



**WORKING PAPERS**

Col·lecció d'Economia E14/319

---

# Sticky income inequality in the Spanish transition (1973-1990)

Sara Torregrosa Hetland



**B** Universitat de Barcelona

## Sticky income inequality in the Spanish transition (1973-1990)

**Abstract:** This paper investigates the evolution of income inequality in Spain during its transition to democracy, suggesting a method for the correction of under-reporting of earnings and profits in the Household Budget Surveys' data. The contribution is twofold: the methodological proposal, based on income-expenditure discrepancy and scaling-up to National Accounts, improves on previous work, and can be convenient for similar historical sources in other countries. Secondly, its application results in an alternative history of the distribution of income in this case, changing the levels and also the observed trend. Previous literature asserted a substantial equalization, related to the democratization process, while after the adjustment inequality in disposable income is shown to have been quite persistent.

JEL Codes: D31, N34.

Keywords: income inequality, income under-reporting, household surveys, democratization.

Sara Torregrosa Hetland  
Universitat de Barcelona

**Acknowledgements:** This paper is part of the PhD dissertation project "Tax system and redistribution: the Spanish fiscal transition (1960-1990)", under the supervision of Alfonso Herranz Loncán and Alejandro Esteller Moré, to whom I am most grateful for their help. I acknowledge financial support from the Spanish Ministry of Education's scholarship program Formación del Profesorado Universitario and the Research Project ECO2012-39169-C03-03. I also thank the team from Carlos III University who worked on making Household Surveys data more easily available for researchers, and the comments of seminar participants in Santiago de Chile and Bogotá. Emmanuel Saez hosted me at UC Berkeley while culminating this research; I am also grateful for his insights. All remaining errors are my own.

# 1 Introduction

Income inequality is at the center of many debates. Political power, economic development or taxation are all related to the distribution of resources in any given country –or the world. This study takes a dynamic national perspective, and investigates how inequality changed during a period of transition from dictatorship to democracy.

The contribution of the paper is twofold. On the one hand, it is inserted on the debate about the distributional consequences of political transitions, providing an example where income inequality did not substantially decrease after democratization. Secondly, it does so by applying a correction methodology to the main historical source, namely the Household Budget Surveys, which leads to results challenging the prior consensus.

The literature on income distribution has undertaken many changes in the last decades. After the popularization of Kuznets (1955)'s theory about structural change and the decrease in inequality in advanced industrial countries, recent work has pointed at a new upsurge. Among its causes, globalization and skill-biased technological change hold pre-eminent places (Krugman, 2000; Atkinson, 2000; Easterly, 2004). The slowdown of economic growth after the oil crises and the rise of unemployment could also have played a role in certain contexts.

This phenomenon, however, cannot be analysed as a purely economic issue. On the contrary, it is connected to political developments, such as the present rise of neo-liberalism and the deep crisis in social democracy in post-industrial societies. Levy and Temin (2007) argue that the widening of income inequality in the US since 1980 is largely related to the institutional context, which is shaped politically. Labour market regulation, the education system and fiscal redistribution all have strong distributive effects, as has also been underlined by Piketty (2003) for the latter.

In this context, transitions from dictatorship to democracy are expected to bring about a decrease in income inequality, as a result of the higher influence of the distributive goals of lower classes (Meltzer and Richard, 1981 and related literature). But, as Acemoglu et al. (2013) note, the issue of transition might be complex and nuanced: the new regime can be "captured" by the elites and not result in fully democratic policies, and it can also lead to economic liberalization and increased market inequalities.

The Spanish transition (1976-82) is an interesting example for this discussion. Democratization came when the oil crises hit the country, and the early period of the new regime was marked by industrial restructuring and international integration, as well as by an unprecedented and dramatic increase in unemployment. The intensification of structural transformation and the development of welfare-state functions brought about by the ascent of social democracy to power could have pushed the income distribution in different directions. So which force prevailed? Was democratization a strong enough driver for equality?

Generally, studies on Spanish income inequality for the period 1970-90 have found that differences between the poor and the rich shrank substantially (e.g. Alcaide, 2000; Ayala et al., 2006). This result is consistent with a positive impact of the political transition and the subsequent development of the Welfare State in the country. This work, however, reaches different conclusions.

The main data source for income distribution in this period are the Household Budget Surveys. Nevertheless, they suffer from a widely known problem of under-reporting of earnings, particularly those coming from self-employment and capital, which can potentially bring about a misrepresentation of the real levels of inequality.<sup>1</sup> I address the issue with an upwards correction of household incomes by revenue sources, using both internal and external information, and ultimately adjusting the flows to the National Accounts. This is common practice in Latin American studies (ECLAC, 1991), and has recently been applied by an extensive literature focusing on inequality measurement issues in several countries (Accardo et al., 2009; Neri and Zizza, 2010; McColl et al., 2010; Fixler and Johnson, 2012).<sup>2</sup>

The majority of these works are very recent, and focus on the latest years available. This paper makes a step forward by adopting a historical perspective, and measuring the inequality trend over several decades.

After scaling up the income data, I find inequality to have been higher than usually assumed, and to have experienced only a very slight decrease in the transition from dictatorship to democracy. This suggests that, in Spain, the democratic transition was not sufficiently strong to impact positively on distributional dynamics. It also implies that under-reporting has to be taken into account in the study of income distribution and its changes over time. Differential rates of income concealing by source will not only mean higher inequality than that directly observed, but may also affect its trend, fundamentally in the presence of significant changes in the factorial distribution.

The rest of the paper is organized as follows. Section 2 discusses the procedures and conclusions of previous literature on the topic, while also presenting the main data source used, the Household Budget Surveys. The methodology and process of correction of the data are exposed in section 3, and the results and some of their implications are reported in section 4. Section 5 concludes.

## 2 The story of personal income equalization

Literature has shown a widespread consensus on the fact that inequality decreased very substantially in Spain between the 1970s and the 1990s. This has been related to structural economic change and to an increase in the redistributive role played by the State in the second half of the period, due to democratization.

Those studies are generally derived from the Household Budget Surveys (henceforth HBSs). These are consumption and income investigations conducted by the National Statistical Insti-

---

<sup>1</sup>This problem has been signalled by the literature as a reason to use tax data for the top incomes, which would allow to perform an upwards correction of the inequality indices (Atkinson, 2007; Alvaredo, 2011).

<sup>2</sup>This orientation can be traced back to the NBER Conference on Research in Income and Wealth of 1975 (Budd and Radner, 1975). At present, there is related work in progress about the US by E. Saez, Th. Piketty and G. Zucman.

tute (INE, from now on) more or less on a ten-year basis since 1964.<sup>3</sup> They provide information on socio-economic classes, total household disposable income and expenditure in different categories of goods and services. The detail and quantity of information have improved over time. Estimations of self-supply and imputed income from owner-occupied housing are also provided (thus indicating if the family rents or owns their house), as are the households' size and some information on their age composition. The income data always refers to disposable income, so each component is net of direct taxes: this is also the definition used all along this paper.

There are significant differences in the results obtained from this source. Some studies use the original income data provided by the surveys, while other rely on different correction procedures, since some troubling problems are widely known to be present in the HBSs. I will first review the results based on the original data, and then proceed to discuss the quality issues in the surveys. Finally, I will show the corrections proposed by previous literature.

## 2.1 Working with the raw HBS data

The studies which use the raw HBS data are surveyed in table 1. They observe a significant reduction of inequality along these decades, attaining by 1990 levels comparable to those of other developed European countries. Many of these studies acknowledge the problems in the data, such as under-reporting, and therefore call for caution or test for possible impacts with techniques such as trimming (for instance, Cowell et al., 1999) or a comparison with National Accounts (Oliver et al., 2001).

The values of the inequality indices vary depending on each author's methodological choices, such as the income definition, the equivalence scale applied, or the weighting unit. All these are important conceptual decisions to be made by the researcher. I deem preferable an income definition as wide as possible (TDI in the table, which includes in kind elements such as imputed owner-occupier income –but, recall, excludes direct taxation), and individual weighting. This last aspect may not have a big quantitative impact on the indices, but implies giving the same value in our calculations to all individuals (weighting by households effectively means attaching less importance to those living in big families). In any case, these choices do not change the qualitative result here: a decrease in inequality along with the political transition.

Alternatively, many authors are interested in working with inequality of consumption instead of income, or along with it (e.g. Del Río and Ruiz-Castillo, 1996; Martín-Guzmán et al., 1996; Goerlich and Mas, 2001; Gradín, 2002; Gradín et al., 2008). They generally also find a decrease in inequality during the decades of 1970-90.<sup>4</sup> The rationale for this approach is that, in the context

---

<sup>3</sup>It is not possible to rely on Personal Income Tax statistics to study income distribution in the general population, given that until 1979 they covered only a very small part of it. As a depiction of top incomes, they have been used by Alvaredo and Saez (2009). Other work has relied on macroeconomic indicators (Prados de la Escosura, 2008).

<sup>4</sup>Martín-Guzmán et al. (1996), for example, give a Gini of per capita expenditure of 35.7, 35.0 and 34.2 respectively for the three HBSs (this includes non-monetary items and is weighted by household). Goerlich and Mas (2001) find 32.7, 31.4 and 30.3 with total equivalized expenditure. Gradín (2002)'s calculations result in 35.9, 33.2 and 31.7 (OECD's equivalence

Table 1: Income inequality in studies using the raw HBSs

| Study                       | Income definition | Equivalence scale | Weights | Income Gini |        |       |
|-----------------------------|-------------------|-------------------|---------|-------------|--------|-------|
|                             |                   |                   |         | 1973        | 1980   | 1990  |
| Ayala et al. (1996)         | MDOI              | OECD              | H       | -           | 0.337  | 0.312 |
|                             | MDOI              | OECD              | I       | -           | 0.330  | 0.305 |
| Martín-Guzmán et al. (1996) | TDI               | per capita        | H       | 0.358       | 0.352  | 0.330 |
| Cowell et al. (1999)        | TDI               | SR                | I       | -           | 0.313  | 0.300 |
| Goerlich and Mas (2001)     | TDI               | B(0.5)            | I       | 0.321       | 0.310  | 0.293 |
|                             | TDI               | no                | H       | 0.362       | 0.342  | 0.330 |
|                             | TDI               | per capita        | I       | 0.342       | 0.337  | 0.316 |
| Oliver et al. (2001)        | TDI               | OECD              | H       | -           | 0.331* | 0.293 |
| Ayala et al. (2006)         | MDI               | B(0.5)            | H       | 0.353       | 0.333  | 0.317 |
|                             | MDI               | OECD              | H       | 0.354       | 0.339  | 0.317 |
|                             | MDI               | OECD mod.         | H       | 0.350       | 0.334  | 0.314 |

Income definition: TDI is Total Disposable Income, MDI is Monetary Disposable Income, MDOI includes only ordinary revenues. Equivalence scales: SR means square root of household size, B(0.5) Buhmann et al. (1988)'s scale with elasticity of 0.5, 'no' means total household income is used with no adjustment, 'per capita' involves dividing it by real household size. Weighting: I stands for individual, H for household.

