

The Life Cycle Dynamics of Wealth Mobility*

Richard Audoly Rory M^cGee Sergio Ocampo Gonzalo Paz-Pardo

March 2023

Preliminary and Incomplete – Do Not Circulate

Abstract

We use the universe of tax records in Norway between 1993 and 2017 to study the movements of individuals across the wealth distribution during their working lives. We find substantial mobility across within-cohort wealth ranks, but an individual's rank is more persistent over the long run than what short-run fluctuations suggest. The persistence of these wealth ranks is substantially higher at the top. To efficiently summarize the underlying life cycle patterns in wealth mobility, we group rank histories adopting clustering tools from the machine learning literature. Many individuals experience reversals of fortune over their life-cycle. While there are persistently poor and persistently rich individuals, many individuals' wealth rank slowly declines over their lifetime, while two large groups climb up the wealth distribution in their late 30s and late 40s, respectively. Ex-ante characteristics have little explanatory power for distinguishing between these groups, even though highly educated individuals are less likely to drop in the wealth distribution later in life, and parental wealth at age 30 is a more important predictor of whether an individual will be lifelong poor or wealthy.

*Audoly: FRB of New York; audolyr@gmail.com; Web: <https://richard-a.github.io>.

M^cGee: Western University; rmcgee4@uwo.ca; Web: <https://sites.google.com/view/rorymcgee>.

Ocampo: Western University; socampod@uwo.ca; Web: <https://sites.google.com/site/sergiocampod>.

Paz-Pardo: ECB; gonzalo.paz_pardo@ecb.europa.eu; Web: <https://www.gonzalopazpardo.com>.

The views below are those of the authors and do not necessarily reflect the position of the Federal Reserve Bank of New York, the Federal Reserve System, the European Central Bank or the Eurosystem. We are grateful for comments and support from Roberto Iacono, Paolo Piacquadio, Alfred Løvgren and the staff at the Oslo Fiscal Studies (OFS) center at the University of Oslo. We thank seminar participants at Stony Brook. We also thank Emmanuel Murray Leclair for research assistance. Ocampo acknowledges financial support from the Research Council of Norway through the project TaxFair, Number 315765. M^cGee and Ocampo acknowledges financial support from the Social Science and Humanities Research Council of Canada through the Insight Development Grant Number 430-2022-00394.

1. Introduction

The movements of individuals across the wealth distribution reflect much of their life experience: the dynamics of their earnings, the businesses they engage in, and their saving and spending choices. However, we know little about the life cycle dynamics of wealth mobility when compared to the dynamics of earnings both within and across generations.¹ This is due, in part, to the limited availability of longitudinal data on individuals' wealth and to the variety of life paths taken by individuals.

Our goal in this paper is to document the lifetime trajectories of individuals through the wealth distribution. To do so, we use administrative data from the Norwegian tax registry that provides us with 25 years of observations on individuals' wealth from 1993 to 2017 for the entire Norwegian population.² This data makes it possible to explore how the position of individuals in the (entire) wealth distribution moves over time, as well as allowing us to relate those movements to a variety of individual characteristics. Accordingly, our object of interest is an individual's percentile rank in the wealth distribution.

We focus on wealth mobility for three main reasons. First, differences in wealth at any age are one of the best proxies for differences in lifetime economic resources. Second, wealth encodes information on preferences, risks, and luck that are important determinants of individual well-being because it is a stock. Consequently, studying wealth complements the measures of cross-sectional inequality that emerge from

¹See, for instance, [Arellano, Blundell, and Bonhomme \(2017\)](#), [De Nardi, Fella, and Paz Pardo \(2019\)](#), [Guvenen, Karahan, Ozkan, and Song \(2021\)](#), and [Guvenen, Kaplan, Song, and Weidner \(2022\)](#) for the dynamics of earnings over the life cycle, and [Solon \(1992\)](#), [Chetty, Hendren, Kline, Saez, and Turner \(2014\)](#), [Chetty, Grusky, Hendren, Hell, Manduca, and Narang \(2017\)](#), and [Halvorsen, Ozkan, and Salgado \(2022\)](#) for accounts of inter-generational mobility.

²These data have been used to study the role of return heterogeneity in generating wealth inequality ([Fagereng, Guiso, Malacrinò, and Pistaferri 2020](#)), the trajectories of the wealthiest individuals in the economy ([Halvorsen, Hubmer, Ozkan, and Salgado 2023](#)), and the relation between wealth and lifetime income ([Black, Devereux, Landaud, and Salvanes 2020](#)), among other questions. We instead focus on characterizing the life cycle paths of individuals across the entire wealth distribution.

studying earnings. Third, understanding intra- and inter-generational mobility in the entire trajectory of individual wealth is crucial for quantifying (in)equality of opportunities.

We divide our analysis in two broad steps. First, we document the persistence of wealth ranks over different horizons and at different ages. This provides us with an overview of how mobile individuals are within their own lifetime. We also document how the distribution of future ranks varies by age, horizon, and, most importantly, an individual's current position in the wealth distribution.³ Second, we exploit machine learning algorithms to elicit the patterns underlying individuals' trajectories through the wealth distribution. Specifically, we use agglomerative hierarchical clustering to group trajectories allowing us to overcome the inherent challenges in characterizing and predicting the nature of wealth mobility.⁴

Our analysis allows us to summarize individual wealth rank histories without resorting to strong parametric assumptions that can mask many of their features, such as non-linear and non-Markovian dynamics. Further, we establish how the dynamics of wealth relate to economic forces that shape life cycle decisions including household formation and dissolution, receiving inheritances, and the dynamics of individual earnings. Finally, we build a new measure of equality of opportunity, by asking to what extent full life cycle histories can be predicted using only information on individuals near the start of their working lives.

