Review of Income and Wealth
Series 56, Number 3, September 2010

# TAX DATA FOR WEALTH CONCENTRATION ANALYSIS: AN APPLICATION TO SPANISH WEALTH TAX

by José Mª Durán-Cabré* and Alejandro Esteller-Moré

*University of Barcelona and Institut d'Economia de Barcelona (IEB)*

As recent studies in different countries show, tax data offer the opportunity to estimate income or wealth shares for the upper income groups. However, several critical points must be considered in order to avoid misleading conclusions: the interpolation technique used, legal amendments, and tax fraud. In this note we take Spanish wealth tax as a case study to assess the importance of these factors, and compare our results with those obtained by Alvaredo and Saez (2009). Although the results of the two analyses are very similar, our approach complements theirs by offering a more precise treatment of the correction of fiscal underassessment and tax fraud in real estate, which is the main asset in Spaniards' portfolios.

## 1. Introduction

Tax data offer an opportunity to estimate income or wealth shares for upper income groups. Based on a seminal study by Kuznets (1953), recent studies on different countries (see, e.g. Atkinson and Piketty, 2007) calculate income concentration using income data from tax returns to compute the income of upper groups, and national accounts data to compute the total income. If a wealth tax is levied, a similar process can be carried out for wealth concentration (e.g. Spant, 1987; Tuomala and Vilmunen, 1988; Ohlsson *et al.*, 2006). In an impressive recent study, Alvaredo and Saez (2009) focus on both wealth and income concentration for Spain.

When only aggregate data are available, it is necessary to interpolate in order to estimate the level of revenue or wealth for a specific percentile (e.g., the top 1 percent). Atkinson (2005) stresses some caveats that must be taken into account when using tax data. These include tax evasion, tax avoidance, legal definitions that are not suitable for distribution studies, and the lack of contextual data. Legal amendments may cause additional problems. Therefore, the use of tax data for distribution analysis is not free of problems.

In this note, we focus on three possible limitations: (i) the reliability of the interpolation method; (ii) the consideration of fiscal changes that may produce discontinuities or level effects in the series; and (iii) the impact of tax fraud. We

*Note*: The authors gratefully acknowledge the comments of X. Oliver, A. Solé, and P. Sorribas, and especially the detailed and useful comments made by the two referees, and by the Editor (Bart van Ark). We also appreciate very much the kind and disinterested feedback received from F. Alvaredo and E. Saez. The usual disclaimer applies. We are grateful for the financial support of the Instituto de Estudios Fiscales, the Spanish Ministry of Science and Technology (ECO2009-12928), and the Government of Catalonia (2009SGR102). A previous longer version was published as WP #460 of *Fundación de las Cajas de Ahorros* (FUNCAS, Madrid).

*Correspondence to: José Mª Durán-Cabré, Department of Public Finance, Faculty of Economics and Business, University of Barcelona, Av. Diagonal, 690, T IV, 2, 08034 Barcelona, Spain (jmduran@ub.edu).

stress the importance of taking these potentially problematic issues into account and propose a methodology, which is applied to the Spanish case for wealth concentration. This is a useful case, since no other statistics on household wealth are available,[1] and the data from the annual wealth tax cover a long period of time (1983–2003). We can also compare the results obtained with our methodology with those presented in Alvaredo and Saez (2009). The considerations in (ii) and (iii) enrich the results obtained by Alvaredo and Saez (2009).[2]

The Spanish wealth tax levies the difference between the value of assets and rights and the value of liabilities and obligations, i.e. the value of an individual's net wealth.[3] This includes, among other things, real estate, goods needed for an individual's business, savings accounts, life insurance, bonds, and closely held and traded shares. However, there are certain exemptions, such as the consolidated rights of members of pension plans; family businesses (individual or incorporated), when certain conditions are met (since 1994); and owner-occupied housing (since 2000 and up to around €150,000). Likewise, the law establishes the valuation rules for taxable assets. This is particularly important in relation to real estate—the main asset in Spaniards' portfolios—whose fiscal value may be very different from the market price.

After addressing the three critical points when using tax data (the interpolation, the underassessment of real estate, housing exemption, the tax unit, and tax fraud), we find that our results are quantitatively not very different from those obtained by Alvaredo and Saez (2009). However, from a qualitative point of view we are able to extract more information from our methodology of analysis, and therefore our results can be seen as complementary to their findings. In particular, our correction of the reported fiscal values of real estate is more precise, and probably most importantly, our methodology permits disentangling this price correction from tax evasion. In general, our methodology can be applied to other countries, where tax data present similar shortcomings to those identified in the Spanish case.

## 2. Interpolation of Top Bracket Wealth Data

Our aim is to estimate the level of wealth of the top 1 percent of the adult population.[4] However, only aggregate data are available from the Spanish wealth

[1]Recently, the Bank of Spain has started to conduct a household wealth survey in order to gain a more accurate idea of the wealth of Spanish households and their financing decisions (Bover, 2004; Bover *et al*., 2005). The aim is to undertake the survey periodically in the future, but data are currently only available for 2002. Therefore, this survey does not overcome the problem of the lack of data for the last few decades.

