Tax Morale and Perceived Intergenerational Mobility: a Machine Learning Predictive Approach

Caferra, Rocco and Morone, Andrea

April 2019

Tax Morale and Perceived Intergenerational Mobility: a Machine Learning Predictive Approach

Rocco Caferra[a], Andrea Morone[b] *

[a] rocco.caferra@gmail.com

[b] a.morone@gmail.com

Abstract

The purpose of this paper is to investigate the linkage between the perceived intergenerational mobility and the preferences for tax payment. Unfortunately, we do not have a unique dataset, however missing data might be predicted by employing different methods. We compare the efficiency of k-nearest-neighbors (kNN), Random Forest (RF) and Tobit-2-sample-2-Stage (T2S2S) techniques in predicting the perceived intergenerational mobility, hence we exploit the predicted values to estimate the relation with tax morale. Results provide evidence of a strong negative relation between perceived mobility and tax cheating, suggesting that fairness in tax payment has also to be seen on the light of the perceived efficiency of the welfare state in providing more opportunities across generations.

1 Introduction

High government spending is strictly dependent to an higher level of fiscal capacity. Equivalently, a high level of taxation is not always well-seen by tax-payers who observe a reduction of their budget constraint. These aspects justify the existence of a trade-off between sustainability and willingness to pay taxes as illustrated in the standard model of tax evasion (Allingham and Sandmo, 1972), where risk-adverse taxpayers maximize the expected utility of income on the basis of the probability to be detected and the total amount of the fine due. The idea that cheating on taxes depends not only on monetary but also on psychological factors took hold in 1990. Luttmer and Singhal

*Università degli Studi di Bari "Aldo Moro"; Dipartimento di Economia, Management e Diritto dell'Impresa Università degli Studi di Bari Aldo Moro. Largo Abbazia Santa Scolastica, 53 70124 Bari (Italy).

(2014) defined tax morale as the totality of non-pecuniary motivations lowering people propensity to pay taxes. That is, investigating the interaction effects between formal and informal institutions in shaping the tax morale might lead to new important insights, for instance inequality might trigger tax cheating. As discussed by Alesina and Rodrik (1994), in a society where there are unequal access to resources, there will be an high demand for redistribution and this conflict obstacles the economic growth. Conversely, Senik (2009) noticed that attitudes towards inequality depend on an individual's calculation of his or her own chances and possible direction of mobility. To wit, whether on one hand mobility might be enhanced by the state, the perception of a quite immobile society might disencourage tax payment, since poorer expect to suffer from greater inequality in the future and they found ineffective the redistribution role of tax levy. Unluckily, we have to merge information from two different datasets (The European Values Survey (EVS) 2008 and the International Social Survey Programme 2009 (ISSP)) to estimate the effect on tax cheating. Literature proposed several methods to treat missing data (see Efron and Hastie (2010)) and in the methodology we propose, we compare k-nearest-neighbors(kNN), Random Forest and Tobit-2-sample-2-Stage (T2S2S) accuracy in the estimation of the perceived social mobility (INT). The variable under consideration is an existing data of the ISSP survey (*"If you compare this job to the job your father had when you were <14,15,16>, would you say that the level of status of your job is (or was)?"*) that is missing in the EVS one. Data have been merged on the basis of a common set of information given.
The remainder of this paper is organized as follows. We first present a Literature Review (section 2) we discuss the accuracy of the methodology proposed (section 3). Then, we report our results in section 4. Finally, section 5 concludes.

2 Literature Review

The relationship between perceived intergenerational mobility and the preferences for tax payment is strictly related, at least, with two stands of literature : the one which analyze social and individual responsibility toward tax payment and the other which focus on the driving factor of the intergenerational mobility. The standard model of Allingham and Sandmo (1972) has been considered as the reference point to explain cheating on taxes, stating that a profit-maximization agents maximized their own monetary pay-offs according to the expected probability to be fined. However it has not been able in itself to explain the whole attitude toward tax evasion, which is why the concept of tax morality has become increasingly importance in defining the intrinsic motivation to pay taxes (Torgler, 2002b, 2005c, 2012; Torgler and Schneider, 2007). Generally, trustness in public authorities, particularly in the fairness and in the effectiveness of the public expenditure potentially improve tax compliance (Torgler and Schneider (2004), Torgler et al. (2007, 2008, 2010), Barone and Mocetti (2011),Trüdinger and Hildebrandt (2013), Vythelingum et al. (2017)).

Moreover, inequality and GDP performance influence citizens attitude toward taxes (Williams and Krasniqi (2017), Williams and Martinez (2014), Doerrenberg and Peichl (2010)), since the higher is perceived the meritocracy and the perception of the social and the economic progress, the higher seems to be human responsibility toward duties, such as tax payment. A recent study of Amendola et al. (2018) pointed out the negative relation between inequality and happiness. To sum up, we can state that satisfaction lays down the basis to the willingness to comply and tax evasion is not merely a question of money, inasmuch a prominent role is played by the interaction effects between formal and informal institutions (Horodnic, 2018).

