Intergenerational Mobility and Life Satisfaction in Spain

Amaia Palencia-Esteban

Universidade de Vigo (Spain), ECOBAS and EQUALITAS

E-mail: apalencia@uvigo.es

Pedro Salas-Rojo

Universidad Complutense de Madrid (Spain), ICAE and EQUALITAS

E-mail: pedsalas@ucm.es

Abstract

This paper explores the relation between personal wellbeing - measured with life satisfaction - and intergenerational mobility in Spain (2017). We start by applying machine learning techniques to overcome traditional data limitations and estimate intergenerational income mobility. Then, by means of several econometric specifications, we find the relation between personal wellbeing and intergenerational income mobility to be non-significant. This result is robust to several measures of educational and occupational mobility. Contrary to the comparison theory, if Spanish citizens derive wellbeing benefits or losses from intergenerational mobility, these effects are not permanent and dissipate with time. We find other variables, such as enjoying good health, higher income levels and marriage, to be positively associated with life satisfaction. Overall, personal wellbeing in Spain is more related to materialistic aspects rather than to the comparison of individuals' current position against the previous generations' socioeconomic status.

Keywords: Wellbeing, Life Satisfaction, Intergenerational Mobility, Machine Learning, Happiness, LASSO.

JEL Code: I14, I31, J62

1. Introduction

The study of wellbeing has received growing attention in the last decade as social scientists have understood that the quality of life includes many aspects beyond income or consumption capabilities (Stiglitz et al., 2009). In this context, understanding the factors associated with long-lasting welfare is crucial to design policy interventions aimed at improving and promoting general life satisfaction and happiness. However, this is not an easy task, because wellbeing is multidimensional. The literature has thereby been limited to studying the interaction of wellbeing with basic demographics such as gender, migration and employment status (European Union, 2016), or the occupational/social status, educational level and relevant aspects of childhood (Di Tella and MacCulloch, 2006; Hadjar and Samuel, 2015). In this paper we take advantage of Machine Learning techniques that allow us to exploit new data and thus explore the relation between personal wellbeing and intergenerational mobility in Spain (2017).

Previous research on the relation between intergenerational mobility and wellbeing has led to inconclusive results. Some authors have found upward mobility to be associated with higher subjective wellbeing, with downward mobility producing the opposite effect (Zhao et al., 2017). In line with the *comparison theory*, the psychological effects derived from achieving a higher or lower socioeconomic status also apply to intergenerational mobility. While upward mobility implies the fulfillment of parental expectations, personal self-realization and higher consumption capabilities, downward mobility may provoke sadness, disappointment and lack of self-confidence. However, the literature has also suggested that these mobility effects may be transitory and dissipate with time (Di Tella and MacCulloch, 2006; Guilbert and Paul, 2009). According to the *hedonic adaptation theory*, individuals rapidly adjust to their new status as they move along the socioeconomic ladder. Thus, wellbeing consequences derived from intergenerational mobility would be non-lasting; they appear after the mobility takes place and vanish afterwards. Finally, other authors (Zang and de Graaf, 2016; Iveson and Deary, 2017) find intergenerational mobility and personal wellbeing to be unrelated, with no effects attributed neither to the short nor the long term.

The scarcity of data collecting information about personal wellbeing and variables used to measure intergenerational mobility has conditioned most analysis to country case-studies.[1] In this paper we focus on Spain, because its intergenerational mobility patterns make it suitable for analyzing whether the effects over wellbeing vary across the three main intergenerational mobility

[1] The literature has usually considered Anglo-Saxon countries. For instance, Nikolaev and Burns (2014) studied the U.S., Guilbert and Paul (2009) Australia, Hadjar and Samuel (2015) Britain, and Iveson and Deary (2017) Scotland. Exceptionally, Clark et al. (2008) used data from Germany, and Zhang and de Graaf (2016) or Zhao et al. (2017) focused on China. Finally, some studies perform cross-country analyses, but they either use old data (1994-2001) from Eurostat (Molina et al., 2011) or restrict their analysis to young individuals and educational mobility (Schuck and Steiber, 2018).

approaches: income, occupational and educational. While Spain holds a median position among the European countries regarding wellbeing levels (Eurostat, 2018), its rank varies across different forms of mobility. Spain is the most educationally mobile country in the European Union, with 60% of its total population experiencing upward education mobility, but it presents a relatively strong occupational persistence and an intermediate intergenerational income correlation (European Union, 2018).

We use data from the Centro de Investigaciones Sociológicas (2017, CIS henceforth), which provides precise information on fathers' occupation and education. However, this database does not include fathers' incomes, hindering a direct approach to intergenerational income mobility. The absence of valid fathers' income information is the norm in intergenerational mobility analyses and has traditionally hampered empirical approaches to the matter. Overcoming this limitation, we take data from three waves of the "Encuesta de Presupuestos Familiares" (1980/81, 1990/91 and 2000/01) to impute the fathers' incomes into the CIS database. In doing so, we apply the Machine Learning techniques recently proposed in Bloise et al. (2021), who upgrade the traditional "Two Samples Two Stages Least Squares (TSTSLS)" imputation method (Bjorklund and Jantti, 1997). We propose a methodological improvement by employing several auxiliary imputation waves instead of just one, hence avoiding strong underlying assumptions such as the invariance of the income structure over time.

We regress individual's life satisfaction against several intergenerational mobility measures and sociodemographic controls. Our results show that the potential permanent effects derived from improving or worsening one's father's socioeconomic situation are not significant in Spain. No mobility coefficient is significant in any of these regressions, suggesting that if intergenerational mobility has an impact on welfare, the effect dissipates with time. Still, and in line with previous literature, other factors such as enjoying good health or being married are positively connected to wellbeing. Furthermore, we find that belonging to higher income quintiles is positively associated with higher life satisfaction levels, with the effects of the occupational and educational levels being always non-significant.

We highlight two main contributions. First, we contribute to the wellbeing debate by showing the absence of a permanent relation between life satisfaction and intergenerational mobility in Spain. Long-lasting wellbeing in this country seems to be more related to materialistic aspects, such as health or economic capabilities, rather than to self-comparisons about one's position against previous generations. Second, we apply Machine Learning algorithms to palliate the incomplete nature of the data and thereby estimate intergenerational income mobility. These computing techniques, combined with the TSTSLS imputation method, allow us to include intergenerational income mobility in welfare analysis, but this is just one of its potential applications. In this way, we provide a reference point for the implementation of these new imputing tools.

The reminder of the paper is structured as follows. Section 2 contextualizes our research and explains the main hypothesis under scrutiny. Section 3 presents the main database, explains how wellbeing is measured and describes the remaining variables employed in the analysis, providing some theoretical background on how they might be related. Section 4 explains the Machine Learning methods and the auxiliary data used to compute intergenerational income mobility. Section 5 studies intergenerational income, educational and occupational mobility in Spain. Section 6 presents our main results, and Section 7 concludes.

2. Literature review and hypothesis.

When assessing the connection between individuals' utility or wellbeing and economic outcomes, the simplest relation can be formulated as the former being a function of the latter. In this way, individuals' welfare would steam from goods, services or capabilities assumed to provide wellbeing or, at least, prevent displeasure (Verme, 2018). However, critics to this reasoning claim that individuals do not evaluate their lives in isolation, as they are also concerned about relative outcomes, this is, their own income, consumption or educational levels put in perspective with others (Miles and Rossi, 2007). In particular, Verme (2011; 2018) proposes that individuals compare themselves with other people in society -considered as reference groups- in the same point in time ("alter" comparisons), but also with their own past status and own expected status in the future ("ego" comparisons). As a result, we could define up to five types of reference groups: comparison with peers (parallel comparisons), richer and poorer individuals in the present (upward and downward comparisons), and past and future own status (past and future self-comparisons). In this article we relate wellbeing with intergenerational mobility that, according to this classification, corresponds to a very particular case where individuals' reference group is defined by her own parents, characterized with certain outcomes that ultimately reflect individuals' socioeconomic status in the past, during childhood.

Empirical applications have fostered the development of an array of theories to explain the relation between individuals' wellbeing (often proxied with life satisfaction and happiness) and intergenerational mobility. For instance, Hadjar and Samuel (2015) find evidence in the UK (but not in Switzerland) supporting the dissociative hypothesis, according to which upward social mobility has negative implications over wellbeing because individuals face difficulties to adapt their behavior to a new social class to which they have not always belonged. However, other authors, such as Nikolaev and Burns (2014) for the U.S. and Zhao et al. (2017) for China, find upward mobility to be positively associated with beneficial outcomes. These findings are framed in the comparison theory, where individuals permanently bear the psychological rewards/punishments of improving/worsening their parents' social or economic position. In the same context, Guilbert and Paul (2009) show that, in Australia, the psychological self-punishment derived from downward mobility is stronger than the potential rewards obtained from upward mobility, providing evidence supporting the so-called falling-from-grace hypothesis.

The permanent welfare effects described by the comparison theory are challenged by the hedonic adaptation theory, which suggest that the relation between the comparison group and individuals' wellbeing fades with time.[2] According to this idea, Di Tella and MacCulloch (2006) claim that positive and negative wellbeing effects derived from economic shocks may disappear as individuals get used to their newly achieved social status. For instance, losing a job and getting another one with lower salary may lead towards lower income and consumption levels, hence decreasing wellbeing, but this effect would eventually dissipate as the affected individual adapts her behavior. Providing more evidence, Di Tella et al. (2010) employ German data to show that, in the case of upward mobility, despite higher income levels increase marginal utility in the short-run, the positive psychological effects vanish relatively soon. Similar results are found by Stutzer (2004) employing Swiss data.

