



Document de suport del PaD FUNDACIÓ JAUME BOFILL

# Panel de Desigualtats Socials a Catalunya

## Manual d'ús de les matrius

Onades de la A (2001/02) a la K (2012)

JUNY 2014



## Índex

Presentació .....	3
1 L'enquesta PaD .....	4
1.1. Dades sobre estructura i desigualtats socials .....	4
1.2. Aspectes objectius i subjectius de la vida dels individus .....	4
1.3. Una enquesta longitudinal .....	4
1.4. Dues unitats d'estudi: la llar i els seus individus .....	6
1.5. El disseny de la mostra (any 2001) .....	7
1.5.1. Criteris de seguiment de la mostra	
1.5.2. Evolució de la mostra	
1.5.3. Remostrejos	
1.5.4. Ponderació	
1.5.4.1. Llars	
1.5.4.2. Individus	
1.6. Els qüestionaris .....	14
2 Les matrius de dades .....	15
2.1. Tipus de matrius .....	15
2.2. Estructura de les matrius .....	16
2.3. Tipus de variables i nomenclatura .....	16
2.4. Valors perduts .....	17
2.5. Els identificadors d'individu i de llar .....	18
3 Els materials de suport .....	19
3.1. L'informe longitudinal de continguts .....	19
3.2. Els llibres de codis .....	20
3.3. Documents metodològics .....	20
Annexos	
Annex 1. Taules de referència dels gràfics sobre mostra .....	21
Annex 2. Construcció dels identificadors numèrics .....	22
Annex 3: Qüestionari .....	24
Annex 4: Mostra .....	27
Annex 5: Treball de camp .....	28
Annex 6: Dades .....	29
Annex 7: Fitxa tècnica de l'anàlisi .....	30
Annex 8: Fitxa tècnica de difusió .....	31

## Presentació

El *Manual d'ús* vol ser una eina de suport per a tots els investigadors i investigadores que treballin amb la base de dades del Panel de Desigualtats Socials a Catalunya (d'ara endavant PaD). Les particularitats metodològiques, tècniques i de contingut d'una enquesta com el PaD fan del tot recomanable la lectura detallada d'aquest document abans d'endinsar-se en l'explotació de les dades. S'ha procurat que fos el més breu possible, seleccionant amb cura la informació que semblava indispensable per poder entendre bé el tipus de dades i l'estructura dels fitxers del PaD.

Per qualsevol consulta o informació complementària sobre el PaD, podeu adreçar-vos a:

Laia Pineda, coordinació general del PaD [lpineda@fbofill.cat](mailto:lpineda@fbofill.cat)

Laura Morató, mostra, dades i anàlisi del PaD [lmorato@fbofill.cat](mailto:lmorato@fbofill.cat)

# 1. L'enquesta PaD

## 1.1. Dades sobre estructura i desigualtats socials

El PaD té com a objectiu principal aportar informació nova, en clau longitudinal, sobre les desigualtats socials al nostre país. L'enquesta està pensada per obtenir informació sobre les oportunitats i limitacions de les persones que viuen a Catalunya en referència a diferents aspectes. Els àmbits principals sobre els quals se centra el PaD, i dels quals se'n recull informació a cada onada, són els següents:

- Situació personal i familiar
- Educació reglada i no reglada
- Treball productiu, reproductiu i conciliació
- Usos del temps
- Habitatge
- Posició econòmica
- Salut i discapacitats
- Posició i relacions socials
- Valors i actituds

Trobareu més informació sobre els continguts del PaD, per blocs temàtics, al document *Síntesi de continguts*

## 1.2. Aspectes objectius i subjectius de la vida dels individus

Una de les particularitats de l'enquesta PaD té a veure amb la recollida d'aspectes subjectius de la vida dels enquestats. Les estadístiques oficials disponibles (tant a Catalunya com a Espanya i a Europa) marquen la tendència a una expansió de les estadístiques sobre fets (dades objectivables) i fan poca atenció a la valoració que fan els individus d'aquests fets, és a dir a la vivència personal que els fets generen. Per això, els qüestionaris del PaD intenten trobar un equilibri entre les preguntes que busquen recollir dades de caràcter objectiu i aquelles que se centren en els aspectes més subjectius de la vida dels enquestats.

## 1.3. Una enquesta longitudinal

El PaD és la primera enquesta longitudinal que s'ha impulsat des de Catalunya amb una mostra estadísticament representativa de les llars catalanes<sup>1</sup>. Optar per una enquesta longitudinal, atesa la minsa experiència que el nostre país ha acumulat en aquest tipus d'enquestes, recomana emmirallar-se en les enquestes europees. Panels com el *British Household Panel Survey* (BHPS) ara

---

<sup>1</sup> A Catalunya hi ha hagut llars que han format part de mostres d'enquestes longitudinals impulsades per l'Institut Nacional de Estadística (INE), com per exemple la Encuesta Continua de Presupuestos Familiares (ECPF), que subministra informació sobre l'origen i el volum de les rendes de les llars a Espanya i la manera com s'utilitzen en les diverses despeses de consum, però mai no han estat mostres estadísticament prou representatives com per fer estudis específics de Catalunya.

anomenat *Understanding Society*,<sup>2</sup> el *German Socio-Economic Panel* (GSOEP)<sup>3</sup> o el *Swiss Household Panel* (SHP)<sup>4</sup> són enquestes longitudinals de solvència contrastada. En el moment inicial del PaD, l'experiència britànica va ser la principal referència, sense que això signifiqués fer una mera exportació de l'eina. En els anys de desenvolupament del PaD, però, ha estat necessari també explorar les experiències d'altres països i d'altres panels, tot ajustant el contingut de l'enquesta a les darreres innovacions que aquests han implementat.

Des d'un punt de vista estrictament tècnic, una enquesta tipus panel pur té com a característica bàsica que es fa sempre sobre la mateixa mostra d'individus al llarg de les diferents onades en què es realitza, en contrast amb les enquestes tipus panel rotatiu que renoven part de la mostra en cada edició. La periodicitat d'un panel no està predeterminada, sinó que s'ajusta a la voluntat i/o possibilitats econòmiques de cada projecte. En aquest sentit, el PaD és una enquesta tipus panel pur i la seva periodicitat és anual.

No obstant, cal dir que tot i que la mostra inicial es manté en les onades successives i l'interès específic rau precisament en la seva continuïtat, el fet d'estar configurada per llars i individus confereix també una certa "vida" a la mostra. És per això que, al llarg de les onades, la mostra experimenta canvis com ara la incorporació de nous individus (naixements, formació de noves parelles, acollida d'avis a la llar, reagrupaments familiars...), la creació de noves llars (amb la sortida de la llar paterna per part dels joves que creen una nova llar, per divorcis o separacions...). En aquest sentit, la mostra d'una enquesta panel combina la dimensió dinàmica (de canvi) amb la dimensió d'estabilitat i continuïtat.

D'altra banda, en una mostra panel també hi ha pèrdues mostrals tant de llars com d'individus (l'*attrition*) per motius diversos: mort dels membres, institucionalització (ingrés en un centre tipus residència d'avis, internats d'estudis, presó...), trasllat de tota la llar fora del territori català, rebuig a seguir participant en l'enquesta... La pèrdua mostral obliga a fer noves captacions de llars durant la vida de l'enquesta, per tal de que la mostra no perdi representativitat. Per aquest motiu, el PaD ha fet captació de noves llars en diferents moments de l'enquesta, per suplir així l'erosió de la mostra. (veure l'apartat 1.5.3. Remostrejos)

Els avantatges de fer un seguiment d'una mostra panel es poden resumir en quatre aspectes fonamentals:

- **Qualitat:** Una mostra panel assegura poder recollir molt millor les dades: el retorn a les mateixes llars en cada onada permet depurar errors i validar les dades any rere any.
- **Quantitat:** La mostra panel permet desenvolupar, amb diferents nivells de profunditat, diferents blocs de preguntes en cadascuna de les onades (un any s'entra a fons en el tema de l'educació, l'altre en el treball domèstic, un altre en temes d'habitatge...), de manera que es pot acumular molta més informació de cada àmbit, a mesura que es van succeint les onades.
- **Perspectiva del canvi:** La mostra panel permet recollir dades molt valuoses sobre tot tipus de canvis que experimenten, tant les llars com els individus (canvis d'habitatge, canvis en

---

<sup>2</sup> <https://www.iser.essex.ac.uk/bhps>

<sup>3</sup> <http://www.diw.de/english/soep/>

<sup>4</sup> <http://forscenter.ch/en/our-surveys/swiss-household-panel/>

l'estructura de les llars, canvis en el treball, etc). El fet que es preguntí als mateixos individus any rere any permet conèixer, no només quins canvis es produeixen a les seves vides, sinó que es pot localitzar en quins grups socials es produeixen especialment, amb quina intensitat, les conseqüències sobre altres aspectes de la vida personal i familiar, etc., sense els biaixos que s'introduïrien pel fet d'estar preguntant a cada nova edició de l'enquesta una mostra nova d'individus. És l'única eina que permet donar compte de la trajectòria dels individus i de les llars, és a dir dels processos que viuen en cada moment de les seves vides. D'aquesta manera permet disposar d'informació privilegiada sobre els esdeveniments més rellevants del cicle vital.

- **Significació:** La mostra panel implica que l'anàlisi dels processos és més significativa estadísticament que la d'una anàlisi transversal. D'aquesta manera, es poden fer afirmacions més robustes sobre els canvis que es produeixen en la realitat social encara que aquests siguin, en termes percentuals, més petits ja que es disminueixen els marges d'error que es deriven de qualsevol mostra.

#### 1.4. Dues unitats d'estudi: la llar i els seus individus

El PaD té la particularitat de tenir una doble unitat d'estudi: les llars i els seus individus. Aquesta característica té a veure amb la voluntat d'ampliar al màxim el potencial explicatiu de les dades recollides. L'anàlisi de les dades del PaD pot ser tan de les llars en si mateixes (mobilitat geogràfica, formació i dissolució de llars, estratègies en l'educació dels menors de 16 anys...) com dels individus que les conformen. A això s'hi ha d'afegir l'interès analític de disposar, per als individus, de dades de les seves llars, és a dir del seu context immediat.

Metodològicament, les implicacions de tenir dues unitats d'estudi són:

- A nivell de mostra: cal fer un mostreig en el que tant el nombre de llars com el d'individus respongui als marges d'error màxims estipulats.
- A nivell de qüestionari: ha calgut dissenyar dos qüestionaris diferents, un per a les llars i un altre per als individus, per poder recollir la informació en funció de la unitat d'anàlisi.
- A nivell de treball de camp:
  - Enquestació: ha calgut desenvolupar un programa que permet codificar els canvis que es produeixen en la composició de les llars i que això es reflecteixi immediatament en l'enquestació.
  - Captació de llars: per formar part de la mostra, les llars han d'estar d'acord a participar-hi en el seu conjunt. En anys de captació de llars (1a, 6a, 7a, 8a i 10a onades) no s'accepten llars amb participació parcial, a més d'haver d'estar d'acord en participar en les successives onades de l'enquesta.
- A nivell de fitxers de dades: per poder seguir els individus i les llars al llarg de les onades ha calgut construir identificadors de llar i d'individu, invariables en el temps, que garanteixen poder unificar fitxers de diferents onades així com afegir informació de les llars als indivi-

dus o a l'inrevés.