Although income inequality might be seen as a signal of intra-generational mobility, the perceived chances to ameliorate social and labor position in the future might be also used as an index of satisfaction and inter-generational mobility. Piketty (2000) revised the key factors that determine the transmission and the persistence of inequality across generations. Mulligan (1997) focused on the perpetuated inequalities in earning, while Becker and Tomes (1986) pointed out the theory of the efficient ability transmission, where each subjects is able to invest in education since it returns profitability regardless of the social status. Franzini and Pianta (2015) identified four "engines of inequality": 1) the power of capital over labor which reduce wages; 2) the rise of oligarchs capitalism, where a limited number of people concentrate a large and increasing share of income and wealth; 3) the individualisation of social and economic conditions,where the identification into a social status create discrimination,different level of wages and the fragmentation of the population slow down the growth; 4)the retreat of politics and its controversial role in the redistribution of wealth. Moreover, Benabou (1993) involved the differences in counties to explain the amplification of the future human capital disparity and self-fulfilling beliefs refers to racial or social discrimination or simply to different cultural attitude. With regard to the preference for redistribution, Alesina, Stantcheva and Teso (2016) evidenced that pessimism about social mobility are correlated with policy preferences and tends to favor more generous redistributive policies, especially about the equality of opportunity.

The aforementioned considerations strengthen our research hypothesis: does a lower perception of intergenerational mobility discourage tax morale? In the next section we provide an extensive analysis of our approach.

3 Data and Methodology

3.1 Some Theoretical Notes

As announced, we do not have a unique dataset therefore we have to exploit information available to relate the dependent variable of our model (tax morale) and the independent one (perceived intergenerational mobility). The problem might be summarized as follow:

- the ISSP dataset contains information on the perceived intergenerational mobility (INT) but unfortunately lacks of the respondent preferences toward Tax Cheating(TC);

- the EVS dataset displays the opposite case;
- Both of the datasets share common information (X) about respondents' characteristics and preferences.

The basic idea is that we can predict the relation between Intergenerational Mobility INT and X, then impute the variable where it is missing by exploiting the same variable. Let us consider

$$INT^T = \{INT_{ISSP}, INT_{EVS}\}$$

as the perceived intergenerational mobility, where INT_{EVS} is missing and

$$X^T = \{X_{ISSP}, X_{EVS}\}$$

as the common information matrix about respondents' characteristics and preferences.

Now we can exploit the relation between:

$$P(INT_{ISSP}) \sim P(INT_{ISSP}|X_{ISSP})$$

to obtain

$$P(INT_{EVS}) \sim P(INT_{EVS}|X_{EVS}, X_{ISSP}, INT_{ISSP})$$

that is, the estimation methods seems to do not be in function of the tax morale, enhancing the validity of the subsequent estimation (henceforth *final model*):

$$P(TC_{EVS}|INT_{EVS}, Z_{EVS})$$

Where Z are some other control variables.
On the other hand, we can not ensure that the instrumental variable choose are independent from the response variable (TM), hence endogeneity problem might arise. As demonstrated in Jerrim, Choi and Simancas (2017), the most powerful will be the prediction, the lower will be the endogeneity bias, we have to be worried about the influence that the estimator has on the dependent variable of the final model. Formally :

$$P(\widehat{INT}) \sim P(INT|X) \perp P(TC)$$

if and only if

$$P(TC) \sim P(TC|X) \sim P(TC)P(X)$$

That it can be defined as the *exogeneity* condition. In our case:

$$\begin{aligned}
P(\widehat{INT}) &\sim P(INT|X) \\
P(TC|\widehat{INT}) &\sim P(TC|INT) \sim P(TC|INT + \widehat{INT} - \widehat{INT}) \\
P(TC|\widehat{INT}) &\sim P(TC|INT) + P(TC|\widehat{INT} - INT)
\end{aligned}$$

Where $P(TC|\widehat{INT} - INT)$ can be seen as the bias of our model. Hence, what we expect to observe is $E(TC|\widehat{INT} - INT) = 0$ to fit the consistency of the model. Basically, it happens whether:

- The estimated data are closer the observed ones (($\widehat{INT} - INT$) approaches to zero);
- The distribution is randomly distributed as the error term.

In short, we first fit the best method to classify data by taking all the variables [1] and after remove the effect of the potential endogenous ones.
Generally, there are different estimation techniques of missing values (for a complete overview see Efron and Hastie (2017) or Hastie, Tibshirani and Friedman (2017)). We check whether:

Proposition 1 Imputing data by classification of observations (for instance by using a machine learning techniques) outperforms classical methods of imputation based on a general shaping of the functional form of variable relationships.

Proposition 2 After having found the technique that best fits our purpose, we aim to identify some relation between the perceived earnings mobility and tax compliance. In the simplest vision of the proof, we wish to confirm that meritocracy influences subjects willingess to contribute to the sustainability of the welfare state;

We employ three different approaches to impute data: 1) Tobit-Two-Sample-Two-Stage (T2S2S), 2) k-Nearest-Neighbors (kNN) and 3) Random forest (RF). The first type of estimation belongs to the class of model based (supervised), methods where the relation between response variables and regression is a priori predetermined, while the last two approaches are based upon the most recent machine learning techniques, that is, no relation is defined a priori, maybe the imputation of the data is made by considering some metrics of "similarity" between data, hence the shape of the distribution is calculated by inferring on them.