Finally, other authors find wellbeing and intergenerational mobility to be non-connected. Zang and de Graaf (2016) use Chinese data to show a significant and positive association between short-distance intra-generational mobility and wellbeing, with the effect for intergenerational mobility being not significant. One possible explanation for this result suggests that, once individuals improve their parents' socioeconomic status, the rise in subjective wellbeing is compensated by detrimental comparisons with intra-generational status (Nikolaev and Burns, 2014). Similarly, Iveson and Deary (2017) use cohort data from Scotland and find that life satisfaction in elder individuals is not significantly associated with childhood or adulthood socioeconomic status, nor with the intensity of the social mobility experience from parental occupational social class. Consequently, other attributes such as health status, consumption capabilities and future prospects are to be blamed for their different welfare levels.

The wide heterogeneity of results found in the literature suggests that this debate is far from being closed. Aimed to provide new evidence, we explore the relation between intergenerational mobility and wellbeing in a -to the best of our knowledge- unexplored country: Spain. Being our data cross-sectional (see Section 3 below), we cannot follow individual's wellbeing through time, such as in Zang and de Graaf (2016), so testing the hedonic adaptation theories is hardly possible. This kind of analysis is better suited for panel data, as the comparison between cohorts in just one

[2] See Lyubomirsky (2011) for a review of the hedonic adaptation theory on a broader context, and Galiani et al. (2015) for a pool analysis of the prevalence of this theory on Latin-American countries.

cross-section could lead to specific generational spurious effects that may distort the analysis. Hence, in this paper, finding no significant association between mobility and wellbeing would provide evidence against the comparison theory in Spain, although we could not distinguish between the hedonic adaptation theory and a complete absence of association. With this idea in mind, our first hypothesis under scrutiny can be stated as follows:

Hypothesis (a): Intergenerational mobility is associated with individual’s wellbeing.

If we found no evidence supporting Hypothesis (a) we may wonder whether the lack of a significant association between intergenerational mobility and wellbeing is driven by the opposite effects of upward and downward mobility over the dependent variable. As previously explained, according to the falling-from-grace hypothesis (Guilbert and Paul, 2009), the latter type of mobility should have a negative and bigger association with wellbeing than the former type, whose effect should be positive and smaller. Hence, we state the second hypothesis as:

Hypothesis (b): Upward and downward mobility are differently associated with individual wellbeing.

Finally, we wonder whether the effect of intergenerational mobility over mobility depends on the intensity of mobility. Consider the case where the individual climbs or falls just one step in the social or income ladder. In such case, her wellbeing may be unaffected because the consumption or saving capabilities could remain similar. On the contrary, if the individual experienced a remarkable climb or fall in the same ladder, drastically changing her position with respect to the comparison group, the effects over wellbeing might be rather strong. Following this reasoning, we state our final hypothesis as:

Hypothesis (c): The higher the intensity of intergenerational mobility, the stronger the association with individual wellbeing.

We consider three different types of intergenerational mobility and their connection with wellbeing. First, employing income mobility allows us to interpret the results as the comparison of individuals’ consumption, savings and economic security relative to that experienced by previous generations (Molina et al., 2011). Second, we analyze social class mobility, proxied by professional occupations (Bukodi and Goldthorpe, 2011; Iveson and Deary, 2017). Climbing the social ladder by working on a profession with higher social recognition than that of the parents may bring positive physiological rewards not fully collected by income mobility, such as the fulfillment of family or personal expectations. Finally, by looking at educational mobility, we focus on other adjacent effects such as those steaming from a more diverse leisure and cultural consumption (Torche, 2015; Michalos, 2017; Schuck and Steiber, 2018).

3. Data

The data comes from the module "Social Inequality and Social Mobility in Spain", a survey conducted by the Centro de Investigaciones Sociológicas (CIS) in 2017 based on the design explained in Betancort et al. (2019).[3] This survey is the latest database available in Spain where the respondents are retrospectively asked about their fathers' information during their adolescence. From the 2500 individuals originally surveyed following a stratified multi-stage sampling procedure, the CIS reports 2482 valid observations representative for the Spanish population by age, region and gender. To exclusively include individuals participating in the labor market, while also following the intergenerational mobility literature, the final sample is restricted to individuals aged between 30 and 60 years. Once we apply these restrictions, we are left with a final sample of 1151 valid observations.

The remainder of the section presents the variables employed in our analysis. First, we focus on the dependent variable, explaining how life satisfaction is used to measure wellbeing. Then, we describe the variables used to estimate intergenerational mobility - income, educational level and occupation of the individuals and their fathers -, and relate them to the main theories proposed to explain their relation to life satisfaction. Finally, we introduce the social and demographic controls used to account for the remaining factors that, according to the literature, affect wellbeing.

3.1 Dependent variable.

The literature usually proxies wellbeing by self-reported life satisfaction (Molina et al., 2011; Iveson and Deary, 2017; Shuck and Steiber, 2018). Despite we acknowledge certain subjectivity in this variable, ample evidence demonstrates that it provides meaningful, reliable and valid information about individuals' wellbeing (Verme, 2018). Life satisfaction is usually related to long-term factors used by the individuals to make judgements about the quality of their lives. Among those factors, the literature highlights positive psychological aspects such as the self-fulfillment of personal ambitions or expectances, enjoying good health and shape, or the benefits derived from stable social interactions, including a successful marriage. In this paper, we measure life satisfaction by collecting the answers to the following question: *On a discrete scale from 0 (completely unsatisfied) to 10 (completely satisfied), do you consider yourself to be satisfied with your life?*. Figure 1 shows how life satisfaction is distributed: the mean reported life satisfaction reaches a value of 7.54 over 10, the standard deviation being 1.72 points.[4] This estimate is close

---

[3] The CIS is a dependent entity of the Spanish Ministry of Presidency whose main task consists on improving the scientific knowledge of the Spanish society. The database, the questionnaire and details of the sample design are available at:
http://www.cis.es/cis/opencm/ES/2_bancodatos/estudios/ver.jsp?estudio=14350

[4] In addition, we test the robustness of our results by repeating the whole analysis with the self-reported happiness, which also ranges between 0 (completely unhappy) and 10 (completely happy). However, the

to that obtained with Eurostat data for 2018 (variable *ilc_pw01*), where a mean value of 7.3 is found for Spain using a similar definition of life satisfaction. This number coincides with the mean value in the EU. The same survey data delivers a mean life satisfaction score of 7.1 points in Italy, 7.3 in France, 7.4 in Germany, 6.7 Portugal and 7.6 the UK, with the highest score being reached in Ireland (8.1 points) and the smallest one in Bulgaria (5.4 points).

[INSERT HERE FIGURE 1]

3.2 Intergenerational mobility variables.

Measuring intergenerational income mobility requires, by definition, two variables: one collecting fathers' income and another reporting income of the children. Unfortunately, the CIS (2017) database does not include information on fathers' incomes. Overcoming this limitation, we follow the mainstream intergenerational mobility literature (Olivetti and Paserman, 2015; Jerrim et al., 2016) and implement the "Two-Sample Two-Stage Least Squares" (TSTSLS) methodology to impute fathers' incomes from previous surveys (Bjorklund and Jantti, 1997).[5] This technique, despite not being difficult to apply, requires a separated explanation, which is assessed in section 4. Consequently, for now, we present the main statistics of income reported in the CIS database, this is, children's income.

The literature on intergenerational income mobility has traditionally focused on personal income to analyze the transmission of opportunities from fathers to sons and daughters (Jantti and Jenkins, 2013). We claim that individuals' household income should be considered when relating intergenerational income mobility to life satisfaction. Using household income does not only control for assortative mating, but also collects more complete information about consumption and saving capacities within the family unit. Relying on personal income would not account for these aspects of household economies, and could lead to biased results in the context of personal wellbeing. Thereby, in this article we use household income, which includes all sources of income perceived by the household where the respondent lives, net of taxes and benefits.[6] Still, for comparative reasons, we divide the household income by the squared root of the household size,

long-lasting effects of happiness over personal wellbeing remain unclear, as it is usually related to external shocks such as unemployment, divorces or deaths of relatives (Oswald and Powdthavee, 2008), so we choose life satisfaction as our preferred dependent variable. Results remain largely unchanged and are available upon request.

[5] Particularly, we employ three waves of the Encuesta de Presupuestos Familiares (EPF): 1980/81, 1990/91 and 2000/01.

[6] All economic units used in the paper are adjusted to 2017 euros.

as this is the scale of equivalence method commonly used for inequality studies in Spain (see Cabrera et al., 2021).[7]

Table 1 shows the summary statistics of per capita adjusted household income by age groups. The relative similitude of the mean values and the standard deviations goes in line with the permanent income hypothesis. Regarding income inequality, the Gini coefficient associated with the household income distribution gets to 0.301 points, close to the 0.315 Gini that Ayala (2016) estimated for Spain in 2014.

