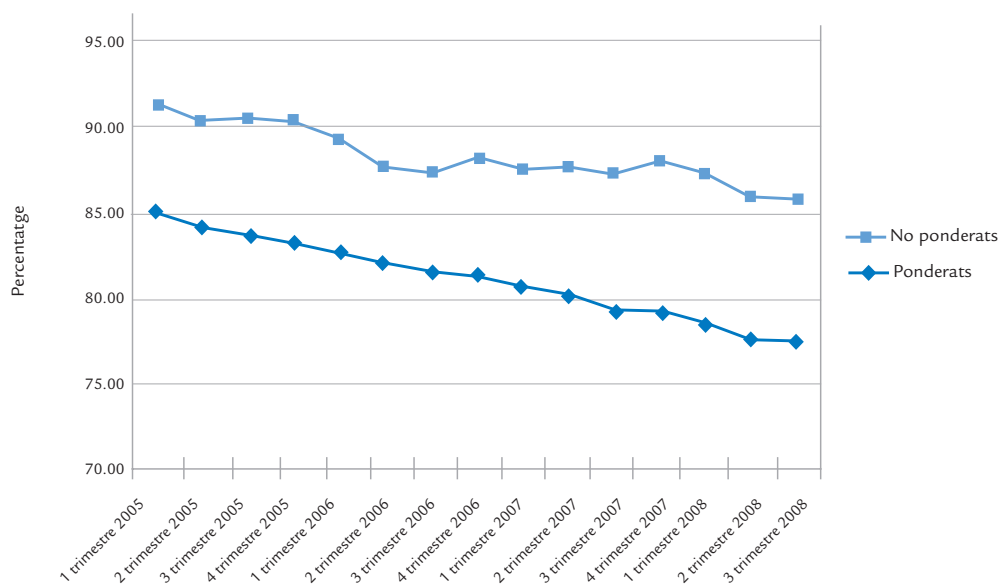
QUADRE 6. EPA (3R TRIMESTRE). DADES NO PONDERADES DE POBLACIÓ, ILLES BALEARS, 2005-2008

	2005	2006	2007	2008	2005	2006	2007	2008
Espanyols	3.407	3.481	3.525	3.452	90,42	87,33	87,27	85,76
Estrangers	361	505	514	573	9,58	12,67	12,73	14,23
Total	3.768	3.986	4.039	4.025	100	100	100	100

QUADRE 7. EPA. PERCENTATGE DE POBLACIÓ PER NACIONALITAT, ILLES BALEARS, 2005-2008

	Dades no ponderades		Dades ponderades	
	Espanyols	Estrangers	Espanyols	Estrangers
1 trimestre 2005	91,37	8,63	85,03	14,97
2 trimestre 2005	90,25	9,75	84,24	15,76
3 trimestre 2005	90,42	9,58	83,73	16,27
4 trimestre 2005	90,30	9,70	83,44	16,56
1 trimestre 2006	89,33	10,67	82,70	17,30
2 trimestre 2006	87,61	12,39	81,94	18,06
3 trimestre 2006	87,33	12,67	81,43	18,57
4 trimestre 2006	88,11	11,89	81,34	18,66
1 trimestre 2007	87,47	12,53	80,67	19,33
2 trimestre 2007	87,60	12,40	80,20	19,80
3 trimestre 2007	87,27	12,73	79,35	20,65
4 trimestre 2007	88,03	11,97	79,32	20,68
1 trimestre 2008	87,19	12,81	78,53	21,47
2 trimestre 2008	85,76	14,24	77,64	22,36
3 trimestre 2008	85,76	14,24	77,64	22,36

GRÀFICA 9. PERCENTATGE DE PERSONES DE NACIONALITAT ESPANYOLA A L'EPA. DADES PONDERADES I NO PONDERADES DE LES ILLES BALEARS, 2005-2008



4.4. Nacionalitats estrangeres

Segons les revisions del Padró Municipal, la distribució de les diferents nacionalitats a les Illes presenta un pes important dels estrangers comunitaris (50,65 % dels estrangers el 2007) i aquells nascuts a Amèrica del Sud (27,92 %). L'EPA no presenta xifres equilibrades en aquesta distribució (vegi's la Gràfica 10). Els valors poblacionals de l'EPA assenyalen també les anteriors zones com les més importants, però amb percentatges molt diferents, representant els estrangers comunitaris en el tercer trimestre de 2008 un 25,24 % dels estrangers, mentre que l'origen sud-americà correspon a un 51,63 %. A més, s'ha de considerar la gran variabilitat de les xifres estimades a l'EPA. En els Quadres 8 (dades ponderades) i 9 (dades no ponderades) es mostren els resultats corresponents al tercer trimestre de l'EPA dels anys 2006, 2007 i 2008. Pot observar-se que els percentatges corresponents a les diferents nacionalitats pateixen fortes variacions. Aquesta variabilitat en la composició per nacionalitats no està recolzada per les dades de les revisions del padró (vegi's el Quadre 10), per la qual cosa s'ha d'atribuir al mostreig.

QUADRE 8. EPA. PERSONES NASCUTES A L'ESTRANGER. DADES PONDERADES. (TERCER TRIMESTRE)

	Número de persones			Percentatges		
	2006	2007	2008	2006	2007	2008
Reste d'Europa (no UE)	15.647	20.573	22.836	8,45	9,69	9,66
UE-15	55.652	57.926	57.609	30,06	27,29	24,37
UE-25 (no UE-15)	2.701	996	2.063	1,46	0,47	0,87
Europa	74.000	79.495	82.508	39,97	37,45	34,90
Àfrica	36.245	29.610	22.798	19,57	13,95	9,65
Amèrica del Nort	2.193	1.397	288	1,18	0,66	0,12
Centreamèrica i Carib	3.009	6.911	5.122	1,63	3,26	2,17
Sudamèrica	63.404	91.690	122.028	34,24	43,19	51,63
Àsia Oriental	3.364	2.407	2.240	1,82	1,13	0,95
Àsia del Sud i Sudest	2.790	771	900	1,51	0,36	0,38
Oceania	161	0	470	0,09	0,00	0,20
Total	185.166	212.281	236.354	100	100	100

QUADRE 9. EPA. PERSONES NASCUTES A L'ESTRANGER. DADES NO PONDERADES. (TERCER TRIMESTRE)

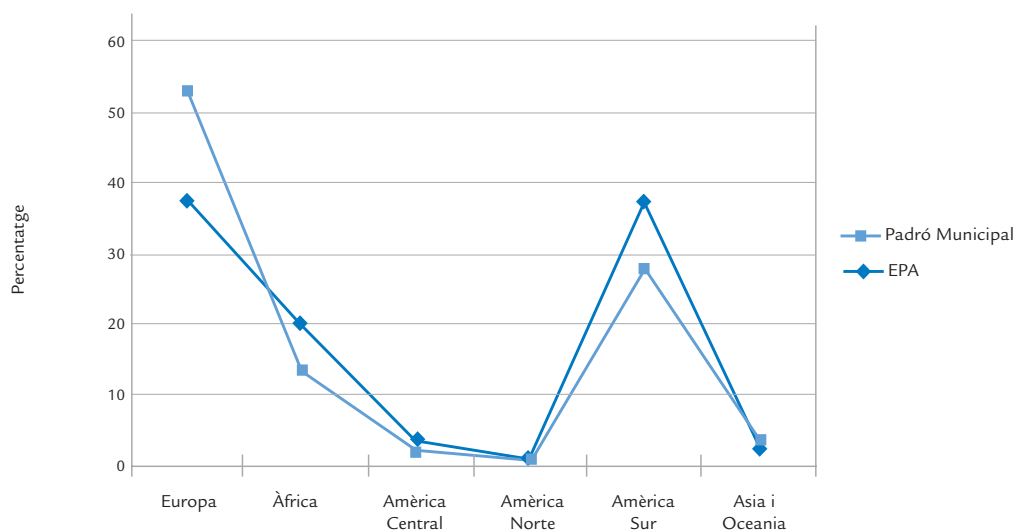
	Número de persones			Percentatges		
	2006	2007	2008	2006	2007	2008
Reste d'Europa (no UE)	37	48	48	7,33	9,34	8,38
UE-15	163	147	139	32,28	28,60	24,26
UE-25 (no UE-15)	6	3	5	1,19	0,58	0,87
Europa	206	198	192	40,80	38,52	33,51
Àfrica	90	65	50	17,82	12,65	8,73
Amèrica del Nort	5	3	1	0,99	0,58	0,17
Centreamèrica i Carib	10	21	16	1,98	4,09	2,79
Sudamèrica	176	218	305	34,85	42,41	53,23
Àsia Oriental	9	7	6	1,78	1,36	1,05
Àsia del Sud i Sudest	8	2	2	1,58	0,39	0,35
Oceania	1	0	1	0,20	0,00	0,17
Total	505	514	573	100	100	100

QUADRE 10. INE. PERSONES NASCUTES A L'ESTRANGER

	Número de persones		Percentatges	
	2006	2007	2006	2007
Unió Europea (25 y 27)	72.637	96.329	43,30	50,65
Reste Europa	13.469	4.605	8,03	2,42
Europa	86.106	100.934	51,33	53,07
Àfrica	23.666	25.495	14,11	13,41
Amèrica Central	3.419	3.784	2,04	1,99
Amèrica Nort	1.197	1.247	0,71	0,66
Amèrica Sud	48.135	53.101	28,69	27,92
Àsia	5.060	5.453	3,02	2,87
Oceania	168	155	0,10	0,08
Total	167.751	190.169	100	100

Font: Revisió del padró municipal

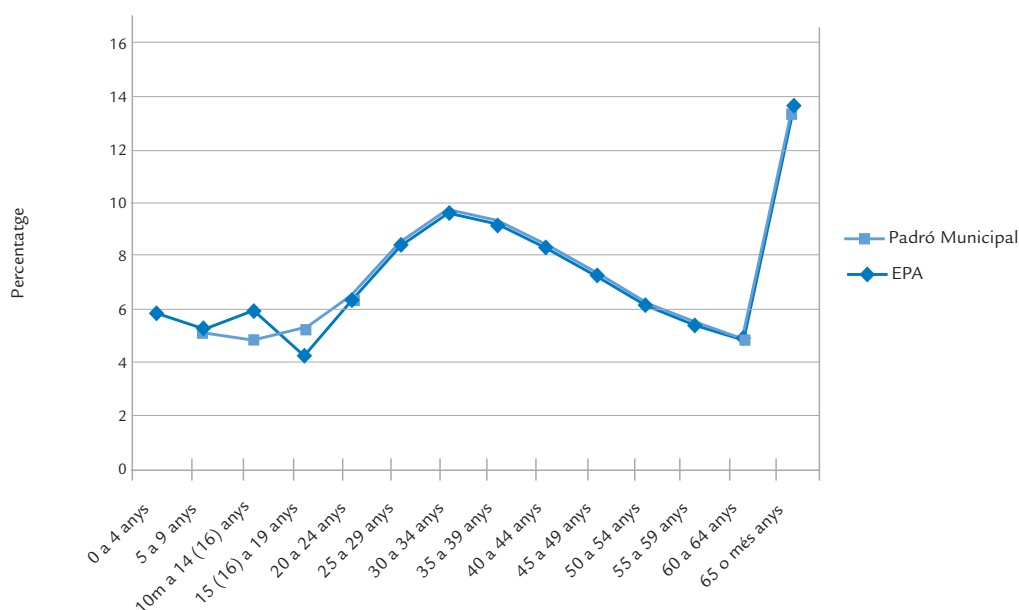
GRÀFICA 10. COMPARACIÓ ENTRE L'EPA (PRIMER TRIMESTRE 2007) I LA REVISIÓ DEL PADRÓ MUNICIPAL (1 DE GENER 2007) DE L'ORIGEN DE LES PERSONES DE NACIONALITAT ESTRANGERA, ILLES BALEARS



4.5. Edat

Com il·lustra la Gràfica 11, amb les dades de l'any 2008, la distribució per edats de l'EPA està equilibrada respecte a l'obtinguda en les revisions del padró municipal. En els Quadres 11 i 12 es mostren les distribucions d'edat corresponents al padró municipal i a l'EPA, per als anys 2005 a 2008.

GRÀFICA 11. DISTRIBUCIÓ PER EDATS SEGONS PADRÓ MUNICIPAL (1 GENER 2008) I EPA (PRIMER TRIMESTRE 2008) A LES ILLES BALEARS



QUADRE 11. DISTRIBUCIÓ DE LA POBLACIÓ PER EDATS A LES ILLES BALEARS SEGONS EL PADRÓ MUNICIPAL, 2005-2008

	2005	2006	2007	2008
0 a 4 anys	5,61	5,68	5,77	5,75
5 a 9 anys	4,98	5,03	5,07	5,18
10 a 15 anys	5,11	5,00	4,91	4,87
16 a 19 anys	5,43	5,34	5,27	5,22
20 a 24 anys	6,81	6,60	6,43	6,33
25 a 29 anys	8,95	8,80	8,59	8,39
30 a 34 anys	9,50	9,60	9,64	9,62
35 a 39 anys	8,81	8,91	9,02	9,16
40 a 44 anys	7,94	8,07	8,19	8,26
45 a 49 anys	6,90	7,01	7,11	7,23
50 a 54 anys	6,01	6,03	6,07	6,14
55 a 59 anys	5,51	5,50	5,45	5,41
60 a 64 anys	4,52	4,65	4,76	4,82
65 o més anys	13,92	13,79	13,71	13,62
Total	100	100	100	100

Font: INE, Revisió del padró municipal

QUADRE 12. DISTRIBUCIÓ DE LA POBLACIÓ PER EDATS A LES ILLES BALEARS SEGONS L'EPA AMB LES DADES PONDERADES, 2005-2008

	2005	2006	2007	2008
0 a 4 anys	5,64	5,71	5,78	5,76
5 a 9 anys	5,00	5,05	5,10	5,20
10 a 15 anys	6,16	6,04	5,94	5,88
16 a 19 anys	4,39	4,30	4,25	4,22
20 a 24 anys	6,79	6,59	6,43	6,33
25 a 29 anys	8,94	8,78	8,59	8,38
30 a 34 anys	9,54	9,63	9,66	9,63
35 a 39 anys	8,84	8,95	9,06	9,19
40 a 44 anys	7,98	8,11	8,22	8,30
45 a 49 anys	6,93	7,04	7,14	7,26
50 a 54 anys	6,03	6,05	6,09	6,17
55 a 59 anys	5,51	5,50	5,45	5,41
60 a 64 anys	4,54	4,67	4,77	4,83
65 o més anys	13,70	13,59	13,50	13,43
Total	100	100	100	100

Font: EPA. Primer trimestre de cada any

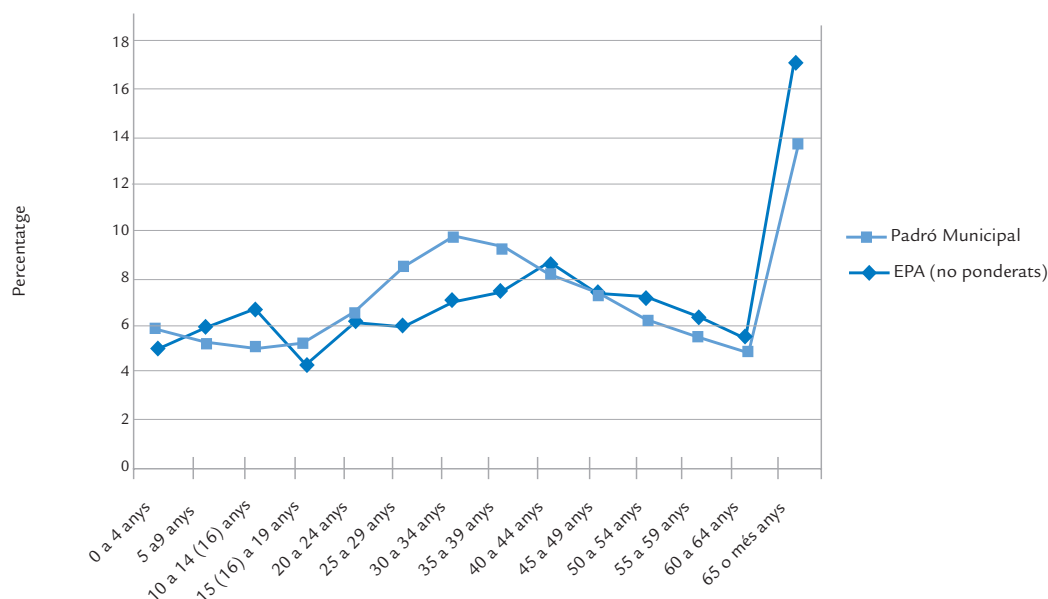
QUADRE 13. DISTRIBUCIÓ PER EDATS A LA MOSTRA DE L'EPA DE LES ILLES BALEARS AMB LES DADES NO PONDERADES, 2005-2008

	2005	2006	2007	2008
0 a 4 anys	5,62	5,61	5,17	5,05
5 a 9 anys	4,92	5,24	5,92	5,81
10 a 15 anys	6,33	6,50	6,28	6,59
16 a 19 anys	5,07	5,06	4,97	4,29
20 a 24 anys	6,09	5,32	5,36	6,20
25 a 29 anys	7,47	6,37	5,92	6,00
30 a 34 anys	7,94	7,36	7,94	6,90
35 a 39 anys	7,72	8,35	8,57	7,37
40 a 44 anys	7,77	7,75	7,97	8,59
45 a 49 anys	7,57	7,83	7,61	7,32
50 a 54 anys	6,98	6,37	5,84	7,05
55 a 59 anys	5,67	5,27	5,46	6,22
60 a 64 anys	4,95	4,93	5,53	5,54
65 o més anys	15,91	18,05	17,46	17,08
Total	100	100	100	100

Font: EPA. Primer trimestre de cada any. Dades no ponderades

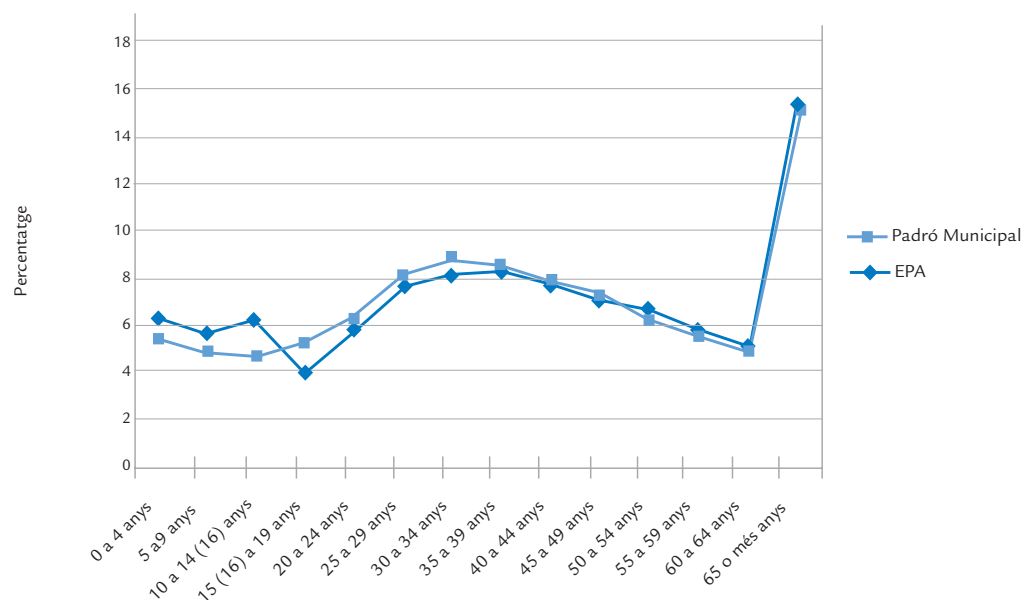
No obstant això, si s'analitza la distribució per edats amb les dades no ponderades de l'EPA (Quadre 13) es detecta un important biaix. A la Gràfica 12 es comparen les distribucions d'edat que s'obtenen de l'EPA (primer trimestre 2008), emprant les dades sense ponderar, amb les del padró municipal (a 1 de gener de 2008). Pot observar-se que prèviament a la seva ponderació, la distribució per edats que proporciona la mostra de l'EPA no estima correctament l'estructura d'edats. Els factors d'elevació poblacional asseguren que, en elevar les dades a població, la distribució per edats sigui coincident. No obstant això, el fet és que el disseny mostral de l'EPA subestima la importància del col·lectiu entre 25 i 39 anys (sobre aquest tema, pot consultar-se Álvarez, 2000). Aquesta subestimació posa de nou en evidència el problema de l'EPA per entrevistar a determinades llars.

GRÀFICA 12. DISTRIBUCIÓ PER EDATS SEGONS PADRÓ MUNICIPAL (1 GENER 2008) I DADES NO PONDERADES DE L'EPA (PRIMER TRIMESTRE 2008)

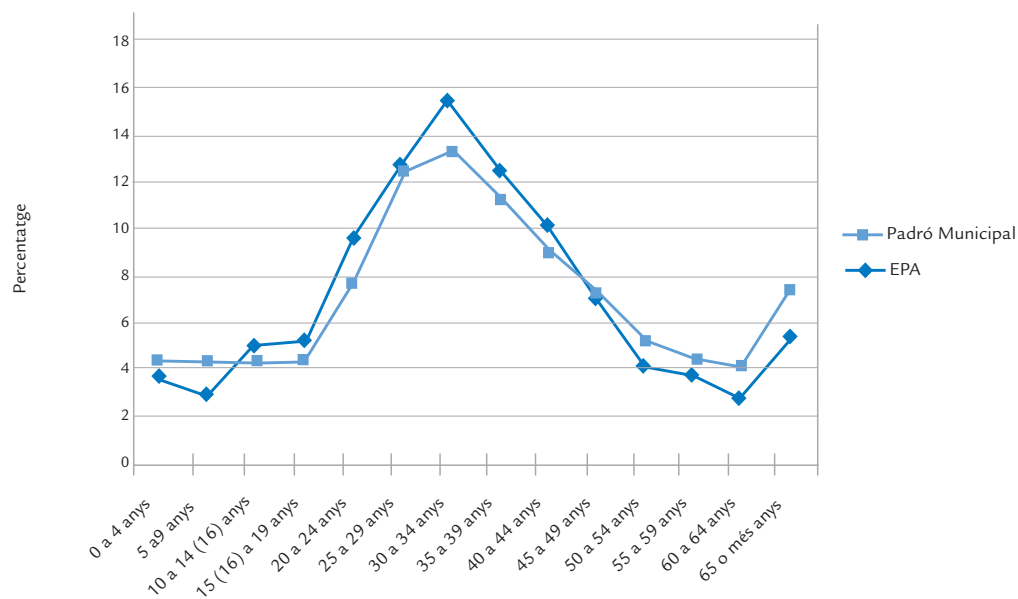


L'anàlisi de la distribució per edat pot fer-se distingint entre la població de nacionalitat espanyola i estrangera. En aquest cas, es posa en evidència que la subestimació detectada en el total poblacional no és uniforme en les dues subpoblacions. Com es mostra a les Gràfiques 13 i 14, mentre que per a la nacionalitat espanyola l'EPA subestima en una lleugera quantia el tram entre 20 i 34 anys, en el cas de la nacionalitat estrangera l'EPA tendeix a sobreestimar la població en els trams d'edat intermèdia (entre 20 i 44 anys).