### 1.5. El disseny de la mostra (any 2001)

La mostra inicial del PaD consta de 2.000 llars de Catalunya seleccionades aleatòriament seguint un mètode estratificat de mostreig en etapes múltiples, amb una selecció aleatòria sistemàtica de les primeres unitats mostrals (seccions censals) i una selecció aleatòria simple de les unitats finals (adreces).

Es decideix utilitzar les seccions censals com a primeres unitats mostrals, ja que són les unitats territorials estadístiques d'ús públic més petites. Com que són agrupacions territorials amb mitjanes entre 1.000 i 1.200 persones, o aproximadament 400 llars, tenen una gran coherència interna i la informació disponible permet fer-ne classificacions d'estructura social i territorial diferents. El nombre idoni de llars per enquestar a cada secció censal es fixa en 10. Es compta que, de mitjana, les llars de cada secció censal derivaran en 270 individus per a la mostra final. En relació a la mostra territorial fixada, cal seleccionar 200 seccions censals de les 5.186 existents l'any 1996 (actualització del cens de 1991).

Per tal de garantir una bona representativitat de la mostra, es combinen 4 variables d'estratificació, dues de tipus territorial i dues de tipus social:

1. província (incloent la subdivisió en 3 subregions en el cas de Barcelona);
2. àmbit urbà-rural (en funció de la grandària municipi);
3. variable construïda de posició social (classificació segons  $\pm$  classe treballadora);
4. situació ocupacional (endreq segons nombre d'actius/inactius)<sup>5</sup>.

El resultat de la selecció de seccions censals es concreta en 203 unitats ubicades en 115 municipis de Catalunya.

La mostra d'individus es deriva de comptabilitzar 2,7 individus de mitjana a cada llar seleccionada pel mostreig, resultant en 5.445 individus pel conjunt de Catalunya.

Per poder garantir una bona representativitat territorial en la mostra del PaD i compensar la gran concentració poblacional a la regió metropolitana de Barcelona (on habiten més de tres quartes parts de les llars catalanes) s'opta per una mostra aporportional pel que fa a les províncies.

És per això que, partint del  $\pm 5\%$  com el marge màxim d'error mostral assumible per a resultats globals (amb un interval de confiança del 95,5%, amb  $k=2$ , i en el supòsit de màxima indeterminació  $p=q=50$ ), cal observar un mínim 400 llars a cadascuna de les 4 províncies catalanes. Pel que fa a la província de Barcelona, la seva diversitat i complexitat exigeix una mostra específica de mida més gran que es fixa finalment en 800 llars, repartides proporcionalment, segons la distribució de la població, entre Barcelona ciutat (261), la seva Regió Metropolitana (470) i Comarques Centrals (69). És així com s'arriba a la mida final de la mostra, que es compondrà de 2.000 llars distribuïdes aporportionalment a escala provincial. Es pot veure el detall de la seva distribució i els errors mostrals associats a la següent taula.

---

<sup>5</sup> La variable de posició social es construeix a partir de les preguntes 12 i 13 del Padró de 1996.



**Taula 1. DISTRIBUCIÓ LLARS I INDIVIDUS DE LA MOSTRA ORIGINAL<sup>6</sup> I ERROR MOSTRAL (1a onada del PaD: 2001- 2)**

ÀREES	NOMBRE LLARS	e* (%)	NOMBRE INDIVIDUS	e* (%)
<b>Barcelona</b>	791	±3,54	2.213	±2,13
Ciutat BCN	252	±6,12		
Regió Metropolitana	470	±4,61		
Comarques centrals	69	±12,04		
<b>Girona</b>	400	±5,00	1.193	±2,89
<b>Lleida</b>	400	±5,00	1.175	±2,91
<b>Tarragona</b>	400	±5,00	1.176	±2,91
<b>CATALUNYA</b>	<b>1991</b>	<b>±2,78</b>	<b>5.757</b>	<b>±1,32</b>

\*e = error mostral (Interval confiança=95,5%, k=2 i p=q=50)  
 Font: elaboració pròpia

Durant el treball de camp es realitzen tres comprovacions de la bondat de la mostra, tan de les llars com dels individus captats (amb l'11%, el 50% i el 100% del treball de camp realitzat) assegurant que no hi ha biaixos importants que demanin rectificacions de la mostra dissenyada.

### 1.5.1. Criteris de seguiment de la mostra

Qualsevol enquesta ha de marcar uns criteris sobre quina és la mostra que necessita per respondre als seus objectius. En el cas del PaD, una enquesta longitudinal pur, això és cabdal. No només per a la captació sinó per saber en cada nova onada a quines llars o individus s'ha de continuar enquestant i a quins no.

Poder formar part de la mostra vol dir se considerat "elegible", i al contrari, no poder formar part de la mostra "inelegible". En cada onada, doncs, s'ha mirat d'enquestar totes les llars, i els individus, elegibles, ja siguin provinents de; la mostra entrevistada durant l'onada anterior; enquestades més enrere en el temps però que es valora que podrien participar en l'onada actual; les llars de nova creació entre l'onada anterior i l'actual.

En el cas de les llars, els criteris d'elegibilitat són que la llar estigui domiciliada a Catalunya i que sigui una llar privada, és a dir, no un centre o institució. En el cas dels individus el criteri és que

<sup>6</sup> Es poden consultar els errors mostrals de la mostra teòrica a l'annex 4. Mostra.



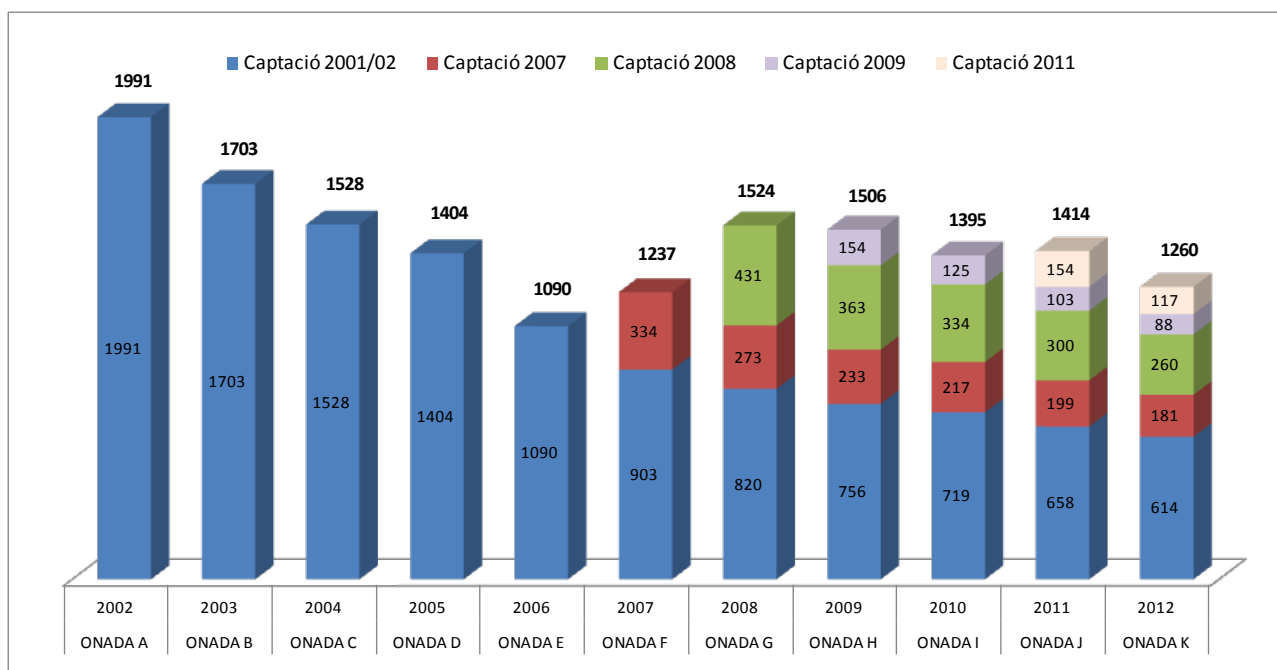
siguin majors de 15 anys i que visquin en una llar elegible.

### 1.5.2. Evolució de la mostra

Com ja s'ha mencionat més amunt, les enquestes longitudinals, degut a la seva naturalesa periòdica, van perdent casos de la mostra any rere any. D'una banda, les llars i els individus deixen de ser elegibles (per institucionalització, mort o trasllat fora de Catalunya). De l'altra, donat que la resposta a l'enquesta del PaD és voluntària, poden rebutjar participar, o ser il·localitzables, entre altres causes.

Des d'un punt de vista de l'anàlisi longitudinal, interessa conèixer quina és l'evolució de la mostra, per saber a quantes llars i/o individus es pot seguir al llarg de les diferents onades del PaD. Amb la voluntat de sintetitzar aquesta evolució, els dos gràfics següents mostren el comportament de les llars (gràfic 1) i els individus (gràfic 2) captats, en les onades en que hi ha hagut captació<sup>7</sup>. S'ha de tenir en compte el total que apareix als gràfics no és el total de llars, ni d'individus, que s'inclouen a les matrius de dades ja que aquí no hi ha les noves llars (creades per independització de la llar paterna, separació o divorci...) ni els nous individus que s'han enquestat (perquè han anat a viure a una llar de la mostra del PaD).