[INSERT HERE TABLE 1]

Regarding intergenerational occupation mobility, we focus on the occupational change between the respondents and their fathers.[8] The CIS employs the ISCO-08 classification, providing occupations disaggregated up to the 3-digit level. However, occupational mobility matrixes require an aggregated dimension; the occupational groups should reflect certain professional status so that moving from one group to another may imply substantial changes in a person's wellbeing. With this aim, we use the ISCO-08 skills classification to create four categories that represent diverse skill levels and constitute different class status, working conditions and wage levels. First, we have unqualified workers (ISCO-08=9); second, semi-qualified and qualified laborers (ISCO-08=4-8); third, technicians and support professionals (ISCO-08=3); fourth, managers and professionals (ISCO-08=1-2).

Table 2 presents the summary statistics of the occupational distribution of the respondents and their fathers. The structural changes of the Spanish labor market are evident: while 34.2% of respondents have high skill occupations (ISCO-08=1-3), this ratio only reaches 22.9% for the fathers. Still, the main difference lies in the semi-qualified and qualified workers, as their proportion is much smaller for the respondents than for their fathers.

[INSERT HERE TABLE 2]

Finally, intergenerational education mobility summarizes how the education of the respondents relates to their fathers'. In the CIS database, the educational categories are defined following the ISCED classification (UNESCO, 2012) but, again, we follow Cabrera et al. (2021) and recode the levels of studies to create four groups. We distinguish those with zero or primary education

[7] The squared root of the household size is an equivalence of scale more common in the U.S. than in Spain, where OECD modified scale is used for instance in EUROSTAT statistics. However, we do not have information in our database to apply this scale, because the ages of household members –besides the head- are not included.

[8] Considering mothers' occupation would substantially reduce the sample size due to the late incorporation of women into the Spanish labor market. Indeed, a big share of the respondents' mothers carried household informal works.

(ISCED=0-1), those with lower-secondary education (ISCED=2), those with upper secondary education and post-secondary (ISCED=3-4) and those with tertiary education (ISCED=5-8).

[INSERT HERE TABLE 3]

Table 3 presents the summary statistics of the educational distribution of the respondents and their fathers, reflecting the expansion of the compulsory secondary education in Spain. While the share of individuals with just primary education is reduced by more than a half, those with post-secondary education double their proportion. Indeed, upper secondary education presents the biggest improvement: 29.9% of the respondents hold tertiary degrees when only 9.8% of the fathers does.

3.3 Control variables.

We include several control variables that account for social and demographic aspects that the literature has found to be related to personal wellbeing (Nicholaev and Burns, 2014; Hadjar and Samuel, 2015). In particular, we consider respondents' age and age squared to control for the life cycle and its non linearities, gender (binary), health status, marital status (being married or not) and the presence of children (binary). Table 4 shows the summary statistics. The sample is evenly distributed across men and women, whose mean age lies around 45 years. Regarding the self-assessed health status, on average, individuals situate themselves at 3.9 on a scale from 1 (very bad) to 5 (very good). Finally, while only 60% of our sample is married, individuals have, on average, 1.27 children.

[INSERT HERE TABLE 4]

Finally, and taking into account the ideas exposed in Section 2, we also control by the economic outcome representing the actual status of the individual. In other words, in the regressions in Section 6, when we relate income mobility with wellbeing, we also control for the income quintile of the individual. Similarly, in the models relating occupation and education mobility with wellbeing, we control for the occupational and educational status, respectively. In this manner, the coefficients capturing the impact that intergenerational mobility has on wellbeing are not capturing specific effects of the occupational, educational or income groups to which the individuals belong.

4. Methods

Computing intergenerational income mobility requires income information from two cohorts or generations. However, fathers' income is not available in the CIS (2017) nor in any other modern Spanish database that also includes wellbeing variables. This lack of data is prevalent in many developed economies and hinders the study of intergenerational income mobility and its associated factors. Overcoming this limitation, Bjorklund and Jäntti (1997) proposed the TSTSLS methodology based on a two-sample instrumental variable estimator (Angrist and Krueger, 1992). This technique has been repeatedly used in the intergenerational income mobility literature, such as in Cervini-Pla (2015) for Spain, Barbieri et al. (2019) for Italy or Bloise et al. (2021) for the US and South Africa.

The TSTSLS estimation method requires two different samples. The main sample must include data on individuals' current income and fathers' socioeconomic variables like their educational level or occupation. However, as the main sample (the CIS in our case) lacks information on fathers' income, a secondary or auxiliary sample must be employed. Precisely, this sample comes from an earlier survey that contains the same information - income and socioeconomic variables- for previous cohorts. The main idea of this procedure consists in considering individuals in the auxiliary sample as pseudo-parents, and estimating their income conditioned on the selected set of common socioeconomic factors. The resulting fitted income values are then imputed into the main sample by matching the fathers' and pseudo-fathers' socioeconomic information that is present in both surveys. Formally, consider Equation (1):

$$y_i^s = \alpha + \beta y_i^f + \varepsilon_i \tag{1}$$

Where y_i^s is the logarithm of the sons' permanent individual income, y_i^f is the logarithm of fathers' permanent earnings, α is the mean income of sons' and ε_i is an error term that collects individual's income not explained by the fathers'. As the CIS dataset does not include y_i^f, we use the auxiliary sample to estimate the following Equation:

$$y_i^{pf} = \varphi + \gamma z_i^{pf} + \delta_i \tag{2}$$

Where y_i^{pf} is individual income of the pseudo-fathers in the auxiliary sample and z_i^{pf} is a vector of time-invariant socioeconomic factors used to predict income. Finally, δ_i is the component of pseudo-fathers' income not explained by the control socioeconomic factors. Equation (2) is estimated by OLS and then used to predict fathers' income: $\hat{y}_i^{pf} = \hat{\gamma} z_i^{pf}$.

This method poses an extra problem. The vector of estimated coefficients ($\hat{\gamma}$) is estimated with imperfect and incomplete data, as fathers' occupation and educational level are the only variables we have to match both samples and use as regressors in Equation (2). Since the exclusion of relevant socioeconomic controls makes the imputation highly dependent on data quality, the resulting fitted values are probably biased.[9] To improve the accuracy of our imputations, we follow Bloise et al. (2021) and apply Machine Learning methods to increase the precision of our intergenerational income mobility estimates. Formally, the best possible imputation is obtained when we reduce at a minimum the squared error (MSE) between $\hat{y}_i^{pf}$ and y_i^f:

$$min\left\{E\left[\left(y_i^f - \hat{y}_i^{pf}\right)^2\right]\right\} = min\left\{E\left[\left(y_i^f - f\left(z_i^{pf}\right)\right)^2\right]\right\} \tag{3}$$

The expected squared error in Equation (3) can be decomposed into three different elements:

$$E\left[\left(y_i^f - \hat{y}_i^{pf}\right)^2\right] = var\left(\hat{f}\left(z_i^{pf}\right)\right) + (bias)^2 + var(\delta_i) \tag{4}$$

The first term on the left, $var\left(\hat{f}\left(z_i^{pf}\right)\right)$, is the error coming from the sensibility of Equation (2) to the random noise in the auxiliary sample. The second term is the bias of the model, which quantifies the error generated by the selection of the variables in the data generation process. The last term is an irreducible error that captures the smallest possible error we must cope with when predicting y_i^{pf}.

By definition, a trade-off exists in Equation (4). Very complex models, such as those including all occupational and educational categories as dummies in vector z_i^{pf}, diminish the bias term but increase the variance, leading to a potential over fitting. On the contrary, too simple models that use highly aggregated variables as controls diminish the variance component at the expense of increasing the bias term. Solving this tension, following Bloise et al. (2021), we estimate Equation (2) with the regularization term first introduced by Zou and Hastie (2005). This Machine Learning method consists in adding up an extra term to the classical least-square regression, so that the estimated coefficients are obtained by minimizing Equation (5):

$$\sum_{i=1}^{n}\left(y_i^{pf} - \sum_{j=1}^{k}\rho_j z_{j,i}^{pf}\right)^2 + \lambda\left(\alpha\sum_{j=1}^{k}\left|\rho_j\right| + (1-\alpha)\sum_{j=1}^{k}\rho_j^2\right) \tag{5}$$

The left-hand side term is a canonical OLS element, with ρ_j being the associated parameter to each -j regressors included in the vector $z_{k,i}^{pf}$. The right-hand side element is a regularization term that penalizes over fitting by shrinking some of the estimated coefficients towards zero. The main idea of the algorithm lies on including as much information as possible and, simultaneously, eliminating the coefficients that do not provide meaningful information to minimize Equation (3).

[9] For a complete formal explanation, see Nybom and Stuhler (2016).

Summing up, if Equation (5) was estimated including a high number of covariates in vector z, in our case all educational and occupational categories, the regularization term would shrink many ρ_j coefficients towards zero, optimizing the predictive capacity and avoiding overfitting. Thus, given the available data, estimating the mincerian Equation (3) in this manner provides the most accurate possible prediction of fathers' income.

4.1 Auxiliary database

The data for the auxiliary sample comes from the Household Budget Survey (Encuesta de Presupuestos Familiares, EPF) conducted by the Spanish National Institute of Statistics (Instituto Nacional de Estadística, INE). First implemented in 1973, this survey is representative of the Spanish population and collects information on incomes, expenses and a wide range of socioeconomic characteristics of the Spanish Households. The INE carried two other waves in 1980 and 1990 before changing its design to a panel structure in1997.