[2]The calculation of the denominator of the wealth share would be another possible caveat, but in our case we take it for granted from Alvaredo and Saez (2009). The denominator is calculated from several statistical sources, mainly from the Bank of Spain, as the authors explain in appendix C.1 of their paper.

[3]This tax was introduced in 1977 and abolished in 2008. It was payable by residents and non-residents. For residents, the liability was *ad personam*, that is, in respect of all assets and rights. For all others, the liability was *in rem*, in respect of assets in Spanish territory and rights which can be exercised therein.

[4]This accounts for an average of about 280,000 people every year. The threshold for this tax is high, so it only affects the wealthiest individuals. That is why the wealth tax has only been levied on a small percentage of the adult population, between 2 and 4 percent. If we consider the adult population instead of just taxpayers, our target group is more stable and the number of members is less sensitive to legal modifications, for instance, in the threshold or the exemptions.

tax, as taxpayers are gathered into several brackets according to their level of reported wealth, regardless of the number of people in each interval.[5] Thus it is necessary to interpolate. The most common method employed in the literature, and the one used by Alvaredo and Saez (2009), is the Pareto Interpolation (PI).[6] Atkinson (2005) argues that the potential error in this interpolation depends on the width of the ranges. This is an important issue for the Spanish case, as the number and width of intervals in the fiscal data are not the same throughout the study period (1983–2003), and the changes are quite significant (there are 14 brackets from 1988 to 1998, then 10 and later on 20). Moreover, it is possible to obtain more than one value for the interpolated share. Therefore, we calculate lower and upper bounds for wealth shares. The reliability of assuming that wealth distribution at the top is Pareto in form is confirmed when the bounds are tight.

The calculation of gross bounds is explained in Cowell (1995). The gross lower bound implies that within the range of the bracket in which we have to interpolate, the mass of the population is concentrated at the average of the range. Hence, this bound implies maximum equality in the distribution of wealth within that interval. In contrast, the gross upper bound implies maximum inequality within the interval, and is calculated by assuming that within that range a certain percentage of taxpayers are concentrated at the minimum amount of wealth of the interval, while the rest are concentrated at the maximum. However, as suggested and explained in detail by Atkinson (2005), it is still possible to obtain tighter bounds if it is simply assumed that density at the top of the whole distribution of taxpayers is non-increasing (Gastwirth, 1972). In Table 1, we show the results of these tighter bounds, which are known as refined bounds.[7]

The bounds are quite tight, as the average width for the whole period is 1.2 percent, and the maximum value is 3.2 percent (for 1992).[8] As expected, the width is generally greater during 1988–98, when there were only 10 brackets (on average, the width is 1.7 percent for this period). Therefore, we can conclude that the PI very closely approximates the wealth distribution of the top 1 percent, even for the widest intervals.[9] Hence, from now on, we will use the values obtained using the PI.

[5]Fiscal data is based on the WT return statistics published yearly by the Spanish Ministry of Economy and Finance in the *Memoria de la Administración Tributaria* (available at www.aeat.es since 1998). This publication excludes two regions, Navarre and the Basque Country, as they have their own WT, due to a particular financing system. Therefore, our analysis does not include these regions (which in fact account for only 6 percent of the adult Spanish population).

[6]Among many others, see Feenberg and Poterba (1993, 2000) and Piketty (2001, 2003). Atkinson (2005) explains that this follows an honorable tradition, since a Pareto interpolation was used in a 1906 report on Income Tax by the House of Commons Committee.

[7]Although data are also available for 1982, its analysis often shows statistical errors. Therefore, our study period begins in 1983.

[8]In contrast, if we calculate the gross bounds, the average width is 3.6 percent (maximum 10.4 percent in 1992).

[9]In order to calculate the average wealth, assuming a Pareto distribution at the top, we follow the same steps as Alvaredo and Saez (2009) in appendix D.1. In the second step, once the wealth threshold above which we have the top 1 percent has been computed ($W_p$), the wealth between $W_p$ and the extreme value ($t$) of the bracket in which $W_p$ lies is calculated. This wealth is added to the declared wealth above $t$, from which the average wealth of the top 1 percent can be calculated. However, on some occasions, the average does not lie within the refined bounds. In such cases, we calculate the wealth between $W_p$ and the lowest value ($s$) of the bracket in which $W_p$ lies. Then, we deduct that amount from the total wealth of that bracket in order to obtain the total wealth between $t$ and $W_p$. Hence, the calculation of these bounds becomes useful when more than one value might be obtained from PI.