GRÀFICA 13. DISTRIBUCIÓ PER EDATS DE LA SUBPOBLACIÓ ESPANYOLA, SEGONS PADRÓ MUNICIPAL (1 GENER 2007) I DADES NO PONDERADES DE L'EPA (PRIMER TRIMESTRE 2007)

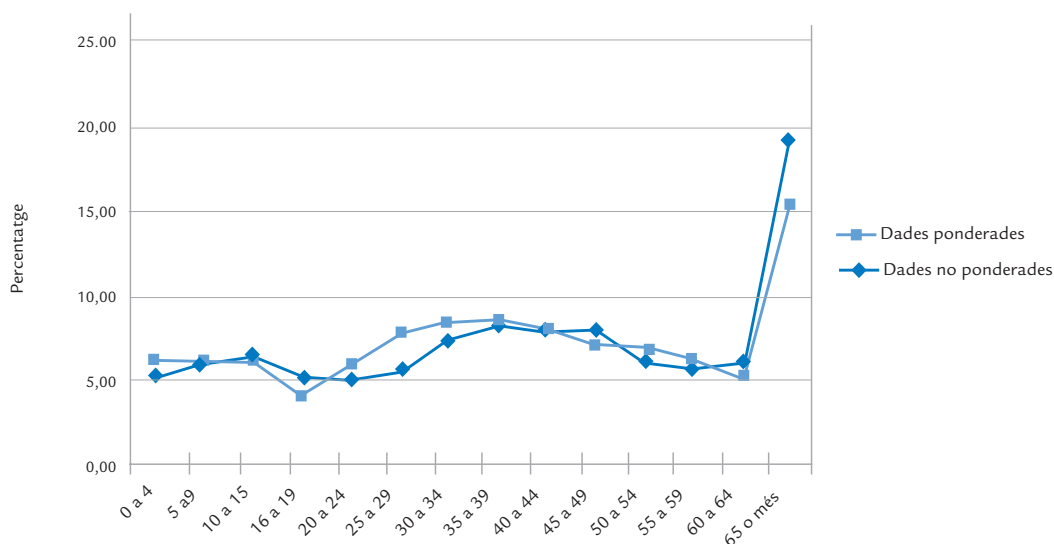


GRÀFICA 14. DISTRIBUCIÓ PER EDATS DE LA SUBPOBLACIÓ ESTRANGERA, SEGONS PADRÓ MUNICIPAL (1 GENER 2007) I DADES NO PONDERADES DE L'EPA (PRIMER TRIMESTRE 2007)

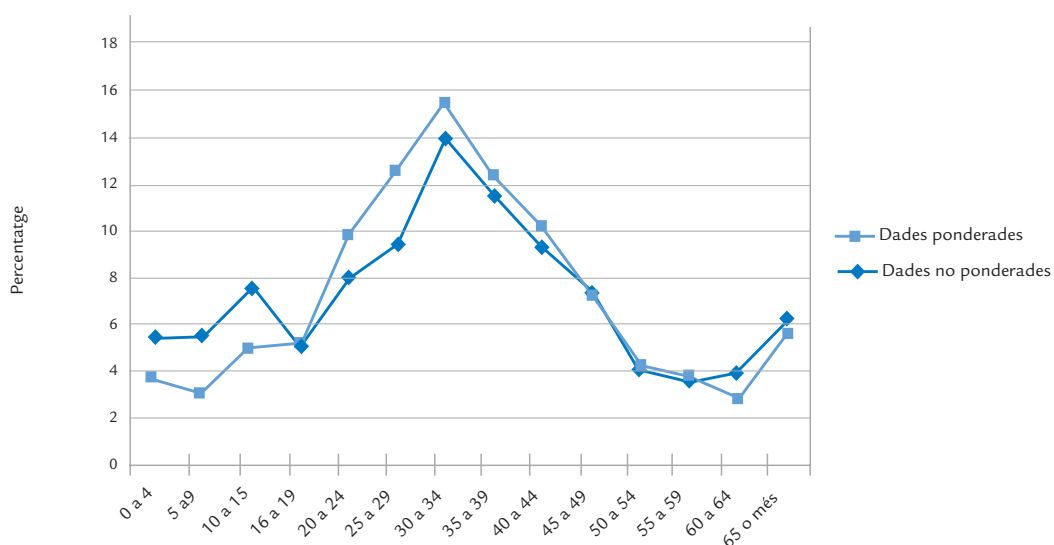


Resulta curiós observar que el procés de ponderació actua d'una manera perversa en el cas de la població estrangera. A la Gràfica 15 es comparen, per a la nacionalitat espanyola, les distribucions de l'EPA amb dades ponderades i sense ponderar. En aquest cas, l'elevació de les dades a la població augmenta el pes dels trams d'edat que estarien subestimats respecte al padró. Tanmateix, en el cas de la nacionalitat estrangera, les dades ponderades (Gràfica 16) augmenten la representació dels trams que s'han detectat com a sobrerepresentats respecte al padró. De fet, la distribució de les dades no ponderades s'ajusta millor a les dades del padró, que la que s'obté després d'elevat les dades (Gràfica 17).

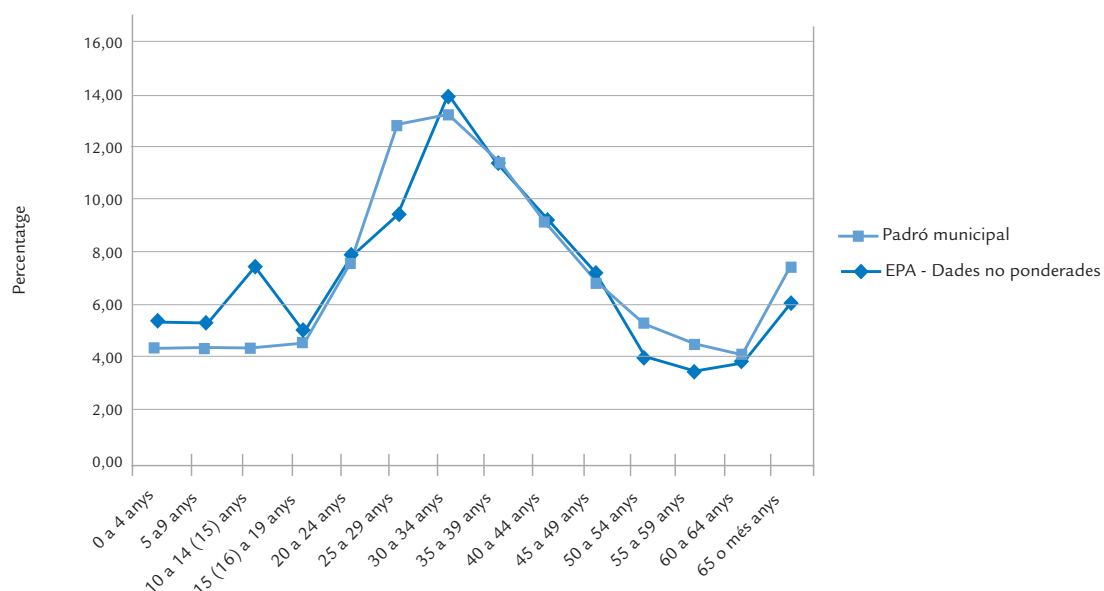
GRÀFICA 15. DISTRIBUCIÓ PER EDATS DE LA SUBPOBLACIÓ ESPANYOLA, SEGONS L'EPA (PRIMER TRIMESTRE 2007)



GRÀFICA 16. DISTRIBUCIÓ PER EDATS DE LA SUBPOBLACIÓ ESTRANGERA, SEGONS L'EPA (PRIMER TRIMESTRE 2007)



GRÀFICA 17. DISTRIBUCIÓ PER EDATS DE LA SUBPOBLACIÓ ESTRANGERA, SEGONS EL PADRÓ MUNICIPAL (1 DE GENER DE 2007) I L'EPA
(DADES NO PONDERADES, PRIMER TRIMESTRE 2007)



4.6. Sexe

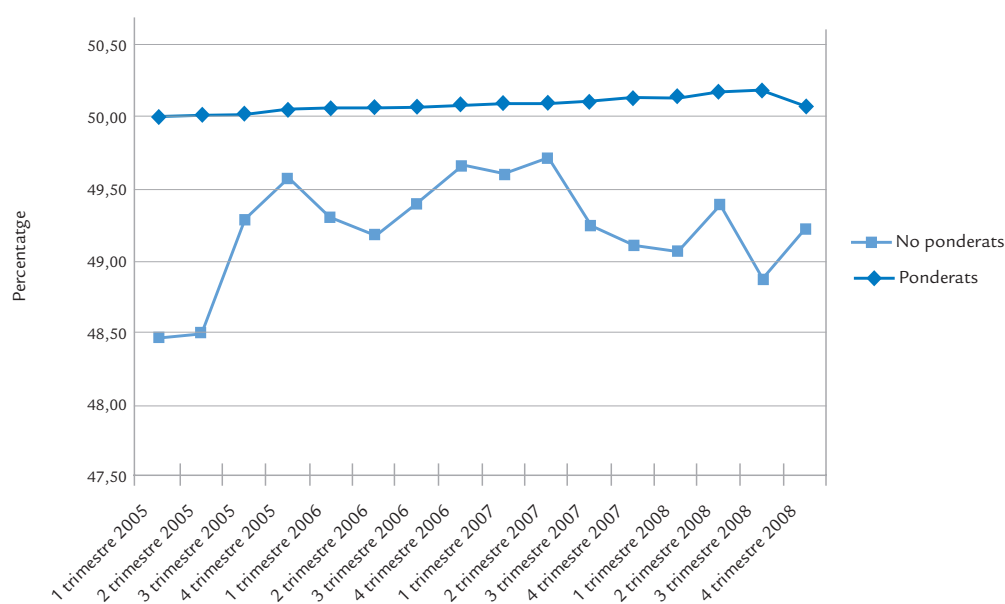
La mostra de l'EPA a l'àmbit nacional sobreestima la proporció de dones (51,6 %, en 2007-2008) davant el padró municipal (50,6 %). Una situació semblant es dona a la mostra de les Balears, en la qual la mostra de l'EPA (vegin-se el Quadre 14 i la Gràfica 18) proporciona (en el tercer trimestre de 2008) un 51,13 % de dones. L'elevació de la mostra a població corregeix aquesta proporció, situant-la en 49,8 %, igual al percentatge que es recull al padró.

QUADRE 14. PERCENTATGE DE POBLACIÓ PER SEXES DE LES ILLES BALEARS

	Dades no ponderades		Dades ponderades	
	Homes	Dones	Homes	Dones
1T-2005	48,47	51,53	50,00	50,00
2T-2005	48,49	51,51	50,00	50,00
3T-2005	49,28	50,72	50,02	49,98
4T-2005	49,57	50,43	50,04	49,96
1T-2006	49,31	50,69	50,06	49,94
2T-2006	49,19	50,81	50,06	49,94
3T-2006	49,40	50,60	50,07	49,93
4T-2006	49,67	50,33	50,08	49,92
1T-2007	49,60	50,40	50,09	49,91
2T-2007	49,73	50,27	50,09	49,91
3T-2007	49,24	50,76	50,11	49,89
4T-2007	49,10	50,90	50,13	49,87
1T-2008	49,06	50,94	50,15	49,85
2T-2008	49,39	50,61	50,16	49,84
3T-2008	48,87	51,13	50,18	49,82

Font: EPA

GRÀFICA 18. PERCENTATGE D'HOMES A L'EPA, 2005-2008



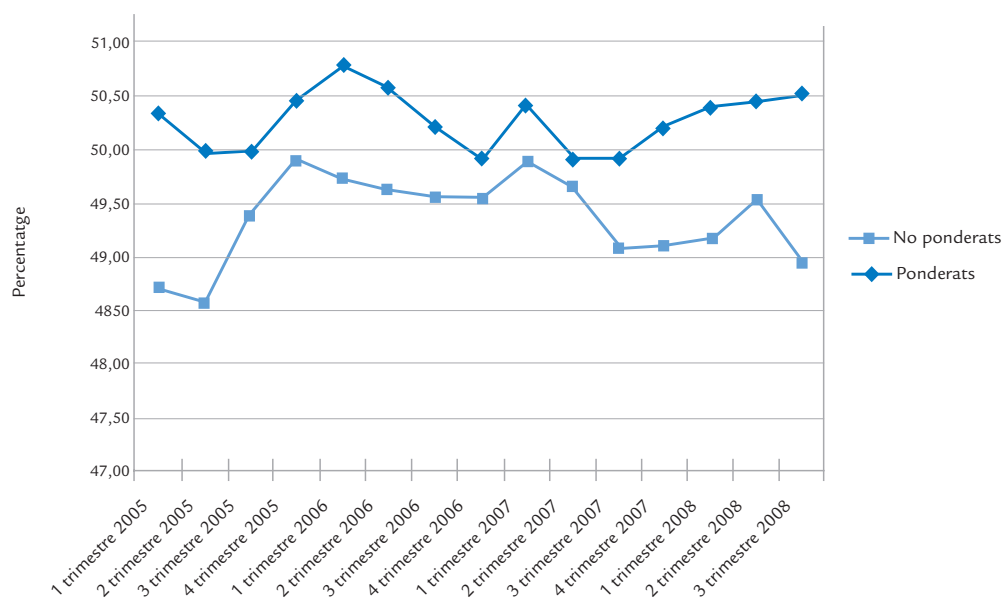
La distribució per sexes quan es distingeix per nacionalitats (espanyols i estrangers) només s'equilibra respecte al padró municipal per a la nacionalitat espanyola (vegi's el Quadre 15). La població estrangera, encara després de ponderar la mostra, presenta a l'EPA un menor percentatge d'homes. Al padró municipal, del total d'estrangers, el 51,5 % són homes. A l'EPA, una vegada ponderada la mostra, aquest percentatge és del 48,92%.

QUADRE 15. POBLACIÓ DE LES ILLES BALEARS PER SEXES, 2007

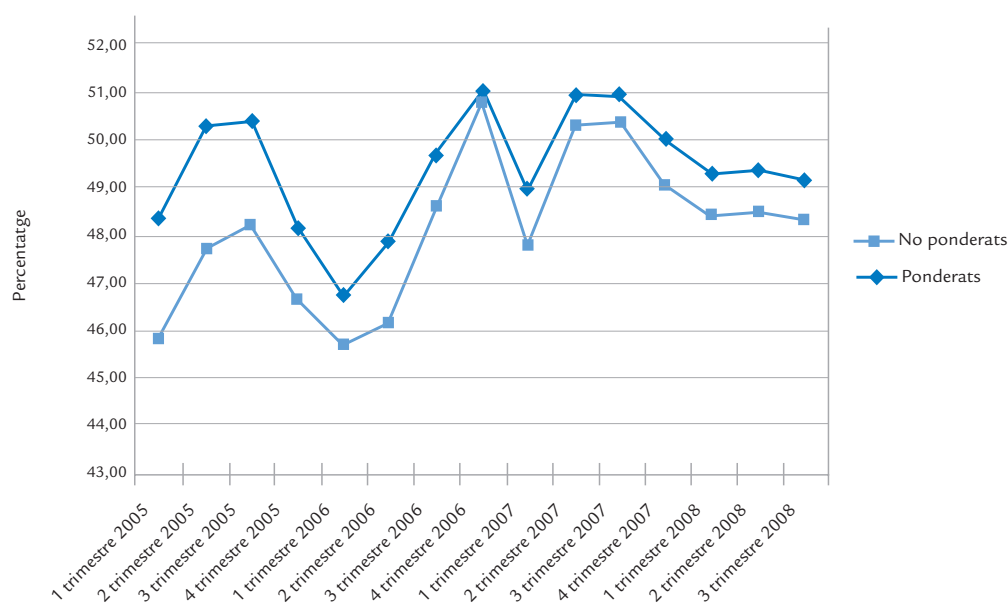
	Revisió padró 2007 (1 gener 2007)				EPA (1 trimestre 2007)			
	Homes	Dones	Total	% Homes	Homes	Dones	Total	% Homes
Total	517.593	513.057	1.030.650	50,22	508.007	506.243	1.014.250	50,09
Espanyols	419.465	421.015	840.480	49,91	412.076	406.078	818.154	50,37
Estrangers	98.128	92.042	190.170	51,60	95.931	100.165	196.096	48,92

En el cas de la nacionalitat estrangera, la falta de precisió de l'estimació de la distribució per sexe a l'EPA es mostra quan s'observa la seva variabilitat en el període analitzat (vegin-se el Quadre 16 i les Gràfiques 19 i 20). Mentre el percentatge d'homes en la subpoblació estrangera al padró es manté estable (en els anys 2005, 2006 i 2007 se situa en 51,67 %, 51,52 % i 51,60 %, respectivament), el percentatge d'homes oscil·la a l'EPA (entre 1-2005 i 3-2008) amb un mínim de 46,67 % i un màxim de 50,96 %. Similar variabilitat es veu a la mostra no ponderada (amb un mínim de 45,72 % i un màxim de 50,83 %).

GRÀFICA 19. PERCENTATGE D'HOMES DE NACIONALITAT ESPANYOLA A L'EPA, ILLES BALEARS 2005-2008



GRÀFICA 20. PERCENTATGE D'HOMES DE NACIONALITAT ESTRANGERA A L'EPA, ILLES BALEARS 2005-2008



QUADRE 16. DISTRIBUCIÓ PER SEXES, SEGONS NACIONALITAT

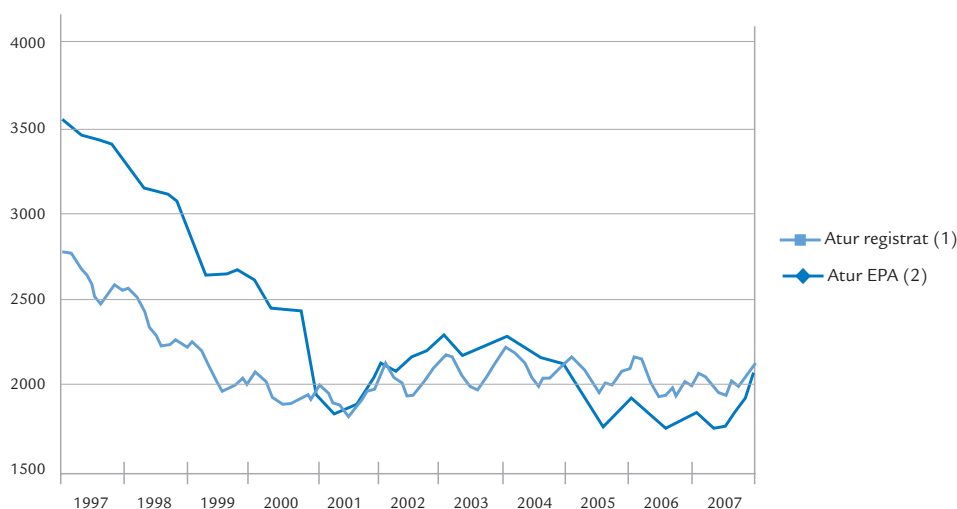
	Espanyola				Estrangera			
	Dades no ponderades		Dades ponderades		Dades no ponderades		Dades ponderades	
	Homes	Dones	Homes	Dones	Homes	Dones	Homes	Dones
1T-2005	48,71	51,29	50,29	49,71	45,85	54,15	48,34	51,66
2T-2005	48,57	51,43	49,96	50,04	47,75	52,25	50,24	49,76
3T-2005	49,40	50,60	49,96	50,04	48,20	51,80	50,35	49,65
4T-2005	49,88	50,12	50,42	49,58	46,70	53,30	48,12	51,88
1T-2006	49,74	50,26	50,76	49,24	45,72	54,28	46,67	53,33
2T-2006	49,62	50,38	50,56	49,44	46,15	53,85	47,82	52,18
3T-2006	49,53	50,47	50,17	49,83	48,51	51,49	49,62	50,38
4T-2006	49,51	50,49	49,88	50,12	50,83	49,17	50,96	49,04
1T-2007	49,86	50,14	50,37	49,63	47,78	52,22	48,92	51,08
2T-2007	49,65	50,35	49,89	50,11	50,29	49,71	50,92	49,08
3T-2007	49,08	50,92	49,90	50,10	50,39	49,61	50,89	49,11
4T-2007	49,11	50,89	50,17	49,83	49,06	50,94	49,98	50,02
1T-2008	49,16	50,84	50,39	49,61	48,38	51,62	49,27	50,73
2T-2008	49,54	50,46	50,42	49,58	48,46	51,54	49,28	50,72
3T-2008	48,96	51,04	50,49	49,51	48,34	51,66	49,10	50,90

Font: EPA

4.7. Situació laboral

Ja que l'EPA ofereix informació sobre les persones inscrites en els serveis públics d'ocupació, és possible quantificar el seu nombre per aconseguir una comparació més homogènia amb l'atur registrat. Prudencio i Toharia (2007) han analitzat l'equilibri entre l'EPA i els aturats inscrits a les oficines d'ocupació públiques, conclouent que l'EPA aproxima correctament les xifres dels serveis públics d'ocupació (SPO). No obstant això, les sèries d'atur registrat i atur EPA per a tot Espanya (Gràfica 21) mostren que, encara a partir de 2005, les sèries discrepen i ho fan en diferents sentits (Ministeri de Treball i Immigració, 2007).

GRÀFICA 21. SÈRIES D'ATUR REGISTRAT SEGONS L'SPO I L'EPA PER A TOT ESPANYA



- (1) A partir de maig de 2005 dades SISPE, les dades de períodes anteriors són estimacions en termes SISPE facilitats per l'INEM.
(2) Dades revisades amb metodologia EPA-2005. A partir de 2001 s'aplica la nova definició d'atur comunitari.

Font: Ministeri de Treball i Immigració, 2007

Les discrepàncies entre l'EPA i l'atur registrat són conegudes (vegi's, per exemple, Pérez 2008), per la qual cosa pot resultar inapropiat comparar l'EPA amb l'atur registrat. De fet, la disparitat entre ambdues fonts no permet saber a priori quina de les dues quantifica una xifra major. Com assenyala Pérez (2008) hi ha dues raons que explicarien que la xifra de l'EPA sigui major:

- (1) No tots els aturats de l'EPA cerquen feina a través de les oficines públiques d'ocupació.
- (2) No tots els inscrits com a aturats a les oficines públiques d'ocupació es consideren aturats registrats (per exemple, estudiants menors de 25 anys, gent que demana una primera ocupació, persones que demanen ocupació de curta durada o de jornada reduïda).

Però hi ha també dues raons que justificarien que l'atur registrat sigui la xifra més alta:

-
- (1) Els aturats registrats que no cerquen activament feina o no estan disponibles per fer feina no són considerats aturats per l'EPA.
- (2) L'EPA tampoc considera els aturats registrats que no han tengut contacte amb l'oficina pública d'ocupació en les quatre setmanes anteriors, amb la finalitat de trobar feina.

Tenint en compte els anteriors comentaris, a continuació es compara la informació de l'SPO sobre atur registrat i la de l'EPA. Les dades de l'EPA sobre el nombre d'aturats registrats es deriva de la següent pregunta de l'enquesta: "El diumenge de la setmana de referència, estava inscrit en una oficina d'ocupació pública?". S'ha de tenir en compte que les xifres de l'SPO corresponen a l'atur registrat el mes de setembre de 2008, mentre que les xifres de l'EPA són les del tercer trimestre de 2008.

Segons dades de l'SPO, l'atur registrat a les Illes Balears el setembre de 2008 era de 46.807, amb la distribució per sexes i edat que es mostra al Quadre 16. Segons l'EPA (3r trimestre de 2008) l'atur registrat és de 46.426. Les distribucions per edat i sexe d'ambdues fonts d'informació es poden comparar a partir dels Quadres 17 i 18. Els percentatges totals per sexe s'aproximen, encara que el percentatge estimat de dones és una mica major a l'EPA (45,47 %) que a l'SPO (43,3 %). La distribució per edats no coincideix. A l'EPA el percentatge d'aturats registrats menors de 25 anys sobrepassa en molt el percentatge de l'SPO (24,25 % a l'EPA, davant el 14,33 % de l'SPO). Els percentatges corresponents als trams de més edat estan subestimats a l'EPA: en el tram de 25 a 44 anys, 48,89 % (EPA) i 55,09 % (SPO); en el tram de més de 45 anys 26,86 % (EPA) i 30,57 % (SPO).