(\*) Oliver et al. (2001)'s source is not the HBS, but a different survey (*Encuesta Continua de Presupuestos Familiares*) which started to be conducted in 1985. The value for that year is introduced in the 1980 column in the table.

of the life-cycle and permanent income theory, consumption is a better indicator of welfare. An excellent survey of the debate is given by Gradín et al. (2008), who compare the results of using income or consumption. Morelli et al. (2014) argue that income is conceptually a better indicator, since it measures potential consumption and therefore does not lead to confuse need with chosen frugality (following Sen, 1992), and because current consumption may not mirror permanent income in presence of obstacles to lifetime smoothing (especially borrowing constraints).<sup>5</sup>

The use of consumption can also arise from the acknowledgement that income is under-assessed, and therefore reported consumption would actually be closer to real income than the stated revenue amounts. Expenditure data is not free of measurement issues, such as the difficulty to correctly capture durable goods consumption. But income is truly known to be under-

---

scale), which is slightly over the values he obtains for income (35.2, 32.7 and 30.6). Some works on expenditure inequality have also been done with the data provided by another survey, the *Encuesta Continua de Presupuestos Familiares*, generally obtaining lower levels; see e.g. Gradín et al. (2008) or Pijoan-Mas and Sanchez-Marcos (2010).

<sup>5</sup>Attanasio (1999), with cohort data from the US and the UK, shows how the variability of disposable income over the life cycle is mirrored by that of consumption, although in a less pronounced way in the case of equivalent non-durables. Borrowing constraints have in fact been found significant for low income households in several studies, e.g. Cutanda (2003).

estimated in many surveys, and remarkably in the case of Spain during the second half of the 20th century. We turn to this now.

## 2.2 Biases in the sources

The quality of the HBSs data is highly uneven. The original micro-data of 1964-65 are lost, so it is only possible to work with aggregate results published by INE. In the other cases (1973-74, 1980-81 and 1990-91),<sup>6</sup> micro-data are available on-line. In this work, I am using the files provided by a team from Carlos III University, which undertook a project to facilitate their usage.

Several issues on the reliability of HBSs (as that of their counterparts in other countries) have been put forward by the literature, starting with the publications of INE itself. As may be seen in table 2, household surveys underestimate inequality for a number of reasons. Some of them seem more worrying than others: the exclusion of the homeless might be quite insurmountable, but its quantitative impact is limited. Undistributed profits can be considered as part of the economic capacity of the individuals they accrue to, but may be left aside from an annual income analysis (as is indeed most common in the literature).<sup>7</sup>

Table 2: Biases in the Household Budget Surveys

|                    | Problem                        | Effect on inequality |
|--------------------|--------------------------------|----------------------|
| Universe           | Excludes the homeless          | -                    |
| Income definition  | Excludes undistributed profits | -                    |
| Sampling procedure | No oversampling                | -                    |
| Non-response       | Bigger in dynamic urban areas  | -                    |
| Under-reporting    | Bigger in non-salary income    | -                    |

Source: Author's compilation, based on Garde et al. (1996).

The remaining issues appear more troubling. Oversampling the higher-income strata (urban areas with wealthy inhabitants) would have helped to provide better estimates of income for rich families, since the variability among them is usually higher (this is a common method in modern statistics). On the other hand, non-response and under-reporting entail a likely under-representation of the rich both in quantity and income levels. Not correcting for these effects implies a potentially important bias. The problem is relatively common in this kind of surveys, stemming from lack of accounting control in the families, hiding of income from informal activities, fear of tax inspection, and so on.

Trying to confront the issue, some statistical work was already undertaken during the 1970s.

<sup>6</sup>The surveys always covered a 12-month period, but it did not coincide with the calendar year.

<sup>7</sup>The bias associated to this exclusion would grow, however, given that National Accounts depict an increase in corporations' share of capital income with respect to households'.

At least part of the unit non-response bias is corrected by INE with the scaling-up factors provided with the results, which give higher population weights to observations in strata where unit non-response was more acute.<sup>8</sup> However, under-reporting clearly remains an issue.<sup>9</sup> A simple comparison of the data on total income and total expenditure (plus net savings) tells us that something is wrong: only 30-40% of the households spend less than their yearly income, while around 10% would consume more than twice its level (table 3).<sup>10</sup>

Table 3: Households and budget constraint:  
distribution according to the ratio (Expenditure + Net Savings) / Income

|         | 0 - 1 | 1 - 2 | 2 - 4 | >4   | Total |
|---------|-------|-------|-------|------|-------|
| 1973-74 | 30.8% | 57.4% | 10.7% | 1.1% | 100%  |
| 1980-81 | 37.2% | 53.5% | 8.4%  | 0.8% | 100%  |
| 1990-91 | 41.9% | 48.5% | 8.8%  | 0.8% | 100%  |

Source: author's calculations based on Household Budget Surveys.

Note: households in the first column spend within their budget constraint.

A ratio of 2, for example, means that the family reported to spend twice as much as her yearly income.

Certainly, not all families in a given year spend less than they earn, but the high ratios in the table seem implausible, specially given that total net household savings in those years were positive, according to the Spanish National Accounts. In fact, another possible evidence of under-reporting is a comparison with National Accounts data, which normally are taken as a more reliable source for the aggregates. Disposable income totals are contrasted in table 4 (a disaggregated comparison can be found in tables A.1, A.2 and A.3 in the Appendix).

Incomes in the surveys are only around 70% of those estimated in national accounting, which reinforces our suspicion that in the HBSs they are under-assessed to a considerable extent. The fact that this problem affects richer areas and non-salary income to a greater degree (as stated

<sup>8</sup>Unit non-response is total lack of answer from one selected household, due to refusal or inability to contact it; it is different from item non-response, which arises when one household participates in the survey but fails to provide answers to selected questions. The re-weighting procedure does not eliminate the whole problem, as it can be argued for example that non-response correction should take into account also the income level of households' strata, which affects the probability of response, as suggested by Mistiaen and Ravallion (2003). Pérez-Duarte et al. (2010), however, show that for the Finnish wealth survey non-response bias was not substantially reduced after applying more refined re-weightings and calibrations using further variables.

<sup>9</sup>It has to be kept in mind that these are not tax data, so the term under-reporting does not equal tax base fraud: it could be related to evasion, but also to lack of adequate accountancy, mistakes and forgetfulness, or errors in calculating yearly totals from the questions in the surveys.

<sup>10</sup>Only the 1990 HBS has data on reported savings. For the 1973 and 1980 surveys, therefore, calculations are directly expenditure/income. Negative savings would reduce the apparent excess spending shown in the table for some cases, but are not reasonably the whole issue.



by e.g. Alcaide and Alcaide, 1974 and Sanz, 1995) should warn us against the use of these data without enough criticism. The under-estimation of incomes in the surveys seems more acute in the seventies than in 1990, which could indicate an improvement in the accuracy of the source and therefore a non-homogeneous bias over time – thus affecting inter-temporal comparisons.<sup>11</sup>

Table 4: Disposable Income in HBSs and National Accounts

|      | HBSs (1)   | Nat. Acc. (2) | (1)/(2) |
|------|------------|---------------|---------|
| 1973 | 2,209,839  | 3,099,302     | 65.5%   |
| 1980 | 7,703,772  | 11,049,326    | 69.7%   |
| 1990 | 25,079,849 | 33,387,093    | 75.1%   |

Millions of nominal pesetas.

Source: author's calculations and Pena and Callealta (1996).

### 2.3 Literature with correction of HBSs

The problems surveyed in the previous subsection were known by both INE and the research community at the time. As a result, some corrections were attempted in the data. Table 5 displays the original distribution from the HBSs, together with the main adjusted estimates available.

The original distributions show a constant increase in the shares of the bottom five deciles, together with a decrease in income accruing to the top (deciles 9-10). The Gini index corresponds to that given by Goerlich and Mas (2001) in their second row in table 1. According to these data, the period of the democratic transition was very positive for the Spanish poor and middle classes.

The other columns in the table present distributions corrected for under-reporting with different procedures. J. Alcaide was the first researcher to tackle the issue, contemporaneously to the surveys. In Alcaide (2000) he showed an abrupt decrease in disposable income inequality starting at some point between 1973 and 1980, and continuing with less intensity in the following years. His corrections on the HBSs are based on the difference between total income and total expenditure data, taking the latter as more reliable (since they adjust better to the National Accounts and households may have felt less reluctant to report them). His first step thus consisted of an upwards adjustment of income to expenditure, with data aggregated by socio-economic groups,

<sup>11</sup>Other household surveys have been contrasted with the magnitudes from National Accounts with similar results. Pou and Alegre (2002) made the comparison for the *Encuesta Continua de Presupuestos Familiares*, a rotating panel stretching from 1985 to 1996, and found that the ratio in gross terms was around 62-69%, with a slight improvement over the period. (This ratio is not directly comparable to those given in the last column of table 4, which are obtained from the net magnitudes. In the case of the HBSs, the corresponding gross values would be 66.8%, 74.7% and 80.2% for the three years respectively: the ECPF thus appears more unreliable for the study of income inequality than the benchmark year surveys I am using here). Andrés and Mercader-Prats (2001) engaged in similar calculations with the 1994 European Community Household Panel. They present ratios for the different income sources which show the high reliability of wages and salaries data, and the troubles with incomes from self-employment and capital. This result coincides with our further exploration of HBSs in section 3.

Table 5: Distribution of disposable income across household deciles  
(percentage over total)

|            | Original HBSs |         |         | Alcaide |         |          | INE     |         | Estruch |         | Pena et al |         |         |
|------------|---------------|---------|---------|---------|---------|----------|---------|---------|---------|---------|------------|---------|---------|
|            | 1973-74       | 1980-81 | 1990-91 | 1973-74 | 1980-81 | 1990 (*) | 1973-74 | 1990-91 | 1973-74 | 1990-91 | 1973-74    | 1980-81 | 1990-91 |
| Decile 1   | 2.07          | 2.48    | 2.71    | 1.76    | 2.41    | 2.69     | 1.8     | 3.91    | 2.58    | 2.57    | 2.80       |         |         |
| Decile 2   | 3.90          | 4.07    | 4.33    | 3.18    | 3.98    | 3.47     | 3.1     | 4.39    | 4.03    | 4.16    | 4.49       |         |         |
| Decile 3   | 5.26          | 5.43    | 5.64    | 4.47    | 5.20    | 5.22     | 4.2     | 5.94    | 5.06    | 5.25    | 5.58       |         |         |
| Decile 4   | 6.51          | 6.69    | 6.78    | 5.11    | 6.31    | 6.32     | 5.3     | 8.09    | 6.07    | 6.26    | 6.54       |         |         |
| Decile 5   | 7.76          | 7.94    | 7.95    | 6.34    | 7.38    | 7.66     | 6.5     | 9.91    | 7.11    | 7.29    | 7.60       |         |         |
| Decile 6   | 9.06          | 9.25    | 9.23    | 8.04    | 8.80    | 8.48     | 7.9     | 9.96    | 8.32    | 8.46    | 8.74       |         |         |
| Decile 7   | 10.55         | 10.67   | 10.65   | 9.06    | 10.01   | 9.75     | 9.8     | 10.27   | 9.78    | 9.88    | 10.10      |         |         |
| Decile 8   | 12.51         | 12.57   | 12.51   | 10.09   | 11.53   | 11.78    | 12.6    | 12.16   | 11.76   | 11.76   | 11.89      |         |         |
| Decile 9   | 15.63         | 15.55   | 15.49   | 12.38   | 15.05   | 15.08    | 16.9    | 15.93   | 15.02   | 14.96   | 14.91      |         |         |
| Decile 10  | 26.75         | 25.35   | 24.70   | 39.57   | 29.23   | 28.35    | 31.9    | 19.42   | 30.78   | 29.41   | 27.35      |         |         |
| Gini index | 36.2          | 34.3    | 33.0    | 44.6    | 36.3    | 34.7     | 42.5    | 29.3    | 38.66   | 37.57   | 34.96      |         |         |

Sources: author's calculations on the basis of HBSs, Alcaide (2000), Instituto Nacional de Estadística (1977), Estruch (1996) and Pena and Callealta (1996).