TABLE 1

AVERAGE WEALTH OF THE TOP 1%: PI AND REFINED BOUNDS

| | Refined Lower Bound | Pareto Interpolation | Refined Upper Bound | Width |
|---|---|---|---|---|
| | | 14 brackets | | |
| 1983 | 579,504 | 581,235 | 581,411 | 0.3% |
| 1984 | 554,180 | 556,694 | 558,149 | 0.7% |
| 1985 | 564,827 | 567,095 | 567,877 | 0.5% |
| 1986 | 610,893 | 612,930 | 618,166 | 1.2% |
| 1987 | 649,361 | 651,248 | 651,873 | 0.4% |
| | | 10 brackets | | |
| 1988 | 661,606 | 661,757 | 661,930 | 0.1% |
| 1989 | 691,109 | 693,859 | 696,958 | 0.9% |
| 1990 | 686,920 | 689,838 | 702,852 | 2.3% |
| 1991 | 690,157 | 696,227 | 710,821 | 3.0% |
| 1992 | 662,234 | 675,713 | 683,650 | 3.2% |
| 1993 | 695,671 | 705,229 | 716,994 | 3.1% |
| 1994 | 684,199 | 688,344 | 700,757 | 2.4% |
| 1995 | 692,202 | 695,570 | 705,269 | 1.9% |
| 1996 | 739,219 | 741,803 | 747,673 | 1.1% |
| 1997 | 815,023 | 816,421 | 818,140 | 0.4% |
| 1998 | 889,300 | 889,646 | 889,870 | 0.1% |
| | | 20 brackets | | |
| 1999 | 960,947 | 965,200 | 965,445 | 0.5% |
| 2000 | 949,720 | 952,984 | 952,984 | 0.3% |
| 2001 | 984,105 | 984,105 | 988,024 | 0.4% |
| 2002 | 1,003,172 | 1,006,424 | 1,006,514 | 0.3% |
| 2003 | 1,047,138 | 1,047,374 | 1,058,573 | 1.1% |

*Notes*: Expressed in 2003€. The data are obtained from fiscal information. They include reported net wealth of debts, and exempted business assets (from 1994 onwards) since they must be reported. Wealth is expressed according to fiscal assessment criteria. In order to calculate the average, we divide wealth from tax data among 1% of the adult population.

However, the data shown in Table 1 cannot be used directly to analyze how wealth concentration has evolved, since fiscal issues must first be taken into consideration.

## 3. FISCAL ISSUES

### *Tax Base: Exemptions and Assessment*

The introduction of new exemptions may alter the value of reported wealth, particularly if the exempted assets form a large part of the taxpayer's capital. In Spain, two new exemptions have had an impact: one for business assets (the "family business exemption"), which has existed since 1994, and another for owner-occupied housing, which has been in place since 2000. Fortunately, the value of exempt family business assets must still be reported, and they can therefore be included in the assessment of wealth. However, there is no public data about the distribution of housing exempted values among taxpayers.[10] Consequently, we will have to estimate the value of the exemption. Housing is the main asset of Spanish families (accounting for around 65 percent of their average

[10]In fact, taxpayers must also report the exempted values, but the statistics published for the 2000–03 period do not include them. These data have only been available since 2004.

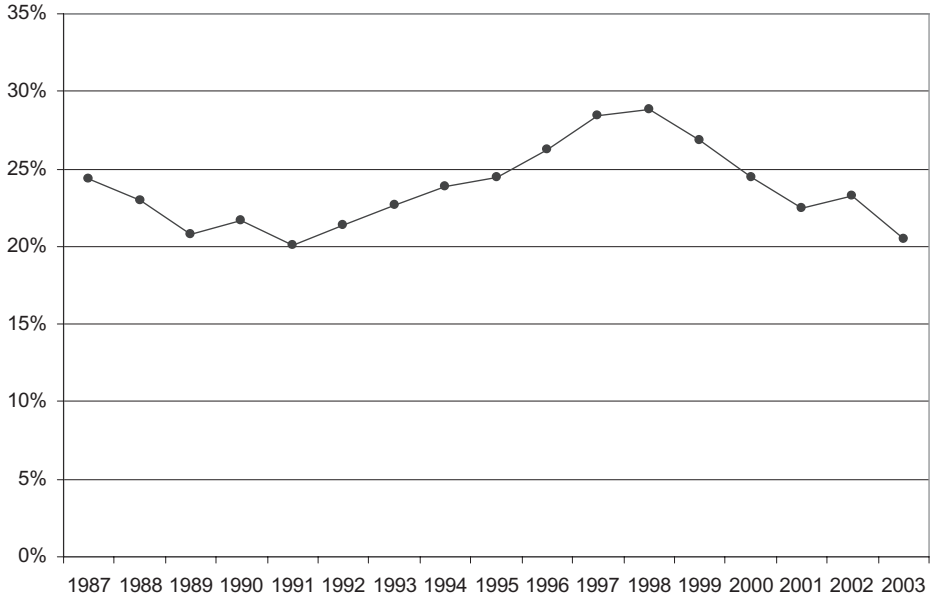


Figure 1. Underassessment of Real Estate Property in Wealth Tax: Ratio of Average Cadastral Value/Average Market Price

*Notes*: In the published statistics, cadastral value is expressed per dwelling. We have assumed that each house has a surface area of 125 $m^2$ in order to compare the two series of prices. In this way, the ratio closely follows the series calculated by García-Vaquero and Martínez (2005).