QUADRE 17. ATURATS REGISTRATS A L'SPO, ILLES BALEARS. SETEMBRE DE 2008

	< 25 anys	25 a 44	45 o més	Total
Homes	4.072	14.537	7.883	26.492
Dones	2.637	11.251	6.427	20.315
Total	6.709	25.788	14.310	46.807
	< 25 anys	25 a 44	45 o més	Total
Homes	8,70	31,06	16,84	56,60
Dones	5,63	24,04	13,73	43,40
Total	14,33	55,09	30,57	100

QUADRE 18. ATURATS REGISTRATS A OFICINES D'Ocupació PÚBLIQUES SEGONS L'EPA

	< 25 anys	25 a 44	45 o més	Total
Homes	6.923	11.603	6.790	25.316
Dones	4.335	11.094	5.681	21.110
Total	11.258	22.697	12.471	46.426
	< 25 anys	25 a 44	45 o més	Total
Homes	14,91	24,99	14,63	54,53
Dones	9,34	23,90	12,24	45,47
Total	24,25	48,89	26,86	100

Font: EPA 3r trimestre 2008

En segmentar simultàniament per edat i sexe, els percentatges només resulten semblants en els dos trams de 45 o més anys i en el de dones de 25 a 44 anys.

Una segona referència de la capacitat de l'EPA per equilibrar la informació de fonts externes la proporciona la distribució dels aturats per nacionalitat. L'Observatori del Treball de les Illes Balears (2008) ofereix dades del SOIB d'octubre de 2008, reproduïdes en el Quadre 19. Les dades més pròximes en el temps per a l'EPA són les que apareixen en el Quadre 20. Encara que s'ha de tenir en compte que es comparen diferents períodes, la distribució percentual d'ambdues fonts amb prou feines presenta discrepància.

QUADRE 19. SOIB OCTUBRE 2008.

	Aturats	Percentatge
Espanyols	41.570	72,95
Estrangers UE-27	4.571	8,02
Estrangers no UE-27	10.841	19,03
Total	56.982	100

QUADRE 20. EPA 3R TRIMESTRE 2008.

	Aturats Registrats	Percentatge
Espanyols	33.075	71,25
Estrangers EU-27	3.961	8,53
Estrangers no UE-27	9.387	20,22
Total	46.423	100

5. FACTORS D'ELEVACIÓ

L'INE exposa en diversos documents els criteris estadístics a partir dels quals s'obtenen els factors d'elevació, però és necessari esbrinar quin és el resultat d'aplicar aquests pesos en les diferents variables de l'EPA.

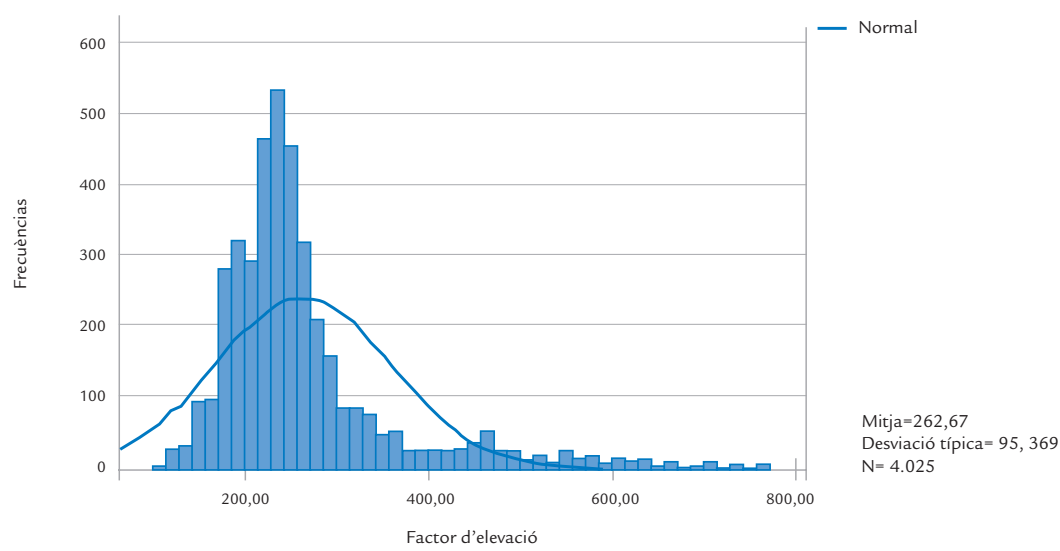
Amb la mostra de les Balears del tercer trimestre de 2008 s'ha analitzat la distribució dels factors d'elevació. Les principals característiques detectades són una important asimetria positiva (o per la dreta) i una forta curtosis (vegi's la Gràfica 22). Aquestes característiques no són diferents a les que presenten els factors d'elevació a la mostra de tot Espanya (en la Gràfica 23). La comparació d'alguns dels estadístics d'ambdues distribucions (Quadres 21 i 22) mostra, tanmateix, que aquestes característiques estan més suavitzades a les Illes Balears, amb una distribució a més que presenta menys dispersió.

QUADRE 21. ESTADÍSTICS DESCRIPTIUS DE LA DISTRIBUCIÓ DELS FACTORS D'ELEVACIÓ

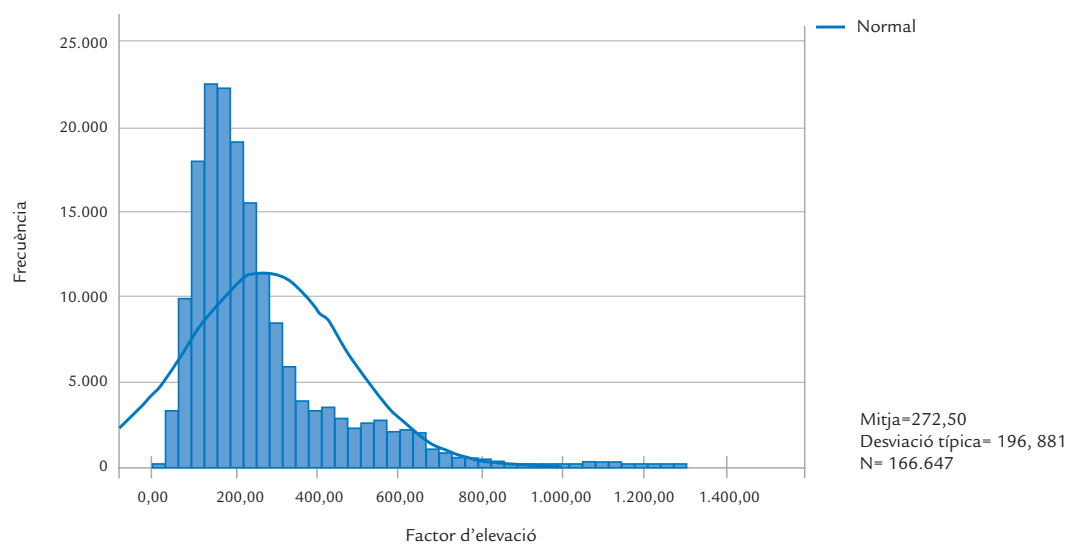
	Illes Balears		Espanya	
	Estadístic	Error Típic	Estadístic	Error típic
Mitja	263	2	273	0
Mitjana	240		215	
Variança	9.095		38.762	
Desv. típ.	95		197	
Mínim	112		13	
Màxim	761		2.960	
Rang	649		2.947	
Amplitud interquartil	69		169	
Asimetria	2	0	2	0
Curtosis	5	0	10	0

	5	10	25	50	75	90	95
Balears	150	163	191	229	281	371	455
Espanya	88	108	150	215	319	540	661

GRÀFICA 22. HISTOGRAMA DELS FACTORS D'ELEVACIÓ DE L'EPA (3R TRIMESTRE 2008), A LA MOSTRA DE LES ILLES BALEARS



GRÀFICA 23. HISTOGRAMA DELS FACTORS D'ELEVACIÓ DE L'EPA (3R TRIMESTRE 2008), A LA MOSTRA TOTAL D'ESPANYA



Les característiques que s'observen en les distribucions dels factors d'elevació poblacional no són en si mateixes negatives. El problema pot aparèixer quan es donin simultàniament dues circumstàncies:

- (1) Que els factors d'elevació situats en el rang superior de la distribució s'apliquin especialment a un nombre limitat de segments de la mostra.
- (2) Que els factors d'elevació es determinin en menor mesura pel disseny mostral, i en major mesura pel procés de calibratge o reponderació posterior.

Si es donen aquestes dues circumstàncies, l'elevació de les dades de la mostra a població pot generar problemes sobre l'estimació de les variables que no han estat tengudes en compte en el procés de calibratge. Concretament, suposem que ja que la subpoblació estrangera no està ben representada a la mostra de l'EPA, es modifiquen a l'alça els factors d'elevació corresponents a aquella població. Cada un dels individus mostrals representarà ara un major nombre d'individus en la població. Ja que aquesta subpoblació no està ben representada a la mostra, tampoc ho estan les característiques associades a aquella. Quan es reponderen els factors d'elevació tenint en compte que a la mostra hi ha pocs estrangers, però no es controla la representativitat d'aquest segment de la mostra, les estimacions en la població poden resultar errònies. El problema no només afectarà les característiques de la subpoblació d'estrangers, sinó les estimacions del conjunt de la població.

On més clarament es manifesten les conseqüències d'aquesta situació és en els factors d'elevació emprats en la subpoblació d'origen estranger, especialment quan es distingeix per la seva situació respecte a l'activitat econòmica.

En el Quadre 23 es mostren els valors mitjans dels factors d'elevació poblacional, calculats per a cada trimestre de l'EPA, distingint entre ocupats, aturats i inactius, i segons nacionalitat. Les sèries, representades en la Gràfica 24, posen en evidència diverses qüestions que resulta d'interès ressaltar:

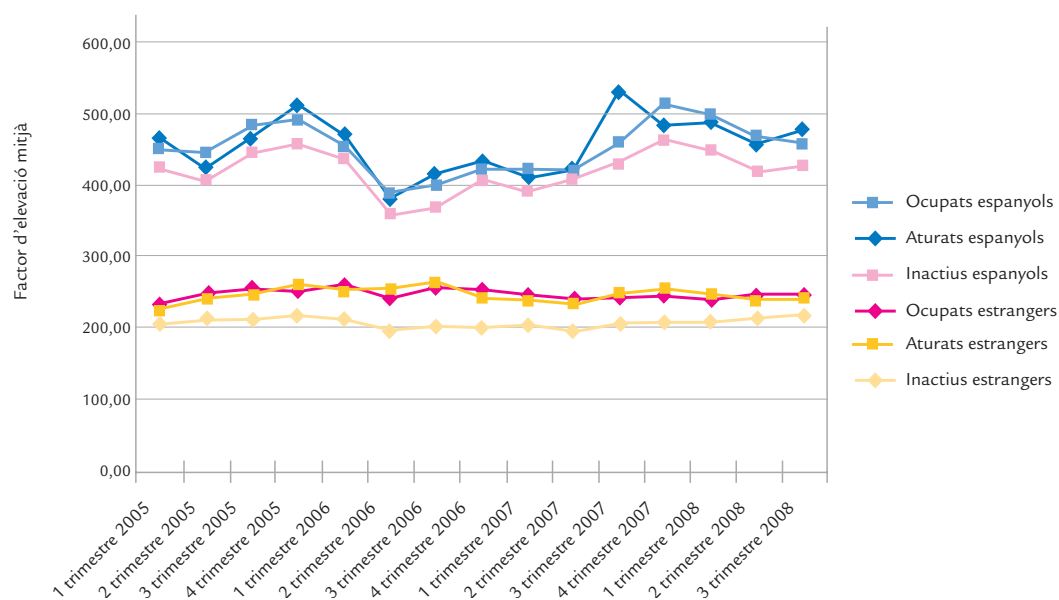
- (1) Els factors d'elevació mitjans corresponents a individus de nacionalitat estrangera pràcticament dupliquen els valors de la nacionalitat espanyola, per a les tres situacions d'activitat.
- (2) Els factors d'elevació corresponents a la nacionalitat estrangera són més volàtils que els de la nacionalitat espanyola.

Ambdues característiques, patents en la visualització de les sèries, es confirmen en els estadístics descriptius que es presenten en el Quadre 24.

QUADRE 23. VALORS MITJANS DELS FACTORS D'ELEVACIÓ DE L'EPA

	Espanyols			Estrangers		
	Ocupats	Aturats	Inactius	Ocupats	Aturats	Inactius
1T-05	227	223	205	449	458	422
2T-05	241	239	212	442	421	407
3T-05	253	244	211	481	461	444
4T-05	253	258	216	489	508	456
1T-06	255	252	213	455	465	435
2T-06	245	252	196	384	379	360
3T-06	253	255	200	399	412	368
4T-06	248	239	199	420	430	405
1T-07	242	239	202	417	409	391
2T-07	236	231	194	419	421	408
3T-07	247	247	203	457	526	429
4T-07	247	254	204	513	484	461
1T-08	241	246	207	492	488	447
2T-08	243	241	212	461	455	419
3T-08	247	248	218	459	475	425

GRÀFICA 24. EVOLUCIÓ DELS FACTORS D'ELEVACIÓ MITJANS (SEGONS ACTIVITAT I NACIONALITAT)



QUADRE 24. ESTADÍSTICS DESCRIPTIUS DE LA SÈRIE DE FACTORS D'ELEVACIÓ (PERÍODE 1-2005 A 3-2008)

		Mínim	Màxim	Mitja	Desv. típ.	Coef. variació
Espanyols	Ocupats	227	255	245	7	3,01%
	Aturats	223	258	245	10	3,89%
	Inactius	194	218	206	7	3,55%
Estrangers	Ocupats	384	513	449	36	8,08%
	Aturats	379	526	453	40	8,94%
	Inactius	360	461	419	30	7,07%

Les conseqüències d'aquesta estructura de ponderació es descobreixen quan s'analitzen les transicions entre diferents trimestres de l'EPA. Fixi's que si el nombre d'aturats estrangers augmenta en una sola unitat, elevar el seu pes en la població suposarà un increment estimat d'un 452 persones en la població d'aturats. L'impacte de les variacions mostrals en successives enquestes sobre les dades poblacionals és, malgrat tot, incert. En el cas que els factors d'elevació no sofreixin una modificació important d'un període a un altre, les transformacions en les dades no ponderades es traslladen uniformement a les dades ponderades. Però, com s'ha indicat, les sèries dels factors d'elevació mitjans mostren una forta variabilitat en el temps. Com a conseqüència, ocorre que les variacions entre períodes de les dades no ponderades suposen variacions molt diferents en qüestió de les dades ponderades. En el Quadre 25 es mostren les variacions intertrimestrals sobre les dades no ponderades i ponderades. En el Quadre 26 es presenten uns "factores de variació" implícits en aquestes variacions. El resultat que s'obté és poc justificable en termes de representativitat de la mostra. Concretament, si s'observen les variacions del darrer trimestre pot comprovar-se que un increment en el nombre d'aturats estrangers de 8 individus a la mostra es correspon amb un increment de l'atur de 4.613 persones. Els efectes dels diferents pesos poden arribar a invertir la variació, com en el cas del nombre d'ocupats espanyols, que encara que en aquell trimestre disminueix en 5 individus a la mostra, això es correspon amb un increment de 5.524 persones en la població.

Ha de tenir-se en compte que el disseny mostral de l'EPA no confereix sentit estadístic als anteriors "factores d'elevació implícits" calculats en les variacions trimestrals. S'entén que les variacions a la mostra no es tradueixen directament en la població, ja que no són únicament 5 nous individus els que adquireixen la condició d'ocupat, sinó que aquest nombre és el saldo final d'entrades i sortides d'individus, tots ells amb diferents ponderacions. El descens del nombre d'ocupats en la mostra de cinc individus és el saldo net de les variacions, que es tradueix en la població en un saldo poblacional indeterminat. Malgrat tot, és difícil justificar que variacions netes d'un nombre tan reduït d'individus tinguin un impacte tan elevat sobre els valors estimats de la població.

QUADRE 25. VARIACIONS RESPECTE AL TRIMESTRE ANTERIOR

		Dades no ponderades			Dades ponderades		
		Ocupats	Aturats	Inactius	Ocupats	Aturats	Inactius
Espanyols	2 trimestre 2005	63	-68	-153	37.745	-13.594	-22.931
	3 trimestre 2005	-22	-25	-41	13.679	-5.579	-9.514
	4 trimestre 2005	-149	32	125	-37.512	9.209	31.979
	1 trimestre 2006	-6	14	-2	988	2.939	-4.493
	2 trimestre 2006	215	-17	-62	38.861	-4.337	-33.587
	3 trimestre 2006	-2	-28	-47	12.221	-6.792	-4.643
	4 trimestre 2006	-90	10	152	-30.389	1.214	29.819
	1 trimestre 2007	-26	35	28	-16.793	8.352	9.222
	2 trimestre 2007	169	-24	-80	30.675	-6.442	-25.910
	3 trimestre 2007	-56	-29	-52	5.204	-5.665	495
	4 trimestre 2007	-124	48	104	-30.063	12.605	21.994
	1 trimestre 2008	15	25	-18	-5.738	5.256	335
	2 trimestre 2008	76	-16	-121	21.604	-4.503	-18.775
	3 trimestre 2008	-5	7	-73	5.524	2.567	-9.374
Estrangers		Dades no ponderades			Dades ponderades		
		Ocupats	Aturats	Inactius	Ocupats	Aturats	Inactius
	2 trimestre 2005	44	-12	-16	18.221	-6.204	-7.864
	3 trimestre 2005	-17	0	7	647	760	5.773
	4 trimestre 2005	-14	1	12	-5.297	1.411	6.348
	1 trimestre 2006	12	7	16	-1.196	2.380	5.047
	2 trimestre 2006	66	1	9	10.827	-1.934	-4.711
	3 trimestre 2006	15	-9	-10	10.061	-2.775	-2.745
	4 trimestre 2006	-31	10	11	-6.945	4.631	8.412
	1 trimestre 2007	2	19	3	57	7.176	-463
	2 trimestre 2007	46	-27	-8	19.677	-10.806	-1.243
	3 trimestre 2007	-11	-4	-9	6.702	98	-1.530
	4 trimestre 2007	-55	25	-1	-11.970	11.382	2.765
	1 trimestre 2008	1	8	19	-4.506	4.091	7.039
	2 trimestre 2008	57	-9	-6	19.023	-5.752	-5.766
	3 trimestre 2008	9	8	-4	3.329	4.613	-1.110

QUADRE 26. FACTORS D'ELEVACIÓ IMPLÍCITS A LES VARIACIONS INTERTRIMESTRALS

	Espanyols			Estrangers		
	Ocupats	Aturats	Inactius	Ocupats	Aturats	Inactius
2 trimestre 2005	599	200	150	414	517	492
3 trimestre 2005	-622	223	232	-38	-	825
4 trimestre 2005	252	288	256	378	1.411	529
1 trimestre 2006	-165	210	2.247	-100	340	315
2 trimestre 2006	181	255	542	164	-1.934	-523
3 trimestre 2006	-6.111	243	99	671	308	275
4 trimestre 2006	338	121	196	224	463	765
1 trimestre 2007	646	239	329	29	378	-154
2 trimestre 2007	182	268	324	428	400	155
3 trimestre 2007	-93	195	-10	-609	-25	170
4 trimestre 2007	242	263	211	218	455	-2.765
1 trimestre 2008	-383	210	-19	-4.506	511	370
2 trimestre 2008	284	281	155	334	639	961
3 trimestre 2008	-1.105	367	128	370	577	278

6. CONCLUSIONS

El primer punt a destacar és la necessitat d'augmentar la mida mostral de l'EPA a les Balears, si aquesta ha de servir com un instrument útil per a l'anàlisi del mercat laboral a les Illes. Com s'ha indicat, el propi INE adverteix del perill d'intentar extreure conclusions quan es basen en segments inferiors a les 5.000 persones. No és difícil que això ocorri quan es creuen algunes de les variables de l'EPA. Igualment s'ha d'assenyalar la necessitat que l'EPA ofereixi informació amb un nivell de desagregació més elevat que la comunitat autònoma, especialment tenint en compte les diferents illes. A més, s'ha de ressaltar la cada vegada més descompensada relació entre la mida de la mostra assignada a les Illes Balears (més o menys estable en el període analitzat) i la població total (amb una taxa de creixement rellevant).

Respecte a l'anàlisi realitzada sobre la capacitat de l'EPA per equilibrar les seves estimacions poblacionals respecte a fonts d'informació externes, s'ha obtingut una evidència mixta.

En primer lloc, encara que és cert que per a algunes de les principals variables les estimacions poblacionals són correctes, això no ocorre a totes les variables. Per exemple, l'equilibri respecte a la informació del Padró Municipal s'aconsegueix per a la població total i per a la subpoblació de nacionalitat estrangera, però no per a la nacionalitat espanyola. El mateix problema apareix en altres variables que també estan equilibrades en el total de la població, però que deixen d'estar-ho quan es valoren en diferents subpoblacions. Això ocorre, per exemple, en la distribució per sexes quan es consideren diversos trams d'edat o per als aturats registrats, on en segmentar-se per trams d'edat tampoc resulta equilibrada respecte a l'obtinguda a l'INEM.

En segon lloc, aquest equilibri s'aconsegueix en gran part gràcies a un procés de calibratge o reponderació que és aliè al disseny mostral. Per a algunes variables, la comparació entre els resultats no ponderats i ponderats de la mostra permet detectar que importants segments de la població estan subrepresentats a la mostra. Això s'aplica, per exemple, al tram d'edat entre 25 i 39 anys, un tram d'edat amb problemes de representació en anteriors períodes de l'EPA i que les modificacions de 2005 no semblen haver resolt.

De manera general, pot concloure's que l'EPA només aconsegueix equilibrar els principals agregats, però no aconsegueix el mateix resultat a les variables desagregades. S'ha de destacar a més que la mostra de l'EPA presenta problemes de subrepresentació per a alguns segments importants de la població.

Respecte al procés de reponderació o calibratge que realitza l'INE, els seus efectes sobre les estimacions de les Balears són problemàtics. El procés de calibratge té per objectiu equilibrar la mostra respecte a fonts d'informació externa. Si bé la reponderació sembla resoldre un problema d'equilibri dirigit a les xifres de les variables agregades, el seu ús és discutible des d'altres perspectives.

En primer lloc, allò que seria desitjable és que el procés de calibratge no hagués d'efectuar-se, ja que d'alguna manera implica que el disseny mostral no aconsegueix reflectir correctament la informació poblacional. L'INE ha justificat el calibratge especialment en relació amb les dificultats per entrevistar la població immigrant (INE, 2005). El procés de calibratge és, per tant, una forma de neutralitzar la falta d'informació que és conseqüència que determinats segments de la població no s'han recollit correctament a la mostra.

En segon lloc, tant si es tracta de solucionar un problema del disseny mostral, com si és un problema de falta de resposta, el procés de calibratge només té sentit si els factors d'elevació poblacionals calculats originalment no es modifiquen en excés. Si la reponderació implica que les ponderacions originals es modifiquen extremadament, es danyarà el disseny probabilístic. La asimetria positiva detectada en la distribució dels factors d'elevació il·lustra que aquest problema pot donar-se a la mostra de l'EPA.

En tercer lloc, el procés de calibratge té per objectiu unes poques variables. Per això, quan s'apliquen els nous factors de ponderació s'ignora quin és el seu efecte sobre altres variables d'interès en la mostra. L'ús de les noves ponderacions requereix el compliment

de la hipòtesi que les variables que intervenen en la seva elaboració presentin una correlació alta amb la resta de variables sobre les quals s'apliquen. Aquesta hipòtesi s'incompleix quan les característiques dels segments de població subestimats no són homogènies respecte a les fonts d'informació externa.

Finalment, quan s'analitza la dinàmica de la població activa, els factors de ponderació juguen un paper poc precís. Per això, per estudiar les transicions en el mercat laboral seria més adequat, en lloc d'analitzar les sèries de dades agregades, emprar les dades individuals de l'EPA, valent-se de l'estructura de plafó que pot construir-se a partir de les microdades.

7. REFERÈNCIES

Álvarez, F., 2000, “La Encuesta de Población Activa: cómo mide el mercado laboral, Estadística Española”, vol. 42, núm. 146, pàg. 205 a 261.

INE, “Método de reponderación empleado en la EPA”, disponible en xarxa a: http://www.ine.es/epa02/documento_tecnico.pdf.

INE, 2004, “Revisión de la Encuesta de Población Activa”, disponible en xarxa a: <http://www.ine.es/prensa/np322.pdf>.