**Gràfic 1. Evolució de la mostra captada de llars**

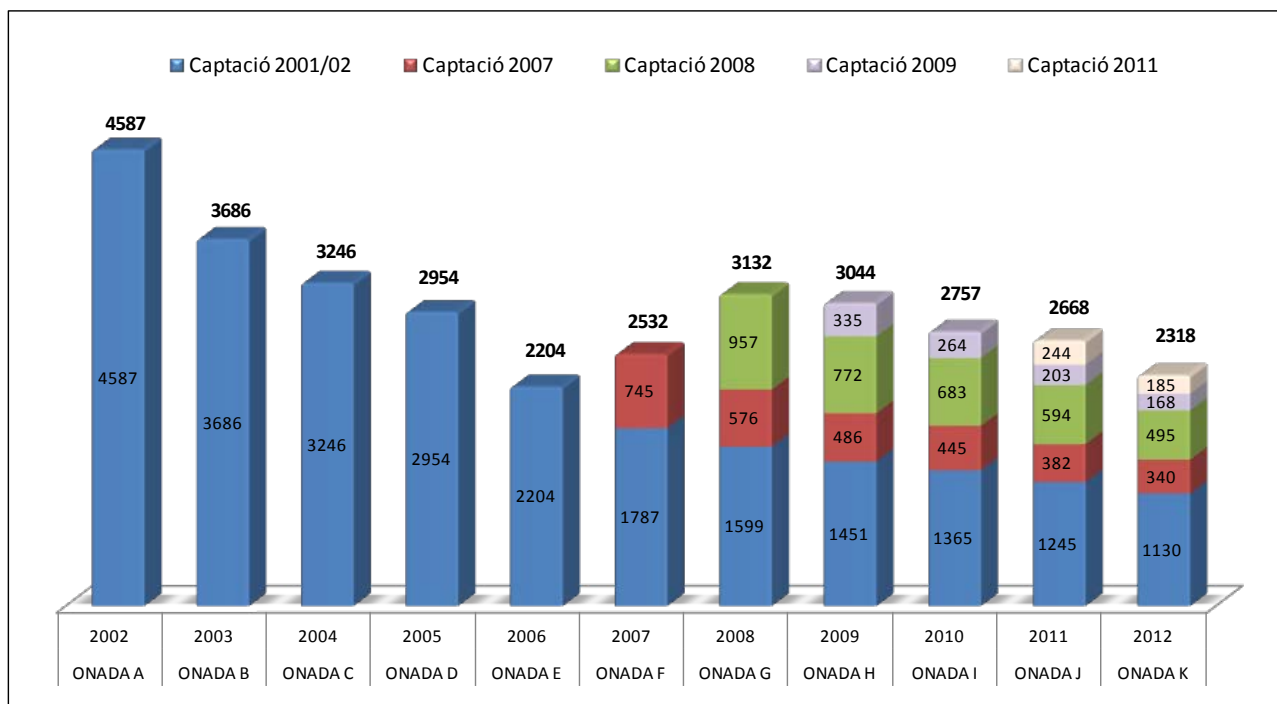


Font: elaboració pròpia amb dades del PaD (1a a 11a onades)

Aquest gràfic inclou únicament les incorporacions de llars a la mostra per captació, no les produïdes per les dinàmiques pròpies de les llars.

<sup>7</sup> Es poden consultar les taules, amb els percentatges de pèrdua corresponents, a l'annex 1.

**Gràfic 2. Evolució de la mostra captada d'individus (entrevistats i proxys\*)**



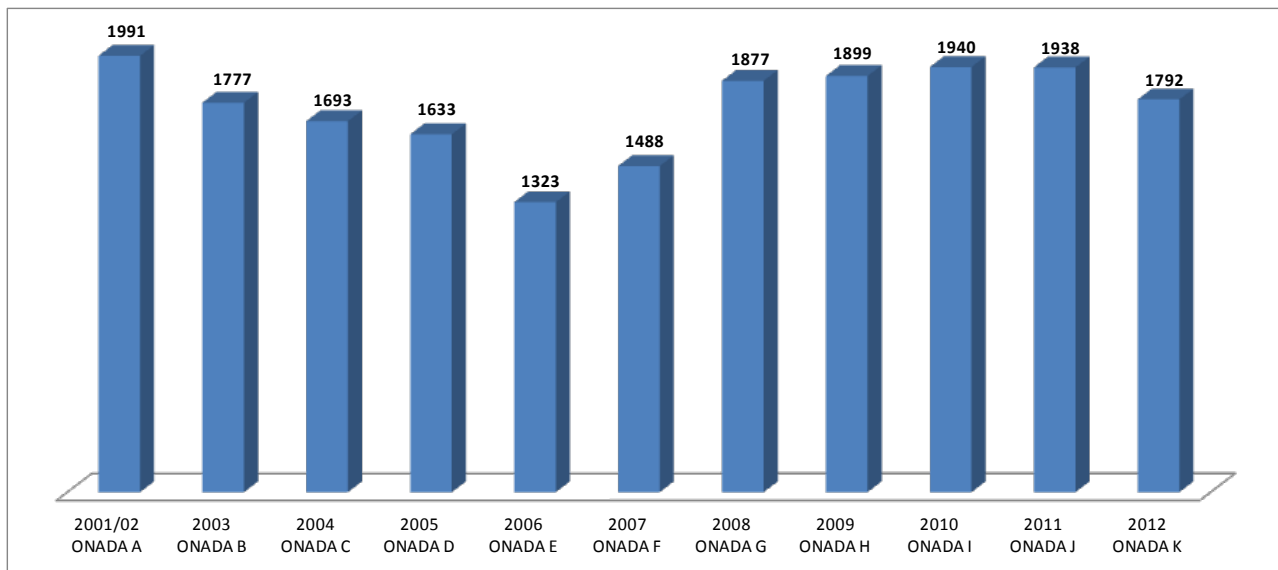
Font: elaboració pròpia amb dades del PaD (1a a 11a onades)

Aquest gràfic inclou únicament les incorporacions de membres a la mostra per captació, no les produïdes per les dinàmiques pròpies de les llars.

\* No s'inclouen els menors de 16 anys, que per criteris de mostra no poden ser entrevistats però dels que també hi ha informació a cada onada, facilitada per l'informant principal de la llar.

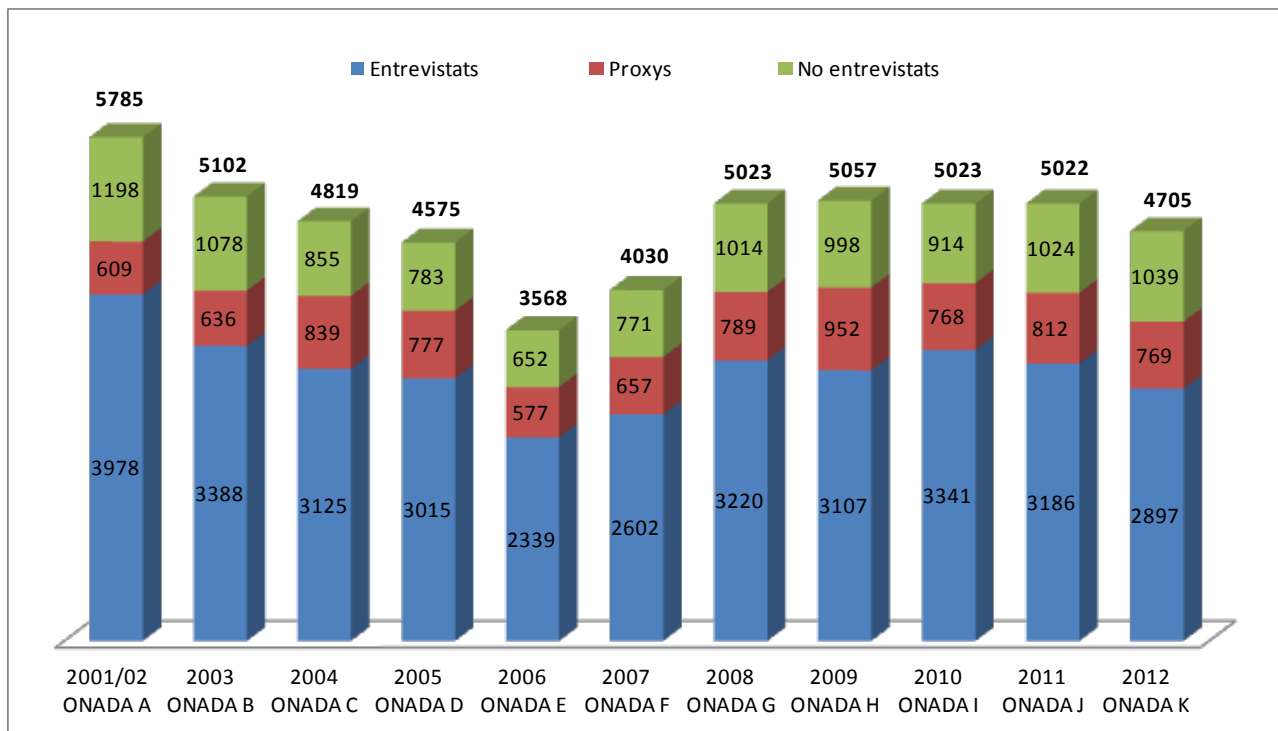
Des d'un punt de vista transversal, interessa el total d'observacions que hi ha a cada onada. Els dos gràfics següents volen presentar aquesta informació de manera clara. En el cas dels individus (gràfic 4), s'ha volgut especificar el tipus de qüestionari (individual, proxy o no entrevistat) ja que cada un d'ells proveeix d'una quantitat d'informació diferent. En l'apartat corresponent a qüestionari, trobareu informació més detallada sobre els tipus de qüestionari.

**Gràfic 3. Llars entrevistades al PaD per onada**



Font: elaboració pròpia amb dades del PaD (1a a 11a onades)

**Gràfic 4. Individus presents al PaD per tipus de qüestionari i onada**



### 1.5.3. Remostrejos

Com es pot veure en els gràfics anteriors, l'erosió mostral va deixar les 1.991 llars captades a primera onada en 1.323 a cinquena. Per això, a la 6<sup>a</sup> onada (any 2007), i gràcies a la disponibilitat de recursos econòmics, s'opta per augmentar els efectius de la mostra.

L'augment de la mostra del PaD es realitza a partir d'un remostreig, seguint el mateix mètode emprat per a la mostra inicial de primera onada (mateixes variables d'estratificació, mateix mètode de selecció de les seccions censals finals) però amb una diferència important: les seccions censals de la mostra inicial amb llars supervivents són el punt de partida del nou mostreig, que caldrà completar amb les seccions censals que garanteixin la representativitat territorial i social desitjades. Aquesta decisió respon a la necessitat de contenir la dispersió d'enquestadors en el territori català i facilitar al màxim la gestió del treball de camp. Quan pels resultats del remostreig no és possible utilitzar les mateixes seccions censals, es seleccionen de forma aleatòria noves seccions censals.

Les dificultats de captació que s'experimenten en el treball de camp obliguen a dilatar la captació prevista de 1.074 llars durant la 6<sup>a</sup> onada en les onades 7<sup>a</sup> (2008), 8<sup>a</sup> (2009) i 9<sup>a</sup> (2010)<sup>8</sup>. La dilatació de la captació al llarg de quatre onades implica que entre la 7<sup>a</sup> i la 8<sup>a</sup> onada s'ha de recalculer el remostreig per ajustar al màxim les noves llars a captar a les característiques de la societat catalana.

### 1.5.4. Ponderació

Tot i que la mostra estigui ben dissenyada i les proves de bondat donin resultats acceptables, és prou conegut que les mostres sempre es desvien de l'univers a estudiar. Per això, és necessari tenir variables de ponderació o pesos.