Our findings contribute to our understanding of wealth mobility, wealth dynamics,

³While we focus on distributions we also formalise these statistical properties by reporting the conditional higher-order moments of innovations to an individual's rank.

⁴See [Hastie, Tibshirani, Friedman, and Friedman \(2009\)](#), ch. 14 for an introduction to clustering. This approach is common in Sociology, where it has been used to summarize sequences of categorical variables ([Dijkstra and Taris 1995](#); [McVicar and Anyadike-Danes 2002](#); [Dlouhy and Biemann 2015](#)) and in economics to study the in-and-outs of self-employment ([Humphries 2021](#)). Clustering approaches have also gained popularity in the study of labour market transitions and sorting (e.g., [Gregory, Menzio, and Wiczer 2021](#); [Bonhomme, Lamadon, and Manresa 2022](#); [Lentz, Piyapromdee, and Robin 2022](#)). Alternative methods that impose dynamics following a hidden Markov chain have also proven useful in the study of labor earnings and labor force participation (e.g., [Ahn, Hobijn, and Şahin 2023](#)).

and inequality in three main ways. First, we find that the annual persistence of wealth ranks is high, but lower than the persistence of earnings ranks (see [Halvorsen, Ozkan, and Salgado 2022](#)). Wealth mobility across ranks decreases with age, and is larger for the poorest agents and for the rich, but not the very richest. For example, an individual at the 20th percentile of wealth at age 35 expects, on average, to reach percentile 40 at age 45, but has a relatively low chance of climbing above the median of the wealth distribution.

Medium and long run mobility are low relative to short run (annual) mobility. In fact, wealth mobility at 5- and 10-year horizons is much lower than what 1-year movements imply, suggesting that wealth dynamics cannot be well captured by a first-order Markovian process. Rank changes also display rich, nonlinear dynamics. These facts cannot be established by observing the cross-sectional wealth distribution and are of interest to policymakers concerned with the distributive effects of wealth taxes and the overall tax-benefit system.

Second, we describe typical life cycle paths through the wealth distribution and link them to observable characteristics at both labour market entry and later in life. We focus on six groups of individuals. There are two large groups that stay persistently poor (20 percent of the population) and persistently rich (13 percent). A significant part of the population divided in two groups suffers a decline in wealth rank over their lifetime (40 percent), whilst the remaining two groups capture individuals who rise through the distribution, with 9 percent of individuals being “early risers”, improving their position in the distribution between the ages of 35 and 40, and 18 percent of individuals being “late risers” whose rank improve between the ages of 45 and 60.

Characterizing these typical paths helps quantify economic mobility and provides important data with which to confront models of wealth accumulation over the life cycle. For instance, among risers and fallers, we find important differences in portfolio composition as they age. An individual’s rise is closely correlated with the increase in

the private equity share of their portfolio. Despite these differences, the majority of risers and fallers hold similar portfolios at age 31.

Third, we exploit our rich micro data on the characteristics of individuals and their parents to predict wealth rank histories. This moves beyond a snapshot approach to inter-generational mobility and asks whether individual characteristics, including parental wealth, can predict life cycle histories.⁵ We find an important and non-linear role for family background. Individuals born to parents at the top of the wealth distribution are almost 30 percentage points more likely to enjoy a trajectory where they are permanently at or near the top of their generation's distribution, when compared to those born to parents at the bottom of the distribution. In contrast, those born to parents at the bottom of the distribution are not only likely to be poorer at any given snapshot, but also much more likely to be persistently poor. However, parental wealth plays almost no role for individuals who experience a rise or fall through the distribution.

Turning to the role of education, we find that highly educated individuals are less likely to fall through the wealth distribution later in life. We also find that the effect of education on the probability of belonging to the group of the persistently rich, although significant, is of a much smaller magnitude than that of parental wealth.

In addition to these contributions, we believe that we provide a data driven approach to extract moments approximating the universe of individual histories from administrative data. We hope that this will allow researchers interested in using high quality administrative data to estimate structural models to use the richness of the data in a tractable way, while also respecting disclosure limits.

⁵The snapshot approach measures inter-generational mobility by comparing the rank of different generations at a similar point in their life cycle, thus relying on a snapshot of their wealth trajectory to infer mobility (see, e.g., Chetty et al. 2014; Fagereng et al. 2020; Fagereng, Mogstad, and Rønning 2021). Our approach makes it possible to take into account the trajectory of wealth of individuals.

2. Data: A panel of wealth histories for the Norwegian population

We employ data from the Norwegian tax registry between 1993 and 2017 and its associated population characteristics files. We are able to link these various data sets at the individual and household levels using unique (anonymized) identifiers. The resulting data contains information on wealth (net worth), assets, debt, and a variety of individual characteristics.⁶ The coverage and properties of the Norwegian administrative data sets it apart from survey and administrative data available in other countries and makes it uniquely suited to the study of the evolution of wealth mobility over the life cycle. We highlight the most important aspects of the data below before giving a more detailed description of the variables we use and our sample selection.

First, Norway has recorded the information in their wealth tax returns since 1993 providing us with a long panel with over 25 years of observations. This allows us to track individuals over important phases of their life cycle. This is crucial to understand mobility over long horizons and to differentiate between different life cycle trajectories experienced by individuals, as we do in our clustering procedure.

Second, the Norwegian income and wealth tax records capture the entire population. We are therefore able to construct accurate measures of individual's position in the wealth distribution, within cohorts and for the population at large. Furthermore, the data covers individuals at the very bottom and the top of the distribution, who are typically difficult to capture in survey data.⁷

⁶The quality and detail of this data has proven useful in a variety of studies. For more information on the Norwegian Administrative data we refer the interested reader to [Fagereng et al. \(2020\)](#), [Fagereng, Mogstad, and Rønning \(2021\)](#), and [Fagereng, Holm, and Natvik \(2021\)](#).