INE, 2005, “Encuesta de Población Activa. Diseño de la Encuesta y Evolución de la calidad de los Dades, Informe técnico”.

Martí, M. i C. Ródenas, 2004, “Migrantes y migraciones: de nuevo la divergencia en las fuentes estadísticas”, vol. 46, núm. 156, pàg. 293-321.

Ministeri de Treball i Immigració, 2007, Anuari del Ministeri de Treball i Immigració 2007, <http://www.mtin.es/estadisticas/anuario2007/index.htm>.

Observatori del Treball de les Illes Balears, “Nota de conjuntura laboral a les Illes Balears, octubre 2008”.

Pérez, J.I., 2008, “¿Por qué el paro registrado supera al paro de la EPA?”, revista del Ministeri de Treball i Afers Socials, núm. 71, març, pàg. 53-65.

Prudencio, C.A. i L. Toharia, 2007, “Análisis regional del comportamiento de las personas inscritas en las oficinas de empleo públicas según la Encuesta de Población Activa. 2004 y 2005”, Estadística Española, vol. 49, núm. 166, pàg. 407-450.

ANÁLISIS DE LA REPRESENTATIVIDAD DE LA ENCUESTA DE POBLACIÓN ACTIVA EN LAS ISLAS BALEARES

OBSERVATORI DEL TREBALL DE LES ILLES BALEARS

*Joaquín Alegre Martín
Departament de Economia Aplicada
Universitat de les Illes Balears*

*Col·lecció d'Estudis Laborals.
Núm.1*

Dirección: Llorenç Pou Garcias
Director General de Planificación Estratégica

Coordinación: Delia Riera Tur

Autor: Joaquín Alegre Martín

ISBN 978-84-693-2893-4

Si desea más información y/o subscribirse a los informes del Observatori del Treball de les Illes Balears, puede dirigirse a otib@dgplaes.caib.es o bien:

Dirección General de Planificación Estratégica
Conselleria de Treball i Formació
Plaça de Son Castelló, 1
Polígono de Son Castelló. 07009 Palma
Tel.: 971 17 63 00
<http://www.caib.es>
<http://treballiformacio.caib.es>

1. INTRODUCCIÓN	3
2. DISEÑO MUESTRAL DE LA EPA	5
3. PROBLEMAS DEL DISEÑO MUESTRAL DE LA EPA	6
4. REPRESENTATIVIDAD DE LA MUESTRA.....	9
4.1. Comparación con fuentes externas	10
4.2. Población	11
4.3. Nacionalidad española y extranjera.....	13
4.4. Nacionalidades extranjeras	17
4.5. Edad.....	18
4.6. Sexo	24
4.7. Situación laboral	27
5. FACTORES DE ELEVACIÓN.....	30
6. CONCLUSIONES	36
7. REFERENCIAS	38

I. INTRODUCCIÓN

En este trabajo se analiza la capacidad que tiene la Encuesta de Población Activa (EPA) del INE para proporcionar información representativa de las Islas Baleares. Las dificultades que ha tenido la EPA para describir correctamente la situación del mercado laboral español han sido la causa de sus diversos cambios metodológicos. Las mejoras introducidas, especialmente tras las modificaciones del año 2005, han reafirmado a la EPA como una de las fuentes de información más fiables sobre la situación laboral de los españoles. En la actualidad el INE entrevista aproximadamente a 200.000 personas en 65.000 hogares, ofreciendo información a nivel nacional y provincial tanto de la población activa y sus diversas categorías, como de la población inactiva. La validez de estas estimaciones, con todos los matices o interpretaciones que quieran realizarse, ha sido confirmada en diversos estudios. No obstante, existen aún pocos trabajos que analicen la validez de las estimaciones de la EPA a nivel de Comunidad Autónoma. Son varios los servicios de estudios e institutos estadísticos regionales que reclaman un incremento del tamaño muestral correspondiente al nivel autonómico y provincial; esta reclamación se justifica porque en este nivel de desagregación es fácil que en los análisis se alcance el tamaño mínimo a partir del cual los resultados de la Encuesta no son fiables¹. Como ya señalaba el INE en 2004, “es necesario proseguir con el contraste de las series de resultados con los de otras fuentes estadísticas y registros administrativos, con el fin de conocer y explicar las causas de ciertas inconsistencias que se han observado, sobre todo cuando se trabaja con desagregaciones espaciales” (INE, 2004). No obstante, aún faltan estudios que valoren la capacidad de la Encuesta para estimar correctamente información relevante a nivel de las Comunidades Autónomas.

El diseño muestral de la EPA tiene como objetivo que las personas seleccionadas en la muestra sean representativas de toda la población a nivel nacional. Para las Comunidades Autónomas (y las provincias) el INE indica que sólo es válida la información con un menor nivel de desagregación, dependiendo de los coeficientes de variación de los estimadores. El diseño muestral de la EPA no garantiza, por tanto, la representatividad estadística provincial en todas las variables. En consecuencia, el objetivo de este trabajo no es averiguar la precisión de las estimaciones a un nivel muy desagregado. Las variables empleadas se definen, por tanto, en un nivel de agregación alto. No obstante, debe tenerse en cuenta que, en el contexto de la EPA, la precisión en el nivel agregado depende del grado de precisión empleado en un nivel de agregación menor. Es posible que los procesos de reponderación que emplea la EPA puedan equilibrar la información agregada que se obtiene de fuentes de información externa, pero que ello sea a costa de desvirtuar la información más desagregada que también se deriva de la encuesta.

Los contenidos de este trabajo se presentan en el siguiente orden. En un primer apartado se explica el diseño muestral de la EPA, tal como es expuesto por el INE. Aunque se trata de un diseño conocido, se ha preferido realizar una corta presentación del mismo, dado que la representatividad de la encuesta depende principalmente de su diseño. En un segundo apartado se discuten los problemas

¹ El INE advierte que resultados inferiores a 5.000 personas no son fiables, al estar sujetos a fuertes variaciones consecuencia de los errores muestrales.

relacionados con el tipo de diseño muestral estratificado que emplea la EPA. Se trata de una crítica específica al diseño, que se basa, esencialmente, en la posibilidad de que los estratos que define el INE carezcan del nivel de precisión suficiente para estimar algunas variables. En un tercer apartado se analiza la representatividad de la muestra de la EPA en las Islas Baleares, entendiendo por representatividad la capacidad de la Encuesta para proporcionar cifras coincidentes con fuentes de información externas (por ejemplo, del Padrón Municipal). Tras exponer el proceso de calibración o reponderación que emplea el INE, se analiza la capacidad de la EPA para equilibrar información referida a población (total y por nacionalidades), distribución por edades, sexo y situación laboral. Estos análisis permiten detectar cuáles son los segmentos en los que la EPA subestima las variables poblacionales. En el último apartado se analizan los factores de elevación poblacional empleados en el INE. Los resultados que se obtienen en este apartado sugieren que el excesivo valor de los factores de elevación poblacional en algunos segmentos de la Encuesta, así como su variabilidad en el tiempo no sólo ocasionan problemas de interpretación de los resultados de la Encuesta, sino que pueden generar estimaciones erróneas. En este apartado se señalan además las dificultades del proceso de calibración cuando éste es una alternativa de facto a la ampliación del tamaño muestral. Finalmente, se exponen las principales conclusiones del trabajo.

Para la realización del trabajo se han empleado los ficheros de microdatos de la EPA, del periodo comprendido entre el primer trimestre de 2005 y el tercer trimestre de 2008. La mayoría de la información empleada se refiere a la identificación de las Islas Baleares, aunque también se han realizado algunos análisis referidos a toda España. Las comparaciones externas se han realizado, esencialmente con respecto a los Padrones Municipales.

2. DISEÑO MUESTRAL DE LA EPA

El diseño de la muestra EPA (INE, 2005) utiliza un muestreo bietápico con estratificación en las unidades de primera etapa. Las unidades de primera etapa están constituidas por las secciones censales. Las unidades de segunda etapa son las viviendas familiares principales (ocupadas permanentemente) y los alojamientos fijos. Dentro de las unidades secundarias no se realiza ningún submuestreo, recogiendo información de todas las personas que viven en ellas.

Las unidades de primera etapa se estratifican según un doble criterio:

- (1) Criterio geográfico (de estratificación). Por el cual las secciones censales se agrupan en estratos dentro de cada provincia, de acuerdo a la importancia demográfica del municipio al que pertenecen.
- (2) Criterio socioeconómico (de subestratificación). Por el cual las secciones censales se agrupan dentro de cada uno de los estratos, según las características de las mismas.

(1) Los estratos se forman distinguiendo:

- Estrato 1: Municipio capital de provincia.
- Estratos 2 y 3: Municipios autorrepresentados.
- Estrato 4: Municipios entre 50.000 y 100.000 habitantes.
- Estrato 5: Municipios entre 20.000 y 50.000 habitantes.
- Estrato 6: Municipios entre 10.000 y 20.000 habitantes.
- Estrato 7: Municipios entre 5.000 y 10.000 habitantes.
- Estrato 8: Municipios entre 2000 y 5.000 habitantes.
- Estrato 9: Municipios menores de 2.000 habitantes.

(2) Dentro de cada estrato se han generado subestratos a partir de dos criterios. Para las secciones de los estratos correspondientes a municipios de menos de 10.000 habitantes se les asigna como subestrato la comarca del municipio al que pertenecen. En el caso de Baleares, este criterio se aplica al 20% de las secciones censales. Para el resto de secciones los subestratos se generan mediante un análisis cluster que emplea información demográfica y socioeconómica de las secciones. El objetivo es formar grupos homogéneos de secciones. La información utilizada para generar los clusters procede del Censo 2001 y de la Agencia Estatal de Administración Tributaria. Según el INE (2005), se emplean las variables siguientes:

- Los porcentajes de parados, inactivos y ocupados.
- El porcentaje de extranjeros.
- Los porcentajes de población en distintos tramos de edad (entre 0 y 19 años; entre 15 y 24; 65 o más años).
- Porcentaje de personas por distintos niveles de estudios (analfabetos o estudios primarios; ESO, EGB, Bachillerato o FP; estudios universitarios).
- Porcentajes de la población ocupada según su condición socioeconómica (empresarios, trabajadores, etc.).
- Renta por vivienda.
- Rentas de capital sobre la renta total.
- Renta agraria sobre la renta total.

Como señala el documento del INE estas variables han sido seleccionadas considerando que están muy correlacionadas con el objeto de estudio de la encuesta.

Dentro de cada provincia la afijación entre estratos es proporcional al tamaño de cada uno de ellos, aunque se han potenciado los municipios de mayor tamaño. Dentro de los estratos, la afijación entre subestratos es estrictamente proporcional al número de viviendas familiares en el mismo. Inicialmente el tamaño de la muestra se fijó, mediante un método de mínima varianza para un coste fijo, en un total de 3000 secciones, con una media de 20 viviendas por sección. Para distribuir estas secciones entre provincias, estratos y subestratos se establecieron diversos criterios, que garantizaran la fiabilidad de las estimaciones a nivel nacional y provincial. El criterio final de afijación es un compromiso entre la afijación uniforme y la proporcional. Las Islas Baleares distribuyen sus secciones censales en cada estrato de la siguiente manera: Estrato 1, 44; Estrato 5, 27; Estrato 6, 12; Estrato 7, 12; Estrato 8, 9.

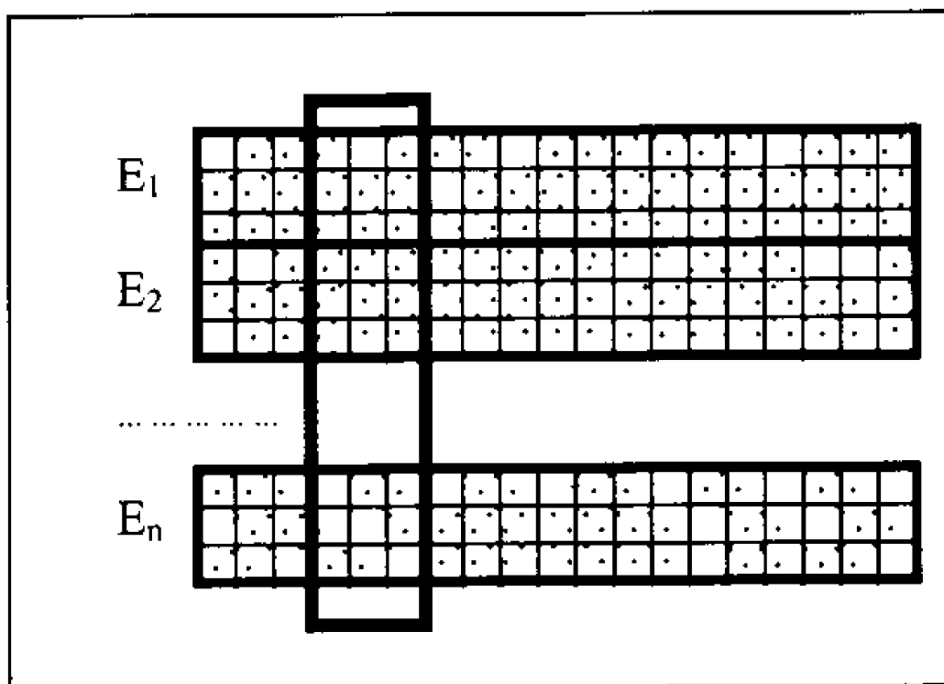
3. PROBLEMAS DEL DISEÑO MUESTRAL DE LA EPA

En el muestreo estratificado, los estratos se definen de manera que, respecto a las variables de interés, cada uno de los estratos presente unidades homogéneas entre sí y, a su vez, sean heterogéneas cuando se comparan entre ellos. La eficacia del muestreo estratificado depende de que los grupos o estratos presenten estas características. Sin embargo, este no parece ser el caso de algunas características de interés en la EPA. Por ejemplo, Martí y Ródenas (2004) señalan que los individuos con nacionalidad extranjera constituyen una categoría poblacional que se presenta por igual en todos los estratos de la muestra y se distribuye de manera homogénea dentro de la vivienda. Ello puede tener como consecuencia que las estimaciones ligadas a este tipo de característica sean imprecisas.

Martí y Ródenas (2004) exponen gráficamente las dificultades de estimar determinadas características mediante un diseño muestral por estratos. En las gráficas (que reproducimos de su artículo), cada estrato está constituido por un conjunto de viviendas, representadas mediante casillas. Dentro de cada vivienda los individuos que poseen la característica estudiada se simbolizan con un punto. La muestra de viviendas seleccionadas en el muestreo se señala con la barra vertical que atraviesa todos los estratos. Martí y Ródenas (2004) distinguen tres posibles situaciones:

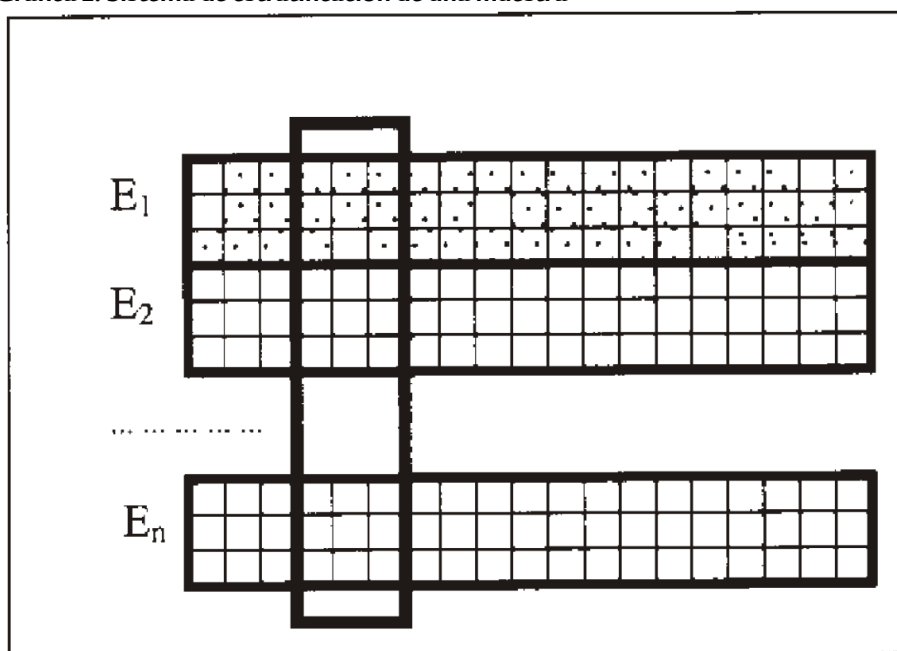
En la primera, se quiere estimar una característica que está presente por igual en todos los estratos y además es una unidad heterogénea en la unidad última de muestreo (en el caso de la EPA, la vivienda). Ejemplos de este tipo de característica son el sexo, la edad o el estado civil. Si además la vivienda está compuesta por personas de distinto sexo, edad y estado civil, la muestra que se extrae en cualquier estrato será representativa de la población. La Gráfica 1 ilustra esta situación.

Gráfica 1. Sistema de estratificación de una muestra



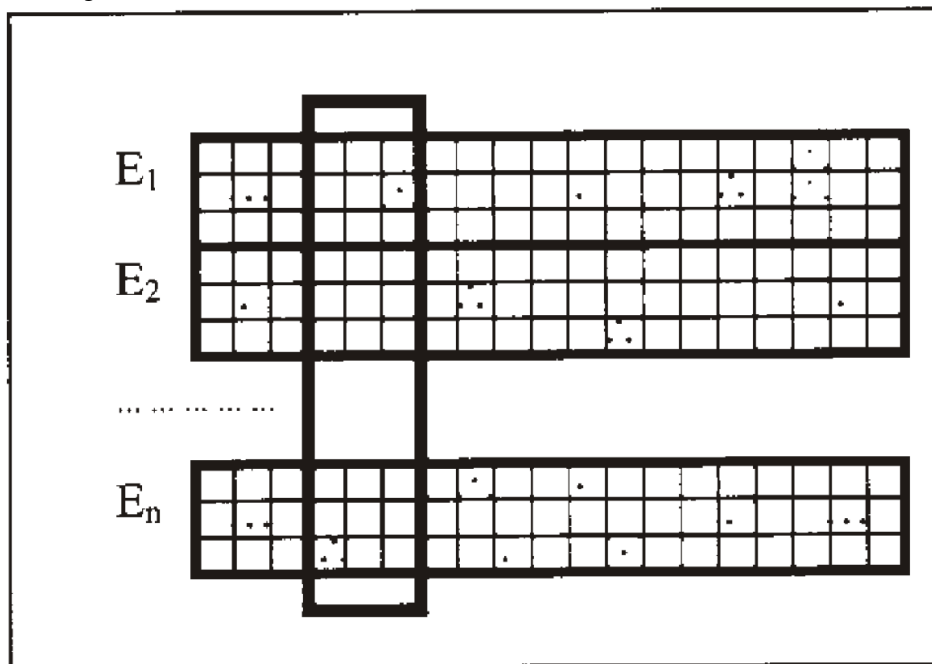
En la segunda situación, la característica a estimar la poseen un número reducido de viviendas, distribuyéndose en las mismas de manera homogénea (véase la Gráfica 2). Si los estratos se construyen teniendo en cuenta esta característica, creando un estrato que contenga las viviendas con individuos que la posean, la muestra resultará representativa.

Gráfica 2. Sistema de estratificación de una muestra



El problema puede surgir cuando la característica de interés no se tiene en cuenta cuando se construyen los estratos. Esta situación quedaría reflejada en la Gráfica 3. Si ningún estrato se diseña para recoger esa característica, la muestra seleccionada incluirá pocas viviendas con individuos que tengan la característica. Una posible consecuencia de este problema es que la presencia de esta característica en la muestra puede presentar gran variabilidad. Pero incluso si un tamaño muestral adecuado consigue que en la muestra la característica poblacional esté bien representada, la representatividad de otras variables asociadas a esta característica será extremadamente baja.

Gráfica 3. Sistema de estratificación de una muestra



Aunque se ha apuntado que la falta de respuesta puede ser una de las razones por las que las personas de nacionalidad extranjera no están bien representadas en la muestra de la EPA, los datos disponibles para el año 2007 no parecen confirmarlo. Según el informe del INE de evaluación de la calidad de la EPA-2007, las negativas a responder a la encuesta por primera vez correspondientes a personas extranjeras es de 4,9% (para toda España, tercer trimestre), algo inferior al porcentaje de personas extranjeras que se encuesta en la EPA (5,7% en el tercer trimestre).

Los posibles problemas de utilización de la EPA pueden resumirse, por tanto, en que la precisión de las estimaciones depende del tamaño de la muestra y la heterogeneidad u homogeneidad de la variable estudiada:

(i) Respecto al tamaño muestral, Martí y Ródenas (2004) señalan que “muestras dimensionadas para extraer conclusiones generales o medir determinados parámetros no se pueden utilizar para el análisis de caracteres muy poco frecuentes, pues el número de entrevistas realizadas a personas que dispongan de dicho rasgo no será suficiente para dar soporte a la estimación, ya que se puede superar ampliamente lo que se podría llamar el error permitido”. Y, por tanto, sólo se

deberían estimar características poblacionales cuya presencia en la muestra es lo suficientemente amplia.

(2) En el diseño de la EPA se incluyen un número limitado de características para definir los estratos de población. Esta definición puede perjudicar la precisión de características que, en primer lugar, se distribuyen de manera uniforme entre los estratos de la población y, en segundo lugar, lo hacen homogéneamente dentro de la última unidad de muestreo (la vivienda).

Como se ha explicado al exponer el diseño muestral de la EPA, los criterios de definición de los estratos han intentado incluir no sólo criterios geográficos, sino socioeconómico, empleando en este último caso información censal correlacionada con el objeto de estudio de la encuesta. Entre las variables empleadas para generar estos estratos se encuentra el porcentaje de extranjeros. Sin embargo, el procedimiento expuesto por el INE deja algunas cuestiones imprecisas. En primer lugar, no se publica cuál es la importancia que cada una de las variables socioeconómicas tiene en la determinación de los subestratos. Dado el conjunto de variables que se utiliza no es descartable que sean aquellas variables con una mayor variabilidad las que más influyan en el proceso de agrupación. En segundo lugar, estas variables sólo juegan un papel secundario, en la definición de los subestratos, estando los estratos definidos por un criterio geográfico.

4. REPRESENTATIVIDAD DE LA MUESTRA

El número de individuos de las Islas Baleares sobre los que la EPA ofrece información efectiva varía en cada ola de la encuesta. En el periodo comprendido entre el primer trimestre de 2005 y el tercer trimestre de 2008 los encuestados en Baleares oscilan entre un mínimo de 3752 y un máximo de 4235. La información de la muestra se eleva a la población empleando los factores de elevación calculados por el INE. En principio, éstos se derivan del diseño muestral y elevan la información recogida en un estrato a lo que éste representa en la población. Además, la EPA corrige los pesos o factores de elevación originales (los que se deducen del diseño muestral), mediante un proceso de calibración o reponderación, que emplea fuentes estadísticas externas. Su objetivo es mejorar la calidad de las estimaciones, de manera que la estimación de las variables para las que se dispone de información externa fiable coincida con los datos de esa fuente.

La base de microdatos de la encuesta ofrece los factores de elevación de cada individuo. Estos valores indican el número de personas de la población a la que representan cada una de las personas encuestadas. A partir del año 2005 estos factores se han modificado para reflejar mejor el peso de la población extranjera. Los factores de elevación se obtienen en dos etapas. En primer lugar, se obtienen los factores de elevación a partir del diseño muestral de la encuesta. Sin embargo, los datos muestrales elevados no se concuerdan con los datos de población. La diferencia entre las estimaciones derivadas de la encuesta y la de otras fuentes externas las justifica el INE por diversas causas (INE, Método de reponderación empleado en la EPA):

- (1) Falta de respuesta, que produce sesgos en los estimadores finales al afectar de forma relativamente distinta a hogares o grupos de población de determinada composición.
- (2) Posible tendencia de la muestra a sobrerrepresentar a hogares de una determinada estructura en detrimento de otros.

