Inside the Black Box of Mobility: Understanding Upward Mobility in Two British Cohorts

Andrew Eyles
University College London

Abstract

I use tools from the machine learning explainability literature, combined with a latent factor framework, to highlight how patterns of upward mobility have changed across two British birth cohorts. The methodology allows for arbitrary patterns of non-linearity, guards against overfitting, is invariant to monotonic transformations of the factors studied, and, unlike previous studies, is not sensitive to the order in which predictors enter the model. I find that origin and gender predict mobility. Those born in the bottom quintile are most likely to be upwardly mobile due to having more 'room to move up'. Females are much less likely than their male counterparts to be upwardly mobile. Years of schooling and cognitive ability are also important. These act as sufficient statistics for a host of early childhood factors; namely, household composition, socioemotional skills, parental time investments, and school quality. Despite the relationship between parental and child income strengthening between the birth cohorts, the determinants of upward mobility are stable.

1. Introduction

In this paper, I ask whether patterns of social mobility have changed over time. Making inferences about trends in mobility is difficult due to a multiplicity of measures that can be used.[1] In the UK, debates about whether social mobility has changed over time centre around the metric used to measure mobility. Measures of occupational change show stasis (Bukodi and Goldthorpe, 2018) whereas income-based measures, such as rank correlations between parental income and that of their offspring, suggest a decline (Blanden et al., 2005).

Even if the aforementioned statistics are stable, it may still be the case that the process linking later life outcomes to one's origin is changing. For instance, the

[1]Simply by varying how, and for whom, income mobility is measured, Engzell and Mood (2023) estimate 82,944 specifications relating parental and child income using Swedish registry data.

drivers of upward mobility—be they educational attainment, household demographics, or early childhood investments—may vary in importance across cohorts. In this paper, I use tools from the machine learning explainability literature to assess whether drivers of mobility have changed over time.

To do so, I use data from two British cohort studies—the 1958 National Child Development Study (NCDS) and the 1970 British Cohort Study (BCS). Each contains a wealth of measurements on conditions from birth until adulthood. In accordance with the literature on human capital formation, I combine these measurements using a latent factor framework to produce measures of cognitive skills, non-cognitive skills, and parental investments at various stages of development. I supplement these with variables capturing family structure and completed years of schooling. These 'features' are then used as inputs in forest-based predictive models that aim to predict upward mobility.[2] I use tools from coalitional game theory to measure how each 'feature' contributes to the predictions made by the models. Examining how features contribute differentially to mobility, across the two cohorts, is the contribution of this paper.

This paper adds to the literature that aims to understand the mechanisms linking parental occupation or income to the occupation and income of their offspring.[3] It also adds to the growing literature that uses machine learning methods in applied economics research (Mullainathan and Spiess, 2017; Athey, 2019).

The papers closest in spirit to this one are Blanden, Gregg, and Macmillan (2007) who look at the factors that mediate the intergenerational earnings elasticity (IGE) for NCDS and BCS cohort members. Similarly, Bolt et al. (2021) combine multilevel mediation analysis with a latent factor framework to decompose the fraction of the IGE explained by a number of early life factors. Methodologically, Blundell and Risa (2018) make use of machine learning methods and argue that the outputs from machine learning models can be used as measures of mobility.[4]

This paper differs from these in both conceptual terms—I aim to assess the relative importance of different contributors to upward mobility rather than explain which factors mediate the IGE—and methodologically.[5] On the latter, I argue that machine learning (ML) methods are particularly appropriate for the problem at hand. Forest-based methods allow for inputs to interact with each other in a man-

[2]In following the ML literature, I will refer to model inputs as features throughout.

[3]Blanden, Doepke, and Stuhler (2022) provide a detailed overview of the link between educational inequality and social mobility. Heckman and Mosso (2014) survey the links between early life skills and parental investments and mobility.

[4]Blundell and Risa (2018) employ many of the same models that I use to predict income for BCS and NCDS cohort members based upon parental income and family characteristics. Their focus is on whether the fraction of variance explained by the models differs across the cohorts.

[5]The final section does include a mediation analysis to help interpretation of the ML results.

ner that linear regression does not. A wealth of evidence highlights the importance of interactions between stocks of skills and investments at different ages (Cunha and Heckman, 2008; Cunha et al., 2010).

Understanding feature importance necessitates a predictive model that allows for non-linearities such as these. The methods used rely on cross validation as a means of model selection. While not a panacea for the small-sample sizes that are inherent with cohort studies, this form of model selection is less prone to overfitting than methods which simply use linear methods on all the available data.

Despite these advantages, machine-learning methods are often opaque and difficult to interpret. Models that provide the best fit are often 'ensemble' methods that fit multiple models and aggregate the predictions (Hastie, Tibshirani, and Friedman, 2009). As such, these methods have no counterpart to a regression coefficient by which the effect of a unit change of some variable can be mapped to a change in the outcome.[6] For this paper—where the primary aim is to understand how the importance of different features changes over time—this presents a problem. To overcome this, I use Shapley additive explanations (SHAP) as introduced in Lundberg and Lee (2017). SHAP is inspired by the use of Shapley values in game theory that measure the contribution of a player to a coalition as the average of all the marginal contributions a player makes to all possible coalitions of players. By replacing 'players' with ML model features, relative contributions of different features to model predictions can be compared in an intuitive manner.

Despite patterns of mobility being different across the two cohorts, the determinants of upward mobility are shown to be surprisingly stable. Being born in the bottom 20% of the income distribution makes upward mobility 'easier'. Intuitively, those at the bottom have more space to move up while those born into the middle of the distribution can only achieve upward mobility by becoming a top 20% earner. For the 1958 and 1970 birth cohorts, the wage premium given to males makes upward mobility less likely for female cohort members.

Cognition and years of schooling are strong predictors of upward mobility in both cohorts. Once one knows a cohort member's cognitive abilities at adolescence, information on family background has little predictive power over upward mobility outcomes. Multi-step mediation analysis highlights that school quality, parental time investments, and family background variables operate indirectly by increasing early stage cognition and completed years of schooling.

The paper proceeds as follows. Section 2 describes the data used. Section 3

[6] As will be explained in the methodology section, the SHAP approach does not suffer from the problem that different scales of the variables can lead to different conclusions regarding the relative importance of the variable.

discusses the measurement framework. Section 4 gives a brief overview of the ML algorithms and SHAP. Section 5 presents the results. Section 6 concludes.

2. Data

I use data from two British cohort studies—the 1958 National Child Development Study (NCDS) and the 1970 British Cohort Study (BCS). These two studies are ideal for the purposes of this study and have been used to examine intergenerational mobility in income (Dearden et al., 1997; Blanden et al., 2004; Gregg et al., 2017), social class (Erikson and Goldthorpe, 2010), and home ownership (Blanden et al., 2023). They track two birth cohorts through their childhood, adolescence, and their working life. Importantly, the studies are structured similarly, and a number of variables are harmonised across the studies.

For the purposes of this study, I use parental income observations at age 16 and cohort members' income at age 42. Upward mobility is taken to mean residing in a higher income bracket—measured by income quintiles—than one's parents. Someone born into the bottom 20% of the income distribution but whose earnings place them above the 20th percentile is upwardly mobile. By construction, those born to parents in the top 20% of the distribution cannot be upwardly mobile. These individuals, unless noted, are removed from what follows.

Tables 1–3 present some well-known facts about income mobility from the cohort studies. A linear regression of log parental income on income at age 42 highlights that between 2000 and 2012 (when age 42 income is measured), the intergenerational persistence of income went from 0.274 to 0.354. This strengthening of the relationship between parental income and one's own income holds when income earlier in the life cycle (age 33/34) is considered. In principle, this could represent a widening of the income distribution as well as changes in the dependency between origin and destination income. The bottom row of Table 1 follows Chetty et al. (2014) and estimates the rank slope between parental and child income. A rise in rank-slopes is evident irrespective of whether age 33/34 or age 42 income is used as an outcome. One's position in the income distribution relates more strongly to parental income in the BCS than the NCDS.

Tables 2 and 3 look at transition probabilities between income quintiles where the origin is the quintile one was born into and the destination is income quintile at age 42. There are several changes between the cohorts. Income persistence at the top strengthens between the NCDS and BCS. Those born into the top 20% of household incomes have a 30% likelihood of remaining there in 2000 and a 36.8% chance in 2012. When looking at the likelihood of upward mobility, this translates into a fall in mobility prospects for those born into the second highest income quintile. These individuals must move into the top quintile to be upwardly mobile, but only 21.79% do so in the BCS as compared to 24% in the NCDS.

In general, even where upward mobility is unchanged, the form it takes differs across cohorts. 54% of those born into the 2nd income quintile achieve upward mobility in both cohorts, but those born in the BCS are more likely to do so by moving to a nearby income quintile, i.e. their prospects of reaching the top 20% have declined over time (going from 14.42 to 11.81). Those born into the lowest income quintile have declining upward mobility prospects between the cohorts and, even for those that are upwardly mobile, the prospects of reaching the top two quintiles falls from 30.71 to 26.92.

In total, persistence in income increases at the top and, for those that are upwardly mobile, they are more likely to move to neighbouring income quintiles than make extreme moves. These patterns mean that the IGE and rank-rank slope in income increase while upward mobility remains stable.

In this paper, my focus is on the drivers of upward mobility. While patterns of mobility are well understood, the determinants of upward mobility are less well studied. Several studies have highlighted factors that may shape mobility outcomes. Kearney (2023) shows how family composition—particularly the presence of two married parents—advantages some children over others. Chetty et al. (2014) discusses how family structure can partially explain cross sectional variation in mobility rates across US commuting zones. Heckman and Mosso (2014) discuss how family background and parental investments combine to produce stocks of cognitive and socio-emotional skills that influence schooling choices and outcomes.

As such, I collect measurements in the cohorts on family composition, cognition, parental investments, and socioemotional skills. For cognition, I use the plethora of tests discussed in Moulton et al. (2020). These tests are administered at ages 7, 11, and 16 in the NCDS and ages 5, 10, and 16 in the BCS. While the tests are not fully comparable, they do aim to capture the same underlying constructs such as mathematical ability, verbal reasoning, and vocabulary. For socio-emotional skills, I follow Attanasio et al. (2020) and use identical Rutter scale items that are reported by cohort member's mothers. For investments, I consider variables that capture parental time investment (whether one reads to the child, number of weekly outings, interest taken in child's education) and school level investments (pupil/teacher ratio, the presence of a parent teacher association in the school). These are chosen to match investment variables used elsewhere in the literature on human capital formation (see Agostinelli and Wiswall, 2024 for a recent example). In each case, the measures are merely proxies of some underlying construct.