The unit of analysis is the household and the income definition used corresponds to total disposable income (not per capita, not equivalized); except for Pena's study, where it is income per capita. The Gini indices given in the cited studies are calculated out of the aggregated data, and thus underestimated with respect to those obtainable from micro-data.

(\*) Alcaide's calculations for 1990 are based on a different survey, the *ECPF*, and therefore not strictly comparable to those of the HBS I am analysing here.

and later he scaled-up the corrected income figures to National Accounts.<sup>12</sup> These results have been widely accepted. Table 5 shows that his procedure provided higher inequality figures than the raw HBSs data. Since the difference is much larger for 1973-74, his calculations would depict a more powerful retreat of inequality during the years of the political transition than in the following decade. Estruch (1996) used a very similar methodology, applying it to the 1990-91 data, in his work about public spending.

Alternatively, it can be accepted as economically normal that some households consume above their yearly income, up to a certain extent. Such an approach was taken by INE's study of the 1973-74 survey, for the volume *La Renta Nacional y su Distribución 1976* (Instituto Nacional de Estadística, 1977): they accepted as "honest" those households where the difference between total expenditure (plus net savings) and income was not bigger than 5%.<sup>13</sup> With those, a log-log relationship was estimated between consumption and income and used to correct the under-reported incomes. The result was also a more unequal distribution. The authors themselves considered it as a lower cap on inequality, since "honest" families were found mostly in the poorer deciles: if expenditure-income elasticity is not constant but decreasing, the concentration of income would be bigger than estimated.

A similar procedure was applied by Pena and Callealta (1996). They first obtained under-reporting correction factors by socio-economic categories, again derived from the relationship of declared income with consumption (ranging from 1.63 to 1.11). But these were not applied directly on the total income of the household: 1.06 was assigned to salaries and 1.03 to public benefits, following the results in Díaz and Fernández (1993); which implies that the correction factors for other income sources resulted higher as a consequence. In a second step, they applied a uniform adjustment to the National Accounts. Their result is also higher inequality than in the original surveys, with a lower reduction over time than according to Alcaide.<sup>14</sup>

In the next section I present an alternative procedure to deal with income under-reporting, which leads to different conclusions.

### 3 Adjusting Household Surveys

My methodology is similar to Alcaide's and Pena's in the basic intuitions, namely using income-expenditure discrepancy and scaling-up to National Accounts. But the specific calculations differ, and so the results. I first follow Pissarides and Weber (1989) and Martínez-López (2012) to ob-

---

<sup>12</sup>The methodology is best explained in Alcaide and Alcaide (1974), where there is reference to other sources used, such as INE's Wage surveys, and to a Pareto simulation in the upper tail of the distribution.

<sup>13</sup>Argimón et al. (1987), for their study on indirect taxation in 1980, followed the same assumption. They used, however, provincial-level data, since the micro-data were not yet available at the time of their work, and I have not therefore included their estimation in the table.

<sup>14</sup>It should be noted that the reported distribution in Pena and Callealta (1996) is of income *per capita*, so the comparison with the other columns is not straightforward.

tain the relative level of under-reporting of the self-employed, using only information from the surveys. Then I resort to comparison with National Accounts, but instead of employing the aggregate disposable income I make separate contrasts for the different sources of household revenue, as suggested by Oliver (1997). This allows to obtain particular adjustment factors and therefore a more realistic view of the distribution.

### 3.1 Relative under-reporting of the self-employed

It is widely believed that the self-employed under-report their incomes both in tax assessments and income surveys. Pissarides and Weber (1989) were the first to suggest an estimation of this concealing of incomes by means of contrasting their expenditure levels (in food) with those of wage earners in household surveys. Their idea rested on the basic assumptions of accurate reporting of: (a) the incomes of wage earners and (b) the food expenditures of both kinds of households. The intuition is that wage earners can more easily know their exact income (because of its regularity) and have also less tax-fraud incentives to hide it in a survey (since they have less capacity to evade anyway, given withholding at source). Expenditures are generally known to be better declared than income in household surveys, and specially in the case of food, with ratios near 90-100% with respect to National Accounts.

Pissarides and Weber (1989) obtained for Britain in 1982 that incomes reported by the self-employed should be multiplied by a factor of 1.55 to obtain their true earnings. After them, a wide literature has undertaken similar calculations for other countries and time-periods, with some further methodological contributions (Lyssiotou et al., 2004; Johansson, 2005; Engström and Holmlund, 2009; Hurst et al., 2010; Tedds, 2010; Martínez-López, 2012). Here I follow Engström and Holmlund (2009), who calculated a factor of 1.30 for Sweden around the year 2000, and Martínez-López (2012), who estimated 1.25 for Spain in 2006-2009. Martínez-López stressed that this coefficient was *relative* to the wage earners' own under-reporting rate –something which is important in the Spanish case and in a historical analysis, where salaried workers might not be completely reliable.

The procedure is based on the estimation of an Engel curve with the following form:

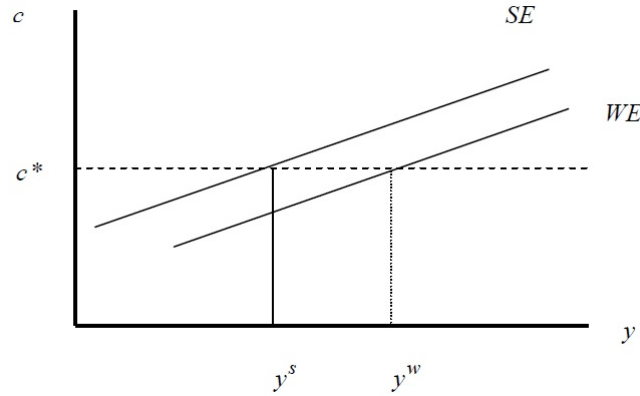
$$\ln F = \alpha + \beta \ln Y^D + \gamma SE + \delta Z + u, \quad (1)$$

$F$  being declared food expenditure,  $\alpha$  the subsistence level,  $Y^D$  total declared income,  $SE$  a dummy for self-employed households (defined as those where at least the household head or the spouse is so),  $Z$  a vector of control variables (family size, town size, and so on), and  $u$  the error term.  $\gamma$  is expected to be positive, implying an apparent higher consumption of food among the self-employed, which is interpreted as income under-reporting. The idea is clear in figure 1, where  $\gamma$  would be the vertical difference between both regression lines, and  $\beta$  the slope (estimated elasticity).

The difference between real income  $Y^R$  and declared income  $Y^D$  (in logs) is given by:

$$\ln Y^R - \ln Y^D = \frac{\beta}{\gamma}, \quad (2)$$

Figure 1: Pissarides-Weber's model



Source: Engström and Holmlund (2009).

because of the formula to calculate the slope of the regression line in figure 1 ( $\beta = \gamma / (\ln Y^R - \ln Y^D)$ ). Then we can further obtain:

$$\frac{Y^R}{Y^D} = \exp\left(\frac{\beta}{\gamma}\right) \equiv k, \quad (3)$$

$k$  being the factor by which the self-employed's declared income should be multiplied in order to obtain their real income (under the assumption that the wage earners' reporting is correct – i.e., relative to it).

Food expenditure is used as the dependent variable for various reasons: it is one of the most accurately reported expenditures in the surveys (in terms of the adjustment with National Accounts of total resulting consumption), and we can safely assume that it is less affected by preferences than other goods. Rural households are excluded from the estimation, since they might obtain a significant part of their food supply out of the market and not report it correctly. The variable  $F$  is defined as expenditure in food (excluding alcohol and tobacco) plus foodstuff self-supply and free meals provided by companies to their employees. It is thus supposed to capture total food consumption, except for meals at restaurants and similar establishments.

In order to make the results more robust, I have made an alternative estimation with energy consumption as the dependent variable. In the surveys, this item was asked for as the last bill, so it could be easier to report correctly, without the need to note down and control purchases that is associated with food expenditure questions. It is also less affected by the issue of eating at the firm, out of home and so on. The energy consumption reported is only that of the household as a family: i.e. explicitly excluding expenditures associated with unincorporated businesses.

The results of the estimation are shown in table 6. Taking the average  $k$  derived from both models, for each year, the self-employed would under-report their incomes by around 14-20% more than the recipients of salary income.<sup>15</sup>

<sup>15</sup>My results are slightly below Martínez-López (2012)'s estimation for later years. However, this should not be directly

Table 6: Regression for relative under-reporting of the Self-Employed

|              | 1973                   |                      | 1980                  |                       | 1990                 |                       |
|--------------|------------------------|----------------------|-----------------------|-----------------------|----------------------|-----------------------|
| Dep. var.    | (1)<br>ln(Food)        | (2)<br>ln(Energy)    | (3)<br>ln(Food)       | (4)<br>ln(Energy)     | (5)<br>ln(Food)      | (6)<br>ln(Energy)     |
| lnY          | 0.268***<br>(0.00841)  | 0.438***<br>(0.0106) | 0.207***<br>(0.00960) | 0.368***<br>(0.0124)  | 0.167***<br>(0.0100) | 0.280***<br>(0.0137)  |
| SE           | 0.0407***<br>(0.00999) | 0.102***<br>(0.0135) | 0.0446***<br>(0.0107) | 0.0892***<br>(0.0145) | 0.0185*<br>(0.0109)  | 0.0558***<br>(0.0148) |
| Observations | 14,442                 | 14,297               | 12,624                | 12,619                | 10,360               | 10,242                |
| R-squared    | 0.371                  | 0.236                | 0.315                 | 0.200                 | 0.312                | 0.133                 |

Robust standard errors in parentheses. \*\*\* p<0.01, \*\* p<0.05, \* p<0.1

Controls include: household size, age of household head, dummies for municipality size and survey seasonality, meals in restaurants in columns (1), (3) and (5), a dummy for cold climate in columns (2), (4) and (6), and a constant.

|           |      |      |      |      |      |      |
|-----------|------|------|------|------|------|------|
| k         | 1.16 | 1.26 | 1.24 | 1.27 | 1.12 | 1.22 |
| average k | 1.21 |      | 1.26 |      | 1.17 |      |

Source: author's calculations.

### 3.2 Scaling-up to National Accounts

The other source of correction is external information: a comparison of the totals for each type of income obtained from the surveys with those in National Accounts, which are considered more reliable for the aggregate results, and supposed to capture at least a part of the black economy. This micro-macro contrast of aggregates is a common and desirable practice at present, as stated by the Canberra Expert Group (2011). The step is usual in analysis of survey or tax data in other countries, as can be seen, for instance, in ECLAC's reports, Engel et al. (1999); Barreix (2011) and Fairfield and Jorratt (2014).

Complete separate income accounts for households are not available in the Spanish National Accounts before 1980. Data for 1973 is taken from Pena and Callealta (1996), with the exception of capital incomes, which have been approximated using the percentage of dividend and interest income in "incomes from property and enterprise" in the household sector in 1969 and 1980 (the two closest available years).<sup>16</sup>

There are some coverage differences between the surveys and the National Accounts data: namely, in the latter households appear aggregated with Private Non-Profit Institutions, and they

---

interpreted as an increase in under-reporting, since the factor is relative to the wage-earners' behaviour. A constant reporting rate of the self-employed with increased compliance of the salaried households would also be consistent with the results.

<sup>16</sup>In Pena and Callealta (1996), capital incomes seem to be underestimated. I have thus used this information provided in the National Accounts publications of the pre-1970 base (where only some household flows are present).

also include people living in collective arrangements (e.g. retirement homes), who are not present in the surveys. For an extensive discussion, see Sanz (1995). These differences are considered minor and not dealt with here.