Por ello, en una segunda etapa, denominada de calibrado o reponderación, los factores de elevación se modifican para que los resultados poblacionales coincidan con los de una o varias fuentes externas de referencia. En el caso de la EPA se ajusta a la pirámide de población por sexo y edad de cada comunidad autónoma, el total de población de cada provincia y el total de población por nacionalidad en cada comunidad autónoma (distinguiendo en este caso entre españoles y extranjeros). El proceso de calibración permite equilibrar la muestra, en el sentido de que las estimaciones de estas variables en la muestra coincidirán aproximadamente con el valor observado en las fuentes externas.

A la hora de valorar el impacto del calibrado sobre las cifras de la EPA se pueden tener en cuenta dos cuestiones. La primera, relacionada con el grado en el que se consigue el objetivo perseguido en el proceso de calibración. La segunda, sobre si la estructura de pesos que se genera en el proceso de calibración ocasiona que algunos segmentos de la muestra estén hiperrepresentados. La primera cuestión puede valorarse comprobando si efectivamente la información elevada a población está equilibrada con respecto a las distintas fuentes externas empleadas. Adicionalmente, esta comprobación lleva aparejado además el análisis de cuáles son los segmentos sobre los que tiene mayor incidencia la reponderación, lo que enlaza con la segunda cuestión. Para contestar a ésta última, es necesario realizar un análisis de la importancia relativa de los factores de elevación en algunos segmentos de la muestra, valorando cuál es el impacto relativo de los elementos muestrales cuando se eleva su información a nivel poblacional.

4. 1. Comparación con fuentes externas

El proceso de calibrado emplea como fuentes externas las siguientes variables en cada Comunidad Autónoma:

- Proyecciones de población por grupos de edad y sexo (22 grupos) quinquenales.
- Proyecciones de población por nacionalidad (españoles y extranjeros).
- Totales de población por provincia.

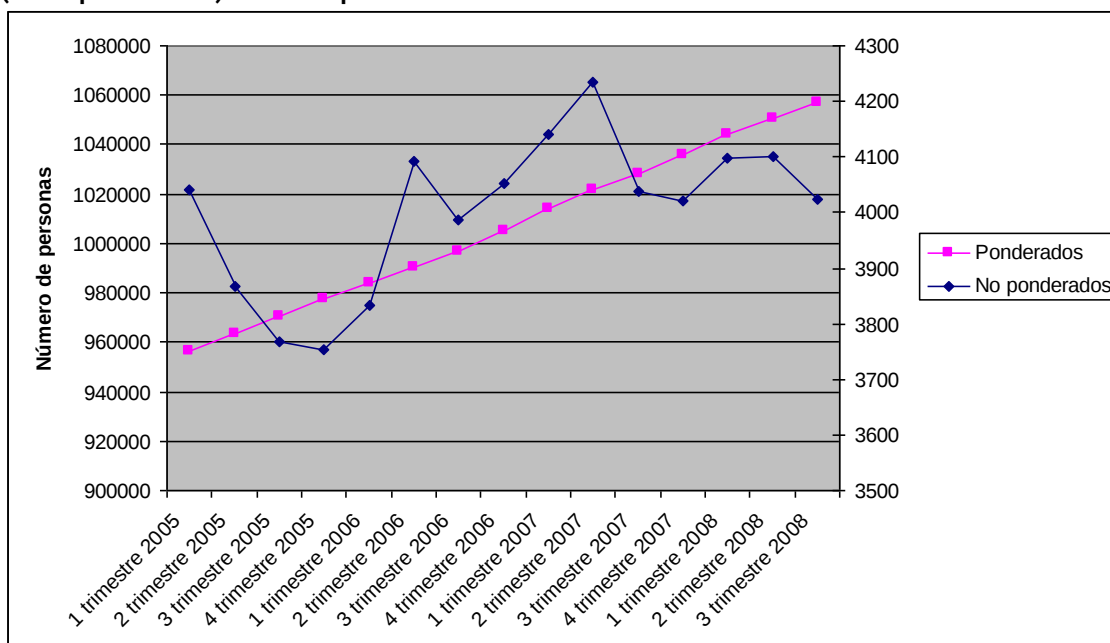
Uno de los objetivos del calibrado es garantizar que la suma de las observaciones ponderadas coincida con los valores de las fuentes externas. El proceso garantiza, por tanto, que cuando se eleva a la población los resultados de la EPA, coincidirán con las fuentes externas la información agregada de los distintos segmentos. Debe tenerse en cuenta que, en principio, el proceso de calibrado debería garantizar el mismo equilibrio en los segmentos cruzados de la variable. Es decir, los resultados por edad y sexo (22 grupos) deben equilibrarse con la fuente externa, igual que el total de población o la cantidad de población española y extranjera.

4.2. Población

Uno de los principales criterios de equilibrio definido por el INE es el de población. Para el conjunto de Comunidades Autónomas el proceso de calibrado tiene en cuenta el total de población por provincia. En el caso de Comunidades con una única provincia no se establece un agregado de menor entidad. En el caso de las Islas Baleares, se considera únicamente el total de la Comunidad, ignorando la división territorial definida por la insularidad.

Las proyecciones de población que realiza el INE en Baleares para el periodo 1-2005 a 3-2008 suponen un incremento de la población desde 956.673 a 1.057.259 personas. Esto implica una tasa trimestral de crecimiento medio acumulativo del 0,62%, con un incremento final de la población del 10,51% en el conjunto del período. A pesar del crecimiento de la población, el tamaño de la muestra correspondiente a Baleares se ha mantenido la mayoría de los trimestres en alrededor de 4000 personas. En la gráfica 4 puede observarse la evolución del tamaño de la muestra (datos no ponderados) y de la proyección de población (datos ponderados a partir de los factores de elevación) durante el periodo.

Gráfica 4. Evolución del tamaño muestral (datos no ponderados) y de la población representada (datos ponderados) en la EPA para las Islas Baleares.

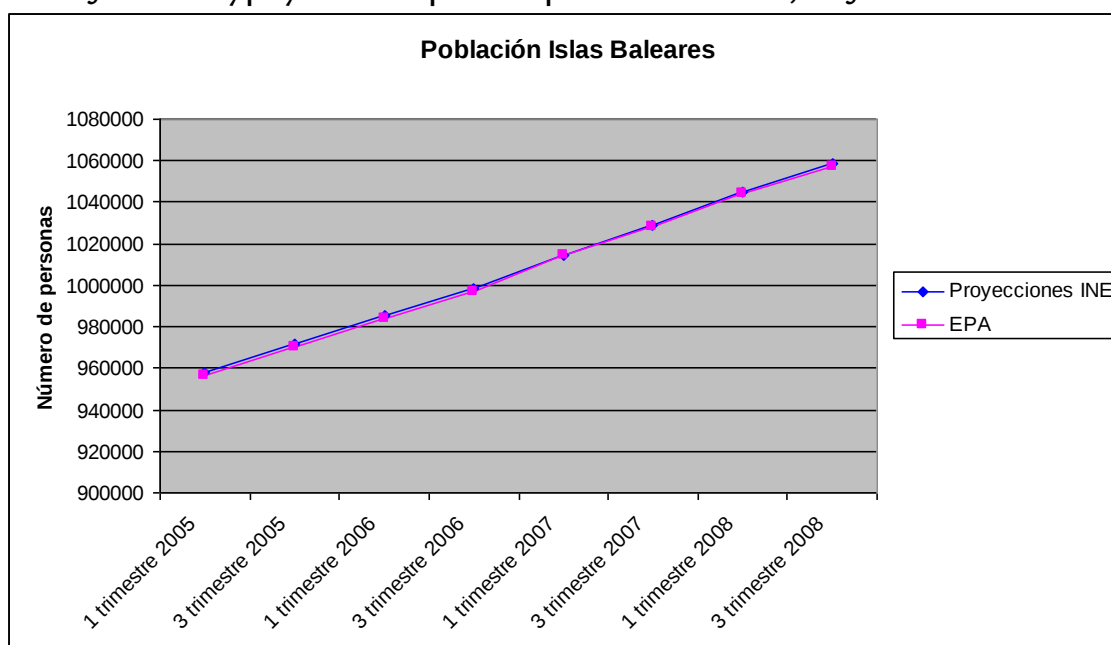


En el Cuadro 1 se muestra la evolución de la población de las Islas (para el 1 de enero y el 1 de julio del periodo 2005-2008) según es estimada a partir de la EPA y las proyecciones del INE de población, calculadas a partir del censo de 2001 (fuente: INE, Estimaciones de la Población Actual, ePOBa). Los datos se representan en la Gráfica 5, resultando evidente que el proceso de calibración cumple el objetivo de equilibrio en el total de población de la Comunidad Autónoma, cuando se compara con la ePOBa.

Cuadro 1. Evolución de la población de las Islas Baleares, 2005-2008

	Población Islas Baleares	
	EPA-Ponderados	Proyecciones INE
1 trimestre 2005	956.673	957.953
3 trimestre 2005	970.188	971.779
1 trimestre 2006	984.223	985.620
3 trimestre 2006	997.000	998.054
1 trimestre 2007	1.014.250	1.014.405
3 trimestre 2007	1.027.886	1.028.635
1 trimestre 2008	1.043.845	1.045.008
3 trimestre 2008	1.057.259	1.058.668

Gráfica 5. Población y proyecciones de población para las Islas Baleares, 2005-2008

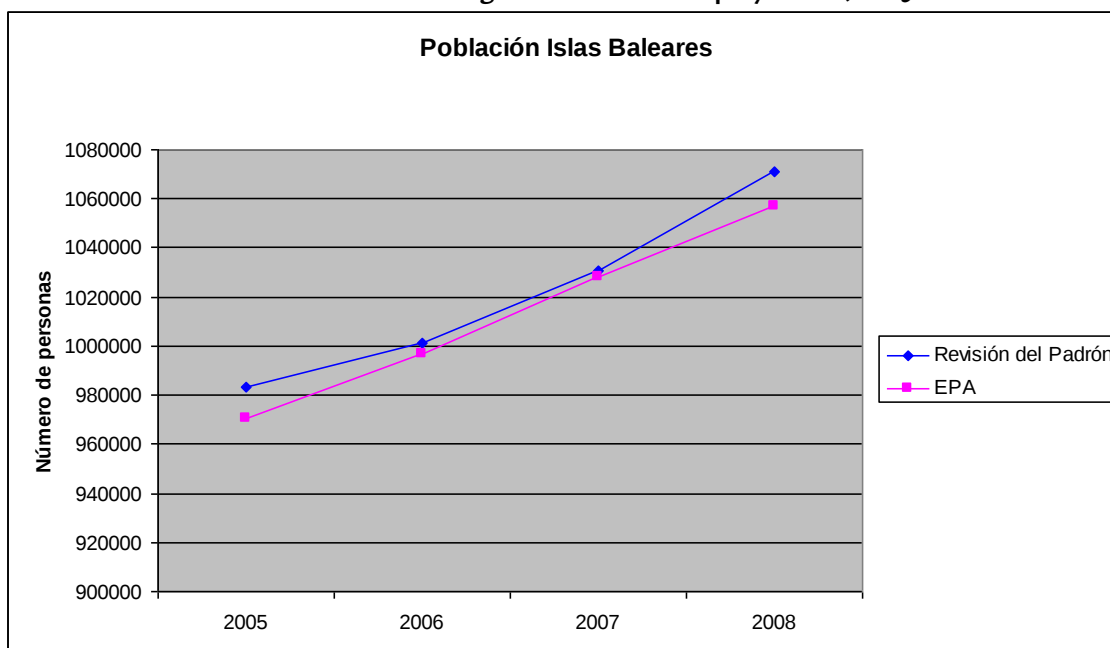


El objetivo de equilibrio, sin embargo, no se alcanza cuando se compara la población de la EPA con las revisiones del padrón municipal con referencia a 1 de enero de cada año (INE. Revisión del Padrón municipal). Los datos del INE correspondientes a las Proyecciones del INE y a la Revisión del Padrón Municipal no son totalmente coincidentes. Como puede comprobarse en el Cuadro 2, en el periodo 2005-2008, las Proyecciones del INE parecen infravalorar la población de Baleares. De manera equivalente, la población estimada que se estima en la EPA es algo menor (véase la Gráfica 6).

Cuadro 2. Población de las Islas Baleares según fuentes, 2005-2008

	Población Islas Baleares	
	Proyecciones INE	Revisión Padrón Municipal
01-ene-05	957.953	983.131
01-ene-06	985.620	1.001.062
01-ene-07	1.014.405	1.030.650
01-ene-08	1.045.008	1.071.221

Gráfica 6. Población de las Islas Baleares según el Padrón municipal y la EPA, 2005-2008



4.3. Nacionalidad española y extranjera

Respecto a las proyecciones de la población derivadas del Censo de 1991, una de las principales características de las nuevas proyecciones, es la revisión de las hipótesis sobre inmigración. Esta revisión se introduce en 2005 y afecta no sólo al número absoluto de inmigrantes (que se habían infraestimado anteriormente), sino que modifica los factores de elevación con el objetivo de compensar las dificultades de entrevistar a la población inmigrante. Según el INE los factores de elevación tienen en cuenta las dificultades de entrevistar a la población inmigrante, como consecuencia de diversas causas: una parte no reside en hogares familiares, su situación es en ocasiones irregular, o las dificultades del idioma que dificultan la entrevista. Uno de los objetivos de la revisión en el año 2005 de la metodología de la EPA era reflejar mejor el peso específico de los extranjeros en determinados sectores o profesiones. Sin embargo, el INE no publica las proyecciones de población correspondientes a las distintas nacionalidades. Por ello, para conocer la distribución por nacionalidades

debe recurrirse a las revisiones del Padrón municipal, con respecto a cuyo total, como se ha indicado, la EPA mantiene una infravaloración.

En el Cuadro 3 se muestran los datos de la revisión del padrón municipal realizado por el INE para el período 2005 a 2008. Los datos correspondientes a la EPA se muestran en Cuadro 4. Puede observarse que entre 2005 y 2008 la población de nacionalidad española ha aumentado en 22029 personas (una tasa anual media acumulativa del 0,88%), mientras que la de nacionalidad extranjera ha crecido en 66061 personas (con una tasa anual media acumulativa del 12,47%), lo que supone un incremento de población de 88090 personas (tasa media del 2,9%). Los datos de la EPA muestran incrementos de 6320 personas de nacionalidad española, 80852 de nacionalidad extranjera, con un total de 87172. Estos incrementos suponen tasas medias acumulativas del 0,26%, 16,09% y 2,95%, respectivamente (véase Cuadro 5).

Cuadro 3. Revisión del padrón municipal del INE para las Islas Baleares, 2005-2008

	Número de personas				Porcentajes			
	2005	2006	2007	2008	2005	2006	2007	2008
Españoles	826.861	833.311	840.480	848.890	84,1	83,2	81,6	79,3
Extranjeros	156.270	167.751	190.170	222.331	15,9	16,8	18,5	20,8
Total	983.131	1.001.062	1.030.650	1.071.221	100	100	100	100

Fuente: Revisión del padrón municipal (a 1 de enero de cada año).

Cuadro 4. EPA. Datos ponderados de población para las Islas Baleares, 2005-2008 (primer trimestre)

	Número de personas				Porcentajes			
	ene-05	ene-06	ene-07	ene-08	ene-05	ene-06	ene-07	ene-08
Españoles	813.463	813.908	818.155	819.783	85,0	82,7	80,7	78,5
Extranjeros	143.210	170.315	196.095	224.062	15,0	17,3	19,3	21,5
Total	956.673	984.223	1.014.250	1.043.845	100	100	100	100

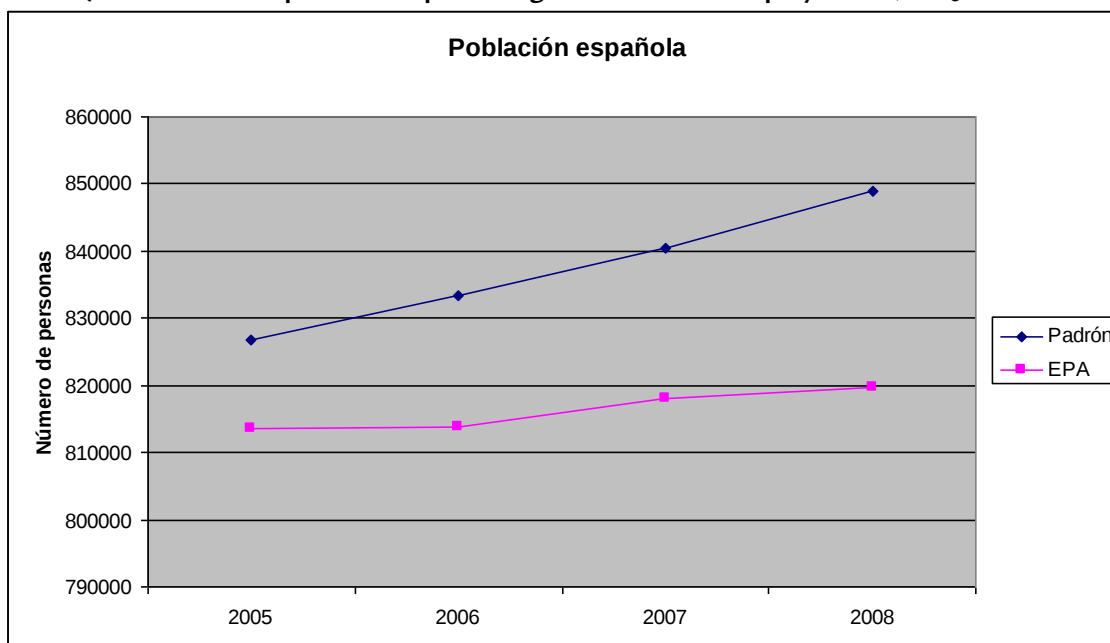
Cuadro 5. Variaciones de población para las Islas Baleares, 2005-2008 Fuente: INE, Revisión del padrón municipal

	Variación 2008-2005		Tasa media acumulativa	
	Padrón	EPA	Padrón	EPA
Españoles	22.029	6.320	0,88%	0,26%
Extranjeros	66.061	80.852	12,47%	16,09%
Total	88.090	87.172	2,90%	2,95%

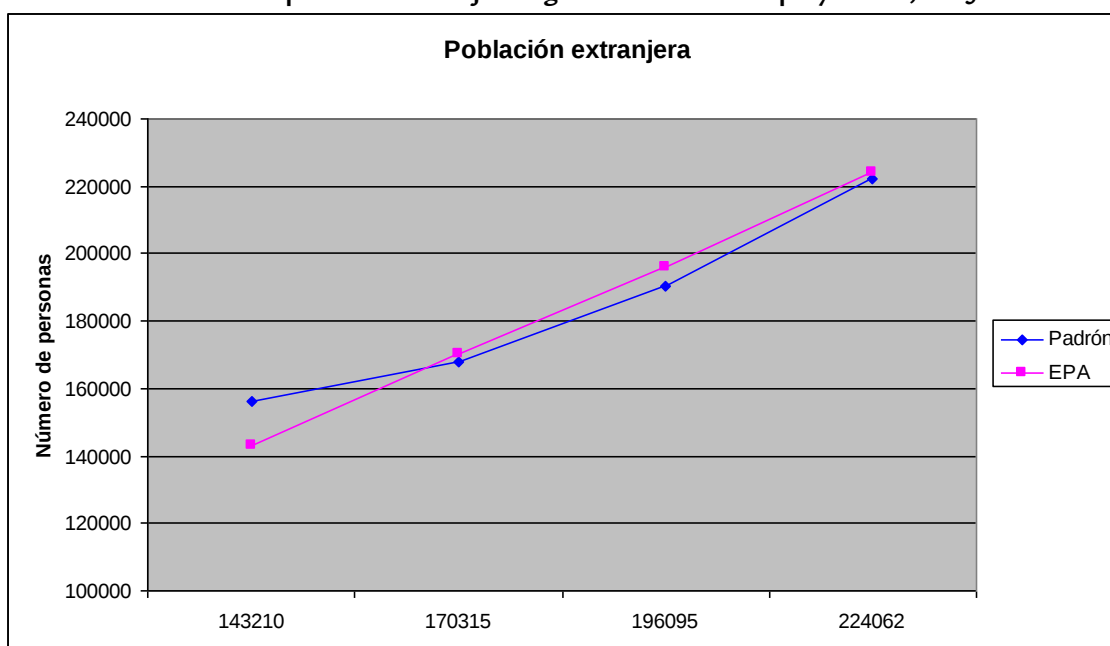
Fuente: INE, Revisión del padrón municipal

Por tanto, las cifras estimadas por la EPA (en el Cuadro 4) difieren de las obtenidas de la revisión del padrón no sólo en términos absolutos, sino también en su evolución temporal. La comparación de ambas fuentes muestra que la EPA (véanse las Gráficas 7 y 8) subestima de manera sistemática la población de nacionalidad española, no corrigiéndose este sesgo en el período (véase el Cuadro 5). No ocurre lo mismo con la población extranjera, que sí se presenta equilibrada respecto al Padrón. Este resultado no parece tener una consecuencia importante en la distribución estimada de la población según nacionalidad, aunque la EPA presenta un porcentaje ligeramente mayor de extranjeros que el Padrón.

Gráfica 7. Evolución de la población española según el Padrón municipal y la EPA, 2005-2008



Gráfica 8. Evolución de la población extranjera según el Padrón municipal y la EPA, 2005-2008



El peso que presenta la nacionalidad extranjera en los datos poblacionales de la EPA son, en una parte importante, resultado de la calibración. El análisis de los datos no ponderados muestra (véanse los Cuadros 6 y 7 y la Gráfica 9) que el porcentaje de encuestados españoles presenta una tendencia clara a su disminución, pasando de un 91,37% en el primer trimestre de 2005 a un 85,76% en el tercer trimestre de 2008. No obstante, la muestra sobrestima excesivamente el peso de la nacionalidad española. Cuando los datos son elevados a nivel poblacional, la tendencia de la serie se agudiza además de disminuir de manera importante el porcentaje, que pasan de 85,03% (primer trimestre de 2005) a 77,64% (tercer trimestre de 2008). Los resultados

indican, por tanto, que parece existir un fuerte sesgo en la representación de la población de nacionalidad extranjera en la muestra. Aunque el uso de los factores de elevación poblacional corrige ese sesgo, la importante diferencia en los porcentajes de los datos no ponderados frente a los ponderados señala la existencia de una muy deficiente representación de la población extranjera en la muestra.