*Source*: Ministry of Public Works for market price and Cadastral Office for cadastral value, several years.

wealth). We include this second exemption by using the estimated values obtained by Durán and Esteller (2007).

Likewise, the assessment of assets must be homogeneous for the distribution analysis, which means that we can only apply one criterion: the market price. This should not present a problem, as in theory a wealth tax ought to fix the same criterion for all assets. Nonetheless, this is not the case for the Spanish wealth tax, as Spanish law establishes different valuation rules for certain assets, due to administrative constraints. For instance, real estate is usually assessed at administrative values (cadastral value), unless there has been a recent acquisition. Cadastral values are far below the market price throughout the period (Figure 1). Therefore, reported values must be converted into market price.

In order to correct the top 1 percent of wealth in terms of real estate, we take into account underassessment and the housing exemption. In Table 2, we show the corrected values of estates. From the PI, the average percentage of total wealth made up of real estate during 1983–2003 is 28.1 percent, which is a relatively small percentage.[11] However, once we convert the fiscal values into market price and

[11]Reported real estate values are divided into urban and rural values. When transforming estate values into market price, we consider only the former, as before 1992 the latter were assessed using different criteria; they only account for a small proportion of all estates (according to reported values of the top 1 percent, rural properties account for less than 10 percent on average); and, to our knowledge, there are no official statistics regarding their market price. Hence, the market price for real estates refers only to urban properties.

TABLE 2

REAL ESTATES: TRANSFORMATION INTO MARKET VALUE AND ESTIMATED VALUE OF OWNER-OCCUPIED HOUSING EXEMPTION

| Year | % Declared Real Estate Over Total Declared Wealth | Value of Real Estate (from PI) | (Inverse of) Conversion Factor into Market Value (%) | | Exemption Owner-Occupied Housing | | Totally Corrected Value of Real Estate (from PI) | | % Reported Real Estate Over Reported Total Wealth |
|---|---|---|---|---|---|---|---|---|---|
| | | | Market Price Used Very Little | Market Price Used Very Frequently | Market Price Used Very Little | Market Price Used Very Frequently | Market Price Used Very Little | Market Price Used Very Frequently | Market Price Used Very Little |
| 1983 | 29.3% | 177,294 | 27.1% | 49.0% | n.i.f. | n.i.f. | 653,884 | 361,685 | 61.8% |
| 1984 | 30.9% | 179,045 | 27.1% | 49.0% | n.i.f. | n.i.f. | 660,343 | 365,258 | 63.6% |
| 1985 | 31.4% | 185,075 | 27.1% | 49.0% | n.i.f. | n.i.f. | 682,580 | 377,558 | 64.1% |
| 1986 | 30.4% | 195,197 | 27.1% | 49.0% | n.i.f. | n.i.f. | 719,913 | 398,209 | 63.3% |
| 1987 | 29.4% | 200,191 | 27.1% | 49.0% | n.i.f. | n.i.f. | 738,333 | 408,397 | 62.1% |
| 1988 | 27.2% | 188,241 | 25.6% | 47.5% | n.i.f. | n.i.f. | 734,626 | 396,060 | 60.8% |
| 1989 | 26.4% | 187,485 | 23.4% | 45.2% | n.i.f. | n.i.f. | 802,119 | 414,782 | 61.3% |
| 1990 | 27.5% | 195,156 | 24.4% | 46.2% | n.i.f. | n.i.f. | 801,536 | 422,231 | 61.8% |
| 1991 | 27.0% | 194,047 | 22.7% | 44.4% | n.i.f. | n.i.f. | 856,715 | 436,739 | 63.0% |
| 1992 | 27.6% | 192,168 | 24.0% | 45.9% | n.i.f. | n.i.f. | 800,466 | 418,979 | 62.3% |
| 1993 | 26.8% | 195,140 | 25.3% | 47.2% | n.i.f. | n.i.f. | 770,710 | 413,265 | 60.2% |
| 1994 | 28.6% | 196,174 | 26.5% | 48.5% | n.i.f. | n.i.f. | 739,141 | 404,900 | 60.0% |
| 1995 | 28.9% | 197,803 | 27.2% | 49.1% | n.i.f. | n.i.f. | 728,007 | 403,066 | 59.4% |
| 1996 | 29.6% | 210,306 | 29.1% | 50.9% | n.i.f. | n.i.f. | 723,460 | 413,016 | 57.7% |
| 1997 | 29.8% | 226,933 | 31.3% | 53.0% | n.i.f. | n.i.f. | 725,678 | 428,240 | 55.2% |
| 1998 | 29.4% | 235,133 | 31.6% | 53.3% | n.i.f. | n.i.f. | 743,494 | 440,997 | 53.2% |
| 1999 | 27.8% | 235,240 | 29.6% | 51.5% | n.i.f. | n.i.f. | 793,573 | 457,078 | 52.1% |
| 2000 | 22.6% | 181,287 | 27.2% | 49.1% | 205,469 | 131,944 | 871,580 | 501,020 | 53.0% |
| 2001 | 23.6% | 186,876 | 25.1% | 47.0% | 231,568 | 144,682 | 975,677 | 542,211 | 55.1% |
| 2002 | 27.9% | 220,134 | 26.0% | 47.9% | 273,088 | 172,528 | 1,119,817 | 632,037 | 58.8% |
| 2003 | 27.4% | 225,869 | 23.1% | 44.9% | 315,835 | 191,493 | 1,294,134 | 694,563 | 61.2% |

*Notes*: Monetary amounts expressed in €1000 (2003).

n.i.f.: not in force.