⁷This problem leads to methods that over-sample the tails of the distribution. These methods are ill suited to the focus of our study. For example, the US Panel Study of Income Dynamics over-samples lower income households (the Survey of Economic Opportunity households) while the Survey of Consumer Finances over-samples wealthier households. Researchers often resort to ad-hoc methods to build more accurate measures of the upper tail of the wealth distribution where wealth is extremely concentrated, for example by augmenting the SCF with the Forbes-400 list of the 400 richest Americans or Estate tax data (see, e.g., [Vermeulen 2016](#)). [Davies and Shorrocks \(2000\)](#) provide an extensive review of these methods.

Third, we are able to link individuals across generations and to their demographic and educational information. This wealth of information lets us not only document trajectories of wealth mobility, but also document their determinants.

Finally, the Norwegian administrative data is uniquely suited for our dynamic analysis because most of the components of income and wealth we rely on are reported by third parties and are not top- or bottom-coded, eliminating concerns about measurement error from self-reporting and censoring that is common in survey data. These sources of measurement error have a limited effect on mean estimates, but represent an important challenge for the study of dynamics of individual observations, representing a key source of bias that typically attenuates persistence measures. In the income dynamics literature this has lead to the popularity of errors in variables estimators (see, e.g., Haider and Solon 2006; Bolt, French, Maccuish, and O'Dea 2021).

We describe key features of our data and the variables of interest below.

Demographic Descriptors. We observe the individuals' immutable characteristics such as date and place of birth, gender, parents' identifiers, date of death, and immigration status.

Civil Status. We observe the individual's civil status as well as their cohabitation status as recorded by the Norwegian Statistical Bureau (SSB). This allows us to classify individuals as married or in cohabitation, as well as other sub-categories like divorced and widowed.

Education. Six-digit classification of the level and field of study according to the Norwegian Standard Classification of Education (NUS2000).⁸ These data provides 9

⁸See <https://www.ssb.no/en/klass/klassifikasjoner/36>.

levels of education ranging from no education to post-graduate PhD level, and 350 coarse fields of study that we use when cataloguing the education of individuals.

Wealth. We observe the individuals' assets, debt, and net worth as reported in their wealth tax return. Housing values are adjusted using the reported values in [Fagereng, Holm, and Torstensen \(2020\)](#). We report levels in 2019 US dollars for convenience.

Portfolio Components. In addition to observing individuals' assets and debt, we are also able to observe sub-components reported in their wealth tax return. We aggregate primary residences, secondary residences and leisure properties, and foreign residences into a property category. Vehicles includes both auto-mobiles and boats. Public equity is defined as directly owned stocks which are traded on the Norwegian Stock Exchange. Private equity includes the value of business assets and unlisted stocks. Our measure of safe assets includes government bonds, chequing accounts, and shares in money market and mutual funds.⁹ We also observe foreign assets and a residual category that includes hard to value assets such as jewellery and paintings – we do not report results for these categories in the paper. We report levels in 2019 U.S. dollars for convenience.

Sample Selection. We begin with the universe of Norwegian residents at any point between 1993 to 2017. We first create a broad cross-cohort sample with individuals born after 1905 (Norwegian independence) and before 1990. We also exclude individuals who ever emigrate from Norway and those who either immigrated after the age of 25 or who arrived after 2011. We use this sample for studying the age-profile of rank persistence in Section [3.1](#).

Our primary sample of interest focuses on the 1960-1964 birth cohort. This birth

⁹We view the last of these items as less safe than government bonds and deposits, however, data restrictions prevent us from considering an alternative definition where we pool this with public equity.

cohort is first observed in their early thirties in our data (1993-2017) and is therefore observed for a significant fraction of their work life. In addition to appearing in our sample for ages of interest, this cohort faces a compulsory school age known at their birth and is not affected by the 1959 reform (this reform was not implemented uniformly across place and time, see [Black, Devereux, and Salvanes 2005](#); [Bhuller, Mogstad, and Salvanes 2017](#), for more details). We use this sample to calculate the within cohort ranks that form our primary outcome of interest.

To illustrate the value of our Norwegian administrative data, our sample selection criteria yields 292,222 individuals in our cohort of interest. By contrast, before we impose any other restrictions there are only 1463 unique households in the US Panel Study of Income Dynamics who satisfy the same birth cohort criteria.

Finally, we further restrict the sample to ensure that it is balanced over the 25 years we observe individuals, leaving a total of 217,383. We will use this sample for the clustering exercise of Section [4.2](#). Our balancing eliminates unavoidable attrition, driven by migration and mortality. Given the age ranges we consider, increasing mortality in late middle-age drives a large share of this sample selection criteria.

3. Mobility and the Persistence of Wealth Ranks

We first construct measures of individual ranks within the wealth distribution. For our cohort of interest we construct yearly individual ranks using the unbalanced sub-sample from 1993 to 2017. We rank all individuals by their net worth in each tax year, producing a measure of each individual's position within their cohort's wealth distribution. Consequently, our measure of rank is not contaminated by spurious

cross-cohort or cross-age comparisons.¹⁰

Formally, given individual i 's net worth in time t , $w_{i,t}$, we compute ranks within our cohort of interest for each calendar year as:

$$y_{i,t} = 100 \times F_w(w_{i,t}|t, i \in BC(i)), \quad (1)$$

where we multiply by 100 to express ranks on a percentile scale. We report the levels of selected percentiles of the wealth distribution at different ages for our cohort of interest in Table A.1 of Appendix A.