We propose a variation from Bloise et al. (2021). These authors consider a single wave as the auxiliary sample, in particular, the 1982 wave of the Panel Survey of Income Dynamics. Employing just one wave implies the strong assumption that pseudo-parents are all alike, regardless of the age of the individual, implicitly neglecting economic cycles or structural changes. Amending this, we employ three different waves (1980/81, 1990/91, 2000) to define our pseudo-parents.[10] Recall that we have restricted the main sample by keeping individuals aged between 30 and 60 years. The CIS data was collected in 2017, and the respondents are retrospectively asked about the fathers' information when they were 16. Thus, if we followed Bloise et al. (2021) and only used the 1980-81 wave to impute fathers' income, the imputation for younger individuals of the CIS might be problematic, as they were not even born in that year. In this manner, based on the age of the individuals, we assign them the most accurate pseudo-father and, thus, account for the relevant structural changes that the Spanish economy experienced during the 80s and 90s.

Table 5 presents the correspondence between respondents' age in the main sample (CIS) and the auxiliary waves employed to impute their respective fathers' income. Younger cohorts (those aged between 30 and 35) receive their fathers' income imputation from the EPF 2000. This is indeed the best possible match. Considering that the CIS data was collected in 2017, those aged

[10] The INE has traditionally carried out two types of EPFs: the structural or basic ones every eight or ten years (our 1980-81 and 1990-91 surveys) and, since 1997, the quarterly ones (our 2000 survey). For this last surveys and for each year, the INE also provides an annual longitudinal database collecting the corresponding four quarter, but relevant variables like education and occupation are too aggregated. Thus, we use the four quarterly databases, which offers more disaggregated classifications.

32 (the median point between the age range 30-35) were 16 years old when the 2000/01 EPF wave was collected. Similarly, middle-aged (36-45 years) and older cohorts (45-60 years) receive, respectively, their fathers' income imputations from the EPF 1990-91 and 1980-81.

[INSERT HERE TABLE 5]

4.2 Imputation.

Once we establish the correspondence between the main sample and the three auxiliary samples, we apply the TSTSLS methodology. Following Equation (5), we use the EPF to regress the pseudo-fathers' socioeconomic factors captured in vector $z_{j,i}^{pf}$ (the educational level and occupation) against their reported incomes. The resulting fitted income values, $\hat{y}_i^{pf}$, are imputed into the CIS by matching fathers' and pseudo-fathers socioeconomic information, which is present in both surveys.[11] However, Equation (5) is not only defined with the usual regression parameters (ρ_j), but also includes a regularization term with other two undefined extra parameters: λ and α.

The values of λ and α should not be arbitrarily selected. Indeed, if their setting was left to the researchers' criteria, they could easily affect the quality of the imputation by implicitly leading to the exclusion of more or less regressors, artificially shrinking their ρ_j coefficients towards zero. Avoiding exogenous alterations on the imputation, the proposed algorithms compute all possible tunings and combinations of λ and α to finally select the one that delivers the smallest MSE in Equation 3. To make their tuning completely transparent, the interested reader may find further information in the Technical Appendix.

Table 6 presents the summary statistics of the income imputations performed for each EPF wave. While the mean imputed income is similar in the three waves, the dispersion is reduced over time, in line with the declining inequality described in Ayala (2016). Once we estimate the vector of income for each wave, and taking Table 5 as a reference, we impute those values to each cohort in the CIS database by matching the occupation and educational level of the fathers and pseudo-fathers. This last step completes the imputation, as fathers have received their correspondent imputed income in the main sample.

[11] Occupational classifications have suffered several updates through the decades. Thus, we convert CNO-79 and CNO-94 (EPF's classifications) into CNO-2011 (CIS's classification) using the correspondence tables published by the INE. Occupation is reported at a two-digits level. We also recode the educational categories into six levels: illiterates, primary, secondary (first stage), secondary (second stage), professional formation and tertiary education. Detailed information is available upon request.

[INSERT HERE TABLE 6]

5. Intergenerational Mobility in Spain

Once we impute fathers' income, we start studying intergenerational income mobility. First, we build a transition matrix by tabulating the quintiles of fathers' imputed income (rows) against the quintiles of household adjusted income (columns). This way, Table 7 reports intergenerational persistence (observations that remain in the main diagonal), upward (observations situated above the main diagonal) and downward mobility (those below the main diagonal).

[INSERT HERE TABLE 7]

Around 36% of our sample experienced upward mobility, while a similar proportion suffered downward mobility. However, when we measure relative mobility, we find a strong persistence at the tails of both distributions. Up to 33.63% (76/226) of the richest fathers have children that stay in the same quintile, whereas only 13.27% (31/226) have descendants in the lowest quintile. On the contrary, while one third of low-income fathers (71/234) have low-income children, only 13.25% (31/234) of them have children in the fifth quintile.

We also look at the β coefficient in Equation 1, which captures the elasticity between father's and offspring's earnings. This parameter is indeed a measure commonly applied in intergenerational mobility studies because it collects the degree to which income capabilities are transmitted through generations. In Spain, this coefficient was equal to 0.4 in 2011, and presented small variations across different subsamples (Cervini-Pla, 2015). Actualizing this estimate for 2017, we get a value of 0.46. The detailed results of this regression are shown in Table A1 in the Appendix.

Regarding occupational mobility. Table 8 tabulates the professional status of the fathers (rows) against that of the children (columns). In line with the literature, we find a strong intergenerational occupational persistence (European Union, 2018). Around 50.4% of our sample has a job with a qualification requirement similar to that of their fathers, while 27.1% experience upward mobility and 22.5% downward mobility. Moreover, when we analyze relative mobility, we find that fathers' position largely conditions the occupation of his descendent. From those fathers who participated in the highest occupational level, 37.3% (59/158) have descendants in that same category, but only 5.1% (8/158) have children in the less qualified jobs. By contrast, while just 13.6% (9/66) of low qualified fathers have descendants who reach the highest occupational level, around one third (20/66) of their children remain in the lowest category.

[INSERT HERE TABLE 8]

Finally, we study intergenerational educational mobility, since the literature has shown that having access to a wider variety of forms of cultural consumption and leisure are also related to higher personal welfare (Torche, 2015; Schuck and Steiber, 2018). Table 9 presents the transition matrix of education, showing that around half of the sample (50.8%) experienced upward mobility, while only 7.6% have less education than their fathers. Indeed, absolute mobility ratios are encouraging, but, once again, the relative mobility analysis highlights the strong persistence of the educational levels between generations and the unequal opportunities that hide behind these results. While 67.9% (95/140) of highly educated fathers have children with the same educational level, this ratio descends to a mere 18.1% (139/770) when we consider fathers with primary or lower studies and highly educated children. Clearly, upward educational mobility has not been homogeneously distributed among the Spanish population.

[INSERT HERE TABLE 9]

6. Results

In this section we delve into the relation between personal wellbeing and intergenerational mobility. We run several OLS regressions including various approaches to intergenerational income, occupational and educational mobility while controlling for other factors that might also affect personal wellbeing.[12] It is necessary to remark that although we cannot claim causal effects, we can still interpret and explain the significant (and non-significant) association between these variables according to the theories we have exposed in Section 2. Our baseline econometric model can be formalized as:

$$Life\ Satis_i = \alpha + \beta_t Mobility_i + \beta_z X_i + \beta_1 Gender_i + \beta_2 Age_i + \beta_3 Age\ Squared_i + \beta_4 Health_i + \beta_5 Married_i + \beta_6 Children_i + e_i \qquad (6)$$

Where the dependent variable "Life Satis" stands for life satisfaction. "Mobility" is our main variable of interest that captures the degree of intergeneration mobility experienced by each individual. We also include a vector X_i that, depending on the type of mobility under study, collects respondents' actual income quintile, occupational group or educational level. The remaining controls include individuals' gender, age, age squared, health, marital status and the presence of children in the household.

The hypotheses stated in Section 2 are assessed with different definitions of the "Mobility" variable in Equation 6. In particular, we specify five models, each one capturing a different

[12] Being the dependent variable defined from 0 to 10, we prefer not to use ordered probit regressions because they would make the results rather cumbersome. Thus, our analysis is based on traditional OLS regressions, although we have checked that all our conclusions are robust to using ordered probits.

potential effect of intergenerational mobility on personal wellbeing. Model 1 defines mobility as a discrete variable, where 1 represents upward mobility, 0 immobility and -1 downward mobility. If the associated coefficient was significant in this model, we would find evidence supporting Hypothesis (a) and the comparison theory. Model 2 includes two dummies, one for upward and the other for downward mobility, with the immobility status being the omitted category. This model is used to disentangle whether the association (or lack of association) found in Model 1 is provoked by those who improve or worsen their situation, thus checking Hypothesis (b), the fallen-from-grace and dissociative hypotheses.[13]

The three remaining models allow us to delve deeper into the intensity of the relation between upward and downward mobility and wellbeing, and check Hypothesis (c). Hence, the "Mobility" term in Model 3 captures the magnitude or intensity of intergenerational mobility. This measure is constructed by counting the number of ladders ascended or descended between generations. For instance, as education is classified in four levels, respondents who achieve the highest educational level and have a low educated father receive a value of 3 (they ascend three steps in the educational ladder), while those whose fathers attended upper secondary education but are low educated receive a -2 (they descend two steps). Note that Model 3 assumes that the effect of intensity is linear in mobility, this is, homogeneous and independent from the number of steps climbed or descended. Thus, Model 4 broadens the analysis and accounts for potential non-linearities by squaring the mobility variable defined for Model 3. Finally, in Model 5 we include all mobility steps as dummies, leaving immobility as the omitted category. With this model we check whether there is a certain mobility intensity level associated with higher or lower life satisfaction.