La ponderació és una transformació que permet donar-li a cada unitat mostral de l'enquesta un valor numèric (pes) representant la contribució que farà la unitat a les estimacions basades en les dades de l'enquesta.

Cada una de les onades del PaD disposa de variables de ponderació transversal, a nivell de llar i d'individu. Aquestes variables les genera el Servei d'Estadística Aplicada de la Universitat Autònoma de Barcelona ([SEA-UAB](#)).

A continuació presentem les característiques bàsiques de cada una d'aquestes variables. Per a més informació, us remetem a l'informe metodològic elaborat pel SEA-UAB, disponible al [web del PaD](#).

#### 1.5.4.1. Llars

En el cas de les llars, s'han construït dues variables de ponderació. La primera les reequilibra en el moment en el que es va dissenyar la mostra, l'any 2001, amb les dades del Cens de població i habitatges 2001 (Instituto Nacional de Estadística, INE). Era la intenció del PaD generar-la amb el

---

<sup>8</sup> Les llars captades el 2010 no es van incorporar efectivament a la mostra fins l'any següent, el 2011.

Cens de 2011 per a les dues últimes onades però la disponibilitat de les dades de l'INE no ho han fet possible. Per això, la variable reequilibra totes i cadascuna de les onades amb referència a les dades de 2001.

Aquesta variable es presenta en dos formats, el primer d'ells només reequilibra la mostra (codi de la variable XW009001L<sup>9</sup>), el segon reequilibra i eleva el nombre de llars a les llars catalanes (XW009002L).

Les variables que s'utilitzen per a la ponderació de les llars del PaD en base al Cens 2001 són:

- Secció censal
- Província
- Nombre de persones a la llar
- Lloc de naixement
- Edat
- Situació respecte el treball remunerat
- Règim de tinença de la llar

Tanmateix, interessava també disposar d'una variable de ponderació que tingués en compte la realitat social de cada un dels anys en que el PaD va fer treball de camp. A la impossibilitat de trobar dades exhaustives de l'univers d'estudi s'hi afegeix la dificultat de disposar de dades sobre llars, no sobre individus. Finalment, es va optar per utilitzar una enquesta d'àmbit estatal, amb representativitat per Catalunya, amb mostra de llars, periodicitat anual i harmonitzada a nivell europeu: la Encuesta de Población Activa. Per tant, en les matrius s'hi inclou una segona variable de ponderació (XW011001L) que reequilibra la mostra a la de l'EPA del 4rt trimestre de cada any.

Les variables que s'utilitzen per a la ponderació de les llars del PaD en base a l'EPA són:

- Província
- Nombre de persones a la llar
- Lloc de naixement
- Edat
- Situació respecte el treball remunerat
- Nivell màxim d'estudis

#### 1.5.5.2. Individus

En el cas dels individus, la base de dades de referència ha estat el Padró continu de cada any. De la mateixa manera que amb les dades del Cens per a les llars, es lliuren en el format que reequilibra la mostra (XW010001E) i en el format que a més d'equilibrar, l'eleva (XW010002E).

Les variables que s'utilitzen per a la ponderació dels individus del PaD en base al Padró continu són:

- Província
- Tamany del municipi
- Sexe
- Edat
- Lloc de naixement

---

<sup>9</sup> La X fa referència a la lletra de l'onada, per 1a onada és A, 2a és B i així successivament.

## 1.6. Els qüestionaris

El fet que el PaD tingui una doble unitat d'anàlisi (llars i individus) ha fet necessari desenvolupar tres **tipus de qüestionaris**:

- El qüestionari de Llar. El respon una única persona (anomenada "Informant Principal"). Aquest qüestionari no només conté preguntes generals de la llar sinó també preguntes relatives a les persones menors de 16 anys de la llar que no entrevistarem individualment. Aquesta informació, a nivell de matrius, s'incorpora a la matriu individual.

D'altra banda, hi ha la fitxa de llar, que no és pròpiament un qüestionari i que permet la gestió de la informació sobre altes i baixes de membres, dades de contacte, entre altres.

- El qüestionari Individual. És el qüestionari més extens i el contesten totes les persones de 16 o més anys de la llar.
- El qüestionari d'aproximació o Proxy. Permet disposar d'informació bàsica d'aquelles persones de 16 o més anys de la llar a qui no es podrà entrevistar durant el treball de camp de l'onada, ja sigui perquè rebutgen participar com perquè no poden fer-ho pel motiu que sigui. Aquest qüestionari conté les preguntes objectives que poden ser contestades per alguna altra persona de la llar, generalment l'informant principal.

Des del punt de vista de **les preguntes** que es fan a l'enquesta, podem distingir diferents tipus, segons la perspectiva que prenguem.

En el que es refereix al període temporal, el PaD inclou preguntes que recullen informació transversal, sobre el moment de l'entrevista, i informació longitudinal. La informació longitudinal es recull en diferents models temporals:

- Preguntes retrospectives, recullen esdeveniments succeïts en algun moment passat de la vida de l'individu.
- Preguntes de canvi des de la darrera entrevista, recullen els canvis que s'han pogut donar des de la darrera vegada que es va entrevistar la llar o la persona. En aquest tipus de preguntes es referencia la data de l'última entrevista i la informació aportada per la persona en aquell moment (per exemple: a la darrera entrevista que li vam fer, el XX de XXX de 20XX, ens va dir que treballava a jornada completa, des d'aleshores fins ara: ha treballat de manera continuada, ha combinat atur amb treball...). És el que s'anomena "entrevista dependent" i provoca un alt nombre de valors perduts ja que, si no hi ha hagut canvis (en l'exemple anterior es completava la informació pels que havien treballat de manera continuada, demanant si hi havia hagut altres canvis, com d'empresa, tipus de contracte, responsabilitat...) ja no es demana la informació de la qual disposem d'onades anteriors (en el nostre exemple, l'activitat de l'empresa, el tipus de contracte...).
- Preguntes sobre períodes concrets en el temps, és a dir que recullen informació sobre un període determinat. Són, bàsicament, les preguntes referents a ingressos i a despeses. En unes d'elles interessa recollir períodes tancats, com l'any anterior (és el cas dels ingressos

mensuals nets de la llar o de l'individu); en d'altres, interessa recollir altres períodes temporals diferents (per exemple, les despeses de l'educació reglada, es fan sobre el curs escolar anterior).

Sobre la periodicitat<sup>10</sup> de les preguntes, en el PaD cada pregunta té una periodicitat pròpia. Concretament:

- Preguntes troncats estructurals. Només es pregunten un cop perquè és informació invariable en el temps (per exemple, la data de naixement pròpia o la del pare o la mare). S'inclouen en el moment de la primera entrevista per a tots els entrevistats i proxys.
- Preguntes troncats. Apareixen en cada edició de l'enquesta per la seva natura estructural però variable en el temps (per exemple, la situació respecte el treball remunerat).
- Preguntes cícliques. Donat el seu contingut i/o la seva possible variabilitat en el temps, es valora que no és necessari incloure-les en cada edició de l'enquesta i es decideix fer-ho cada x anys (per exemple, qui fa les tasques domèstiques de la llar).

Tots els qüestionaris del PaD poden consultar-se seguint aquest [vincle](#).

## 2. Les matrius de dades

El present apartat descriu les diferents matrius de dades del PaD, en presenta els elements bàsics i necessaris per utilitzar-les i permet familiaritzar-s'hi d'una manera àgil. Considerem que aquesta és la informació bàsica que s'ha de tenir en compte si es volen explotar les dades de l'enquesta PaD.

### 2.1. Tipus de matrius

Les matrius s'ofereixen en dos formats, per facilitar tot tipus d'anàlisi: transversal, de cada onada, i longitudinal, inclou la informació de totes les onades possibles.

**Matrius d'onada.** N'hi ha tantes com onades i unitats d'anàlisi. Fruit de les dues maneres de recollir la informació a les llars –de la llar en el seu conjunt i de cada un dels individus que la integren– es generen dues matrius, o fitxers de dades, per a cada onada del PaD:

- Matriu de llar: recull les variables procedents de l'enquesta de llar. Prenen el nom A\_LLAR per primera onada (treball de camp entre 2001 i 2002), B\_LLAR per segona (2003) i així successivament fins a la K\_LLAR (2012).
- Matriu individual: recull les variables de l'enquesta individual i la informació recollida al qüestionari de llar sobre els menors de 16 anys. S'anomenen A\_INDI, B\_INDI ... K\_INDI.

Inclouen la informació estructural de la fitxa de la llar, la informació pròpia del qüestionari de l'onada i, en el cas de les variables harmonitzades, la informació harmonitzada disponible.

**Matrius longitudinals.** S'ha optat per generar-ne dues per a cada unitat d'anàlisi. D'una banda, la sèrie completa de totes les onades del PaD (prenen el nom A\_K INDI LONG o A\_K LLAR LONG) i, de l'altra, la sèrie completa tret de la primera onada (B\_K INDI LONG o B\_K LLAR LONG). Aquestes segones matrius han permès incloure les variables que recullen el canvi d'una onada a l'altra i que van començar a recollir-se a segona onada (per exemple, situació respecte el treball remunerat des de la darrera entrevista).

---

<sup>10</sup> Hi ha més detall a l'annex 3: Qüestionari



Les diferències entre unes i altres matrius, és que les A\_K contenen totes les onades però no les variables de canvi (inexistents a primera onada) mentre que les B\_K contenen totes les variables (incloses les de canvi d'una onada a l'altre) però no la primera onada.

Les matrius longitudinals s'ofereixen en format llarg (*long*), és a dir, els individus o les llars es repeteixen en tantes files com el nombre d'onades en què han participat. Hi ha una variable (P107001E en la matriu individual o P107001L en la matriu de llar), que informa sobre l'any, i en conseqüència l'onada, en què s'han recollit les dades.

## 2.2. Estructura de les matrius

Com a norma general, totes les matrius –tant les d'onada com les longitudinals– tenen com a capçalera les variables identificatives. Si treballem amb la matriu de llar les primeres són: els identificadors de llar, els identificadors de la persona de referència i els de l'informant principal, totes elles en format alfanumèric i numèric. A la matriu individual hi trobem els identificadors individuals i els identificadors de llar. A aquestes els segueixen, en les matrius de llar, una sèrie de variables resum o de recompte. Ens referim a variables del tipus nombre de membres de la llar, tipologia de la llar, nombre de menors de 16 anys, entre altres. En el cas de la matriu individual es tracta, més aviat, de variables de context tipus quina ubicació té el membre a la llar, el seu parentiu amb la persona de referència, quin tipus de membre de la mostra és... seguides de les variables estructurals dels individus.

A partir d'aquí, en les matrius d'onada segueixen les variables del qüestionari pròpiament dit. En la matriu longitudinal, segueixen les variables que han estat harmonitzades i omplertes (en el següent apartat s'explicita què és el procés d'harmonització i ompliment).