3.1. Persistence in wealth ranks

We begin our analysis by documenting the persistence in wealth rank over the life cycle measured by their h^{th} -order autoregressive coefficient. Specifically, we take our measure of an individual's within cohort wealth rank $y_{i,t}$ and estimate the following regression separately at each age t allowing the horizon parameter, h , to take values $h \in \{1, 5, 10\}$:

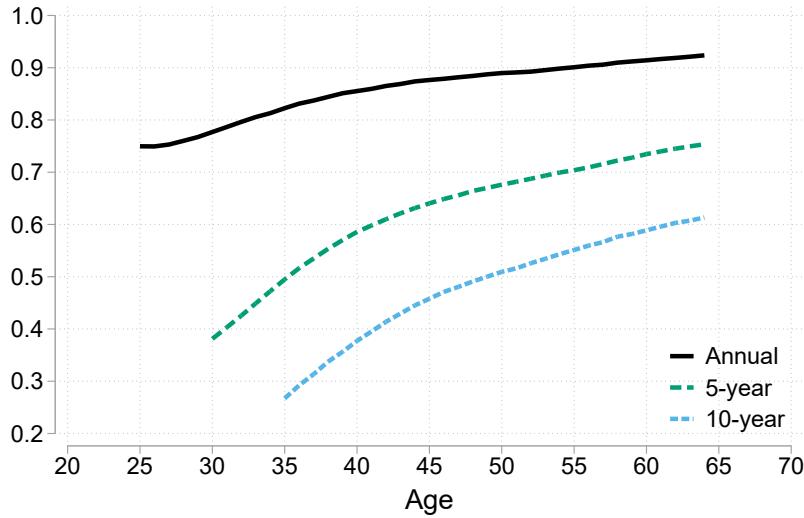
$$y_{i,t} = \alpha_t(h) + \rho_t(h)y_{i,t-h} + u_{i,t}, \quad (2)$$

where $\rho_t(h)$ captures the persistence of wealth rank at age/year t over a horizon of h years. We report the results of these regressions for the full sample in Figure 1 and report results for our selected sample corresponding to the 1960-1964 birth cohort in Figure A.1 of Appendix A. All ranks are within cohort as described in (1).

For the one-year persistence measure, Figure 1 reveals a pattern of increasing persistence over the first 20 years of the life cycle. Persistence is relatively low at age 25

¹⁰Importantly, this also purges ranks from time effects varying by age. For instance, all members of our sample experience the effects of the 2008 global recession at approximately the same age. We therefore do not consider scarring effects or cross cohort differences in patterns of life cycle wealth accumulation (see, e.g., [Gale, Gelfond, Fichtner, and Harris 2021](#); [Paz Pardo 2022](#), who document these changes in the US).

FIGURE 1. Persistence in Wealth Rank



Notes: The figure plots the within cohort persistence measure ($\rho_t(h)$ in equation 2) against individual's age. The dark continuous line corresponds to a one year horizon ($h = 1$). The dashed line corresponds to a five year horizon ($h = 5$). The short-dashed line corresponds to a ten year horizon ($h = 10$).

before increasing over the life cycle. While it continues to increase, by age 55 it is close to 0.9—the value it stabilizes around. Interestingly, the increasing then stabilizing pattern bears similarity to the persistence of earnings over the life cycle, but the absolute value of the annual persistence is lower and slower moving than that of earnings which typically stabilizes around 0.9 by age 35.¹¹ This indicates higher volatility in wealth-ranks than in earnings, but not necessarily higher mobility.

At longer horizons, using 5 and 10 year auto-regressive coefficients, the persistence measure takes longer to stabilize. We also find a decline in the absolute level of persistence. However, the persistence over a 5 (or 10) year horizon is substantially higher than a naive measure computed as the product of the 1-year persistence over the

¹¹For instance, [Karahan and Ozkan \(2013\)](#) and [De Nardi, Fella, and Paz Pardo \(2019\)](#) document this pattern for earnings levels the U.S., and [De Nardi, Fella, Knoef, Paz Pardo, and Van Ooijen \(2021\)](#) document it for the Netherlands. In Appendix A.3 we reproduce the analysis for income ranks in Norway and the U.S. using data from the Global Repository of Income Dynamics database ([Guvenen, Pistaferri, and Violante 2022](#)), see [Halvorsen, Ozkan, and Salgado \(2022\)](#) and [McKinney, Abowd, and Janicki \(2022\)](#) for details on the income dynamics of each country.

same time period and can reject the null hypothesis of equivalence. For instance, looking at 36-year olds, their 5-year persistence is 0.54, while the persistence implied by the annual 1-year coefficients is 0.36. Overall, the 5-year persistence is 30-50 percent higher than iterating the annual persistence, this pattern is starker when considering the 10-year persistence, which is 50-250 percent higher. These results indicate that short run mobility in wealth rank is larger than long run mobility and imply non-Markovian dynamics in ranks.¹²

3.2. The distribution of change in wealth ranks

To better understand how and why individuals transition across wealth ranks, we now turn to documenting the properties of rank changes. We consider the distribution of 1, 5, and 10 year rank changes at age 35 and age 45 conditional on their current wealth percentile rank.¹³ We present results for the evolution of wealth levels in Section 4.2.

