HOUSEHOLD INDEBTEDNESS ACCORDING TO THE SPANISH SURVEY OF HOUSEHOLD FINANCES AND THE CENTRAL CREDIT REGISTER: A COMPARATIVE ANALYSIS

2022

BANCO DE **ESPAÑA**

Eurosistema

Documentos Ocasionales N.º 2205

Olympia Bover, Laura Crespo and Sandra García-Uribe

HOUSEHOLD INDEBTEDNESS ACCORDING TO THE SPANISH SURVEY OF HOUSEHOLD FINANCES AND THE CENTRAL CREDIT REGISTER: A COMPARATIVE ANALYSIS

HOUSEHOLD INDEBTEDNESS ACCORDING TO THE SPANISH SURVEY OF HOUSEHOLD FINANCES AND THE CENTRAL CREDIT REGISTER: A COMPARATIVE ANALYSIS

Olympia Bover

BANCO DE ESPAÑA

Laura Crespo

BANCO DE ESPAÑA

Sandra García-Uribe

BANCO DE ESPAÑA

The Occasional Paper Series seeks to disseminate work conducted at the Banco de España, in the performance of its functions, that may be of general interest.

The opinions and analyses in the Occasional Paper Series are the responsibility of the authors and, therefore, do not necessarily coincide with those of the Banco de España or the Eurosystem.

Reproduction for educational and non-commercial purposes is permitted provided that the source is acknowledged.

The Banco de España disseminates its main reports and most of its publications via the Internet on its

© BANCO DE ESPAÑA, Madrid, 2022

ISSN: 1696-2230 (on-line edition)

website at: http://www.bde.es.

Abstract

The aim of this study is to analyse the quality of the information on indebtedness gathered by the Spanish Survey of Household Finances ("EFF" by its Spanish initials). To this end, we match EFF data with the administrative data from the Central Credit Register ("CIR" by its Spanish initials), which every month details all outstanding loans in excess of €6,000 arranged by individuals with financial institutions in Spain. Given the differences between the two sources in terms of the information they gather, we construct and compare various measurements of household indebtedness. In order to minimise the differences associated with the discrepancies in household composition according to the municipal population register and the EFF, we analyse both the total linked sample and a subset of comparable households. Our findings show that, after controlling for the limitations of the link, indebtedness calculated with the EFF and the CIR is similar. 25.8% of households have mortgage debt according to the EFF, versus 29.9% according to the CIR. Within indebted households, the median mortgage debt recorded in the EFF is only 0.5% lower than the figure according to the CIR. Non-mortgage debt differences are bigger, but not substantial. 18% of households have non-mortgage debt according to the EFF, versus 23% according to the CIR, and the median debt is 10% lower in the EFF. Moreover, the detailed information provided by the survey on the characteristics of households and their respective debts makes it possible to identify the age of the reference person and the existence of debts shared with individuals who are not members of the household as being the characteristics that have the most bearing on the discrepancies between the EFF and the CIR. The findings of this analysis will help improve the gathering of information and the protocols for interviewing households for the EFF.

Keywords: indebtedness, households, comparative analysis, survey data, administrative data.

JEL classification: C81, C83, G51.

Resumen

El objetivo de este estudio es analizar la calidad de la información sobre endeudamiento recogida por la Encuesta Financiera de las Familias (EFF). Para ello, implementamos un cruce de los datos de esta con los datos administrativos de la Central de Información de Riesgos (CIR), que contiene, con carácter mensual, todo el universo de préstamos vivos de más de 6.000 euros contraídos por personas físicas con entidades financieras en España. Teniendo en cuenta las diferencias entre ambas fuentes en relación con la información que recogen, construimos y comparamos diversas medidas de endeudamiento a nivel de hogar. Para minimizar las diferencias asociadas a las discrepancias en la composición del hogar según el padrón y la EFF, realizamos el análisis tanto en la muestra enlazada total como en un subconjunto de hogares comparables. Nuestros resultados muestran que, una vez controladas por las limitaciones del enlace, las medidas de endeudamiento calculadas con la EFF y con la CIR son similares. La proporción de hogares con deuda hipotecaria es del 25,8 % en la EFF y del 29,9 % en la CIR. Dentro de los hogares endeudados, la deuda hipotecaria mediana recogida en la EFF es solo un 0,5 % inferior a la recogida en la CIR. Por lo que se refiere a deuda no hipotecaria, las diferencias son mayores, pero no sustanciales. La tenencia es del 18 % en la EFF y del 23 % en la CIR, y la deuda mediana es un 10 % inferior en la EFF. Además, la disponibilidad de información detallada que proporciona la encuesta sobre las características de los hogares y sus respectivas deudas permite identificar que la edad de la persona de referencia y la existencia de deudas compartidas con personas que no pertenecen al hogar son las características que se relacionan en mayor medida con las discrepancias entre la EFF y la CIR. Los resultados de este análisis contribuirán a la mejora de la recogida de información y de los protocolos de entrevista a los hogares de la EFF.

Palabras clave: endeudamiento, hogares, análisis comparativo, datos de encuesta, datos administrativos.

Códigos JEL: C81, C83, G51.

Contents

| Abstract 5 |
|---|
| Resumen 6 |
| 1 Introduction 8 |
| 2 Comparability of debts according to the EFF and the CIR 10 |
| 2.1 CIR debts 12 2.2 EFF debts 12 |
| 3 Proportion of indebted households 14 |
| 4 Analysis of the discrepancies in the proportion of indebted households 17 |
| 5 Indebted households' debt level 27 |
| 6 Conclusions 32 |
| Annex 1 Methodology for linking the Spanish Survey of Household Finances and the Central Credit Register 33 |
| Annex 2 Analysis of discrepancies with the municipal population register 35 |
| Annex 3 Basic definitions of the characteristics of the risks in the Central Credit Register 37 |

Introduction

There are two sources of micro data on household indebtedness in Spain. First, the Spanish Survey of Household Finances ("EFF" by its Spanish initials) has been gathering information on wealth, debt, income and expenditure from a representative sample of households in Spain since 2002.1 Second, the Central Credit Register ("CIR" by its Spanish initials) records monthly information on all risks in excess of €6,000 incurred by natural and legal persons with credit institutions in Spain.

This paper sets out an analysis of household indebtedness in Spain captured by the EFF, comparing this information with that provided by the CIR. First of all, we describe the methodology developed to compare the information on household debt gathered by these two sources. Specifically, we build a database at household and interview-date level, containing the most comparable measurements of indebtedness from the CIR and the EFF. To this end, we use the ID numbers of the persons registered in the households in the EFF sample, which are then matched with the information from the CIR using a procedure that safeguards the confidentiality of the data from both sources. Nonetheless, we find a significant number of households where the number of registered individuals and the number of household members reported in the EFF does not tally. With this in mind, we design a comparable sample to account for these differences in the analysis. Using this sample, we analyse the differences in the indebtedness metrics calculated with the EFF and the CIR. Second, we analyse which demographic and debt characteristics are more closely related to the differences between household indebtedness according to the EFF and the CIR.

Our findings show that, after controlling for the limitations of the link, indebtedness calculated with the EFF and the CIR is similar. In the comparable sample, 41.8% of households are indebted according to the CIR, versus 36.3% according to the EFF. 25.8% of households have mortgage debt according to the EFF, versus 29.9% according to the CIR. Within indebted households, the median mortgage debt recorded in the EFF is only 0.5% lower than the figure according to the CIR. Non-mortgage debt differences are bigger, but not substantial. 18% of households have non-mortgage debt according to the EFF, versus 23% according to the CIR, and the median debt is 10% lower in the EFF. In addition, analysing the characteristics of the households and their debts and the probability of a household reporting them shows that the discrepancies are largely due to households whose reference person is over 55 failing to respond and the existence of debts shared with individuals who are not members of the household. Furthermore, the findings also enable us to learn certain lessons to help us improve how the EFF measures household debt, e.g. in terms of the protocols for interviewing older households. Specific proposals could include assigning specialist interviewers for this type of household, strengthening survey methods, in addition to warnings or reminders that pop up during the interview so as to avoid debts being forgotten, especially those shared with individuals who are not members of the household.

¹ A detailed description of the survey can be found at https://www.bde.es/bde/en/areas/estadis/estadisticas-por/ encuestas-hogar/relacionados/Encuesta Financi/.

In Section 2 we present the methodology for comparing the information on household debt in Spain using the two sources of available micro data (the EFF and the CIR). In Section 3 we examine the percentage of indebted households in Spain according to the two sources. In Section 4 we analyse the debt and household characteristics that have the biggest bearing on the differences in indebtedness between the two sources of information. In Section 5 we detail the differences between the outstanding debt of the indebted households in the two sources. Section 6 sets out the main conclusions of the study.

2 Comparability of debts according to the EFF and the CIR

For this analysis, we linked the data on households collected in the various editions of the EFF with the data on debtors of credit institutions in Spain recorded on the CIR. In this section we describe the methodology used to this end and examine what features of the debts included in the EFF can be compared with the information from the CIR. Specifically, we only used data from the 2002, 2005, 2008 and 2011 editions of the EFF, focusing on the period 2002-2011 for two reasons. First, the information required to match the two data sources is only available up to 2011. Moreover, in 2013 significant changes were made to the recording of risks on the CIR, making it harder to compare with the previous series in our analysis.

In order to reassure participating households that their responses are confidential, to date the EFF has not collected their members' ID numbers, making it impossible to perform a direct link using an individual's national identity card number. As an alternative, and as a means of safeguarding the confidentiality of the data of the EFF participants and the debtors on the CIR, the data are linked via a rigorous blind matching process. First, the National Statistics Institute (INE by its Spanish initials) furnishes the Banco de España with the encrypted identifiers of the individuals registered at the address where the household resided when selected. Second, the encrypted identifiers are used, together with the EFF interview date, to gather information on the household risks registered on the CIR and link it with the information in the EFF. The methodology for linking the two samples is detailed in Annex 1. In total, 21,396 out of 23,415 households could be linked. The reasons why 2,019 households could not be linked are set out in Annex 1, and have to do with sampling frame deficiencies (in particular, as the 2002 and 2005 municipal population register data from the Basque Country and Navarre are no longer available). The population weights are adjusted on the basis of the households that can ultimately be linked. The date on which the EFF interview was conducted with a household is used to identify the corresponding risk on the CIR.

While the use of municipal population register data is thus key when linking the two sources of information on debt, in turn, it poses two challenges to comparability. An initial challenge lies in the fact that the municipal population register data do not always offer up-to-date, accurate information on the persons living at a particular address at the time of the interview. Meanwhile, the EFF records the number of household members at the exact time of the interview. A comparison of the number of household members and the number of persons registered at the address where that household resided when the sample was selected reveals certain differences (see Table 1). For 42.8% of households, there are no differences between the number of registered persons and the number of household members. However, 29.2% of households have more registered persons than there are members. This difference could be due to: several households residing at the same address, but only one of them being interviewed for the EFF; the existence of domestic help registered at the address, who are not deemed part of the household for the EFF and are not interviewed; changes in the household composition that have yet to be reported to the municipal population register; differences in the household composition between the sample selection date and the interview date; and,

Table 1

DIFFERENCE IN THE NUMBER OF HOUSEHOLD MEMBERS BETWEEN

THE EFF AND THE MUNICIPAL POPULATION REGISTER (a) (b) (c)

Percentage of the total sample (d)

| | Total | Panel | Refreshment |
|----------------------------|--------|-------|-------------|
| At least four more | 1.3 | 1.3 | 1.3 |
| Three more | 3.4 | 4.1 | 2.9 |
| Two more | 9.4 | 10.8 | 8.3 |
| One more | 13.9 | 14.3 | 13.7 |
| No difference | 42.8 | 36.0 | 48.0 |
| One fewer | 16.7 | 18.5 | 15.2 |
| Two fewer | 7.5 | 9.5 | 6.0 |
| Three fewer | 2.9 | 3.4 | 2.5 |
| At least four fewer | 2.1 | 2.1 | 2.0 |
| Number of observations (e) | 21,396 | 9,319 | 12,077 |

SOURCE: Banco de España.

