

El Panel de Desigualtats Socials a Catalunya (PaD)

Apunts metodològics
de l'enquesta

*FUNDACIÓ JAUME BOFILL
(EQUIP DEL PAD)*



El Panel de Desigualtats Socials a Catalunya (PaD)

Apunts metodològics de l'enquesta

FUNDACIÓ JAUME BOFILL

(EQUIP DEL PAD)

El Panel de Desigualtats Socials a Catalunya (PaD)

Apunts metodològics de l'enquesta

FUNDACIÓ JAUME BOFILL
(EQUIP DEL PAD)

La col·lecció *Polítiques* és la principal col·lecció de la Fundació Jaume Bofill. S'hi publiquen aquelles investigacions que han estat encarregades i que es consideren més rellevants pel seu interès social i polític. Les opinions que s'hi expressen corresponen als autors.

La reproducció total o parcial d'aquesta obra per qualsevol procediment, compresos la reprografia i el tractament informàtic, resta rigorosament prohibida sense l'autorització dels propietaris del *copyright*, i estarà sotmesa a les sancions establertes a la llei.

Primera edició: juny de 2008

© Fundació Jaume Bofill, 2008
Provença, 324
08037 Barcelona
fbofill@fbofill.cat
<http://www.fbofill.cat>

© d'aquesta edició: Editorial Mediterrània, S. L.
Guillem Tell, 15, entl. 1a
08006 Barcelona
Tel. 93 218 34 58 – Fax 93 237 22 10
editorial@editorialmediterrania.cat
www.editorialmediterrania.cat

Disseny de la col·lecció: Martí Abril
Disseny de la coberta: Amador Garrell
Fotografia de la coberta: Lluís Salvadó
Maquetació: Jordi Vives

ISBN de l'Obra Completa: 978-84-8334-922-9
ISBN del volum Apunts metodològics: 978-84-8334-925-0
DL: B-27.642-2008

Impressió: Romanya Valls, Capellades (Anoia)
Imprès a Catalunya – *Printed in Catalonia*

Índex

PRESENTACIÓ	9
1. ELS ORÍGENS DE L'ENQUESTA PAD	13
Una enquesta longitudinal	16
Dues unitats d'estudi: la llar i els seus individus	17
Aspectes objectius i subjectius de la vida dels individus	18
2. EL QÜESTIONARI	19
Perspectiva transversal <i>versus</i> longitudinal	22
Perspectiva de llar <i>versus</i> individual	23
El pes dels blocs temàtics	24
Programa d'enquestació CAPI	25
3. EL TREBALL DE CAMP, UNA APOSTA PER LA QUALITAT DE LES DADES	27
Plantejaments generals i objectius del treball de camp	29
La formació dels enquestadors, element clau per a la qualitat	30
El paper de l'enquestador	31
Els controls de qualitat de l'enquestació	32
La continuïtat dels enquestadors del PaD	33
El calendari d'enquestació	34
4. LA MOSTRA	37
El disseny de la mostra i els seus resultats finals	39
Les variables de ponderació	42
L'evolució de la mostra al llarg de les onades	45
L'evolució de la mostra de llars del PaD	46
L'evolució de la mostra d'individus del PaD	51
5. L'ELABORACIÓ DE LES MATRIUS DE DADES	57
La consolidació i depuració de les dades	59

Creació i actualització de les variables generades	61
Control de les freqüències de les variables	62
Anonimització de les matrius de dades	63
Organització de les variables a les matrius de dades	63
Tipologia de les matrius del PaD	64
Els llibres de codis	66
Creació de matrius longitudinals	67
Materials de suport: els informes longitudinals	67
 6. L'ACCÉS A LES DADES DEL PAD	 69
Criteris de cessió de les dades del PaD	71
Com formular la demanda de dades	71
La cessió de dades i el contracte	72
Finalització de la recerca	72

Presentació

Sovint quan llegim els resultats d'una enquesta o d'un treball d'investigació, desconeixem com s'han recollit les dades i el procés al qual han estat sotmeses abans de ser analitzades. El Panel de Desigualtats Socials a Catalunya (PaD) és una enquesta complexa i amb algunes particularitats que val la pena conèixer abans d'endinsar-se en els seus resultats.

El fet que es tracti d'una enquesta longitudinal, amb vocació de resseguir la trajectòria de les llars, dels individus que les conformen i de conèixer tant els aspectes objectius com subjectius de la vida de les persones, té implicacions que no són menors. Aquestes tres dimensions que caracteritzen el PaD afecten en gran mesura els continguts dels qüestionaris que s'elaboren a cada onada, el desenvolupament del treball de camp, les depuracions que s'han d'aplicar sobre les dades recollides i la manera com es poden analitzar.

Les particularitats de l'enquesta junt a la voluntat de transparència de la Fundació Jaume Bofill al voltant del PaD, ens han portat a considerar l'oportunitat d'adjuntar un petit annex metodològic a aquesta publicació.

Hem seleccionat amb cura la informació d'aquest annex, pensant en allò que pot interessar el lector o lectora dels informes de resultats. Aquest annex no vol ni pot

substituir el *Manual d'ús* del PaD (disponible a www.obdesigualtats.cat), que és l'eina indispensable per a qualsevol que es plantegi endinsar-se en l'explotació de les dades del PaD.

Aprofitem l'avinentesa per animar tots els investigadors i investigadores a fer ús de les dades del PaD, ja que està en mans de la comunitat investigadora evidenciar la utilitat del PaD com a instrument de coneixement de la realitat. Estem convençuts que l'explotació de les dades del PaD és l'actuació que més pot contribuir garantir-ne la continuïtat.

Equip del PaD
 Barcelona, primavera 2008

EQUIP PaD		
Direcció de l'enquesta	Jordi Sánchez	jsanchez@fbofill.cat
	Laia Pineda	lpineda@fbofill.cat
Coordinació general	Meritxell Argelagués	margelagues@fbofill.cat
Qüestionari	Anna Tarrés	atarres@fbofill.cat
Treball de camp	Carla Domínguez	pad@obdesigualtats.cat
	Mercè Mauri	pad2@obdesigualtats.cat
	Jep Jorba	jjorba@obdesigualtats.cat
	M. Ona Sanmartí	padborsa@obdesigualtats.cat
Depuració i confecció de matrius	Juli Carrere	jcarrere@fbofill.cat
Mostra i explotació de les dades	Àngels Llorens	allorems@fbofill.cat

1 Els orígens de l'enquesta PaD

L'enquesta PaD neix l'any 2001 de la necessitat d'aportar informació rellevant al nostre país sobre l'estructura social i les desigualtats de la societat catalana. Però, a més d'aportar dades per conèixer millor com operen els mecanismes de les desigualtats, el PaD neix també amb la voluntat de ser una eina útil per als responsables de les polítiques públiques —teòrics, tècnics i polítics— en la mesura que conèixer la situació real de la població és, sota qualsevol consideració, el pas previ per orientar una actuació.

El PaD parteix de la definició bàsica de la desigualtat com l'expressió social de les diferents possibilitats dels individus per accedir als béns i als recursos socials com ara l'educació, el treball, la reproducció, l'habitatge, la salut o la riquesa. El PaD intenta fer emergir no només les diferències entre la població enquestada, sinó també els mecanismes que generen o reforcen les desigualtats socials.

L'enquesta PaD té tres particularitats bàsiques que ens ajuden a comprendre com ha estat estructurada:

1. és una enquesta longitudinal,
2. recull informació a un doble nivell: llar i individus, i
3. busca l'equilibri entre la recollida d'aspectes objectius i subjectius de la vida dels individus.

UNA ENQUESTA LONGITUDINAL

Un dels potencials més grans de l'enquesta del PaD és que es tracta d'una enquesta longitudinal o enquesta panel. Cal subratllar que aquest és un tret diferencial respecte d'altres enquestes de Catalunya, ja que es tracta de la primera enquesta longitudinal que s'ha impulsat des de Catalunya amb una mostra estadísticament prou representativa de llars catalanes.¹

Des d'un punt de vista estrictament tècnic, una enquesta panel té com a característica bàsica que es fa sempre sobre la mateixa mostra d'individus al llarg de les diferents onades. Els avantatges de fer el seguiment d'una mostra panel es poden resumir en quatre: incorpora la perspectiva del canvi, les dades tenen major significació, millor qualitat i se n'obté una major quantitat.

- **Perspectiva del canvi.** La mostra panel ens permet recollir dades molt valuoses sobre el canvi: canvis d'habitatge, canvis en l'estructura de les llars, canvis en el treball, etc. El fet que preguntem als mateixos individus any rere any, ens permet conèixer no només quins canvis es produeixen en les seves vides, sinó que ens permet localitzar en quins grups socials es donen els canvis, amb quina intensitat, les conseqüències d'aquests canvis sobre altres aspectes de la vida personal i familiar, etc.
- **Significació.** La mostra panel, en la mesura que enquestem les mateixes persones onada rere onada, implica que, en termes estadístics, els canvis registrats siguin més significatius que si enquestéssim una mostra nova d'individus cada vegada.
- **Qualitat.** Una mostra panel ens assegura una recollida de dades molt millor: el retorn a les mateixes llars any rere any ens permet depurar errors i validar moltes de les dades recollides sobre els individus que les componen.

1. A Catalunya hi ha hagut llars que han format part de mostres d'enquestes longitudinals impulsades per l'Institut Nacional de Estadística (INE). Un exemple és l'"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. Aquestes mostres, però, mai no han estat estadísticament prou representatives com per fer estudis específics de Catalunya.

- **Quantitat.** La mostra panel ens permet desenvolupar amb profunditats diferents els diversos blocs de preguntes en cadascuna de les onades (un any entrem a fons en el tema de l'educació, l'altre en el treball domèstic, un altre en temes d'habitatge...).

Consegüentment, l'interès d'una enquesta panel no està tant en la tècnica de mostreig (que només actua com a mitjà) com en el seu fi: maximitzar el potencial analític de les dades que se n'obtenen.

L'enquesta panel incorpora la possibilitat d'una visió longitudinal de la realitat. Ens permet anar més enllà de la "instantània" i visionar un període llarg de temps. La perspectiva longitudinal permet donar compte de la trajectòria dels individus, analitzar com experimenten els canvis i la resposta que els donen, en el seu entorn socioeconòmic, els individus i les llars. L'enquesta panel, a més, permet disposar d'informació sobre els moments especialment rellevants en la vida dels individus en les nostres societats: processos de transició escola-treball, treball-jubilació; emancipació de la llar dels pares a una de pròpia; naixement dels fills; pas cap al llinar per sota del qual es considera que s'està en situació de pobresa monetària, entre molts altres processos de canvi.

DUES UNITATS D'ESTUDI: LA LLAR I ELS SEUS INDIVIDUS

La segona particularitat del PaD és que té una doble unitat d'estudi: les llars i els seus individus. Aquesta característica també es remet a la voluntat d'ampliar al màxim el potencial explicatiu de les dades recollides. L'anàlisi de les dades del PaD és primordialment individual, però esdevé molt més rica si podem comparar les situacions individuals comptant amb les característiques de les seves llars, el context immediat on es desenvolupen els individus. Alhora, la consideració de les llars com a unitats d'anàlisi, possibilita l'estudi dels processos de canvi de les llars. Una altra informació que es pot obtenir és la referida als processos de mobilitat geogràfica i als processos de formació i dissolució de llars a Catalunya.