In "Market Price Used Very Little," we suppose market prices is used on 1 out of 10 occasions; while in "Market Price Used Very Frequently" we suppose that it is used on 1 out of 2 occasions. "% Reported Real Estate Over Reported Total Wealth" is the ratio between real estate wealth (obtained using a linear interpolation) and total wealth for the top 1% (using PI).

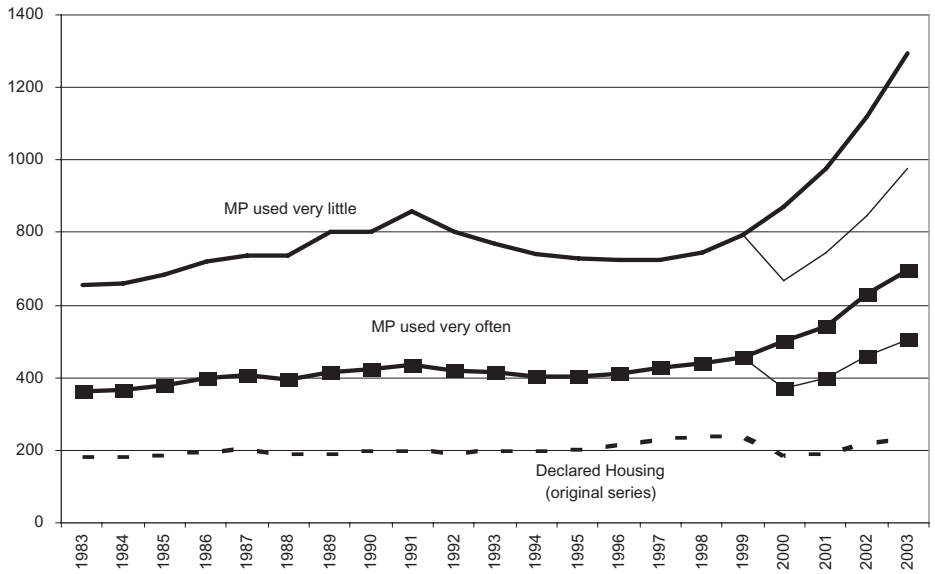


Figure 2. Transformed Value of Real Estate, Taking Into Account Underassessment and the Estimated Value of the Exemption (since 2000)

*Notes*: Monetary amounts expressed in €1000 (2003). See also the Notes for Table 2.
MP = market price.

take into account the exemption (in both cases, assuming market price is used on very few occasions), this average increases by up to 59.5 percent, i.e. more than twofold. Therefore, in order to obtain sensible conclusions about wealth shares, it is important to estimate the corresponding corrections of reported real estate values as accurately as possible. In Figure 2, we show the evolution of the reported real estate values, depending on the assumption of how to correct them for market value and for the exemption. The bottom line shows the evolution of declared values, that is, those according to fiscal criteria and when the exemption is not considered. The other two series correct for both issues. Obviously, as long as we suppose that the market price is used very frequently, the correction is not as severe as when we assume that the market price is rarely used when taxpayers fill in their tax returns.

Certainly, it is impossible to ascertain which assumption is the right one. In fact, there are many other choices, as the labels "little use" or "much use" are arbitrary and could have been parameterized in other ways. For instance, an extreme case would be to suppose that the market price is not used at all (either because it does not legally apply or because taxpayers simply record cadastral value regardless of tax rules), and the (inverse of the) conversion factor would have simply been the share between cadastral value and market price (see Figure 1). In any case, the advantage of our correction is that it can be modulated according to the fiscal use of market price, to obtain a range of transformed values of declared

real estate.[12] In contrast, as explained in Alvaredo and Saez (2009, appendix D.2, section "Top Wealth Shares Estimation"), the conversion factor is the ratio between the "total real estate from the Flow of Funds accounts divided by total cadastral value reported in aggregate real estate statistics." Undoubtedly, this is also a sensible procedure, but it implicitly assumes that all real estate is reported according to cadastral value, which is only close to the reality. Moreover, it mixes underassessment and tax fraud.

A similar reasoning applies to the calculation of the housing exemption, whose value, according to our methodology, again depends on the hypothesized use of the market price in the wealth tax. Basically, we have estimated this value by regressing the reported value with respect to cadastral value and other controls (see table 8 and equation (10) in Durán and Esteller, 2007). The value of the exemption implies a lower estimate of the impact of cadastral value on reported values. Again, this approach enables us to obtain a declared monetary value of the exemption that is greater (lower) if the market price is used very frequently (little), which we then transform using the corresponding conversion factor. Nevertheless, given an estimated reduction in the impact of cadastral value on declared values, the real monetary amount of the exemption becomes greater in terms of market price the lower the use of market price. In Table 2, we show two sets of reported values of the exemption, which again depend on the fiscal use of market price. As the published statistics for 2004 give the values of the exempted housing among taxpayers (see footnote 10), we have estimated the value of the exemption for the top 1 percent using a linear interpolation, and obtained a value of €59,158 (2003€), which should still be transformed into market price. Therefore, if we suppose that market price is used very little and we use the conversion factor for 2003, that amount would be €256,206, which is relatively close to €315,835, the value we estimated for 2003.