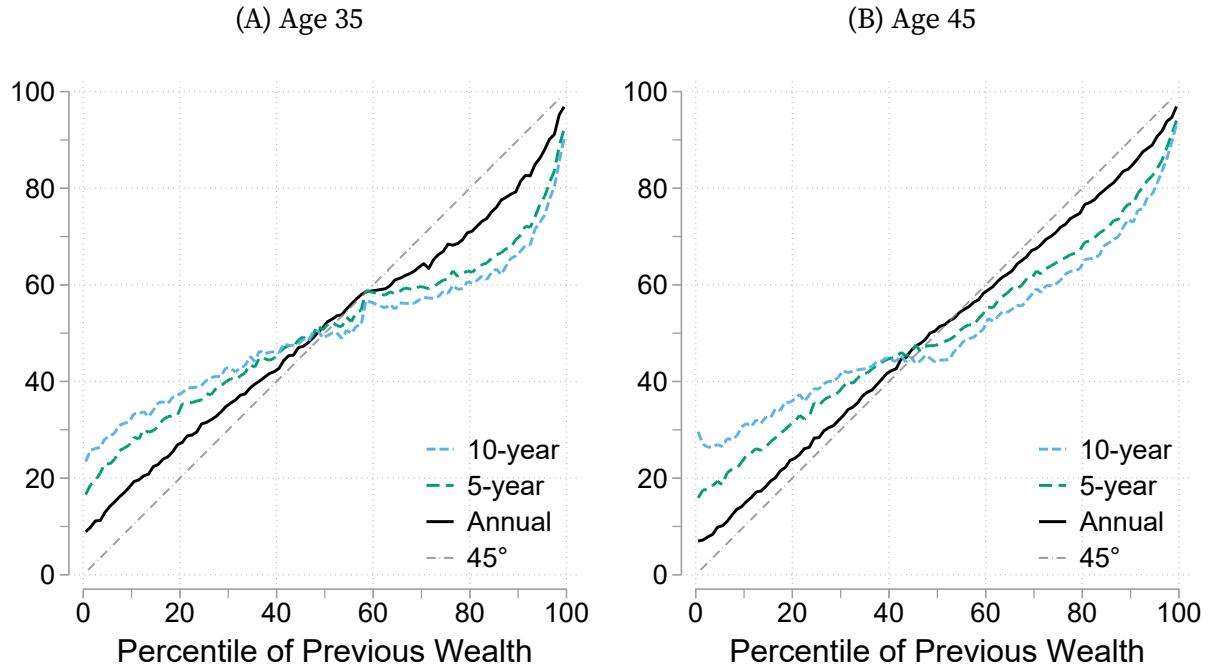
We start by considering the average rank of individuals 1, 5, and 10 years ahead conditional on their previous wealth percentile rank and plot them in Figure 2. To aid comparison, we also plot the 45-degree line which would represent no change in ranks and, absent risk, perfect persistence in wealth positions. The less persistent wealth ranks are, the further away from the 45-degree line their future values will be.

Looking first at the annual lines, we observe relatively small deviations from the 45-degree line, consistent with very little changes year-on-year in wealth ranks. However, as the horizon grows we see more evidence of mean reversion. At age 31, for example, those in the bottom decile see their rank increase by 7 percentiles on average over a single year and by almost 20 percentiles over half a decade (Figure A.4, Appendix A). This still places them well below the median, suggesting that there is some (upwards)

¹²Similar results exist for the inter-generational correlation of income across multiple generations, e.g., [Adermon, Lindahl, and Palme \(2021\)](#).

¹³In Figure A.4 of Appendix A we also report moments at age 31. These results capture the youngest and oldest ages we can use with our wealth panel, which are necessary because we are concerned with transitions across long periods of time.

FIGURE 2. Average Wealth Rank h Years Ahead



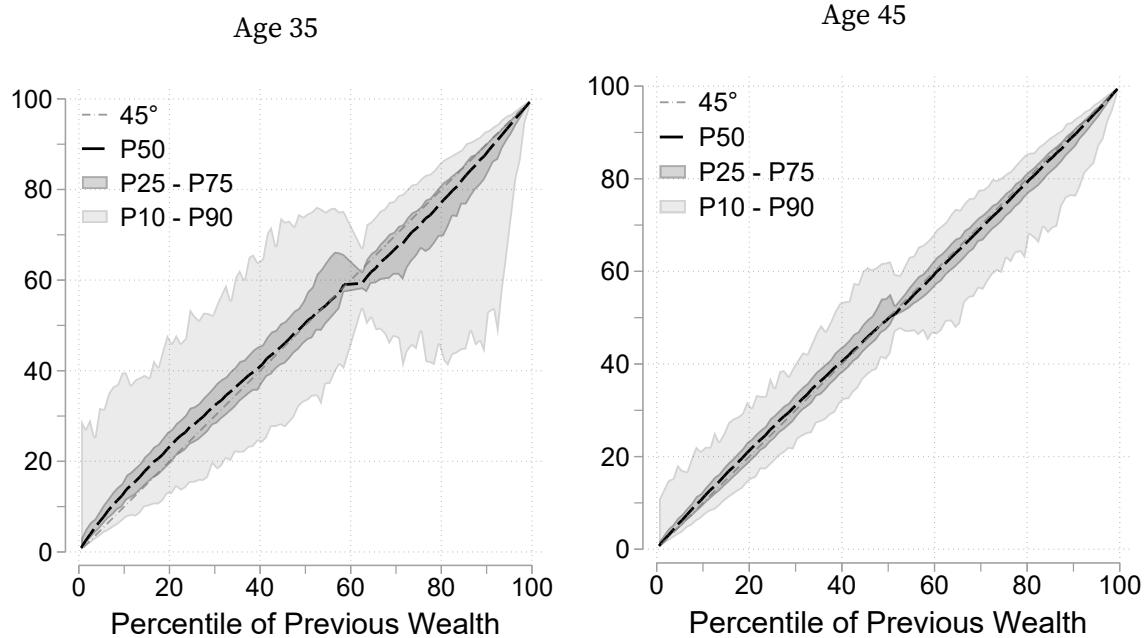
Notes: The figures plot the average future wealth rank of individuals in the 1960-1964 birth cohort conditional on their rank at ages 35, panel 2A, and 45, panel 2B. The dark-continuous line corresponds to a one year horizon ($h = 1$). The dashed line corresponds to a five year horizon ($h = 5$). The short-dashed line corresponds to a ten year horizon ($h = 10$). The light-line corresponds to the 45 degree line.

wealth mobility, but that it is not very strong. Comparing across panels, the extent of this mean reversion declines as individuals age.