In the next section, I discuss how I combine these using a latent factor framework to produce interpretable indices.[7][8]

For family background, I choose a set of variables that capture conditions at birth and throughout childhood. Tables 4 and 5 show averages across three distinct groups—those who are upwardly mobile, those who are not, and those born into the top 20% of the income distribution. The latter serve as a reference group as, by definition, they cannot achieve upward mobility. Alongside these, residualised differences between those who do and do not achieve upward mobility are presented. The differences condition out the income quintile group means due to the mechanical effect of one's origin quintile on the likelihood of upward mobility.

While most of the differences are significantly different, a few stand out numerically. Firstly, most of the variables that capture household demographics are small in the NCDS sample. Having a teen mother, an absent father, and parents who are not married at birth have limited effect on upward mobility. This is not entirely surprising given the limited variability in these variables in the sample. Few cohort members are born to teen mothers and even fewer are born into single parent families. Family structure appears uniform across the income distribution. Having a parent attend university is more prevalent amongst the upwardly mobile as is home ownership. The majority of those in the top 20% have parents who own a home and almost a quarter have a parent with a degree. Amongst the bottom 80%, the upwardly mobile, while not matching those in the top 20%, are edging towards them in terms of home ownership rates and parental education. Upwardly mobile individuals, conditional on origin quintile, are 3.3 percentage points more likely to have a university educated parent and are 8.2 percentage points more likely to grow up in owner occupied housing.

Numerically, the results from the BCS are somewhat similar. The differences between the upwardly mobile and not upwardly mobile in terms of parental education dwarfs differences in other family background characteristics such as parental divorce and the absence of a father figure. This is in spite of the fact that parental divorce increases sharply between the cohorts. In the NCDS, 8.7% of parents divorce and this rises to 17.1% in the 1970 cohort. Many more cohort members are born into owner occupied housing in the latter cohort but the difference between

[7] Appendix Tables A2.1–A2.14 give a full breakdown of the confirmatory factor analysis underlying the indices.

[8] The structure of the latent factor framework is also useful when imputing missing data. For many observations, full data are not available but measures at earlier and/or later ages are alongside information on family background. I use Multiple Imputation by Chained Equations (MICE) to impute missing data for the sample of individuals who have age 42 income data and parental income at age 16 data.

those who are upwardly mobile and those who are not remains similar in magnitude (the residualised difference in the BCS is 6.5 percentage points). One of the largest changes is in parental education. The proportion of parents with degrees doubles across the cohorts across the income distribution. 54% of those born into the top quintile have a parent with a degree in the BCS (as compared with 23% in the NCDS), while 16% of those in the bottom 4 quintiles do (as compared with 6% in the NCDS). The gap in parental education between the mobile and the not widens between the cohorts. Upwardly mobile individuals are 8.8 percentage points more likely to have a degree educated parent than those who are not upwardly mobile. The corresponding difference in the NCDS is 3.3 percentage points.

In both cohorts, gender plays a large role in mobility prospects. Males are more likely to be upwardly mobile than their female counterparts reflecting a sizable gender pay gap in the cohort data. While all the differences in Tables 4 and 5 are 'origin' variables, I do include one's own education as well. In each cohort, upwardly mobile sample members are much more likely to be degree educated than their non-mobile peers. The difference in the fraction of individuals having degrees doubles between the cohorts. In the NCDS, upwardly mobile individuals are 18.2 percentage points more likely to attend university than non-mobile individuals. This gap rises to 32 percentage points in the BCS.

3. Latent Variable Measurement

While family background can be aptly summarised as above, many of the predictors that I consider are less well defined. The cohort studies collect many different measurements that proxy underlying latent variables of interest. I follow the approach taken by Heckman et al. (2013) and reduce the plethora of proxy variables down into low dimensional indices.[9]

$$M_{\omega,i,t,j} \;=\; \mu_{\omega,t,j} + \lambda_{\omega,t,j}\,\omega_{i,t} + \varepsilon_{\omega,i,t,j} \tag{1}$$

Here M denotes measure, ω denotes the underlying construct (time investment, cognition, socio-emotional skill, and school quality), t denotes age, and j indexes the measure. Error terms are assumed to be uncorrelated across individuals, measures, and age.

As the latent variables have no natural scale, normalizations are needed to

[9] Appendix Tables A2.1–A2.14 give the results of the exploratory factor analysis. I first conduct an EFA to determine the underlying structure of the data i.e. how many constructs are needed to describe covariation in the data. I then estimate a series of dedicated factor models where each measure loads onto a single factor.

identify the model. For the continuous measures that I use in the cognition equation, I set the constant to zero and the intercept to 1 for one of the test score loadings.[10] For the equations governing investments, socio-emotional skills, and school quality, the variables have an ordinal scale. Here I take the approach discussed in Muthén (1984) that uses polychoric correlation matrices to derive factor loadings. In doing so, I normalise the model by fixing the variance of the latent factors to 1.

Exploratory factor analysis suggests that for the NCDS there is a single cognitive construct at each age (7, 11, and 16), a single socio-emotional construct at age 7, two socio-emotional constructs at ages 11 and 16 (that I label internalising and externalising skills), a parental time investment factor at all ages, and a school quality factor at age 7. The BCS yields a similar structure with a cognitive factor, two socio-emotional factors, and a parental time investment factor at each age (5, 10, and 16), alongside a school quality factor at age 10.[11]

Factor scores, which later serve as inputs to the ML algorithm described in the next section, are plotted in Figures 2–8. As the scores alone are difficult to interpret, I plot the distribution separately for upwardly mobile, not upwardly mobile, and those born into the top quintile. I also report rank differences between both groups—those who are mobile and those who are not—and those born into the top 20%.

Figures 2 and 3 show the distribution of scores for the socio-emotional skill measures. Several studies highlight the role that non-cognitive skills play in shaping outcomes and the increasing importance of non-cognitive skills in determining wages (Heckman et al., 2006; Deming, 2017). Interestingly, there appears to be less of a role played by these skills in enhancing mobility prospects. The distributions of both externalising and internalising skills are similar for those who are and who are not upwardly mobile in both cohorts. Both groups have lower scores, on average, than those born in the top quintile. In line with previous studies that document an increase in inequality in socio-emotional skills (Attanasio et al., 2020), the gap between those born outside of the top 20% of the income distribution and the rest widens between the NCDS and the BCS. This is particularly the case for externalising skills where the difference in mean rank between those who are not upwardly mobile and the top 20% goes from 3.94 to 9.78. The differences between the top 20% and those who achieve mobility exhibits a similar rise from 2.30 to

[10] The cognition equation uses standardised test scores as the measures.

[11] The existence of a single school quality factor alongside multiple parental time investment factors (one for each age) is due to data collection. Only at ages 7 and 10 are detailed questions given to teachers and heads about the school. Conversely, parents are asked multiple questions about their engagement with their children in all childhood/adolescent waves.

10.20. Internalising skills exhibit a somewhat weaker relationship with socioeconomic background.

A different picture emerges when cognition at age 16 is looked at (Figures 4 and 5). For both cohorts, the score distribution of the upwardly mobile is shifted away from those who remain in the quintile into which they were born. For the NCDS, the cognitive score distribution of mobile individuals looks more like the distribution amongst those born into the top quintile. A regression of ranked factor scores on group dummies (with top quintile as the omitted category) gives coefficients of -18.02 and -4.60 for those who are not and who are upwardly mobile in the NCDS cohort. For the BCS, the same coefficients are -23.37 and -11.91. While the top quintile pulls further away from the bottom 80% in the BCS, inequality in cognitive achievement between those who move up the income distribution and those who do not remains stable across the cohort.

Finally, I plot scores for two well-known inputs into cognition—school quality and time investments. For both these investments, and across both cohorts, there is a clear divide between the bottom 80% of the parental income distribution and the top 20%. Those in the top quintile have a score distribution that lies clearly to the right of the rest. Looking within the bottom 80% and splitting by mobility status shows little difference between those who are upwardly mobile and those who are not. For parental time investments, the mean rank of scores is 9.02 lower for those who are not mobile and 7.98 lower for those who are when compared with those born into the top quintile in the NCDS. These gaps widened further to 17.31 and 17.96 in the BCS cohort. With a notable exception, a similar pattern follows with the school quality factor scores. Here the gap between those born in the top quintile and those who are not/are upwardly mobile respectively are -15.44 and -11.34 in the NCDS. These gaps widened in the BCS cohort to -22.97 and -22.41. In the NCDS cohort, the rank difference is significant here at 4.10. The score distribution is bimodal for all three categories with a greater density around the second (higher school quality) maxima for those who are upwardly mobile.

Figures 7 and 8 consolidate the evidence above as follows. I run a logistic regression of an upward mobility dummy on each individual's cognitive score, their socio-emotional scores, their school quality score, and their parental investment score. I use the fitted model to vary each factor from the 5th to 95th percentile holding other factors at their mean values. The resultant curves trace out how the predicted probability of upward mobility changes with each factor. The curves highlight that raising cognition has an outsized impact on mobility relative to socioemotional skills. Varying cognition from the 5th to 95th percentile raises mobility prospects from 7% to 82% in the NCDS and 8% to 68% in the BCS. School quality raises mobility prospects more than parental investments of time, but again the change is more pronounced in the NCDS than the BCS. In the NCDS the impact is a 20 percentage points rise (going from just under 30% to 50%) while the BCS rise is 12 percentage points (36% to 48%).

Taken together, the evidence shows that on several determinants of wages—cognition and externalising skills—those born into the top quintile of household income have pulled away from their counterparts between the two cohorts. Similar holds true when one considers factors that serve as inputs into cognition and socio-emotional skills—school quality and parental time investments.[12] Amongst those that are born into the bottom 80% of the income distribution, individuals who are and who are not upwardly mobile are more comparable to each other than they are to those in the top quintile. Cognition, and to a lesser extent school quality, are exceptions to this suggesting that these variables are important determinants of upward mobility.

4. Methodology

4.1 XGBoost

The descriptive evidence above suggests that certain factors play a more crucial role in driving mobility than others; however, it is desirable to combine them in a holistic way when predicting the likelihood of upward mobility. To do so, I use an algorithm called extreme gradient boosting (XGBoost).

There are several reasons why this method is preferred to, for instance, running a linear regression. Firstly, the fitted model allows for arbitrary patterns of nonlinearity. Features such as socioemotional skills are allowed to interact with other features to predict outcomes. Secondly, although not specific to this method, the algorithm has a number of 'tuning' parameters that can be chosen to prevent the model from overfitting. This is particularly important given the data that is used to estimate intergenerational mobility. Cohort data typically has few observations relative to the number of features included. Not explicitly cross validating the models can lead to overfitting.

For this paper, I aim to explain what drives upward mobility. Any kind of model explainability will ultimately explain model fit. Understanding models that are less sensitive to outliers and generalise well to other samples is of more interest than understanding models that perform poorly when new data is introduced. Finally, unlike parametric models, the fit of XGBoost is invariant to monotonic transformations of the features used. This is useful given that many of the features are latent constructs that have no natural scale.

[12]Later I will formally link these inputs in a production function framework.