In general, once we have taken into account the fiscal issues dealt with in this section, we achieve a rectification of the original series of real estate wealth declared by the top 1 percent, and we can use this figure to recalculate total wealth from tax data. However, this correction brings the series within a range of values that depend on the fiscal use of market price, which are probably closer to "little use" than to "very frequent use."

*The Tax Unit*

The tax unit in Spain changed in 1988, with a shift to individual taxation. Before this time, joint taxation was compulsory. Therefore, the reported wealth in a tax return dated pre-1988 may belong to more than one person. As we only have

[12]Analytically, the transformation works as follows. Taking "MP" as market price, and "CV" as average cadastral value, we suppose $CV = MP^{\delta}$, where $\delta$ identifies the underassessment (i.e., if $\delta = 1$, the underassessment is null), and it is calculated to replicate real data (see Figure 1). The declared price per unit of housing is $CV^{1-\alpha}MP^{\alpha}$, where $\alpha$ identifies the frequency of use of MP (i.e., if $\alpha = 1$, MP is used on all occasions). From both equations, we obtain the conversion factor—that is, the ratio between MP and declared price—as $MP^{(1-\alpha)(1-\delta)}$. Then, the factor ($\geq 1$) depends both on the frequency of the use of MP (necessarily hypothesized) and on the degree of underassessment (obtained from real data). This is explained in more detail in Durán and Esteller (2007). Finally, statistics on market price are only available from 1987 onwards. Thus, for the period previous to 1987, like Alvaredo and Saez (2009), we assume that the ratio CV/MP is the same as in 1987.

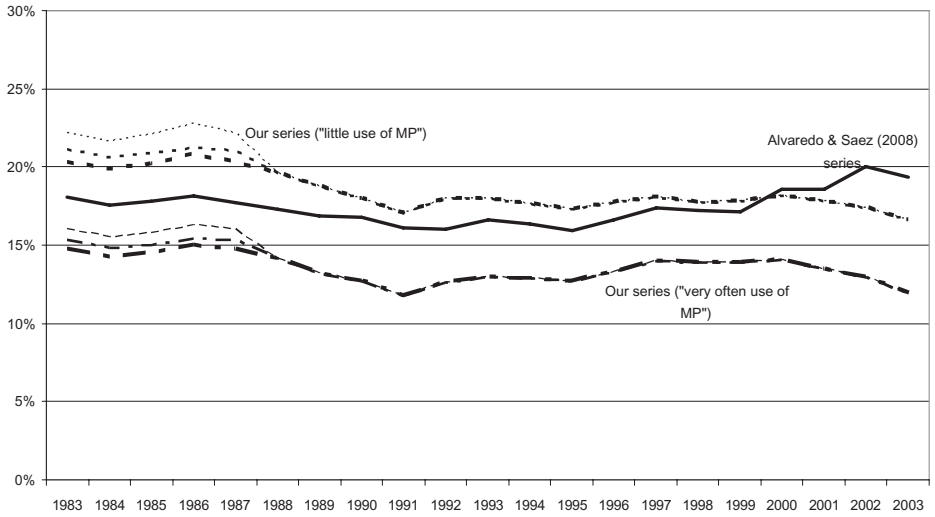


Figure 3. Wealth Share for the Top 1%: An Initial Comparison with Alvaredo and Saez (2009)
*Note*: MP = market price.

aggregate data, it is not possible to identify when there is more than one person in a tax return or how wealth is split. Therefore there is discontinuity in 1988. If this legal amendment were not considered in the distribution study, the results would show a significant decrease in wealth concentration in Spain in 1988, which would lead to confusing conclusions.

Alvaredo and Saez (2009) also take this fiscal issue into account, which is explained in their appendix D.2.[13] Basically, they correct the 1987 wealth share and assume that growth between 1987 and 1988 is the average growth between 1988–89 and 1986–87. Once they obtain the corrected value for 1987, each share between 1983 and 1986 is obtained using 1987 as the base year and employing the original growth rates between years. We also use this methodology, but apply it both to wealth shares and to wealth amounts for the top 1 percent (before correcting by market price and by owner-occupied housing; afterwards we make these corrections). Although both ways of approaching the problem of the tax unit imply a logical reduction in the top 1 percent of wealth share, the value of the reduction differs in the two cases. This is shown in Figure 3.

Our series (dotted lines) are corrected according to the method outline above, depending on the frequency of use of the market price. The top line for each one of our series does not correct for the problem of the tax unit between 1983 and 1986, while each of our thickest dotted series corrects for this problem in the same way as Alvaredo and Saez (2009). Between both lines, we construct a new series that employs the same correction for the tax unit, calculated on the amount of wealth instead of on wealth shares. The differences are insignificant.