- a The differences are calculated by subtracting the number of registered individuals from the number of household members.
- **b** Data on registered persons in each household correspond to the first time a household was included in the EFF.
- c The number of household members is addressed in question p.1 during the interview.
- d The figures in the table are percentages of the total observations by type of sample.
- e The number of observations corresponds to the total households in the linkable sample for each type of sample.

finally, the existence of potential measurement errors inherent in the survey data. Elsewhere, 28% of households have more members according to the survey than there are registered persons. This difference may be due to children or relatives that have not been registered at the address for a range of reasons. To control for these discrepancies between household members and registered persons, we analysed the subset of households for which there are no differences between the number of adult members residing in the household and those registered at the household's address (including households where this difference coincides with the number of minors). We have called this data subset (representing 50.75% of the total sample) the comparable sample.

A second challenge is that, for each household in the EFF sample, the population register data available correspond to the date on which each household was first selected, and have not been updated for any subsequent editions in which the household has participated. This affects the EFF panel households whose members have changed over time, e.g. where children leave home, spouses separate, members pass away or offspring return. To account for these problems, we have presented our findings using only the refreshment set of household data, since the gap between the sample selection date and the interview date is as narrow as possible.

Annex 2 explores how some household characteristics may have a bearing on the discrepancies between household members and the municipal population register data.

Having linked the two databases (EFF and CIR), we then identified a series of criteria that we used to obtain comparable measurements of debt between the two sources.

2.1 CIR debts

The CIR data used for this study come from the database known as the "Household CIR", which contains information on risks incurred by individuals only. In order to identify which debts are comparable with the EFF, we used the definitions and categories of risks in Circulars 3/1995 and 1/2013, page 37, the latter for the breakdown of classes of risk.

In general, the CIR records risks in excess of €6,000, with the exception of non-performing risks, to which this lower limit does not apply. Specifically, the following set of variables is used to define a risk: counterparty, institution, country, joint and several nature and risk key. In turn, five variables are used to define the risk key: type of risk, type of currency, maturity, collateral and status. See Annex 3 for further details on the definitions of these variables.

The risks that can be compared with the information from the EFF are those classed as "Finance credit" (B), "Finance lease transactions" (K) and "Loans or credit transferred to third parties" (Q). Figure A3.1 details the equivalence between these classes of risk and the types of debt in the EFF. We have not included risks in the form of collateral in this study, since these do not constitute debt incurred.

Whether risks incurred by several persons are joint and several or, alternatively, several in nature was considered. In the case of joint and several risks, the CIR records the full amount for each counterparty. Thus, in order to avoid duplications, we have distributed the proportional part of such risks among the counterparties. Several risks are recorded on the CIR in respect of the proportion corresponding to each counterparty under the contract giving rise to the risk, so this is the figure we used for the comparison.

Based on these comparability criteria, we calculated the comparable outstanding debt (whether mortgage debt or otherwise)² according to the EFF and the CIR at household level.³

2.2 EFF debts

The EFF details the overall debts of a household's members, but provides no information on the debtor in each debt.⁴ Specifically, the EFF provides information on debts incurred to purchase real estate assets (section 2 of the questionnaire), other types of debt incurred by

² To identify mortgage debt with the CIR information, we used a definition provided by the Financial Stability Department. According to this definition, mortgage risk refers to finance credit (type of risk = B or Q) that is fully collateralised (collateral = A) and with a maturity of over five years (maturity = E).

³ The EFF does not include information on the debtor in each loan, and we cannot therefore perform a loan-to-loan link. Another option would be to link loans based on the likelihood that they are the same according to a series of characteristics.

⁴ Even if such information were included, it could not be used to link to the CIR since we cannot identify the persons residing in the household as national identity card numbers have never been requested for the EFF.

the household and/or the sole proprietorships owned by household members (section 3 of the questionnaire) and credit card debt (section 8 of the questionnaire).

To compare EFF and CIR debts, first we used questions on loan type (mortgage loan, other collateralised loan, personal loan, line of credit, deferred payment, advance, loan from relatives or friends and other). Specifically, we used questions on the type of loan arranged by a household to purchase its main residence (p2_9) or other properties (p2_52) and on other debts incurred by the household (p3_2). We discarded loans classed as "Loans from relatives or friends", since these are not included in the CIR. Then, we used questions on the type of institution that arranged the relevant loan (p2_10, p2_52 and p3_4, respectively). In this case, we discarded loans from "Non-financial corporation", "Enterprise or business for which the borrower works" and "Other". We compared another class of debt included in the EFF with the information included in the CIR; namely, credit card debt (included in question p8_5a since the 2005 edition of the EFF). We also included in the comparison debts incurred to purchase additional properties over and above the fourth property, for which we do not know the type of loan or the lending institution, although such loans are in all likelihood comparable.⁵

We have also included in the comparison debts incurred with credit institutions with an outstanding balance of less than €6,000, for two reasons. First, the CIR may include debts for an amount of less than €6,000 where they have been recorded together with other risks that share the same risk key and debtor or where they are classed as non-performing. Second, because we believe that including this type of debt in the comparison enhances the analysis, as it reveals the full range of discrepancies between the indebtedness metrics from the two sources, and this characteristic can be fully controlled for in the comparative analysis.

⁵ The results of the analysis have been reproduced without including debts on other properties above and beyond the fourth property as comparable debt. The results barely differ from those set out in this paper. The same can be said if debts with non-financial corporations are included as comparable debts.

3 Proportion of indebted households

Table 2 shows the proportion of indebted households (households with at least one debt with credit institutions)⁶ in Spain from 2002 to 2011. This proportion is 7.3 percentage points (pp) higher according to the CIR than to the EFF. The percentages of households with mortgage debt according to the two sources of information are shown in the second column: according to the EFF, households with mortgage debt account for 31.1% of the total, while the equivalent percentage stands at 41.6% according to the CIR. 26.5% and 29% of households have other types of debt according to the EFF and the CIR, respectively, representing a smaller difference than in the case of mortgage debt. The standard errors shown in brackets are also similar in both. The higher rate of household indebtedness according to the CIR data can be explained by there being households with more registered individuals than household members according to the survey. Thus, CIR debts that do not correspond to such households may have been included.

Table 3 shows the proportion of indebted households by age group of the reference person.⁷ The proportions of indebted households by age group follow the same pattern in both sources. Nonetheless, compared with the CIR, the EFF shows a higher percentage of indebtedness among households under the age of 45, whereas the CIR identifies a higher percentage than the EFF among the over-44s, and the differences are greater. The greatest differences can be observed in households whose reference person is over the age of 54. Such households may still have registered offspring who no longer live at home, several registered households⁸ or, alternatively, domestic help who do not form part of the household.

To control for the problems referred to above when linking the two databases, we generated two subsets of observations. The first subset includes households that reported to the survey a number of adults identical to the number of persons on the municipal population register and those where the difference matches exactly the number of members under 16. As noted above, we have called this subset of observations the comparable sample. The population weights are not representative for this subset of observations. We therefore compared the percentage of observations with debts in sample terms according to both sources of information. The second subset of observations seeks to control for the error deriving from linking panel households; the INE does not update panel household members for addresses after the first edition in which they take part. Specifically, we have called this subset the refreshment sample, since it only includes each edition's new respondents, excluding any that have responded more than once.

Table 4 shows the sample percentages of indebted households for the different subsets of observations. Indebted households account for 37% of the total sample in the EFF and 49.7% according to the CIR. This 12.7 pp difference is higher than that calculated

⁶ All of the statistics are calculated using the five imputations of the EFF and the population weights adjusted to the linkable sample.

⁷ Member of the household most familiar with the household finances and who responds to the questionnaire.

⁸ In the case of addresses where several households reside, only one of them is interviewed and included in the EFF, whereas the debts of all registered individuals are accounted for on the CIR.

Table 2 PERCENTAGE OF INDEBTED HOUSEHOLDS (a)

Percentage of the population, pp

| | | Type of debt | | | |
|-----|------------|--------------|-----------|--|--|
| | Total | Mortgage (c) | Other (d) | | |
| EFF | 46.2 | 31.1 | 26.5 | | |
| | (0.57) (b) | (0.54) | (0.52) | | |
| CIR | 53.5 | 41.6 | 29.0 | | |
| | (0.55) | (0.57) | (0.52) | | |

SOURCE: Banco de España.

- a Percentage of indebted households in the population. The weights have been adjusted for the households that can be linked with the information from the municipal population register.
- b The standard error is shown in brackets.
- c In CIR data, "mortgage debt" is defined as financial credit that is fully collateralised and with a maturity of over five years.
- d "Other" includes all types of non-mortgage debt.

Table 3 INDEBTED HOUSEHOLDS BY AGE OF THE REFERENCE PERSON (a) (b)

| Percentage of the population, pp | Sou | irce |
|----------------------------------|------|------|
| | EFF | CIR |
| Under-35s | 68.1 | 66.3 |
| 35-44 | 68.2 | 66.2 |
| 45-54 | 57.3 | 59.9 |
| 55-64 | 43.0 | 57.1 |
| 65-74 | 19.2 | 41.7 |
| Over-74s | 5.9 | 18.0 |

SOURCE: Banco de España.

- a Percentage of indebted households by age group of the reference person.
- b The weights have been adjusted for the households that can be linked with the information from the municipal population register.

with the population weights (7.3 pp) in Table 2, probably due to an oversampling in the EFF of households at the higher end of the wealth distribution, which are typically more indebted than those at the lower end. Meanwhile, in the comparable sample, these sample proportions are more similar, i.e. 36.3% in the EFF and 41.8% according to the CIR. This greater similarity in the proportion of indebted households according to the CIR and the EFF is consistent with the fact that the total sample includes households where there are generally more registered persons than persons interviewed in the survey. Lastly, among the refreshment households, these percentages stand at 37% and 47.9%, respectively. The difference between the two is therefore smaller than in the total sample, which is consistent with the population register data on panel households not being updated between editions. It is nonetheless worth noting that including only households with the same number of adults in the survey and the municipal population register (the comparable sample) is what most reduces the differences between the EFF and the CIR. Moreover, this 5.5 pp sample difference could be expected to be even smaller in the population, as seen above for the total sample.

Table 4 **INDEBTED HOUSEHOLDS (a)**

| Sample percentage, pp | | Type of sample (d) | | | | | |
|-----------------------|------------|-----------------------------|--------|--|--|--|--|
| | Total | Total Comparable Refreshmen | | | | | |
| EFF | 37.0 | 36.3 | 37.0 | | | | |
| | (0.33) (b) | (0.46) | (0.44) | | | | |
| CIR | 49.7 | 41.8 | 47.9 | | | | |
| | (0.34) | (0.47) | (0.45) | | | | |
| Test statistic (c) | -36.0 | -13.3 | -18.1 | | | | |

SOURCE: Banco de España.

- a Percentage of indebted households for various sets of observations.
- **b** The standard error is shown in brackets.
- c The test statistic tests the difference in proportions according to the EFF and CIR household indebtedness figures. This statistic uses the standard t distribution.
- d "Total" refers to the total number of linkable household observations, "Comparable" refers to the set of observations for which the difference between household members and members registered at that household's address is zero or matches the number of children residing in the household. "Refreshment" comprises observations of households that have taken part in the survey only once.

Table 5 **INDEBTED HOUSEHOLDS BY TYPE OF DEBT (a)**

| Sample percentage, pp | | Mortgage debt | | | Other | |
|-----------------------|------------|---------------|-------------|--------|------------|-------------|
| | Total (d) | Comparable | Refreshment | Total | Comparable | Refreshment |
| EFF | 25.2 | 25.8 | 25.6 | 19.5 | 18.5 | 18.9 |
| | (0.30) (b) | (0.42) | (0.40) | (0.27) | (0.37) | (0.36) |
| CIR | 37.0 | 29.9 | 34.8 | 29.1 | 23.5 | 27.6 |
| | (0.33) | (0.44) | (0.43) | (0.31) | (0.41) | (0.41) |
| Test statistic (c) | -37.6 | -11.9 | -19.4 | -27.0 | -11.2 | -14.6 |

SOURCE: Banco de España.

- a Percentage of indebted households with each type of debt and for various sets of observations.
- **b** The standard error is shown in brackets.
- c The test statistic tests the difference in proportions according to the EFF and CIR household indebtedness figures. This statistic uses the standard t distribution.
- d "Total" refers to the total number of linkable household observations. "Comparable" refers to the set of observations for which the difference between household members and members registered at that household's address is zero or matches the number of children residing in the household. "Refreshment" comprises observations of households that have taken part in the survey only once.