The algorithm fits a series of decision trees before averaging their predictions. Each fitted tree is a weak learner that performs marginally better at classifying the observation than a guess. Intuitively, weak learners have low variance in that they perform similarly on different datasets. Each tree is fit sequentially on the errors of the previous tree so that each additional tree is able to predict where the previous one performed poorly. Formally:

$$
\begin{aligned}
Y &= F_0(x) + e_0 \\
\hat{Y} &= \hat{F}_0(x) \\
e_0 &= F_1(x) + e_1 \\
\hat{e}_0 &= \hat{F}_1(x)
\end{aligned}
$$

Here $F_n(x)$ is a prediction from the nth tree-based model.[13] A final prediction is then made by summing the initial prediction along with the predicted errors:

$$
\text{Prediction} = \hat{Y} + \hat{e}_0 + \hat{e}_1 + \cdots \tag{2}
$$

In practice there are a number of decisions to be made about how the underlying trees are built. I vary the maximum depth of the underlying trees i.e. how deep the $F(x)$ predictions can go, the number of boosting rounds (the size of n), and the learning rate. The latter governs the weight that each additional tree adds to the final prediction. To select these parameters, I perform 5-fold cross validation and a grid search over the parameter space before selecting the parameters that provide the highest predictive accuracy.

4.2 Shapley Additive Explanations (SHAP)

While the above algorithm provides a convenient way to predict mobility, model predictions are often opaque. Tree based models, and models that average over tree-based predictions, give no counterpart to a regression coefficient that explains how a unit change in a feature changes the probability of the outcome studied.

To understand the output of the model I use a tool from coalitional game theory—Shapley values—as suggested in Lundberg and Lee (2017). Each feature used in the model is assigned a value that is that feature's marginal contribution across all possible combinations of features used in the model. Formally,

[13] Tree splits are selected so as to minimise the log loss criterion in each node. Log loss penalises predictions that are confidently wrong. In this case, instances where upward mobility has a high probability, but observations remain in the class into which they were born.

$$\phi_i(f) = \sum_{S \subseteq N \setminus \{i\}} \frac{|S|(|N| - |S| - 1)!}{|N|!} [f(S \cup \{i\}) - f(S)] \tag{3}$$

Here, the contribution of feature i depends upon the difference in prediction between a model with and without that feature, $f(S \cup \{i\}) - f(S)$, for a given subset of features S that is weighted and summed across all feature combinations.

Shapley values have several properties that make them well suited to explaining model predictions. Firstly, for a categorisation problem such as predicting upward mobility, the log odds of the model's predicted probability can be expressed as a sum of the baseline prediction (the sample average probability) and the Shapley value for each included feature. The sum of Shapley values for the features thus account for the deviation of the model's prediction from the sample average of the outcome. In other words, the values show how to distribute the differences between individual predictions and the sample average among all features within the model.

Secondly, unlike previous analysis that use regression-based methods to assess feature importance in similar contexts (Blanden et al., 2007), SHAP is, by construction, invariant to the order in which variables enter the model. SHAP is also invariant to any variable transformations that leave the model fit unchanged. The underlying model that I explain is invariant to monotonic transformations of the model features and computed Shapley values will inherit this characteristic.

5. Results

5.1 Model Fit

Table 6 presents several fit measures for the model. As a point of comparison, fit measures are also presented for a naive model—one that predicts each observation will fall in the most common mobility class (no mobility in both samples)—and a linear probability model (LPM) that predicts upward mobility for those with fitted probabilities above 0.5. The fit metrics reported are the fraction of correct predictions (accuracy), the ratio of correct positive predictions to total positive predictions (precision), and a measure of the model's ability to distinguish between positive and negative outcomes (AUC).

The results highlight that the features considered have a significant predictive power for upward mobility. Both the LPM and XGBoost perform well in terms of overall accuracy. Interestingly, there is little difference in predictive accuracy between a linear model and the more flexible boosting model in terms when predicting the main upward mobility measure used in this paper—whether one resides in a higher income quintile than the quintile one was born in. This suggests a limited role for interactions between the features considered.[14]

In addition to the simple mobility definition, I also consider a second definition model that predicts the likelihood of moving from either of the bottom two quintiles to the top two. This 'rags to riches' style of mobility is rare in both the cohorts. 8%/9% of those in the bottom quintiles reach the top as compared with 47%/45% who are upwardly mobile according to the first definition. In this case, there is a large performance discrepancy between the tuned ML model and the LPM. While the LPM is accurate in this case, its accuracy is driven by the naive prediction that no one will achieve upward mobility. In absence of parameter tuning, the ML model also predicts no upward mobility for either the NCDS or the BCS. The tuned model on the other hand does predict upward mobility for some observations and does so correctly in 59% (NCDS) and 86% (BCS).

5.2 Feature Importance

The tuned ML model does well at correctly classifying individuals as mobile vs not irrespective of the definition of mobility. In this section, I explain the model predictions using Shapley values.

Figures 9 and 10 plot the mean absolute SHAP value of each feature over all observations in the dataset. The features are ordered by their mean absolute value and features with limited impact on model predictions are pooled at the bottom. The results are presented on a log odds scale.

A consistent picture emerges across both the cohorts. Despite evidence of different transition probabilities between the 1958 and 1970 birth cohorts, the determinants of upward mobility are stable. In each case, one's gender and origin are of crucial importance. This is due to a sizable wage penalty for females in the cohorts and greater ease of upward mobility for those in the bottom quintiles. Upward mobility for those born in the 'middle' means breaking into the top income quintile in adulthood whereas mobility can more easily be achieved by those who can, for instance, move from the bottom quintile into the middle of the income distribution.

Aside from these effects, education, and cognition (particularly at ages 10/11 and 16) play a large role. In the NCDS the total effect of cognition variables is 0.39 while the effect of schooling variables is 0.19.[15] For the BCS, the total effect of cognition related variables is 0.22 whereas for education variables it is 0.43. Completed

[14]The result above is likely to be driven by data processing done prior to fitting the models. Reducing the large amount of cognitive, socio-emotional, and investment proxies into indices guards against overfitting at the outset and minimises the role played by hyperparameter tuning. This, coupled with weak interaction effects, drives the results.

[15]These are obtained by summing the SHAP values on the education dummies used in the model.

education has higher predictive power in the latter cohort than the former where cognition has more predictive power, but taken together gender, origin, cognition, and schooling are the main drivers of mobility in both cohorts.

Figures 11 and 12 give a more complete picture of how each feature contributes to model fit. The previous figures plot absolute values whereas Figures 11 and 12 colour the values according to whether they are positive or negative. As the model allows for arbitrary interactions it may be the case that high levels of a feature lead to high predicted probabilities for some individuals and lower probabilities for others. However, looking at the plots the features that are of most importance are uniform in sign. Higher levels of cognition have uniformly positive SHAP values for both the NCDS and BCS whereas being female pushes down the likelihood of upward mobility. Similarly, being born into lower income quintiles makes it easier to achieve mobility (at least according to the definition in this paper) as does going on to have more years of schooling. For the latter, educational levels of NVQ level 3 and above boost mobility prospects whereas NVQ2 and below decrease them. This result is consistent across both cohorts.

While education and cognition play a large role in upward mobility, they are somewhat secondary to the mechanical role played by origin (and to a lesser extent gender). Figures 13 and 14 assess whether the role of cognition varies by gender and origin. In each case, Shapley values for cognition are plotted separately for those born into the bottom 20% and for females. As the figures show, the effect of cognition is uniform across quintiles and genders. Higher measured cognition is associated with higher Shapley irrespective of gender and income origin and the relationship between the level of cognition and the extent to which the model predicts upward mobility is constant across origin/gender.

5.3 Interpretation

The results above suggest that cognition and schooling play a primary role in shaping upward mobility. Several factors—particularly family background measures, school quality, and time investments—have minimal predictive value. There are a few potential explanations for this; firstly, these factors could have no explanatory power; or secondly, their effects can be fully captured by schooling and cognition i.e. educational outcomes act as a sufficient statistic for the effect of family composition and investments.

To better understand how investment and background contribute to mobility outcomes, I follow Bolt et al. (2021) and run a multi-level mediation analysis to understand the relationship between parental income and one's own income for both birth cohorts. In the first stage, I allow only for direct mediation. In the second, I allow inputs such as time investment and school quality, to mediate the

IGE both directly and indirectly. Formally, in the first stage I estimate:

$$\ln Y_i = \alpha_S S_i + \alpha_C C_i + \alpha_{SE} SE_i + \alpha_I I_i + \alpha_F F_i + \alpha_{Y_p} \ln Y_{\text{Parent},i} + u_i^Y \tag{4}$$

Where own income is regressed onto years of schooling, the cognitive measures, socio-emotional measures, investment measures, family background characteristics, and parental income.[16] For the simple mediation, I then relate each measure to parental income. If ρ denotes the IGE and κ_j denotes the coefficient from a regression of measure j on log parental income, then the share of the IGE explained by measure j can be written as $\alpha_j \kappa_j / \rho$.

A more complex model allows for measures to have indirect as well as direct effects. Here, I allow for a richer set of dynamics as illustrated in the path diagram in Figure 1.

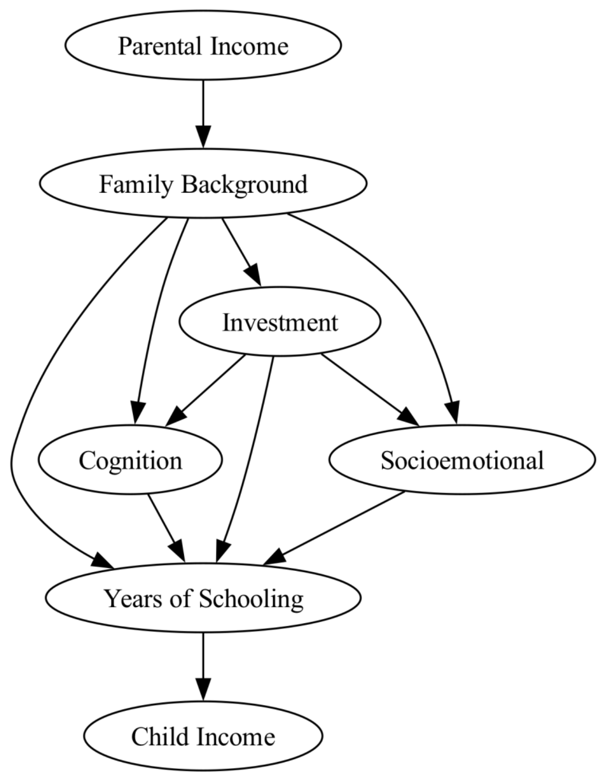

Figure 1: Path diagram for the complex mediation model.

Factors such as investment operate through multiple channels. They have a direct effect on income. They also have an indirect effect through years of schooling,

[16]In practice, the multilevel mediation is made simpler by only including final levels of cognition and socio-emotional skills. As I aggregate these across ages, and because early skills operate through increasing later skills, the results are unchanged by their exclusion.

cognition, and socioemotional skills. The indirect effects themselves can be indirect with investments raising cognition, which then raises schooling, which itself shifts child income.