The adjustment procedure needs to take into account that Household Surveys provide net incomes, while the figures in National Accounts are in gross terms. The corresponding taxes have therefore been subtracted from the latter before calculating the relationship between magnitudes. Imputed incomes are not corrected, since they do not mostly derive from the respondents' answers but were estimated by INE; hence, they are also extracted a priori from both sources.<sup>17</sup> Scaling-up factors for each source of income have been calculated with the following formula:

$$m_i = (X_{i,NA} - I_i - T_i)/(X_{i,HBS} - I_i), \quad (4)$$

with  $X_{i,NA}$  meaning the gross amount in National Accounts,  $I_i$  the imputed (non-monetary) incomes in category  $i$  if there are,  $T_i$  the associated taxes, and  $X_{i,HBS}$  the net amount given by the Household Budget Survey. See table 7 for the correspondence between magnitudes in both sources.

The resulting scaling-up factors  $m_i$  are shown in table 8. The procedure with the different components of the Net Operating Surplus (from now on, NOS) is slightly different. NOS includes self-employment monetary income, self-employment imputed income (not corrected) and income from real estate rentals. A total adjustment of the flows to National Accounts would be incorrect, since these include undistributed profits of unincorporated enterprises, which are not present in the surveys (recall section 2.2). The procedure applied here yields a difference of around 10% under total adjustment. It is based on the factor for self-employment obtained thanks to the Pissarides-Weber regressions:  $m_{SE}$  is the product of the previously estimated  $k$  (see subsection 3.1) and the factor for Labour income (since the equation yielded under-estimation relative to wage-earners). It has also been applied to rental incomes, which make part of the same category in National Accounts.<sup>18</sup>

As can be seen, the correction factors tend to decline over time, showing what seems to be the increasing reliability of the surveys. However, this is not the case with capital income, which has the highest estimated factors (together with private transfers), causing a strong upward adjustment on the affected households. This may be a reflection of structural and regulatory change. A decrease in capital income concentration could be accompanied by growing non-reporting: a rising number of households receiving small quantities of capital income and neglecting to include them in the surveys' questionnaires.<sup>19</sup> On the other hand, the increase in the associated tax

---

<sup>17</sup>These are non-monetary flows accruing to households, related either to wage-earning activities (in-kind compensation and meals at the workplace) or to self-employment (self-supply and housing services in owner-occupied housing).

<sup>18</sup>Income from real estate rentals is in fact only available separately for 1990.

<sup>19</sup>This source of misrepresentation of incomes in HBSs is dealt with in Engel et al. (1999) with a random imputation procedure (by deciles), the effect of which would presumably be a slight decrease in measured inequality. However, this choice is not taken here because it would be necessary to establish first what share of total misrepresentation corresponds

Table 7: Matching Household Surveys with National Accounts

| <b>Household Surveys (net)</b>  | <b>Direct Taxes</b>                 | <b>National Accounts (gross)</b> |
|---------------------------------|-------------------------------------|----------------------------------|
| Self-Employment monetary income | Self-Employed Social Contributions  | Net                              |
| Income from real estate rentals | Real Estate tax (80%)               | Operating Surplus                |
| Self-Employment imputed income  | Income tax from mixed income        | Workers' Remuneration            |
| Wages and salaries              | Income tax employment income        | Interests                        |
| Employment imputed income       | Workers' and Employers' Soc. Contr. | Land rents                       |
| Capital income                  | Income tax from capital income      | Dividends                        |
| Social Benefits                 | Unemployed Social Contributions     | Social Benefits                  |
| Private transfers               |                                     | Private transfers                |
|                                 |                                     | Insur. compensation              |

Source: Author's compilation, based on Sanz (1995) and Pena and Callealta (1996).

n/a: non applicable.



Table 8: Correction factors by sources of income

|                  | 1973 | 1980 | 1990 |
|------------------|------|------|------|
| Wages & salaries | 1.35 | 1.19 | 1.12 |
| Self-employment  | 1.63 | 1.51 | 1.31 |
| Rental income    | 1.63 | 1.51 | 1.31 |
| Capital income   | 4.77 | 3.64 | 6.85 |
| Transfers        | 2.51 | 2.05 | 4.02 |
| Social benefits  |      |      | 1.43 |

Source: Author's calculations.

The table displays the factors  $m_i$ , obtained with expression (4), which serve to scale-up the income data to the totals in National Accounts.

burden could have implied higher concealing of such incomes.<sup>20</sup>

From the surveys of 1973 and 1980 it is only possible to obtain a joint correction factor for the total of transfers (which includes social benefits together with all sorts of private flows). However, applying this number to all households equally would underestimate inequality, because benefits are better reported in the surveys than the rest of transfers, and both kinds of revenue have very distinct distributions (as shown by the separate estimates for 1990). To account for this problem, I have used a different correction factor for each decile, based on the results in 1990: for this year the survey provides both variables, allowing to obtain a different ratio total corrected transfers / total reported transfers in each decile (by composition effect). The relation of this ratio by deciles with the general factor for transfers is used in 1980 and 1973. This entails that the correction factor of transfers increases with income.<sup>21</sup>

The application of these coefficients to each type of income, at the micro-data level, yields a different compound correction factor to each household, as well as to every possible socio-economic sub-group, by composition effect. Table 9 shows the resulting factors by deciles. The profiles have a J-shape, being lower at the middle part of the distribution and attaining the highest values at the very top.

---

to each problem (under-reporting versus non-reporting).

<sup>20</sup>Another possible explanation for the raise in  $m$  would be that the total reference gross magnitudes used in 1973 and 1980 are underestimated, but it seems unlikely. An examination of the factor shares shows that the participation of capital income in household revenue increased during the decades considered here, from 5.5% in 1973 to 6.2% in 1980 and 10% in 1990. The accounts for both households and non-financial enterprises show a similar trend. However, it is possible that the data for 1973 reflect an extraordinary, circumstantial situation, since wage remuneration was increasing strongly in national income during the first half of the seventies and profits were decreasing. The 1980 data can also be thought of as depicting an economy with low profits, given the context of crisis.

<sup>21</sup>Obviously, such a procedure is not completely accurate, since the distribution of both kinds of transfers, and specially public benefits, may have changed across the period. It is however preferred to applying a single factor to all households.

Table 9: Correction factors by deciles

|           | 1973 | 1980 | 1990 |
|-----------|------|------|------|
| Decile 1  | 1.71 | 1.49 | 1.33 |
| Decile 2  | 1.58 | 1.48 | 1.28 |
| Decile 3  | 1.54 | 1.38 | 1.26 |
| Decile 4  | 1.49 | 1.37 | 1.27 |
| Decile 5  | 1.49 | 1.35 | 1.25 |
| Decile 6  | 1.47 | 1.34 | 1.27 |
| Decile 7  | 1.49 | 1.36 | 1.28 |
| Decile 8  | 1.50 | 1.39 | 1.29 |
| Decile 9  | 1.57 | 1.38 | 1.33 |
| Decile 10 | 1.74 | 1.44 | 1.55 |
| Top 1%    | 2.05 | 1.66 | 2.39 |
| Total     | 1.56 | 1.40 | 1.31 |

Source: author's calculations.

Mean factor applied (weighted average). Deciles are built on the corrected resulting disposable income.

The final outcome of the correction is a set of higher inequality estimates, compared with those resulting from the original INE data, as was originally expected. Table 10 displays the Gini indices obtained, following two different calculations. The first row shows inequality of disposable income across households, with no adjustment for household size and using them as the unit of analysis (thus giving the same importance in the estimation to a 1-member household and to a 6-member one); the second uses equivalent income and individual weights (i.e., each person is assigned the equivalent per capita income of its household and has the same importance in the estimation). The latter approach provides a better measure of inequality between individuals, but it requires some assumptions about the distribution of resources within the family and economies of scale in consumption.<sup>22</sup> Unsurprisingly, inequality is lower between individuals than between households, because larger families tend to have higher aggregate incomes.

As can be seen, the correction of under-reporting also implies a change in the observed trend of inequality. While the unadjusted data and the corrections from previous literature reviewed earlier showed an abrupt improvement in the distribution over time, the new corrected incomes show a much more slight change across these decades (around 1.5 Gini points). We can thus talk

<sup>22</sup>It is assumed that all members of a household are entitled to the same level of material well-being (that they share their income equally). Regarding the elasticity of "needs" to household size and composition, here I use the OECD scale, which attaches value 1 to the first adult, 0.7 to the subsequent ones and 0.5 to the minors in the household (up to 14 years old). The choice is consistent with empirical results based on Spanish data (Bosch-Domenech, 1991; Duclos and Mercader-Prats, 1999; Labeaga et al., 2004).

about considerable persistence in inequality.

Table 10: Spanish income inequality (1973-1990): Gini index in scaled-up data

|  | 1973  | 1980  | 1990  |
|--|-------|-------|-------|
| Disposable total income (households)       | 36.88 | 33.58 | 34.84 |
| Disposable equivalent income (individuals) | 34.59 | 32.59 | 32.95 |

Source: Author's calculations on HBSs. Equivalent incomes are obtained using the OECD scale.

## 4 The evolution of the Spanish income distribution (1973-1990): an alternative picture

After addressing the bias present in the household surveys, I have obtained a distribution of incomes that involves higher inequality levels than the original data, and a very slight equalising trend in the period 1973-1990. Based on the corrected HBSs, it cannot be said that inequality fell very substantially in the period. This result contrasts with most of the literature presented in section 2, but is not necessarily at odds with studies based on tax or macroeconomic data, which are reviewed in the following subsection.

### 4.1 Comparing with other approaches to Spanish inequality

The results in this paper differ from those previously obtained using the survey data, because I estimate a higher level of inequality and a smaller decrease over time. This includes the studies reviewed in section 2, based on the HBSs, and also those which have used a different source, the *Encuesta Continua de Presupuestos Familiares*. This is a rotating household panel also provided by INE, with quarterly data and households staying in the sample for a maximum of 2 years.

The *ECPF* consistently displays lower levels of inequality than the HBSs (*EPFs*). One reason for it might be that it suffers from a larger downward bias, because of sample size and the definition of income employed (notably excluding certain capital incomes). According to some reputable sources, this results in its low reliability for the study of inequality (Eurostat, 1999; Goerlich and Mas, 1999). Its higher discrepancy with respect to National Accounts can be seen in Pou and Alegre (2002).

For this reason, differences between my results and those of analyses based on the *ECPF* are to some extent not surprising. Both Oliver et al. (2001) and Pijoan-Mas and Sanchez-Marcos (2010), for example, provide an account of falling inequality between 1985 and 1996/2000, partially overlapping with the period analysed here.<sup>23</sup> The difference between my results and theirs, however,

---

<sup>23</sup>Oliver et al. (2001) do acknowledge the limitation of under-reporting in their source, around 30% of the National Accounts data for disposable income, although decreasing over time.

is not completely irreconcilable: in fact, a decrease in inequality in the second half of the eighties could be compatible with a general stability, when the whole decade is considered (specially knowing that a whole cycle of recession and growth took place during the eighties, and rates of unemployment were similar at the beginning and the end of the decade, at around 15%). Falling inequality in labour market revenues of household heads is also found for the entire decade in the HBSs by Abadie (1997) and in my scaled-up data.<sup>24</sup>

My results can also be compared with studies on the evolution of inequality based on other kinds of data. Prados de la Escosura (2008) provided a long-run estimation based on a macroeconomic approach, calculating dispersion within and between the incomes of “workers” and “capitalists”. His series show a rapid decrease in inequality in Spain between the mid 1950s and the mid 1960s, followed by a much slower diminishing trend since then and until the second half of the 1990s, when inequality would have started to go upwards again. The persistence I obtain is therefore quite consistent with Prados de la Escosura’s calculations.