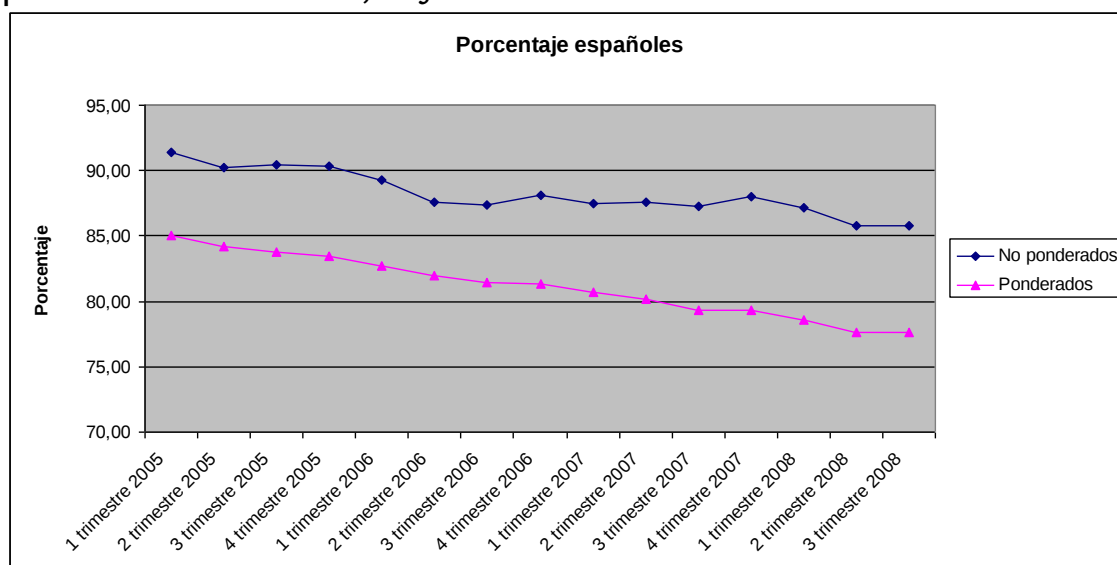
Cuadro 6. EPA (3 trimestre). Datos no ponderados de población, Islas Baleares, 2005-2008

	2005	2006	2007	2008	2005	2006	2007	2008
Españoles	3.407	3.481	3.525	3.452	90,42	87,33	87,27	85,76
Extranjeros	361	505	514	573	9,58	12,67	12,73	14,23
Total	3.768	3.986	4.039	4.025	100	100	100	100

Cuadro 7. EPA. Porcentaje de población por nacionalidad, Islas Baleares, 2005-2008

	Datos no ponderados		Datos ponderados	
	Españoles	Extranjeros	Españoles	Extranjeros
1 trimestre 2005	91,37	8,63	85,03	14,97
2 trimestre 2005	90,25	9,75	84,24	15,76
3 trimestre 2005	90,42	9,58	83,73	16,27
4 trimestre 2005	90,30	9,70	83,44	16,56
1 trimestre 2006	89,33	10,67	82,70	17,30
2 trimestre 2006	87,61	12,39	81,94	18,06
3 trimestre 2006	87,33	12,67	81,43	18,57
4 trimestre 2006	88,11	11,89	81,34	18,66
1 trimestre 2007	87,47	12,53	80,67	19,33
2 trimestre 2007	87,60	12,40	80,20	19,80
3 trimestre 2007	87,27	12,73	79,35	20,65
4 trimestre 2007	88,03	11,97	79,32	20,68
1 trimestre 2008	87,19	12,81	78,53	21,47
2 trimestre 2008	85,76	14,24	77,64	22,36
3 trimestre 2008	85,76	14,24	77,64	22,36

Gráfica 9. Porcentaje de personas de nacionalidad española en la EPA. Datos ponderados y no ponderados de las Islas Baleares, 2005-2008



4.4. Nacionalidades extranjeras

Según las revisiones del Padrón municipal la distribución de las distintas nacionalidades en las Islas presenta un peso importante de los extranjeros comunitarios (50,65% de los extranjeros en 2007) y aquellos nacidos en Sudamérica (27,92%). La EPA no presenta cifras equilibradas en esta distribución (véase la Gráfica 10). Los valores poblacionales de la EPA señalan también las anteriores zonas como las más importantes, pero con porcentajes muy distintos, representando los extranjeros comunitarios en el tercer trimestre de 2008 un 25,24% de los extranjeros, mientras que el origen sudamericano corresponde a un 51,63%. Además, debe considerarse la gran variabilidad de las cifras estimadas en la EPA. En los Cuadros 8 (datos ponderados) y 9 (datos no ponderados) se muestran los resultados correspondientes al tercer trimestre de la EPA de los años 2006, 2007 y 2008. Puede observarse que los porcentajes correspondientes a las distintas nacionalidades sufren fuertes variaciones. Esta variabilidad en la composición por nacionalidades no está respaldada por los datos de las revisiones del Padrón (véase el Cuadro 10), por lo que debe atribuirse al muestreo.

Cuadro 8. EPA. Personas nacidas en el extranjero. Datos ponderados. (Tercer trimestre)

	Número de personas			Porcentajes		
	2006	2007	2008	2006	2007	2008
Resto de Europa (no UE)	15.647	20.573	22.836	8,45	9,69	9,66
UE-15	55.652	57.926	57.609	30,06	27,29	24,37
UE-25 (no UE-15)	2.701	996	2.063	1,46	0,47	0,87
Europa	74.000	79.495	82.508	39,97	37,45	34,90
África	36.245	29.610	22.798	19,57	13,95	9,65
América del Norte	2.193	1.397	288	1,18	0,66	0,12
Centroamérica y Caribe	3.009	6.911	5.122	1,63	3,26	2,17
Sudamérica	63.404	91.690	122.028	34,24	43,19	51,63
Asia Oriental	3.364	2.407	2.240	1,82	1,13	0,95
Asia del Sur y Sudoeste	2.790	771	900	1,51	0,36	0,38
Oceanía	161	0	470	0,09	0,00	0,20
Total	185.166	212.281	236.354	100	100	100

Cuadro 9. EPA. Personas nacidas en el extranjero. Datos no ponderados. (Tercer trimestre)

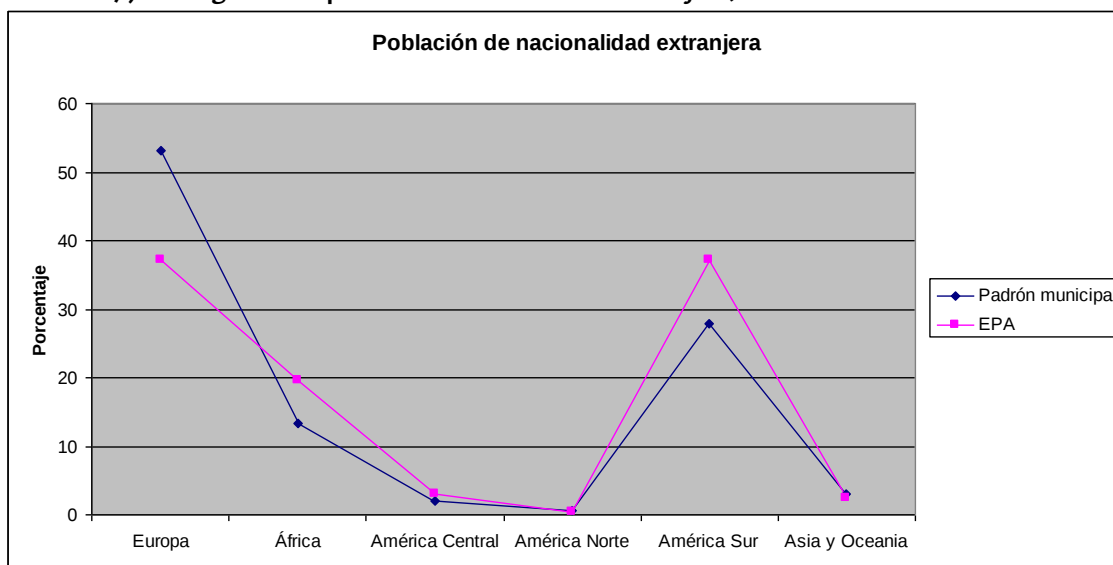
	2006	2007	2008	2006	2007	2008
Resto de Europa (no UE)	37	48	48	7,33	9,34	8,38
UE-15	163	147	139	32,28	28,60	24,26
UE-25 (no UE-15)	6	3	5	1,19	0,58	0,87
Europa	206	198	192	40,80	38,52	33,51
África	90	65	50	17,82	12,65	8,73
América del Norte	5	3	1	0,99	0,58	0,17
Centroamérica y Caribe	10	21	16	1,98	4,09	2,79
Sudamérica	176	218	305	34,85	42,41	53,23
Asia Oriental	9	7	6	1,78	1,36	1,05
Asia del Sur y Sudoeste	8	2	2	1,58	0,39	0,35
Oceanía	1	0	1	0,20	0,00	0,17
Total	505	514	573	100	100	100

Cuadro 10. INE. Personas nacidas en el extranjero

	Número de personas		Porcentajes	
	2006	2007	2006	2007
Unión Europea (25 y 27)	72.637	96.329	43,30	50,65
Resto Europa	13.469	4.605	8,03	2,42
Europa	86.106	100.934	51,33	53,07
África	23.666	25.495	14,11	13,41
América central	3.419	3.784	2,04	1,99
América Norte	1.197	1.247	0,71	0,66
América Sur	48.135	53.101	28,69	27,92
Asia	5.060	5.453	3,02	2,87
Oceanía	168	155	0,10	0,08
Total	167.751	190.169	100	100

Fuente: Revisión del Padrón municipal

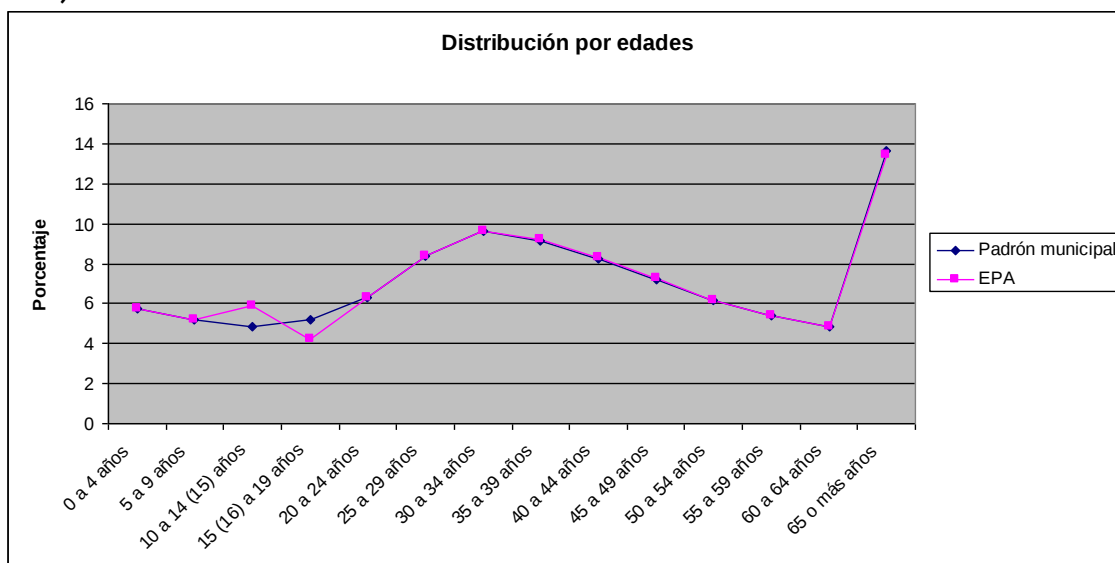
Gráfica 10. Comparación entre la EPA (primer trimestre 2007) y la Revisión del Padrón Municipal (1 de enero 2007) del origen de las personas de nacionalidad extranjera, Islas Baleares.



4.5. Edad

Como se ilustra en la Gráfica 11, con los datos del año 2008, la distribución por edades de la EPA está equilibrada respecto a la obtenida en las revisiones del Padrón Municipal. En los Cuadros 11 y 12 se muestran las distribuciones de edad correspondientes al Padrón Municipal y a la EPA, para los años 2005 a 2008.

Gráfica II. Distribución por edades según Padrón municipal (1 enero 2008) y EPA (primer trimestre 2008) en las Islas Baleares



Cuadro II. Distribución de la población por edades en las Islas Baleares según el Padrón municipal, 2005-2008

	2005	2006	2007	2008
0 a 4 años	5,61	5,68	5,77	5,75
5 a 9 años	4,98	5,03	5,07	5,18
10 a 14 años	5,11	5,00	4,91	4,87
15 a 19 años	5,43	5,34	5,27	5,22
20 a 24 años	6,81	6,60	6,43	6,33
25 a 29 años	8,95	8,80	8,59	8,39
30 a 34 años	9,50	9,60	9,64	9,62
35 a 39 años	8,81	8,91	9,02	9,16
40 a 44 años	7,94	8,07	8,19	8,26
45 a 49 años	6,90	7,01	7,11	7,23
50 a 54 años	6,01	6,03	6,07	6,14
55 a 59 años	5,51	5,50	5,45	5,41
60 a 64 años	4,52	4,65	4,76	4,82
65 o más años	13,92	13,79	13,71	13,62
Total	100	100	100	100

Fuente: INE, Revisión del Padrón municipal

Cuadro 12. Distribución de la población por edades en las Islas Baleares según la EPA con los datos ponderados, 2005-2008

	2005	2006	2007	2008
de 0 a 4	5,64	5,71	5,78	5,76
de 5 a 9	5,00	5,05	5,10	5,20
de 10 a 15	6,16	6,04	5,94	5,88
de 16 a 19	4,39	4,30	4,25	4,22
de 20 a 24	6,79	6,59	6,43	6,33
de 25 a 29	8,94	8,78	8,59	8,38
de 30 a 34	9,54	9,63	9,66	9,63
de 35 a 39	8,84	8,95	9,06	9,19
de 40 a 44	7,98	8,11	8,22	8,30
de 45 a 49	6,93	7,04	7,14	7,26
de 50 a 54	6,03	6,05	6,09	6,17
de 55 a 59	5,51	5,50	5,45	5,41
de 60 a 64	4,54	4,67	4,77	4,83
65 o más	13,70	13,59	13,50	13,43
Total	100	100	100	100

Fuente: EPA. Primer trimestre de cada año

Cuadro 13. Distribución por edades en la muestra de la EPA de las Islas Baleares con los datos no ponderados, 2005-2008

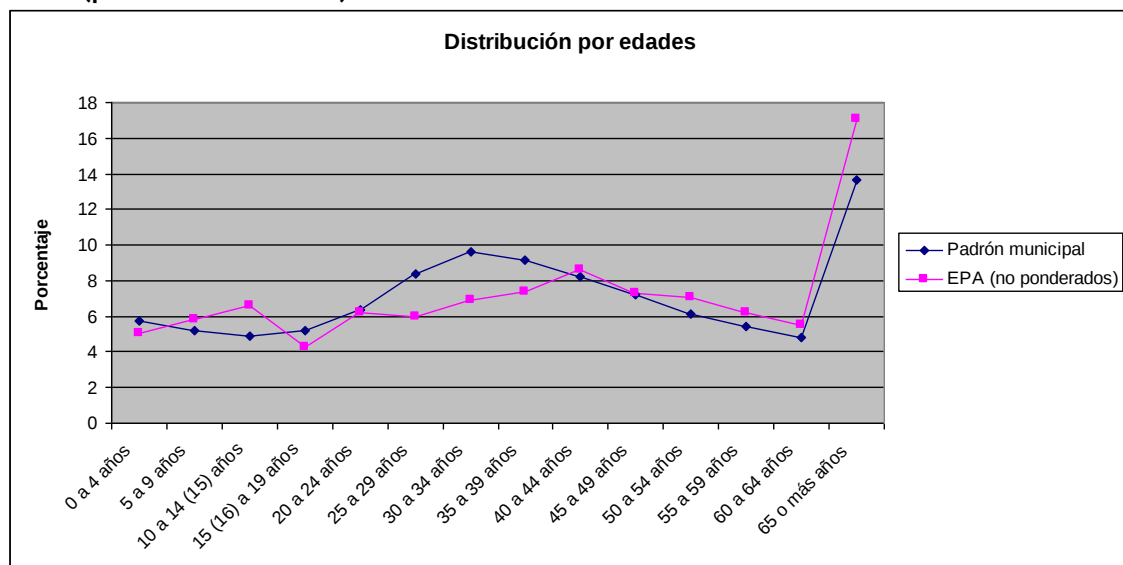
	2005	2006	2007	2008
0 a 4 años	5,62	5,61	5,17	5,05
5 a 9 años	4,92	5,24	5,92	5,81
10 a 15 años	6,33	6,50	6,28	6,59
16 a 19 años	5,07	5,06	4,97	4,29
20 a 24 años	6,09	5,32	5,36	6,20
25 a 29 años	7,47	6,37	5,92	6,00
30 a 34 años	7,94	7,36	7,94	6,90
35 a 39 años	7,72	8,35	8,57	7,37
40 a 44 años	7,77	7,75	7,97	8,59
45 a 49 años	7,57	7,83	7,61	7,32
50 a 54 años	6,98	6,37	5,84	7,05
55 a 59 años	5,67	5,27	5,46	6,22
60 a 64 años	4,95	4,93	5,53	5,54
65 o más años	15,91	18,05	17,46	17,08
Total	100	100	100	100

Fuente: EPA. Primer trimestre de cada año. Datos no ponderados

No obstante, si se analiza la distribución por edades con los datos no ponderados de la EPA (en el Cuadro 13) se detecta un importante sesgo. En la Gráfica 12 se comparan las distribuciones de edad que se obtienen de la EPA (primer trimestre 2008), empleando los datos sin ponderar, con la del Padrón Municipal (a 1 de enero de

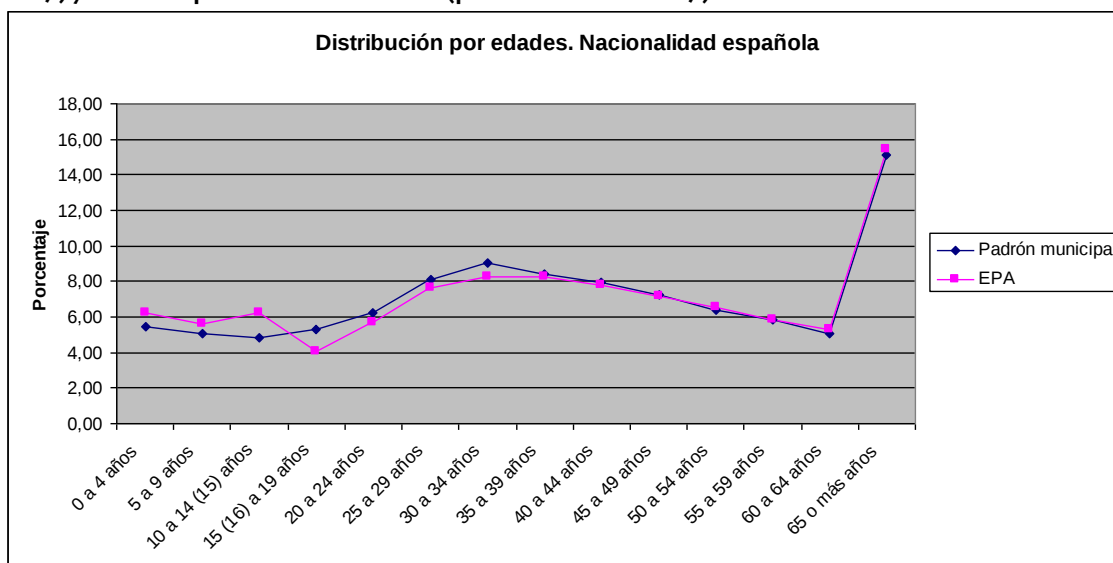
2008). Puede observarse que previa a su ponderación, la distribución por edades que proporciona la muestra de la EPA no estima correctamente la estructura de edades. Los factores de elevación poblacional aseguran que, al elevar los datos a población, la distribución por edades sea coincidente. No obstante, el hecho es que el diseño muestral de la EPA subestima la importancia del colectivo entre 25 y 39 años (sobre este tema, puede consultarse Álvarez, 2000). Esta subestimación pone de nuevo en evidencia el problema de la EPA para entrevistar a determinados hogares.

Gráfica 12. Distribución por edades según Padrón municipal (1 enero 2008) y datos no ponderados de la EPA (primer trimestre 2008).

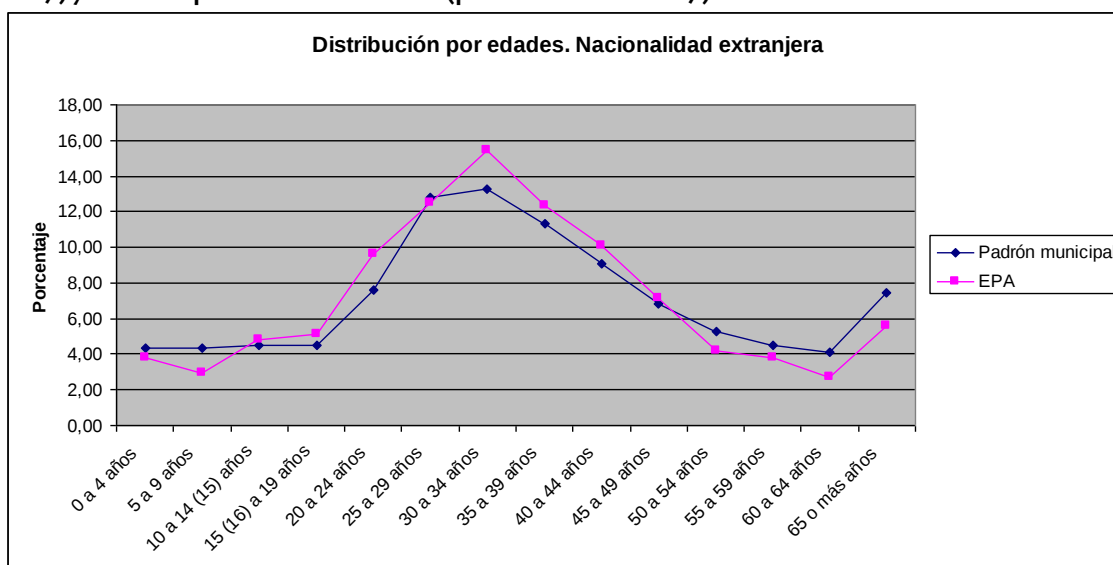


El análisis de la distribución por edad puede hacerse distinguiendo entre la población de nacionalidad española y extranjera. En este caso, se pone en evidencia que la subestimación detectada en el total poblacional no es uniforme en las dos subpoblaciones. Como se muestra en las Gráficas 13 y 14, mientras que para la nacionalidad española la EPA subestima en una ligera cuantía el tramo entre 20 y 34 años, en el caso de la nacionalidad extranjera, la EPA tiende a sobreestimar la población en los tramos de edad intermedia (entre 20 y 44 años).

Gráfica 13. Distribución por edades de la subpoblación española, según Padrón municipal (1 enero 2007) y datos no ponderados de la EPA (primer trimestre 2007).

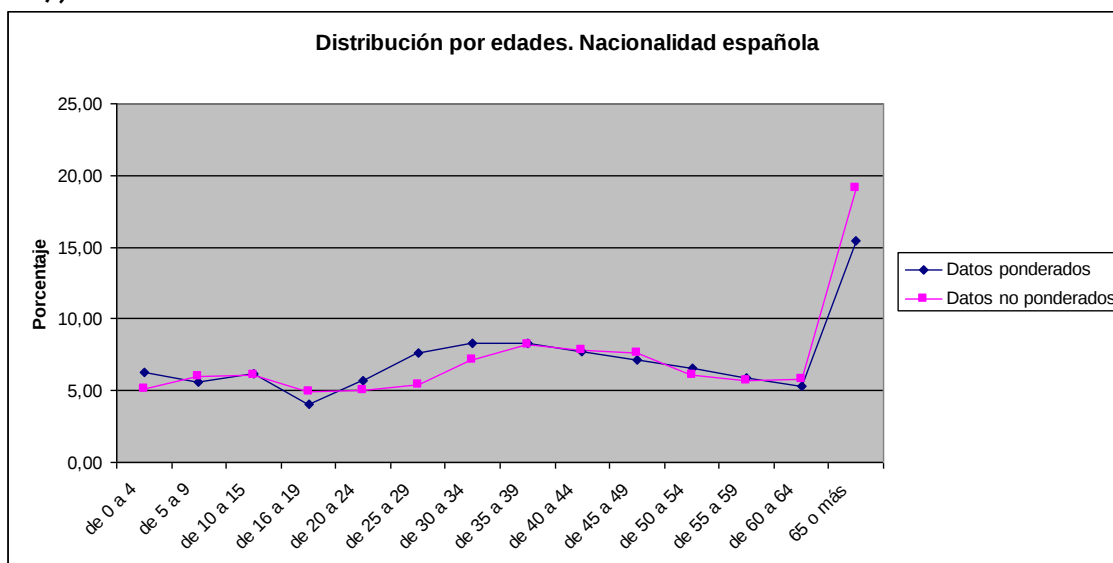


Gráfica 14. Distribución por edades de la subpoblación extranjera, según Padrón municipal (1 enero 2007) y datos no ponderados de la EPA (primer trimestre 2007).