Table 7 shows the fraction of the IGE explained by each of the factors. Columns (1) and (3) show the direct effects estimated from equation (4). In line with the XGBoost results, both schooling and cognition play a large role in explaining the association between child and parent income. 75% of the IGE in the NCDS can be explained by these two factors alone, although this falls to 49% in the BCS. Interestingly, there is a larger role played by school quality and family background in the BCS. Each contributed equally with a total contribution of 22% against a statistically insignificant 9% in the BCS. This is not in direct contradiction with the findings presented earlier. The ML algorithm predicts upward mobility while the mediation analysis highlights factors that strengthen the persistence between parental and child income. The relatively larger role played by school quality and family background in the latter analysis highlight that while these variables are not strongly predictive of upward mobility, they do strengthen income persistence at the top. This is highlighted in the earlier discussion where school quality—particularly in the BCS—differs between those born into the top 20% and the bottom 80% but is very similar between those who are and who are not upwardly mobile. A similar point can be made about several family background variables in Tables 4 and 5.

Columns (2) and (4) allow for more indirect channels between measures. While the direct effect of schooling and cognition remain significant across both cohorts (a combined effect of 22% in the NCDS and 23% in the BCS), the roles of investments and family background grow in importance once indirect channels are allowed for. This is true for parental time investments in the BCS where the share explained increases from -2.2% to 21.4%. In both cohorts, the role played by family background variables rises so that around 25% of the IGE can be explained by these variables across both cohorts.

As noted earlier, variables that explain the IGE need not be the same as those that explain the likelihood of upward mobility. However, the mediation analysis does highlight an increased role for investments and family background once they are allowed to influence income via cognition and schooling. Table 8 adds to this by focusing solely on the 80% of observations on whom the ML model is estimated. For these, I estimate a simple linear model of skill production where years of schooling is regressed on standardised cognitive and socioemotional factors scores and family background variables. I then estimate how cognition, socioemotional skills, and investments combine to produce skills (in this case cognition at later ages).

The estimates show several patterns. Firstly, socioemotional skills only weakly relate to years of schooling once cognition is conditioned on. A standard deviation change in cognition increases years of schooling by between 1.8 (NCDS) and 1.2 years (BCS) whereas externalising skills increase years of education by 0.23 (NCDS) and 0.09 (BCS). Despite this, externalising skills are important in so far as they increase cognition. Across both cohorts externalising skills at age 11 have a positive effect upon cognition at age 16. A unit increase in the standardised measure leads to a 0.044 increase in cognition in the NCDS and a 0.149 increase in the BCS. For the BCS, this effect is sizable (the coefficient on age 11 cognition is 0.598 as a comparison).

In line with earlier studies (see for instance Cunha and Heckman, 2008), investments—in this case school quality and parental time—also shift cognition. However, they only have a quantitatively important effect on early cognition (age 11) which then increases later attainment. Investments at age 11 have minimal effect on cognition at age 16.[17]

6. Conclusion

In this paper, I use tools from the machine learning explainability literature to highlight how patterns of upward mobility have changed across two British birth cohorts. While mobility in the cohorts has been studied extensively, the approach here adds to an emerging literature that aims to understand the mechanisms driving mobility outcomes (Chetty and Hendren, 2018; Blanden et al., 2007; Bolt et al., 2021).

The approach makes use of the multiplicity of measures at multiple ages available in cohort studies by using a latent factor framework to condense multiple proxy measurements into low dimensional indices that capture parental time investments, skills, and school quality. These are used in a predictive ML model whose parameters are explicitly chosen to avoid overfitting—a problem inherent in cohort studies that typically have small numbers of observations relative to the number of 'features' used in the predictive model.

Tools from coalitional game theory—Shapley values—are then used to explain how each feature contributes to the model's prediction. One's origin plays an important role, but this is somewhat mechanical—those born in the bottom quintile have more 'room to move up' than those whose only means of being upwardly

[17] While investments at later age do increase socioemotional skills, these only weakly relate to years of schooling and (as shown in Figures 7 and 8) have limited effect on the likelihood of upward mobility.

mobile is to break into the top 20 percent of income earners. Aside from this, cognition and years of schooling play an important role irrespective of gender and the income quintile one was born into. While education and cognition are equally important in each cohort, their relative roles reverse between 1958 and 1970. Years of education have less additional predictive power once cognition is conditioned on for NCDS members than BCS members. For the latter, years of education have significant additional predictive power even when cognition at adolescence is controlled for. Rising returns to education in the UK labour market (for example see Gosling et al., 2000) may well contribute to this result.

For both the NCDS and the BCS, educational outcomes and cognition at adolescence act as sufficient statistics for a range of early childhood factors. Once these are known, there is little predictive role for family composition, school quality, parental time investments, or socioemotional skills. In line with the vast literature on human capital production, these factors are shown to influence cognition, and through this, educational outcomes; however, my results highlight that they play little to no direct role in predicting upward mobility even in a model that allows for arbitrary interactions between these factors, adolescent cognitive ability, and completed years of schooling.

Table 1: Mobility in the Cohorts

	NCDS 1958		BCS 1970	
	Age 33	Age 42	Age 34	Age 42
Elasticity	0.289	0.274	0.389	0.354
	(0.035)	(0.037)	(0.033)	(0.035)
Rank Slope	0.163	0.158	0.280	0.229
	(0.016)	(0.016)	(0.020)	(0.020)
Sample Size	4096	3740	2348	2426

Notes: Robust standard errors in parenthesis. The independent variable in each case is the log/rank of parental income at age 16. For a detailed description of parental income data in the NCDS, see Dearden, Machin, and Reed (1997). Sample sizes refer to those used in the training of the ML models later. Ranks are computed based on the full sample of those with observed earnings.

Table 2: Transition Matrices, NCDS

	Cohort member income quintile, age 42, 2000					% Upwardly
Parental quintile	1	2	3	4	5	Mobile
1	23.59	22.40	23.29	17.95	12.76	76.41
	(23.59)	(45.99)	(69.29)	(87.24)	(100)	
2	22.83	23.36	21.23	18.16	14.42	53.81
	(22.83)	(46.19)	(67.42)	(85.58)	(100)	
3	21.40	19.33	20.75	20.62	17.90	38.52
	(21.40)	(40.73)	(61.48)	(82.10)	(100)	
4	18.44	19.43	18.19	19.93	24.01	24.01
	(18.44)	(37.87)	(56.06)	(75.99)	(100)	
5	16.12	15.18	16.94	21.68	30.08	0
	(16.12)	(31.30)	(48.24)	(69.92)	(100)	

Notes: Quintile 1 refers to those in the lowest income grouping. Numbers in parentheses are cumulative percentages across columns. The final column gives the percentage from each parental income quintile to achieve upward mobility. Overall, 47% of the NCDS sample and 45% of the BCS sample are upwardly mobile. For the secondary measure (bottom two quintiles to top two), the likelihood of upward mobility is 23% in each cohort.

Table 3: Transition Matrices, BCS

	Cohort member income quintile, age 42, 2012					% Upwardly
Parental quintile	1	2	3	4	5	Mobile
1	24.45	26.37	22.25	15.66	11.26	75.55
	(24.45)	(50.82)	(73.08)	(88.74)	(100)	
2	20.89	25.53	23.21	18.57	11.81	53.59
	(20.89)	(46.41)	(69.62)	(88.19)	(100)	
3	21.73	22.54	16.30	20.32	19.11	39.44
	(21.73)	(44.27)	(60.56)	(80.89)	(100)	
4	18.99	17.32	18.81	23.09	21.79	21.79
	(18.99)	(36.31)	(55.12)	(78.21)	(100)	
5	14.26	13.18	14.62	21.12	36.82	0
	(14.26)	(27.44)	(42.06)	(63.18)	(100)	

Notes: See Table 2.

Table 4: Descriptives, Demographics, NCDS

	Bottom 80% Parental Income				Top 20%
	Overall	Upwardly Mobile	Not Upwardly Mobile	Residualised Difference	Overall
Female	0.500	0.289	0.672	−0.471 (0.016)	0.527
Birthweight (Kg)	3.349	3.406	3.310	0.137 (0.021)	3.386
BMI (Age 10)	17.464	17.416	17.489	−0.121 (0.107)	17.383
Teenage mother	0.052	0.046	0.060	−0.016 (0.009)	0.041
Parents married at birth	0.973	0.975	0.970	0.014 (0.007)	0.980
Father figure	0.947	0.947	0.949	0.020 (0.009)	0.965
Parents divorced	0.087	0.090	0.076	−0.028 (0.010)	0.067
Home owners	0.432	0.435	0.433	0.082 (0.019)	0.709
Household size	5.074	4.958	5.071	−0.140 (0.062)	4.996
Parent attended university	0.060	0.068	0.053	0.033 (0.010)	0.232
Attended university	0.176	0.262	0.125	0.182 (0.017)	0.325

Notes: Upwardly mobile is measured as being in a higher income quintile (at age 42) than one's parents where parental income is measured at age 16. By construction, those born in the top quintile cannot be upwardly mobile. Differences by mobility status are confounded by origin; I present regression adjusted differences that partial out the effect of parental income quintile. Robust standard errors in parentheses.

Table 5: Descriptives, Demographics, BCS

	Bottom 80% Parental Income				Top 20%
	Overall	Upwardly Mobile	Not Upwardly Mobile	Residualised Difference	Overall
Female	0.540	0.360	0.705	−0.414 (0.022)	0.510
Birthweight	3.335	3.371	3.306	0.102 (0.026)	3.428
BMI (Age 10)	16.919	16.832	16.975	−0.215 (0.116)	16.715
Teenage mother	0.093	0.096	0.087	−0.005 (0.014)	0.046
Parents married at birth	0.956	0.959	0.956	0.019 (0.010)	0.974
Father figure	0.929	0.921	0.935	0.013 (0.013)	0.963
Parents divorced	0.171	0.214	0.133	0.010 (0.018)	0.097
Home owners	0.732	0.698	0.771	0.065 (0.020)	0.974
Household size	4.525	4.569	4.457	−0.036 (0.061)	4.313
Parent attended university	0.158	0.175	0.151	0.088 (0.019)	0.540
Attended university	0.351	0.482	0.270	0.320 (0.023)	0.621

Notes: See Table 4.