En aquest sentit és fonamental determinar què entenem per *llar* i la manera com la construïm. D'entrada, la llar té un fort component d'espai i podem definir-la com “l'agregació d'individus que conviuen plegats sota un mateix sostre”. Però des del PaD aquest no és l'únic criteri utilitzat. Si bé els elements objectius resulten fonamentals per no confondre *llar* amb *família*, els elements subjectius resulten igualment significatius quan interessa fer emergir les relacions familiars de la llar. Incloure aspectes subjectius en la construcció de la llar a més dels objectius (conviure sota un mateix sostre) implica que, de vegades, es considerin membres d'una llar persones que n'estan absents (per estudis, per treball, per absències temporals) i/o a considerar persones que conviuen sota un mateix sostre com a membres de llars diferenciades (rellogats dins d'un mateix habitatge o habitatges on conviuen membres amb economies separades, pisos d'estudiants, etc.). Tot i partir d'aquestes consideracions, a la pràctica, en el 95% dels casos la definició de la llar coincideix amb les persones que conviuen sota un mateix sostre i que hi estan presents en el període d'enquestació.

Finalment cal assenyalar que la informació de llar possibilita recollir les relacions de filiació i parentiu entre els membres que la componen, la qual cosa obre un ventall de possibilitats per a l'anàlisi de qüestions relacionades amb la família.

ASPECTES OBJECTIUS I SUBJECTIUS DE LA VIDA DELS INDIVIDUS

L'última particularitat té a veure amb la recollida d'aspectes subjectius de la vida dels enquestats. Les estadístiques oficials disponibles (tant respecte a Catalunya, com a Espanya i Europa) marquen la tendència a una expansió de les estadístiques sobre fets (dades objectivables), i fan poca atenció als aspectes subjectius dels individus. La contrapartida d'aquesta prioritització de les dades objectives genera una mancança recurrent de fonts de dades que s'aturin en els aspectes més subjectius de la vida dels individus, com ara les motivacions, percepcions, valoracions, sentiments i opinions.

En coherència amb 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.

2 El qüestionari

El qüestionari està pensat des de quatre perspectives diferents, que poden aparellar-se en dos eixos:

- perspectiva transversal i longitudinal
- perspectiva de llar i individual

El primer eix combina la perspectiva dinàmica amb l'estructural. D'una banda tenim l'anàlisi estàtica o transversal (de foto fixa) de la realitat social dels individus en un moment determinat de la seva història personal i de la història del conjunt del país. De l'altra, l'anàlisi longitudinal que permet el format panel i, per tant, el coneixement del canvi, especialment del canvi microsocial. Les implicacions d'aquest primer eix sobre el qüestionari tenen a veure amb els continguts de les enquestes que, onada rere onada, es van repetint amb periodicitats diferents segons el tipus de pregunta.

El segon eix d'anàlisi combina la realitat familiar amb la individual. Com ja s'ha dit, el PaD permet, d'una banda, l'anàlisi dels individus de la mostra i, de l'altra, l'anàlisi de les llars. És per això que un mateix tema pot aparèixer segmentat entre el qüestionari de llar i el d'individu per fer les preguntes adients a cada tipus de qüestionari.

A la pràctica, els continguts dels qüestionaris del PaD estan dividits en apartats que agrupen preguntes sobre un mateix tema o amb un mateix objectiu d'anàlisi. Els àmbits

principals sobre els quals se centra el PaD i dels quals se'n recull informació a cada onada són: dades sociodemogràfiques, educació, treball productiu, treball domèstic i familiar, salut, habitatge, llengua, posició i relacions socials, ingressos i cultura política. A més d'aquests grans blocs, podem trobar-ne d'altres, com per exemple mitjans de comunicació o temps de lleure, que apareixen de manera cíclica cada tres o quatre onades del panel.

PERSPECTIVA TRANSVERSAL *VERSUS* LONGITUDINAL

La combinació de les perspectives transversal i longitudinal ens obliga a assignar una periodicitat a cada una de les variables del qüestionari. Diferents experts i analistes, han orientat sobre la periodicitat idònia per a cada variable. Aquesta periodicitat és la que defineix cada quan apareix la variable; és a dir, en quines onades farem una pregunta determinada. La freqüència amb què apareixerà una variable al qüestionari depèn de la combinació de dos criteris:

1. El criteri clau a l'hora d'establir la periodicitat d'una variable és la seva susceptibilitat al canvi, tot i que no és l'únic. També cal tenir en compte la importància de la variable (central en el tema que s'estudia o més marginal), l'equilibri entre les variables objectives i subjectives en el conjunt del bloc, etc. Quant a la susceptibilitat al canvi, les variables amb un alt grau de possibilitats de canvi cal preguntar-les a cada onada o bé amb periodicitats curtes per assegurar la recollida d'aquest canvi. A les variables que proporcionen informacions menys susceptibles al canvi, se'ls dona una periodicitat més llarga.
2. El segon dels criteris observats per designar la periodicitat d'una variable és el necessari equilibri entre la periodicitat recomanada i les limitacions en el nombre màxim de variables que pot albergar cada qüestionari.

Com a resultat de l'aplicació dels dos criteris esmentats, totes les variables del PaD tenen assignada una periodicitat, que pot ser de quatre tipus diferents:

Quadre 1.**Temporalitat de les variables**

Troncals	Variables 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. Per exemple: • Està cursant estudis reglats actualment? quins? • Situació respecte el treball remunerat...
Troncals estructurals	Existeixen dos tipus de variables troncals estructurals: 1. Variables de nul·la variació en la vida de l'enquestat/da que 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... 2. Variables de poca variació en la vida de l'enquestat/da i que, per tant, no ens interessa recollir a cada onada. Aquest és el cas de variables com ara: l'edat de finalització dels estudis, els anys que fa que viu a l'habitatge, el nivell màxim d'estudis acabats...
Cícliques	Variables que apareixen de forma cíclica perquè s'estima que té sentit plantejar-se l'anàlisi del seu canvi només cada nombre determinat d'anys. S'indica C2, C3, C4... en funció de cada quantes onades ha d'aparèixer la variable al qüestionari. Per exemple: • Qui organitza i coordina el treball domèstic i familiar?(C2) • Quines coses milloraria del seu habitatge?(C4) • Quins dels següents familiars viuen a prop seu?(C8)
Esporàdiques	Variables 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-ne cap anàlisi en termes d'estudi del canvi. Per exemple: • Quina feina fa/feia el seu millor amic/ga? • Quin és el nivell màxim d'estudis acabats del seu millor amic/ga?

Font: Elaboració pròpia.

PERSPECTIVA DE LLAR *VERSUS* INDIVIDUAL

Una de les implicacions de fer una enquesta que vol resseguir la realitat dels individus i de les llars consisteix en la necessitat de tenir diversos qüestionaris. En concret n'hi ha tres: el qüestionari de llar, el qüestionari individual i el qüestionari d'aproximació (*proxy* d'ara endavant).

El **qüestionari de llar** el contesta una sola persona que es denomina informant principal. L'informant principal pot variar d'una onada a una altra, sempre i quan compleixi dues condicions: 1) estar suficientment qualificat per respondre el qüestionari de llar

(conèixer les dades de la llar: ingressos, habitatge, informació general sobre els seus membres...) 2) ser un membre original de la mostra, atès que són els individus que se segueixen fins al final del panel.

El **qüestionari individual** és el qüestionari més extens. El responen tots els membres de la llar de setze anys o més, a fi de recollir la informació personalitzada de cada membre de la llar. També és el qüestionari que permet recollir més informació respecte als aspectes subjectius de la vida dels individus.

La importància de disposar d'un mínim d'informació de cada un dels membres de la llar fa que sigui necessari un tercer tipus de qüestionari, el **qüestionari proxy**. El qüestionari *proxy* recull les preguntes objectives que es considera que poden ser contestades per algun altre membre de la llar, quan resulta del tot impossible enquestar el membre directament (perquè està absent de la llar durant el treball de camp, perquè es troba hospitalitzat o incapacitat per contestar l'enquesta, entre altres situacions possibles).

EL PES DELS BLOCS TEMÀTICS

Els àmbits temàtics centrals del PaD apareixen a cada onada, però amb un pes relatiu diferenciat. Aquesta variació en el nombre de variables directes finals és el resultat d'aplicar la periodicitat marcada a cada variable. És a dir, que en l'onada on apareguin les variables de treball domèstic de periodicitat cíclica 4 (que està indicant que cal que apareguin cada quatre anys), el nombre final de variables per al bloc serà molt més gran que en les onades anteriors o posteriors.

Hem d'assenyalar que a mesura que avancen les onades de l'enquesta, es recull un nombre de variables molt similar o superior al de les onades anteriors i, en canvi, els temps d'enquestació es mantenen o fins i tot es redueixen. Aquesta tendència es produeix degut a un major treball intern en l'elaboració dels salts i filtres del qüestionari i al fet d'assumir plenament, a partir de la quarta onada, la lògica longitudinal de l'enquesta que marca una periodicitat i una temporalitat per a cada variable.

Quadre 2.

Nombre de variables per onada del PaD i bloc temàtic*

Blocs temàtics del qüestionari	1a onada	2a onada	3a onada	4a onada
Dades demogràfiques	19	36	37	20
Dades dels pares	46	24	24	22
Situació de parella	3	9	8	5
Fecunditat	4	4	15	2
Educació	24	32	45	39
Treball productiu	179	144	63	119
Treball domèstic i familiar	27	19	62	
Conciliació**			19	18
Salut	11	8	20	34
Dependència**			12	
Habitatge	65	48	57	36
Llengua	20	1	1	50
Posició i relacions socials	34	34	63	9
Posició econòmica	76	138	149	190
Cultura política	5	5	4	14
Mitjans de comunicació			16	
Temps de lleure			19	
Immigració				
Nombre total de variables	513	502	614	558

* Només s'han comptat les variables pel primer membre de les graelles de llar per evitar l'efecte graella que cada onada afegeix més posicions a la llar i, per tant, més variables.

** Els blocs de conciliació i dependència no existeixen com a tals fins a la tercera onada, tot i que en les onades anteriors (primera i segona) es recull informació de forma desagregada als apartats d'educació, treball, i treball domèstic i salut.

Font: Elaboració pròpia.

PROGRAMA D'ENQUESTACIÓ CAPI

Una altra particularitat dels qüestionaris del PaD que mereix ser esmentada rau en el fet que les enquestes anuals es realitzen mitjançant el sistema d'enquestació CAPI (*Computer Assisted Personal Interview*). L'enquestació amb CAPI implica que els enquestadors realitzen les enquestes a les llars amb un ordinador portàtil i a partir d'un programa que digitalitza les respostes directament.

Els avantatges del sistema d'enquestació amb CAPI són múltiples en la mesura que aquest sistema permet minimitzar molts dels errors que són habituals amb el sistema d'enquestació PAPI (*Pen and Paper Personal Interview*):

1. Control sobre els temps d'enquestació: el temps d'administració dels qüestionaris es redueix considerablement, en comparació amb el que es triga a administrar un qüestionari en paper.
2. Control sobre les rutes d'enquestació: el programa d'enquestació té programada la ruta a seguir en funció de la resposta gravada a cada pregunta. Això suposa un gran avantatge pel que fa a la qualitat de les dades (minimització dels errors de ruta) i la feina dels enquestadors, ja que el control sobre la ruta la fa l'ordinador.
3. Minimització dels valors perduts: amb el CAPI es redueix el risc de deixar preguntes sense resposta com a resultat d'un salt o un filtre mal aplicat per l'enquestador.
4. Impossibilitat d'introduir valors inconsistents: mitjançant els controls, es minimitzen els errors produïts per la introducció de valors no permesos en una determinada pregunta.
5. Codificació automàtica de les respostes: pel fet de digitalitzar les dades directament, els errors que es deriven de codificar les respostes sobre el paper i la seva posterior informatització es veuen minimitzats.

3 El treball de camp, una aposta per la qualitat de les dades

Onada rere onada, el treball de camp del PaD concentra tots els seus esforços a enquestar totes les llars i tots els individus de setze anys o més que formen part de la mostra. Si bé aquest és l'objectiu final i indiscutible del treball de camp, el mètode definit per adreçar-se a les llars i als individus de la mostra no és menys important, en la mesura que respon a una sèrie de criteris que aspiren a garantir no només l'obtenció del nombre màxim d'enquestes, sinó també la màxima qualitat en les dades recollides.