Per conèixer el contingut concret de cada una de les matrius, convé consultar el *llibre de codis* corresponent.

## 2.3. Tipus de variables i nomenclatura

Les variables de les matrius de dades segueixen un sistema de numeració lògic. Aquesta nomenclatura permet identificar el bloc temàtic del qüestionari, la dimensió llar/individual de la variable i si la variable és d'observació directa o és generada. La suma de tots aquests dígits fa que totes les variables tinguin una extensió mínima de 9 dígits. Hi ha un quadre explicatiu de la lògica de tots i cadascun dels dígits del nom d'una variable a l'annex 3. Per al que interessa per a l'anàlisi de les dades, volem destacar la lògica de l'inici i del final del nom de la variable.

Sobre **l'inici del nom de les variables**, podem trobar:

- Les variables pròpies d'una onada, que prenen com a primer dígit la lletra de l'onada. En aquest sentit, en una matriu d'onada (per exemple, 5a onada) les variables tenen el primer dígit corresponent (en l'exemple, la E).
- Les variables troncals estructurals (com l'identificador individual o el lloc de naixement), que pel seu caire d'invariabilitat en el temps, no tenen aquest primer dígit i comencen per la lletra que identifica el bloc temàtic corresponent. Poden trobar-se en totes les matrius, en

les d'onada pels individus presents en aquella onada i en les longitudinals per a tots els individus.

- Les variables longitudinals pures, que s'han generat per permetre específicament l'anàlisi longitudinal, tampoc tenen lletra d'onada. Són variables generades amb la informació disponible, tant transversal com retrospectiva, disponible durant les onze onades del PaD. Són bàsicament les referents a la trajectòria educativa, encara que es disposa també d'altres sobre treball productiu o dades dels pares. Aquestes variables poden trobar-se a les matrius longitudinals però no a les d'onada.

Sobre el final del nom de les variables, és a dir les **terminacions**, els tipus són:

- M1, M2... provenen de la partició d'una multiresposta. Poden ser categòriques o dicotòmiques. Generalment, el criteri per fer una partició o una altra ha estat el nombre de variables a generar. S'han dicotomitat les que contenen un màxim de 15 respostes possibles (és a dir, les que generen un màxim de 15 variables) i s'han categoritzat les que obliguen a fer més de 15 variables. Com conseqüència d'aquesta decisió, la majoria de multiresposta són dicotòmiques.
- G1, G2... són generades a partir de dues o més variables o no van aparèixer en el qüestionari de l'onada en qüestió. En el procés de depuració longitudinal, en alguns casos, ha estat possible generar una variable no existent a l'onada gràcies a la informació existent en les onades anteriors i/o posteriors.
- R1, R2, R3... són fruit del tancament d'una pregunta oberta, d'una recodificació d'una altra variable o d'una agrupació de la variable original.
- I informa sobre si el valor de la variable original ha estat imputat.
- H1, H2, s'aplica a aquelles variables que s'han harmonitzat i omplert longitudinalment. El procés d'harmonització suposa homogeneïtzar els codis de resposta de totes les onades, per tal de que siguin els mateixos i es pugui generar una única variable longitudinal. Paral·lelament, és necessari el procés d'ompliment. Tal i com s'ha exposat anteriorment, el fet d'utilitzar l'"enquesta dependent" per minimitzar el temps d'enquestació i evitar cansament (i rebuig a seguir participant) per part dels enquestats, genera molts valors perduts (de la forma -8 "Filtrat per base"). L'ompliment serveix per, cas que no hi hagi hagut canvis, omplir la informació de la qual disposem en onades anteriors. Així doncs, amb l'harmonització i l'ompliment, s'aconsegueix homogeneïtzar codis de resposta de les variables per a totes les onades i disposar del màxim nombre d'observacions per a cada un dels individus. Aquest tipus de variables són el contingut principal de les matrius longitudinals.

A més, aquests finals de nom de variable poden trobar-se combinats (per exemple H1M1) quan han passat per dos d'aquests processos (en l'exemple, una variable harmonitzada i posteriorment partida com a multiresposta).

## 2.4. Valors perduts

Per tal de poder reflectir la varietat de situacions en què una pregunta queda sense respondre,

s'han adoptat una sèrie de codis específics, tots ells amb valors negatius. Tot i això, les matrius es lliuren amb la definició com a valor perdut només del valor -8, corresponent als filtrats per base o no pertinents.

Els codis possibles d'absència de resposta que es poden trobar a les matrius són:

- - **1** No sap
- - **3** No imputable. Correspon a uns pocs casos d'unes onades concretes en els quals no ha estat possible imputar els ingressos.
- - **9** No contesta
- - **8** Filtrat per base/no pertinent. Correspon als casos en que la persona, o la llar, no ha de respondre la pregunta. Es troba des del cas més senzill de que per una pregunta filtre (règim de tinença de la llar) no s'ha de respondre una pregunta posterior (any de compra de l'habitatge per persones que viuen de lloguer) fins al cas més complex que té a veure amb la base de la pregunta. En el procés d'elaboració del qüestionari es defineixen les rutes que han de seguir les persones, i les llars, en funció de les seves característiques personals o familiars. Això genera el que anomenem la base de la pregunta, és a dir qui l'ha de respondre, a partir d'un joc complex de filtres. Per exemple, el nivell màxim d'estudis es demana a les persones que s'enquesten per primera vegada i a aquelles persones que van ser enquestades en la darrera onada i van declarar que estaven estudiant i que ara no estudien.

## 2.5. Els identificadors d'individu i de llar

Són les variables clau que permeten vincular la informació de les llars i dels individus o la informació de diferents onades. Per posar en relació la informació de llar i la individual és necessari fusionar els fitxers a partir de l'identificador de llar. Els identificadors de llar i dels individus es mantenen en el temps, és a dir que són invariables.

Els identificadors originals, que prenen forma alfanumèrica, s'assignen automàticament en el transcurs del treball de camp per a totes i cadascuna de les llars i dels individus de la mostra. Les variables que recullen l'identificador de llar són la P017001L (en la forma alfanumèrica) i la P017001LG2 (la traducció numèrica) i les que recullen l'identificador individual són la P008001E i la P008001EG2. La traducció numèrica de l'identificador es genera per permetre la fusió dels diferents fitxers amb qualsevol programa informàtic d'anàlisi de dades. Es recomana el seu ús per evitar problemàtiques derivades de la natura de l'identificador original. Podeu trobar informació detallada de la traducció de l'identificador alfanumèric al numèric a l'annex 2.

Els identificadors estan dotats d'una nomenclatura pròpia que proporciona informació sobre l'individu i les seves llars. La lògica dels identificadors de llar és la següent:

- la primera lletra (o els dos primers números, en la versió numèrica) indica l'onada en que la llar ha entrat a formar part de l'enquesta o s'ha creat. Així, les que comencen per A van entrar a la mostra PaD a 1a onada (2001-2002), les que comencen per B ho van fer a 2a onada (2003) i així successivament.
- La segona lletra indica, en els casos que pertoca, si la llar prové d'una altra llar que ja formava part del panel. Per exemple, una llar EA vol dir que es va crear a 5a onada (2006) i que provenia d'una llar de 1a onada. En aquest sentit, les llars captades amb posterioritat a la 1a onada tenen les dues primeres lletres coincidents amb les de l'onada en qüestió (una llar captada a 6a onada començarà per FF).

- Si d'una llar original, en una onada concreta, s'ha creat més d'una llar nova, aquestes es distingeixen amb una lletra petita al final de l'identificador (a, b..., per exemple DA0267a i DA0267b per dues llars que s'han creat a 4a onada i que provenien d'una llar de la mostra inicial). S'ha previst que alguna d'aquestes llars també pogués generar més d'una llar en una mateixa onada, pel que podria haver alguna llar amb dues lletres petites al final (la primera seria la corresponent a la llar original i la segona a l'ordre de creació de la nova llar).

Tot i això, a nivell de matrius, cal tenir en compte que l'onada de creació d'una llar i que genera un determinat identificador de llar, no vol dir que es pogués enquestar en aquella onada. És a dir, que és possible que una llar es generés, a nivell de treball de camp i de generació del nou identificador, en una onada però que no pogués ser efectivament enquestada fins una onada posterior. Per això, s'han generat les variables P040001LG1 per llars i la P040001EG2 per individus (variables sense lletra d'onada) per especificar en quina onada hi ha la primera entrevista efectiva, o l'alta a nivell de dades.

L'identificador individual es crea afegint al de llar la posició de l'individu dins la llar (dos dígit). Això vol dir que un individu que s'hagi incorporat a la mostra a 6a onada (perquè s'ha casat amb un membre original de la mostra, per exemple) en una llar AA tindrà un identificador que començarà per AA.

### 3. Els materials de suport

Amb l'objectiu de facilitar l'accés i l'ús de les dades s'han desenvolupat una sèrie de materials de suport. Els dos documents bàsics de consulta són els que es detallen a continuació.

#### 3.1. L'informe longitudinal de continguts

L'informe longitudinal de continguts s'ha elaborat amb l'objectiu d'aproximar als usuaris de les dades del PaD els continguts dels diferents blocs temàtics de l'enquesta PaD.

Aquest informe permet:

- Conèixer quines variables han aparegut a les diferents onades del PaD per cadascun dels àmbits temàtics del qüestionari i
- Identificar la correspondència entre el codi de la variable del qüestionari i el codi de la variable a la matriu de dades. Aquesta informació és cabdal, ja que en força casos la nomenclatura de les variables durant la recollida de dades es modifica després de la depuració i tractament de les dades a les matrius. També perquè algunes variables generades no tenen una correspondència directa amb les variables de qüestionari, perquè integren la informació de diverses d'elles.

L'informe conté totes les variables del PaD tret de les que només van aparèixer a la primera onada.

Aquest document proporciona la següent informació de les variables:

- l'etiqueta,
- el codi de la variable en el qüestionari,
- el codi de la variable en la matriu,
- la periodicitat,
- les onades en què ha aparegut la variable,

- si la variable ha estat eliminada o reabsorbida per una altra o per unes altres variables.

Podeu accedir a l'informe a través d'aquest [vincl](#)

### 3.2. Els llibres de codis

Un cop feta la tria sobre els àmbits d'anàlisi i les variables d'interès cal un document més específic sobre el contingut concret de cada matriu de dades. Els llibres de codis han estat dissenyats per respondre a aquesta necessitat i, en aquest sentit, són una eina indispensable per l'explotació del PaD. Cada matriu de dades disposa del seu llibre de codis propi.

L'estructura dels llibres de codis és la següent:

**Encapçalament:** en el qual s'especifica quants registres (individus o llars) conté.

**Cos central:** mostra les característiques de cadascuna de les variables en el mateix ordre en què apareixen a la matriu de dades. La informació que es dona és la següent:

- el codi de la variable,
- l'etiqueta,
- la base, és a dir quins individus o llars disposen de la informació,
- els codis de resposta,
- el nombre de casos disponible
- el nombre de no sap o no contesta que inclou la n.