Table 10 relates life satisfaction and intergenerational income mobility. Model 1 shows that there is no significant association between both variables and, similar to Zang and de Graaf (2016) and Iveson and Deary (2017), provides evidence against the comparison theory and Hypothesis (a). If mobility had an initial impact on life satisfaction, as proposed by the hedonic adaptation theory, our results show that the effect vanishes with time. Model 2 suggests that this absence of relation is neither prevalent nor in upward nor in downward mobility, so we can also reject Hypothesis (b). Although the coefficient in downward mobility is negative and higher than that for upward mobility, the non-significance leads us to rule out the fallen-from-grace hypothesis, as in (Guilbert and Paul, 2009).

[INSERT HERE TABLE 10]

[13] We have also checked that our results are robust to the interaction between "Mobility" and vector "X_i" in Equation 6, this is, between intergenerational mobility and the income quintile or the occupation and education position the individual belongs to. All results and conclusions remain, and are available upon request.

When looking at Models 3 and 4, where mobility intensity is assessed, we again find no significance in the coefficients, meaning that the non-relation between life satisfaction and intergenerational income mobility is not caused by the heterogeneity of the latter variable. In other words, improving or worsening one's income status with respect to one's fathers' has no permanent effects on personal wellbeing regardless the intensity of that movement. This idea is finally confirmed in Model 5, where all steps climbed or descended in the income ladder are included as dummies. Only upward mobility (+4 steps) is weakly significant. The negative sign in this coefficient might tempt us to follow Hadjar and Samuel (2015) and claim to have evidence supporting the dissociative theory, this is, that individuals who experience a big climb in the social ladder find it hard to get used to their newly acquired position. However, given the small significance (10%) and the remarkable standard error, we prefer to be conservative and consider the result too weak to make such strong affirmations. In any case, these three models provide evidence against Hypothesis (c).

Individuals in the third- and fourth-income quintiles, followed by those at the fifth- and second-quintiles, experience higher levels of life satisfaction than those located in the first quintile, which is the omitted category. This result is unsurprising, because, higher income may lead to higher consumption, savings and economic security, which can in turn result in higher utility or wellbeing. Although one might have expected this effect to be monotonically increasing with the income quintile (probably following a concave shape), we always find the biggest coefficients in the third- and fourth-income quintiles, never in the fifth. Data limitations prevents us from exploring this issue in further detail, but future research may investigate whether wellbeing in this group (among the richest) is more affected by another reference group, maybe reflecting a detrimental effect of intra-generational comparisons as proposed in Nikolaev and Burns (2014).

The coefficients, signs and significance of the remaining control variables are robust among the five different Models. Gender, age and having children are never significantly associated with life satisfaction, with their coefficients being rather small and often close to zero. On the contrary, and in line with the literature, enjoying good health is one of the most contributing variables to personal wellbeing (Deaton and Arora, 2009; Iveson and Deary, 2017). Unsurprisingly, healthy people are less likely to report sadness, physical pain, stress and anger, which are emotions related to lower levels of life satisfaction. Same as Hamermesh (2020), the effect of being married (or coupled) over life satisfaction is significant, positive and with a coefficient often higher than that of health status. This positive effect of marriage is usually associated with a higher emotional stability, positive reliance on others and also favorable thoughts regarding future prospects.

Tables 11 and 12 repeat the analysis, respectively focusing on the effect of educational and occupational mobility over life satisfaction. All results reinforce the ideas we have previously explained. Same as for Table 10, there is no significant connection between life satisfaction and any type nor approach to intergenerational mobility. It seems that, in Spain, the psychological effects captured by these two approaches, such as the fulfillment of family expectations in the case of occupational mobility and the more diverse leisure or cultural consumption in the case of educational mobility, are not permanently associated to higher life satisfaction. Once more, our results show evidence against Hypothesis (a), (b) and (c), although we cannot disentangle whether positive or negative effects steaming from intergenerational mobility appear and vanish in the short-run.

Contrary to Table 10, where the income quintile coefficients were significant and positive, the occupation and education categories are never significant, neither in Table 11 (for occupational mobility) nor in 12 (for educational mobility). In Table 11, the coefficient associated to 'Directives and Professionals', the highest occupational category, is always positive and much bigger than the other two, but it is still insignificant. Similar results are found in Table 12, where Tertiary education is positively -but never significantly- associated with life satisfaction. The remaining control factors maintain their overall sign, size and significance, so the previous interpretation applies.

[INSERT HERE TABLE 11]

[INSERT HERE TABLE 12]

7. Conclusions

In this paper, we explore the relation between wellbeing and intergenerational income, occupational and educational mobility in Spain (2017). First, we apply Machine Learning techniques to overcome data limitations and update estimates of intergenerational income mobility for Spain. Afterwards, we run several regressions that, after controlling for some classical sociodemographic factors, allow us to show that no measure of intergenerational mobility is associated with higher nor lower levels of personal wellbeing. These findings provide evidence against the comparison theory in Spain, according to which mobility provokes permanent effects over individuals' wellbeing. However, data limitations prevent us from discerning whether these effects never appear, or whether they emerge in the short run but vanish as individuals adapt to the newly acquired socioeconomic status, as stated by the hedonic adaptation theory. Similar to the related literature, other factors such as enjoying good health, marriage and higher income levels are positively associated with higher wellbeing levels. Overall, our results show that once individuals cover some basic necessities, including economic, physical and emotional stability, the comparison with the own parents' socioeconomic status (which ultimately reflects the individuals' own status during childhood) has non-lasting effects on their life satisfaction.

All in all, these findings suggest that comparisons regarding individuals' status -in terms of income, occupation and education- are not as simple as those stated by comparison theories. As explained in Verme (2018), self-evaluations with respect to the reference group can lead to complex systems of 'ego' and 'alter' comparisons, which may be dynamic and evolve through time. For instance, it is not hard to imagine the experience of joy and proud when becoming the first member in a family to obtain a university degree, or when getting a certain occupation associated with higher living standards and status. However, this effect would probably vanish as individuals start comparing themselves with their peers from this newly acquired social status, and start developing other consumption, leisure or culture necessities. In such case, the relevant reference group would evolve from inter-generational to intra-generational. Understanding these dynamics is beyond the scope of this paper and current data possibilities, but should be assessed in future research.

We conclude by remarking that social mobility, even in situations of intense upward and downward intergenerational movements, does not seem to necessarily produce more satisfied nor unsatisfied citizens. Although this result may seem discouraging, because upward mobility is not a permanent life-satisfaction generator per-se, it could also have a positive interpretation, as downward mobility does not generate long-term sadness or personal resentment. In this regard, we leave normative judgements to the reader. Still, our results highlight the remarkable association that income, which largely reflects consumption capabilities and economic security, and heath have with personal wellbeing. Overall, governments concerned about the general welfare of their citizens should focus on improving their health systems and ensuring working conditions capable of increasing their economic stability or capacity for consumption. Aspects related to self-comparison with others is probably within individuals' own responsibility.

References

Angrist, J. D. and Krueger, A. B. (1992) "The Effect of Age at School Entry on Educational Attainment: An Application of Instrumental Variables with Moments from Two Samples", *Journal of the American Statistical Association*, 87, 328-336.

Ayala, L. (2016). *La Desigualdad En España: Fuentes, Tendencias y Comparaciones Internacionales*. Madrid: FEDEA.

Barbieri, T., Bloise, F. and Raitano, M. (2019) "Intergenerational Earnings Inequality: New Evidence from Italy", *Review of Income and Wealth*, https://doi.org/10.1111/roiw.12420

Betancort, M., Darias, S., Cabrera, L., Nieves, C., Marrero, G.A., Rodríguez, J.G., and Sanchez, D. (2019): "Inequality of Opportunity in an Outermost Region: The Case of Canary Islands". *Island Studies Journal*, 14: 23-42.

Bjorklund, A. and Jantti, M. (1997) "Intertenerational Income Mobility in Sweden Compared to the United States", *American Economic Review*, 87, 1009-1018.

Bloise, F., Brunori, P. and Piraino, P. (2021) "Estimating Intergenerational Income Mobility on Sub-Optimal Data: A Machine Learning Approach", The Journal of Economic Inequality, 1, 1-23.

Bukodi, E. and Goldthorpe, J. H. (2011) "Class Origins, Education and Occupational Attainment in Britain: Secular Trends or Cohort-Specific Effects?", *European Societies*, 3, 347-375.

Cabrera, L., Marrero, G. A., Rodríguez, J. G. and Salas-Rojo, P. (2021) "Inequality of Opportunity in Spain: New Insights from New Data", *Review of Public Economics*, 237, 153-185.

Cervini-Pla, M. (2015) "Intergenerational Earnings and Income Mobility in Spain", *Review of Income and Wealth*, 61, 812-828.

Clark, A. E., Dienere, E., Georgellis, Y. and Lucas, R. E. (2008) "Lags and Leads in Life Satisfaction: A Test of the Baseline Hypothesis", *The Economic Journal*, 118, 222-243.

Deaton, A. and Arora, R. (2009) "Life at the Top: the Benefits of Height", *Economics and Human Biology*, 7, 133-126.

Di Tella, R. and MacCulloch, R. (2006) "Some Uses of Happiness Data in Economics", *Journal of Economic Perspectives*, 20, 25-46.