3.1.1 Tobit-Two-Sample-Two-Stage (T2S2S)

Given a censored response variable (in our case tax morale lies between 1 and 10) we might exploit a regression model for prediction. Let X^T be a vector of characteristics $X^T = \{X_1, X_2, ..., X_p\}$, we first predict $\hat{\beta}^T = \{\hat{\beta}_0, \hat{\beta}_1, ..., \hat{\beta}_p\}$ by exploiting:

$$INT_{ISSP} = \hat{\beta}_0 + \sum_{j}^{p} X_{ISSPj}\hat{\beta}_j$$

and after

$$INT_{EVS} = \hat{\beta}_0 + \sum_{j}^{p} X_{EVSj}\hat{\beta}_j$$

[1] Roughly speaking, we states that, for instance, *if it might be considered true* that a part-time elder employed respondent with and low level of income comes from a family with a full-employed father perceive its labor position lower than the one of his parent, *it makes some sense* to extend this perception to subjects sharing the same characteristics.

might be predicted by using the same coefficient. As discussed above, one limit of such approach is that we are conjecturing the global shape of the relation.

3.1.2 k-Nearest-Neighbors (kNN)

Nearest-neighbor methods use those observations in the ISSP dataset closest in input space to X to form INT_{ISSP}. Hence, to fit the predicted INT_{EVS} we have to consider the same cluster of characteristics linked to INT_{ISSP} and then impute $\widehat{INT_{EVS}}$

$$\widehat{INT_{EVS}} = \frac{1}{k} \sum_{x_i \in N_k(x)} INT_{ISSPi}$$

where $N_k(x)$ is the neighborhood of x defined by the k closest points x_i in the sample. Closeness implies a metric, which we assume to be the Euclidean distance.

3.1.3 Random forest (RF)

Since the response variable is continuous, we employ an unsupervised regression threes model. We adapt the description of the algorithm of Hastie, Tibshirani and Friedman (2017). Random forest is based on the bagging, that is, it both bootstrap from the original sample and aggregate the estimation obtained. In detail:

1. Bootstrapped samples (from b=1 to B) are drawn from the original one:
2. A random forest T_b is grown for each subsample, by creating nodes through the m out of p variables selected at random;
3. to predict $\widehat{INT_{EVS}}$ we have:

$$\widehat{INT_{EVS}} = \hat{f}^B_{rf}(X) = \frac{1}{B} \sum_{b=1}^{B} T_b(X)$$

Even in this case, the advantage is that the functional form of the regression is locally determined through nonlinear estimators (trees).

3.2 Data Description

The EVS Longitudinal Data Files 1981-2008 include data and documentation of 48 countries/regions that participated in the four EVS waves. In our case, we restrict the analysis to the Italian survey of EVS 2008, using 1519 observations. On the other side, The International Social Survey Programme 2009 (ISSP) focused on social inequality perception surveyed 1084 italian respondents. In order to do not lose observation, we impute missing data in the two datasets , coming ahead with a final dataset of 2603 independent observation. The X^T matrix is formed by[2]:

$$
\begin{aligned}
X^T =& \{X_1, X_2, ..., X_p\} \\
X^T =& \{edu, age, occ, edu, hhmembers, region, religion, , laborunion, \\
& , sexdiscr, occ_r, occ_f, income, inequality, unemp\text{-}perc, hardwork, \\
& election\}
\end{aligned}
$$

where:

- *Age* is a numeric variable reporting the respondent's age ;
- *sex* is a categorical variable reporting the respondent's age ;
- *education* is a categorical variable reporting the respondent's educational level ;
- *hh members* is a continuous variable reporting the household members ;
- *region* is a categorical variable reporting the twenty italian regions;
- *religion* is a categorical variable controlling for the religiosity of the respondent ;
- *labor union* is a dummy that accounts for respondent's belonging to union membership ;
- *sexdiscr* is a dummy that considers the disparities perception to be born woman or man;
- occ_r is a categorical variable reporting the occupational status;
- occ_f is a categorical variable reporting the previous occupational status of father respondent;
- *election* is a categorical variable reporting the wing of the parties voted at the latest election;
- *income* is a continuous variable reporting the wing of the monthly family income;
- *hardwork* is a continuous variable reporting the importance of hard work in daily life;
- $unemp-perc$ is a continuous variable reporting the importance of helping unemployed people;
- *inequality* is a continuous variable reporting the perception of income inequality in the Country;

The perceived intergenerational mobility (INT) is calculated by using a the following question from the ISSP survey: *"If you compare this job to the job your father had when you were <14,15,16>, would you say that the level of status of your job is (or was)?"* and moves from 1 (lower mobility) to 5 (higher mobility). In the final model, we relate tax morale and the estimated perceived intergenerational mobility by controlling for:

- the economic position of the respondent according to its level of income;
- post-materialism attitude of the society;
- trustworthiness in the development of democracy;
- macro-area effects (north,center and south)

Now we can move on evaluating the accuracy of the estimation.