Es pot accedir a tots els llibres de codis seguint aquest [vincl](#)

### 3.3. Documents metodològics

També hi ha una sèrie de documents metodològics que poden ser d'interès per als usuaris, i que recullen el mètode emprat en diferents processos de preparació de les dades. Concretament:

- Informe metodològic sobre la generació de les variables de ponderació.
- Informe metodològic sobre les variables d'imputació d'ingressos.
- Informe metodològic sobre la generació de les variables de trajectòries educatives.

Aquests documents són accessibles al [web del PaD](#). Volem recordar, també que al mateix web hi ha un recull de les recerques que s'han fet amb les dades de l'enquesta, en aquest [vincl](#).

## Annexos

### Annex 1: taules de referència dels gràfics núm. 1 i 2 sobre evolució de la mostra

#### Llars

	Captació 1a onada	Pèrdua interanual	Pèrdua acumulada	Captació 6a onada	Pèrdua interanual	Pèrdua acumulada
Onada A (2002)	1.991					
Onada B (2003)	1.703	14,47%	14,47%			
Onada C (2004)	1.528	10,28%	23,25%			
Onada D (2005)	1.404	8,12%	29,48%			
Onada E (2006)	1.090	22,36%	45,25%			
Onada F (2007)	903	17,16%	54,65%	334		
Onada G (2008)	820	9,19%	58,81%	273	18,26%	18,26%
Onada H (2009)	756	7,80%	62,03%	233	14,65%	30,24%
Onada I (2010)	719	4,89%	63,89%	217	6,87%	35,03%
Onada J (2011)	658	8,48%	66,95%	199	8,29%	40,42%
Onada K (2012)	614	6,69%	69,16%	181	9,05%	45,81%

	Captació 7a onada	Pèrdua interanual	Pèrdua acumulada	Captació 8a onada	Pèrdua interanual	Pèrdua acumulada	Total individus	Pèrdua acumulada
Onada A (2002)							1.991	
Onada B (2003)							1.703	14,47%
Onada C (2004)							1.528	23,25%
Onada D (2005)							1.404	29,48%
Onada E (2006)							1.090	45,25%
Onada F (2007)							1.237	37,87%
Onada G (2008)	431						1.524	23,46%
Onada H (2009)	363	15,78%	15,78%	154			1.506	24,36%
Onada I (2010)	334	7,99%	22,51%	125	18,83%	18,83%	1.395	29,93%
Onada J (2011)	300	10,18%	30,39%	103	17,60%	33,12%	1.260	36,72%
Onada K (2012)	260	13,33%	39,68%	88	14,56%	42,86%	1.143	42,59%

## Individus

	Captació 1a onada	Pèrdua interanual	Pèrdua acumulada	Captació 6a onada	Pèrdua interanual	Pèrdua acumulada
<b>Onada A (2002)</b>	4.587					
<b>Onada B (2003)</b>	3.686	19,64%	19,64%			
<b>Onada C (2004)</b>	3.246	11,94%	29,23%			
<b>Onada D (2005)</b>	2.954	9,00%	35,60%			
<b>Onada E (2006)</b>	2.204	25,39%	51,95%			
<b>Onada F (2007)</b>	1.787	18,92%	61,04%	745		
<b>Onada G (2008)</b>	1.599	10,52%	65,14%	576	22,68%	22,68%
<b>Onada H (2009)</b>	1.451	9,26%	68,37%	486	15,63%	34,77%
<b>Onada I (2010)</b>	1.365	5,93%	70,24%	445	8,44%	40,27%
<b>Onada J (2011)</b>	1.245	8,79%	72,86%	382	14,16%	48,72%
<b>Onada K (2012)</b>	1.130	9,24%	75,37%	340	10,99%	54,36%

	Captació 7a onada	Pèrdua interanual	Pèrdua acumulada	Captació 8a onada	Pèrdua interanual	Pèrdua acumulada	Total individus	Pèrdua acumulada
<b>Onada A (2002)</b>							4.587	
<b>Onada B (2003)</b>							3.686	19,64%
<b>Onada C (2004)</b>							3.246	29,23%
<b>Onada D (2005)</b>							2.954	35,60%
<b>Onada E (2006)</b>							2.204	51,95%
<b>Onada F (2007)</b>							2.532	44,80%
<b>Onada G (2008)</b>	957						3.132	31,72%
<b>Onada H (2009)</b>	772	19,33%	19,33%	335			3.044	33,64%
<b>Onada I (2010)</b>	683	11,53%	28,63%	264	21,19%	21,19%	2.757	39,90%
<b>Onada J (2011)</b>	594	13,03%	37,93%	203	23,11%	39,40%	2.424	47,16%
<b>Onada K (2012)</b>	495	16,67%	48,28%	168	17,24%	49,85%	2.133	53,50%

NOTA: Només es tenen en compte els individus entrevistats o proxys



## Annex 2: construcció dels identificadors numèrics

L'identificador numèric es construeix de la següent manera:

- Dos dígits per a la primera lletra i dos dígits per a la segona. El primer dígit, tant d'una lletra com de l'altra, fa referència a la dècada del panel i el segon a l'any. D'aquesta manera, assegurem que el codi no comença per zero i es reconeixerà bé per qualsevol programa informàtic, a més de preveure els codis que es generaran més enllà de la 10a onada.
- Els nombres de l'identificador
- Un dígit per a la primera lletra petita
- Un dígit per a la segona lletra petita
- Dos dígits per a la posició del membre, en el cas dels identificadors individuals

D'aquesta manera la llar AA0154 es traduirà en 1111015400; la llar BA1854a en 1211185410 i la llar FA2342ab en 1611234212. En el cas dels identificadors individuals, a aquests codis només caldrà afegir la posició del membre: per exemple AA015401 o 111101540001.

### EXEMPLES D'IDENTIFICADORS I EQUIVALÈNCIA ENTRE ELS ALFANUMÈRICS I ELS NUMÈRICS

Exemple	Tipus id*	Onada d'incorporació de la llar	Onada de creació de la llar original	Nº de la llar	Indicatiu que d'una llar original se'n han derivat dues o més	Codi llar Resultant	Ubicació que ocupa l'individu a la llar	Codi individual Resultant (dos dígits més)
1	O	A	A	0012	-	AA0012	02	AA001202
	G	11	11	0012	00	1111001200	02	111100120002
2	O	B	A	0365	-	BA0365	05	BA036505
	G	12	11	0365	00	1211036500	05	121103650005
3	O	D	A	0128	a	DA0128a	07	DA0128a07
	G	14	11	0128	10	1411012810	07	141101281007
4	O	F	C	1642	b	FC1642b	04	FC1642b04
	G	16	13	1642	20	1613164220	04	161316422004

\*ID= identificador; \*\*O=original (alfanumèric); \*\*\*G=generat (numèric)  
Font: elaboració pròpia

### Annex 3: Fitxa tècnica del qüestionari

<b>Principi metodològic</b>	<p>Els qüestionaris del PaD es treballen en base als següents criteris metodològics:</p> <ul style="list-style-type: none"> <li>▪ <b>Simplicitat:</b> per contenir els canvis a introduir al qüestionari.</li> <li>▪ <b>Objectivació i fiabilitat:</b> per donar compte dels motius finals que han motivat el canvi i l'elecció d'una determinada acció vers la variable.</li> <li>▪ <b>Rellevància:</b> per assegurar l'ús potencial de la informació que es recollirà.</li> <li>▪ <b>Documentació:</b> per garantir l'acreditació en les modificacions adoptades.</li> </ul>
<b>Unitats d'anàlisi</b>	Llar i Individu
<b>Tipus de qüestionaris</b>	<p><b>Qüestionari Llar:</b> el contesta l'informant principal (una sola persona per llar).</p> <p><b>Qüestionari Individual:</b> el contesten tots els membres de la llar que tenen més de 15 anys.</p> <p><b>Qüestionari Proxy:</b> pensat per a les persones a qui no es pot enquestar perquè durant el treball de camp no estan presents a la llar; no són aptes per contestar l'enquesta; etc. El qüestionari proxy s'administra a una persona disposada a facilitar dades bàsiques (preguntes objectives) de la persona a qui no podem enquestar.</p>
<b>Contingut</b>	Dades identificatives i de gestió; Educació; Treball productiu, reproductiu i conciliació; Usos del temps ; Ingressos; Habitatge; Salut i dependència; Situació personal i familiar; Posició i relacions socials; Valors i actituds.
<b>Tipus de preguntes</b>	Combinació perspectiva objectiva i subjectiva. Combinació de preguntes troncal (enquesta longitudinal) i cíclics (apareixen cada x anys).
<b>Calendari de treball</b>	<p><b>Inici</b> elaboració qüestionaris: març</p> <p><b>Tancament</b> qüestionaris onada: desembre</p> <p><b>Prova pilot:</b> gener-febrer</p>
<b>Activitats principals</b>	<ul style="list-style-type: none"> <li>▪ Discussió dels continguts</li> <li>▪ Documentació dels canvis de les variables</li> <li>▪ Informatització dels qüestionaris</li> <li>▪ Testeig del qüestionari (prova pilot)</li> <li>▪ Control de Qualitat de l'Enquestació</li> </ul>
<b>Outputs</b>	<ul style="list-style-type: none"> <li>▪ Qüestionaris per onada (Llar, Individu, Proxy)</li> <li>▪ Informe longitudinal de continguts (General i per blocs)</li> <li>▪ Informe longitudinal tècnic</li> </ul>

## Periodicitat de les preguntes del pad

La **periodicitat** diferenciada de cada una de les preguntes del PaD implica que a cada onada els qüestionaris siguin diferents.