Di Tella, R., Haiksen-De New, J. and MacCulloch, R. (2010) "Happiness Adaptation to Income and to Status in an Individual Panel", *Journal of Economic Behavior and Organization*, 76, 834-852.

European Union (2016) "Analytical Report on Subjective Well-Being", Statistical Working Papers Eurostat, no KS-TC-16-005.

European Union (2018) "Inter-generational Mobility in Social Status, Education and Occupation", *Special Eurobarometer*, 471, Summary (Chaper IV).

Eurostat (2018) Quality of life indicators: Overall experience of life. Publications Office of the European Union, Luxembourg.

Galiani, S., Gertler, P. and Undurraga, R. (2015) "The Half-Life of Happiness: Hedonic Adaptation in the Subjective Well-Being of Poor Slum Dwellers to a Large Improvement in Housing. NBER Working Paper series, no 21098.

Guilbert, D. and Paul, S. (2009) "Income and Happiness: An Analysis of Adaptation and Comparison Income Effects", *HILDA Survey Research Conference*, 16-17.

Iveson, M. H. and Deary, I. J. (2017) "Intergenerational Social Mobility and Subjective Wellbeing in Later Life", *Social Science and Medicine*, 188, 11-20.

Hadjar, A. and Samuel, R. (2015) "Does Upward Social Mobility Increase Life Satisfaction? A Longitudinal Analysis Using British and Swiss Panel Data", *Research in Social Stratification and Mobility*, 39, 48-58.

Hamermesh, D. (2020) "Life Satisfaction, Loneliness and Togetherness, with and Application to Covid 19 Lock-downs", NBER Working Paper, no 27018 (revised version forthcoming in Review of Economics of the Household).

Jantti, M. and Jenkins, S. P. (2013) "Income Mobility", Handbook of Income Distribution, 2, 807-935.

Jerrim, J., Choi, A. and Simancas, R. (2016) "Two-Sample Two-Stage Least Squares Estimates of Earnings Mobility: How Consistent Are They?", *Survey Research Methods*, 10, 85-102.

Lyubomirsky, S. (2011) "Hedonic Adaptation to Positive and Negative Experiences". In S. Folkman, *The Oxford handbook of stress, health, and coping*. Oxford University Press, Oxford.

Michalos, A. C. (2017) "Education, Happiness and Wellbeing", *Connecting the Quality of Life Theory to Health, Well-being and Education*, Springer, New York.

Miles, D. and Rossi, M. (2007) "Learning About One's Relative Position and Subjective Wellbeing", *Applied Economics*, 39, 1711-1718.

Molina, J. A., Navarro, M. and Walker, I. (2011) "Intergenerational Well Being Mobility in Europe", *Kyklos*, 64, 253-270.

Nikolaev, B. and Burns, A. (2014) "Intergenerational Mobility and Subjective Well-Being, Evidence from the General Social Survey", *Journal of Behavioral and Experimental Economics*, 53, 82-96.

Nybom, M. and Stuhler, J. (2016) "Heterogeneous Income Profiles and Lifecycle Bias in Intergenerational Mobility Estimation", *Journal of Human Resources*, 51, 239-268.

Olivetti, C. and Paserman, D. (2015) "In the Name of the Son (and the Daughter): Intergenerational Mobility in the United States, 1850-1940". *American Economic Review*, 105, 2695 - 2724.

Oswald, A. J. and Powdthavee, N. (2008) "Death, Happiness and the Calculation of Compensatory Damages", *The Journal of Legal Studies*, 37, 217-251.

Schuck, B. and Steiber, N. (2018) "Does Intergenerational Educational Mobility Shape the Well Being of Young Europeans? Evidence from the European Social Survey", *Social Indicators Research*, 139, 1237-1255.

Stiglitz, J., Sen, A. and Fitoussi, J. (2009) "The Measurement of Economic Performance and Social Progress Revisited: Reflections and Overview". Available at Science Po Publications, no 200933.

Stutzer, A. (2004) "The Role of Income Aspirations in Individual Happiness", *Journal of Economic Behavior and Organization*, 54, 89-109.

Torche, F. (2015) “Analyses of Intergenerational Mobility: An Interdisciplinary Review”, *The ANNALS of The American Academy of Political and Social Science*, 657, 37-62.

Varian, H. R. (2014) “Big Data: New Tricks for Econometrics”, *Journal of Economic Perspectives*, 28, 3-28.

Verme, P. (2011) “Life Satisfaction and Income Inequality”, *Review of Income and Wealth*, 57, 111-127.

Verme, P. (2018) “Relative Deprivation and Satisfaction: Empirical Findings”, in C. D’Ambrosio (ed.), Handbook of Research on Economic and Social Well-Being, Chapter 16, 356-376.

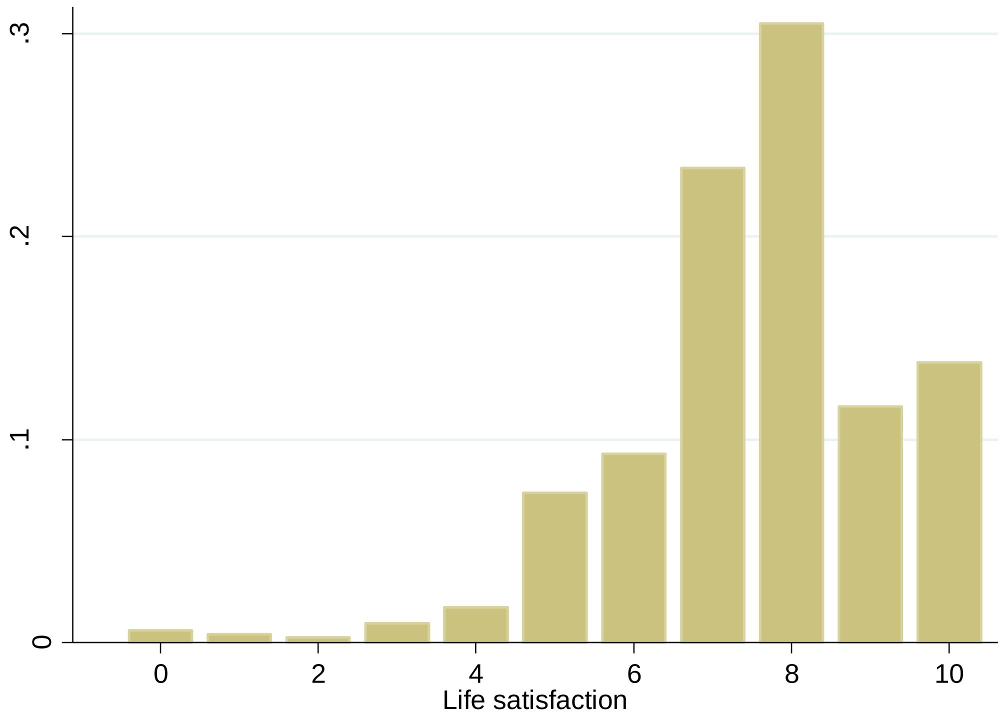
Zhao, Y., Li, Y., Heath, A. and Shryane, N. (2017) “Inter and Intra Generational Social Mobility Effects on Subjective Well-Being, Evidence from Mainland China”, *Research in Social Stratification and Mobility*, 48, 54-66.

Zang, E. and De Graaf, N. D. (2016) “Frustrated Achievers or Satisfied Losers? Inter and Intergenerational Social Mobility and Happiness in China”, *Sociological Science*, 3, 779-800.

Zou, H. and Hastie, T. (2005) “Regularization and Variable Selection Via the Elastic Net”, *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 67, 301-320.

Figures

Figure 1: Density of life satisfaction



Note: Source: Own elaboration, data from CIS 2017.

Tables

Table 1: Summary Statistics of the Household per capita adjusted income, by cohorts.

Age	Mean household adjusted income	Sd. of the household adjusted income
30-40	23,777.26	13,442.23
41-50	24,974.19	13,149.09
51-60	22,498.83	13,858.61

Note: Sd. Stands for Standard Deviation. All values in €2017. Source: Own elaboration, data from CIS 2017

Table 2: Summary Statistics of the occupation of the respondents and their fathers

	Respondents	Fathers
ISCO-08=9, Unqualified workers (1)	12.16%	5.73%
ISCO-08=4-8, Semi-qualified and qualified workers (2)	53.61%	71.42%
ISCO-08=3, Technicians and support professionals (3)	14.16%	9.12%
ISCO-08=1-2, Managers and professionals (4)	20.07%	13.73%
Mean	2.42	2.30
Standard Deviation	0.94	0.78

Note: Source: Own elaboration, data from CIS 2017

Table 3: Summary Statistics of the education of the respondents and their fathers.

	Respondents	Fathers
ISCED=0-1, Primary education (1)	29.63%	66.90%
ISCED=2, Low secondary education (2)	13.12%	11.12%
ISCED=3-4, Upper secondary education (3)	29.89%	9.82%
ISCED=5-8, Post-secondary (4)	27.37%	12.16%
Mean	2.54	1.67
Standard Deviation	1.18	1.07

Note: Source: Own elaboration, data from CIS 2017

Table 4: Summary Statistics of the control variables.

	Mean	Standard Deviation
Age	45.05	8.33
Gender (Men=1)	0.50	0.50
Health Status	3.91	0.81
Civil Status (Married=1)	0.60	0.49
Children (Have children=1)	1.27	0.45

Note: Health Status ranges from 1 (very bad) to 5 (very good). Source: Own elaboration, data from CIS 2017

Table 5: Relation between the age of the respondent and wave employed to impute fathers' income.