Table 5 shows a comparison by types of debt. The results reveal that for the total sample there are greater differences in the sample percentages of indebted households with mortgage debt than in the percentage of indebted households with non-mortgage debt (11.8% vs 9.6%, respectively). In the comparable sample, such differences are greatly reduced, the sample percentage of households with mortgage debt standing at 25.8% and 29.9% according to the EFF and the CIR, respectively, while the same percentage for other types of debt stands at 18.5% and 23.5%, respectively.

4 Analysis of the discrepancies in the proportion of indebted households

This section looks at whether certain demographic and household debt characteristics have a bearing on the differences observed in the proportions of indebted households between the CIR and the EFF. The aim is to better understand such discrepancies and how they relate to various measurement errors that can be found in the indebtedness figures. This information will make it possible to identify possible ways to improve how household debt is measured in the EFF.

Specifically, we identified two types of discrepancy that arise when comparing household indebtedness in the two statistical sources. The first type concerns households that have debt according to the CIR but not the EFF. The second concerns households that have debt according to the EFF but not the CIR.

Table 6 shows the average marginal effects of certain demographic and household debt characteristics on the probability that a household has no debts in the EFF, and yet does according to the CIR. The two columns are the result of using the total sample to estimate a single logit model in which all of the regressors are interacted with a variable indicating whether the household belongs to the comparable sample and in which population weights have been used. The second column shows the results obtained for the total sample, while the first corresponds to the results where a household belongs to the comparable sample. Specifically, we focused on the results in the first column, since this sample enables us to more confidently identify debts omitted in the EFF that are recorded on the CIR.

The probability of featuring on the CIR but not in the EFF is greater for households that only have non-mortgage debt of some sort (6 pp higher on average than those with only mortgages). This probability is 5 pp lower for households with a mortgage and some other form of debt than for those with mortgage debt only. These results indicate that omissions of household debt are less common where a household has significant debt, such as a mortgage, and where it has several forms of debt, perhaps as such debts are harder to forget when responding to the survey. Meanwhile, the probability of being indebted according to the CIR but not the EFF is 8 pp higher for debts owed joint and severally with persons who are not household members. One interpretation of this set of results is that the bigger household debts are, the less likely they are to be included on the CIR but not reported in the EFF.

In terms of the demographic characteristics of a household, we found that the probability of debt being included on the CIR but not reported in the EFF is greater the older the household's reference person. Specifically, this probability is 8 pp, 16 pp, 33 pp and 44 pp higher for households between the ages of 45-54, 55-64, 65-74 and those over 74, respectively, as compared with the under-35s. These findings could be explained in part by the fact that elderly households find it harder to recall their debts when interviewed. We also found that households in the lowest percentile of gross wealth are more likely not to report

CHARACTERISTICS OF THE HOUSEHOLDS AND THEIR DEBTS, AND PROBABILITY OF A HOUSEHOLD HAVING DEBT ACCORDING TO THE CIR BUT NOT THE EFF (a) (b)

The probability of having debt according to the CIR but not the EFF is higher for households with non-mortgage debt, lower for households with mortgages and with more varied types of debt, and higher for households with debt shared with people who are not members of the household.

Decimals (c)

| ecimals (c) | Comparable (d) | Total |
|---|----------------|----------|
| Characteristics of the debt | | |
| Types of household debt (compared with just mortgage debt) | | |
| Mortgage and other forms of debt | -0.05*** | -0.06*** |
| | (0.02) (f) | (0.01) |
| Just other forms of debt | 0.06** | 0.02 |
| | (0.02) | (0.02) |
| Joint and several risks with people who are not household members | 0.08*** | 0.00 |
| | (0.03) | (0.02) |
| Demographic characteristics | | |
| Educational attainment level of the reference person | | |
| Secondary education | -0.01 | -0.02 |
| | (0.02) | (0.02) |
| Tertiary education | -0.01 | -0.01 |
| | (0.02) | (0.02) |
| Age of the reference person | | |
| 35-44 | 0.02 | 0.03* |
| | (0.02) | (0.02) |
| 45-54 | 0.08*** | 0.10*** |
| | (0.03) | (0.02) |
| 55-64 | 0.16*** | 0.22*** |
| | (0.03) | (0.02) |
| 65-74 | 0.33*** | 0.40*** |
| | (0.04) | (0.03) |
| Over-74s | 0.40*** | 0.49*** |
| | (0.07) | (0.04) |
| Gender of the reference person | | |
| Female | 0.05*** | 0.03** |
| | (0.02) | (0.01) |
| Gross wealth percentile of the household (e) | | |
| 20-40 | -0.23*** | -0.23*** |
| | (0.05) | (0.03) |
| 40-60 | -0.25*** | -0.22*** |
| | (0.04) | (0.03) |
| 60-80 | -0.24*** | -0.20*** |
| | (0.05) | (0.03) |
| 80-90 | -0.21*** | -0.21*** |
| | (0.05) | (0.03) |
| 90-100 | -0.25*** | -0.19*** |
| | (0.05) | (0.04) |
| | | |

SOURCE: Banco de España.

- a The effects are measured for the different sets of observations, using the five imputations related to the EFF and all the editions.
- b The estimates are based on a logit model of the probability of a household having debt according to the CIR but not the EFF, conditional on the variables shown interacting with a variable denoting whether the observation is part of the comparable sample and using the population weights adjusted to the linkable sample.
- c Average marginal effects in terms of predicted probabilities.
- d The first column presents the effects under the condition that the household is part of the comparable sample, while the second column shows the effects for any type of household.
- e The income and wealth percentiles are defined vis-à-vis the overall distribution for all the editions.
- f Standard errors robust to heteroscedasticity are shown in brackets. Significance level at 1%, 5% and 10% as indicated by ***, ** and *, respectively.

CHARACTERISTICS OF THE HOUSEHOLDS AND THEIR DEBTS, AND PROBABILITY OF A HOUSEHOLD HAVING DEBT ACCORDING TO THE CIR BUT NOT THE EFF (a) (b) (cont'd)

The probability of having debt according to the CIR but not the EFF is higher for households with non-mortgage debt, lower for households with mortgages and with more varied types of debt, and higher for households with debt shared with people who are not members of the household.

| ecimals (c) | Comparable (d) | Total |
|--|----------------|----------|
| emographic characteristics | | |
| Total income percentile of the household | | |
| 20-40 | -0.07 | -0.08*** |
| | (O.O4) (f) | (0.03) |
| 40-60 | -0.13*** | -0.16*** |
| | (0.04) | (0.03) |
| 60-80 | -0.13*** | -0.16*** |
| | (0.05) | (0.03) |
| 80-90 | -0.12** | -0.20*** |
| | (0.05) | (0.04) |
| 90-100 | -0.10** | -0.19*** |
| | (0.05) | (0.04) |
| Number of adults in the household (e) | | |
| Two | -0.01 | -0.07*** |
| | (0.02) | (0.02) |
| Three | 0.05 | -0.08*** |
| | (0.03) | (0.03) |
| More than three | 0.01 | -0.11*** |
| | (0.03) | (0.03) |
| At least one self-employed worker | 0.01 | 0.01 |
| | (0.02) | (0.02) |
| Edition | | |
| 2005 | -0.08*** | -0.07*** |
| | (0.02) | (0.02) |
| 2008 | -0.06** | -0.08*** |
| | (0.03) | (0.02) |
| 2011 | -0.09*** | -0.10*** |
| | (0.02) | (0.02) |
| Comparable observation | | -0.12*** |
| | | (0.01) |
| bservations | 4,539 | 10,627 |
| Dependent variable average | 0.20 | 0.31 |

SOURCE: Banco de España.

- a The effects are measured for the different sets of observations, using the five imputations of the EFF and all the editions.
- b The estimates are based on a logit model of the probability of a household having debt according to the CIR but not the EFF, conditional on the variables shown interacting with a variable denoting whether the observation is part of the comparable sample and using the population weights adjusted to the linkable sample.
- $\boldsymbol{c}\,$ Average marginal effects in terms of predicted probabilities.
- d The first column presents the effects under the condition that the household is part of the comparable sample, while the second column shows the effects for any type of household.
- e The number of adults in the household is that reflected in the EFF.
- f Standard errors robust to heteroscedasticity are shown in brackets. Significance level at 1%, 5% and 10% as indicated by ***, ** and *, respectively.

to the EFF debts that are registered on the CIR as compared with those in higher wealth percentiles. Households whose income falls between percentiles 0 and 40 are more likely not to report debts in the survey (but which are on the CIR) than households with income above the 40th percentile. The reference person's education has no significant bearing on

the likelihood of their having debts according to the CIR but not the EFF. However, this probability is 5 pp higher for households in which the reference person is female. Moreover, the probability of having a debt according to the CIR but not the EFF is 6-9 pp lower in all editions when compared with the 2002 EFF. This could be explained by the fact that the first edition of the survey did not include credit card debt and that the measurement error was greater in the 2002 edition.

Table 7 shows the average marginal effects of the same demographic and household debt characteristics on the probability of a household reporting in the survey debt not recorded on the CIR. As in the case of Table 6, the two columns are the result of estimating a single logit model in which all of the control variables interact with an indicator of whether the household belongs to the comparable sample, using the population weights. The second column shows the results for the total sample, while the first sets out the results where the household belongs to the comparable sample, thus enabling us to more confidently identify debts included in the EFF but not featuring on the CIR. With this in mind, we focused on the results in the first column. We found that the probability of featuring in the EFF but not on the CIR is 9 pp higher for households that only have nonmortgage debt of some sort as compared with mortgaged households. On average, the type of debt included in the EFF but not on the CIR also tends to be non-mortgage debt, more so than for the probability of being indebted on the CIR but not in the EFF. We found that the probability of reporting in the survey a debt not included on the CIR is 15 pp lower for households with both mortgage and non-mortgage debt, and this figure also exceeds that found in the analysis in Table 6. Overall, these findings can be taken to mean that the type of debt has a larger impact on the probability of featuring in the EFF but not on the CIR than vice versa. The existence of debts of less than €6,000 increases the probability of not being indebted on the CIR but having debts in the EFF by 27 pp when compared with households without this type of debt in the EFF. This is consistent with the lower limit on the risks recorded on the CIR.9

Turning to demographic characteristics, we found that the probability of reporting a debt in the survey that was not reflected in the CIR was 5 pp, 10 pp and 11 pp lower in households with two, three or more than three members, respectively, than in one-person households. Furthermore, the probability of a household debt featuring in the EFF but not on the CIR was 5 pp higher where the reference person was female. Lastly, we observed that the probability of a household debt featuring in the EFF but not on the CIR was 9 pp, 7 pp and 6 pp lower in the 2005, 2008 and 2011 editions, respectively, than in the 2002 edition.¹⁰

⁹ When debts of less than €6,000 are removed from the comparison, the impact of the type of debt on the probability of a household being indebted according to the CIR but not the EFF is reduced. Specifically, the existence of only other types of debt entails a 5.6 pp higher probability of featuring in the EFF but not on the CIR compared with households with mortgage debt only, whereas the existence of mortgage and non-mortgage debt reduces the probability of featuring in the EFF but not on the CIR by 3 pp. The existence of self-employed workers in a household has a significant bearing on this probability. Specifically, this characteristic is estimated to increase the probability by 3.4 pp. The impact of the other demographic characteristics does not vary significantly.

¹⁰ Excluding the 2002 edition, the findings do not differ significantly.