Table 6: Algorithm Performance, NCDS and BCS

	NCDS				BCS			
	Naive	LPM	ML Boost	ML Boost (tuned)	Naive	LPM	ML Boost	ML Boost (tuned)
Mobility 1								
Accuracy	0.53	0.77	0.76	0.78	0.55	0.77	0.74	0.79
Precision	0	0.77	0.75	0.77	0	0.75	0.72	0.80
AUC	0.50	0.86	0.84	0.86	0.50	0.85	0.82	0.79
Mobility 2								
Accuracy	0.91	0.91	0.91	0.92	0.92	0.92	0.91	0.96
Precision	0	0	0.51	0.59	0	0	0.25	0.86
AUC	0.50	0.78	0.76	0.79	0.50	0.74	0.67	0.98

Notes: Sample sizes for the NCDS and BCS are 2844 and 1536 respectively. Mobility 1 takes value 1 if the cohort member is born in any of the bottom 4 income quintiles and resides in a higher quintile at age 42. Mobility 2 focuses on mobility from quintiles 1 and 2 to quintiles 4 and 5. The boosting algorithm is XGBoost with Bayesian hyperparameter optimisation. Performance is assessed using 5-fold cross validation. The naive estimator assigns each observation to the most common class.

Table 7: Mediation Analysis, NCDS and BCS

	NCDS		BCS	
	Simple	Complex	Simple	Complex
Years of schooling	0.457	0.073	0.384	0.123
	[0.337, 0.606]	[0.021, 0.141]	[0.293, 0.493]	[0.07, 0.191]
Cognition	0.304	0.147	0.104	0.107
	[0.196, 0.444]	[0.086, 0.219]	[0.034, 0.187]	[0.061, 0.16]
Socio-emotional skills	−0.021	0.012	−0.032	−0.010
	[−0.054, 0.006]	[−0.01, 0.039]	[−0.09, 0.02]	[−0.038, 0.011]
Time investments	−0.022	0.214	0.003	0.003
	[−0.103, 0.062]	[0.139, 0.316]	[−0.01, 0.017]	[−0.006, 0.015]
School quality	0.038	0.064	0.115	0.138
	[−0.027, 0.105]	[0.022, 0.111]	[0.039, 0.194]	[0.081, 0.211]
Family background	0.054	0.248	0.101	0.250
	[−0.03, 0.157]	[0.141, 0.377]	[0.007, 0.200]	[0.145, 0.359]
Sample Size	2877		1995	

Notes: The table shows the fraction of the IGE between parental income at age 16 and one's own income at 42 that is explained by each variable. Years of schooling is derived via converting NVQ levels into a continuous variable. Cognition, socio-emotional skills, and time investments are pooled estimates over the course of the cohort member's childhood. School quality is measured at age 10 in the BCS and age 7 in the NCDS. 95% confidence intervals are derived from 500 bootstrap replications.

Table 8: Linear Skill Production Function, NCDS and BCS

	NCDS			BCS		
	Years of Schooling	Cognition (16)	Cognitive (11)	Years of Schooling	Cognition (16)	Cognitive (11)
Cognition $(t-1)$	1.816	0.666	0.648	1.177	0.598	0.440
	(0.058)	(0.010)	(0.014)	(0.097)	(0.015)	(0.022)
Internalising $(t-1)$	-0.059	-0.033	—	0.057	-0.104	0.004
	(0.073)	(0.012)		(0.117)	(0.015)	(0.025)
Externalising $(t-1)$	0.234	0.044	0.016	0.090	0.149	0.041
	(0.081)	(0.013)	(0.014)	(0.109)	(0.018)	(0.028)
Time Invest $(t-1)$	—	0.023	0.037	—	0.009	0.086
		(0.010)	(0.018)		(0.014)	(0.023)
School Quality $(t-1)$	—	0.031	0.132	—	0.018	0.153
		(0.010)	(0.018)		(0.014)	(0.021)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Sample Size		2312			1534	

Notes: Estimation is based on those in the bottom 80% of the parental income distribution. Latent variables are standardised within the sample. Controls are all family background variables alongside dummies for parental income quintiles. Standard errors in parentheses. For the NCDS, I use a single socio-emotional skill measure at age 7; the coefficient is reported in the externalising $(t-1)$ row.

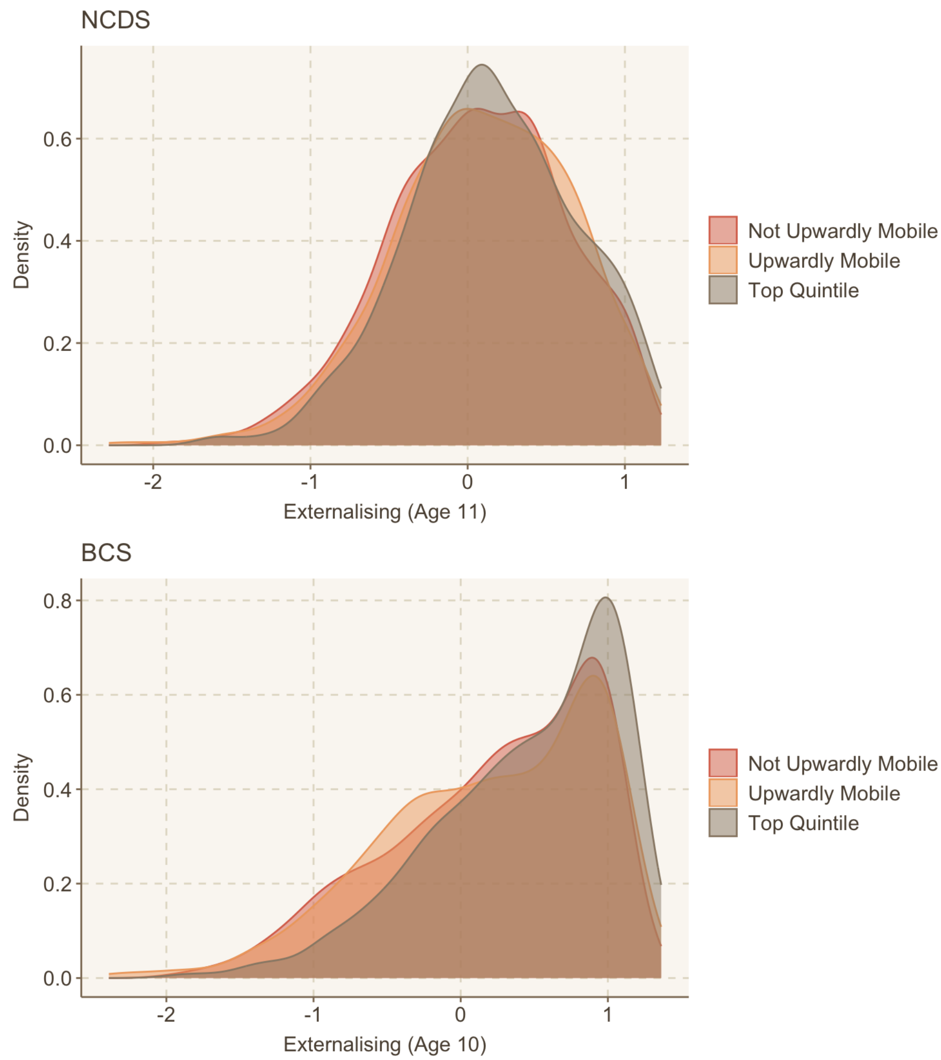

Figure 2: Externalising Skills, Mid-Childhood.

Notes: Distribution of factor scores plotted for those who achieve upward mobility, those who do not, and those born into the top quintile of family incomes. A regression of ranked factor scores on two group indicators (top quintile omitted) gives coefficients of -3.94 (1.29) and -2.30 (1.30) for not upwardly mobile and upwardly mobile respectively in the NCDS. The same results for the BCS are -9.78 (1.54) and -10.20 (1.62).

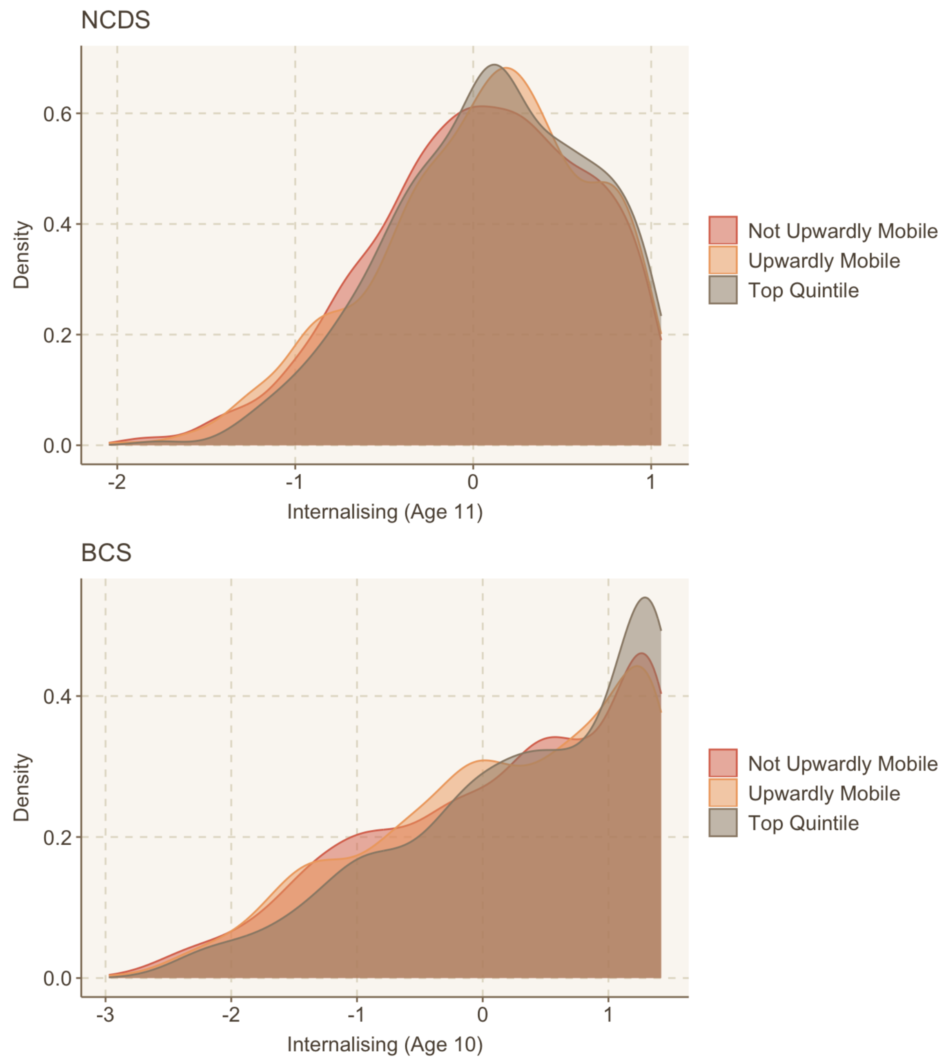

Figure 3: Internalising Skills, Mid-Childhood.

Notes: See Figure 2. Coefficients are −3.08 (1.29) and −1.80 (1.31) for the NCDS; −5.68 (1.58) and −5.69 (1.64) for the BCS.