	Group 1	Group 2	Group 3
Age of the respondent in the main sample	30-35	36-45	45-60
EPF wave	2000	1990-91	1980-81
Year when the age-median observation had 16	2000	1992	1980

Note: Source: Own elaboration, data from CIS 2017

Table 6: Summary statistics of the imputed fathers' income for the three EPF waves.

Wave	1980-1981	1990-1991	2000
Number of observations	14,987	15,567	15,567
Log(λ*)	-7.2844	-6.5427	-5.8132
Mean income imputed with α=1	23,205.84	24,333.60	24,236.88
Sd of income imputed with α=1	7,829.52	6,046.32	5,462.52

Note: Sd stands for Standard Deviation. All monetary values in €2017. Source: Own elaboration, data from EPF 1980/81, 1990/91 and 2000.

Table 7: Intergenerational income transition matrix.

		Household per capita adjusted income					
		Quintile 1	Quintile 2	Quintile 3	Quintile 4	Quintile 5	Total
Fathers income	Quintile 1	71	51	51	30	31	234
	Quintile 2	59	49	41	49	32	230
	Quintile 3	46	42	54	40	45	227
	Quintile 4	43	49	50	44	48	234
	Quintile 5	31	24	49	46	76	226
	Total	250	215	245	209	232	1,151

Note: Source: Own elaboration, data from CIS 2017

Table 8. Transition matrix of occupation.

		Respondents occupation				
		ISCO-08=9, Unqualified workers	ISCO-08=4-8, Semi-qualified and qualified workers	ISCO-08=3, Technicians and support professionals	ISCO-08=1-2, Managers and professionals	Total
Fathers occupation	ISCO-08=9, Unqualified workers	20	31	6	9	66
	ISCO-08=4-8, Semi-qualified and qualified workers	105	478	103	136	822
	ISCO-08=3, Technicians and support professionals	7	48	23	27	105
	ISCO-08=1-2, Managers and professionals	8	60	31	59	158
	Total	140	617	163	231	1,151

Note: Source: Own elaboration, data from CIS 2017

Table 9. Transition matrix of education.

		Education of the respondents				
		ISCED=0-1, Primary education	ISCED=2, Low secondary education	ISCED=3-4, Upper secondary education	ISCED=5-8, Post-secondary education	Total
Education of the fathers	ISCED=0-1, Primary education	314	104	213	139	770
	ISCED=2, Low secondary education	17	21	48	42	128
	ISCED=3-4, Upper secondary education	8	17	49	39	113
	ISCED=5-8, Post-secondary education	2	9	34	95	140
	Total	341	151	344	315	1,151

Note: Source: Own elaboration, data from CIS 2017

Table 10: Life Satisfaction and Income Mobility

Variables	Model 1	Model 2	Model 3	Model 4	Model 5
Mobility	0.040				
	(0.066)				
Up Mob.		-0.053			
		(0.122)			
Down Mob.		-0.134			
		(0.128)			
Mob. Intensity			0.022		
			(0.034)		
Mob. Intensity (squared)				-0.003	
				(0.011)	
Mob. (-4 steps)					0.393
					(0.322)
Mob. (-3 steps)					-0.334
					(0.245)
Mob. (-2 steps)					-0.336
					(0.188)
Mob. (-1 steps)					-0.005
					(0.149)
Mob. (+1 steps)					-0.127
					(0.161)
Mob. (+2 steps)					0.024
					(0.148)
Mob. (+3 steps)					0.148
					(0.215)
Mob. (+4 steps)					-0.350*
					(0.208)
Income quintile 2	0.354**	0.358**	0.351**	0.364**	0.445**
	(0.169)	(0.170)	(0.172)	(0.174)	(0.182)
Income quintile 3	0.518***	0.523***	0.510***	0.537***	0.558***
	(0.152)	(0.153)	(0.157)	(0.155)	(0.171)
Income quintile 4	0.509***	0.514***	0.495***	0.545***	0.479***
	(0.169)	(0.169)	(0.177)	(0.160)	(0.183)
Income quintile 5	0.437**	0.432**	0.421**	0.493***	0.473**
	(0.178)	(0.178)	(0.195)	(0.157)	(0.197)
Sex	0.032	0.029	0.032	0.032	0.022
	(0.094)	(0.094)	(0.094)	(0.094)	(0.094)
Age	-0.095	-0.094	-0.095	-0.096	-0.087
	(0.061)	(0.061)	(0.061)	(0.061)	(0.061)
Age squared	0.001	0.001	0.001	0.001	0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Health	0.550***	0.551***	0.550***	0.549***	0.544***
	(0.068)	(0.068)	(0.069)	(0.069)	(0.068)
Married	0.566***	0.571***	0.566***	0.573***	0.572***
	(0.118)	(0.118)	(0.118)	(0.117)	(0.118)
Kids	-0.054	-0.050	-0.054	-0.053	-0.061
	(0.122)	(0.123)	(0.122)	(0.122)	(0.122)
Constant	7.405***	7.441***	7.426***	7.397***	7.317***
	(1.350)	(1.351)	(1.349)	(1.351)	(1.352)
Observations	1,151	1,151	1,151	1,151	1,151
R-squared	0.147	0.148	0.147	0.147	0.155

*** p<0.01, ** p<0.05, * p<0.1

Note: Source: Own elaboration, data from CIS 2017. Mob. Stands for "Mobility". Robust standard errors in parentheses.

Table 11: Life Satisfaction and Occupational Mobility

Variables	Model 1	Model 2	Model 3	Model 4	Model 5
Mobility	-0.053				
	(0.088)				
Up Mob.		0.063			
		(0.141)			
Down Mob.		0.153			
		(0.141)			
Mob. Intensity			-0.031		
			(0.059)		
Mob. Intensity (squared)				-0.004	
				(0.031)	
Mob. (-3 steps)					-0.793
					(0.559)
Mob. (-2 steps)					0.265
					(0.190)
Mob. (-1 steps)					0.082
					(0.182)
Mob. (+1 steps)					0.014
					(0.170)
Mob. (+2 steps)					0.180
					(0.186)
Mob. (+3 steps)					-0.917
					(0.609)
Qualified	0.006	0.069	-0.008	-0.036	-0.036
	(0.185)	(0.199)	(0.179)	(0.175)	(0.228)
Technicians	-0.001	-0.012	-0.025	-0.074	-0.074
	(0.226)	(0.226)	(0.215)	(0.196)	(0.238)
Directives and professionals	0.319	0.318	0.309	0.239	0.192
	(0.234)	(0.235)	(0.235)	(0.187)	(0.269)
Sex	0.049	0.050	0.049	0.048	0.041
	(0.093)	(0.093)	(0.093)	(0.093)	(0.093)
Age	-0.086	-0.090	-0.086	-0.084	-0.088
	(0.061)	(0.061)	(0.061)	(0.061)	(0.061)
Age squared	0.001	0.001	0.001	0.001	0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Health	0.573***	0.573***	0.573***	0.575***	0.576***
	(0.069)	(0.069)	(0.069)	(0.069)	(0.069)
Married	0.612***	0.610***	0.611***	0.605***	0.616***
	(0.121)	(0.120)	(0.121)	(0.119)	(0.121)
Children	-0.029	-0.032	-0.029	-0.030	-0.032
	(0.122)	(0.122)	(0.122)	(0.122)	(0.122)
Constant	7.320***	7.325***	7.332***	7.338***	7.364***
	(1.377)	(1.377)	(1.375)	(1.373)	(1.382)
Observations	1,151	1,151	1,151	1,151	1,151
R-squared	0.136	0.137	0.136	0.136	0.142

*** p<0.01, ** p<0.05, * p<0.1

Note: Source: Own elaboration, data from CIS 2017. Mob. Stands for "Mobility". Robust standard errors in parentheses.

Table 12: Life Satisfaction and Educational Mobility

Variables	Model 1	Model 2	Model 3	Model 4	Model 5
Mobility	0.108				
	(0.082)				
Up Mob.		0.059			
		(0.128)			
Down Mob.		-0.190			
		(0.188)			
Mob. Intensity			0.043		
			(0.043)		
Mob. Intensity (squared)				0.010	
				(0.017)	
Mob. (-3 steps)					1.340
					(0.743)
Mob. (-2 steps)					-0.247
					(0.431)
Mob. (-1 steps)					-0.248
					(0.204)
Mob. (+1 steps)					0.078
					(0.185)
Mob. (+2 steps)					-0.037
					(0.156)
Mob. (+3 steps)					0.118
					(0.162)
Compulsory Secondary	-0.041	-0.000	-0.002	0.016	0.002
	(0.197)	(0.217)	(0.188)	(0.185)	(0.242)
Post-compulsory Secondary	0.004	0.043	0.023	0.059	0.111
	(0.144)	(0.161)	(0.144)	(0.137)	(0.171)
Tertiary	0.100	0.127	0.104	0.140	0.117
	(0.148)	(0.157)	(0.156)	(0.152)	(0.162)
Sex	0.055	0.057	0.054	0.054	0.051
	(0.094)	(0.094)	(0.094)	(0.094)	(0.094)
Age	-0.091	-0.091	-0.089	-0.089	-0.093
	(0.062)	(0.062)	(0.062)	(0.062)	(0.062)
Age squared	0.001	0.001	0.001	0.001	0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Health	0.574***	0.574***	0.575***	0.575***	0.570***
	(0.069)	(0.069)	(0.069)	(0.069)	(0.069)
Married	0.609***	0.610***	0.611***	0.614***	0.613***
	(0.119)	(0.119)	(0.119)	(0.119)	(0.119)
Children	-0.035	-0.034	-0.028	-0.025	-0.027
	(0.122)	(0.122)	(0.122)	(0.122)	(0.123)
Constant	7.426***	7.425***	7.378***	7.347***	7.473***
	(1.392)	(1.392)	(1.393)	(1.394)	(1.395)
Observations	1,151	1,151	1,151	1,151	1,151
R-squared	0.135	0.135	0.134	0.134	0.137

*** p<0.01, ** p<0.05, * p<0.1

Note: Source: Own elaboration, data from CIS 2017. Mob. Stands for "Mobility". Robust standard errors in parentheses.