CHARACTERISTICS OF THE HOUSEHOLDS AND THEIR DEBTS, AND PROBABILITY OF A HOUSEHOLD HAVING DEBT ACCORDING TO THE EFF BUT NOT THE CIR (a) (b)

The probability of a household having debt according to the EFF but not the CIR is higher when it only has non-mortgage debt as compared with households that have mortgage debt and debts under €6,000. The probability is higher in one-person households and where the reference person is female.

| Decimals (| |
|------------|--|
| | |

| Parametristics of the debt Types of household debt (compared with just mortgage debt) -0.15** -0.14*** -0.02*** (0.02) (0.02) (0.02) (0.02) (0.02) (0.02) (0.02) (0.02) (0.02) (0.03) (0.0 | Decimals (c) | Comparable (d) | Total |
|--|--|----------------|----------|
| Mortgage and other forms of debt -0.15" -0.14" Just other forms of debt 0.02" 0.10" Debts under €6,000 0.27" 0.22" Debts under €6,000 0.27" 0.22" Debts under €6,000 0.03 0.02" Debrouge pilot characteristics 0.03 0.02 Educational attainment level of the reference person —0.02 0.01 Secondary education 0.01 0.02 0.02 Territary education 0.01 0.02 0.02 Age of the reference person 0.01 0.02 0.02 35-44 0.01 0.00 0.02 0.02 45-54 0.02 0.03 0.02 <t< td=""><td>Characteristics of the debt</td><td></td><td></td></t<> | Characteristics of the debt | | |
| | Types of household debt (compared with just mortgage debt) | | |
| Just other forms of debt 0.08** 0.10** Debts under €6,000 0.27** 0.22** Debts under €6,000 0.27** 0.22** Debts under €6,000 0.02** 0.02 Demographic characteristics Use of the reference person Use of the reference person Secondary education 0.02 0.01 Tertiary education 0.01 0.02 Age of the reference person Use of the reference person 35-44 0.01 0.00 45-54 0.02 0.02 45-54 0.03 0.02* 65-74 0.03 0.02* 65-74 0.05 0.00 0xer-74s 0.05 0.03 0xer-74s 0.05 0.03 Cender of the reference person 0.05** 0.04** Female 0.05** 0.04** 20-40 0.00** 0.02** 40-60 0.02 0.02** 40-60 0.02 0.02** 40-60 0.02 0.02** < | Mortgage and other forms of debt | -0.15*** | -0.14*** |
| Debts under €6,000 0.02*** 0.02**** 0.02**** 0.02***** 0.02****** 0.02******* 0.02******* 0.02******* 0.02****** 0.02****** 0.02****** 0.02****** 0.02***** 0.02**** 0.02**** 0.02**** 0.02**** 0.02**** 0.02 | | (0.02) (f) | (0.02) |
| Debts under €6,000 0,27*** 0,22*** Commographic characteristics Commographic characteristics Educational attainment level of the reference person −0,02 0,01 Secondary education −0,02 0,01 Tertiary education −0,01 0,02 Age of the reference person −0,01 0,00 35-44 −0,01 0,00 45-54 −0,02 −0,03 45-54 −0,02 −0,03 66-74 −0,03 −0,02 66-74 −0,05 −0,00 00-74s −0,05 −0,03 00-75 −0,03 −0,02 Penale −0,05 −0,03 Cender of the reference person −0 −0,04** Fenale −0,05 −0,03 −0,02* Cender of the reference person −0 −0 −0 Cercas wealth percentile of the household (e) −0 −0 −0 −0 −0 −0 −0 −0 −0 −0 −0 −0 −0 | Just other forms of debt | 0.09*** | 0.10*** |
| Demographic characteristics | | (0.03) | (0.02) |
| Demographic characteristics | Debts under €6,000 | 0.27*** | 0.22*** |
| Page of the reference person Page of the reference person | | (0.03) | (0.02) |
| Secondary education -0.02 (0.02) 0.01 (0.02) Tertiary education 0.01 (0.02) 0.02 (0.02) Age of the reference person -0.01 (0.02) 0.02 35-44 -0.01 (0.02) 0.02 45-54 -0.02 (0.03) 0.02 55-64 -0.03 (0.03) 0.02 65-74 0.05 (0.03) 0.02 0ver-74s 0.05 (0.03) 0.02 Deender of the reference person 0.05 (0.02) 0.05 Female 0.05 (0.02) 0.01 20-40 0.00 (0.02) 0.01 40-60 0.00 (0.02) 0.02 40-60 0.02 (0.02) 0.02 60-80 0.02 (0.02) 0.02 60-80 0.02 (0.02) 0.02 80-90 0.02 (0.03) 0.02 80-90 0.02 (0.03) 0.02 90-100 0.00 (0.03) 0.03 | Demographic characteristics | | |
| Content Cont | Educational attainment level of the reference person | | |
| Tertiary education 0.01 (0.02) 0.021 (0.02) Age of the reference person CO.02 (0.02) 0.02 (0.02) 35-44 -0.02 (0.02) 0.02 (0.02) 45-54 -0.02 (0.03) 0.02 (0.02) 55-64 -0.03 (0.02) 0.02 (0.02) 65-74 0.05 (0.03) 0.02 (0.02) Over-74s 0.05 (0.06) 0.05 (0.02) Female 0.05 (0.02) 0.05 (0.02) Gross wealth percentile of the household (e) 0.00 (0.02) 0.01 20-40 0.02 (0.02) 0.02 (0.02) 40-60 0.02 (0.02) 0.02 (0.02) 60-80 0.02 (0.02) 0.02 (0.02) 80-90 -0.02 (0.02) 0.02 (0.02) 80-90 -0.02 (0.02) 0.02 (0.02) 80-90 -0.02 (0.02) 0.02 (0.02) 90-100 -0.02 (0.03) 0.02 | Secondary education | -0.02 | 0.01 |
| (0.02) (0.01) Age of the reference person -0.01 0.00 35-44 -0.02 -0.02 45-54 (0.03) (0.02) 55-64 -0.03 -0.08* 65-74 0.05 0.00 0ver-74s 0.05 0.02 5 ender of the reference person 0.06 0.05 Female 0.05*** 0.04*** 5 cross wealth percentile of the household (e) 0.00 0.01 Scross wealth percentile of the household (e) 0.02 0.04*** 20-40 0.02 0.02 0.02 40-60 0.02 0.02 0.02 60-80 0.02 0.02 0.02 80-90 0.02 0.02 0.02 80-90 0.02 0.02 0.02 80-90 0.02 0.02 0.02 90-100 0.03 0.03 0.03 | | (0.02) | (0.02) |
| Age of the reference person 35-44 -0.01 0.00 35-44 -0.02 -0.03 45-54 -0.03 -0.03* 55-64 -0.03 -0.06*** 65-74 -0.05 0.00 0ver-74s 0.05 0.03 0ver-74s 0.05 0.03 Female 0.05*** 0.04*** 6-74 0.00 0.05*** 0ver-74s 0.05 0.03 6-74 0.00 0.05** 6-74 0.00 0.05** 0.00 0.05** 0.00** 0.00 0.05** 0.00** 6-0-14 0.00 0.05*** 0.00** 6-0-15 0.00 0.00*** 0.00*** 6-0-2 0.00 0.00** 0.00** 40-60 0.02 0.01 0.02** 60-80 0.02 0.01 0.02** 60-80 0.02 0.01 0.02** 60-90 0.02 0.01 0.02** 60-90 0.00 0.00** 0.00** | Tertiary education | 0.01 | 0.02 |
| 35-44 | | (0.02) | (0.01) |
| (0.02) (0.03) (0.02) (0.03) (0.02) (0.03) (0.02) (0.03) (0.02) (0.03) (0.02) (0.03) (0.02) (0.03) (0.02) (0.03) (0.02) (0.03) (0.02) (0.03) (0.02) (0.03) (0.02) (0.03) (0.03) (0.02) (0.03) (0.03) (0.03) (0.03) (0.03) (0.03) (0.03) (0.03) (0.03) (0.03) (0.03) (0.03) (0.03) (0.03) (0.03) (0.03) (0.03) (0.02) (0.03) (0.03) (0.02) (0.03) (0.03) (0.02) (0.03) | Age of the reference person | | |
| 45-54 | 35-44 | -0.01 | 0.00 |
| (0.03) (0.02) 55-64 | | (0.02) | (0.02) |
| 55-64 | 45-54 | -0.02 | -0.03 |
| (0.03) (0.02) 65-74 | | (0.03) | (0.02) |
| 65-74 0.05 0.00 Cover-74s 0.05 0.03 Gender of the reference person 0.05*** 0.04*** Female 0.05*** 0.01 Gross wealth percentile of the household (e) 0.00 -0.04* 20-40 0.00 -0.02 -0.02 40-60 0.02 -0.02 -0.02 60-80 0.02 -0.01 -0.02 -0.01 60-80 0.02 -0.01 -0.02 -0.01 80-90 -0.02 -0.02 -0.02 80-90 -0.02 -0.02 -0.02 90-100 0.00 0.00 0.03 | 55-64 | -0.03 | -0.06*** |
| Over-74s (0.03) (0.02) Gender of the reference person (0.06) (0.05) Female (0.02) (0.01) Forest wealth percentile of the household (e) 0.00 -0.04* 20-40 (0.03) (0.02) 40-60 (0.02) -0.02 60-80 (0.02) -0.02 60-80 (0.03) (0.02) 80-90 -0.02 -0.02 80-90 -0.02 -0.02 90-100 0.00 0.00 | | (0.03) | (0.02) |
| Over-74s 0.05 (0.06) 0.03 (0.05) Gender of the reference person Female 0.05*** (0.02) 0.04*** (0.02) Female 0.00 (0.02) 0.01 Gross wealth percentile of the household (e) 20-40 0.00 -0.04* (0.02) 40-60 0.02 -0.02 60-80 0.02 -0.01 60-80 0.02 -0.01 80-90 -0.02 -0.02 80-90 -0.02 -0.02 90-100 0.00 0.00 | 65-74 | 0.05 | 0.00 |
| (0.06) (0.05) Gender of the reference person 0.05*** 0.04*** Female 0.05*** 0.04*** Gross wealth percentile of the household (e) Union 0.00 -0.04* 20-40 0.00 0.03 (0.02) 40-60 0.02 -0.02 -0.02 60-80 0.02 -0.01 (0.03) (0.02) 80-90 -0.02 -0.02 -0.02 80-90 -0.02 -0.02 -0.02 90-100 0.00 0.00 0.03) | | (0.03) | (0.02) |
| Gender of the reference person Female 0.05*** (0.02) 0.04*** (0.02) Gross wealth percentile of the household (e) 0.00 -0.04* 20-40 0.00 (0.03) (0.02) 40-60 0.02 -0.02 (0.02) 60-80 0.02 -0.01 (0.03) (0.02) 60-80 0.02 -0.01 (0.03) (0.02) 80-90 -0.02 -0.02 (0.03) -0.02 (0.03) 90-100 0.003 (0.03) (0.03) | Over-74s | 0.05 | 0.03 |
| Female 0.05*** 0.04*** (0.02) (0.01) Gross wealth percentile of the household (e) 20-40 0.00 -0.04* (0.03) (0.02) 40-60 0.02 -0.02 60-80 0.02 -0.01 (0.03) (0.02) 80-90 -0.02 -0.02 90-100 0.03) (0.03) 90-100 0.00 0.00 | | (0.06) | (0.05) |
| (0.02) (0.01) Gross wealth percentile of the household (e) 0.00 -0.04* 20-40 0.00 (0.03) (0.02) 40-60 0.02 -0.02 (0.02) 60-80 0.02 -0.01 (0.03) (0.02) 80-90 -0.02 -0.02 -0.02 -0.02 -0.02 -0.02 -0.02 -0.02 -0.02 -0.02 -0.02 -0.02 -0.03 0.03) 90-100 0.00 | Gender of the reference person | | |
| (0.02) (0.01) Gross wealth percentile of the household (e) 0.00 -0.04* 20-40 0.00 (0.03) (0.02) 40-60 0.02 -0.02 (0.02) 60-80 0.02 -0.01 (0.03) (0.02) 80-90 -0.02 | Female | 0.05*** | 0.04*** |
| 20-40 0.00 -0.04* (0.03) (0.02) 40-60 0.02 -0.02 (0.02) (0.02) 60-80 0.02 -0.01 (0.03) (0.02) 80-90 -0.02 -0.02 (0.03) (0.03) 90-100 0.00 0.00 | | (0.02) | (0.01) |
| (0.03) (0.02) 40-60 0.02 -0.02 (0.02) (0.02) 60-80 0.02 -0.01 (0.03) (0.02) 80-90 -0.02 -0.02 (0.03) (0.03) (0.03) 90-100 0.00 0.00 | Gross wealth percentile of the household (e) | | |
| 40-60 0.02 -0.02 (0.02) (0.02) 60-80 0.02 -0.01 (0.03) (0.02) 80-90 -0.02 -0.02 (0.03) (0.03) 90-100 0.00 0.00 | 20-40 | 0.00 | -0.04* |
| (0.02) (0.02) 60-80 0.02 -0.01 (0.03) (0.02) 80-90 -0.02 -0.02 (0.03) (0.03) (0.03) 90-100 0.00 0.00 | | (0.03) | (0.02) |
| (0.02) (0.02) 60-80 0.02 -0.01 (0.03) (0.02) 80-90 -0.02 -0.02 (0.03) (0.03) (0.03) 90-100 0.00 0.00 | 40-60 | 0.02 | -0.02 |
| 60-80 0.02 -0.01 (0.03) (0.02) 80-90 -0.02 -0.02 (0.03) (0.03) 90-100 0.00 0.00 | | | (0.02) |
| (0.03) (0.02) 80-90 -0.02 -0.02 (0.03) (0.03) 90-100 0.00 0.00 | 60-80 | | , , |
| 80-90 -0.02 -0.02 (0.03) (0.03) 90-100 0.00 0.00 | | | (0.02) |
| (0.03) (0.03) 90-100 0.00 0.00 | 80-90 | | . , , |
| 90-100 0.00 0.00 | | | |
| | 90-100 | | |
| | | | |

SOURCE: Banco de España.