TRONCAL	<p>Preguntes que apareixen a cada onada des del moment de la seva inclusió al qüestionari. Serveixen per analitzar el present de les llars i els individus i estan enfocades a l'anàlisi del microcanvi any rere any.</p> <p>Per exemple: Està cursant estudis reglats actualment? Quins? Quina és la seva situació respecte el treball remunerat?</p>
TRONCAL ESTRUCTURAL	<p>Les preguntes troncal estructurals són aquelles que recullen informació amb variació nul·la en la vida de l'enqu Coastat/da i que, per tant, només es recullen una vegada en tota la vida del panel. Aquest és el cas de variables com ara l'any de naixement, el sexe, l'edat d'inici de la vida laboral, el lloc i l'any de naixement del pare o la mare...</p> <p>Tot i això, i per tal de contrastar dades i assegurar la consistència de la informació, hi ha preguntes troncal estructurals que s'han inclòs al qüestionari en més d'una ocasió.</p>
CÍCLICA	<p>Preguntes que apareixen de forma cíclica, ja que s'estima que té sentit plantejar-se l'anàlisi del canvi només cada X anys. S'indica C2, C3, C4... en funció de cada quantes onades ha d'aparèixer cada pregunta cíclica al qüestionari.</p> <p>Per exemple: Qui organitza i coordina el treball domèstic i familiar? (C2) Quines coses milloraria del seu habitatge? (C4) Quins dels familiars següents viuen a prop seu? (C8)</p>
ESPORÀDICA	<p>Preguntes que apareixen de forma esporàdica en alguna de les onades del panel i per a les quals no es concreta data de nova inclusió. El seu caràcter esporàdic implica que no s'espera fer cap anàlisi en termes d'estudi del canvi.</p> <p>Per exemple: Mostri el seu grau d'acord o desacord amb la següent afirmació: a Catalunya hi ha massa immigrants.</p>

## Descripció de la numeració lògica de les variables

DÍGIT Matrius d'onada	DÍGIT Matriu long	NOM DEL DÍGIT	CONTINGUTS
1		ONADA DE LA VARIABLE	Cada onada s'anomena per la lletra de l'abecedari que li pertoca. D'aquesta manera, totes les variables dels fitxers de la 1a Onada estan encapçalades per la lletra A, les de la 2a Onada per la lletra B, les de la 3a Onada per la lletra C i així successivament. Així, la matriu individual de 1a onada és la A_INDI, la de segona B_INDI, etc. El mateix per les matrius de llar.
2	1	BLOC DEL QÜESTIONARI	La correspondència entre el nom del bloc del qüestionari i la lletra és la següent: <b>P</b> , portada; <b>L</b> , llar; <b>E</b> , educació; <b>F</b> , fecunditat; <b>H</b> , habitatge; <b>G</b> , llengua; <b>R</b> , posició i relacions socials; <b>S</b> , salut; <b>T</b> , treball; <b>D</b> , treball domèstic; <b>I</b> , posició econòmica; <b>U</b> , temps de lleure; <b>M</b> , mitjans de comunicació; <b>C</b> , cultura política; <b>O</b> , immigració; <b>K</b> , dades de seguiment; <b>W</b> , ponderació.
3, 4, 5	2, 3, 4	DÍGITS PER NUMERAR LA VARIABLE	A mesura que es crea una nova variable se li assigna el número que pertoca, correlativament.
6, 7, 8	5, 6, 7	CANVI UNITATS O POSSIBILITATS DE RESPOSTA QUE TÉ LA VARIABLE	En algunes preguntes i per tal d'aconseguir el màxim nombre de respostes possible, es dona l'opció de respondre en unitats diferents (per exemple, euros/mes i euros/any). Aquestes variables mantenen tots els dígit del seu nom lògic a excepció del dígit 7, que serveix per distingir la unitat en que es recull la resposta.
9	8	TIPUS QÜESTIONARI	Per diferenciar el qüestionari al què pertany la variable: <b>E</b> : per a les variables dels qüestionaris individuals i proxys <b>L</b> : per a les variables dels qüestionaris de llar Cal recordar l'excepcionalitat de les variables que contenen informació individual però que es recullen en el qüestionari de llar i que, en les matrius, canvien la lletra final de L a E.
>9	>8	IDENTIFICADOR DEL TIPUS DE VARIABLE	A més, s'afegeixen les següents lletres i, eventualment, un nombre a les variables en els següents casos: - <b>M1, M2...</b> , provenen de la partició d'una multiresposta; - <b>G1</b> , són generades a partir de dues o més variables; - <b>R1, R2, R3...</b> són fruit d'una recodificació o d'una agrupació de la variable original; - <b>I</b> , informa, dicotòmicament, sobre si el valor de la variable prové de la variable original o ha estat imputat; - <b>H1</b> , s'aplica a aquelles variables que s'han harmonitzat i omplert longitudinalment. Poden trobar-se combinats (per exemple H1M1) quan han passat per dos d'aquests processos (en l'exemple, una variable harmonitzada i posteriorment partida com a multiresposta).

## Annex 4: Mostra

### Fitxa tècnica de la mostra

<b>Principis metodològics</b>	<p><b>Enquesta Panel pur:</b> seguiment al llarg del temps d'una mateixa mostra de llars i individus.</p> <p><b>Remostrejos</b> periòdics per tal de garantir el sostre mostral de 2.000 llars i 4.600 individus enquestats.</p>																														
<b>Univers</b>	<p><b>Llars privades</b> de Catalunya i els <b>individus</b> que les conformen.</p>																														
<b>Mostra original (1a onada)</b>	<p>→ <b>Mostra teòrica:</b> 2000 llars de Catalunya; 5.456 individus (2,7 individus/llar de mitjana; 4.600 dels quals enquestats =&gt; majors 15 anys)</p> <p>→ <b>Mostra captada</b> (fi 1a onada): 1991 llars; 5.757 individus, 4.589 enquestats.</p>																														
<b>Tipus de mostreig</b>	<p>1a etapa: Distribució apropiada a nivell de província i proporcional a nivell de grandària de municipi del nombre de llars a entrevistar.</p> <p>2a etapa: Selecció aleatòria sistemàtica i polietàpica de les primeres unitats mostrals (seccions censals) segons variables d'estratificació territorial i social:</p> <ul style="list-style-type: none"> <li>▪ àrea geogràfica</li> <li>▪ grandària municipi</li> <li>▪ estructura social (= &gt; situació ocupacional)</li> </ul> <p>3a etapa: Selecció aleatòria simple de les unitats mostrals finals =&gt; domicilis de les llars a entrevistar en les primeres unitats mostrals.</p>																														
<b>Remostreigs</b>	<p>Els remostreigs es realitzen seguint els mateixos criteris d'aleatorietat i segmentació de la mostra de la 1a Onada.</p> <table border="1" style="width: 100%; border-collapse: collapse;"> <thead> <tr> <th></th> <th style="text-align: center;">Noves llars captades</th> <th style="text-align: center;">Llars de seguiment</th> <th style="text-align: center;">Llars totals mostra PaD</th> </tr> </thead> <tbody> <tr> <td><b>Resultats 1r Remostreig</b></td> <td></td> <td></td> <td></td> </tr> <tr> <td>6a Onada (2007):</td> <td style="text-align: center;">333</td> <td style="text-align: center;">1154</td> <td style="text-align: center;">1487</td> </tr> <tr> <td>7a Onada (2008):</td> <td style="text-align: center;">431</td> <td style="text-align: center;">1446</td> <td style="text-align: center;">1877</td> </tr> <tr> <td><b>Resultats 2n Remostreig</b></td> <td></td> <td></td> <td></td> </tr> <tr> <td>8a Onada (2009):</td> <td style="text-align: center;">154</td> <td style="text-align: center;">1745</td> <td style="text-align: center;">1899</td> </tr> <tr> <td>10a Onada (2011):</td> <td style="text-align: center;">154</td> <td style="text-align: center;">1784</td> <td style="text-align: center;">1938</td> </tr> </tbody> </table>				Noves llars captades	Llars de seguiment	Llars totals mostra PaD	<b>Resultats 1r Remostreig</b>				6a Onada (2007):	333	1154	1487	7a Onada (2008):	431	1446	1877	<b>Resultats 2n Remostreig</b>				8a Onada (2009):	154	1745	1899	10a Onada (2011):	154	1784	1938
	Noves llars captades	Llars de seguiment	Llars totals mostra PaD																												
<b>Resultats 1r Remostreig</b>																															
6a Onada (2007):	333	1154	1487																												
7a Onada (2008):	431	1446	1877																												
<b>Resultats 2n Remostreig</b>																															
8a Onada (2009):	154	1745	1899																												
10a Onada (2011):	154	1784	1938																												

## Detalls tècnics sobre la mostra teòrica

### Univers

Llars privades de Catalunya

### Mostra teòrica i marges d'error

Llars

ÀREES	NOMBRE DE LLARS	e* (%)
<b>Barcelona</b>	<b>800</b>	<b>±3,54</b>
Ciutat BCN	261	±6,12
Regió Metropolitana	470	±4,61
Comarques centrals	69	±12,04
<b>Girona</b>	<b>400</b>	<b>±5,00</b>
<b>Lleida</b>	<b>400</b>	<b>±5,00</b>
<b>Tarragona</b>	<b>400</b>	<b>±5,00</b>
<b>CATALUNYA</b>	<b>2.000</b>	<b>±2,79</b>

\*e = error mostral (interval confiança=95,5%, k=2 i p=q=50)  
Font: elaboració pròpia

### Individus

ÀREES	NOMBRE D'INDIVIDUS	e* (%)
Barcelona	2.178	±2,14
Girona	1.054	±3,08
Lleida	1.120	±2,98
Tarragona	1.104	±3,01
<b>CATALUNYA</b>	<b>5.456</b>	<b>±1,67</b>

\*e = error mostral (interval confiança=95,5%, k=2 i p=q=50)  
Font: elaboració pròpia

## Annex 5: El treball de camp

### Fitxa tècnica del treball de camp

<b>Principi metodològic</b>	El principi metodològic que regeix el treball de camp, és desenvolupar un treball personalitzat i de màxima proximitat amb l'equip d'enquestadors i les llars col·laboradores per tal d'optimitzar la qualitat de les dades recollides i la fidelització al projecte. En el context d'una enquesta longitudinal, la continuïtat tant de llars com d'enquestadors és un aspecte clau.
<b>Periodicitat de l'enquesta</b>	Anual
<b>Calendari d'enquestació</b>	<ul style="list-style-type: none"><li>▪ 1a etapa: març-juliol: 80% enquestació</li><li>▪ 2a etapa*: setembre -octubre: 20% enquestació *etapa destinada a la recuperació de llars que a la 1a etapa no s'han pogut enquestar.</li></ul>
<b>Mode d'administració de l'enquesta</b>	Presencial (1a a 8a onades). Mètode Mixt: Telefònic-Presencial (9a onada en endavant): <ul style="list-style-type: none"><li>▪ Mostra enquestada telefònicament: 80%.</li><li>▪ Mostra enquestada presencialment: 20% (inclou llars de nova incorporació a l'enquesta).</li></ul>
<b>Univers</b>	Persones aptes majors de 15 anys considerades de la llar
<b>Sistema d'administració de l'enquesta</b>	CAPI ( <i>Computer Assisted Personal Interview</i> ) => Entrevista assistida per ordinador
<b>Equip d'enquestadors/es</b>	Formació en ciències socials, amb alta motivació per participar a l'enquesta => 80 % continuïtat enquestadors (mitjana 1a a 8a onades PaD)
<b>Participació de les llars</b>	Voluntària
<b>Incentiu participació</b>	Obsequi per llar
<b>Controls de qualitat de la informació</b>	<ul style="list-style-type: none"><li>▪ COE: Control de qualitat de l'enquestació (mesura qualitat de la informació recollida)→ <i>veure apartat Qüestionari</i></li><li>▪ CT: Control telefònic a les llars (comprova la realització de l'entrevista): 15% de les llars enquestades</li></ul>
<b>Principals activitats</b>	<ul style="list-style-type: none"><li>▪ Borsa On-Line d'enquestadors (oferta permanent).</li><li>▪ Identificació de les llars a enquestar.</li><li>▪ Gestió del treball de camp: enquestadors/es i llars: formació personalitzada i seguiment via reunions de coordinació.</li><li>▪ Ampliació de la mostra de llars (captació).</li><li>▪ Tancament de l'onada de treball de camp.</li><li>▪ Preparació de les dades de seguiment per l'onada següent.</li></ul>