Technical Appendix

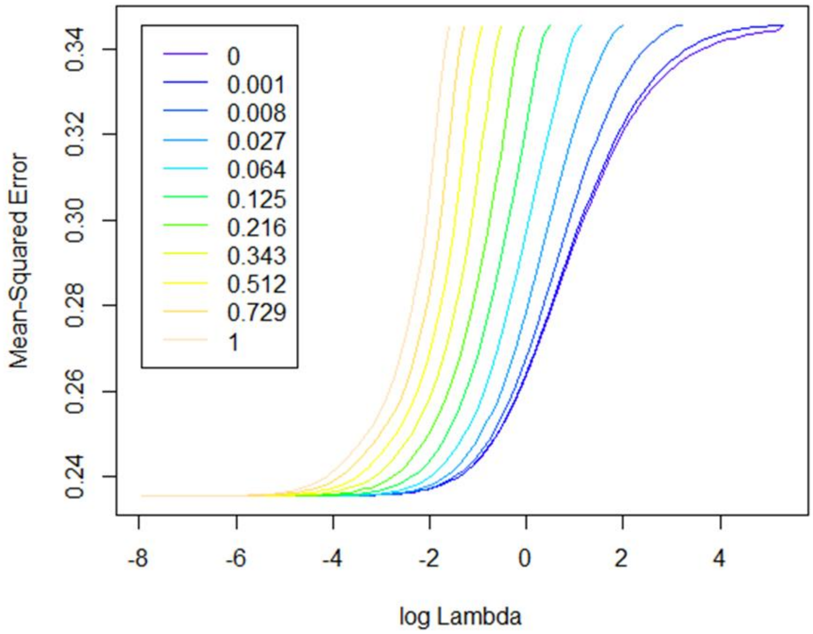
The values of λ and α in Equation (5) should not be arbitrarily selected. The former (λ), controls the importance of the regularization term and is equal or higher than zero. The latter (α), is the elastic net regulator obtained from a linear combination of two standard Machine Learning techniques, and its possible values range between 0 and 1. In particular, Equation (5) is equivalent to the Least Absolute Shrinkage and Selection Operator (LASSO) when α=1, but it is equal to a ridge regression when α=0 (see Varian, 2014). Different combinations of these parameters might lead to different imputations, leading to non-robust results. Solving this problem, the proposed algorithms compute all possible tunings and combinations of λ and α to finally select the one that delivers the smallest Mean Squared Errors (MSE).

To make the tuning as transparent as possible, we plot the relation between the MSEs and several values of λ and α. This way, we show that our tuning provides the smallest possible associated MSEs. Figures TA1 to TA3 correspond to the imputations performed on the EPF 1980/81, EPF 1990/91 and EPF 2000, respectively.

Figure TA1 should be interpreted as follows. The MSE produced by Equation (5) is stable at less than 0.24 for any α value lower than $\log(\lambda) \approx$ -5. At that point, the MSE associated with α = 1 starts rising, while the MSEs associated with the rest of possible α values remain constant. Clearly, for values of α smaller than 1, the associated MSEs sequentially take off from $\log(\lambda) \approx$ -5 onwards, until the MSE produced by the ridge regression (α = 0) rises at $\log(\lambda) \approx$ -3. Figures TA2 and TA3 are interpreted similarly, with diverse MSEs associated with different parameter settings.

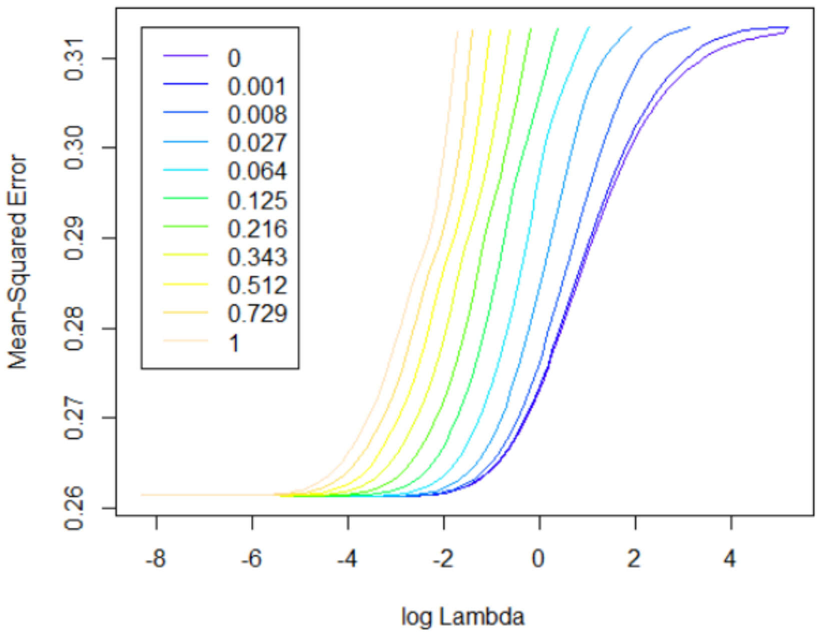
As the algorithm searches for the values of the parameters associated with the lowest stable MSE, it provides the combination of λ and α that leads to the most accurate imputation. For instance, in Figure TA1, although it cannot be graphically distinguished, this occurs when $\log(\lambda)$ equals -7.2844. At that point, no matter the value of α we select, the MSE is constant and has the lowest possible value. For any other combination, the associated MSE is higher (for higher λ) or stable (for lower λ). Similarly, in Figure TA2, the optimal $\log(\lambda)$ equals -6.5427, and in Figure TA3, it reaches -5.8132. Then, these values conform the parameter tuning, as summarized in Table 6, being the selected α equal to 1, as the LASSO regression has been used more often in the literature. However, results for α equal to 0 (and the same lambda values) are available upon request, varying the resulting imputations in around 20€.

Figure TA1: Mean squared error provoked by the regularization term in an OLS regression



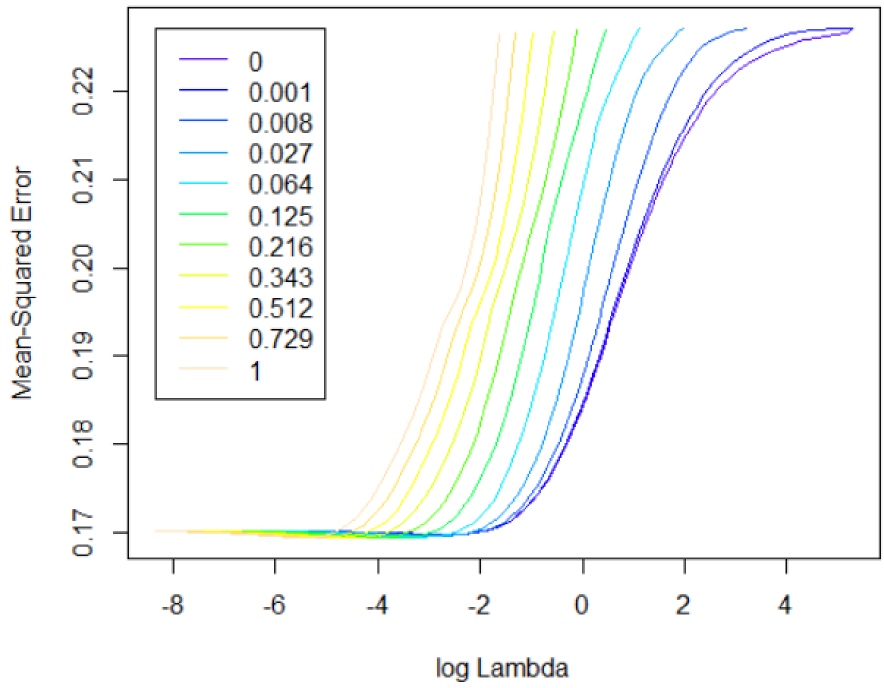
Note: Source: Own elaboration, data from EPF 1980/81.

Figure TA2: Mean squared error provoked by the regularization term in an OLS regression.



Note: Source: Own elaboration, data from EPF 1990/91.

Figure TA3: Mean squared error provoked by the regularization term in an OLS regression



Note: Source: Own elaboration, data from EPF 2000.

Appendix

Table A.1: Intergenerational Elasticity

Variable	Dependent: Child's income
Fathers' Income	0.459***
	(0.064)
Constant	4.788***
	(0.641)
Observations	1,151
R-squared	0.048

*** p<0.01, ** p<0.05, * p<0.1

Note: Source: Own elaboration, data from CIS 2017. Robust standard errors in parentheses.