- a The effects are measured for the different sets of observations, using the five imputations relating to the EFF and all the editions.
- **b** The estimates are based on a logit model of the probability of a household having debt according to the EFF but not the CIR, conditional on the variables shown interacting with a variable denoting whether the observation is part of the comparable sample and using the population weights adjusted to the linkable sample.
- c Average marginal effects in terms of predicted probabilities.
- d The first column presents the effects under the condition that the household is part of the comparable sample, while the second column shows the effects for any type of household.
- e The income and wealth percentiles are defined vis-à-vis the overall distribution for all the editions.
- f Standard errors robust to heteroscedasticity are shown in brackets. Significance level at 1%, 5% and 10% as indicated by ***, ** and *, respectively.

CHARACTERISTICS OF THE HOUSEHOLDS AND THEIR DEBTS, AND PROBABILITY OF A HOUSEHOLD HAVING DEBT ACCORDING TO THE EFF BUT NOT THE CIR (a) (b) (cont'd)

The probability of a household having debt according to the EFF but not the CIR is higher when it only has non-mortgage debt as compared with households that have mortgage debt and debts under €6,000. The probability is higher in one-person households and where the reference person is female.

Decimals (c)

| | Comparable (d) | Total |
|--|----------------|----------|
| emographic characteristics | | |
| Total income percentile of the household | | |
| 20-40 | 0.01 | -0.03 |
| | (0.03) (f) | (0.03) |
| 40-60 | -0.03 | -0.05 |
| | (0.04) | (0.03) |
| 60-80 | -0.05 | -0.07*** |
| | (0.03) | (0.03) |
| 80-90 | -0.09** | -0.11*** |
| | (0.04) | (0.03) |
| 90-100 | -0.04 | -0.07** |
| | (0.04) | (0.03) |
| Number of adults in the household (e) | | |
| Two | -0.05** | -0.03 |
| | (0.03) | (0.02) |
| Three | -0.10*** | -0.05** |
| | (0.03) | (0.02) |
| More than three | -0.11*** | -0.06** |
| | (0.03) | (0.02) |
| At least one self-employed worker | 0.03 | 0.01 |
| | (0.02) | (0.02) |
| Edition | | |
| 2005 | -0.09*** | -0.06*** |
| | (0.02) | (0.02) |
| 2008 | -0.07*** | -0.10*** |
| | (0.02) | (0.02) |
| 2011 | -0.06** | -0.08*** |
| | (0.02) | (0.02) |
| Comparable observation | | -0.03** |
| | | (0.01) |
| bservations | 3,936 | 7,904 |
| ependent variable average | 0.18 | 0.20 |

SOURCE: Banco de España.

- a The effects are measured for the different sets of observations, using the five imputations of the EFF and all the editions.
- b The estimates are based on a logit model of the probability of a household having debt according to the EFF but not the CIR, conditional on the variables shown interacting with a variable denoting whether the observation is part of the comparable sample and using the population weights adjusted to the linkable sample.
- c Average marginal effects in terms of predicted probabilities.
- d The first column presents the effects under the condition that the household is part of the comparable sample, while the second column shows the effects for any type of household.
- e The number of adults in the household is that reflected in the EFF.
- f Standard errors robust to heteroscedasticity are shown in brackets. Significance level at 1%, 5% and 10% as indicated by ***, ** and *, respectively.

As a robustness analysis, we included in these regressions other factors that could explain the differences between EFF/CIR indebtedness, although these findings are not shown in the paper.¹¹ First, it is important to highlight that the CIR records the position of credit institutions' risks at the end of the month, whereas the EFF reflects the household's level of indebtedness at the time of the interview. Therefore, this difference in the reference period could help explain some of the discrepancies found between the two sources. In this regard, we found that the probability of a comparable household having debt according to the EFF but not the CIR was 5.3 pp lower if the interview was conducted in the last week of the month and 4.7 pp lower if it was conducted in the third week than in the first. However, the effect of the other demographic and debt factors does not change significantly when adding this variable. In relation to the other possible discrepancy between the two sources, we found that the week of the interview had no significant impact on the probability of a household not reporting in the survey the debts recorded on the CIR.

We also explored other factors that could affect a household's response during the interview. One was household availability to answer the survey. We constructed indicators to control for the fact that the household works more than 45 hours a week or has children aged under 16, but we did not find that these had a significant impact on the findings for the comparable sample. Another factor that could be a possible source of discrepancies would be having children aged 25 and above. 12 However, the estimates show that this factor does not significantly affect any of the above-mentioned findings. A further variable that we explored is whether there are children aged 45 and over who are members of the household, since the survey respondent may be unaware of their debts. This could therefore affect the probability of a household debt featuring on the CIR but not in the EFF. Once again, the findings are robust to the inclusion of this variable.

It is also worth considering that some EFF respondents may be confused as to the type of institution with which a debt has been arranged. Therefore, there may be debts reported as arranged with non-financial corporations that have in actual fact been arranged with credit institutions. However, including debts arranged with non-financial corporations as comparable debts between the CIR and the EFF has no impact on the findings. Along the same lines, debts arranged for the purchase of the fourth property may also not be comparable, as we do not have accurate information on the type of institution with which they have been arranged. To test the robustness of the findings to this factor, we eliminated these debts from the comparison. This did not affect the findings either.

Furthermore, while some households could be linked to the CIR as we have the national identity card number of at least one of the registered people, they could not be included in this analysis because they also comprise people registered using only a passport or an unknown ID. To study how this affects the findings, we constructed indicators for

¹¹ These findings are available from the authors upon request.

¹² For example, because these children are yet to register at their main residence or because the parents confuse their children's debts with their own.

Chart 1

DISTRIBUTION OF THE DISCREPANCY IN THE PERCENTAGE OF INDEBTED HOUSEHOLDS, BY AGE GROUP

This chart depicts the breakdown, by age group, of the difference between the percentage of indebted households according to the CIR and the EFF for the comparable sample. For the percentage of indebted households, we observe that the households aged 65-74 are those which make the biggest contribution to the discrepancy between the CIR and the EFF. However, for the percentage of mortgagor households, households aged 55-64 make the main contribution.



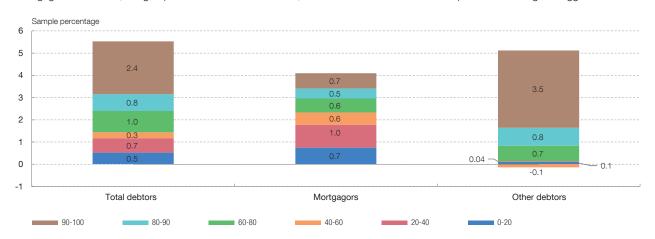
SOURCE: Banco de España.

households with at least one person registered using an unknown ID or a passport. Using the total sample, drawing on the same socio-demographic and debt factors as in Tables 6 and 7, we found that the probability of a household debt featuring on the CIR but not in the EFF was 7 pp lower for households with at least one person registered with an unknown ID. We also found that the probability of a household debt featuring in the EFF but not on the CIR was 7 pp lower for those households where at least one person is registered with an unknown ID. However, the effects of the other demographic and debt factors do not change with the inclusion of these new variables. Lastly, households having people registered with a passport has no significant impact on the discrepancies between the two sources.

Finally, to ascertain which groups of households contribute the most to the differences in the percentage of indebted households between the CIR and EFF, we broke down that difference linearly as the sum of the differences weighted by population groups defined using a specific characteristic, i.e. $d=\sum_g \pi_g d_g$, where d is the total difference, d_g is the difference in the group of households defined by the characteristic g and π_g is the proportion of the comparable sample that this group of households represents. The characteristics we analysed separately were age, total household income, gross wealth and the edition to which they belong because these were the most relevant characteristics. This enabled us to describe certain groups of households' participation in, or contribution to, the total discrepancy, $\pi_g d_g$. Chart 1 depicts the breakdown by age group. Indeed, households whose reference person belongs to the highest age groups contribute more to the discrepancy in the percentage of total indebted households, although it is the 65-74 age group that makes the biggest contribution (2.3 pp). In the discrepancies for other debt, this is also the group that makes the biggest contribution (1.9 pp); however, in the mortgage debt discrepancies it is the 55-64 age group that makes the biggest contribution (1.4 pp).

Chart 2 DISTRIBUTION OF THE DISCREPANCY IN THE PERCENTAGE OF INDEBTED HOUSEHOLDS, BY TOTAL INCOME GROUP

This chart depicts the breakdown, by income group, of the difference between the percentage of indebted households according to the CIR and the EFF for the comparable sample. For the percentage of total indebted households, we observe that the households that are above the 90th income percentile are those which make the biggest contribution to the discrepancy between the CIR and the EFF. However, for the percentage of mortgagor households, the groups make similar contributions, with the households in 20th-40th percentile making the biggest contribution.

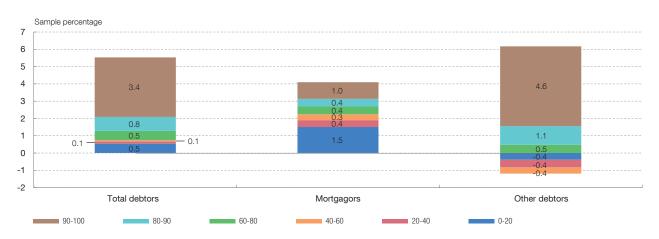


SOURCE: Banco de España.

Chart 3

DISTRIBUTION OF THE DISCREPANCY IN THE PERCENTAGE OF INDEBTED HOUSEHOLDS, BY GROSS WEALTH GROUP

This chart depicts the breakdown, by gross wealth group, of the difference between the percentage of indebted households according to the CIR and the EFF for the comparable sample. For the percentage of total indebted households, we observe that the households that are above the 90th wealth percentile are those which make the biggest contribution to the discrepancy between the CIR and the EFF. However, for the percentage of mortgagor households, the biggest contribution is made by the 0-20th percentile group.

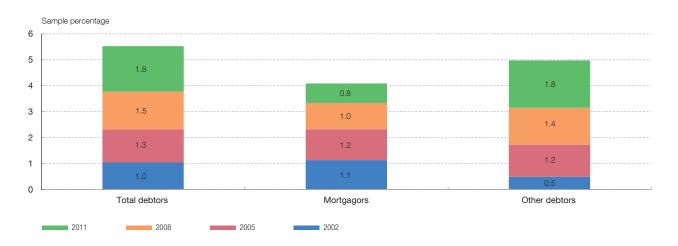


SOURCE: Banco de España.

Chart 2 depicts the breakdown by income group. In this case, higher income groups make a greater contribution to the total discrepancy. However, as shown above, their marginal contribution was smaller. This is because they are the most oversampled households. Specifically, their contribution to the total discrepancy in other debt stands out (3.5 pp). In Chart 3 we show the breakdown by gross wealth group. The findings

Chart 4 DISTRIBUTION OF THE DISCREPANCY IN THE PERCENTAGE OF INDEBTED HOUSEHOLDS, BY EDITION

This chart depicts the breakdown of the difference between the percentage of indebted households according to the CIR and the EFF in the different editions for the comparable sample. For the total percentage, we observe that households interviewed in 2011 are those which make the biggest contribution to this discrepancy. However, for the percentage of mortgagor households, the biggest contribution to this discrepancy is that of the group of households interviewed in 2002.



SOURCE: Banco de España.

also indicate that the highest wealth group makes the biggest contribution to the other debt discrepancy (4.6 pp); however, it is the lowest wealth group that makes the biggest contribution to the discrepancy in mortgage debt (1.5 pp). In this case we know that this is not the most oversampled group. Consequently, its contribution has more to do with the fact that the discrepancy in the percentage of mortgage debt between the CIR and the EFF is greater in lower income households. Lastly, the discrepancies by edition are less noteworthy (see Chart 4). However, we observed that, while the probability of discrepancy was greater in the 2002 edition, overall there were fewer observations in that edition than in subsequent ones. This is because the EFF's sample size has grown over time.