[2]In the appendix (Table 3A) might be found the correspondent questions and the description of the variable .

4 Analysis and Results

For the sake of soundness, firstly we discuss the accuracy of the model by applying a train/test split of the ISSP sample, where the INT variable is already known. Basically, we aim to consider the validity of our results in a twofold ways: firstly by controlling the normalized root mean squared error (NRMSE) and the correlation between the observed value and the fitted one in the test data and after by checking the robustness of the results with respect to the size of the training dataset. Intuitively, we are going to consider the variable at issue as known in the training data and as a missing value in the test one, even if we known the real observed values, hence they can be used post-hoc to evaluate the forecasting. Since the ISSP dataset covers only the 41 % of the complete dataset, this might be seen as a real menace to the validity of the model, this is why we test whether somethings changes according to the dimension of the training test.

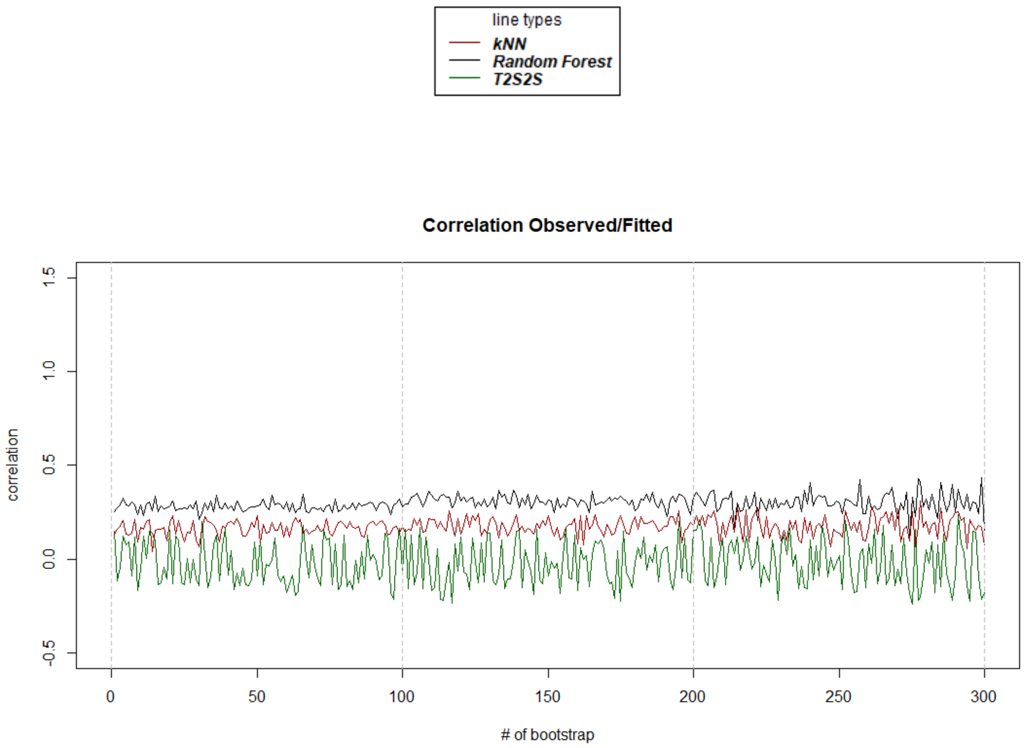

Figure 1: Correlation plot. The dashed lines divided the graph area into three different sections: left (training data 25%), center (training data 50%) and right (training data 75%). Random Forest outperforms both kNN (T-value: -43.942,p-value=0.00) and T2S2S (T-value: -45.589,p-value=0.00).