For the post-transition period it is also possible to use income tax data and assess the evolution of inequality in taxable income. By definition, however, the levels and trend do not need to coincide with those of disposable income: between both lie direct taxes, transfers and the impact of fraud. There are also other methodological differences, discussed in Ayala and Onrubia (2001): generally, tax-based studies use the taxpayer as the unit of analysis (as opposed to the household, and without applying equivalence scales) and have different universes (given by the effective income threshold to personal direct taxation). This category of taxpayers was also changing over the years: new taxpayers were coming in because the tax was being introduced, and also as an effect of fiscal drag. All of this explains why tax data generally show a higher level of inequality than survey data, and a worsening in (reported taxable) income distribution during the eighties (Castañer, 1991; Lambert and Ramos, 1997; Ayala and Onrubia, 2001). The study closest to our discussion is that by Onrubia et al. (2007), which includes calculations for the “fiscal household” (thus homogenizing the periods before and after the introduction of the separate filing option for married couples). The pre-tax income Gini index (taxable base with some adjustments) was found to increase continuously from 1982 to 1991 (31.68 to 42.00).

Alvaredo and Saez (2009) studied top income shares, obtaining the revenues from tax data and the population total (denominator) from National Accounts (therefore, their approach has the same comparability problems with my estimates, namely different income concept and no equalization). Their results show that the top 0.1% share was fairly constant over the 1960s and 1970s (around 1.87%), with concentration starting to increase in the second half of the eighties

---

<sup>24</sup>I have calculated the indicator defined by Pijoan-Mas and Sanchez-Marcos (2010): head of household earnings (considering labour income and 2/3 of self-employment income). The results show decreasing Gini indices: from 34 to 32 in market earnings, and from 32 to 29 considering also unemployment benefits. These numbers are quite compatible with those of the cited authors in trends, but at a slightly higher level. Since these results do not contrast with Abadie (1997)’s with the original data (which, on the other hand, focus on individuals not suffering from unemployment), we can infer that differences in under-reporting do not affect the evolution of *labour earnings* inequality as much as it does for other incomes.

(2.14% in 1990).<sup>25</sup> The same trend is shown in the share of the top 1% (7.5% to 8.37% in 1981 and 1990) and the top 10% (32.61% to 35.35% in the same years). It should come as no surprise that the figures are dissimilar for disposable income: in my work, I obtain for the top 1% of households 6.47% in 1973, 5.99% in 1980 and 7.15% in 1990.

The international comparison picture is also affected. Previous work has usually concluded that Spain started with a higher-than-average inequality level with respect to developed countries and converged during this period. This was a consequence of the equalising trend found in Spain, together with the opposite evolution in several other economies in the eighties (Ayala et al., 1996; Pena and Callealta, 1996). The conclusion is thus in line with the process of Europeanization during the years of the democratic transition and the entry at the EEC.

Using my new data, this result can be nuanced. In table 11 I show the Gini indices for disposable per capita income in selected OECD countries (this is not my preferred estimate shown previously, but it is used here for the sake of comparison). It can be seen that Spain was similar to Italy, the US and Canada at the beginning, with significantly higher inequality than the UK, Germany and Sweden; by 1990, the country remained in the first positions of the ranking. The most significant evolution was the growth in income concentration in the UK, which joined the more unequal countries in the group, and the reverse trend in Italy and Canada. Throughout the period, Spain remained more unequal than the European chore, with no significant convergence to it.<sup>26</sup>

Table 11: Gini index of disposable per capita income among individuals

| Year | Canada | France | Germany | Italy | Spain | Sweden | UK   | US   |
|------|--------|--------|---------|-------|-------|--------|------|------|
| 1973 | 36.6   | -      | 29.9    | 40.0  | 36.6  | 24.1   | 29.4 | 40.0 |
| 1980 | 32.3   | 31.6   | 28.3    | 33.2  | 34.8  | 22.8   | 28.7 | 40.3 |
| 1990 | 31.2   | 31.7   | 28.8    | 30.5  | 34.8  | 24.9   | 36.5 | 42.8 |

Source: Milanovic (2013a), except for Spain (author's calculations).

Notes: International data correspond to the nearest available year. The indices for the US are not strictly comparable, since they refer to households instead of individuals.

<sup>25</sup>Tables B2 and B3 in Alvaredo and Saez (2009). There are only three observations in the period 1961-1981, because of problems in the availability of the original data.

<sup>26</sup>Note that the Spanish data are the result of the under-reporting correction in this study, while those of other countries are compiled by Milanovic (2013a) from official agencies or international income distribution databases. This, of course, might affect the comparison if incomes are also significantly miss-reported in other countries. However, the problem in the Spanish historical HBSs seems unlikely to affect the surveys of the more developed nations to a similar extent, with the possible exception of Italy.

## 4.2 Relative inequality and its composition

The inspection of decile shares based on the corrected disposable income data allows a deeper analysis of the evolution of income distribution. In table 12 inequality among households is shown to have been quite stable over these decades (consistent with the Gini indices in the first row of table 10). Interpersonal inequality, which is approached by the distribution of equivalent income in columns 5-7, is slightly lower. In any case, the absence of a clear trend remains. The bottom-half deciles increased their share, but the changes are small and erratic.

Table 12: Shares of disposable income among deciles

|           | Total income - households |        |        | Equivalent income - individuals |        |        |
|-----------|---------------------------|--------|--------|---------------------------------|--------|--------|
|           | 1973                      | 1980   | 1990   | 1973                            | 1980   | 1990   |
| Decile 1  | 2.36%                     | 2.82%  | 2.49%  | 3.01%                           | 3.01%  | 2.79%  |
| Decile 2  | 4.07%                     | 4.42%  | 4.25%  | 4.51%                           | 4.75%  | 4.76%  |
| Decile 3  | 5.22%                     | 5.58%  | 5.44%  | 5.56%                           | 5.85%  | 5.82%  |
| Decile 4  | 6.31%                     | 6.71%  | 6.50%  | 6.52%                           | 6.81%  | 6.75%  |
| Decile 5  | 7.42%                     | 7.81%  | 7.61%  | 7.50%                           | 7.83%  | 7.74%  |
| Decile 6  | 8.66%                     | 9.01%  | 8.84%  | 8.62%                           | 8.90%  | 8.80%  |
| Decile 7  | 10.15%                    | 10.42% | 10.29% | 9.98%                           | 10.23% | 10.03% |
| Decile 8  | 12.15%                    | 12.25% | 12.13% | 11.87%                          | 12.02% | 11.73% |
| Decile 9  | 15.42%                    | 15.16% | 15.12% | 14.74%                          | 14.78% | 14.62% |
| Decile 10 | 28.23%                    | 25.81% | 27.32% | 27.69%                          | 25.81% | 26.96% |
| Top 1%    | 6.47%                     | 5.99%  | 7.12%  | 6.64%                           | 6.07%  | 7.31%  |

Source: Author's calculations.

Equivalent incomes are obtained using the OECD scale.

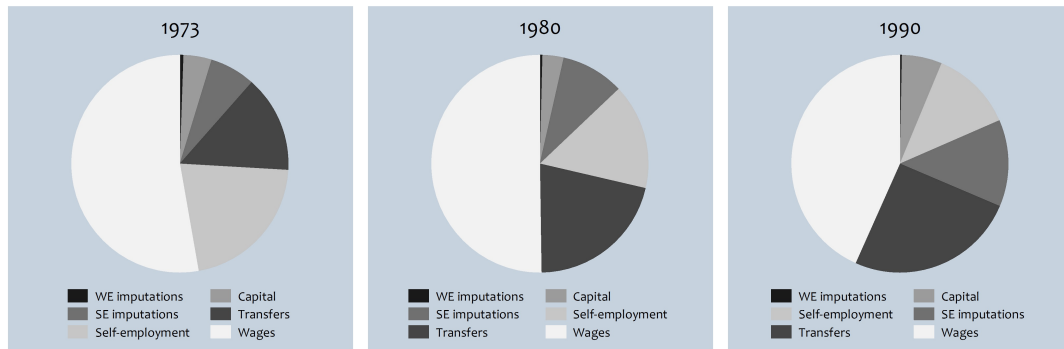
It is nonetheless most likely that the *roots* of inequality in the economy changed during these decades. The capital-labour ratio had been decreasing in the last years of the dictatorship as a short-term response to the crisis, and could have increased again later because of liberalization. Most advanced industrial economies have experienced a recent increase in wage and salaries dispersion. These trends, together with the increase in unemployment, could have counteracted to some extent the equalizing force of public benefits expansion and the introduction of progressivity in the tax system.<sup>27</sup>

Entering such debate in depth is out of the scope of this paper, but the decomposition of dis-

<sup>27</sup>The tax reform starting in 1977 is an important cornerstone in the political transition, since it meant to base the collection of taxes on the principles of equity and efficiency, the first meaning progressive taxation. The implementation of the new system, however, does not seem to have been very successful in this aspect (Torregrosa, nd). On the other hand, it should be noted that our three observations are to a certain extent also a result of short-term fluctuations: 1973 was the culmination of the pre-oil-shock growth in the country, while 1980 was a period of economic distress, and in 1990 the country was back on the ascending side of the cycle. We cannot make strong conclusions out of them.

possible income can provide an idea of the forces behind inequality change. This is done in figure 2. Apart from wages, self-employment income, capital income and transfers, two kinds of imputed incomes are included. These are non-monetary flows accruing to households, that have been given an approximate value in the survey. 'WE imputations' (those related to wage-earning activities) include in-kind compensation and meals at the workplace, while 'SE imputations' (related to self-employment) are self-supply and housing services in owner-occupied housing.<sup>28</sup>

Figure 2: Composition of disposable income



Source: Author's calculations, using *sgini* module for STATA by Philippe van Kerm.

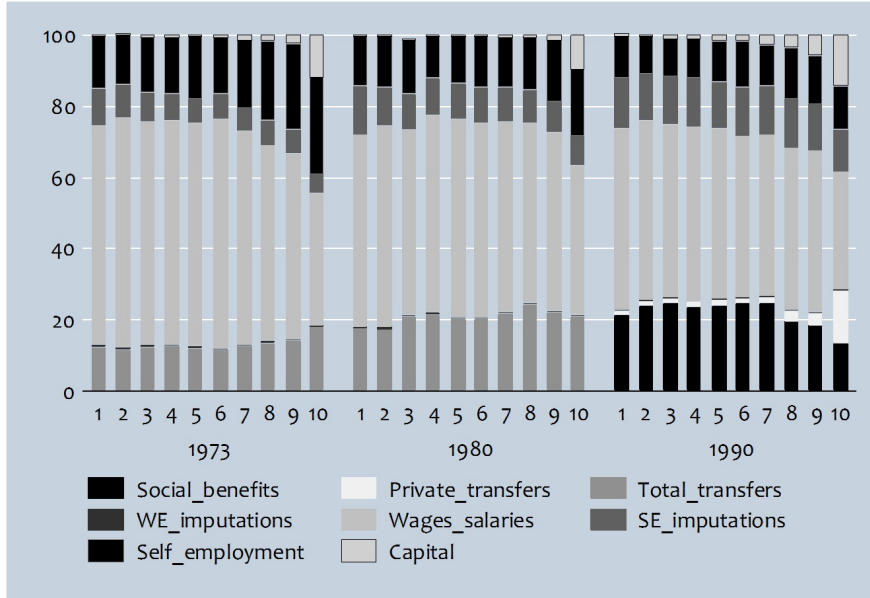
Employment incomes were clearly the main components of disposable household resources. But their share decreased over time (accounting each year for 53%, 50% and 43%, respectively). The items gaining weight were mainly transfers (due to the development of the welfare state: total transfers increased from 14% to 25%) and capital income (from 4% to 6%). Because capital income is concentrated at the top, while public benefits go mainly to the bottom, both changes could have counteracting effects on total inequality. Also self-employment imputations had a growing participation (from 7% to 13%), mainly due to the imputed rentals from owner-occupied housing.