5 Indebted households' debt level

In this section we study the size of the differences between indebted households' outstanding debt in the EFF and the CIR for total debt and by debt type. Table 8 presents some percentiles of the distribution of these differences. The negative discrepancies measure debt amounts that are not reflected in the EFF but are in the CIR, while the positive discrepancies denote those which are reflected in the EFF but not in the CIR. Note again that when we use the population weights the EFF/CIR differences are smaller. For the total sample, the unweighted median discrepancy for any type of debt is -€3,700, whereas it is -€2,500 when weighted. The unweighted 10th and 90th percentiles are -€131,800 and €38,900, respectively. When weighted by the households' population weight, the values for these percentiles decrease to -€90,600 and €30,600, respectively. The results could suggest that the EFF understates the outstanding debt of indebted households according to the two sources in the total sample. The third row for each type of debt presents the resulting statistics for the comparable sample. We observed that, compared with the total sample, the degree of asymmetry decreases and there is less probability mass for negative values of the distribution (i.e. the number of cases where the debt reflected in the EFF is lower than that according to the CIR falls). In the comparable sample the median difference for any debt

Table 8 **DEBT DIFFERENCES BETWEEN THE EFF AND THE CIR (a) (b)**

In more than half of the sample, outstanding debt is higher according to the CIR, and this is observed in all debt types. The distribution of the differences is skewed to the left of zero. In the comparable sample, the discrepancies are smaller and the distribution is more centred around the median. At population level, the differences are also smaller.

Thousands of euro

| | 10th percentile | 25th percentile | Median | 75th percentile | 90th percentile |
|-------------------|-----------------|-----------------|--------|-----------------|-----------------|
| All debt | | | | | |
| Population (c) | -90.6 | -28.9 | -2.5 | 4.4 | 30.6 |
| Total sample | -131.8 | -41.4 | -3.7 | 4.6 | 38.9 |
| Comparable sample | -80.8 | -21.2 | -2.0 | 4.0 | 29.5 |
| Mortgage debt | | | | | |
| Population | -66.5 | -15.0 | -2.5 | 4.4 | 30.6 |
| Total sample | -97.9 | -23.0 | -3.7 | 4.6 | 38.9 |
| Comparable sample | -57.2 | -12.0 | -2.0 | 4.0 | 29.5 |
| Non-mortgage debt | | | | | |
| Population | -31.9 | -10.8 | -2.3 | 1.3 | 8.6 |
| Total sample | -44.5 | -12.3 | -2.2 | 2.0 | 11.1 |
| Comparable sample | -36.0 | -10.0 | -1.5 | 1.8 | 10.5 |

SOURCE: Banco de España.

- a The differences are the discrepancy between the outstanding amounts according to the EFF and the CIR in the interview month.
- ${f b}$ The differences are calculated for households that have debt according to both the EFF and the CIR.
- c "Population" refers to the total linkable sample weighted by the adjusted population weights. "Total sample" refers to the total number of linkable household observations, and "Comparable sample" refers to the observations for which the difference between household members and members registered at that household's address is zero or matches the number of children residing in the household.

RELATIVE DIFFERENCES IN DEBT AMOUNTS BETWEEN THE EFF AND THE CIR (a) (b)

In more than half the sample, outstanding debt is higher according to the CIR. This is observed for all types of debt. The differences are relatively bigger for non-mortgage debt than for mortgage debt, where the difference is barely 1% of the amount according to the CIR. The distribution of the relative differences is skewed to the left of zero. In the comparable sample, the discrepancies are 0.5% for the median, and the 10 and 90th percentiles are more similar in absolute terms than where we use the total sample. At population level, the differences are also smaller.

| % | 10th percentile | 25th percentile | Median | 75th percentile | 90th percentile |
|-------------------|-----------------|-----------------|--------|-----------------|-----------------|
| All debt | | | | | |
| Population (c) | -86.7 | -50.7 | -6.5 | 12.3 | 100.0 |
| Total sample | -87.8 | -54.6 | -8.3 | 12.0 | 100.4 |
| Comparable sample | -75.8 | -36.7 | -3.8 | 9.4 | 83.7 |
| Mortgage debt | | | | | |
| Population | -62.2 | -25.0 | -0.6 | 17.8 | 102.6 |
| Total sample | -69.8 | -30.3 | -1.0 | 19.1 | 103.9 |
| Comparable sample | -56.3 | -18.4 | -0.5 | 11.3 | 94.6 |
| Non-mortgage debt | | | | | |
| Population | -89.4 | -64.6 | -20.0 | 11.2 | 99.0 |
| Total sample | -89.6 | -64.3 | -18.1 | 13.4 | 107.8 |
| Comparable sample | -85.6 | -54.9 | -10.2 | 12.1 | 95.1 |

SOURCE: Banco de España.

- a The differences are the outstanding amounts according to the EFF minus those according to the CIR, divided by the level of debt according to the CIR for the corresponding type of debt.
- b The differences are calculated for households that have debt according to the EFF and the CIR.
- c "Population" refers to the total linkable sample weighted by the adjusted population weights. "Total sample" refers to the total number of linkable household observations, and "Comparable sample" refers to the observations for which the difference between household members and members registered at that household's address is zero or matches the number of children residing in the household.

type is -€2,000 (compared with -€3,700 in the total sample) and the 10th and 90th percentiles are -€80,800 and €29,500 (-€131,800 and €38,900 in the total sample), respectively. In other words, by excluding from the comparison the households with discrepancies vis-àvis the municipal population register, the differences between the outstanding debt in the two sources decrease. This was also the case when comparing the percentage of indebted households according to the two sources.

Turning to mortgage debt, the distribution of the discrepancy for the comparable sample is more symmetric around zero, the 10th and 90th percentiles (-€57,200 and €29,500) and the median is -€2,000. With regard to non-mortgage debt, the median for the total sample is very similar to the median for mortgage debt. At the ends of the distribution, smaller differences are observed than in those for mortgage debt. However, for the comparable sample the difference between the EFF and the CIR is notably smaller: the median discrepancy is -€1,500 and the 10th and 90th percentiles are -€36,000 and €10,500, respectively. These findings confirm that mortgage debt discrepancies tend to be greater than non-mortgage debt ones in absolute terms. This may be due to mortgage debts also tending to be larger than non-mortgage ones.

HOUSEHOLD CHARACTERISTICS AND RELATIVE DISCREPANCY BETWEEN OUTSTANDING DEBT ACCORDING TO THE EFF AND THE CIR (a) (b)

The results for the first and second quartiles show that households where the reference person is 55 and above, that belong to the 0-20th wealth percentile and that have more than two members have less outstanding debt according to the EFF than the CIR.

Marginal effects

| viai gii iai enects | (1) | | (3) |
|---|--------------|---------------------|--------------|
| | 1st quartile | (2) 2nd quartile | 3rd quartile |
| Educational attainment level of the reference person (baseline = primary education) | - 4 | | |
| Secondary education | -0.06* | -0.02 | -0.04 |
| • | (0.04) (d) | (0.02) | (0.04) |
| Tertiary education | 0.00 | -0.01 | -0.03 |
| | (0.03) | (0.02) | (0.04) |
| Age of the reference person (baseline = under 35) | | | |
| 35-44 | -0.01 | 0.00 | 0.05 |
| | (0.04) | (0.04) | (0.04) |
| 45-54 | -0.12** | -0.02 | 0.02 |
| | (0.04) | (0.04) | (0.04) |
| 55-64 | -0.22*** | -0.05* | 0.03 |
| | (0.04) | (0.04) | (0.04) |
| 65-74 | -0.25*** | -0.07** | 0.13* |
| | (0.04) | (0.04) | (0.04) |
| Over-74s | -0.01 | 0.03 | 0.36*** |
| | (0.04) | (0.04) | (0.04) |
| Gender of the reference person (baseline = male) | | | |
| Female | 0.03 | 0.00 | 0.03 |
| | (0.03) | (0.01) | (0.03) |
| Gross wealth percentile of the household (c) | | | |
| 20-40 | 0.31*** | 0.27*** | 0.04 |
| | (0.06) | (0.06) | (0.07) |
| 40-60 | 0.37*** | 0.28*** | 0.07 |
| | (0.06) | (0.05) | (0.07) |
| 60-80 | 0.34*** | 0.27*** | 0.02 |
| | (0.06) | (0.06) | (0.07) |
| 80-90 | 0.35*** | 0.28*** | 0.04 |
| | (0.07) | (0.06) | (0.08) |
| 90-100 | 0.43*** | 0.32*** | 0.12 |
| | (0.07) | (0.05) | (0.08) |
| Total income percentile of the household | | | |
| 20-40 | 0.06 | 0.03 | 0.01 |
| | (0.08) | (0.04) | (0.08) |
| 40-60 | 0.07 | 0.01 | 0.00 |
| | (0.08) | (0.03) | (0.08) |
| 60-80 | 0.08 | 0.01 | -0.01 |
| | (0.07) | (0.03) | (0.08) |
| 80-90 | 0.11 | 0.01 | -0.02 |
| | (0.08) | (0.04) | (0.09) |
| 90-100 | 0.09 | 0.03 | 0.03 |
| | (80.0) | (0.04) | (0.09) |

SOURCE: Banco de España.

- a The effects are measured for all households that have debt according to both the CIR and the EFF, using the five imputations and the population weights.
- b The estimates are based on a quantile regression model where the dependent variable is outstanding debt according to the EFF less outstanding debt according to the CIR divided by outstanding debt according to the CIR. Columns (1), (2) and (3) show the results for the first, second and third quartiles, respectively.
- **c** The income and wealth percentiles are defined vis-à-vis the overall distribution for all the editions.
- d Standard errors robust to heteroscedasticity are shown in brackets. Significance level at 1%, 5% and 10% as indicated by ***, ** and *, respectively.

2008

2011

Constant

Observations

HOUSEHOLD CHARACTERISTICS AND RELATIVE DISCREPANCY BETWEEN OUTSTANDING DEBT ACCORDING TO THE EFF AND THE CIR (a) (b) (cont'd)

The results for the first and second quartiles show that households where the reference person is 55 and above, that belong to the 0-20th wealth percentile and that have more than two members have less outstanding debt according to the EFF than the CIR.

| Marginal effects | | | |
|--|--------------|--------------|--------------|
| via gina onoco | (1) | (2) | (3) |
| | 1st quartile | 2nd quartile | 3rd quartile |
| Number of adults in the household (baseline = one) (c) | | | |
| Two | -0.08* | -0.03 | 0.00 |
| | (0.05) (d) | (0.02) | (0.05) |
| Three | -0.26*** | -0.08*** | -0.03 |
| | (0.05) | (0.03) | (0.06) |
| More than three | -0.27*** | -0.13*** | -0.08 |
| | (0.06) | (0.03) | (0.07) |
| Self-employed | -0.03 | -0.02 | 0.01 |
| | (0.03) | (0.02) | (0.04) |
| Edition (baseline = 2002) | | | · |
| 2005 | -0.01 | -0.01 | -0.07 |
| | (0.04) | (0.02) | (0.05) |

0.03

(0.04)

0.10**

(0.04)

-0.58**

3.198

0.01

(0.02)

0.03

(0.02)

-0 29**

3.198

-0.05

(0.05)

-0.05

(0.05)

0.07

3.198

SOURCE: Banco de España.

- a The effects are measured for all households that have debt according to both the CIR and the EFF, using the five imputations and the population weights.
- b The estimates are based on a quantile regression model where the dependent variable is outstanding debt according to the EFF less outstanding debt according to the CIR divided by outstanding debt according to the CIR. Columns (1), (2) and (3) show the results for the first, second and third quartiles, respectively.
- c The number of adults in the household is that reflected in the EFF.
- d Standard errors robust to heteroscedasticity are shown in brackets. Significance level at 1%, 5% and 10% as indicated by ***, ** and *, respectively.