PLANTEJAMENTS GENERALS I OBJECTIUS DEL TREBALL DE CAMP

La primera de les decisions que la Fundació Jaume Bofill va prendre a l'hora d'impulsar l'enquesta PaD, va ser la d'internalitzar el treball de camp, és a dir, d'assumir l'organització i el desenvolupament de totes les fases del treball de camp amb un equip propi i des de les seves dependències. Aquest ha estat un objectiu perseguit a consciència per tres motius principals:

1. Seguir de prop tot el procés de recollida de dades.
2. Disposar de la flexibilitat suficient per introduir els canvis que s'estimessin necessaris a cada moment.
3. Vetllar globalment per la qualitat de la recollida de la informació.

Per aconseguir aquests objectius es van prendre un seguit de decisions de gran transcendència metodològica:

- La participació de les llars és voluntària, es demana un compromís mínim de participació de quatre onades i no és remunerada, tot i que es lliura un petit obsequi per agrair la participació.
- La captació i el seguiment de la mostra són personalitzats a totes les onades; no es fa per via telefònica sinó mitjançant una visita personal de l'enquestador/a a la llar.
- Al llarg de tot el cicle del panel (anual) es manté un flux d'informació i comunicació amb les llars per mitjà de l'enviament d'informació sobre alguns dels resultats que es van obtenint de l'anàlisi.
- Al llarg del treball de camp es manté un contacte periòdic amb els diferents equips d'enquestadors (la formació inicial es complementa amb reunions de coordinació mensuals).

LA FORMACIÓ DELS ENQUESTADORS, ELEMENT CLAU PER A LA QUALITAT

Des del PaD sempre s'ha tingut clar el paper clau que juga la formació dels enquestadors en pro d'optimitzar la qualitat en el procés de recollida de dades. En aquest sentit, s'esmercen molts esforços en la formació inicial de l'equip d'enquestació i es proporcionem materials de suport per resoldre els dubtes de l'enquestació. A la formació inicial es persegueix accomplir cinc objectius principals:

1. Transmetre la filosofia del projecte i de les implicacions metodològiques de treballar en una enquesta longitudinal.
2. Assegurar una bona comprensió del contingut del qüestionari i de totes les seves preguntes.
3. Conèixer les eines informàtiques amb les quals es treballa (bases de dades, programa d'enquestació CAPI).
4. Donar un mateix punt de partida sobre la tècnica d'enquestació a persones amb formacions i experiències diverses.

5. Conèixer l'equip, la figura del coordinador d'equip provincial, i la metodologia de coordinació del treball de camp i de gestió, seguiment i control de les dades.

Probablement aquest últim objectiu sigui el més important: seguir dels equips provincials al llarg de tot el període d'enquestació. Les reunions de coordinació són espais per ampliar la formació, fer un retorn dels principals errors d'enquestació detectats als controls de qualitat realitzats i un espai privilegiat per a la motivació i la fidelització dels enquestadors amb el projecte.

En els intervals de temps que separen les reunions de coordinació, els enquestadors/es disposen de la figura del coordinador d'equip amb la funció d'atendre dubtes, solucionar situacions d'enquestació i vehicular tota la informació rellevant cap a l'equip general del PaD.

EL PAPER DE L'ENQUESTADOR

El paper de l'enquestador és un punt crucial en la recollida de les dades. Sobretot si tenim en compte que el perfil general d'enquestadors i enquestadores del PaD respon al d'un recent llicenciat en ciències socials i amb una motivació per participar en el projecte que, moltes vegades, transcendeix la finalitat econòmica (guanyar experiència en el camp de la sociologia, treballar en un projecte amb vocació social, conèixer la Fundació Jaume Bofill, entre altres). Si bé aquest perfil té certs avantatges sobre el perfil professionalitzat, també pot ser una font de biaix important, i cal fer un treball específic per traslladar les regles bàsiques de l'enquestació:

1. L'enquestador ha de respectar sempre la informació que dona l'enquestat malgrat que aquesta sembli poc versemblant o estranya.
2. L'enquestador no ha de mostrar la seva opinió davant de la persona enquestada.
3. L'enquestador s'ha de posar al nivell de la persona enquestada per crear un clima de confiança.

ELS CONTROLS DE QUALITAT DE L'ENQUESTACIÓ

Des d'un bon inici, al PaD es plantegen dos tipus de controls de qualitat:

1. El control de qualitat de l'enquestació: control sobre la tasca feta per part dels enquestadors que es fa des de l'equip central del PaD.
2. El control telefònic al 20% de la mostra enquestada: control per assegurar que totes les enquestes s'han fet de forma presencial.

El **control de qualitat de l'enquestació** constitueix el pas previ a la depuració de les dades. La finalitat del control de qualitat és la de detectar, individualment i de manera simultània al treball de camp, els principals errors d'enquestació i comunicar-los als enquestadors per tal que siguin corregits, però també per tal d'evitar que es tornin a cometre. La ràpida interacció amb l'enquestador és un dels punts clau d'aquest procés, ja que interessa detectar com més aviat millor els errors d'enquestació. Els beneficis principals derivats d'aquest tipus de control són:

1. Detectar i possibilitar la correcció dels errors principals de cada enquestador (això és possible gràcies a la simultaneïtat d'aquest control amb la realització del treball de camp), amb la consegüent millora de la qualitat de les dades.
2. Interactuar amb els enquestadors per resoldre els dubtes que no s'han pogut plantejar a les reunions de coordinació.
3. Aportar millores als continguts dels qüestionaris, quan es detecta un error que es reproduceix de manera sistemàtica (per exemple, afegint un comentari al costat d'una pregunta o una nota per l'enquestador...)
4. Disminuir una quantitat important dels errors més freqüents.

El **control telefònic de les llars** per assegurar que totes les enquestes s'han fet de forma presencial es realitza a un 20% de les llars de cada enquestador. Els objectius del control telefònic són:

1. Verificar la fiabilitat de les dades recollides.
2. Aproximar-se a la llar i saber quin és el grau de satisfacció o insatisfacció respecte al tracte rebut per l'enquestador.

3. Conèixer l'opinió de la llar respecte als continguts de l'enquesta.
4. Contrastar algunes de les dades clau recollides durant el treball de camp: número de membres de la llar i situació respecte el treball.
5. Advertir els canvis d'adreça que hi ha hagut a la llar des que l'enquestador hi ha anat.

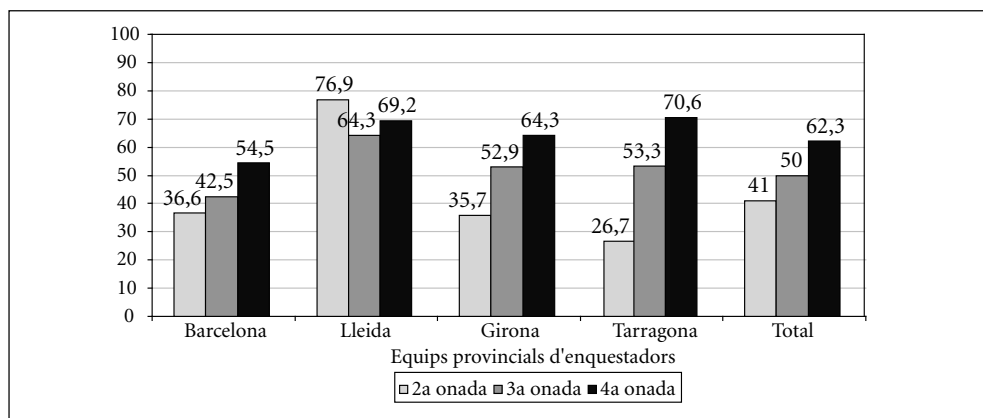
Val a dir que, després de quatre anys de realització del PaD i de més de 23.000² enquestes realitzades, només s'han detectat problemes de caràcter menor i una vintena d'enquestes mal realitzades (0,07% del conjunt global) que han estat eliminades de les matrius.

LA CONTINUÏTAT DELS ENQUESTADORS DEL PAD

La fidelització dels enquestadors del PaD també és un element important per a la qualitat de les dades. Afortunadament, al llarg de les onades del PaD s'ha aconseguit

Gràfic 1.

Nivell de continuïtat dels enquestadors/es del PaD, per onades



Font: Elaboració pròpia a partir de les dades del PaD. Segona onada, 2003; tercera onada, 2004, i quarta onada, 2005.

-
2. En finalitzar la quarta onada, hi havia 23.457 enquestes realitzades al llarg de la vida del PaD.

incrementar la continuïtat del gruix de l'equip, tot i que el nivell de rotació continua sent important. Al gràfic següent es pot observar quin ha estat, al llarg de les onades, el grau de continuïtat dels diferents equips d'enquestadors:

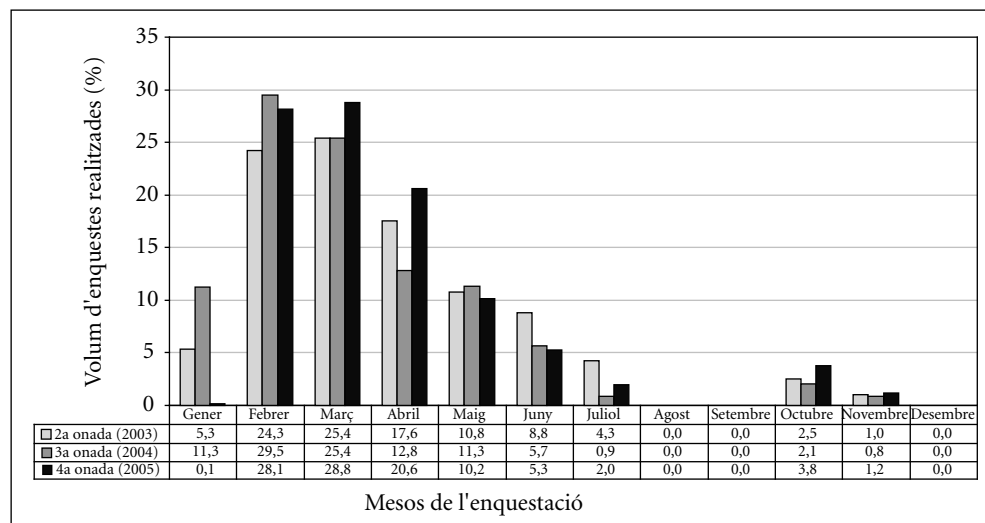
Possiblement el perfil d'enquestadors del PaD (majoritàriament llicenciats recents que estan en procés d'inserció al mercat de treball) estigui a l'origen de l'alt grau de rotació que es pot observar al gràfic. La simultaneïtat del treball com a enquestadors amb altres feines impossibilita, en molts casos, la seva continuïtat al PaD.

EL CALENDARI D'ENQUESTACIÓ

El calendari d'enquestació és un altre factor clau per garantir la qualitat de les dades. En aquest sentit, si es fa un repàs als temps d'enquestació de les diferents onades del PaD, s'obtenen els resultats següents:

Gràfic 2.

Calendaris d'enquestació, per onades del PaD



Font: Elaboració pròpia a partir de les dades del PaD. Segona onada, 2003; tercera onada, 2004, i quarta onada, 2005.

El gràfic no recull el calendari de la primera onada perquè va ser molt diferent del de la resta d'onades. La primera onada es va allargar des de finals de setembre del 2001 fins a principis de juny del 2002.³ En les subsegüents onades, s'ha treballat per compactar més els períodes d'enquestació tot i que ha estat molt difícil, i la durada ha estat des de finals de gener fins a principis de juliol. Cal atribuir les dificultats principals als problemes per localitzar enquestadors en determinades zones de Catalunya i a les incidències amb els equips informàtics, que impedièren treballar en algunes zones durant un cert temps.