[13]In previous versions of their paper, they did not take into account this change, which demonstrates the importance of considering tax regulations when working with tax data.

On the whole, once we correct for the fiscal issues regarding the tax base and the tax unit and compare our series with Alvaredo and Saez's (2009) series, we can conclude that the two series follow a similar pattern until 1999. The variations in the levels might be explained by the different way of correcting real estate, and even the calculation of the percentage of housing wealth (see footnote 11), to transform it into market price and/or the ways of estimating the value of the owner-occupied exemption. Hence according to our series, it seems that the boom in housing prices since the end of the 1990s has been progressive (i.e., wealth concentration has decreased), while according to Alvaredo and Saez (2009), it has been regressive. In the next section, we reconcile these contradictory results.

## 4. Tax Fraud

Tax data only provides comparable information over time if the pattern of tax evasion in the richest groups of taxpayers remains equal over time. Likewise, the share of wealth for a given year would only be reliable if tax fraud were evenly distributed among taxpayers regardless of their level of wealth, which seems unlikely. Therefore, it is important to consider the possible impact of tax fraud, particularly in the Spanish wealth tax. Indeed a common belief among tax specialists is that the level of fraud in wealth tax has been high, and has also increased over time.

Since the mid-1980s, the regional governments have obtained all the revenue raised by wealth tax and are also responsible for its administration. However, the number of tax audits is extremely low (on average, the yearly number of audits between 1986 and 2003 is about 900, accounting for only 0.1 percent of all tax returns).[14] Furthermore, the average revenues raised by wealth tax audits are €1462 (expressed in 2003 terms), which are much lower than the revenues collected from auditing other taxes. Therefore, the common belief regarding the high level of fraud seems to be correct.

Consequently, we try to estimate tax fraud and recalculate wealth shares for the top 1 percent, which are shown in Figure 4 (data are only available from 1987, as we need to use market price, which has only been available since then; in fact, we have eliminated the year 1987 in order to remove the correction of the tax unit).[15] Once we take tax fraud into consideration, our series and that of Alvaredo and Saez (2009) show a very similar pattern (the coefficient of correlation between the two series is 91 percent). This is consistent with the way they correct the declared values of real estate, as the ratio between total real estate from the Flow of Funds accounts and total cadastral value simultaneously takes into account tax

[14]This result is calculated using data from an annual report on the cession of taxes to the ACs, Spanish National Budget, Ministry of Economy and Finance.

[15]According to footnote (12), the declared market price (MP) is $MP^{\alpha+\delta(1-\alpha)}$, but in the presence of tax fraud it transforms into $MP^{\theta[\alpha+\delta(1-\alpha)]}$, where $\theta$ is the degree of tax compliance, $-\infty \leq \theta \leq 1$. Therefore, the conversion factor due to tax fraud is $MP^{(\theta-1)[\alpha+\delta(1-\alpha)]}$, that is, the ratio between the declared price and the price in absence of tax fraud but assessed according to fiscal criteria. As we do not know $\theta$, we have estimated the elasticity of declared values with respect to MP, $\hat{\beta}$, that is, $\hat{\beta} = \theta[\alpha + \delta(1-\alpha)]$. Given that from footnote (12), we know the value of $\delta$, we immediately see that the conversion factor due to tax fraud is equal to $MP^{\hat{\beta}-\alpha-\hat{\delta}(1-\alpha)}$. Again, as in footnote (12), where we calculated the conversion factor due to underassessment, this factor depends on the frequency of use of market price, $\alpha$. See Durán and Esteller (2007, section 5.4.1) for more details.

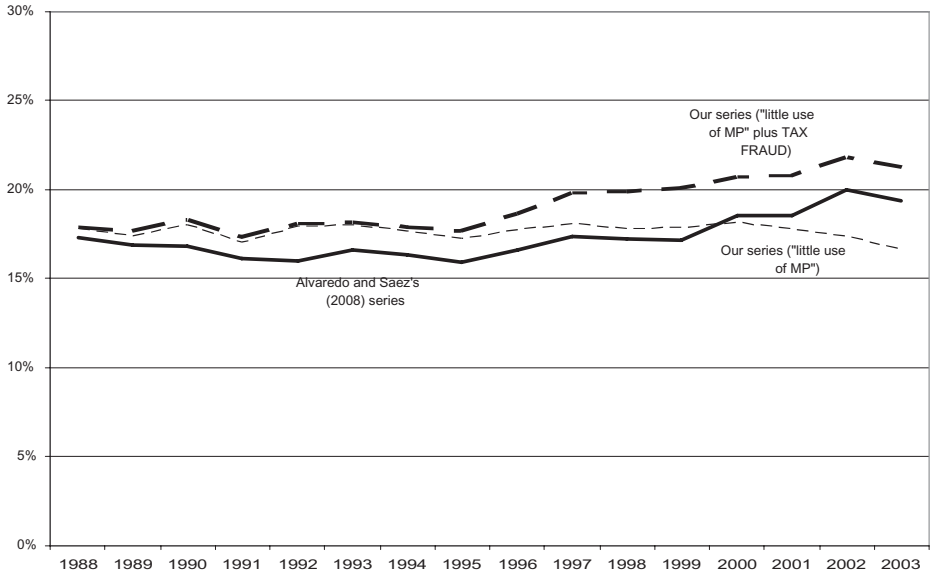


Figure 4. Wealth Share for the Top 1%: A Last Comparison (and Reconciliation) with Alvaredo and Saez (2009)