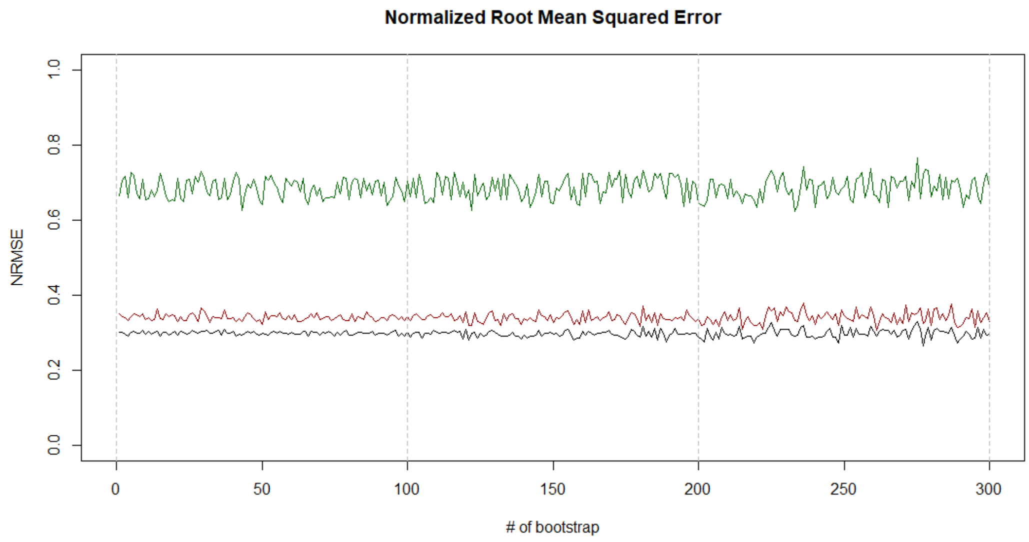

Figure 2: NRMSE plot. The dashed lines divided the graph area into three different sections: left (training data 25%), center (training data 50%) and right (training data 75%). Random Forest outperforms both kNN (T-value:80.262,p-value=0.00) and T2S2S (T-value: 241.92, p-value=0.00)

Proposition 1 As displayed, Random Forest best fits both in terms of correlation and normalized root mean squared error[3]. Moreover, test accuracy do not significantly improved according to the growth of the training data. Jointly considering the positivity of these aspects, we can advance with the rest of the analysis.

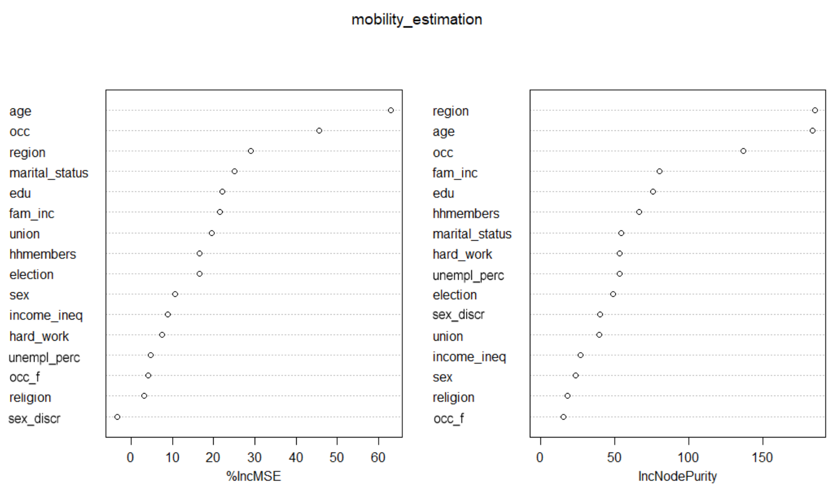

Figure 3: Variance Importance Plot in the estimation of the perceived social mobility

As expected, *age,type of occupation and education, family members and the level of income* influence the attitude toward the inequality. Firstly, we

[3]The analysis has been conducted in R though missForest, RandomForest, kNN and AER packages

report the correlation matrix to find the continuous variable which potentially affect both tax cheating and have been used to classify the estimated perceived intergenerational mobility. Hence we remove their effect from the variable by taking the residual terms.

As can be seen, the educational level, the perception of the inequality and the attitude toward unemployment people are correlated. In general we find no higher correlation, maybe we can remove the average effect of age (-0.1780) and of the perceived inequality of income (-0.1301) making a pairwise regression [4]

	hhmembers	edu	income ineq	hard_work	age	family income	INT	TC
hhmembers	1.0000							
education	0.1053	1.0000						
income ineq	-0.0332	-0.0649	1.0000					
hard work	-0.0168	-0.0261	0.0871	1.0000				
age	-0.4898	-0.4073	0.1346	0.0157	1.0000			
family inc	0.1419	0.2962	-0.0906	0.0164	-0.1964	1.0000		
INT	-0.2936	-0.1702	0.0462	0.0188	0.5794	0.1686	1.0000	
TC	0.0710	0.0187	-0.1301	0.0151	-0.1780	-0.0631	-0.1485	1.000

Table 1: correlation matrix

In fact, by running a first tobit regression with the reduced form of the model, we have that:

TC	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
INT	-2.821	.443	-6.36	0.000	-3.692	-1.951
cons	8.855	1.401	6.32	0.000	6.105	11.605
σ	4.425	0.151			4.129	4.722

Table 2:Log likelihood = -2244.3681, LR chi2(1) =41.15, Prob > chi2 = 0.0000

The estimate seems to be upward biased, in particular we noted that the error correlation $Cov(\widehat{INT}, u)$=-0.1246^{*}, confirming the endogeneity already stated from the previous correlation matrix. We pairwise each other the regressors in order to delete dependency and get them uncorrelated. Since Age is related with the other variable, we pairwise regressed them on this regressor. At the end of the procedure, we succeded in eliminating the correlation between variables. As final step, we look forward to remove their influence and the one of the categorical variables inserted in the final model (Geographical Area, Democracy Perception, Post Materialism Index and occupational status) from the perceived intergenerational mobility (INT).