En part per això i també per minimitzar les pèrdues mostrals, s'han impulsat segones etapes del treball de camp⁴ —molt compactades i localitzades en el temps— per tal de recuperar llars que, per problemes de calendari o circumstàncies diverses, no ens podien atendre entre els mesos de febrer i juliol.

3. La llargada d'aquest període s'explica perquè va ser l'any de captació de la mostra.

4. Les segones etapes del treball de camp, es fan amb un equip molt limitat d'enquestadors que gestionen un volum petit i seleccionat de llars que per diferents circumstàncies de l'enquestació no han pogut ser enquestades abans del mes juliol. La segona etapa neix en la segona onada i es planteja a la tardor, un cop acabat el gruix del treball de camp.

4 La mostra

EL DISSENY DE LA MOSTRA I ELS SEUS RESULTATS FINALS⁵

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), que és **aproporcional** pel que fa a les **províncies**.

Per tal de garantir una bona representativitat territorial es van prendre dues decisions importants:

1. Dissenyar una mostra representativa no només del conjunt de Catalunya sinó també de les quatre províncies.
2. Utilitzar les seccions censals com a primeres unitats mostrals a partir de les quals seleccionar les llars que formen part de la mostra original.

La primera decisió s'explica com a conseqüència de la concentració poblacional a la regió metropolitana de Barcelona, on habiten més de tres quartes parts de les llars

5. El procés d'elaboració de la mostra es descriu íntegrament al *Manual d'ús*, disponible a www.obdesigualtats.cat.

catalanes, i que ha provocat que, massa sovint, els estudis no hagin pogut recollir informació prou significativa de les persones que viuen a la resta del país.

La segona decisió presa, la d'utilitzar les seccions censals com a primeres unitats mostrals, es justifica pel fet que aquestes seccions són les unitats territorials estadístiques més petites d'ús públic. Com que són agrupacions territorials amb mitjanes d'entre 1.000 i 1.200 persones, o aproximadament 400 llars, les seccions censals proporcionen una gran coherència interna i la informació disponible permet fer-ne classificacions d'estructura social i territorial diferents.

En la construcció de la mostra es parteix del $\pm 5\%$ com a marge d'error mostral màxim assumible per a resultats globals (amb un interval de confiança del 95,5%, amb $k=2$, i en el supòsit de màxima indeterminació $p=q=50$). Consegüentment, cal observar com a mínim 400 llars a cada una de les 4 províncies catalanes. Amb tot, la diversitat i complexitat de la província de Barcelona demana d'una mostra específica 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 les comarques centrals (69). És així com s'arriba a la mida final de la mostra que es compona de 2.000 llars distribuïdes aproporcionalment a nivell provincial.

Taula 1.
Distribució territorial del nombre de llars de la mostra original

Àrees	Nombre de llars	e* (%)
Barcelona	800	$\pm 3,54$
Barcelona ciutat	261	$\pm 6,12$
Regió metropolitana	470	$\pm 4,61$
Comarques centrals	69	$\pm 12,04$
Girona	400	$\pm 5,00$
Lleida	400	$\pm 5,00$
Tarragona	400	$\pm 5,00$
Total	2.000	$\pm 2,78$

* e = error mostral (interval confiança=95,5%, $k=2$ i $p=q=50$)

Font: Elaboració pròpia.

Per tal de garantir una bona representativitat mostral, les variables d'estratificació són territorials i socials: províncies (incloent una subdivisió en el cas de Barcelona), àmbit urbà-rural (grandària del municipi), estructura social (classe treballadora) i situació ocupacional (actius/inactius).

A partir d'un llistat estratificat i jerarquitzat de totes les seccions de Catalunya, des d'un inici aleatori en cada estrat (menor que el període) i amb un interval fix (període) es van seleccionar 203 unitats primàries de la mostra o seccions censals dins de 36 grups. Com a resultat d'aquest procés cal captar llars en 115 municipis diferents dels 946 del país.

Abans de seleccionar les llars concretes que cal entrevistar, es va fer una comprovació sobre la bondat de la mostra de seccions censals comparant totes les variables agregades disponibles de les 203 seccions elegides i el total de la població (5.186 seccions) així com de les mostres en els quatre territoris (81 de 3.833 a Barcelona, 41 de 478 a Girona, 41 de 384 a Lleida i 40 de 491 a Tarragona). Les variables utilitzades són: estructura de sexe i edat, origen, nacionalitat, estat civil, nivell educatiu, anys residència a Catalunya, coneixement del català, ocupació, activitat, posició social, parentiu, nombre de persones per llar, tipologia de la llar i grandària del municipi de residència. D'aquesta comprovació no es va detectar cap desviació significativa.

Durant el treball de camp de la primera onada es van fer dues comprovacions de la bondat de la mostra de llars i individus. Amb l'11% i el 50% del treball de camp realitzat no es van observar biaixos importants que demanessin rectificacions ni a nivell de llar (nombre de persones per llar i tipologia de llar) ni a nivell individual (estructura de sexe, edat i ocupació). Tot i que es va fer una tercera i última comprovació en el tancament del treball de camp (amb 1.991 llars) i amb resultats igualment satisfactoris, és evident que per tal d'ajustar correctament els resultats obtinguts cal introduir reequilibratges i utilitzar variables de ponderació en les matrius de dades finals.

LES VARIABLES DE PONDERACIÓ

La ponderació de la mostra és un pas fonamental que respon a la necessitat d’ajustar la població apropiacional de la mostra, a la distribució real de la població a Catalunya.

Les variables de ponderació no tenen una construcció predeterminada sinó que es fan a la mida de cada enquesta. En el cas del PaD es construeixen les variables de ponderació mitjançant un sistema de quotes creuades a partir de les distribucions segons el sexe, l’edat, l’origen geogràfic, la província de residència i la grandària del municipi de residència del conjunt de la població de Catalunya. D’aquesta manera, l’aplicació de les variables de ponderació sobre les dades originals permet corregir els biaixos o desviacions de la mostra a nivell d’estructura de la població.

La decisió de construir variables de ponderació des del PaD i posar-les a disposició dels analistes respon a dues finalitats: d’una banda, generar economies d’escala als analistes i, de l’altra, alertar d’una ponderació indispensable, sempre i quan no es treballin les dades a nivell provincial, que és la ponderació territorial (província de residència i grandària del municipi de residència) degut al fet que el punt de partida és una distribució de llars i individus apropiacional pel que fa al territori.

Taula 2.

Distribució apropiacional de la mostra del PaD *versus* distribució real de la població de Catalunya

Distribució apropiacional	Distribució real*
Barcelona: 791 llars (40% de la mostra)	Barcelona: 75,77% de la població
Girona: 400 llars (20% de la mostra)	Girona: 8,91% de la població
Lleida: 400 llars (20% de la mostra)	Lleida: 5,71% de la població
Tarragona: 400 llars (20% de la mostra)	Tarragona: 9,61% de la població
TOTAL: 1.991 llars (100% de la mostra)	TOTAL: 100% població de Catalunya

Font: Elaboració pròpia amb dades del PaD, primera onada i l’Idescat, 2001: Distribució de la població per províncies a Catalunya.

La mostra aproporcional implica que, abans d'abordar qualsevol anàlisi de l'àmbit català, cal ajustar el pes de les dades de les províncies de Girona, Lleida i Tarragona al pes que tenen realment en el conjunt de Catalunya. Aquest procés queda garantit mitjançant l'aplicació de les variables de ponderació.

Les variables de ponderació són de dos tipus:

1. Les variables de **ponderació transversal** (construïdes des de la primera onada) són tres: una a nivell de llar i dues a nivell individual (una pels individus enquestats i una altra per tots els individus de la mostra, amb independència de si han estat enquestats o no).
2. Les variables de **ponderació longitudinal** (construïdes a partir de la segona onada) són dues de nivell individual⁶ (una per aquells individus enquestats a totes les onades i una altra per a tots els individus de la mostra, independentment de si han estat enquestats o no).⁷

El càlcul de la variable de ponderació transversal a nivell de llars segueix un sistema de càlcul mitjançant quotes creuades. El càlcul es fa combinant quatre factors: província, grandària del municipi, tipus de llar i nombre de membres de la llar.⁸ Per saber els

.....

6. Una qüestió important és decidir si ponderar els individus o ponderar les llars. Al PaD s'ha optat per construir la variable de ponderació a nivell d'individus perquè tant la desaparició com la creació de noves llars són una conseqüència de la mateixa dinàmica de les llars que, en la mesura que són unitats de convivència vives, van canviant. En canvi, els individus són sempre els mateixos. Poden viure en llars diferents, però en si mateixos, no canvien, només canvia la unitat de convivència en què viuen. Per tant, els canvis de llars són un tema d'estudi més, entre les moltes anàlisis que es poden fer des d'un punt de vista longitudinal.

7. Els individus enquestats amb un qüestionari *proxy* no tenen valor de ponderació a les variables de ponderació transversal i longitudinal on només es considera els individus enquestats, però sí que tenen valor de ponderació a les variables de ponderació en què es considera a tots els individus de la mostra amb independència de si han estat enquestats o no.

8. Les categories considerades a cada factor de ponderació són: província (Barcelona, Lleida, Girona, Tarragona), grandària del municipi (menys de 2.000 habitants, de 2.000 a 9.999 habitants, de 10.000 a 49.999 habitants, de 50.000 a 99.999 habitants, de 100.000 a 499.999 habitants, i 1.000.000 d'habitants o més), tipus de llar (llars unipersonals, més d'una persona sense nucli, parella sola, parella amb fills, parella

valors de ponderació, es calcula el percentatge que cada grup poblacional representa respecte el conjunt de la població de Catalunya (amb dades del cens del 2001).

Les dues variables de ponderació transversal a nivell individual es construeixen a partir dels valors de l'estructura de la població de corresponents al cens de 2001 i mitjançant el sistema d'introducció de pesos per quotes creuades. És a dir, que s'introdueixen valors (pesos) de ponderació combinant cinc factors: sexe, edat, província, origen geogràfic i grandària del municipi.⁹

L'ajustament longitudinal és necessari per resoldre una problemàtica derivada de la naturalesa del panel: és molt probable que la mostra inicial hagi perdut representativitat com a conseqüència de la desaparició d'individus, fruit de la negativa de l'individu a continuar respondent i/o de l'esgotament mostral. Per als càlculs de les variables de ponderació longitudinals només són d'interès els individus de la mostra que responen a totes les onades. Per tant, només aquests són els casos ponderats per donar compte de la pèrdua mostral a l'onada anterior. D'altra banda, tampoc s'ha d'aplicar cap ajust per les pèrdues d'aquells individus que se sap que s'han tornat inelegibles (és a dir, que ja no formen part de la mostra: com per exemple els individus que han mort, els que han desaparegut, els que han estat institucionalitzats en centres d'internament, geriàtrics, etc. o els que han fixat la seva residència fora de Catalunya), però sí que cal ponderar les pèrdues dels individus elegibles que, per algun motiu, no responen l'enquesta en l'onada corresponent.

Les variables de ponderació longitudinal són el resultat de multiplicar la ponderació transversal de l'onada anterior per la probabilitat de ser encara elegible en l'onada en

amb altres persones, parella amb fills i altres persones, mare sola amb fills, mare sola amb fills i altres persones, pare sol amb fills, pare sol amb fills i altres persones, dos nuclis o més) i nombre de membres de la llar (1, 2, 3, 4, 5, 6, 7, 8, 9, 10 o més membres).