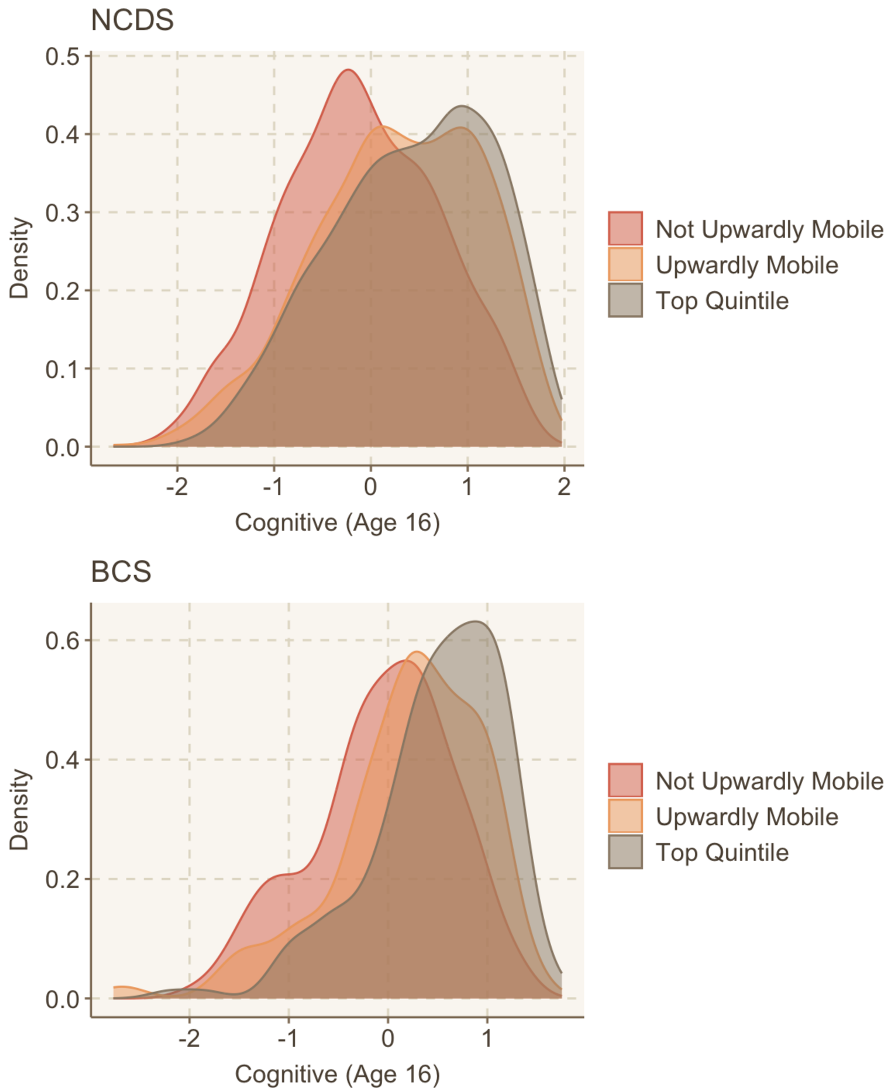

Figure 4: Cognitive Skills, Age 16.

Notes: See Figure 2. Coefficients are -18.02 (1.22) and -4.60 (1.28) for the NCDS; -23.37 (2.74) and -11.91 (3.01) for the BCS.

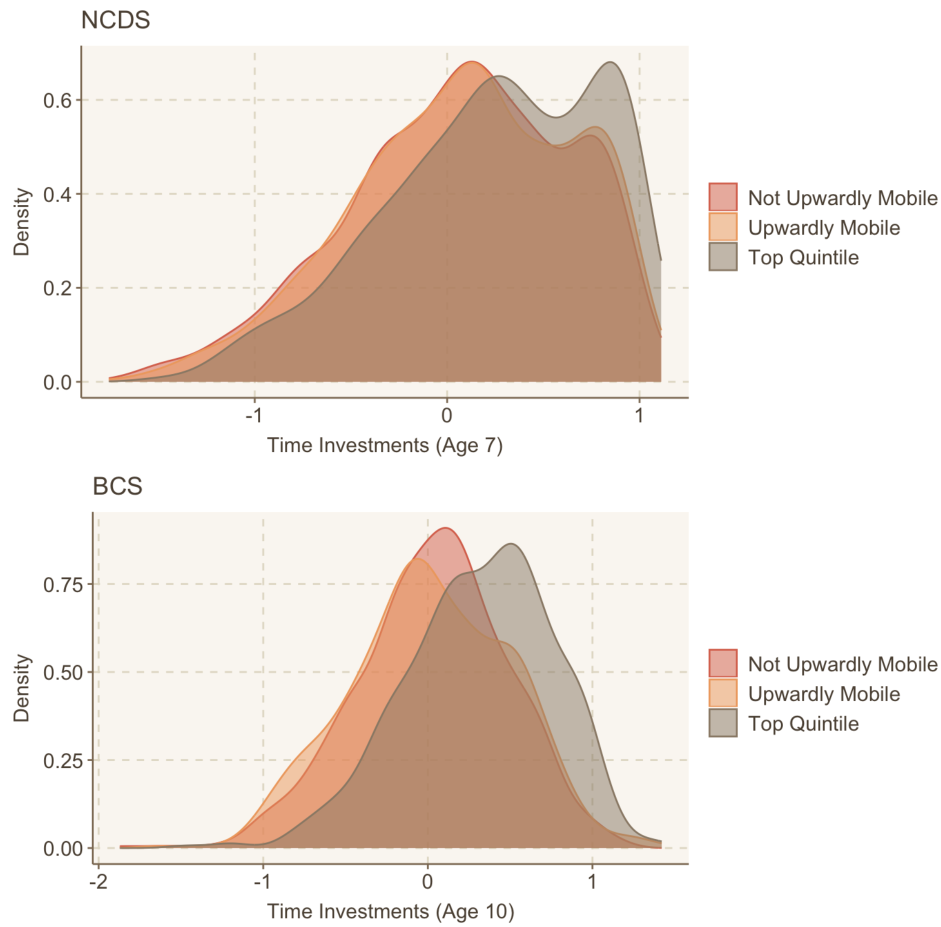

Figure 5: Time Investments, Mid-Childhood.

Notes: See Figure 2. Coefficients are -9.02 (1.34) and -7.98 (1.37) for the NCDS; -17.31 (1.51) and -17.96 (1.62) for the BCS.

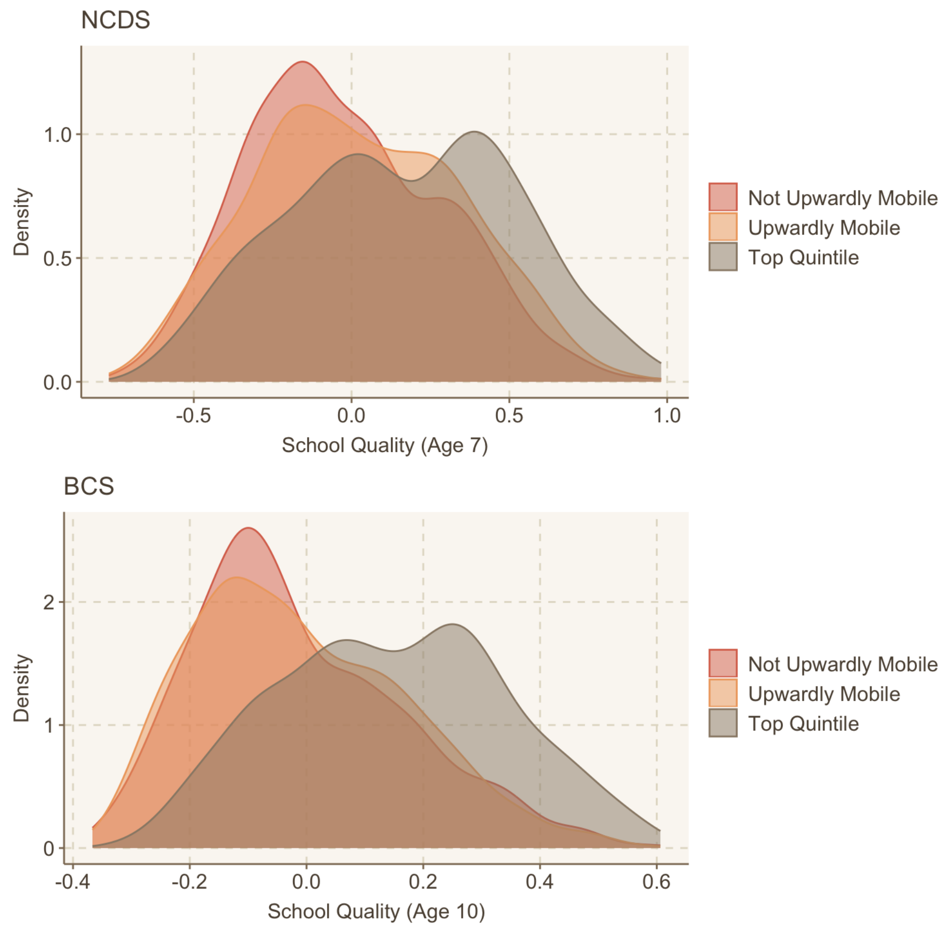

Figure 6: School Quality, Mid-Childhood.

Notes: See Figure 2. Coefficients are −15.44 (1.33) and −11.34 (1.38) for the NCDS; −22.97 (1.46) and −22.41 (1.54) for the BCS.

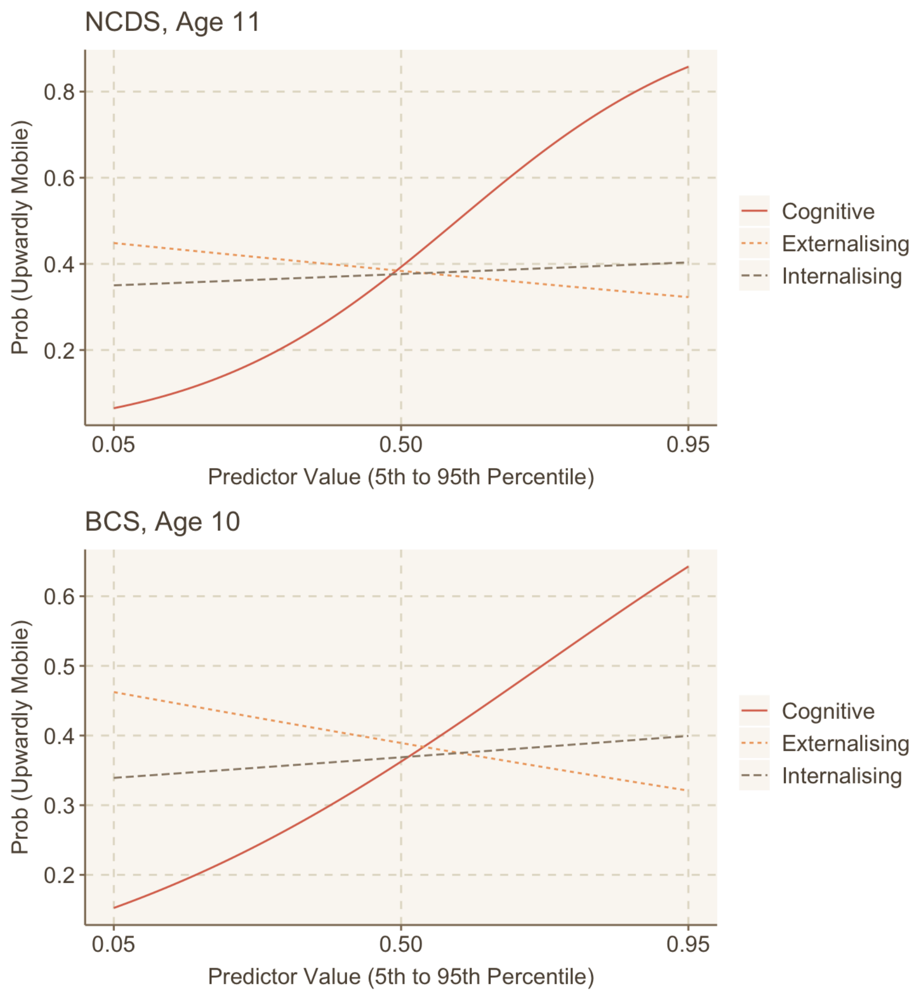

Figure 7: The Comparative Roles of Cognition and Socio-Emotional Skills.