This general composition of disposable income is of course very variable along the social ladder, as can be seen in figure 3. In the bottom deciles transfers and salary income make up most revenues. Social benefits and private transfers are regrettably not disaggregated in the 1973 and 1980 data, but the progressive nature of the first can be seen clearly in 1990. Wages have maximum participation in the middle deciles, and self-employment income is somewhat skewed to the top in the first years. Revenues from capital are the most concentrated: almost absent in the

<sup>28</sup>The imputation of income from owner-occupied housing is conceptually important but empirically complicated. Using only monetary components to measure the possible standards of living of the households can be highly misleading if renters coexist with owner-occupiers, which is of course the case here (although around 80% of the households fall in the second category). The survey includes this variable, which is an approximation to the rent a household would pay if it rented its dwelling. The calculations are certainly not flawless, and moreover it should only be imputed in the percentage that the house is paid (i.e., 100% if the family totally owns its house, 50% if half of it is still owed to the bank). This adjustment is not possible here because of lack of data. However, I consider necessary to include this element, specially in a context marked by rising prices of dwelling and with the housing bubble in the horizon.

lower classes, they constitute over 10% of income for the upper deciles. This pattern is similar in other countries (e.g. Piketty, 2003).

Figure 3: Income composition by deciles



Source: author's calculations, using sgini module for STATA by Philippe van Kerm.  
In all cases incomes are equivalised by household size, using the OECD scale.

Figure 4 plots inequality for each component of income, following the decomposition method originally developed by Lerman and Yitzhaki (1985). It shows that employment incomes got slightly more concentrated over the period: wages and salaries went from 50.6 to 53.2 Gini points, and self-employment income from 84.1 to 86.6.<sup>29</sup> The element with the most uneven distribution is capital income (99-95 Gini points), the increasing participation of which also pulled up total inequality. These forces were offset by a more homogeneous distribution of SE imputations and transfers, income sources which, as we have seen, experienced substantial growth over the period.

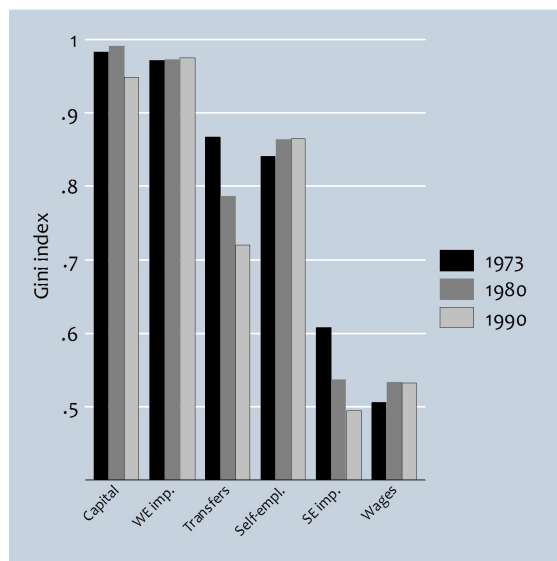
### 4.3 Further nuances: levels and extraction

The apparent stability of the Gini index does not imply the absence of several interesting distributive changes. A different image emerges if we depart from inequality and take a look at the levels of income. Table 13 displays mean disposable per capita equivalent income by deciles, in constant 1990 *pesetas*. It allows observing that all deciles experienced an increase in their purchasing

<sup>29</sup>Recall that this is the Gini index for each component *over the whole population*, not only over the households which do have each kind of income. If we consider only families with salary income, for example, we get a quite constant Gini index of 34-36. Both computations are showing different facts (and none of them is wage inequality among the workforce, which would require working with individual-level data).



Figure 4: Gini indices for components of disposable income



Source: Author's calculations, using *sgini* module for STATA by Philippe van Kerm. In all cases incomes are equivalised by household size, using the OECD scale. 'WE imp.' means imputed incomes from labour activity, while 'SE imp.' refers to non-monetary self-employment incomes such as that from owner-occupied housing.

power during the period.<sup>30</sup> The profiles of growth are however dissimilar in the two sub-periods: while during the seventies growth was higher at the lower-mid levels, in the nineties it was the extremes which benefitted the most (pointing towards the expansion in Welfare State's transfers in the case of the low-income households). If we look at the top 1%, we even find stagnation in the first sub-period (the oil-crisis decade) and a very significant recovery in the second. The ratios in the last rows confirm the same impression of a weak decrease in economic distances.

Let us recall that the Gini index and other related indicators measure *relative* inequality (i.e., independent from the scale: constant if all incomes are multiplied by the same factor). If absolute differences in income are also thought to be important, we can calculate an absolute Gini, which is the same index without normalization to the mean (as put forward by Ravallion, 2003, following Kolm, 1976). Doing this exercise with the three years, we get an increase in the absolute inequality index of 25% between 1973 and 1990. Relative economic distances did not change that much, but they did in absolute terms, in actual consumption capacity.<sup>31</sup>

Relative inequality can in turn be put in perspective by evaluating the "extractive" nature of the economy. Following the intuition in Milanovic et al. (2011), as income per capita grows,

<sup>30</sup>Disposable income is an indicator of consumption capacity of the households, but it should be borne in mind that it is still subject to indirect taxes, so changes in consumption taxation also affect final material well-being.

<sup>31</sup>The issue is more complicated. If we take into account diminishing utility of incomes, it can be argued that absolute differences in income would be more accurately expressed after some kind of functional transformation to reflect it.

Table 13: Levels and growth of Disposable Per Capita Equivalent Income

|            | Mean (constant 1990 ptas) |           |           | Increase |         |         |
|------------|---------------------------|-----------|-----------|----------|---------|---------|
|            | 1973                      | 1980      | 1990      | 1980/73  | 1990/80 | 1990/73 |
| Decile 1   | 256,396                   | 277,955   | 359,219   | 8.4%     | 29.2%   | 40.1%   |
| Decile 2   | 384,786                   | 438,816   | 520,182   | 14.0%    | 18.5%   | 35.2%   |
| Decile 3   | 473,659                   | 540,912   | 652,221   | 14.2%    | 20.6%   | 37.7%   |
| Decile 4   | 556,450                   | 630,107   | 754,046   | 13.2%    | 19.7%   | 35.5%   |
| Decile 5   | 639,019                   | 723,320   | 863,084   | 13.2%    | 19.3%   | 35.1%   |
| Decile 6   | 735,176                   | 822,136   | 974,459   | 11.8%    | 18.5%   | 32.5%   |
| Decile 7   | 851,305                   | 945,044   | 1,111,928 | 11.0%    | 17.7%   | 30.6%   |
| Decile 8   | 1,011,968                 | 1,111,931 | 1,302,864 | 9.9%     | 17.2%   | 28.7%   |
| Decile 9   | 1,256,145                 | 1,366,004 | 1,615,610 | 8.7%     | 18.3%   | 28.6%   |
| Decile 10  | 2,362,682                 | 2,385,219 | 3,004,124 | 1.0%     | 25.9%   | 27.1%   |
| Top 1%     | 5,688,115                 | 5,681,407 | 8,444,164 | -0.1%    | 48.6%   | 48.5%   |
| Total      | 852,657                   | 924,084   | 1,117,712 | 8.4%     | 21.0%   | 31.1%   |
| Ratio 10/1 | 9.21                      | 8.58      | 8.36      | 0.11     | 0.89    | 0.68    |
| Ratio 10/5 | 3.70                      | 3.30      | 3.48      | 0.07     | 1.34    | 0.77    |
| Ratio 5/1  | 2.49                      | 2.60      | 2.40      | 1.57     | 0.66    | 0.87    |

Source: Author's calculations, using GDP deflators from Prados de la Escosura (2003). Deciles of individuals based on Disposable per capita equivalent income, obtained with the OECD scale.

society's surplus gets bigger and therefore a higher inequality becomes feasible. This is the idea behind the concept of the "inequality possibility frontier" (IPF), the maximum level of inequality where all surplus (income above the physiological subsistence minimum of the population) is appropriated by one individual. The "inequality extraction ratio" (IER) can then be calculated as actual inequality divided by the IPF. In our case, given the substantial economic growth seen in the country during these years, it seems quite risk-free to venture that a persistent Gini was compatible with a decreasing extraction ratio.

It can be argued that using a physiological subsistence minimum is not very accurate, since human 'needs' would increase as does the general living standard in a society. In that sense, Milanovic (2013b) introduces a variation in the calculation of the IPF, in which the maximum possible Gini is a function of the "social minimum", which takes into account the mean income of the population. The definition for the maximum possible Gini is:

$$Ginimax = 1 - (\sigma/m), \quad (5)$$

where  $m$  is mean income (per capita GDP) and  $\sigma$  is the social minimum, defined as:

$$\sigma = s\alpha^b, \quad (6)$$

with  $\alpha = m/s$  representing how far a society is from physiological subsistence  $s$  (considered as 400\$ PPP 1990), and  $b$  being the elasticity of the social minimum with respect to  $\alpha$ . We can then

obtain:

$$Ginimax = 1 - (\alpha^b/\alpha) \quad (7)$$

Milanovic uses the value  $b=0.5$  as an approximation (based on some evidence from official poverty lines and subjective poverty surveys), which simplifies the expression to:

$$Ginimax = 1 - (1/\sqrt{\alpha}) \quad (8)$$

Using that formula, I could confirm the intuition that persistent relative inequality coexisted with a decreasing extraction ratio (table 14).

Table 14: Inequality and Extraction

|          | Spain |       |       | UK    |       |       | USA   |       |       |
|----------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
|          | 1973  | 1980  | 1990  | 1964  | 1979  | 1991  | 1967  | 1979  | 1991  |
| $\alpha$ | 26.16 | 32.18 | 42.98 | 31.9  | 43.9  | 53.9  | 47.8  | 62.6  | 76.2  |
| Gini     | 36.35 | 34.70 | 35.18 | 33.6  | 28.8  | 35.9  | 39.7  | 40.4  | 42.8  |
| Ginimax  | 80.45 | 82.37 | 84.75 | 82.2  | 84.9  | 86.4  | 85.5  | 87.4  | 88.5  |
| IER (%)  | 45.18 | 42.13 | 41.51 | 40.88 | 33.92 | 41.55 | 46.43 | 46.22 | 48.36 |

Source: for Spain, author's calculations with GDP per capita from Prados de la Escosura (2003). For UK and the US, Milanovic (2013b).

$\alpha$  is calculated as the relation between GDP per capita and the physiological subsistence level. IER, the Inequality Extraction Ratio, is the actual Gini as a percentage of the maximum "feasible" Gini (both with per capita income).

At the beginning of the period, Spain displayed a level of inequality just a little above that of the UK, but higher extraction (because of the "size of the pie" being smaller). Around 1990, IERs in Spain and the UK had got very similar, as a result of a different trend: while the English-speaking countries show an increase in both indicators during the eighties, in Spain the capacity of the elites to extract surplus seems to have been getting slightly weaker. Most of the change happened during the first decade, which witnessed the political transition, but still maybe very soon to evaluate the effects of the associated reforms. We could therefore ask ourselves whether there is a significant relationship between extraction and regime change. The question will be tackled by future research.

## 5 Concluding remarks

In this paper I have analysed the sources available on disposable income distribution in Spain during the years surrounding the transition to democracy. The main data come from the Household Budget Surveys conducted by the INE, which contain very rich information but need to be used with caution. It is widely known that they suffer from severe under-reporting –and, furthermore, that this is not homogeneous across income sources. Such a problem entails that the under-estimation of incomes is not homogeneous across income levels, biasing the inequality indices readily obtained from the data.