[4] the variable are jointly correlated, which is why we firstly remove the effect they have on each other

	$\widehat{ineq}$	$\widehat{education}$	$\widehat{INT}$	age	TC	$\widehat{familyincome}$	$\widehat{hhmembers}$	hard work
$\widehat{ineq}$	1.0000							
$\widehat{education}$	0.0000	1.0000						
$\widehat{INT}$	0.0000	-0.0000	1.0000					
age	-0.0000	0.0000	-0.0000	1.0000				
TC	-0.1078	-0.0589	-0.0314	-0.1780	1.0000			
$\widehat{familyincome}$	-0.0000	-0.0000	0.0000	0.0000	-0.0956	1.0000		
$\widehat{hhmembers}$	-0.0000	0.0000	0.0000	-0.0000	-0.0133	-0.0000	1.0000	
hard work	0.0855	-0.0216	0.0083	0.0157	0.0151	0.0316	-0.0191	1.0000

Table 3: Correlation Matrix with estimated residuals

Here again, we consider the reduced version of the model:

TC	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
$\widehat{INT}$	-2.069	0.728	-2.84	0.005	-3.497	-0.641
cons	-0.151	0.164	-0.92	0.359	-0.474	0.172
σ	4.489	0.153			4.188	4.790

table 4. Reduced form. Log likelihood = -2263.0187 , Prob > chi2 = 0.0499

TC	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
INT	-2.087	0.698	-2.99	0.003	-3.457	-.718
ineq perc	-0.589	0.145	-4.04	0.000	-0.875	-0.303
education	-0.255	0.160	-1.60	0.111	-0.569	0.058
age	-0.045	0.0111	-4.09	0.000	-0.0672	-0.023
family income	-0.302	0.087	-3.44	0.001	-0.475	-0.130
IT: North	0.156	0.373	0.42	0.675	-0.575	0.889
IT: South	0.554	0.381	1.46	0.146	-0.192	1.302
Mixed	-.0266	0.339	-0.79	0.431	-0.932	0.398
Post-materialist	-0.949	0.441	-2.15	0.032	-1.815	-0.082
Not very satisfied	0.933	0.368	2.53	0.011	0.210	1.656
Rather satisfied	0.947	0.403	2.35	0.019	0.155	1.739
Very satisfied	1.149	0.979	1.17	0.241	-0.771	3.069
home duties	-0.807	0.495	-1.63	0.103	-1.78	0.164
other	-1.306	1.640	-0.80	0.426	-4.524	1.911
part time	0.107	0.492	0.22	0.827	-0.859	1.074
retired	-0.563	0.494	-1.14	0.254	-1.532	0.405
student	-0.143	0.535	-0.27	0.789	-1.194	0.907
unemployed	0.275	0.535	0.51	0.607	-0.774	1.325
constant	1.517	0.728	2.08	0.037	0.088	2.945
σ	4.255	0.144			3.971	4.539

Table 5. Complete form. Log likelihood = -2201.6264 Prob > chi2 = 0.0000

Here the correlation $Cov(\widehat{INT}, u) = -0.0033$ and it is not statistically significant at 5 % level.That is,jointly considering the correlation matrix and the residuals correlation, no omitted variable problems seems to arise. Tobit models suggest us that that tax cheating attitude increases:

- the lower is perceived the social mobility;

- the higher is the claim to re-estabilish income inequality ;
- the lower is the family income;
- higher is the unsatisfaction toward the democracy progress;

Proposition 2 As motivated before, meritocracy is a hint of tax morale. Particularly, we promote the idea that higher level of tax cheating are linked to lower level of meritocracy perception, since perceived mobility might be a valid proxy of the perception of welfare quality from a dynamic point of view.

4.1 Consideration on the Estimation Technique

The novelty of the approach we propose is based on the possibility of using non-linear estimation techniques and several instrumental variables useful to predict missing data. Moreover, since the common variables could brought out endogenous problems, we use a pairwise regression taking the residuals, getting ahead 1) simulating data by using all the common information, obtaining a more precise estimate and 2) building up a consistent model. Such method might be seen as a bridge-builder between classical econometric methodologies and machine learning techniques, contributing to the developing idea that their joint application might be a useful tool in uncovering generalizable patterns (see Mullainathan and Spiess, 2017) and make the data analysis affordable. In our research, we find crucial that the imputation of missing data might have some sense regardless of the statistical significance of the relation obtained. To wit, in our case we estimate the perceived social mobility on the basis of a not spurious relation: it is plausible that education, income level and relation between fathers-sons type of occupation and inequality perception might be a valid proxy to estimate the perceived mobility. On the contrary, exploiting the same variable to estimate tax cheating might be quite unrealistic even if the relation might get some statistical sense, since it is a view too personal to be simulated.