Notes: Probabilities estimated via logistic regression of an upward mobility dummy on cognitive ability, externalising and internalising skills, school quality, and time investments alongside parental income quintile. All factors held at mean; factor of interest varied from 5th to 95th percentile. Fitted values are for those in the 3rd income quintile.

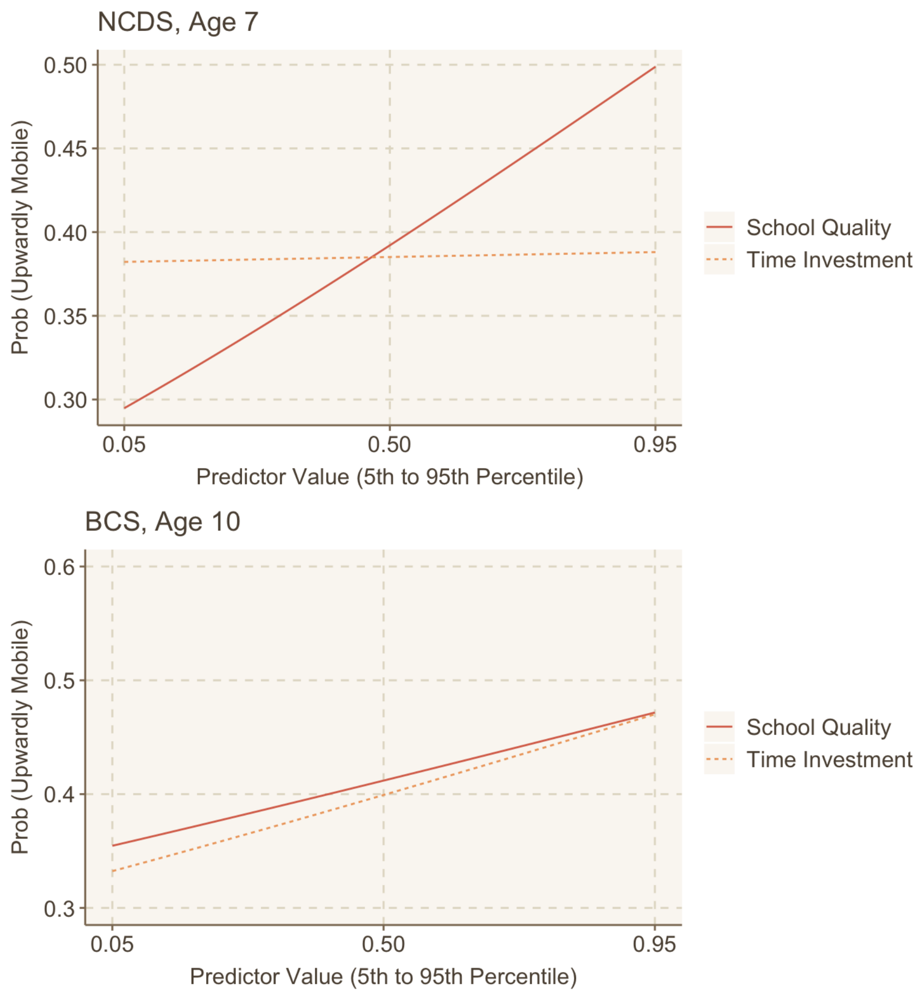

Figure 8: The Comparative Roles of School and Time Investments.
Notes: See Figure 7.

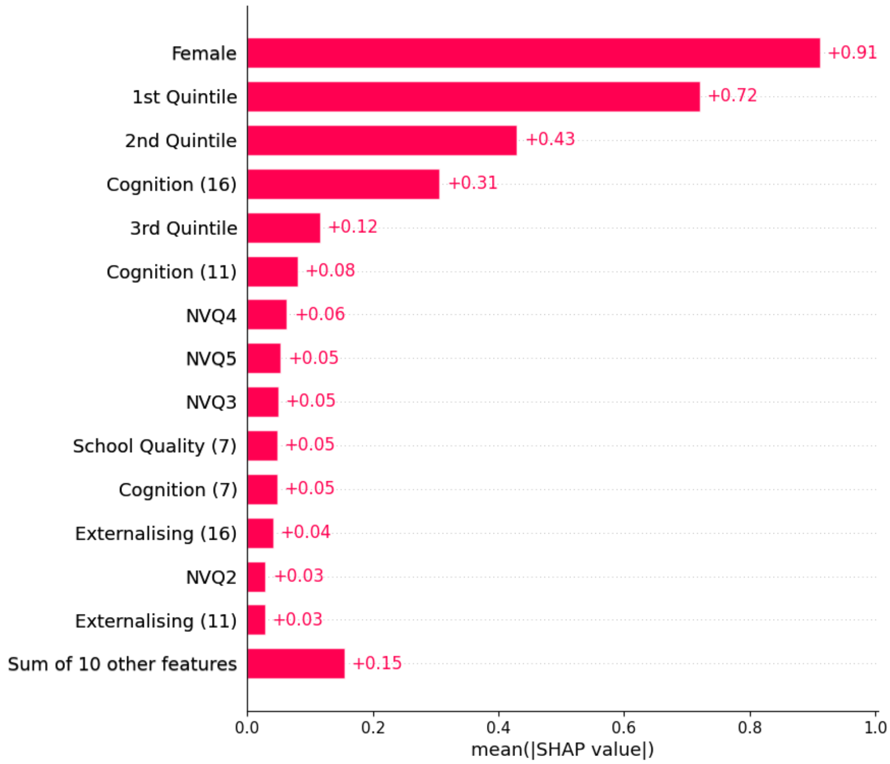

Figure 9: SHAP Values, NCDS.

Notes: Shapley values based on the boosting model fitted on the entire sample with optimal parameters. Features with negligible average absolute Shapley values are excluded. Output is measured as the log odds ratio. Mobility is defined as residing in a higher income quintile at age 42 than one's parental quintile.

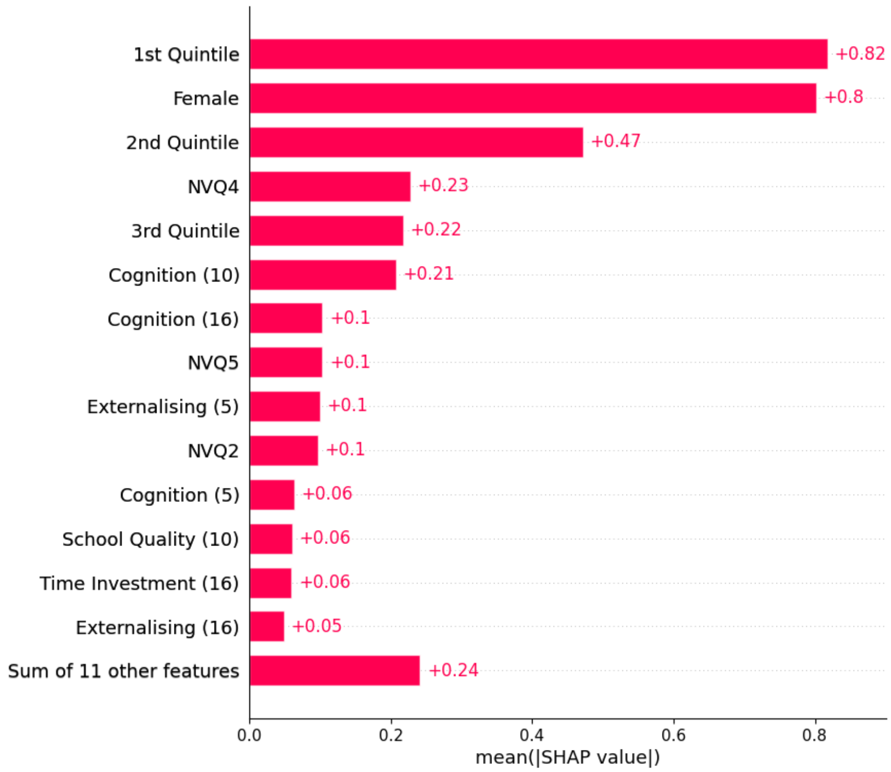

Figure 10: SHAP Values, BCS.

Notes: See Figure 9.

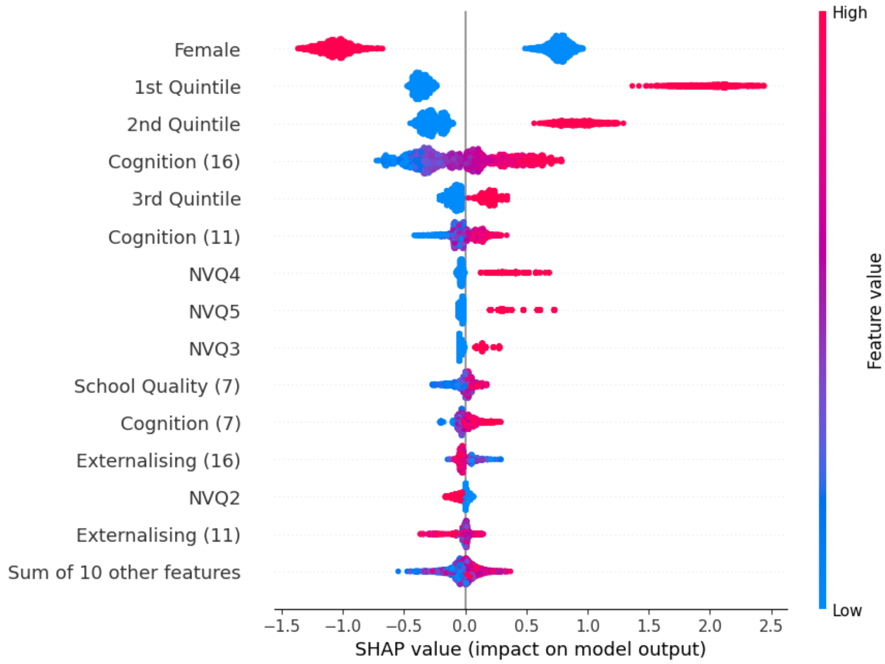

Figure 11: Beeswarm Plot, NCDS.