9. Les categories considerades a cada factor de ponderació són: sexe (home, dona), edat (interval 0-15 anys, 16-24 anys, 25-34 anys, 35-49 anys, 50-64 anys, 65 i més anys), província (Barcelona, Lleida, Girona, Tarragona), origen geogràfic (nascuts a Catalunya, nascuts a Espanya i nascuts en un altre país) i grandària del municipi (menys de 2.000 habitants, de 2.000 a 9.999 habitants, de 10.000 a 49.999 habitants, de 50.000 a 99.999 habitants, de 100.000 a 499.999 habitants, 1.000.000 d'habitants o més).

curs. Alhora, la probabilitat de ser elegible s'estima utilitzant una regressió logística, on la variable dependent pren el valor 0 si és inelegible i 1 si és elegible; i les variables explicatives que es prenen com a referència són: el canvi d'adreça, el sexe, l'edat, l'estat civil, el lloc de naixement, les discapacitats, el nivell d'estudis, la situació laboral i la classe social (subjectiva).¹⁰

L'ús apropiat de les diferents variables de ponderació que es proposen des del PaD, no es pot definir a priori sinó que dependrà de cada anàlisi plantejada.

L'EVOLUCIÓ DE LA MOSTRA AL LLARG DE LES ONADES

Les llars de la mostra del PaD són unitats vives de convivència en les quals es produeixen canvis que impliquen la desaparició de llars completes, la creació de llars noves o, a nivell individual, la desaparició de membres (baixes) i l'aparició de membres nous (altes). És fonamental reflectir tots aquests canvis a l'estructura de les llars per poder-ne fer un seguiment correcte. Òbviament, aquests canvis interns també impliquen canvis en les xifres totals de la mostra, tant si s'agafa com a unitat d'anàlisi la llar, com si s'agafa a l'individu.

A part dels canvis que es produeixen en el si de les llars, hi ha altres tipus de pèrdues mostrals que també cal comptabilitzar. Se'n poden distingir dos tipus:

1. El primer tipus de pèrdues mostrals són les “connaturals” a les enquestes longitudinals i no depenen de l'eficàcia del treball de camp. Es tracta de les pèrdues de llars o d'individus “inevitables” per l'efecte de la mort, institucionalització, desaparició i/o trasllat fora del territori objecte d'enquestació (Catalunya en el cas del PaD). Aquest tipus de pèrdua pot afectar llars senceres o individus concrets que pertanyen a llars que continuen existint. Per norma metodològica de les enquestes panel, quan es donen aquest tipus de circumstàncies, no cal enquestar els

10. Les variables explicatives finalment escollides en la regressió dependran de la significació estadística que tenen en l'onada que correspongui i s'utilitzen només aquelles que són significatives ($p \leq 0,05$).

individus (o les llars, si les circumstàncies de pèrdua es donen en tots els individus que conformen la llar).

2. El segon tipus de pèrdua sí que és atribuïble a l'eficàcia del projecte, tot i que són pèrdues molt difícils de controlar. Es tracta de les pèrdues mostrals “evitables”, que alhora es poden agrupar en dos grans grups: les pèrdues d'individus o llars senceres que es neguen a continuar col·laborant malgrat el compromís inicial; i les pèrdues mostrals fruit d'individus o llars que no s'aconsegueix localitzar en el territori perquè han canviat d'adreça i no ho han comunicat a l'equip del treball de camp.

Dit això, el primer exercici que cal fer a l'hora d'analitzar l'evolució de la mostra és distingir entre la **mostra total d'una onada**, que és el nombre total d'individus o llars abans de considerar cap de les pèrdues mostrals esmentades, i la **mostra total elegible**, que és el resultat de restar, de la mostra total de l'onada, les pèrdues mostrals de caràcter “inevitable” i que per normativa de panel, no cal seguir enquestant ni comptabilitzar en l'anàlisi.

L'evolució de la mostra del PaD es fa per partida doble: per una banda, es fa un seguiment de l'evolució de la mostra agafant les llars com a unitat d'estudi i, per una altra, es fa un anàlisi de l'evolució de la mostra agafant l'individu com a unitat d'estudi.

L'EVOLUCIÓ DE LA MOSTRA DE LLARS DEL PAD

La **mostra total de llars del PaD** la conformen dos grups de llars diferenciades: el de les llars inelegibles i el de les llars elegibles.

El grup de **llars inelegibles** de la mostra el conformen les llars que es perden de la mostra a cada onada per motius “inevitables” a saber:

- 1) les llars en què han mort tots els seus membres
- 2) les llars en què s'han institucionalitzat tots els seus membres (geriàtrics, centres d'internament...)
- 3) les llars que s'han desfet perquè tots els seus membres s'han integrat en altres llars del panel

4) les llars que s’han desplaçat fora de Catalunya i que per normativa panel no hem de seguir enquestant

Tal i com mostra la taula 3, aquestes llars representen un percentatge molt petit de la mostra total de llars i, si bé la pèrdua mostral de llars de caràcter “inevitable” és una constant, hi ha pocs canvis al llarg de les onades:

Taula 3.
Evolució de la mostra total de llars del PaD¹¹

		2a onada		3a onada		4a onada	
Mostra total de llars		2.149	100%	2.044	100%	1.880	100%
Llars Inelegibles	Mort de tots els membres	7	0,33%	6	0,29%	6	0,32%
	Institucionalització de tots els membres	1	0,05%	14	0,68%	3	0,6%
	Llar marxa fora de Catalunya	22	1,02%	5	0,24%	19	1,01%
	Incorporació de tots els membres en una altra llar del PaD	0	0%	3	0,15%	4	0,21%
Mostra de llars elegibles*		2119	98,6%	2016	98,6%	1848	98,3%

* El nombre de llars elegibles s’obté restant de la mostra total de llars el grup de llars inelegibles.
Font: Elaboració pròpia a partir de les dades del PaD. Segona onada, 2003, tercera onada, 2004 i quarta onada 2005.

El grup de **llars elegibles** el formen les llars susceptibles de ser enquestades a cada onada, i són les resultants de restar el grup de llars inelegible de la mostra total.

Un cop determinada quina és la mostra de llars elegible a cada onada, interessa veure quins són els resultats d’enquestació, és a dir, a quines llars s’ha pogut entrevistar i a quines llars no s’ha pogut entrevistar per motius de pèrdua o rebuig. En les enquestes longitudinals, la no enquestació de llars per motius pèrdua o de rebuig són la causa més important d’erosió de la mostra.

11. L’anàlisi comença a partir de la segona onada, atès que la primera onada és la de construcció de la mostra i no s’experimenten pèrdues mostrals.

Si agrupem, dins de les llars elegibles, les llars enquestades i les no enquestades, distingim els quatre subgrups següents:

Llars enquestades:

1. Llars enquestades que formen part de la mostra original captada.
2. Llars enquestades que s’han creat al llarg de l’onada, fruit dels canvis que es produeixen en el sí de les llars (per independització d’algun dels seus membres, separació de la parella, etc.).

Llars no enquestades:

3. Llars originals o de nova creació que no s’han enquestat i que es consideren com a pèrdua o rebuig definitiu del PaD (llars irrecuperables).
4. Llars originals o de nova creació no enquestades i que es consideren com a pèrdua o rebuig no definitiu per diferents motius (no s’han pogut cobrir en el treball de camp per falta d’enquestadors; no participen a l’onada en curs però demostren bona predisposició per ser enquestades a l’onada següent; o els motius de pèrdua o rebuig al·legats no són prou contundents i des del PaD es vol fer un treball de recuperació de la llar). Tot i ser un grup percentualment petit, són llars que es tornen a mirar d’enquestar a l’onada següent i que en un nombre important són recuperades.

Taula 4.

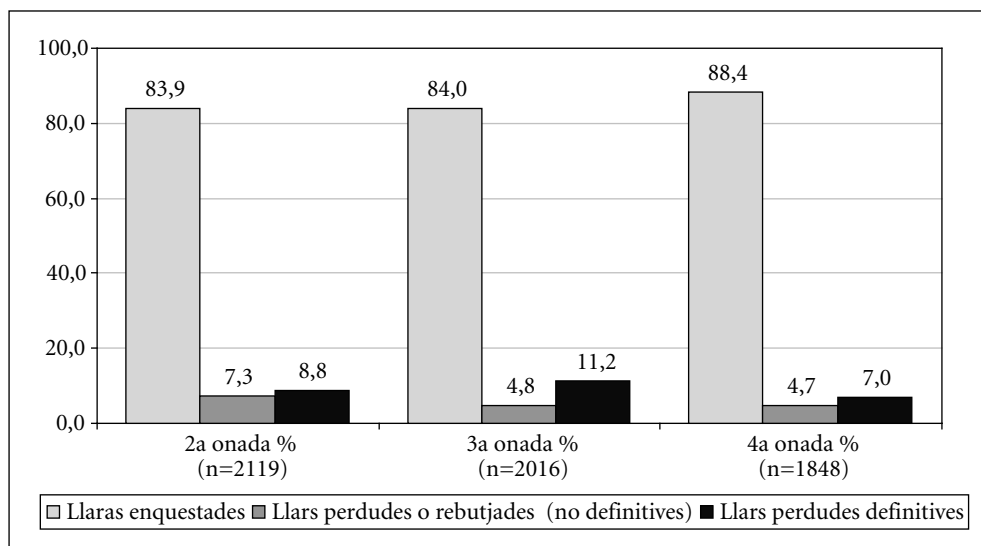
Grau d’enquestació de les llars elegibles del PaD

		2a onada		3a onada		4a onada	
Llars enquestades	Llars que ja existien en altres onades	1703	80,3%	1624	80,5%	1576	85,2%
	Llars de nova creació a l’onada	74	3,35%	69	3,4%	57	3,08%
Llars no enquestades	Llars perdudes o rebutjades no definitives	153	7,2%	97	4,8%	86	4,7%
	Llars perdudes/ rebutjades definitives	189	9%	226	11,2%	129	7%
Total mostra llars elegibles		2119	100%	2016	100%	1848	100%

Font: Elaboració pròpia a partir de les dades del PaD. Segona onada, 2003, tercera onada, 2004, i quarta onada 2005.

Gràfic 3.

Grau d'enquestació de les llars elegibles del PaD



Font: Elaboració pròpia a partir de les dades del PaD. Segona onada, 2003, tercera onada, 2004, i quarta onada 2005.

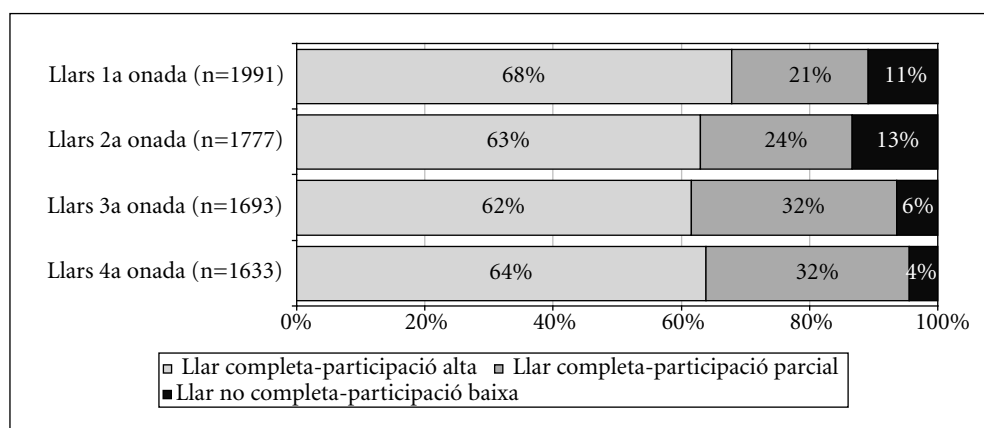
A més de distingir entre les llars enquestades i les llars no enquestades cal veure quin **nivell d'enquestació interna** s'ha obtingut en les llars enquestades. Tot i els esforços que s'hi destinen, no sempre és possible enquestar tots els individus de setze anys i més de les llars (que és la norma que marca el panel). Per veure quin és el grau d'enquestació, les llars s'agrupen en tres categories:

1. Llaras completes amb participació de tots els membres: són les llars on s'ha pogut enquestar tots els membres majors de setze anys amb una enquesta individual. Són llars amb un nivell alt d'enquestació interna.
2. Llaras completes amb participació parcial: són les llars on s'han entrevistat alguns membres amb la seva enquesta individual i altres mitjançant una enquesta *proxy* (enquesta que algú de la llar contesta en representació d'un membre absent o que rebutja participar). Són llars amb un nivell intermedi d'enquestació interna.