To account for this, we calculated the differences relative to outstanding debt according to the CIR. Once again, the negative discrepancies are debt amounts not reflected in the EFF but which are in the CIR, while the positive discrepancies denote those which are reflected in the EFF but not in the CIR. Table 9 presents some percentiles of the distribution of the relative discrepancies. In the comparable sample the median difference between the debt amounts in the EFF and the CIR is -3.8%. By comparing the median and the interquartile ranges of the distributions of the differences in the comparable sample for mortgage and non-mortgage debt, we see that the differences are larger in relative terms for non-mortgage debt. The first quartile of the differences in mortgage debt is -18.4%, while the third quartile is 11.3% and the median difference is only -0.5%. However, in non-mortgage debt the first quartile of the differences is -54.9% in relative terms, while the third quartile is 12.1% and the median is -10.2%. Lastly, we note that in 141 households in the comparable sample between 2002 and 2011 the difference between total debt measured via the EFF and the CIR was zero, while for mortgage debt the differences were zero in 146 households and for non-mortgage debt they were zero in 64 households. Table 10 shows

the marginal effects of demographic characteristics corresponding to quantile regressions of the relative discrepancy between outstanding debt amounts according to the EFF and the CIR. Column 1 presents the results of the regression for the first quartile. Households whose reference person is 55 or older, that belong to the 0-20th wealth percentile and that have more than two members have less outstanding debt in the EFF than according to the CIR. These findings are in line with those shown in Table 6. Therefore, we found that older, less wealthy groups are more likely to report fewer or lower debts. These findings are qualitatively similar to those for the second quartile (column 2 of Table 10), while they are not applicable to the third quartile given that the third quartile of the relative error relates to positive values, i.e. to higher debt amounts according to the EFF than to the CIR, as described in Table 9.

6 Conclusions

In this paper we compare the data on debt collected in the EFF with those according to the CIR. A significant difficulty in linking the two datasets is the existence of differences between the number of members that make up the household in the EFF and the number of people registered at the household's address. To control for these differences, we construct a comparable sample of households by selecting those observations with the same number of adult members in the EFF as in the municipal population register.

For the comparable sample, we find that the difference in the sample percentage of indebted households between the EFF and the CIR is 5.5 pp. This mainly stems from a greater difference in the percentage of households with non-mortgage debt, which is 5 pp higher according to the CIR, while the percentage of households with mortgage debt is 4.1 pp higher. However, all these sample differences appear to diminish at population level.

In addition, we present an analysis of the differences in the probability of a household debt featuring in the EFF and on the CIR on the basis of demographic and debt characteristics. The results of this analysis in the subset of comparable households reveal that the probability of a household debt featuring on the CIR but not in the EFF is higher for households that only have non-mortgage debt. This probability is also higher for those households with debts shared with people who are not members of the household. In terms of the effects of the demographic characteristics, we find that the households where the reference person in the survey is 45 and above have a higher probability of being indebted according to the CIR but not the EFF. Furthermore, we find that where the reference person is female the probability of being indebted according to the CIR but not the EFF is also higher. The wealthiest households have a lower probability of their debt featuring on the CIR but not being reported in the EFF. Lastly, we note that this probability is higher in the first edition of the EFF than in subsequent ones. Also, households with only non-mortgage debt have a higher probability of their debt being reported in the EFF but not featuring on the CIR than those with only mortgage debt, although this is less prevalent than in the case of the former probability. In turn, this probability is higher for households with outstanding debt amounting to less than €6,000. However, this comes as little surprise given that the CIR does not reflect this type of debt. In terms of the household's demographic characteristics, we find that the probability is higher where the reference person is female. This probability is also higher in one-person households.

We can conclude that once the discrepancies between the number of household members and the number of people registered at that address on the municipal population register have been controlled for, the EFF correctly reflects the proportion of indebted households and, largely, their outstanding debt. Also, we can learn a series of lessons to better measure household debt in future editions of the EFF. It would clearly be well worth collecting the national identity card numbers of the respondent household members. Lastly, the EFF could also benefit from including protocols aimed at obtaining more accurate information when the respondent is older.

Annex 1 Methodology for linking the Spanish Survey of Household Finances and the Central Credit Register

To link the information at household level in the Spanish Survey of Household Finances ("EFF" by its Spanish initials) to the information at individual level according to the Central Credit Register ("CIR" by its Spanish initials), we used the encrypted national identity card numbers of the individuals registered at the addresses of the households interviewed for the EFF. The process for obtaining the encrypted national identity card numbers has two stages. First, the National Statistics Institute ("INE" by its Spanish initials) links the sample design information on the households interviewed in the EFF with the municipal population register information.^{1, 2, 3} This is possible under the National Statistics Plan ("PEN" by its Spanish initials), to which the EFF belongs. Second, these personal identification numbers are encrypted in a cloud via an automatic procedure following the same protocols used by the CIR. Under these conditions, at no point do the Banco de España analysts have access to the original identifying data. As a result, in accordance with the PEN, those in charge of the EFF at the Banco de España receive the encrypted identification numbers for the inhabitants registered at the addresses of the households interviewed for the EFF.

However, this procedure has several limitations. First, at present the INE does not have the municipal population registers for households in the Basque Country and Navarre for the 2002 and 2005 editions. 271 and 373 households could not be linked in the 2002 and 2005 editions, respectively, for this reason. Second, valid municipal population register data are also unavailable for those addresses where, at the time of the interview, the panel households interviewed in prior editions no longer resided and new households were interviewed instead. 355, 636 and 374 households could not be linked in the 2005, 2008 and 2011 editions, respectively, for this reason. As these two groups of households needed to be eliminated, we adjusted the population weights provided by the EFF to ensure the representativeness of the sample of households vis-à-vis the population.

Third, we found some cases of people registered at two addresses in the same edition. This problem affected eight, six and three pairs of households in 2005, 2008 and 2011, respectively. Most of the cases were refreshment households that shared identifiers with panel households. To avoid this type of inconsistency, for the 2005, 2008 and 2011 editions we retained the most recently selected household in the sample, as we considered the related municipal population register data more up to date, and we removed the panel household. Overall, 15 observations were removed from the four editions. In turn, we found two refreshment households that shared identifiers of registered members. In these

¹ This is performed for each of the sample addresses selected, linking them to the existing municipal population register information at the time those addresses were first selected to participate in the EFF.

² The EFF does not directly collect the households' identifying data.

³ Since 1996 the data on registered inhabitants available at the INE have been based on municipal population register data, which are updated with the information on new entries and changes of residence sent regularly by the local municipal authorities. In turn, the INE rectifies errors in the identification card numbers by cross-checking them against the numbers in the National Police Corps' database.

cases, we did not remove either household from the sample since we had no clear way of ascertaining which of them provided more accurate information. Furthermore, we checked for the existence of households where some of the identification numbers were repeated within the household. This could be due to minors registered without their own national identity card number. Therefore, we used this list of identifiers to count the number of registered individuals within the household, but removed the duplicate identity card numbers within the household so that debts were not duplicated when linking the data to the CIR.

Lastly, a further limitation on this analysis was that, in some cases, information such as "unknown" or "passport" appears in the field corresponding to the national identity card number, which prevents CIR linkage. This issue affected 78 households in 2008 and 96 in 2011. Therefore, it is possible that for some households the number of total debts according to the CIR is understated since it was not possible to link the two data sources for all individuals registered at the household; however, in Section 4, we have shown that this does not significantly affect the findings shown in Tables 6 and 7.

Annex 2 Analysis of discrepancies with the municipal population register

This annex presents the results of the analysis of the correlation between a household's socio-demographic characteristics and the discrepancies observed between the number of household members at the time of responding to the EFF and the number of registered individuals. Specifically, for each household we defined three different binary variables to characterise this type of discrepancy:

- Whether the number of household members is equal to the number of people registered at that address on the municipal population register.
- Whether the discrepancy between the number of household members and the number of people registered at that address on the municipal population register amounts to just one person.
- Whether the discrepancy between the number of household members and the number of people registered at that address on the municipal population register coincides with the number of minors under 16 who are members of the household.

Table A2.1 shows the results of the analysis of each of the three binary variables. Here we describe the results corresponding to column 1 referring to the probability of household members and the number of people registered at that address on the municipal population register matching exactly. The probability of them matching is higher for households in the refreshment sample; a household in the refreshment sample has a 20 pp higher probability of matching the municipal population register. The probability of the number of household members reported in the survey matching the number of people registered at the household's address decreases as the number of adults reflected in the survey increases. There is a lower probability of matching the municipal population register when the reference person's or their partner's children are no longer members of the household (a 13 pp lower probability). The older the reference person, the higher the probability of matching the municipal population register. In turn, the probability of the number of household members reported in the survey matching the number of people registered at that address is 5 pp lower when the reference person is female. Furthermore, compared with those who rent, the probability of household members matching in the two sources is 3 pp higher among those who own their main residence and 6 pp higher among those who use it for free.

The probability of a discrepancy between the number of household members and the number of people registered at that address has fallen with each edition of the EFF.

Table A2.1

CHARACTERISTICS OF THE HOUSEHOLDS AND THEIR DEBTS, AND PROBABILITY OF A DISCREPANCY BETWEEN THE NUMBER OF HOUSEHOLD MEMBERS AND THE NUMBER OF PEOPLE REGISTERED AT THAT ADDRESS (a) (b) (c)

The probability that the number of household members and the number of people registered at the household's address match is higher the older the reference person, if the reference person is male and in smaller households. The property being owned or used for free increases the probability of the numbers matching. The probability is also higher for refreshment households, since their data are drawn from a municipal population register closer to the time of the EFF's field work. With each edition, the two figures have drawn nearer.

| | | | (3) |
|--|------------|---------------|----------------------|
| | Match | Differ by one | Differ due to minors |
| Age of the reference person (baseline = under 35) | | | |
| 35-44 | -0.04** | -0.07*** | 0.08*** |
| | (0.02) (e) | (0.02) | (0.02) |
| 45-54 | 0.07*** | 0.10*** | 0.01 |
| | (0.02) | (0.02) | (0.02) |
| 55-64 | 0.14*** | 0.14*** | -0.02 |
| | (0.02) | (0.02) | (0.02) |
| 65-74 | 0.19*** | 0.15*** | 0.01 |
| | (0.02) | (0.02) | (0.02) |
| Over-74s | 0.32*** | 0.24*** | 0.16*** |
| | (0.02) | (0.02) | (0.02) |
| Gender of the reference person (baseline = male) | | | |
| Female | -0.05*** | -0.04*** | -0.06*** |
| | (0.01) | (0.01) | (0.01) |
| Number of adults in the household (baseline = one) (d) | | | |
| Two | -0.15*** | -0.06*** | -0.02 |
| | (0.02) | (0.01) | (0.02) |
| Three | -0.19*** | -0.12*** | -0.20*** |
| | (0.02) | (0.02) | (0.02) |
| More than three | -0.21*** | -0.20*** | -0.24*** |
| | (0.02) | (0.02) | (0.02) |
| Children that no longer reside in the household | -0.13*** | -0.08*** | -0.16*** |
| · · | (0.02) | (0.02) | (0.02) |
| enure status of the main residence (baseline = rented) | | , | , |
| Owned | 0.03* | 0.02 | 0.09*** |
| | (0.02) | (0.02) | (0.02) |
| Free use | 0.06** | 0.05* | 0.11*** |
| | (0.03) | (0.03) | (0.03) |
| Other | -0.03 | -0.16* | 0.03 |
| | (0.07) | (0.08) | (0.08) |
| Refreshment household | 0.20*** | 0.16*** | 0.14*** |
| | (0.01) | (0.01) | (0.01) |
| Edition (baseline = 2002) | , | , , | , |
| 2005 | 0.09*** | 0.09*** | 0.06*** |
| | (0.01) | (0.02) | (0.02) |
| 2008 | 0.11*** | 0.12*** | 0.04** |
| | (0.01) | (0.02) | (0.02) |
| 2011 | 0.18*** | 0.14*** | 0.09*** |
| - | (0.02) | (0.02) | (0.02) |
| Dbservations | 21,396 | 21,396 | 21,396 |
| Dependent variable average | 0.41 | 0.72 | 0.53 |

SOURCE: Banco de España.

- a The effects are measured for the total sample, using the five imputations and the population weights.
- **b** Average marginal effects, i.e. conditional on all variables using the average value for the sets of observations.
- c The estimates are based on a logit model of the probability of the household having the same number of members as the number of people registered at that address (1); of the numbers matching or of the difference being +/- 1 (2); and, of the numbers being the same or of the differences matching the number of minors under 16 (3).
- d The number of adults in the household is that reflected in the EFF.
- e Standard errors robust to heteroscedasticity are shown in brackets. Significance level at 1%, 5% and 10% as indicated by ***, ** and *, respectively.