Notes: Each point represents an observation and its SHAP value for that feature. Output is measured as the log odds ratio. Mobility 1 takes value 1 if the cohort member resides in a higher income quintile at age 42 than the parental quintile.

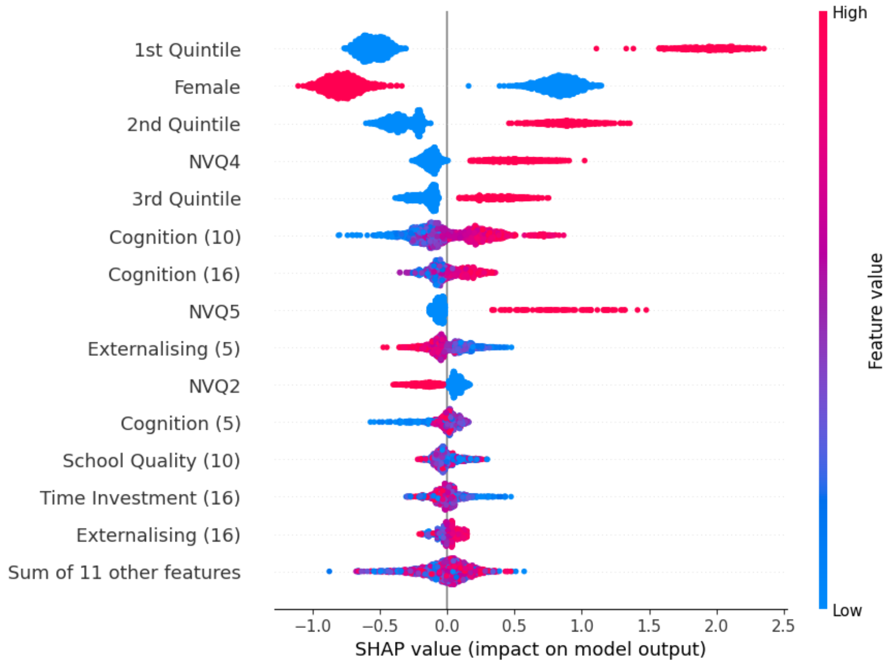

Figure 12: Beeswarm Plot, BCS.

Notes: See Figure 11.

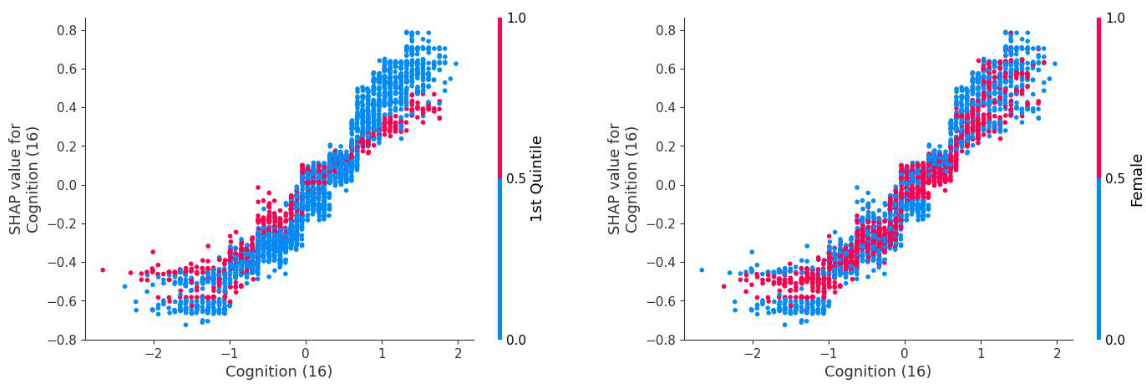

Figure 13: Interaction Plots, NCDS.

Notes: This plots the size of the interaction effects and as such does not account for the direct effect of either cognition, origin quintile, or gender.

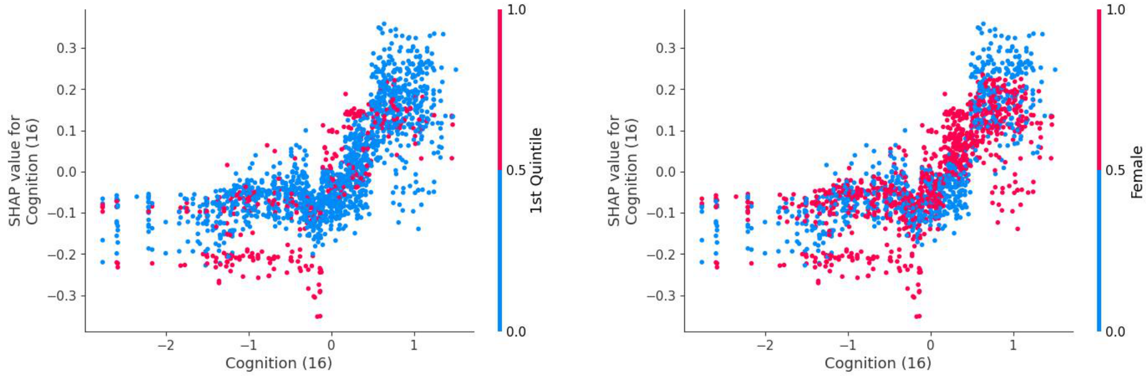

Figure 14: Interaction Plots, BCS.

Notes: See Figure 13.

References

Agostinelli, F. and Wiswall, M. (2025). Estimating the technology of children's skill formation. *Journal of Political Economy*, 133(3), 846–887.

Athey, S. and Imbens, G.W. (2019). Machine learning methods that economists should know about. *Annual Review of Economics*, 11(1), 685–725.

Attanasio, O., Blundell, R., Conti, G., and Mason, G. (2020). Inequality in socio-emotional skills: A cross-cohort comparison. *Journal of Public Economics*, 191, 1041–1071.

Blanden, J., Doepke, M., and Stuhler, J. (2022). Educational inequality. IZA Discussion Paper No. 15225.

Blanden, J., Eyles, A., and Machin, S. (2023). Intergenerational home ownership. *Journal of Economic Inequality*, 21, 251–275.

Blanden, J., Gregg, P., and Machin, S. (2005). *Intergenerational Mobility in Europe and North America*. Sutton Trust.

Blanden, J., Gregg, P., and Macmillan, L. (2007). Accounting for intergenerational income persistence: Noncognitive skills, ability and education. *The Economic Journal*, 117, 43–60.

Blanden, J. and Machin, S. (2004). Educational inequality and the expansion of UK higher education. *Scottish Journal of Political Economy*, 51, 230–249.

Blundell, J. and Risa, E. (2018). Intergenerational mobility as a prediction problem: Theory and evidence from Norway and Britain. Unpublished manuscript.

Bolt, U., French, E., Hentall-MacCuish, J., and O'Dea, C. (2021). Intergenerational altruism and transfers of time and money: A life cycle perspective. Institute for Fiscal Studies.

Bukodi, E. and Goldthorpe, J.H. (2018). *Social Mobility and Education in Britain*. Cambridge University Press.

Chetty, R. and Hendren, N. (2018). The impacts of neighborhoods on intergenerational mobility II: County-level estimates. *The Quarterly Journal of Economics*, 133(3), 1163–1228.

Chetty, R., Hendren, N., Kline, P., and Saez, E. (2014). Where is the land of opportunity? *The Quarterly Journal of Economics*, 129(4), 1553–1623.

Cunha, F. and Heckman, J.J. (2008). Formulating, identifying and estimating the technology of cognitive and noncognitive skill formation. *Journal of Human Resources*, 43(4), 738–82.

Cunha, F., Heckman, J.J., and Schennach, S.M. (2010). Estimating the technology of cognitive and noncognitive skill formation. *Econometrica*, 78, 883–931.

Dearden, L., Machin, S., and Reed, H. (1997). Intergenerational mobility in Britain. *The Economic Journal*, 107, 47–66.

Deming, D. (2017). The growing importance of social skills in the labor market. *The Quarterly Journal of Economics*, 132(4), 1593–1640.

Engzell, P. and Mood, C. (2023). Understanding patterns and trends in income mobility through multiverse analysis. *American Sociological Review*, 88(4), 600–626.

Erikson, R. and Goldthorpe, J.H. (2010). Has social mobility in Britain decreased? *The British Journal of Sociology*, 61, 211–230.

Gosling, A., Machin, S., and Meghir, C. (2000). The changing distribution of male wages in the UK. *Review of Economic Studies*, 67, 635–66.

Gregg, P., Macmillan, L., and Vittori, C. (2017). Moving towards estimating sons' lifetime intergenerational economic mobility in the UK. *Oxford Bulletin of Economics and Statistics*, 79, 79–100.

Hastie, T., Tibshirani, R., and Friedman, J. (2009). *The Elements of Statistical Learning* (2nd ed.). Stanford University.

Heckman, J.J. and Mosso, S. (2014). The economics of human development and social mobility. *Annual Review of Economics*, 6(1), 689–733.

Heckman, J.J., Pinto, R., and Savelyev, P. (2013). Understanding the mechanisms through which an influential early childhood program boosted adult outcomes. *American Economic Review*, 103(6), 2052–86.

Heckman, J.J., Stixrud, J., and Urzua, S. (2006). The effects of cognitive and noncognitive abilities on labor market outcomes and social behavior. *Journal of Labor Economics*, 24(3), 411–482.

Kearney, M.S. (2023). *The Two-Parent Privilege*. University of Chicago Press.

Lundberg, S.M. and Lee, S. (2017). A unified approach to interpreting model predictions. *Proceedings of the 31st International Conference on Neural Information Processing Systems*, 4768–4777.

Moulton, V. et al. (2020). A guide to the cognitive measures in five British birth cohort studies. CLOSER.

Mullainathan, S. and Spiess, J. (2017). Machine learning: An applied econometric approach. *Journal of Economic Perspectives*, 31(2), 87–106.

Muthén, B. (1984). A general structural equation model with dichotomous, ordered categorical, and continuous latent variable indicators. *Psychometrika*, 49(1), 115–132.

Revelle, W. and Rocklin, T. (1979). Very Simple Structure: An alternative procedure for estimating the optimal number of interpretable factors. *Multivariate Behavioral Research*, 14(4), 403–414.

Velicer, W.F. (1976). Determining the number of components from the matrix of partial correlations. *Psychometrika*, 41, 321–327.