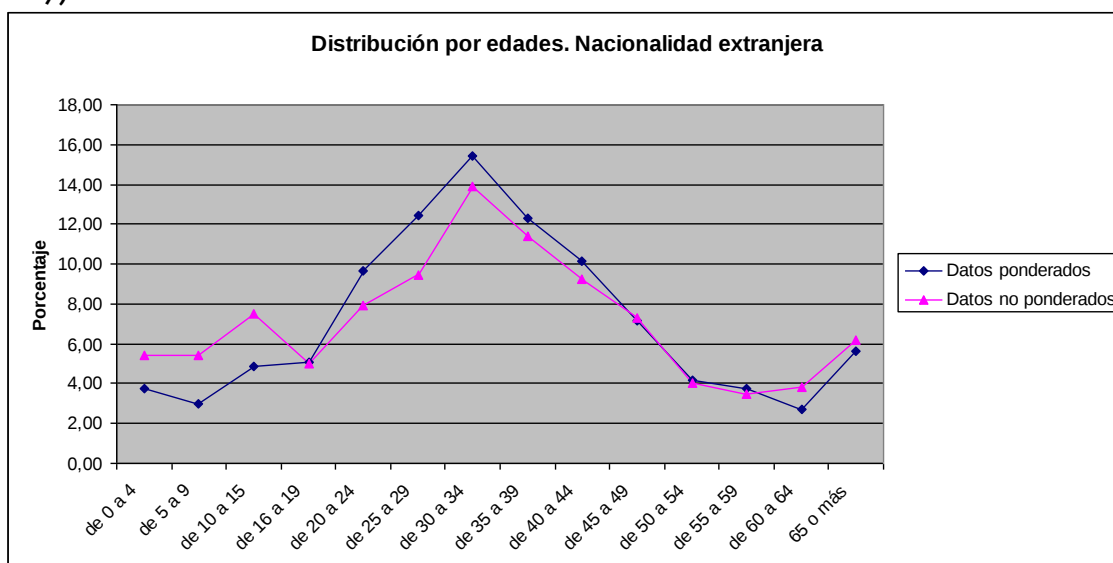


Resulta curioso observar que el proceso de ponderación actúa de una manera perversa en el caso de la población extranjera. En la Gráfica 15 se comparan, para la nacionalidad española, las distribuciones de la EPA con datos ponderados y son ponderar. En este caso, la elevación de los datos a la población aumenta el peso de los tramos de edad que estarían subestimados respecto al Padrón. Sin embargo, en el caso de la nacionalidad extranjera, los datos ponderados (Gráfica 16) aumentan la representación de los tramos que se han detectado como sobrerrepresentados respecto al Padrón. De hecho, la distribución de los datos no ponderados se ajusta mejor a los datos del Padrón, que la que se obtiene tras elevar los datos (Gráfica 17).

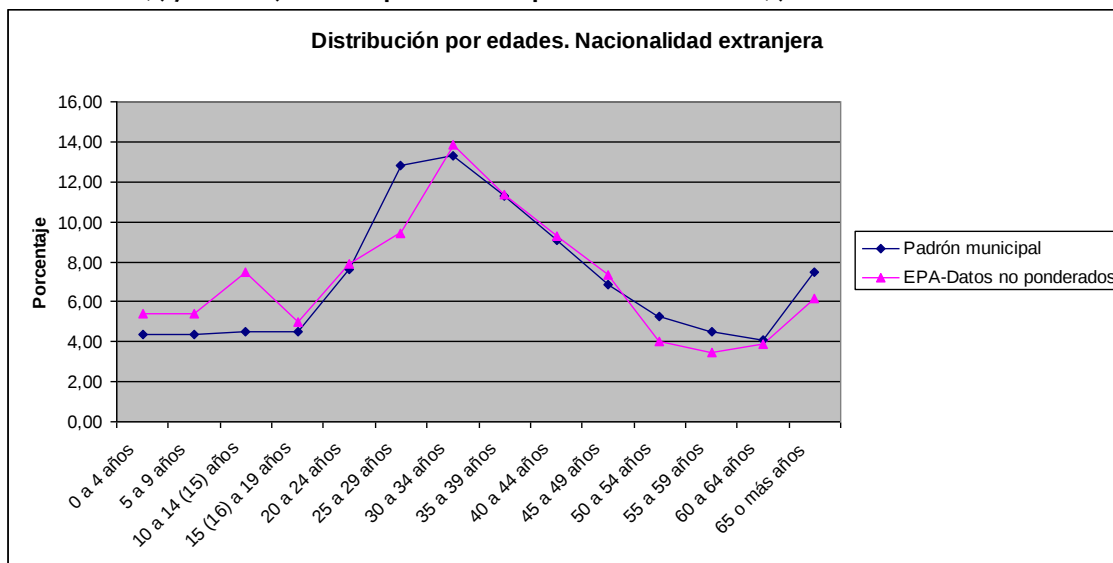
Gráfica 15. Distribución por edades de la subpoblación española, según la EPA (primer trimestre 2007).



Gráfica 16. Distribución por edades de la subpoblación extranjera, según la EPA (primer trimestre 2007).



Gráfica 17. Distribución por edades de la subpoblación extranjera, según el Padrón Municipal (1 de enero de 2007) y la EPA (datos no ponderados, primer trimestre 2007).



4.6. Sexo

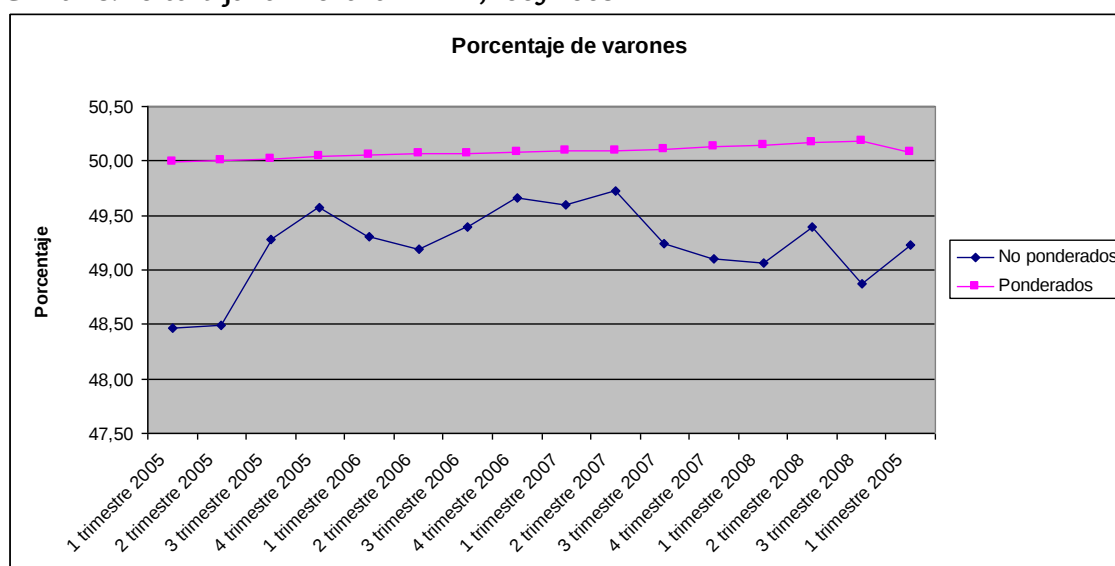
La muestra de la EPA a nivel nacional sobrestima la proporción de mujeres (51,6%, en 2007-2008) frente al Padrón Municipal (50,6%). Una situación parecida se da en la muestra de Baleares, en donde la muestra de la EPA (véanse el Cuadro 14 y la Gráfica 18) proporciona (en el tercer trimestre de 2008) un 51,13% de mujeres. La elevación de la muestra a población corrige esta proporción, situándola en 49,8%, igual al porcentaje que se recoge en el Padrón.

Cuadro 14. Porcentaje de población por sexos de las Islas Baleares

	Datos no ponderados		Datos Ponderados	
	Hombres	mujeres	hombres	Mujeres
01-ene-05	48,47	51,53	50	50
01-feb-05	48,49	51,51	50	50
01-mar-05	49,28	50,72	50,02	49,98
01-abr-05	49,57	50,43	50,04	49,96
01-ene-06	49,31	50,69	50,06	49,94
01-feb-06	49,19	50,81	50,06	49,94
01-mar-06	49,40	50,60	50,07	49,93
01-abr-06	49,67	50,33	50,08	49,92
01-ene-07	49,60	50,40	50,09	49,91
01-feb-07	49,73	50,27	50,09	49,91
01-mar-07	49,24	50,76	50,11	49,89
01-abr-07	49,10	50,90	50,13	49,87
01-ene-08	49,06	50,94	50,15	49,85
01-feb-08	49,39	50,61	50,16	49,84
01-mar-08	48,87	51,13	50,18	49,82

Fuente: EPA

Gráfica 18. Porcentaje de varones en la EPA, 2005-2008



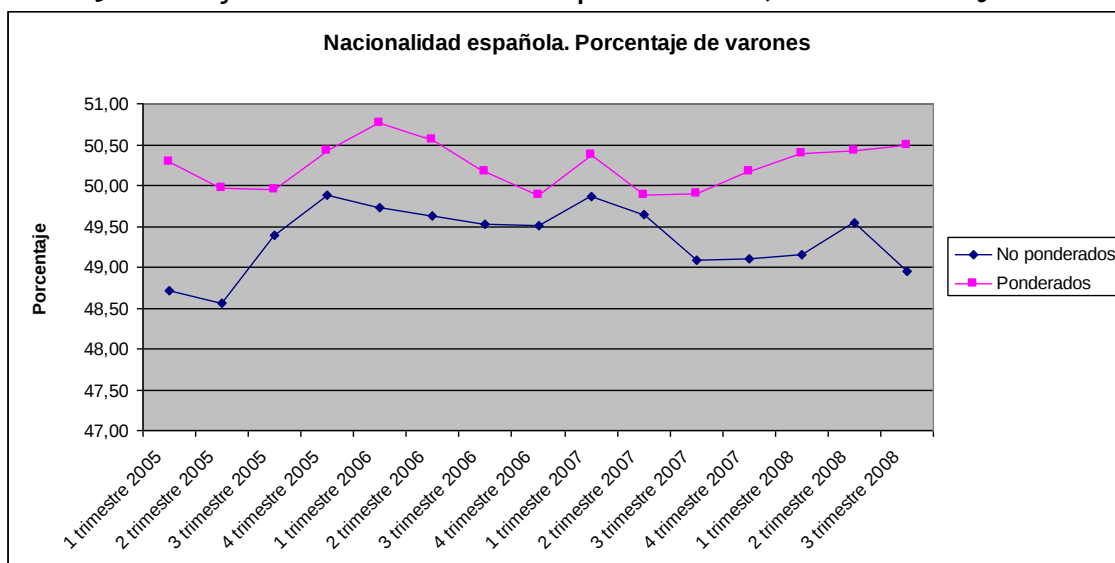
La distribución por sexos cuando se distingue por nacionalidades (españoles y extranjeros) sólo se equilibra con respecto al Padrón municipal para la nacionalidad española (véase el Cuadro 15). La población extranjera, aún después de ponderar la muestra presenta en la EPA un menor porcentaje de varones. En el Padrón municipal, del total de extranjeros, el 51,5% son varones. En la EPA, una vez ponderada la muestra este porcentaje es del 48,92%.

Cuadro 15. Población de las Islas Baleares por sexos, 2007

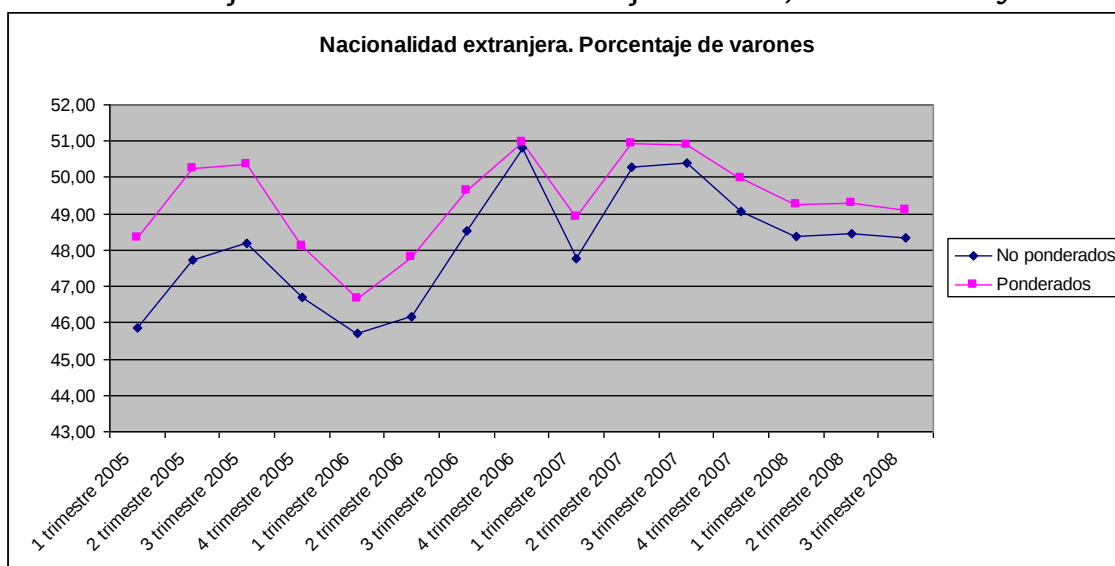
	Revisión Padrón 2007 (1 enero 2007)				EPA (1 trimestre 2007)			
	Varones	Mujeres	Total	% Varones	Varones	Mujeres	Total	% Varones
Total	517.593	513.057	1.030.650	50,22	508.007	506.243	1.014.250	50,09
Espanoles	419.465	421.015	840.480	49,91	412.076	406.078	818.154	50,37
Extranjeros	98.128	92.042	190.170	51,60	95.931	100.165	196.096	48,92

En el caso de la nacionalidad extranjera, la falta de precisión de la estimación de la distribución por sexo en la EPA se muestra cuando se observa su variabilidad en el periodo analizado (véanse el Cuadro 16 y las Gráficas 19 y 20). Mientras el porcentaje de varones en la subpoblación extranjera en el Padrón se mantiene estable (en los años 2005, 2006 y 2007 se sitúa en 51,67%, 51,52% y 51,60%, respectivamente), el porcentaje de varones oscila en la EPA (entre 1-2005 y 3-2008) con un mínimo de 46,67% y un máximo de 50,96%. Similar variabilidad se muestra sobre la muestra no ponderada (con un mínimo de 45,72% y un máximo de 50,83%).

Gráfica 19. Porcentaje de varones de nacionalidad española en la EPA, Islas Baleares 2005-2008.



Gráfica 20. Porcentaje de varones de nacionalidad extranjera en la EPA, Islas Baleares 2005-2008.



Cuadro 16. Distribución por sexos, según nacionalidad

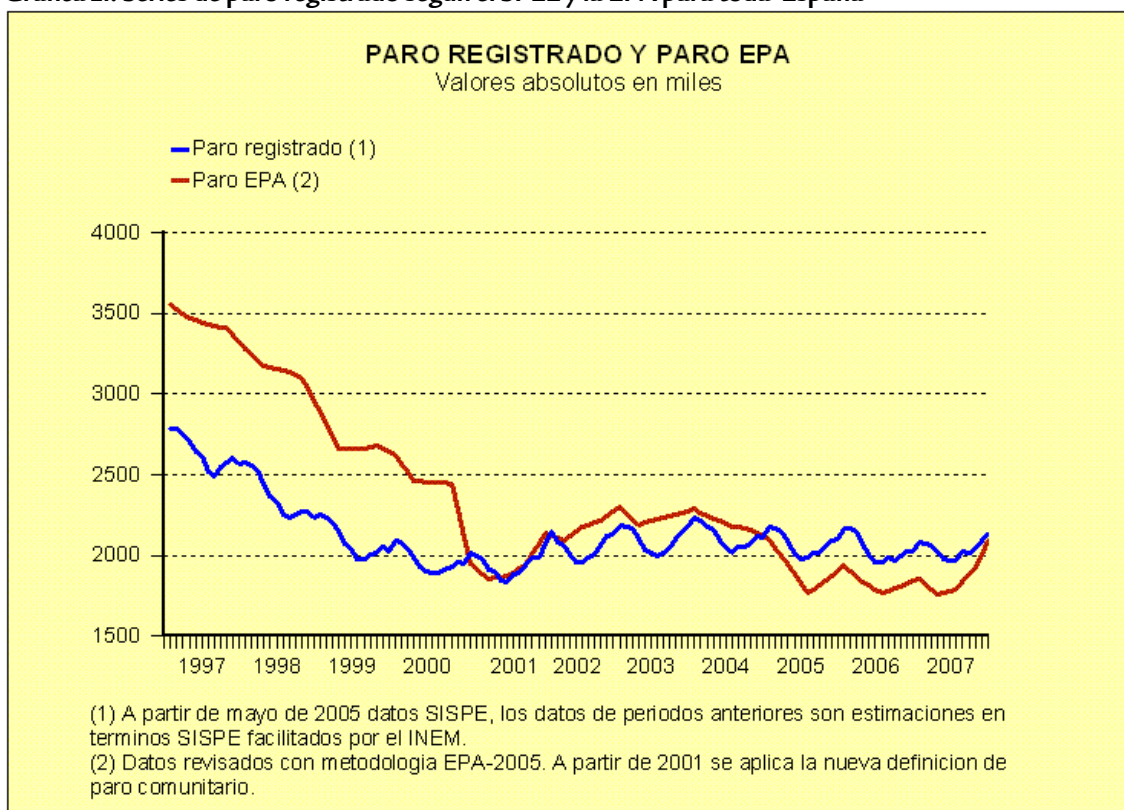
	Española				Extranjera			
	Datos no ponderados		Datos ponderados		Datos no ponderados		Datos ponderados	
	hombres	mujeres	hombres	mujeres	hombres	mujeres	hombres	mujeres
ene-05	48,71	51,29	50,29	49,71	45,85	54,15	48,34	51,66
feb-05	48,57	51,43	49,96	50,04	47,75	52,25	50,24	49,76
mar-05	49,40	50,60	49,96	50,04	48,20	51,80	50,35	49,65
abr-05	49,88	50,12	50,42	49,58	46,70	53,30	48,12	51,88
ene-06	49,74	50,26	50,76	49,24	45,72	54,28	46,67	53,33
feb-06	49,62	50,38	50,56	49,44	46,15	53,85	47,82	52,18
mar-06	49,53	50,47	50,17	49,83	48,51	51,49	49,62	50,38
abr-06	49,51	50,49	49,88	50,12	50,83	49,17	50,96	49,04
ene-07	49,86	50,14	50,37	49,63	47,78	52,22	48,92	51,08
feb-07	49,65	50,35	49,89	50,11	50,29	49,71	50,92	49,08
mar-07	49,08	50,92	49,90	50,10	50,39	49,61	50,89	49,11
abr-07	49,11	50,89	50,17	49,83	49,06	50,94	49,98	50,02
ene-08	49,16	50,84	50,39	49,61	48,38	51,62	49,27	50,73
feb-08	49,54	50,46	50,42	49,58	48,46	51,54	49,28	50,72
mar-08	48,96	51,04	50,49	49,51	48,34	51,66	49,10	50,90

Fuente: EPA

4.7. Situación laboral

Dado que la EPA ofrece información sobre las personas inscritas en los servicios públicos de empleo, es posible cuantificar su número para lograr una comparación más homogénea con el paro registrado. Prudencio y Toharia (2007) han analizado el equilibrio entre la EPA y los parados inscritos en las oficinas de empleo públicas, concluyendo que la EPA aproxima correctamente las cifras de los Servicios Públicos de Empleo (SPE). No obstante, las series de paro registrado y paro EPA para toda España (Gráfica 21) muestran que, aún a partir de 2005, muestran que las series discrepan y lo hacen en distintos sentidos (Ministerio de Trabajo e Inmigración, 2007).

Gráfica 21. Series de paro registrado según el SPEE y la EPA para toda España



Fuente: Ministerio de Trabajo e Inmigración, 2007

Las discrepancias entre la EPA y el paro registrado son conocidas (véase, por ejemplo, Pérez 2008), por lo que puede no resultar apropiado comparar la EPA con el paro registrado. De hecho, la disparidad entre ambas fuentes no permite saber a priori cuál de las dos cuantifica una cifra mayor. Como señala Pérez (2008) hay dos razones que explicarían que la cifra de la EPA sea mayor:

- (1) No todos los parados de la EPA buscan empleo a través de las oficinas públicas de empleo.
- (2) No todos los inscritos como parados en las oficinas públicas de empleo se consideran parados registrados (por ejemplo, estudiantes menores de 25 años, demandantes de primer empleo, personas que demandan empleo de corta duración o de jornada reducida).

Pero existen también dos razones que justificarían que el paro registrado sea la cifra más alta:

- (1) Los parados registrados que no buscan activamente empleo o no están disponibles para trabajar no son considerados parados por la EPA.
- (2) La EPA tampoco considera los parados registrados que no han tenido contacto con la oficina pública de empleo en las cuatro semanas anteriores, con la finalidad de encontrar trabajo.

Teniendo en cuenta los anteriores comentarios, a continuación se comparan la información del SPEE sobre paro registrado y la de la EPA. Los datos de la EPA sobre el número de parados registrados se deriva de la siguiente pregunta de la

encuesta: “El domingo de la semana de referencia, ¿estaba inscrito en una oficina de empleo pública?”. Debe tenerse en cuenta que las cifras del SPEE corresponde al paro registrado en el mes de Septiembre de 2008, mientras que la cifras de la EPA son las del tercer trimestre de 2008.

Según datos del SPEE, el paro registrado en las Islas Baleares en septiembre de 2008 era de 46807, con la distribución por sexos y edad que se muestra en el Cuadro 16. Según la EPA (3 trimestre de 2008) el paro registrado es de 46426. Las distribuciones por edad y sexo de ambas fuentes de información se pueden comparar a partir de los Cuadros 17 y 18. Los porcentajes totales por sexo se aproximan, aunque el porcentaje estimado de mujeres es algo mayor en la EPA (45,47% que en el SPEE (43,3%). La distribución por edades no coincide. En la EPA el porcentaje de parados registrados menores de 25 años sobrepasa en mucho el porcentaje del SPEE (24,25% en la EPA, frente al 14,33% del SPEE). Los porcentajes correspondientes a los tramos de más edad están subestimados en la EPA: en el tramo de 25 a 44 años, 48,89% (EPA) y 55,09% (SPEE); en el tramo de más de 45 años 26,86% (EPA) y 30,57% (SPEE).

Cuadro 17. Parados registrados en el SPEE, Islas Baleares. Septiembre de 2008.

	< 25 años	25 a 44	45 o más	Total
Hombres	4.072	14.537	7.883	26.492
Mujeres	2.637	11.251	6.427	20.315
Total	6.709	25.788	14.310	46.807
	< 25 años	25 a 44	45 o más	Total
Hombres	8,70	31,06	16,84	56,60
Mujeres	5,63	24,04	13,73	43,40
Total	14,33	55,09	30,57	100

Cuadro 18. Parados registrados en oficinas de empleo públicas según la EPA

	< 25 años	25 a 44	45 o más	Total
Hombres	6.923	11.603	6.790	25.316
Mujeres	4.335	11.094	5.681	21.110
Total	11.258	22.697	12.471	46.426
	< 25 años	25 a 44	45 o más	Total
Hombres	14,91	24,99	14,63	54,53
Mujeres	9,34	23,90	12,24	45,47
Total	24,25	48,89	26,86	100

Fuente: EPA 3 trimestre 2008

Al segmentar simultáneamente por edad y sexo, los porcentajes sólo resultan parecidos en los dos tramos de 45 o más años y en el de mujeres de 25 a 44 años.

Una segunda referencia de la capacidad de la EPA para equilibrar la información de fuentes externas la proporciona la distribución de los parados por nacionalidad. El Observatori del Treball de les Illes Balears (2008) ofrece datos del SOIB de Octubre de 2008, reproducidos en el Cuadro 19. Los datos más próximos en el tiempo para la

EPA son los que aparecen en el Cuadro 20. Aunque debe tenerse en cuenta que se comparan diferentes periodos, la distribución porcentual de ambas fuentes apenas presenta discrepancia.

Cuadro 19. SOIB Octubre 2008.

	Parados	Porcentaje
Españoles	41.570	72,95
Extranjeros UE-27	4.571	8,02
Extranjeros no UE-27	10.841	19,03
Total	56.982	100

Cuadro 20. EPA 3 trimestre 2008.