As we move up the wealth distribution the expected annual growth in their rank falls to zero, before becoming negative in the top half of the distribution, consistent with mean reversion in ranks. This relationship holds across the bottom 90 percent of the wealth distribution, the pattern is different in the top decile. While, on average, those in the top decile still fall to lower ranks (in part the role of truncation), their fall is much smaller than for individuals in the previous decile. On average, individuals above the 95th percentile are still in the top decile 1-year later and, furthermore, those in the top 1 percent of the distribution fall half as far as individuals at the 95th percentile over any of the horizon-age combinations we consider.

FIGURE 3. Distribution of Wealth Rank Changes

(A) Distribution of 1 Year Rank Changes



(B) Distribution of 5 Year Rank Changes



Notes: The figures plot the 10th, 25th, 50th, 75th, and p90th percentiles of the wealth rank of individuals in the 1960-1964 birth cohort conditional on their rank at ages 35, left panels, and 45, right panels. Panel 3A plots the distribution of 1 year rank changes. Panel 3B plots the distribution of 5 year rank changes.

This striking feature of the data is robust to examining different ages and is more extreme for longer horizons. It also pervasive across the distribution of changes in wealth ranks, as we show in Figure 3 that presents the distribution for 1 and a 5 year rank changes.¹⁴ Most notably, there is a clear compression of the distribution of rank changes at the top. The high persistence of ranks for those at the top already evidenced in Figure 2 is representative of the whole distribution of rank changes for those individuals. This compression starts at about the 90th percentile and is reflected in a strong reduction in the standard deviation of rank changes and a marked increase in the kurtosis of the distribution for the top percentiles.¹⁵

Looking further at the distribution of rank changes conditional on initial rank, we find that the median one year change in wealth rank is close to no change at all, despite there being significant average changes in rank as shown in Figure 2. Even at a 5 year horizon, the median change in rank is close to zero, although there is more mean reversion. The presence of mean reversion is reflected in a negative skewness of the distribution of rank changes, as the figure make clear by the asymmetry in the inter-quartile and p90–p10 ranges.

Despite mean reversion, most of the distribution is concentrated around the 45 degree line as captured by the (narrow) inter-quartile range for one year changes, on average of 7.9 at age 35 and 4.6 at age 45. This reflects the fact that most of the population experiences small changes in their relative position in the wealth distribution, with a minority of individuals experiencing large changes, as captured by the much wider p90–p10 range, on average of 35 at age 35 and 17.6 at age 45. This property of the distribution of rank changes results in a high kurtosis.¹⁶

Taken together, our results evidence important non-linearities in the dynamics

¹⁴Figure A.2 in Appendix A presents the results for a 10 year horizon. The results are qualitatively similar to those for a 5 year horizon but with a higher dispersion in outcomes.

¹⁵We report higher order moments for the distribution of rank changes in Figure A.5 of Appendix A.

¹⁶The dispersion in outcomes is of course higher for 5 year changes with an average inter-quartile range of 27.6 at age 35 and a p90–p10 range of 63.7 at age 45.

of individual mobility and persistent heterogeneity across wealth groups, which we capture below with our clustering approach.

4. Ex-post Analysis: Patterns of wealth mobility

Motivated by the evidence above, which highlights that wealth rank dynamics are non-Markovian and non-linear, we now turn to documenting the full histories of wealth mobility over the life cycle. This entails following the trajectory of individuals through the wealth distribution during the years of our sample and studying the distribution across these outcomes, which are described by each individual's vector of wealth ranks,

$$\mathbf{Y}_i = (y_{i,1993}, y_{i,1994}, \dots, y_{i,2016}, y_{i,2017}) \in [0, 100]^{25}, \quad (3)$$

where $y_{i,t}$ is the (in-cohort) wealth rank of an individual as defined in equation (1).

The distribution of \mathbf{Y}_i across the population is of course a high-dimensional object, making its direct analysis infeasible. We instead proceed by reducing its dimensionality, defining a set of $G > 1$ disjoint groups (or clusters) of individuals, so that each individual i is assigned to one of these groups, $g_i \in \{1, \dots, G\}$. This induces a partition $\mathcal{G}_G = \{g_i\}_{i=1}^N$. In what follows we describe our clustering procedure and then describe the results for wealth mobility and the characteristics of groups in terms of their demographics, portfolio composition, and other relevant variables.

4.1. Methodology: Eliciting patterns of wealth histories

We define groups of individuals following an agglomerative hierarchical clustering approach. Specifically, we use Ward's Method to agglomerate clusters and adopt the

total within-cluster variance as the dissimilarity metric.¹⁷ Thus, letting g and g' denote disjoint groups, the dissimilarity metric between them is

$$d(g, g') = \sqrt{\frac{2N_g N_{g'}}{N_g + N_{g'}}} \times \left\| \mu_g - \mu_{g'} \right\|_2, \quad (4)$$

where N_g is the number of observations in group g and μ_g is the centroid (average) of the observations in group g .

The algorithm works recursively, starting the lowest level of hierarchy, where $G = N$ and each observation is assigned to its own group, and sequentially combining one pair of groups in each iteration. At each level of hierarchy $G > 1$, the algorithm creates the partition at the next level, \mathcal{G}_{G-1} , by selecting the two groups whose combining minimises the dissimilarity metric across groups, that is,

$$\operatorname{argmin}_{g, g' \in G, g \neq g'} d(g, g'). \quad (5)$$

This produces a hierarchy of partitions ranging from $G = N$ to $G = 1$.