I have performed a two-step correction procedure, trying to identify under-reporting first with an Engel's curve approach (contrasting the self-employed with the wage-earners in their incomes and food/energy expenditure) and then with an aggregate adjustment to the magnitudes of the household sector given in National Accounts. The results allow to question the conventional wisdom that inequality strongly diminished in Spain during these decades. The Gini indices of all surveys are pulled up by the correction, and the trend across the years significantly weakens.

This leads directly to asking another question. Did transition to democracy not introduce significant distributive improvements? Political economy theory would expect from democracy an inclination to favour the lower classes (at least, relative to a right-wing dictatorship as Spain had recently suffered), via labour market regulation, welfare state benefits, and progressive taxation. We do witness an increase in the importance of transfers received by the households at the bottom of the distribution, reflecting Welfare State development in the years after 1977. But they did not outdo forces pulling in the opposite direction. The tax system did not turn out progressive (Torregrosa, nd). The absolute gains from growth went mostly to the upper classes.

Economic growth and decline in inequality in the years after 1950 were suggested to facilitate the transition in the 1970s. Prados de la Escosura (2008) interpreted in this way the elimination of absolute poverty and the growth of the middle class, which would have permitted the stabilization of democracy, contrary to what happened in the interwar period. But that evolution does not seem to have gone much further. Liberalization brought about new distributive forces, while in the context of general economic distress in Europe, the new political system did not turn out to disproportionately favour the less well-off. At least, it could not effectively counteract market forces towards growing inequality.

This is, of course, a political choice, reflecting the equilibrium between classes or interest groups in the young parliamentary state. In that sense, the lack of economic equalization could be enlightening about the access to political power. Future work will explore the relationship between inequality and political transitions in a broader comparative context, with special attention to Latin American countries, which could provide valuable insights.

## References

- Abadie, A. (1997). Changes in spanish labor income structure during the 1980's: a quantile regression approach. *Investigaciones Económicas* 21, 253–272.
- Accardo, J., V. Bellamy, G. Consales, M. Fesseau, S. Laidier, and E. Raymaud (2009). Inequalities between households in the national accounts – Breakdown of household accounts. INSEE Working Paper.
- Acemoglu, D., S. Naidu, P. Restrepo, and J. A. Robinson (2013, Dec). Democracy, Redistribution and Inequality. NBER Working Papers 19746, National Bureau of Economic Research, Inc.
- Alcaide, A. and J. Alcaide (1974). Metodología para la estimación de la distribución personal de la renta en España en 1970. *Hacienda Pública Española* (26), 55–63.
- Alcaide, J. (2000). La renta nacional de España y su distribución. Serie años 1898 a 1998. In J. V. Fuertes (Ed.), *1900-2000. Historia de un esfuerzo colectivo: cómo España superó el pesimismo y la pobreza*, pp. 810, 741. Madrid: Fundación BSCH.
- Alvaredo, F. (2011). A note on the relationship between top income shares and the gini coefficient. *Economics Letters* 110(3), 274 – 277.
- Alvaredo, F. and E. Saez (2009). Income and wealth concentration in Spain from a historical and fiscal perspective. *Journal of the European Economic Association* 7, 1140–1167.
- Andrés, L. and M. Mercader-Prats (2001, Jan). Sobre la fiabilidad de los datos de renta en el Panel de Hogares de la Unión Europea (PHOGUE, 1994). Working Papers wp0102, Department of Applied Economics at Universitat Autònoma of Barcelona.
- Argimón, I., J. M. González-Páramo, and R. Salas (1987). ¿Quién pagaba los impuestos indirectos en España? Un ejercicio de reparto de la carga impositiva indirecta en 1980. *Hacienda Pública Española* 104, 99–127.
- Atkinson, A. (2000). The changing distribution of income: evidence and explanations. *German Economic Review* 1, 3–18.
- Atkinson, A. (2007). Measuring top incomes: Methodological issues. In A. Atkinson and T. Piketty (Eds.), *Top Incomes over the 20th Century: a Contrast Between Continental European and English-Speaking Countries*. Oxford: Oxford University Press.
- Attanasio, O. P. (1999). Consumption. In J. B. Taylor and M. Woodford (Eds.), *Handbook of Macroeconomics*, Volume 1 of *Handbook of Macroeconomics*, Chapter 11, pp. 741–812. Elsevier.
- Ayala, L., A. Jurado, and F. Pedraja (2006). Desigualdad y bienestar en la distribución intrateritorial de la renta, 1973-2000. *Investigaciones Regionales* 8, 5–30.

- Ayala, L., R. Martínez, and J. Ruiz-Huerta (1996). La distribución de la renta en España desde una perspectiva internacional: tendencias y factores de cambio. In *La desigualdad de recursos. II Simposio sobre Igualdad y Distribución de la Renta y Riqueza*, Volume VI, pp. 315–440.
- Ayala, L. and J. Onrubia (2001). La distribución de la renta en España según datos fiscales. *Papeles de Economía Española* 88, 89–112.
- Barreix, A. (Ed.) (2011). *Equidad fiscal en Brasil, Chile, Paraguay y Uruguay*. Washington: BID-Eurosocial.
- Bosch-Domenech, A. (1991). Economies of scale, location, age, and sex discrimination in household demand. *European Economic Review* 35, 1589–1595.
- Budd, E. and D. Radner (1975). The Bureau of Economic Analysis and Current Population Survey Size Distributions: Some Comparisons for 1964. In J. D. Smith (Ed.), *The Personal Distribution of Income and Wealth*, Volume 39 of *Studies in Income and Wealth*, pp. 449–558. New York: Columbia University Press.
- Buhmann, B., L. Rainwater, G. Schmaus, and T. Smeeding (1988). Equivalence Scales, Well-Being, Inequality, and Poverty: Sensitivity Estimates across Ten Countries Using the Luxembourg Income Study (LIS) Database. *Review of Income and Wealth* 34(2), 115–42.
- Canberra Expert Group (2011). *Handbook on Household Income Statistics* (second ed.). Geneva: United Nations Economic Commission for Europe.
- Castañer, J. (1991). El efecto redistributivo del IRPF 1982-1988. *Hacienda Pública Española* II, 131–143.
- Cowell, F. A., J. Litchfield, and M. Mercader-Prats (1999). Income Inequality Comparisons with Dirty Data: The UK and Spain during the 1980s. STICERD - Distributional Analysis Research Programme Papers 45, Suntory and Toyota International Centres for Economics and Related Disciplines, LSE.
- Cutanda, A. (2003). An empirical investigation of the effect of borrowing constraints on Spanish consumption. *Spanish Economic Review* 5, 63–84.
- Díaz, C. and R. Fernández (1993). El fraude en las rentas del trabajo. Documentos de trabajo, Instituto de Estudios Fiscales.
- Del Río, C. and J. Ruiz-Castillo (1996). *Ordenaciones de bienestar e inferencia estadística. El caso de las EPF de 1980-81 y 1990-91*, Volume 6. Madrid: Fundación Argentaria.
- Duclos, J.-Y. and M. Mercader-Prats (1999). Household Needs and Poverty: With Application to Spain and the U.K. *Review of Income and Wealth* 45(1), 77–98.
- Easterly, W. R. (2004). Globalization, inequality, and development: The big picture. *Monetary and Economic Studies* 22(S1), 57–87.

- ECLAC (1991). *Magnitud de la pobreza en América Latina en los años ochenta*. Number 81 in Estudios e informes de la CEPAL. Santiago de Chile.
- Engel, E., A. Galetovic, and C. Raddatz (1999). Taxes and income distribution in Chile: some unpleasant redistributive arithmetic. *Journal of Development Economics* 59(1), 155–192.
- Engström, P. and B. Holmlund (2009). Tax evasion and self-employment in a high-tax country: evidence from Sweden. *Applied Economics* 41, 2419–2430.
- Estruch, A. (1996). *Los efectos redistributivos del gasto social de las administraciones centrales y autonómicas*, Volume VII, Sector Público y Redistribución, pp. 261–290. Madrid: Fundación Argentaria.
- Eurostat (1999). ECHP data quality. DOC.PAN 108/99, ECHP Working Group, European Commission, Luxemburg.
- Fairfield, T. and M. Jorratt (2014). Top income shares, business profits, and effective tax rates in contemporary Chile. Technical Report 17, International Centre for Tax and Development.
- Fixler, D. and D. Johnson (2012). Accounting for the Distribution of Income in the U.S. National Accounts. Paper prepared for the NBER Conference on Research in Income and Wealth “Measuring Economic Stability Progress Conference”, September 30.
- Garde, J. A., R. Martínez, and J. Ruiz-Huerta (1996). *Los estudios sobre distribución de la renta en España: fuentes, resultados y perspectivas de futuro*, Volume VI, pp. 257–314. Madrid: Fundación Argentaria.
- Goerlich, F. and M. Mas (1999, september). Medición de la desigualdad: contribución a una base de datos regional. Technical report, Instituto Valenciano de Investigaciones Económicas.
- Goerlich, F. J. and M. Mas (2001). Inequality in Spain 1973-91: Contribution to a Regional Database. *Review of Income and Wealth* 47(3), 361–78.
- Gradín, C. (2002). Polarization and Inequality in Spain: 1973-1991. *Journal of Income Distribution* 11(1-2), 34–52.
- Gradín, C., O. Cantó, and C. del Río (2008, May). Inequality, poverty and mobility: Choosing income or consumption as welfare indicators. *Investigaciones Económicas* 32(2), 169–200.
- Hurst, E., G. Li, and B. Pugsley (2010, Nov). Are Household Surveys Like Tax Forms: Evidence from Income Underreporting of the Self Employed. NBER Working Papers 16527, National Bureau of Economic Research, Inc.
- Instituto Nacional de Estadística (1977). *La Renta nacional y su distribución 1976*. Madrid: INE.
- Instituto Nacional de Estadística (1983). *Contabilidad Nacional de España. Base 1970. Años 1970-1980, 1981 provisional y avance de 1982*. Madrid: Instituto Nacional de Estadística.

- Instituto Nacional de Estadística (2014).
- Johansson, E. (2005). An estimate of self-employment income underreporting in Finland. *Nordic Journal of Political Economy* 31, 99–109.
- Kolm, S.-C. (1976). Unequal inequalities I. *Journal of Economic Theory* 12, 416–442.
- Krugman, P. R. (2000). Technology, trade and factor prices. *Journal of International Economics* 50(1), 51–71.
- Kuznets, S. (1955). Economic Growth and Income Inequality. *The American Economic Review* 45(1), 1–28.
- Labeaga, J., I. Preston, and J. Sanchis-Llopis (2004). Demanda y escalas de equivalencia: evidencia para España. *Cuadernos Económicos del ICE* 68, 63–87.
- Lambert, P. and X. Ramos (1997). Horizontal inequity and vertical redistribution. *International Tax and Public Finance* 4, 25–37.
- Lerman, R. I. and S. Yitzhaki (1985). Income Inequality Effects by Income Source: a new approach and applications to the United States. *The Review of Economics and Statistics* 67(1), 151–56.
- Levy, F. and P. Temin (2007). Inequality and Institutions in 20th Century America. NBER Working Papers 13106, National Bureau of Economic Research, Inc.
- Lyssioutou, P., P. Pashardes, and T. Stengos (2004, 07). Estimates of the black economy based on consumer demand approaches. *Economic Journal* 114(497), 622–640.
- Martín-Guzmán, P., M. Toledo, N. Bellido, J. López, and N. Jano (1996). *Encuesta de Presupuestos Familiares. Desigualdad y Pobreza en España. Estudio basado en las Encuestas de Presupuestos Familiares de 1973-74, 1980-81 y 1990-91*. Madrid: Instituto Nacional de Estadística and Universidad Autónoma de Madrid.
- Martínez-López, D. (2012). The underreporting of income by self-employed workers in Spain. *SERIEs* 4(4), 353–371.
- McColl, B., J. Billing, B. Kindermann, and H. Burgess (2010). Micro and macro economic estimates for Australian households: Recent developments and future directions. Paper presented at the IARIW annual meetings, August 2010.
- Meltzer, A. and S. Richard (1981). A Rational Theory of the Size of Government. *Journal of Political Economy* 89, 914–927.
- Milanovic, B. (2013a). All the Ginis database, The World Bank, <http://econ.worldbank.org/projects/inequality>. Downloaded 25/10/2013.
- Milanovic, B. (2013b). The inequality possibility frontier: extensions and new applications. Policy Research Working Paper Series 6449, The World Bank.