3. Llars no completes amb participació baixa: són les llars on no s'ha pogut obtenir informació individual de tots els membres de setze anys o més. En aquestes llars es combinen les enquestes individuals, amb les enquestes *proxy* i amb membres no entrevistats. Són llars amb un nivell baix d'enquestació interna.

Gràfic 4.

Grau d'enquestació interna de les llars del PaD



Font: Elaboració pròpia a partir de les dades del PaD. Primera onada 2001-2002; segona onada, 2003, tercera onada, 2004, i quarta onada 2005.

Si s'observa el gràfic es pot concloure que l'estabilitat és el tret més característic pel que fa als diferents nivells d'enquestació. A la primera onada la condició necessària per participar era obtenir la participació de tots els membres de la llar, la qual cosa explica el nivell més elevat d'enquestació interna de les llars. A partir de la segona onada, les llars que formen part de la mostra no es poden canviar i s'han d'anar assumint els individus que rebutgen participar a títol individual malgrat la continuïtat de la llar. Aquest fet explica, en part, que el nombre d'enquestes *proxy* s'hagi anat incrementant a cada onada. D'altra banda, fruit de la fidelització d'alguns dels membres de la llar amb el PaD, també s'aconsegueix que algunes llars passin del grup de llars amb participació baixa a pertànyer al grup de llars amb participació intermèdia, substituint els individus no entrevistats per enquestes *proxy*.

L'EVOLUCIÓ DE LA MOSTRA D'INDIVIDUS DEL PAD

La **mostra total d'individus del PaD** la conformen la suma de dos grups d'individus diferenciats: els individus inelegibles i els individus elegibles.

El grup d'**individus inelegibles** de la mostra el conformen els individus que es perden de la mostra a cada onada per motius “inevitables” i de forma definitiva. Es tracta, per tant, d'individus que no podran ser enquestats ni en l'onada en curs ni en les posteriors. Aquests són:

1. Els individus que s'han mort, han desaparegut o han estat institucionalitzats.
2. Els individus que s'han desplaçat fora del territori català.
3. Els individus temporals¹² que causen baixa d'una llar i deixen de conviure amb un membre original de la mostra.

A aquests tres grups d'individus que conformen la mostra inelegible cal sumar-hi un quart grup (el més important numèricament parlant) que per la seva particularitat cal comptabilitzar a part dels anteriorment descrits:

4. Els individus menors de setze anys en l'onada d'enquestació.

El grup d'individus menors de setze anys no són enquestables en l'onada en curs i s'han de comptabilitzar com a individus inelegibles, però es tracta d'una inelegibilitat “temporal”. Aquest és l'únic cas d'inelegibilitat no definitiva, ja que l'any en què es compleixen els setze anys, l'individu passa a ser enquestable. Per tant, en l'anàlisi d'onades, és un grup comptabilitzat sempre com a mostra individual inelegible, però

12. Els individus temporals són aquells que s'han donat d'alta a la mostra del PaD amb posterioritat a l'onada de captació. A l'onada de captació (primera onada) tots els individus que s'incorporen a la mostra són considerats membres originals de la mostra i cal seguir-los al llarg de tota la vida del panel. En canvi, els que es donen d'alta a les llars de la mostra a partir de la segona onada del panel se'ls considera temporals (amb l'única excepció de si són fills nascuts de membres originals de la mostra, que és l'únic cas on se'ls considerarà també com membres originals de la mostra). Si un membre temporal deixa de formar part d'una llar amb membres originals del PaD, no caldrà seguir-lo enquestant. És per això que passa a formar part de la mostra d'individus inelegible.

que com a grup d'individus no desapareix de la mostra amb el canvi d'onada sinó que s'arrossega onada rere onada. Cada any alguns dels seus individus (els que en l'any de l'onada compleixen setze anys) són traspassats a la mostra individual elegible.

El volum total que representen els individus inelegibles a cada onada del PaD es pot observar a la taula següent:

Taula 5.
Evolució de la mostra total individual del PaD

		1a onada		2a onada		3a onada		4a onada	
Mostra total individual		5785	100,0%	6040	100,0%	5743	100,0%	5254	100,0%
Mostra individual inelegible	Individus menors de setze anys	872	15,1%	853	14,1%	805	14,0%	757	14,4%
	Individus de setze o més anys*	0***	0,0%	96	1,6%	103	1,6%	126	2,4%
Mostra individual elegible**		4913	84,9%	5091	84,3%	4844	84,3%	4373	83,2%

*A la categoria individus inelegibles majors de setze anys agrupem els individus morts, institucionalitzats o desapareguts, els individus que s'han desplaçat fora del territori català i els individus temporals que causen baixa d'una llar i deixen de conviure amb un membre original de la mostra.

** El nombre d'individus elegibles s'obté restant de la mostra total d'individus els individus inelegibles.

***A la primera onada aquest grup és zero perquè és l'onada de construcció de la mostra i, per tant, no s'experimenten pèrdues individuals definitives.

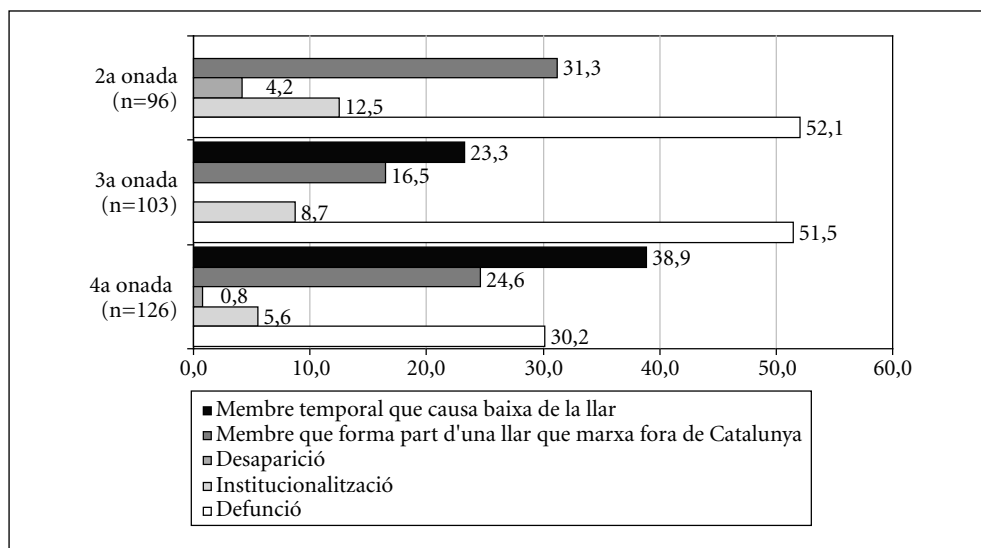
Font: Elaboració pròpia a partir de les dades del PaD. Segona onada, 2003, tercera onada, 2004, i quarta onada 2005.

Una breu mirada a la mostra individual inelegible d'individus de setze anys o més ens permet veure el pes de cada subgrup d'aquesta categoria que a la taula anterior s'ha comptabilitzat de forma agregada (vegeu el gràfic 5):

El grup d'**individus elegibles** està format per tots els individus que són susceptibles de ser enquestats durant l'onada. Per tant, s'obté de restar de la mostra total individual, el grup d'individus inelegibles.

Gràfic 5.

Evolució de la mostra individual inelegible del PaD*



*La desproporció entre el nombre d'individus inelegibles menors de setze i el de setze o més anys, fa preferible descartar el grup de menors de setze anys del gràfic.

Font: Elaboració pròpia a partir de les dades del PaD. Segona onada, 2003, tercera onada, 2004, i quarta onada 2005.

Una vegada determinada quina és la mostra individual elegible interessa veure els **resultats d'enquestació** d'aquesta mostra d'individus. Aquests resultats s'agrupen en els individus enquestats i els individus no enquestats; i dins d'aquests dos grups també cal fer una nova distinció:

Individus enquestats:

1. Individus enquestats a nivell individual.
2. Individus no enquestats a títol individual sinó a través d'una enquesta *proxy*.

Individus no enquestats:

3. Individus que formen part d'una llar que ha participat però que no han estat entrevistats ni disposem de la seva informació objectiva a través d'una enquesta *proxy*.

- Individus que formen part d'una llar que no ha participat a l'onada i, per tant, no disposem d'informació de cap dels individus que la conformen, però que s'intentarà recuperar a l'onada següent (llar amb pèrdua o rebuig no definitiu).
- Individus que formen part d'una llar que no ha participat a l'onada, es considera una llar perduda definitiva, i també es consideren perduts definitius tots els membres que la conformen.

A la taula 6 es mostra l'evolució de la mostra individual elegible del PaD amb la distinció entre els enquestats i els no enquestats, i cada un dels subgrups descrits:

Taula 6.
Grau d'enquestació dels individus elegibles del PaD

		1a onada		2a onada		3a onada		4a onada	
Individus enquestats	Enquestats individuals	3.979	81,0%	3.403	66,8%	3.124	64,5%	3.014	68,9%
	Enquestats proxy	609	12,4%	623	12,2%	839	17,3%	776	17,7%
Individus no enquestats	No enquestats a nivell individual (la llar participa)	325	6,6%	326	6,4%	156	3,2%	112	2,6%
	No enquestats a nivell individual (la llar no participa però no es considera rebuig o pèrdua definitiva)	0	0%	357	7,0%	216	4,5%	195	4,5%
	No enquestats a nivell individual (la llar no participa i es considera rebuig o pèrdua definitiva)	0	0%	381	7,5%	509	10,5%	276	6,3%
Mostra total d'individus elegibles		4.913	100%	5.091	100%	4.844	100%	4.373	100%

Font: Elaboració pròpia a partir de les dades del PaD. Primera onada 2001-2002; segona onada, 2003, tercera onada, 2004, i quarta onada 2005.

A la taula 6 es pot observar com l'enquestació individual és el comportament predominant de la mostra individual elegible, que si bé pateix una davallada de vint punts entre la primera i la segona onada, es manté relativament constant en un 66% en les onades posteriors.

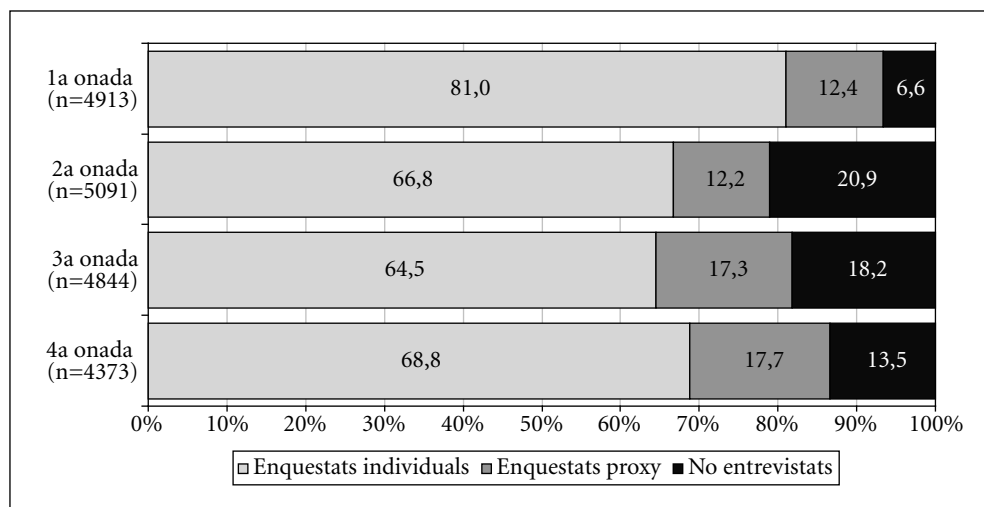
Quan no podem accedir als individus perquè es perd la llar en el seu conjunt es genera el grup de no enquestats individuals més important (se sumen els casos de llars amb pèrdua o rebuig no definitiu amb les llars amb pèrdua definitiva). En conjunt suposa aproximadament el 15% de la no enquestació de la segona i tercera onada, mentre que la no enquestació per motius individuals és un factor que va disminuint a mesura que avancen les onades (del 6,6% de la primera onada al 2,6% de no enquestació de la quarta onada).