Mostra de treball de camp (inclou totes les llars treballades en cada onada)

	LLARS INICI ONADA*	LLARS TREBALLADES**	LLARS ENQUESTADES***
1 <sup>a</sup> onada 2001/02	0	1991	1991
2 <sup>a</sup> onada 2003	1991	2149	1777
3 <sup>a</sup> onada 2004	1931	2044	1693
4 <sup>a</sup> onada 2005	1791	1880	1633
5 <sup>a</sup> onada 2006	1724	1791	1323
6 <sup>a</sup> onada 2007	1473	1862	1488
7 <sup>a</sup> onada 2008	1749	2268	1877
8 <sup>a</sup> onada 2009	2089	2332	1899
9 <sup>a</sup> onada 2010	2114	2219	1940
10 <sup>a</sup> onada 2011	2043	2266	1938
11 <sup>a</sup> onada 2012	2069	2127	1792

\*NOMBRE LLARS INICI ONADA El nombre de llars a l'inici de l'onada inclou: llars entrevistades onada anterior + llars no entrevistades que relancem al camp (rebuig/perdut no definitiu)

\*\*NOMBRE LLARS TREBALLADES ONADA El nombre de llars treballades inclouen les llars de l'inici de l'onada (llars enquestades onada anterior + llars perdut/rebuig no definitiu onada anterior) més les llars que es creen durant l'onada (per la pròpia dinàmica de la mostra PaD: separacions, individus que s'independitzen,...). En anys de Captació (6a, 7a, 8a i 10a) també s'inclouen les llars captades durant l'onada.

\*\*\*NOMBRE LLARS ENQUESTADES El nombre de llars enquestades inclou les llars treballades durant el TC que s'ha aconseguit enquestar.

Percentatge acumulat d'entrevistes (per mes de treball de camp)

	2a Onada 2003	3a Onada 2004	4a Onada 2005	5a Onada 2006	6a Onada 2007	7a Onada 2008	8a Onada 2009	9a Onada 2010	10a Onada 2010	11a Onada 2010
Gener	5,3	11,3	0,1		3,4	1,8	0,4			
Febrer	29,6	40,8	28,3	24,7	21,0	23,8	15,4			0,7
Març	55,0	66,2	57,0	56,7	42,0	41,8	33,8		7,8	6,8
Abril	72,5	79,0	77,6	68,4	55,4	54,8	50,5	26,3	37,9	35,3
Maig	83,3	90,4	87,8	81,1	65,2	65,3	68,6	68,3	70,6	66,2
Juny	92,2	96,0	93,1	91,6	72,8	77,6	81,4	91,0	86,8	87,6
Juliol	96,5	96,9	95,0	98,4	80,4	87,6	89,0	93,8	89,8	92,1
Agost	96,5	96,9	95,0	98,4	81,4	88,4	89,5	93,8	91,5	95,0
Setembre	96,5	96,9	95,0	99,5	84,7	92,2	93,2	93,9	95,1	97,5
Octubre	99,0	99,1	98,8	100,0	98,7	98,5	98,5	98,6	99,4	100,0
Novembre	100,0	100,0	100,0		100,0	100,0	100,0	100,0	100,0	

## Annex 6: Dades

### Fitxa tècnica de depuració i preparació de les matrius de dades

<b>Principi metodològic</b>	<p>El principi bàsic que regeix la feina de depuració de les dades del PaD és no modificar les dades facilitades pels enquestats si no hi ha un error evident que es pugui corregir. Si una persona enquestada ofereix respostes inconsistentes i no hi ha cap evidència que permeti fer algun tipus de correcció, es manté la inconsistència i pertoca a cada analista decidir què fer amb ella.</p>
<b>Programes tractament dades</b>	<p>SPSS Microsoft Excel Barbwin</p>
<b>Calendari de depuració (una Onada)</b>	<p>Inici depuració: octubre/novembre (coincidint amb fi del treball de camp). Finalització de la depuració: desembre (any següent).</p> <p>El calendari està subjecte a variacions en funció de les característiques de l'onada i de la disponibilitat de recursos humans.</p>
<b>Principals activitats</b>	<ul style="list-style-type: none"><li>▪ Consolidació de les dades.</li><li>▪ Preparació de les matrius individual i de llar per a la depuració de les variables.</li><li>▪ Codificació de les preguntes obertes del qüestionari.</li><li>▪ Depuració per blocs de les variables individuals i de llar.</li><li>▪ Encàrrec i seguiment de la imputació de les variables d'ingressos (individuals i de llar).</li><li>▪ Control de les bases de les variables de les matrius individuals i de llar.</li><li>▪ Comprovació final del contingut de totes les variables a través de les freqüències (tabulació).</li><li>▪ Controls de consistència de les variables estructurals amb altres fonts estadístiques (per avaluar la consistència de les dades del PaD).</li><li>▪ Càlcul de les variables generades del PaD (variables creades per l'equip del PaD).</li><li>▪ Encàrrec i seguiment de la generació de les variables de ponderació.</li><li>▪ Elaboració dels materials de suport.</li></ul>
<b>Outputs</b>	<p>Matrius de dades</p> <ul style="list-style-type: none"><li>• Transversals: X_INDI i X_LLAR (una per cada onada)</li><li>• Longitudinals: A_K INDI_LONG, B_K INDI LONG i A_K LLAR_LONG, B_K LLAR LONG</li></ul> <p>Llibres de Codis:</p> <ul style="list-style-type: none"><li>• Un llibre de codis per cada onada del PaD.</li><li>• Un llibre de codis per cada una de les matrius longitudinals</li></ul> <p>Manual d'ús i informe sobre Canvis i Millores a les matrius (es pot descarregar seguint aquest <a href="#">vinçle</a>)</p>

## Nombre de variables a les matrius

	Llar	Indi
1a onada 2001/02	236	236
2a onada 2003	179	418
3a onada 2004	229	863
4a onada 2005	205	432
5a onada 2006	223	569
6a onada 2007	226	487
7a onada 2008	275	936
8a onada 2009	246	518
9a onada 2010	60	145
10a onada 2011	148	344
11a onada 2012	148	743
A_K LONG	42	42
B_K LONG	52	239

## Annex 7: Fitxa tècnica de l'anàlisi

<b>Principi metodològic</b>	<p>El principi bàsic que regeix l'anàlisi de dades del PaD és posar a l'abast de la comunitat investigadora totes les dades del PaD de forma gratuïta. La FJB impulsa l'anàlisi de les dades (producció interna), alhora que dona suport i fa un seguiment de l'ús extern de les dades (producció externa).</p>
<b>Calendari de treball</b>	<p>Permanent (no subjecte a dates)</p>
<b>Principals activitats</b>	<ul style="list-style-type: none"> <li>▪ Atenció a les demandes d'informació d'institucions que s'interessen per les dades del PaD</li> <li>▪ Cessió de les dades del PaD</li> <li>▪ Encàrrecs interns d'anàlisi</li> <li>▪ Seguiment ús extern de les dades</li> <li>▪ Gestió grup d'usuaris del PaD</li> <li>▪ Publicació d'informes del PaD</li> </ul>
<b>Outputs</b>	<p>Publicacions:</p> <ul style="list-style-type: none"> <li>▪ Col·lecció Polítiques</li> <li>▪ Col·lecció informes Breus</li> <li>▪ Documents metodològics (on-line)</li> <li>▪ Informes de resultats (on-line)</li> </ul> <div data-bbox="995 898 1286 1339" style="text-align: right;"> </div> <p>Àmbits d'ús de les dades: Universitats, Administració Pública, Agents Socials, Mitjans de comunicació de Catalunya, Espanya, Europa.</p> <p>Producció compartida: Alguns dels informes resultants s'han publicat de forma conjunta amb altres institucions.</p>

## Annex 8: Fitxa tècnica de difusió

<p><b>Principi metodològic</b></p>	<p>El principi bàsic que regeix la difusió de dades del PaD és arribar a diferents públics: l'acadèmic, els agents socials i polítics responsables del disseny i implementació de les polítiques públiques i la societat en general.</p> <p>També es té una especial atenció a la difusió de resultats entre les llars col·laboradores del PaD, i entre els enquestadors del PaD en senyal d'agraïment i com a demostració de la utilitat de la seva participació a l'enquesta.</p>
<p><b>Calendari de treball</b></p>	<p>Difusió en l'àmbit acadèmic i dels agents polítics i socials:</p> <ul style="list-style-type: none"> <li>▪ Permanent (subjecte a la disponibilitat dels resultats dels estudis)</li> </ul> <p>Difusió en l'àmbit de les llars del PaD, els enquestadors i la societat en general</p> <ul style="list-style-type: none"> <li>▪ Tríptics (inici- mig i final del treball de camp)</li> <li>▪ Dossiers de premsa (març: inici onada)</li> </ul>
<p><b>Principals activitats</b></p>	<ul style="list-style-type: none"> <li>▪ Elaboració tríptics del PaD (comunicació amb llars de la mostra).</li> <li>▪ Elaboració de presentacions de l'enquesta i dels seus resultats en diferents formats.</li> <li>▪ Suport a les publicacions (pre i post edició).</li> <li>▪ Organització de jornades i seminaris per la presentació dels resultats.</li> <li>▪ Manteniment i actualització de la pàgina web del PaD.</li> <li>▪ Manteniment i actualització de les bases de dades per la difusió del PaD.</li> <li>▪ Elaboració de butlletins electrònics periòdics (Butlletí FJB i Butlletins PaD) amb les novetats del PaD.</li> </ul>
<p><b>Outputs</b></p>	<ul style="list-style-type: none"> <li>▪ Web del PaD =&gt; <a href="http://www.obdesigualtats.cat/">http://www.obdesigualtats.cat/</a></li> <li>▪ Col·lecció tríptics del PaD =&gt; <a href="http://www.obdesigualtats.cat/index.php?cm=16.01">http://www.obdesigualtats.cat/index.php?cm=16.01</a></li> <li>▪ Presentacions, en diferents formats, <i>ad hoc</i>.</li> <li>▪ Publicacions =&gt; <a href="http://www.obdesigualtats.cat/index.php?cm=15">http://www.obdesigualtats.cat/index.php?cm=15</a></li> </ul>