5 Policy Implication and Conclusions

Aimed to contribute to the existing theory, we conclude that perceived labor mobility negatively affect the promotion of a welfare system. In our analysis we have to recognize some limits due to the lack of data availability, however, by means of a machine learning approach such as Random Forest, we have been able to impute observation in the most efficient way on the basis of the data at hand. Until now, the state of art explains tax morale in terms of intra-generational perception of unfairness and happiness, as discussed in the Literature Review section. Conversely, there are few works attempting to explain the direction of tax payment from a dynamic point of view, that is, the basic idea is that cheating is *not only* given by the actual perceived unfairness, *but also* it is still persistent on the light of the future perspective. As proposed in the econometric model, the perception of an immobile society strongly discourage subject to contribute, since they can not positive evaluate the efficiency of the welfare state. To conclude, attention should be focused not only on the impact of the tax system as a whole, but it also should be taken into account the expectations of individuals when taxes are paid. The lower will be the return, the lower will be the willingness to pay. Agreeing with recent works we might state that the propensity of paying taxes has to be seen not only as a wealthy optimization problem, but it is also in function of the perceived social status.

Appendix

In the first part of the appendix we report an instrumental variable approach to solve endogeneity. After having simulated data, we use *edu* variable as instrument to estimate the perceived intergenerational mobility. Since it is negatively correlated with the regressor but uncorrelated with the response variable, we invert the order to turn the relation positive. That is, here a positive variation of education has to be read as a reduction of years spent in instruction. We solved correlation between variables as before and ruled out occupation because the estimation algorithm does not converge in that case and *age* and family income since they are correlated with the instrument.

	age	tax cheating	int	$\widehat{familyincome}$	$\widehat{incomeineq}$	$\widehat{edu}$
age	1.0000					
tax cheating	-0.1780	1.0000				
int	0.5794	-0.1485	1.0000			
$\widehat{familyincome}$	0.0035	-0.0699	0.2523	1.0000		
$\widehat{incomeineq}$	0.0000	-0.1071	-0.0321	-0.0776	1.0000	
$\widehat{edu}$	0.5250	-0.0170	0.1853	0.0018	-0.0015	1.0000

Table 1A. Correlation matrix

	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
INT	-4.726	2.678203	-1.76	0.078	-9.975966	.522398
Not very satisfied	0.847	0.383	2.21	0.027	0.096	1.598
Rather satisfied	0.966	0.442	2.18	0.029	0.098	1.833
Very satisfied	1.052	1.037	1.01	0.310	-0.981	3.086
Mixed	-0.280	0.380	-0.74	0.461	-1.026	0.465
Post-materialist	-1.017	0.500	-2.03	0.042	-1.998	-0.035
IT: Isole	-22.208	4189.56	-0.01	0.996	-8233.595	8189.178
IT: Nord	0.482	0.507	0.95	0.341	-0.512	1.47
IT: Sud	1.060	0.453	2.34	0.019	0.172	1.948
$\widehat{incomeineq}$	-0.626	0.150	-4.17	0.000	-0.920	-0.331
constant	13.933	8.388	1.66	0.097	-2.508	30.375
Instrumented	INT					
Instruments	2.demo	3.demo	4.demo	2.materialism	3.materialism	2.area
	3.area	4.area	$\widehat{incomeineq}$	$\widehat{education}$		

Table 2A. Wald test of exogeneity ($\alpha = 0$): chi2(1) = 0.34 Prob > chi2 = 0.5623

Since endogeneity does not persist, the estimation seems to be upwardly biased regarding to the size of the relation.

In the table below it is reported a summary table of the common dataset exploited to estimate the perceived social mobility.

Variable	**Questionnaire**		**Transformed Variable**
	ISSP	**EVS**	
Age	AGE	X003	from 0 to 100
Sex	SEX	X001	Male, Female
family members	HOMPOP	from X002_02B to _06B	from 0 to 100
Education	IT_DEGR	X025	from 0 to 5
Occupation	WRKST	X028	retired,full_time, part_time,home_duties, students,other,unemployed
Marital Status	MARITAL	X007	married, single, separated,widow
Region	IT_REG	X048F	factor levels from 1 to 20
Religion	RELIG	F025	catholic,islam, others
Union	UNION	A067	0,1
Income Equality	V33	E033	1,2,3,4
Sexual Discrimination	V16	C001	0,1,2
father occ	V56	V005	employed,self-employed,unemployed
hard work	V75	A005	0,1,2,3
Election	IT_PRTY	E179	left wing, right wing, other
Family income	IT_INC	X047A	from 1 to 8
Perception of unemployment	V34	E160	from 1 to 5
Tax Cheating	NA	F116	from 1 (never justified) to 10 (justified)
Perceived mobility	V46	NA	from 1 (worsening) to 5 (improvement)

Table 3A.Common Variables Description.