Això ens permet constatar que la dificultat principal de l'enquestació individual rau precisament en l'efecte de no poder enquestar una llar sencera, ja que la llar arrossega tots els individus que la conformen i cada un d'ells suma una enquesta individual no realitzada. En canvi, dins de les llars enquestades, cada vegada són menys els individus que a títol individual decideixen no seguir participant al PaD. Semblaria que la participació d'algun membre de la llar arrossega la participació dels altres i que la no participació d'un membre posa en perill l'enquestació de tota la resta amb la consegüent pèrdua de la llar sencera.

Si l'anàlisi es fa per tipus de qüestionari administrat, s'obtenen els resultats que veiem en el gràfic 6:

El nivell d'enquestació segons tipus de qüestionari administrat reflecteix el que se sol donar a la majoria d'enquestes panel: una forta davallada en el nivell d'entrevistats individuals entre la primera (81%) i la segona onada (67%). Amb tot, l'enquestació individual és sempre la més important a totes les onades i es manté al voltant del 65% dels enquestats.

Una part de les pèrdues d'enquestats individuals s'amortitza amb les enquestes *proxy* que ens contesta un altre membre de la llar en nom de qui no pot o no vol ser enquestat. La xifra global de *proxys* és molt similar a totes les onades (entre el 12 i el 17%) amb una lleugera tendència a l'increment. També cal dir, però, que no tots els motius que hi ha darrere de l'administració del qüestionari *proxy* són de rebuig individual a participar, ans al contrari, n'hi ha un nombre important que s'administren per motius justificats com ara una malaltia, un accident, una hospitalització, una absència temporal de la llar per motius d'estudis o treball, una discapacitat psíquica, entre altres.

Gràfic 6.**Tipus d'enquestació dels individus elegibles del PaD**

Font: Elaboració pròpia a partir de les dades del PaD. Primera onada 2001-2002; segona onada, 2003, tercera onada, 2004, i quarta onada 2005.

A les primeres onades del PaD és també on es concentra el rebuig individual més important. Amb tot, ja s'ha esmentat que l'efecte més devastador sobre les xifres de no enquestats la provoca el fet de perdre llars senceres. Probablement, les pèrdues mostrals més elevades es concentren entre la segona i tercera onades perquè són les onades de consolidació del projecte en què llars i individus copsen amb més claredat el sentit de participar en una enquesta panel i es decideixen a seguir-hi participant o a abandonar el projecte de forma definitiva.

5 L'elaboració de les matrius de dades

Les matrius de dades del PaD¹³ s'elaboren en finalitzar el treball de camp de cada onada seguint una sèrie de passos que inclouen la consolidació de les dades dels qüestionaris, la depuració dels fitxers, la creació de les variables generades, els controls de freqüències de les variables, l'anonimització dels fitxers i la ponderació de la mostra. Un cop assolits tots aquests passos, les matrius de dades es posen a disposició dels analistes. Habitualment, podem oferir les noves matrius de dades al cap d'una any i mig d'haver finalitzat el treball de camp.

LA CONSOLIDACIÓ I DEPURACIÓ DE LES DADES

La consolidació és el procés de neteja dels fitxers que contenen les entrevistes de cada enquestador, tot descartant les entrevistes nul·les i incorporant les entrevistes vàlides en fitxers únics (un per les dades de llar i l'altre per les dades individuals) que acabaran configurant les matrius de dades definitives.

La consolidació de les dades es fa amb el programa d'anàlisi estadístic Gandia Barbwin, un software de la mateixa família que el programa d'enquestació CAPI. El fet que tots dos programes siguin productes de la mateixa família permet la importació automàtica de les entrevistes des del Gandia Capi al Gandia Barbwin sense passos intermedis.

13. Els termes matrius (de dades) i fitxers (de dades) s'utilitzen indistintament, com a sinònims.

La consolidació de les dades de les entrevistes genera dos fitxers de dades (que acabaran configurant les matrius definitives):

1. El fitxer amb les variables corresponents als qüestionaris individual i *proxy*.
2. El fitxer amb les variables corresponents al qüestionari de llar.

Quadre 3.

Passos de la depuració de les dades

1. Control de recepció de les dades	Controls per verificar que disposem de totes les dades (entrevistes) procedents del treball de camp
2. Controls estructurals primaris	Controls per verificar la consistència dels identificadors de llar i individuals a nivell transversal i longitudinal. Controls per verificar que la mobilitat dels individus (altes i baixes) ha quedat codificada segons els principis metodològics de l'enquesta panel.
3. Controls estructurals secundaris	Controls de consistència transversal i longitudinal de les variables demogràfiques individuals (sexe, lloc de naixement, any de naixement...): correcció i actualització de les dades demogràfiques de les onades anteriors Controls de les variables en què la resposta fa referència a algun membre de la llar (controls d'integritat de les ubicacions).
4. Controls de continguts	Controls encaminats a corregir possibles errors en la codificació de les respostes o bé a modificar alguna resposta a petició del propi enquestat/da a partir de les observacions de l'enquestador/a.
5. Controls de consistència interna	Controls per validar la consistència de les variables a partir de la lògica del qüestionari: contrast entre informació del qüestionari individual i de la llar, examen de la consistència d'un grup de variables d'un mateix bloc temàtic... Control de combinacions de resposta poc plausibles (per exemple, en variables múltiples o entre dues o més variables). Controls de valors anòmals o <i>outliers</i> de les variables numèriques.
6. Controls de consistència longitudinal	Controls per validar la consistència de les variables de determinats blocs (treball productiu, fecunditat, educació, habitatge...) a partir de la perspectiva longitudinal: aquest tipus de controls permeten corregir moltes inconsistències (no només de l'onada en curs sinó d'onades anteriors) i recuperar un percentatge important d'informació.
7. Controls de consistència de rutes	Controls per assegurar el bon funcionament dels salts i filtres programats al qüestionari: assegurar que cada variable tingui el mateix nombre de casos plens que els definits a la seva base.

Font: Elaboració pròpia.

La depuració de les dades es fa sobre aquests dos fitxers originals, i consisteix a aplicar una sèrie de controls que garanteixen la producció d'unes matrius de dades amb consistència interna transversal (com a dades específiques de cada onada) i longitudinal.

El principi bàsic de la depuració és no modificar les dades facilitades pels enquestats a no ser que hi hagi un error evident que es pugui corregir. Si una persona enquestada ofereix respostes inconsistents i no hi ha cap evidència que permeti fer algun tipus de correcció, es manté la inconsistència i queda en mans de l'analista destriar els casos a eliminar en funció de l'anàlisi que s'estigui desenvolupant.

El temps i recursos invertits en la depuració dels fitxers de dades ha variat a cada onada, però des de la primera onada s'ha treballat a partir d'uns protocols que estableixen una sèrie de passos seqüencials, de manera que per passar d'una fase a l'altra de la depuració cal finalitzar abans els controls de cada etapa.

CREACIÓ I ACTUALITZACIÓ DE LES VARIABLES GENERADES

Es consideren variables generades totes aquelles variables que no es recullen en el moment de realitzar les entrevistes sinó en finalitzar el treball de camp. La major part de les variables generades es construeixen des del PaD, però hi ha un petit grup de variables que provenen de l'anàlisi que han realitzat diferents experts i que han posat a disposició d'altres analistes (variables socialitzades).

A cada onada, a banda de construir les variables generades per les matrius de dades de l'onada, es tornen a calcular les variables generades de les matrius de dades de les onades anteriors, ja que en alguns casos les depuracions longitudinals de les dades modifiquen valors de variables que intervenen en la construcció de les variables generades.

Quadre 4.**Tipologia de variables generades des del PaD**

Variables per programar les referències dels qüestionaris	Variables que generem abans de l'inici del treball de camp, és a dir abans de realitzar les entrevistes, amb informació d'onades anteriors. A partir de la majoria d'aquestes variables programem les rutes dels qüestionaris. Tot i que aquestes variables no tenen cap valor analític per si mateixes, les incorporem a les matrius de dades per poder fer contrastos a nivell de bases de variables.
Variables codificades	Variables elaborades a partir del tancament (o codificació) de les respostes literals de les variables obertes dels qüestionaris. La codificació de les respostes literals no obeeix a un model preestablert a excepció de les variables que fan referència a l'activitat, l'ocupació de l'enquestat o el problema de salut més important, per a les quals s'han adoptat les classificacions oficials.
Variables d'imputació	Variables generades a partir dels controls de consistència longitudinal, que permeten depurar i recuperar informació. En el cas que hi hagi variables en les quals s'ha pogut recuperar un volum important d'informació, s'elaboren noves variables on s'indica quins valors s'han imputat i quins provenen directament del treball de camp.
Variables de ponderació	Variables per equilibrar la mostra i permetre l'anàlisi transversal i longitudinal de les dades.
Resta de variables	Aquesta darrera categoria constitueix el grup més nombrós de variables generades. Acull variables de diferent temàtica i complexitat: algunes són recodificacions senzilles a partir d'una o més variables d'observació directa (edat, membre que forma part del nucli de la llar, província de residència...), però altres obeeixen a uns models teòrics més complexos (tipus de llar, posició social...).

Font: Elaboració pròpia.

CONTROL DE LES FREQUÈNCIES DE LES VARIABLES

En finalitzar la depuració de les dades es procedeix al control de freqüències de totes les variables de les matrius per garantir uns estàndards mínims de qualitat en la informació que es proporciona. Aquests controls estan orientats a detectar variables amb un percentatge significatiu de casos perduts.

En cas que la variable presenti un percentatge significatiu de casos perduts¹⁴ es fa un contrast de freqüències amb la variable original (abans de ser depurada) per assegurar

.....
14. Generalment les variables amb un percentatge significatiu de casos perduts tenen a veure amb errors en la programació de les rutes dels qüestionaris.

que no ha quedat malmesa pel propi procés de depuració. Si la variable original no presenta un percentatge anòmal de casos perduts es recupera la variable original i se substitueix per la variable que ha quedat malmesa.

En cas contrari, es procedeix de la manera següent:

- Quan el percentatge de casos perduts de la variable oscil·la entre un 5% i un 25% del total de casos, s'adverteix del percentatge de casos perduts de la variable al llibre de codis de la matriu, per tal que sigui cada analista qui decideixi què fer amb la variable (trobareu informació sobre els llibres de codis més endavant).
- Quan el percentatge de casos perduts de la variable és superior al 25% del total de casos, s'invalida, és a dir, s'elimina la variable de la matriu de dades.

ANONIMITZACIÓ DE LES MATRIUS DE DADES

L'anonimització de matrius consisteix a eliminar totes les variables que poden posar en risc l'anonimat de les persones enquestades com ara: variables obertes amb respostes literals, variables d'identificació del municipi, noms i cognoms de les persones enquestades o dels membres que integren les llars i dades de contacte i seguiment de les persones enquestades.

ORGANITZACIÓ DE LES VARIABLES A LES MATRIUS DE DADES

L'organització de les variables de les matrius del PaD es fa en primer lloc amb una classificació en tres grans grups en funció d'on prové la informació que contenen: 1) si es tracta d'informació que s'emplena de manera automàtica (identificadors individuals i de llar); 2) si són d'observació directa (variables estructurals i variables del qüestionari) o 3) si es generen en finalitzar el treball de camp (variables generades).