Annex 3 Basic definitions of the characteristics of the risks in the Central Credit Register

The risks relating to each debtor and the risk's origin (institution and country) are presented together based on five characteristics or risk keys: class, currency, maturity, collateral, and status. These risk keys are defined as follows:

- Class: this defines the type of risk. It is classified into the values A-S and X. Among these risks, those that are comparable with the debts reflected in the Survey of Household Finances (EFF by its Spanish initials) are "Finance credit" (B); "Finance lease transaction" (K); and "Loans or credit transferred to third parties" (Q). Most households have category B debts. Figure A3.1 shows how the classes of risk defined in the Central Credit Register (CIR by its Spanish initials) relate to the types of debt in the EFF.
- Currency: this determines the risk's currency. It is categorised into values A-T, but all the risks relating to household debts in the EFF are denominated in euro.
 Therefore, this category does not change in our analysis.
- Maturity: this variable reflects the average maturity of the risks in categories A-E and M. "A" refers to an average risk maturity of up to 3 months; "B", 3 to 6 months; "C", 1 to 3 years; "D", 3 to 5 years; "E" more than 5 years; and "M" an unspecified maturity. The average maturity is the arithmetic mean weighted by the capitals sharing the same risk keys of the maturities at the risks' origination. These maturities are not directly comparable with those reflected jointly by the questions in the EFF on the years for which the loans for the purchase of the main residence (p2_16) and other properties (p2_59), and other types of debt (p3_9) were applied for since the CIR risks can combine more than one debt per risk key.
- Collateral: this variable reflects the existence and types of collateral for the risks. It is divided into the categories A-H and V. These categories are different from those used in the EFF, except for category A (fully collateralised by Treasury bills, real estate properties or ship mortgages, deposits, securities with official market prices and goods) and B (fully collateralised by assets other than those included under A), which are comparable with categories 1 and 2 of the questions on the type of loan arranged for the purchase of the main residence (p2_9), acquisition of other properties (p2_52) and other types of debt (p3_2).
- Status: this refers to the default status of the risk, classified as A-L. At origination all risks are categorised as A and are then reclassified based on the level of default. "B": performing in arrears; "C": non-performing for reasons other than customer arrears; "D": non-performing past-due for up to three months; "E": non-performing past-due for between three and six months; "F": non-performing past-due for between three and six months.

Figure A3.1

CIR RISKS BY TYPES OF DEBT IN THE EFF

Finance credit (B) and Loans or credit transferred to third parties (Q)

- Mortgage loan
- Other collateralised loans
- Personal loan
- Credit line
- Instalment loan
- Current account overdrafts
- Advances
- Reverse mortgage
- Credit card debt

Finance lease transactions (K)

- Leasing or renting

SOURCE: Banco de España.

performing past-due for between six and 12 months; "G": non-performing pastdue for between 12 and 18 months; "H": non-performing past-due for between 18 and 21 months; "I": non-performing past-due for more than 21 months; "J": suspended; "K": rediscounted bills; "L": loan pending a creditors' agreement. The EFF does not collect information on the default status of debts. This risk characteristic cannot therefore be used for the comparison.

BANCO DE ESPAÑA PUBLICATIONS

OCCASIONAL PAPERS

- 2030 ÁNGEL GÓMEZ-CARREÑO GARCÍA-MORENO: Juan Sebastián Elcano 500 años de la Primera vuelta al mundo en los billetes del Banco de España. Historia y tecnología del billete.
- 2031 OLYMPIA BOVER, NATALIA FABRA, SANDRA GARCÍA-URIBE, AITOR LACUESTA and ROBERTO RAMOS: Firms and households during the pandemic: what do we learn from their electricity consumption?
- 2032 JÚLIA BRUNET, LUCÍA CUADRO-SÁEZ and JAVIER J. PÉREZ: Contingency public funds for emergencies: the lessons from the international experience. (There is a Spanish version of this edition with the same number).
- 2033 CRISTINA BARCELÓ, LAURA CRESPO, SANDRA GARCÍA-URIBE, CARLOS GENTO, MARINA GÓMEZ and ALICIA DE QUINTO: The Spanish Survey of Household Finances (EFF): description and methods of the 2017 wave.
- 2101 LUNA AZAHARA ROMO GONZÁLEZ: Una taxonomía de actividades sostenibles para Europa.
- 2102 FRUCTUOSO BORRALLO, SUSANA PÁRRAGA-RODRÍGUEZ and JAVIER J. PÉREZ: Taxation challenges of population ageing: comparative evidence from the European Union, the United States and Japan. (There is a Spanish version of this edition with the same number).
- 2103 LUIS J. ÁLVAREZ, M.ª DOLORES GADEA and ANA GÓMEZ LOSCOS: Cyclical patterns of the Spanish economy in Europe. (There is a Spanish version of this edition with the same number).
- 2104 PABLO HERNÁNDEZ DE COS: Draft State Budget for 2021. Testimony before the Parliamentary Budget Committee, 4 November 2020. (There is a Spanish version of this edition with the same number).
- 2105 PABLO HERNÁNDEZ DE COS: The independence of economic authorities and supervisors. The case of the Banco de España. Testimony by the Governor of the Banco de España before the Audit Committee on Democratic Quality / Congress of Deputies, 22 December 2020. (There is a Spanish version of this edition with the same number).
- 2106 PABLO HERNÁNDEZ DE COS: The Spanish pension system: an update in the wake of the pandemic. Banco de España contribution to the Committee on the Monitoring and Assessment of the Toledo Pact Agreements.2 September 2020. (There is a Spanish version of this edition with the same number).
- 2107 EDUARDO BANDRÉS, MARÍA-DOLORES GADEA and ANA GÓMEZ-LOSCOS: Dating and synchronisation of regional business cycles in Spain. (There is a Spanish version of this edition with the same number).
- 2108 PABLO BURRIEL, VÍCTOR GONZÁLEZ-DÍEZ, JORGE MARTÍNEZ-PAGÉS and ENRIQUE MORAL-BENITO: Real-time analysis of the revisions to the structural position of public finances.
- 2109 CORINNA GHIRELLI, MARÍA GIL, SAMUEL HURTADO and ALBERTO URTASUN: The relationship between pandemic containment measures, mobility and economic activity. (There is a Spanish version of this edition with the same number).
- 2110 DMITRY KHAMETSHIN: High-yield bond markets during the COVID-19 crisis: the role of monetary policy.
- 2111 IRMA ALONSO and LUIS MOLINA: A GPS navigator to monitor risks in emerging economies: the vulnerability dashboard.
- 2112 JOSÉ MANUEL CARBÓ and ESTHER DIEZ GARCÍA: El interés por la innovación financiera en España. Un análisis con Google Trends.
- 2113 CRISTINA BARCELÓ, MARIO IZQUIERDO, AITOR LACUESTA, SERGIO PUENTE, ANA REGIL and ERNESTO VILLANUEVA: Los efectos del salario mínimo interprofesional en el empleo: nueva evidencia para España.
- 2114 ERIK ANDRES-ESCAYOLA, JUAN CARLOS BERGANZA, RODOLFO CAMPOS and LUIS MOLINA: A BVAR toolkit to assess macrofinancial risks in Brazil and Mexico.
- 2115 ÁNGEL LUIS GÓMEZ and ANA DEL RÍO: The uneven impact of the health crisis on the euro area economies in 2020. (There is a Spanish version of this edition with the same number).
- 2116 FRUCTUOSO BORRALLO EGEA and PEDRO DEL RÍO LÓPEZ: Monetary policy strategy and inflation in Japan. (There is a Spanish version of this edition with the same number).
- 2117 MARÍA J. NIETO and DALVINDER SINGH: Incentive compatible relationship between the ERM II and close cooperation in the Banking Union: the cases of Bulgaria and Croatia.
- 2118 DANIEL ALONSO, ALEJANDRO BUESA, CARLOS MORENO, SUSANA PÁRRAGA and FRANCESCA VIANI: Fiscal policy measures adopted since the second wave of the health crisis: the euro area, the United States and the United Kingdom. (There is a Spanish version of this edition with the same number).
- 2119 ROBERTO BLANCO, SERGIO MAYORDOMO, ÁLVARO MENÉNDEZ and MARISTELA MULINO: Impact of the COVID-19 crisis on Spanish firms' financial vulnerability. (There is a Spanish version of this edition with the same number).
- 2120 MATÍAS PACCE, ISABEL SÁNCHEZ and MARTA SUÁREZ-VARELA: Recent developments in Spanish retail electricity prices: the role played by the cost of CO₂ emission allowances and higher gas prices. (There is a Spanish version of this edition with the same number).

- 2121 MARIO ALLOZA, JAVIER ANDRÉS, PABLO BURRIEL, IVÁN KATARYNIUK, JAVIER J. PÉREZ and JUAN LUIS VEGA: The reform of the European Union's fiscal governance framework in a new macroeconomic environment. (There is a Spanish version of this edition with the same number).
- 2122 MARIO ALLOZA, VÍCTOR GONZÁLEZ-DÍEZ, ENRIQUE MORAL-BENITO and PATROCINIO TELLO-CASAS: Access to services in rural Spain. (There is a Spanish version of this edition with the same number).
- 2123 CARLOS GONZÁLEZ PEDRAZ and ADRIAN VAN RIXTEL: The role of derivatives in market strains during the COVID-19 crisis. (There is a Spanish version of this edition with the same number).
- 2124 IVÁN KATARYNIUK, JAVIER PÉREZ and FRANCESCA VIANI: (De-)Globalisation of trade and regionalisation: a survey of the facts and arguments.
- 2125 BANCO DE ESPAÑA STRATEGIC PLAN 2024: RISK IDENTIFICATION FOR THE FINANCIAL AND MACROECONOMIC STABILITY: How do central banks identify risks? A survey of indicators.
- 2126 CLARA I. GONZÁLEZ and SOLEDAD NÚÑEZ: Markets, financial institutions and central banks in the face of climate change: challenges and opportunities.
- 2127 ISABEL GARRIDO: The International Monetary Fund's view of social equity throughout its 75 years of existence. (There is a Spanish version of this edition with the same number).
- 2128 JORGE ESCOLAR and JOSÉ RAMÓN YRIBARREN: European Central Bank and Banco de España measures against the effects of COVID-19 on the monetary policy collateral framework, and their impact on Spanish counterparties. (There is a Spanish version of this edition with the same number).
- 2129 BRINDUSA ANGHEL, AITOR LACUESTA and FEDERICO TAGLIATI: 2021 Survey of Small Enterprises' Financial Literacy: Main Results. (There is a Spanish version of this edition with the same number).
- 2130 PABLO HERNÁNDEZ DE COS: Testimony before the Congress of Deputies Budget Committee on 25 October 2021 and before the Senate Budget Committee on 30 November 2021 in relation to the Draft State Budget for 2022. (There is a Spanish version of this edition with the same number).
- 2131 LAURA AURIA, MARKUS BINGMER, CARLOS MATEO CAICEDO GRACIANO, CLÉMENCE CHARAVEL, SERGIO GAVILÁ, ALESSANDRA IANNAMORELLI, AVIRAM LEVY, ALFREDO MALDONADO, FLORIAN RESCH, ANNA MARIA ROSSI and STEPHAN SAUER: Overview of central banks' in-house credit assessment systems in the euro area.
- 2132 JORGE E. GALÁN: CREWS: a CAMELS-based early warning system of systemic risk in the banking sector.
- 2133 ALEJANDRO FERNÁNDEZ CEREZO and JOSÉ MANUEL MONTERO: A sectoral analysis of the future challenges facing the spanish economy. (There is a Spanish version of this edition with the same number).
- 2201 MANUEL A. PÉREZ ÁLVAREZ: New allocation of special drawing rights. (There is a Spanish version of this edition with the same number).
- 2202 PILUCA ALVARGONZÁLEZ, MARINA GÓMEZ, CARMEN MARTÍNEZ-CARRASCAL, MYROSLAV PIDKUYKO and ERNESTO VILLANUEVA: Analysis of labor flows and consumption in Spain during COVID-19.
- 2203 MATÍAS LAMAS and SARA ROMANIEGA: Designing a price index for the Spanish commercial real estate market. (There is a Spanish version of this edition with the same number).
- 2204 ÁNGEL IVÁN MORENO BERNAL and TERESA CAMINERO GARCÍA: Analysis of ESG disclosures in Pillar 3 reports.

 A text mining approach.
- 2205 OLYMPIA BOVER, LAURA CRESPO and SANDRA GARCÍA-URIBE: Household indebtedness according to the Spanish survey of household finances and the central credit register: a comparative analysis. (There is a Spanish version of this edition with the same number).