A key advantage of our approach is that we are able to consider different moments of, and differences between, the latent representation we recover for any $G \in \{1, \dots, N\}$. This complements alternative approaches, common in the literature on income dynamics, that reduce the dimensionality of the object of interest to a single summary statistic such as the rank-persistence, or that assume that only a subset of the elements of \mathbf{Y}_i are of interest. We believe these approaches are complementary to our approach and return to them later in our analysis of life cycle patterns.

¹⁷Alternative specifications of the dissimilarity metric, including maximum or median distance, are also possible. See [Humphries \(2021\)](#) for another application of Ward's method in the context of Sequence Analysis, where it is used to cluster panel data with discrete states. We produce our Agglomerative Hierarchical Cluster Tree using Matlab, see <https://www.mathworks.com/help/stats/linkage.html>.

Relationship to Sequence Analysis. The agglomerative hierarchical clustering method we employ is closely related to applications of Sequence Analysis tools that also approximate individual histories by membership of a lower dimensional group. These tools, which originate in quantitative sociology (Dijkstra and Taris 1995; McVicar and Anyadike-Danes 2002; Dlouhy and Biemann 2015), are designed to succinctly summarize paths (or sequences) of categorical outcomes and have growing applications in economics (e.g., Humphries 2021). Our approach chiefly differs because we directly use the continuous variation available in our rank measure, while sequence analysis tools seek to minimise a measure of dissimilarity based on transitions across discrete states.

We exploit the continuous variation of the ranks in two ways. First, we avoid arbitrarily categorising individual mobility into discrete groups—removing a role of researcher freedom.¹⁸ Second, we exploit the cardinality of our rank measurement. We view two people located at the 95th and 5th percentiles of the wealth distribution as further apart than either of those same people and a person at the median. In this way, our approach captures both the ordinal and cardinal information captured in ranks.

Relationship to K-means Clustering. An alternative procedure to construct our latent groups would be to employ the K-means clustering (Bonhomme and Manresa 2015; Bonhomme, Lamadon, and Manresa 2022). In this case, we would use each element of our vector as a data feature and partition the data by minimising the within cluster sum of squared errors. Conceptually, this is an alternative distance metric to (5), leading to our agglomerative algorithm and the K-means approach to produce distinct clusters.

However, the hierarchical clustering approach offers two important advantages for large administrative data-sets such as ours. First, K-means clustering requires solving

¹⁸That is not to say that Sequence Analysis tools are not a valuable approach to the analysis of categorical outcomes such as occupations.

an optimisation problem through local optimization techniques that often require many multi-start evaluations of the objective function for a given K (or G in our case). While the solution is typically fast for a given number of clusters and an initial guess of the partition, the computational demands become larger for data-sets such as ours.¹⁹ Second, the procedure must be repeated whenever the number of possible group changes—including to select the optimal number of groups or for robustness analysis. By contrast, once the dissimilarity measure is specified, hierarchical clustering recovers all optimal clusters for $G \in \{1, \dots, N\}$ simultaneously. [Hastie et al. \(2009\)](#) provide a further overview of alternative clustering techniques.

4.2. Inferring typical trajectories

In this application, we select $G = 6$ groups. This choice captures the key data patterns while minimising the cognitive burden for both us and the reader. In Appendix B we provide heuristic selection criteria from the computer science literature and the underlying dendrogram for our hierarchical clustering, and show results for alternative numbers of clusters, see Figures B.3 and B.4.

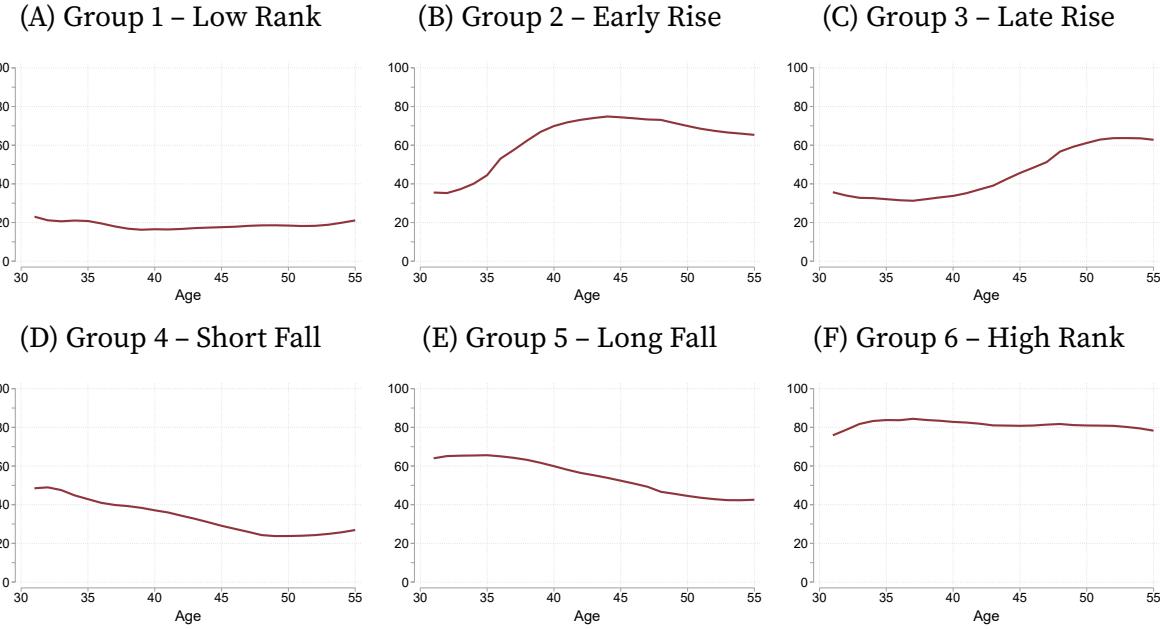
Figure 4 presents the age-profile of the mean rank for each group. The clustering algorithm we employ acts by grouping individuals on the basis of the similarity of their trajectories, so as to generate distinct average trajectories across groups, see (5). In this sense, we recover typical histories as the age-profile of the average rank in each group, μ_g . Table 1 presents the share of each group in the sample and their average ranks at selected ages.