Les variables se situen a les matrius seguint l'ordre d'aquests tres grups esmentats. En totes elles se segueix el sistema de numeració lògic¹⁵ i, a més, es pot identificar si són alfanumèriques, numèriques, categòriques o multiresposta.

Quadre 5.
Tipologia de variables del les matrius del PaD

Tipus de variable	Origen de la informació	Observacions
Variables clau	Construcció automàtica	Funció identificativa (llars/individus) + vinculativa (dimensió individual/llar o bé longitudinal). Els identificadors s'assignen automàticament (a través d'una programació interna).
Variables estructurals	Observació directa	Contenen informació directa d'enquestació facilitada per l'informador principal al CampPanel (base de dades que s'omple prèviament al passi dels qüestionaris)
Variables de qüestionari		Contenen la informació recollida al qüestionari individual/proxy/llar. Estan ordenades seguint la lògica d'enquestació (blocs temàtics). Se situen a continuació de les variables estructurals.
Variables generades	Creació a partir de variables d'observació directa	Tenen diferents tipus de variables i complexitat tècnica. A nivell de fitxers, se situen a continuació de la variable d'observació directa a partir de la qual s'han generat, excepte les variables de ponderació que se situen al final de tot del fitxers.

Font: Elaboració pròpia.

TIPOLOGIA DE LES MATRIUS DEL PAD

Com ja s'ha dit, des del PaD s'asseguren dos nivells d'anàlisi, l'individual i el de llar, a través de dos qüestionaris diferenciats: el qüestionari individual i el qüestionari de llar. A partir d'aquí, es construeixen quatre matrius o fitxers de dades per a cada onada del PaD. La descripció de cada una d'aquestes matrius la podeu trobar al quadre següent:

.....
15. Les variables de les matrius de dades segueixen un sistema de numeració lògic. Aquesta nomenclatura permet identificar l'onada a la qual correspon la variable, el bloc temàtic del qüestionari, el membre de la llar al qual es refereixen les dades de la variable, el nombre d'unitats de la variable, la dimensió individual/de llar de la variable i si la variable és d'observació directa o és generada. Per a més informació, consulteu el *Manual d'ús* al web <www.obdesigualtats.cat>.

Quadre 6.**Descripció de les matrius del PaD**

Nom	Unitat d'anàlisi	Continguts
X_INDI	Individual	• Conté un registre (fila) per cada membre que forma part d'una llar enquestada (independentment de si el membre ha estat enquestat o no). • Conté variables (columnes) que provenen del qüestionari individual i proxy, i variables estructurals dels individus (sexe, edat, lloc de naixement...). • Els registres corresponents a individus que no han estat enquestats només contenen la informació de les variables estructurals (facilitada per l'informant principal de la llar). • La variable clau del fitxer és el codi del membre.
X_LLAR	Llar	• Conté un registre (fila) per cada llar enquestada a l'onada. • Conté variables que provenen del qüestionari de la llar, i les variables estructurals de la llar (nombre de membres, tipus de llar, província de residència...) • La variable clau del fitxer és el codi de la llar.
X_MOSTRA_INDI	Individual	• Conté un registre (fila) per cada membre que formava part d'una llar tancada a l'onada anterior més tots els membres que s'han incorporat a la mostra a l'onada en curs (membres que s'han donat d'alta a la mostra). • Conté variables (columnes) que permeten fer un seguiment del moviment dels individus de la mostra (altes i baixes) en el transcurs d'una onada. • La variable clau del fitxer és el codi del membre. • Aquest fitxer comença a existir a partir de la segona onada (a la primera onada no hi ha altes ni baixes de membres).
X_MOSTRA_LLAR	Llar	• Conté un registre per: 1) Les llars que s'han llançat al camp (llars que existien en onades anteriors). 2) Les llars que es creen en el transcurs de l'onada (llars noves). • El tipus de variables que conté aquest fitxer són: estat de la llar, motiu pel qual no s'ha entrevistat la llar, nombre de membres de la llar, nombre d'altes/baixes a la llar, canvi d'adreça de la llar respecte l'onada anterior, entre altres (veure el llibre de codis de la matriu corresponent). • Aquest fitxer ens permet fer un seguiment del treball de camp i la dinàmica de les llars en termes de participació a l'enquesta. • La variable clau del fitxer és el codi de la llar. • Comença a existir a partir de la segona onada, perquè és quan comença a tenir sentit estudiar la dinàmica de les llars de la mostra.

Font: Elaboració pròpia.

ELS LLIBRES DE CODIS

El llibre de codis d'una matriu conté la informació necessària de cada variable per poder-ne fer l'anàlisi. Cada matriu de dades va necessàriament acompanyada del seu llibre de codis.

Els llibres de codis són una rèplica de l'esquelet de cada fitxer de dades, és a dir, contenen les mateixes variables que les matrius de dades, a excepció de les variables que fan referència a cada individu de la llar (en tractar-se de les mateixes variables repetides per cada individu s'opta per fer aparèixer només la variable corresponent a la primera posició de l'individu). Tots els llibres de codis tenen la mateixa estructura:

- Encapçalament: paràgraf introductori que sintetitza l'estructura de la matriu i adverteix de possibles incidències.
- Cos central: mostra les característiques de cada variable organitzada segons l'ordre que ocupa a la matriu. Per a cada variable es mostra: el text de la pregunta, la tipologia de la variable, els codis de resposta, la base (etiqueta i sintaxi), l'àmbit temàtic i les onades on apareix la variable.

Els llibres de codis van acompanyats de tres annexos¹⁶ per a cada onada:

- ANNEX A: conté totes les variables amb més de vint-i-cinc codis de resposta (per exemple, la nacionalitat, el parentiu, el municipi de naixement...).
- ANNEX B: conté totes aquelles variables generades (internament o per part dels analistes) que degut a la seva complexitat metodològica mereixen ser documentades amb detall (procés de generació, sintaxi de la construcció, font, referències bibliogràfiques...).
- ANNEX C: conté estadístics descriptius de totes les variables de les matrius i un indicador de la qualitat de la variable (mesurat a partir dels estadístics).

16. Els annexos B i C estan en construcció.

CREACIÓ DE MATRIUS LONGITUDINALS

Un dels elements més complicats de gestionar en una enquesta panel és el canvi, que es pot produir a dos nivells: a nivell de llars (pèrdua de llars o creació de noves llars) o a nivell d'individus (baixes i altes d'individus dins de les llars). Per codificar bé tots aquests canvis resulta fonamental assignar un codi per identificar cada llar i cada individu.

Aquests codis són invariables al llarg de tot el panel a fi i efecte de poder resseguir la trajectòria de totes les llars i de tots els individus malgrat els canvis que es vagin produint al llarg de les onades del PaD.

En l'anàlisi longitudinal cal fusionar dades d'un mateix individu (o llar) que es troben ubicades en matrius d'onades diferents. L'única manera de fer la fusió és a partir del codi de la llar i el codi d'individu que els identifiquen al llarg de les diferents onades.

MATERIALS DE SUPORT: ELS INFORMES LONGITUDINALS

Els informes longitudinals són uns materials útils i necessaris per a l'anàlisi longitudinal o transversal de les dades del PaD.

Aquest tipus d'informes permet fer una lectura en clau longitudinal de totes les variables del PaD. Quan parlem de lectura longitudinal de les variables ens referim a la possibilitat de veure les característiques de cada variable, les onades en què la podem trobar, i el comportament i els canvis que ha sofert a cada onada.

L'informe longitudinal està format per totes les variables que han sortit alguna vegada al PaD. Hi ha tants informes longitudinals com blocs temàtics té l'enquesta.

⑥ L'accés a les dades del PaD

CRITERIS DE CESSIÓ DE LES DADES DEL PAD

La Fundació Jaume Bofill posa les dades del PaD a disposició de tota la comunitat investigadora que desitgi analitzar-les. Per poder-hi accedir, cal la presentació d'un projecte de recerca que justifiqui l'ús de les dades a través d'un full de sol·licitud disponible al web del PaD (www.obdesigualtats.cat).

La Fundació garanteix l'accés gratuït a les dades per a tots aquells projectes que no tinguin pressupost assignat per a l'obtenció de dades. En cas contrari, la Fundació estableix en cada cas la quantia que cal abonar per a l'obtenció de les dades. També es poden acollir a les cessions de dades els projectes de tesi de doctorat, tesines i activitats docents de forma totalment gratuïta.

COM FORMULAR LA DEMANDA DE DADES

Per sol·licitar les dades del PaD cal presentar una memòria científica breu (projecte de recerca) que inclogui els apartats següents:

- Resum de la recerca (*abstract*)
- Objectius

- Metodologia
- Justificació de l'ús de les dades del PaD, especificant les onades amb què es vol treballar
- Àmbits d'anàlisi i/o variables que se sol·liciten
- Calendari de realització de la recerca (o bé una estimació)
- Pressupost per a l'obtenció de dades (en cas que el projecte hagi rebut finançament)
- Experiència prèvia en l'ús de bases de dades (especificar quines)

La Fundació valora els projectes d'acord amb les informacions aportades a la memòria científica i es reserva el dret de demanar aclariments o ampliació d'informacions quan ho considera necessari.

LA CESSIÓ DE DADES I EL CONTRACTE

Un cop valorada la proposta de recerca i dels criteris ja esmentats es procedeix a la cessió de les dades que consisteix en la signatura del contracte i en la recollida de les dades en suport CD-ROM. Durant la cessió de dades s'ofereix als analistes que ho desitgin una introducció a l'estructura de les dades i a les seves possibilitats així com l'assessorament inicial en l'estructura i combinació dels diferents fitxers de dades i en el treball amb la documentació pròpia del PaD.

Tots els analistes usuaris de les dades del PaD entren a formar part automàticament de la xarxa d'usuaris del PaD. Periòdicament, a través del correu electrònic, s'avisa de les actualitzacions de les dades, les novetats i altres informacions d'interès del PaD.

FINALITZACIÓ DE LA RECERCA

Tal i com s'estipula al contracte, un cop finalitzada la recerca, els autors han de lliurar un exemplar de la recerca a la Fundació Jaume Bofill així com una petita ressenya (d'unes dotze línies, aproximadament) amb les aportacions principals de la recerca.

La voluntat de la Fundació Jaume Bofill en aquest aspecte és doble: per una banda, millorar la qualitat del PaD amb les aportacions dels usuaris de les dades i, de l'altra, conèixer els resultats de les diferents perspectives d'anàlisi de la realitat social i de les desigualtats socials amb les dades del PaD.

La Fundació, prèvia valoració de l'interès de la recerca presentada, s'esforça per fer difusió dels resultats i, si s'escau, publicar-los.



El Panel de Desigualtats Socials a Catalunya (PaD) és una enquesta longitudinal sobre estructura i desigualtats socials que es va posar en marxa l'any 2001-2002, i que des del seu inici s'edita amb periodicitat anual. La característica fonamental d'un panel és que la mostra original, en el cas del PaD 2.000 llars catalanes i tots els seus integrants, es manté al llarg dels anys. Aquesta característica mostrал permet recollir dades amb caràcter longitudinal, la qual cosa converteix el PaD en un instrument, probablement imprescindible, per estudiar els canvis que s'estan produint a l'interior de la societat catalana que tenen a veure amb les desigualtats.

El PaD parteix de la definició bàsica de la desigualtat com l'expressió social de les diferents possibilitats dels individus per accedir als béns i als recursos socials com ara l'educació, el treball, la reproducció, l'habitatge, la salut o la riquesa. La seva vocació és la de fer emergir no només les diferències entre la població enquestada, sinó també els mecanismes que generen o reforcen les desigualtats socials.

El PaD té el suport de

CAIXA CATALUNYA
OBRA SOCIAL



Editorial
Mediterrània