- Milanovic, B., P. H. Lindert, and J. G. Williamson (2011). Pre-industrial inequality. *Economic Journal* 121(551), 255–272.
- Mistiaen, J. and M. Ravallion (2003). Survey compliance and the distribution of income. Policy Research Working Paper 2956, World Bank, Washington DC.
- Morelli, S., T. Smeeding, and J. Thompson (2014, Mar). Post-1970 Trends in Within-Country Inequality and Poverty: Rich and Middle Income Countries. CSEF Working Papers 356, Centre for Studies in Economics and Finance (CSEF), University of Naples, Italy.
- Neri, A. and R. Zizza (2010). Income reporting behaviour in the SHIW. Temi di Discussione 777, Bank of Italy.
- Oliver, J. (1997). Ingreso, consumo y ahorro de las familias: propuesta de una metodología para la explotación de la encuesta de presupuestos familiares. Documentos de trabajo, FIES.
- Oliver, J., X. Ramos, and J. Raymond-Bara (2001). Anatomía de la distribución de la renta en España, 1985-1996: la continuidad de la mejora. *Papeles de Economía Española* 88, 67–88.
- Onrubia, J., M. C. Rodado, S. Díaz, and C. Pérez (2007). Progresividad y redistribución a través del IRPF español: Un análisis de bienestar social para el periodo 1982-1998. *Hacienda Pública Española* 183, 81–124.
- Pena, J. B. and F. J. Callealta (1996). *Distribución personal de la renta en España: corrección y modelización de la información básica: desigualdad y análisis*. Madrid: Pirámide.
- Pijoan-Mas, J. and V. Sanchez-Marcos (2010). Spain is different: Falling trends of inequality. *Review of Economic Dynamics* 13, 154–178.
- Piketty, T. (2003). Income Inequality in France, 1901–1998. *Journal of Political Economy* 111(5), 1004–1042.
- Pissarides, C. A. and G. Weber (1989). An Expenditure-Based Estimate of Britain's Black Economy. *Journal of Public Economics* 39(1), 17–32.
- Pou, L. and J. Alegre (2002). La Encuesta Continua de Presupuestos Familiares: Descripción, Representatividad y Propuestas de Metodología para la Explotación de la Información de los Ingresos y el Gasto. Documento de Trabajo 172/2002, Fundación de las Cajas de Ahorro Confederadas.
- Prados de la Escosura, L. (2003). *El progreso económico de España (1850-2000)*. Bilbao: Fundación BBVA.
- Prados de la Escosura, L. (2008). Inequality, poverty and the Kuznets curve in Spain, 1850-2000. *European Review of Economic History* 12, 287–324.
- Pérez-Duarte, S., C. Sánchez-Muñoz, and V. Törmälehto (2010). Re-weighting to reduce unit non-response bias in household wealth surveys: a cross-country comparative perspective illustrated by a case study. European conference on quality in official statistics.

- Ravallion, M. (2003). The debate on globalization, poverty and inequality: Why measurement matters. *International Affairs* 79(4), 739–754.
- Sanz, B. (1995). La Articulación micro-macro en el Sector Hogares: de la Encuesta de Presupuestos Familiares a la Contabilidad Nacional. Papeles de trabajo, Instituto de Estudios Fiscales.
- Sen, A. (1992). *Inequality Reexamined*. Oxford: Clarendon Press.
- Tedds, L. (2010, June). Estimating the income reporting function for the self-employed. *Empirical Economics* 38(3), 669–687.
- Torregrosa, S. (n.d.). Did Democracy bring Redistribution? Insights from the Spanish tax system (1960-1990). *European Review of Economic History* (forthcoming).

# Appendices

## A Data aggregates comparison

Table A.1: HBS and NA household income aggregates, 1973

|                   | HBS (net)   | Direct taxes | HBS grossed | NA (gross) | Gross ratio | Net ratio | m     |
|-------------------|-------------|--------------|-------------|------------|-------------|-----------|-------|
| Self-empl. income | Monetary    | 397,269      |             | GOS        | 1,173,484   |           |       |
|                   | Imputations | 210,105      |             | FKC        | 91,708      |           |       |
|                   | Rentals     | 8,713        |             | NOS        | 1,081,776   | 60.2%     | 58.8% |
|                   | Total       | 616,086      | Total       | 34,650     | 650,736     |           | 2.06  |
| Labor income      | Monetary    | 1,184,062    |             |            |             |           |       |
|                   | Imputations | 15,467       |             |            |             |           |       |
|                   | Total       | 1,199,529    | Total       | 419,511    | 1,619,040   | 79.8%     | 74.5% |
| Capital income    | Total (mon) | 26,732       | Total       | 11,425     | 38,158      | 27.5%     | 21.0% |
| Transfer income   | Total       | 187,310      | Total       | 2,347      | 189,657     | 40.2%     | 39.7% |
| Disposable income | Total       | 2,029,658    | Total       | n/a        | 2,029,658   | 65.5%     | 65.5% |

Source: author's calculations with 1973 HBS and NA (Instituto Nacional de Estadística, 1983) and Pena and Callealta (1996).

m: scaling-up factor obtained (for net incomes). GOS: Gross Operating Surplus. FKC: Fixed Capital Consumption. NOS: Net Operating Surplus. WC: Wage Compensation. n/a: non-applicable.

Table A.2: HBS and NA household income aggregates, 1980

|                   | HBS (net)   | Direct taxes | HBS grossed | NA (gross) | Gross ratio | Net ratio | m    |
|-------------------|-------------|--------------|-------------|------------|-------------|-----------|------|
| Self-empl. income | Monetary    |              |             | GOS        |             |           |      |
|                   | Imputations |              |             | FKC        |             |           |      |
|                   | Rentals     |              |             | NOS        |             |           |      |
|                   | Total       | Total        | 2,316,443   |            | 65.1%       | 62.6%     | 2.14 |
| Labor income      | Monetary    |              |             |            |             |           |      |
|                   | Imputations |              |             | WC         |             |           |      |
|                   | Total       | Total        | 6,955,781   |            | 89.2%       | 83.8%     | 1.19 |
| Capital income    | Total (mon) | Total        | 152,488     | Total      |             |           |      |
|                   | Total       | Total        | 1,236,428   | Total      | 40.9%       | 27.4%     | 3.64 |
| Transfer income   | Total       | Total        | 7,669,397   | Total      |             |           |      |
|                   | Total       | Total        | 7,669,397   | Total      | 69.4%       | 69.4%     | 1.44 |

Source: author's calculations with 1980 HBS and NA (Instituto Nacional de Estadística, 2014) and Pena and Callealta (1996).

m: scaling-up factor obtained (for net incomes). GOS: Gross Operating Surplus. FKC: Fixed Capital Consumption. NOS: Net Operating Surplus. WC: Wage Compensation. n/a: non-applicable.

Table A.3: HBS and NA household income aggregates, 1990

|                   | HBS (net)   | Direct taxes | HBS grossed | NA (gross) | Gross ratio | Net ratio | m    |
|-------------------|-------------|--------------|-------------|------------|-------------|-----------|------|
| Self-empl. income | Monetary    |              |             | GOS        |             |           |      |
|                   | Imputations |              |             | FKC        |             |           |      |
|                   | Rentals     |              |             | NOS        |             |           |      |
|                   | Total       | 2,954,395    | 1,072,907   | 11,600,153 | 72.1%       | 69.2%     | 2.06 |
| Labor income      | Monetary    |              |             |            |             |           |      |
|                   | Imputations | 102,544      |             | WC         |             |           |      |
|                   | Total       | 12,383,012   | 9,209,310   | 23,108,029 | 93.4%       | 89.1%     | 1.12 |
| Capital income    | Total (mon) | 141,557      | 756,572     | Total      | 1,725,920   | 52.0%     | 6.85 |
|                   | Public      | 4,797,545    |             | Public     | 7,177,969   | 71.2%     | 1.43 |
|                   | Private     | 439,019      |             | Private    | 1,766,759   | 24.8%     | 4.02 |
| Transfer income   | Total       | 5,236,563    | 310,962     | Total      | 8,944,728   | 62.0%     | 1.65 |
|                   | Total       | 25,048,741   | n/a         | Total      | 32,908,556  | 76.1%     | 1.31 |

Source: author's calculations with 1990 HBS and NA (Instituto Nacional de Estadística, 2014), Sanz (1995) and Oliver (1997).

m: scaling-up factor obtained (for net incomes). GOS: Gross Operating Surplus. FKC: Fixed Capital Consumption. NOS: Net Operating Surplus. WC:

Wage Compensation. n/a: non-applicable.

## B Robustness: other equivalence scales and indices

Table B.1: The distribution of disposable income according to other equivalence scales and indices

|               | 1973     |           | 1980     |           | 1990     |           |
|---------------|----------|-----------|----------|-----------|----------|-----------|
|               | Original | Corrected | Original | Corrected | Original | Corrected |
| Gini          |          |           |          |           |          |           |
| Total (Hh)    | 36.19    | 36.88     | 34.26    | 33.58     | 33.02    | 34.84     |
| OECD          | 32.56    | 34.59     | 31.99    | 32.59     | 29.99    | 32.95     |
| OECD_mod      | 31.97    | 33.79     | 31.27    | 31.60     | 29.35    | 32.13     |
| Sqroot        | 32.12    | 33.88     | 30.96    | 31.38     | 29.37    | 32.04     |
| Top 10% share |          |           |          |           |          |           |
| Total (Hh)    | 26.75    | 28.23     | 25.35    | 25.81     | 24.70    | 27.13     |
| OECD          | 25.72    | 27.69     | 25.09    | 25.81     | 23.97    | 26.87     |
| OECD_mod      | 25.32    | 27.19     | 24.61    | 25.20     | 23.54    | 26.35     |
| Sqroot        | 25.22    | 27.05     | 24.22    | 24.88     | 23.31    | 26.09     |
| Top 1% share  |          |           |          |           |          |           |
| Total (Hh)    | 5.58     | 6.47      | 5.20     | 5.99      | 4.85     | 7.15      |
| OECD          | 5.44     | 6.64      | 5.20     | 6.07      | 4.80     | 7.30      |
| OECD_mod      | 5.40     | 6.45      | 5.06     | 5.90      | 4.65     | 7.23      |
| Sqroot        | 5.32     | 6.24      | 4.90     | 5.77      | 4.52     | 7.08      |

Source: author's calculations. In all indices, the first row represents distribution between household, while the other three depict distribution between individuals according to different equivalence scales. The OECD original and modified scales apply weights 0.7-0.5 and 0.4-0.3 respectively to subsequent adults and minors in the household. The 'Sqroot' scale uses the square root of the household size.

All calculations refer to disposable income. Gross incomes are expected to be more unequally distributed under a progressive direct tax system (which applies specially to 1990).