*Note*: MP = market price.

fraud and underassessment. The advantage of our procedure is that we can disentangle the impact of the two factors, although the final result is very similar.

Therefore, we can conclude: (i) the boom in housing prices since the late 1990s has been regressive, and has contributed to an increase in wealth concentration at the top of the distribution, even though the maximum concentration seems to have been attained in 2002; and (ii) this effect has not been reflected in reported fiscal values; in other words, there seems to be growing fraud in the wealth tax when the housing market is booming. Interestingly, our methodology, which is more parsimonious than the one of Alvaredo and Saez (2009), enables us to infer a growing level of tax fraud during the housing market boom.

The difference in levels between our series and that of Alvaredo and Saez (2009) is not very important. According to our methodology (which, among other factors, implies "little use of market price," as we think this is the most realistic), during 1988–2003 the average wealth share for the top 1 percent was 19.1 percent (maximum 21.8 percent in 2002; and minimum 17.3 percent in 1991), while according to Alvaredo and Saez (2009), the average was 17.3 percent (maximum 20.0 percent in 2002; and minimum 15.9 percent in 1995).

## 5. Conclusion

Tax data can be useful when carrying out a concentration analysis. However, it is important to bear in mind the problems that may be involved. In this note, we observe that the assessment of assets (real estate), the introduction of exemptions (for owner-occupied housing), the tax unit (individual vs. joint taxation), and the level of tax fraud have a significant impact on our case study, the application to

Spanish wealth tax. This methodology may be useful for studies in other countries in which the use of tax data faces similar difficulties.

## References


Alvaredo, F. and E. Saez, "Income and Wealth Concentration in Spain in a Historical and Fiscal Perspective," *Journal of the European Economic Association*, 7, 1140–67, 2009.

Atkinson, A. B., "Measuring Top Incomes: Methodological Issues," Mimeo, Nuffield College, University of Oxford, 2005.

Atkinson, A. B. and T. Piketty, *Top Incomes over the 20th Century. A Contrast Between Continental European and English-Speaking Countries*, Oxford University Press, Oxford, 2007.

Bover, O., "The Spanish Survey of Household Finances (EFF): Description and Methods of the 2002 Wave," Documentos Ocasionales, No. 0409, Bank of Spain, 2004.

Bover, O., C. Martínez-Carrascal, and P. Velilla, "The Wealth of Spanish Households: A Microeconomic Comparison with the United States, Italy and the United Kingdom," *Economic Bulletin*, Banco de España, 2005.

Cowell, F., *Measuring Inequality*, 2nd edition, Prentice Hall, London, 1995.

Durán, J. M. and A. Esteller, "An Empirical Analysis of Wealth Taxation: Equity vs. Tax Compliance," Papeles de Trabajo, 4/07, Instituto de Estudios Fiscales, Madrid, 2007.

Feenberg, D. R. and J. M. Poterba, "Income Inequality and the Incomes of Very High-Income Taxpayers: Evidence from Tax Returns," *Tax Policy and the Economy*, 7, 145–77, 1993.

———, "The Income and Tax Share of Very High-Income Households, 1960–1995," *American Economic Review*, 90, 264–70, 2000.

García-Vaquero, V. and J. Martínez, "Fiscalidad de la vivienda en España," Documento ocasional no. 0506, Banco de España, 2005.

Gastwirth, J. L., "The Estimation of the Lorenz Curve and Gini Index," *Review of Economics and Statistics*, 54, 306–16, 1972.

Kuznets, S., *Shares of Upper Income Groups in Income and Savings*, National Bureau of Economic Research, Cambridge, MA, 1953.

Ohlsson, H., J. Roine, and D. Waldenström, "Long-Run Changes in the Concentration of Wealth," WIDER Research Paper 103, UNU-WIDER, Helsinki, 2006.

Piketty, T., *Les hauts revenus en France au 20eme siècle—Inegalities et redistribution, 1901–1998*, Editions Grasset, Paris, 2001.

———, "Income Inequality in France, 1901–1998," *Journal of Political Economy*, 111, 1004–42, 2003.

Spant, R., "Wealth Distribution in Sweden: 1920–1983," in E. Wolff (ed.), *International Comparison of the Distribution of Household Wealth*, Oxford University Press, Oxford, 1987.

Tuomala, M. and J. Vilmunen, "On the Trends Over Time in the Degree of Concentration of Wealth in Finland," *Finnish Economic Papers*, 1, 184–90, 1988.