References

[1] Alesina A. and Rodrik D. , *"Distributive Politics and Economic Growth"*, The Quarterly Journal of Economics ,Vol. 109, No. 2 , pp. 465-490, May 1994;

[2] Allingham M.G. and Sandmo A., *"Income tax evasion: a theoretical analysis"*, Journal of Public Economics, Vol. 1 Nos 3-4, pp. 323-338, 1972;

[3] Amendola A., Dell'Anno R. and Parisi L., *"Happiness and inequality in European countries: is it a matter of peer group comparisons?"*, Economia Politica, 2018;

[4] Barone G. and Mocetti S. , *"Tax morale and public spending inefficiency"*, International Tax and Public Finance, Vol. 18 No. 6, pp. 724-749.,2011;

[5] Becker G. S. and Tomes N., *Human Capital and the Rise and Fall of Families*, Journal of Labor Economics, University of Chicago Press, vol. 4(3), pages 1-39, July 1986;

[6] Belmonte A., Dell'Anno R. and Teobaldelli D., *"Tax morale, aversion to ethnic diversity, and decentralization"*, European Journal of Political Economy, 2018;

[7] Bénabou R., *Human capital, inequality, and growth: A local perspective*, European Economic Review, Volume 38, Issues 3–4, Pages 817-826, April 1994;

[8] Checchi D.,*Inequality in Incomes and Access to Education: A Cross-Country Analysis* (1960-95), Labour, Vol. 17, pp. 153-201, June 2003;

[9] Choi A., Jerrim J. and Simancas R., Two-Sample Two-Stage Least Squares (TSTSLS) estimates of earnings mobility: how consistent are they?, Survey Research Methods, Vol. 10, No. 2, pp. 85-102, 2016;

[10] Dell'Anno R., *"Tax evasion, tax morale and policy maker's effectiveness"*, The Journal of Socio-Economics, Vol. 38 No. 6, pp. 988-997,2009;

[11] Doerrenberg P. and Peichl A. , *"Progressive taxation and tax morale"*, IZA Discussion Paper,No. 5378, The Institute for the Study of Labor (IZA), Bonn, December, 2010;

[12] European Values Study, EVS 1981 - 2008 Variable Report, Longitudinal data Files;

[13] Franzini M. and Pianta M., *Explaining Inequality*,Routledge, 2015;

[14] Hastie T., Tibshirani R. and Friedman J.,*The Elements of Statistical Learning Data Mining, Inference, and Prediction*, Springer, 2017;

[15] Horodnic, I.A., *Tax morale and institutional theory: a systematic review.* International Journal of Sociology and Social Policy, 38 (9/10). pp. 868-886. ISSN 0144-333X, 2018;

[16] International Social Survey Programme,ISSP 2009 - Social Inequality IV;

[17] Luttmer E.F.P. and Singhal M. , *"Tax morale"*, Journal of Economic Perspectives, Vol. 28 No. 4,pp. 149-168, 2014;

[18] Mullainathan S. and Spiess J., *"Machine Learning: An Applied Econometric Approach"*, Journal of Economic Perspectives, American Economic Association, 2017;

[19] Mulligan C., *Some Evidence on the Role of Imperfect Capital Markets for the Transmission of Inequality*, Mimeo, University of Chicago, 1995;

[20] Pellizzari M., *Employers' Search and the Efficiency of Matching*, British Journal of Industrial Relations, Vol. 49, No. 1, pp. 25-53, 2011;

[21] Piketty T.,Theories of Persistent Inequality and Intergenerational Mobility, CNRS-CEPREMAP, 2000;

[22] R Core Team, *R: A language and environment for statistical computing*, R Foundation for Statistical Computing, Vienna, Austria, 2013;

[23] Roderick and J. A.Little Donald and B. Rubin,*Statistical Analysis with Missing Data*, Wiley Series in Probability and Statistics, 1987;

[24] Torgler B. Schneider F. and Schaltegger C.A. , *"With or against the people? The impact of a bottom-up approach on tax morale and the shadow economy"*, Working/Discussion Paper No. 211, QUT – School of Economics and Finance, Brisbane, February 2007b;

[25] Torgler B. Schneider F. and Schaltegger C.A., *"Local autonomy, tax morale, and the shadow economy"*, Public Choice, Vol. 144 Nos 1-2, pp. 293-321,2010;

[26] Torgler B. , *"Cross-culture comparison of tax morale and tax compliance: evidence from Costa Rica and Switzerland"*, International Journal of Comparative Sociology, Vol. 45 Nos 1-2, pp. 17-43, 2004a;

[27] Torgler B. , *"Tax morale in Asian countries"*, Journal of Asian Economics, Vol. 15, pp. 237-266, 2004b;

[28] Torgler B. , *"Tax morale, trust and corruption: empirical evidence from transition countries"*, Working Paper No. 2004-05, Center for Research in Economics, Management and the Arts, Basel, 2004c;

[29] Trüdinger E.M. and Hildebrandt A., *"Causes and contexts of tax morale: rational considerations, community orientations, and communist rule"*, International Political Science Review, Vol. 34 No. 2, pp. 191-209, 2013;

[30] Vythelingum P., Soondram H. and Jugurnath B. , *"An assessment of tax morale among Mauritian taxpayers"*, Journal of Accounting and Taxation, Vol. 9 No. 1, pp. 1-10, 2017;

[31] Williams C.C. and Krasniqi B., *"Evaluating the individual- and country-level variations in tax morale: evidence from 35 Eurasian countries"*, Journal of Economic Studies, Vol. 44 No. 5, pp. 816-832, 2017;

[32] Williams C.C. and Martinez Á. , *"Explaining cross-national variations in tax morality in the European Union: an exploratory analysis"*, Studies of Transition States and Societies, Vol. 6 No. 1, pp. 5-18, 2014.