	Parados Registrados	Porcentaje
Españoles	33.075	71,25
Extranjeros EU-27	3.961	8,53
Extranjeros no UE-27	9.387	20,22
Total	46.423	100

5. FACTORES DE ELEVACIÓN

El INE expone en diversos documentos los criterios estadísticos a partir de los cuáles se obtienen los factores de elevación, pero es necesario averiguar cuál es el resultado de aplicar esos pesos en las diferentes variables de la EPA.

Con la muestra de Baleares del tercer trimestre de 2008 se ha analizado la distribución de los factores de elevación. Las principales características detectadas son una importante asimetría positiva (o por la derecha) y una fuerte curtosis (véase la Gráfica 22). Estas características no son distintas a las que presentan los factores de elevación en la muestra de toda España (en la Gráfica 23). La comparación de algunos de los estadísticos de ambas distribuciones (Cuadros 21 y 22) muestra, sin embargo, que estas características están más suavizadas en las Islas Baleares, con una distribución además que presenta menos dispersión.

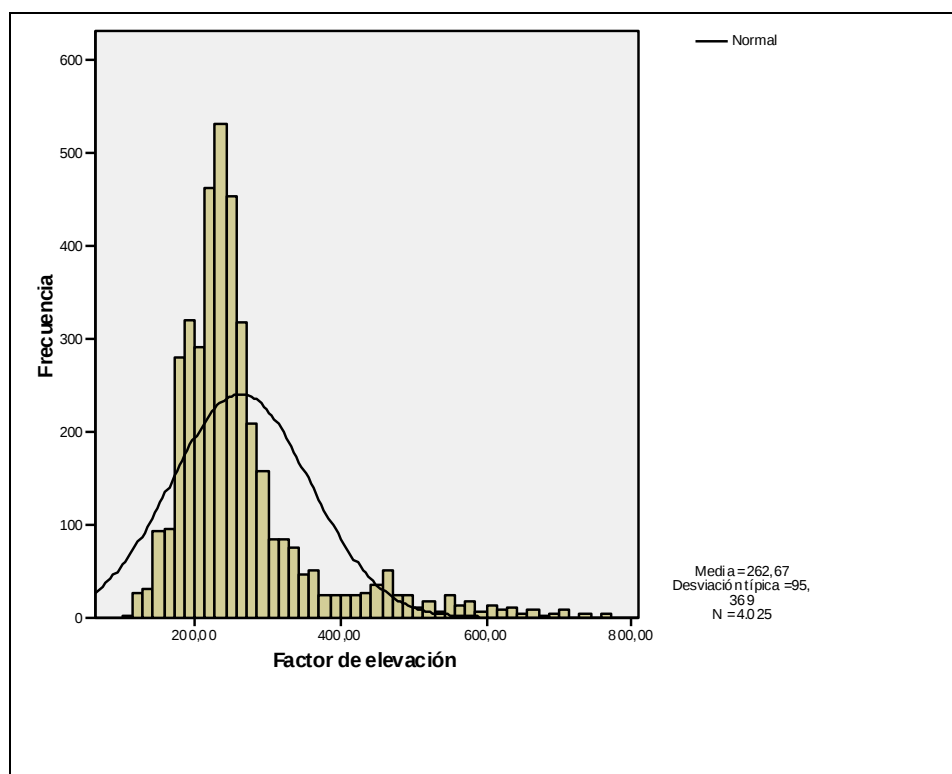
Cuadro 21. Estadísticos descriptivos de la distribución de los factores de elevación.

	Islas Baleares		España	
	Estadístico	Error Típico	Estadístico	Error típico
Media	263	2	273	0
Mediana	240		215	
Varianza	9.095		38.762	
Desv. típ.	95		197	
Mínimo	112		13	
Máximo	761		2.960	
Rango	649		2.947	
Amplitud intercuartil	69		169	
Asimetría	2	0	2	0
Curtosis	5	0	10	0

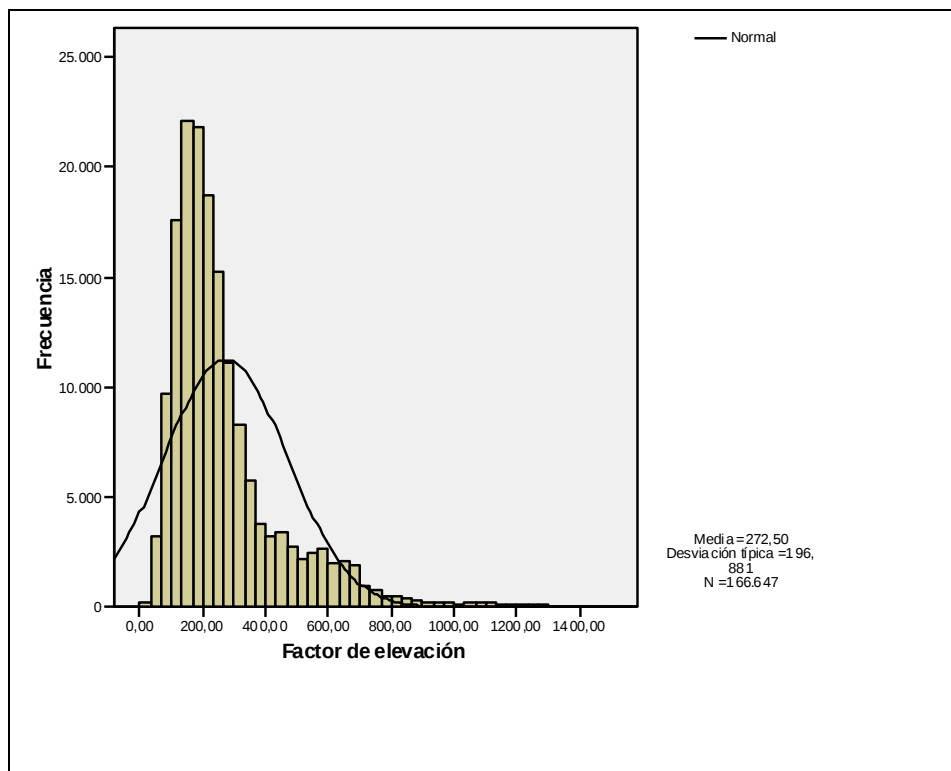
Cuadro 22. Percentiles de las distribuciones.

	5	10	25	50	75	90	95
Baleares	150	163	191	229	281	371	455
España	88	108	150	215	319	540	661

Gráfica 22. Histograma de los factores de elevación de la EPA (3 trimestre 2008), en la muestra de las Islas Baleares.



Gráfica 23. Histograma de los factores de elevación de la EPA (3 trimestre 2008), en la muestra total de España.



Las características que se observan en las distribuciones de los factores de elevación poblacional no son en sí mismas negativas. El problema puede aparecer cuando se den simultáneamente dos circunstancias:

- (1) Que los factores de elevación situados en el rango superior de la distribución se apliquen especialmente en un número limitado de segmentos de la muestra.
- (2) Que los factores de elevación se determinen en menor medida por el diseño muestral, y en mayor medida por el proceso de calibración o reponderación posterior.

Si se dan estas dos circunstancias, la elevación de los datos de la muestra a población puede generar problemas sobre la estimación de las variables que no han sido tenidas en cuenta en el proceso de calibración. Concretamente, supongamos que dado que la subpoblación extranjera no está bien representada en la muestra de la EPA, se modifican al alza los factores de elevación correspondientes a esa población. Cada uno de los individuos muestrales representará ahora un mayor número de individuos en la población. Puesto que esta subpoblación no está bien representada en la muestra, tampoco lo están las características asociadas a la misma. Cuando se reponderan los factores de elevación teniendo en cuenta que en la muestra hay pocos extranjeros, pero no se controla la representatividad de ese segmento de la muestra, las estimaciones en la población pueden resultar erróneas. El problema no sólo afectará a las características de la subpoblación de extranjeros, sino a las estimaciones del conjunto de la población.

Donde más claramente se manifiestan las consecuencias de esta situación es en los factores de elevación empleados en la subpoblación de origen extranjero, especialmente cuando se distingue por su situación respecto a la actividad económica.

En el Cuadro 23 se muestran los valores medios de los factores de elevación poblacional, calculados para cada trimestre de la EPA, distinguiendo entre ocupados, parados e inactivos, y según nacionalidad. Las series, representadas en la Gráfica 24, ponen en evidencia varias cuestiones que resulta de interés resaltar:

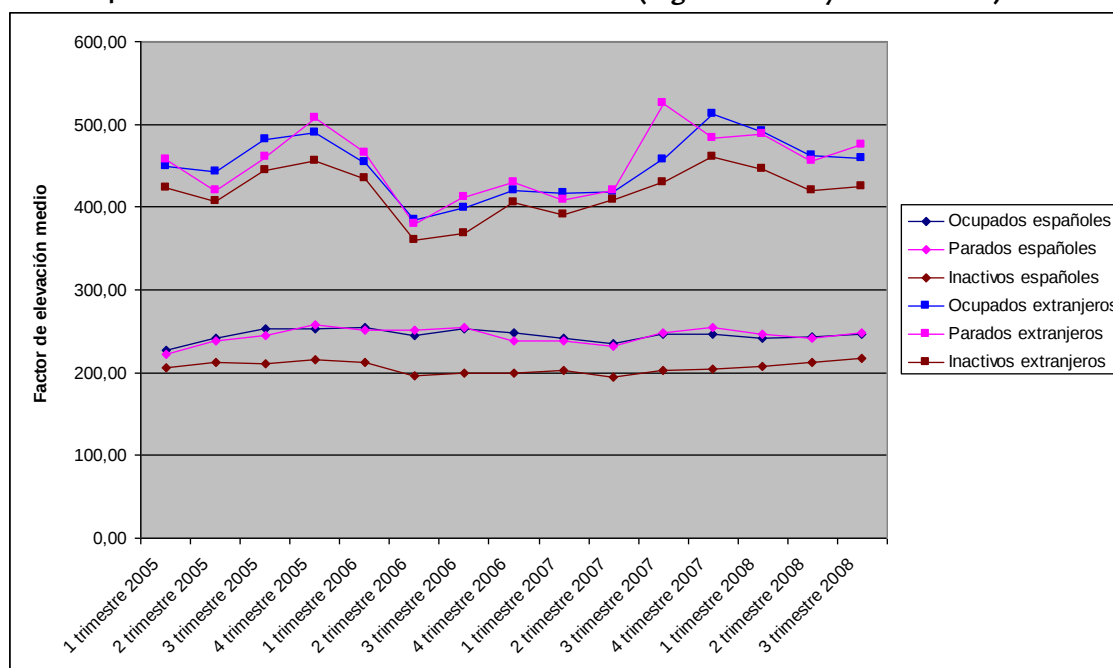
- (1) Los factores de elevación medios correspondientes a individuos de nacionalidad extranjera prácticamente duplican los valores de la nacionalidad española, para las tres situaciones de actividad.
- (2) Los factores de elevación correspondientes a la nacionalidad extranjera son más volátiles que los de la nacionalidad española.

Ambas características, patentes en la visualización de las series, se confirman en los estadísticos descriptivos que se presentan en el Cuadro 24.

Cuadro 23. Valores medios de los factores de elevación de la EPA.

	españoles			extranjeros		
	Ocupados	Parados	Inactivos	Ocupados	Parados	Inactivos
ene-05	227	223	205	449	458	422
feb-05	241	239	212	442	421	407
mar-05	253	244	211	481	461	444
abr-05	253	258	216	489	508	456
ene-06	255	252	213	455	465	435
feb-06	245	252	196	384	379	360
mar-06	253	255	200	399	412	368
abr-06	248	239	199	420	430	405
ene-07	242	239	202	417	409	391
feb-07	236	231	194	419	421	408
mar-07	247	247	203	457	526	429
abr-07	247	254	204	513	484	461
ene-08	241	246	207	492	488	447
feb-08	243	241	212	461	455	419
mar-08	247	248	218	459	475	425

Gráfica 24. Evolución de los factores de elevación medios (según actividad y nacionalidad).



Cuadro 24. Estadísticos descriptivos de la serie de factores de elevación (periodo 1-2005 a 3-2008).

		Mínimo	Máximo	Media	Desv. típ.	Coef. variación
Españoles	Ocupados	227	255	245	7	3,01%
	Parados	223	258	245	10	3,89%
	Inactivos	194	218	206	7	3,55%
Extranjeros	Ocupados	384	513	449	36	8,08%
	Parados	379	526	453	40	8,94%
	Inactivos	360	461	419	30	7,07%

Las consecuencias de esta estructura de ponderación se descubren cuando se analizan las transiciones entre diferentes trimestres de la EPA. Advierta que si el número de parados extranjeros aumenta en una sola unidad, al elevar su peso a nivel de población, supondrá un incremento estimado de alrededor de 452 personas en la población de parados. El impacto de las variaciones muestrales en sucesivas encuestas sobre los datos poblacionales es, sin embargo, incierto. En el caso de que los factores de elevación no sufran una modificación importante de uno a otro periodo, las transformaciones en los datos no ponderados se trasladan uniformemente a los datos ponderados. Pero, como se ha indicado, las series de los factores de elevación medios muestran una fuerte variabilidad en el tiempo. Como consecuencia, ocurre que las variaciones entre periodos de los datos no ponderados suponen variaciones muy diferentes en términos de los datos ponderados. En el Cuadro 25 se muestran las variaciones inter-trimestrales sobre los datos no ponderados y ponderados. En el Cuadro 26 se presentan unos “factores de variación” implícitos en esas variaciones. El resultado que se obtiene es poco justificable en términos de representatividad de la muestra. Concretamente, si se observan las variaciones del último trimestre puede comprobarse que un incremento en el número de parados extranjeros de 8 individuos en la muestra se corresponde con un

incremento del paro de 4613 personas. Los efectos de los distintos pesos pueden llegar a invertir la variación, como en el caso del número de ocupados españoles, que aunque en ese trimestre disminuye en 5 individuos en la muestra, esto se corresponde con un incremento de 5524 personas a nivel de población.

Debe tenerse en cuenta que el diseño muestral de la EPA no confiere sentido estadístico a los anteriores “factores de elevación implícitos” calculados las variaciones trimestrales. Se entiende que las variaciones en la muestra no se traducen directamente en la población, puesto que no son únicamente 5 nuevos individuos los que adquieren la condición de ocupado, sino que este número es el saldo final de entradas y salidas de individuos, todos ellos con distintas ponderaciones. El descenso del número de ocupados en la muestra de cinco individuos es el saldo neto de las variaciones, que se traduce a nivel de población en un saldo poblacional indeterminado. Sin embargo, es difícil justificar que variaciones netas de un número tan reducido de individuos tengan un impacto tan elevado sobre los valores estimados de la población.

Cuadro 25. Variaciones respecto al trimestre anterior.

	Datos no ponderados			Datos ponderados			
		Ocupados	Parados	Inactivos	Ocupados	Parados	Inactivos
Españoles	2 trimestre 2005	63	-68	-153	37.745	-13.594	-22.931
	3 trimestre 2005	-22	-25	-41	13.679	-5.579	-9.514
	4 trimestre 2005	-149	32	125	-37.512	9.209	31.979
	1 trimestre 2006	-6	14	-2	988	2.939	-4.493
	2 trimestre 2006	215	-17	-62	38.861	-4.337	-33.587
	3 trimestre 2006	-2	-28	-47	12.221	-6.792	-4.643
	4 trimestre 2006	-90	10	152	-30.389	1.214	29.819
	1 trimestre 2007	-26	35	28	-16.793	8.352	9.222
	2 trimestre 2007	169	-24	-80	30.675	-6.442	-25.910
	3 trimestre 2007	-56	-29	-52	5.204	-5.665	495
	4 trimestre 2007	-124	48	104	-30.063	12.605	21.994
	1 trimestre 2008	15	25	-18	-5.738	5.256	335
	2 trimestre 2008	76	-16	-121	21.604	-4.503	-18.775
	3 trimestre 2008	-5	7	-73	5.524	2.567	-9.374
Extranjeros		Datos no ponderados			Datos ponderados		
		Ocupados	Parados	Inactivos	Ocupados	Parados	Inactivos
	2 trimestre 2005	44	-12	-16	18.221	-6.204	-7.864
	3 trimestre 2005	-17	0	7	647	760	5.773
	4 trimestre 2005	-14	1	12	-5.297	1.411	6.348
	1 trimestre 2006	12	7	16	-1.196	2.380	5.047
	2 trimestre 2006	66	1	9	10.827	-1.934	-4.711
	3 trimestre 2006	15	-9	-10	10.061	-2.775	-2.745
	4 trimestre 2006	-31	10	11	-6.945	4.631	8.412
	1 trimestre 2007	2	19	3	57	7.176	-463
	2 trimestre 2007	46	-27	-8	19.677	-10.806	-1.243
	3 trimestre 2007	-11	-4	-9	6.702	98	-1.530
	4 trimestre 2007	-55	25	-1	-11.970	11.382	2.765
	1 trimestre 2008	1	8	19	-4.506	4.091	7.039
	2 trimestre 2008	57	-9	-6	19.023	-5.752	-5.766
	3 trimestre 2008	9	8	-4	3.329	4.613	-1.110

Cuadro 26. Factores de elevación implícitos en las variaciones inter-trimestrales.

	Españoles			Extranjeros		
	Ocupados	Parados	Inactivos	Ocupados	Parados	Inactivos
2 trimestre 2005	599	200	150	414	517	492
3 trimestre 2005	-622	223	232	-38	-	825
4 trimestre 2005	252	288	256	378	1.411	529
1 trimestre 2006	-165	210	2.247	-100	340	315
2 trimestre 2006	181	255	542	164	-1.934	-523
3 trimestre 2006	-6.111	243	99	671	308	275
4 trimestre 2006	338	121	196	224	463	765
1 trimestre 2007	646	239	329	29	378	-154
2 trimestre 2007	182	268	324	428	400	155
3 trimestre 2007	-93	195	-10	-609	-25	170
4 trimestre 2007	242	263	211	218	455	-2.765
1 trimestre 2008	-383	210	-19	-4.506	511	370
2 trimestre 2008	284	281	155	334	639	961
3 trimestre 2008	-1.105	367	128	370	577	278

6. CONCLUSIONES

El primer punto que debe destacarse es la necesidad de aumentar el tamaño muestral de la EPA en Baleares, si ésta debe servir como un instrumento útil para el análisis del mercado laboral en las Islas. Como se ha indicado, el propio INE advierte del peligro de intentar extraer conclusiones cuando se basan en segmentos inferiores a las 5000 personas. No es difícil que esto ocurra cuando se cruzan algunas de las variables de la EPA. Igualmente debe señalarse la necesidad de que la EPA ofrezca información con un nivel de desagregación más elevado que la Comunidad Autónoma, especialmente teniendo en cuenta las diferentes islas. Además, debe resaltarse la cada vez más descompensada relación entre el tamaño de la muestra asignado a las Islas Baleares (más o menos estable en el periodo analizado) y la población total (con una tasa de crecimiento relevante).

Respecto al análisis realizado sobre la capacidad de la EPA para equilibrar sus estimaciones poblacionales con respecto a fuentes de información externas, se ha obtenido una evidencia mixta.

En primer lugar, aunque es cierto que para algunas de las principales variables las estimaciones poblacionales son correctas, eso no ocurre en todas las variables. Por ejemplo, el equilibrio respecto a la información del Padrón Municipal se consigue para la población total y para la subpoblación de nacionalidad extranjera, pero no para la nacionalidad española. El mismo problema aparece en otras variables que también están equilibradas en el total de la población, pero que dejan de estarlo cuando se valoran en distintas subpoblaciones. Esto ocurre, por ejemplo, en la distribución por sexos cuando se consideran diversos tramos de edad o para los parados registrados, en donde al segmentarse por tramos de edad tampoco resulta equilibrada respecto a la obtenida en el INEM.

En segundo lugar, ese equilibrio se consigue en gran parte gracias a un proceso de calibración o reponderación que es ajeno al diseño muestral. Para algunas variables, la comparación entre los resultados no ponderados y ponderados de la muestra

permite detectar que importantes segmentos de la población están subrepresentados en la muestra. Esto se aplica, por ejemplo, al tramo de edad entre 25 y 39 años, un tramo de edad con problemas de representación en anteriores periodos de la EPA y que las modificaciones de 2005 no parecen haber resuelto.

De manera general, puede concluirse que la EPA sólo consigue equilibrar los principales agregados, pero no consigue el mismo resultado en las variables desagregadas. Debe destacarse además que la muestra de la EPA presenta problemas de subrepresentación para algunos segmentos importantes de la población.

Respecto al proceso de reponderación o calibración que realiza el INE, sus efectos sobre las estimaciones de Baleares son problemáticos. El proceso de calibración tiene por objetivo equilibrar la muestra con respecto a fuentes de información externa. Si bien la reponderación parece resolver un problema de equilibrio dirigido a las cifras de las variables agregadas, su uso es discutible desde otras perspectivas.

En primer lugar, lo deseable es que el proceso de calibración no tuviera que efectuarse, puesto que de alguna manera implica que el diseño muestral no logra reflejar correctamente la información poblacional. El INE ha justificado la calibración especialmente en relación a las dificultades para entrevistar a la población inmigrante (INE, 2005). El proceso de calibración es, por tanto, una forma de neutralizar la falta de información consecuencia de que determinados segmentos de la población no se han recogido correctamente en la muestra.

En segundo lugar, tanto si se trata de solucionar un problema del diseño muestral, como si es un problema de falta de respuesta, el proceso de calibrado sólo tiene sentido si los factores de elevación poblacionales calculados originalmente no se modifican en exceso. Si la reponderación implica que las ponderaciones originales se modifican extremadamente, se dañará el diseño probabilístico. La asimetría positiva detectada en la distribución de los factores de elevación ilustra que este problema puede darse en la muestra de la EPA.

En tercer lugar, el proceso de calibración tiene por objetivo unas pocas variables. Por ello, cuando se aplican los nuevos factores de ponderación se ignora cuál su efecto sobre otras variables de interés en la muestra. El empleo de las nuevas ponderaciones requiere el cumplimiento de la hipótesis de que las variables que intervienen en su elaboración presenten una correlación alta con el resto de variables sobre las que se aplica. Esta hipótesis se incumple cuando las características de los segmentos de población subestimados no son homogéneas respecto a las fuentes de información externa.

Finalmente, cuando se analiza la dinámica de la población activa, los factores de ponderación juegan un papel poco preciso. Por ello, para estudiar las transiciones en el mercado laboral sería más adecuado, en lugar de analizar las series de datos agregados, emplear los datos individuales de la EPA, valiéndose de la estructura de panel que puede construirse a partir de los microdatos.

7. REFERENCIAS

Álvarez, F., 2000, La Encuesta de Población Activa: cómo mide el mercado laboral, Estadística Española, vol. 42, nº 146, págs. 205 a 261.

INE, Método de reponderación empleado en la EPA, disponible en red en: http://www.ine.es/epa02/documento_tecnico.pdf.

INE, 2004, Revisión de la Encuesta de Población Activa, disponible en red: <http://www.ine.es/prensa/np322.pdf>.

INE, 2005, Encuesta de Población Activa. Diseño de la Encuesta y Evolución de la calidad de los datos, Informe Técnico.

Martí, M. y C. Ródenas, 2004, Migrantes y migraciones: de nuevo la divergencia en las fuentes estadísticas, vol. 46, núm. 156, págs. 293-321.

Ministerio de Trabajo e Inmigración, 2007, Anuario del Ministerio de Trabajo e Inmigración 2007, <http://www.mtin.es/estadisticas/anuario2007/index.htm>.

Observatori del Treball de les Illes Balears, Nota de conjuntura laboral a les Illes Balears, Octubre 2008.

Pérez, J.I., 2008, ¿Por qué el paro registrado supera al paro de la EPA?, Revista del Ministerio de Trabajo y Asuntos Sociales, nº 71, marzo, págs. 53-65.

Prudencio, C.A. y L. Toharia, 2007, Análisis regional del comportamiento de las personas inscritas en las oficinas de empleo públicas según la Encuesta de Población Activa. 2004 y 2005, Estadística Española, vol. 49, nº. 166, págs. 407-450.