The resulting groups are revealing of the patterns underlying wealth mobility. Of the six groups, there are two groups of individuals that (on average) maintain their position in the wealth distribution. Group 1 is comprised by “*low rank*” individuals

¹⁹Given the necessary data security requirements for using data-sets such as ours, it is not easy to parallelize away this problem.

FIGURE 4. Life Cycle Dynamics of Wealth Mobility:

Average Rank by Group



Notes: The figures plot the average wealth rank of the individuals in each clustered group against the cohort's average age. All individuals belong to the 1960-1964 birth cohort. The clusters are constructed from the balanced sample using hierarchical agglomerative clustering and Ward's method with a dissimilarity measure (5).

and group 6 by “*high rank*” individuals. These individuals maintain, on average, their place in the distribution. Another two groups represent individuals who rise through the wealth distribution. Group 2 has “*early risers*,” starting below the 40th percentile of the distribution and rising to the 75th percentile between the ages of 35 and 40, while group 3 has “*late risers*,” starting also below the 40th percentile and rising to the 60th percentile between the ages of 45 and 50. Finally, there are two groups of individuals who experience downward mobility. These are groups 4 and 5 formed by individuals experiencing a “*short*” and “*long*” falls, respectively.

Our results highlight that the rank an individual occupies in the wealth distribution is far from being constant over the life-cycle, instead displaying slow-moving patterns that are hard to capture with methods that focus on a shorter snapshot of wealth dynamics.

TABLE 1. Clusters

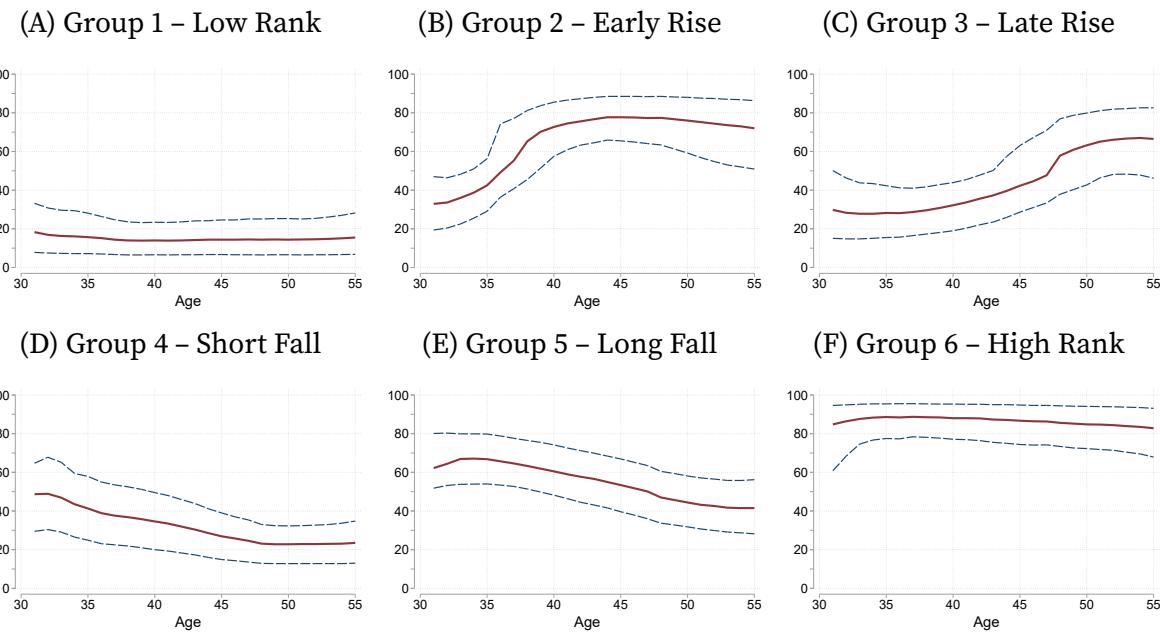
Label	Share	Average Rank (pct)		
		Age 35	Age 45	Age 55
1 Low Rank	19.83	21	18	21
2 Early Rise	9.39	44	74	65
3 Late Rise	17.88	32	46	63
4 Short Fall	19.72	43	29	27
5 Long Fall	20.01	66	52	43
6 High Rank	13.16	84	81	78

Notes: The table reports the shares of each clustered group along with the average wealth rank of the individuals in the group at ages 35, 45 and 55. All individuals belong to the 1960-1964 birth cohort. The clusters are constructed from the balanced sample using hierarchical agglomerative clustering and Ward's method with a dissimilarity measure (5).

The fact that only Group 1 and Group 6 show very stable rank patterns over the life-cycle highlights that, whilst there is large mobility within the middle class, being very poor and very rich are close to being permanent types.

Figure 5 explores the dispersion in the histories of individuals within each group by reporting the age-profile of the median, 25th, and 75th percentile of the wealth rank. The median tracks the mean rank in Figure 4 closely and the inter-quartile range is on average 25.5 percentiles and symmetric around the median. Further examining the distribution reveals other interesting patterns. For instance, the first decile of group 6 is comprised of individuals at the median of the wealth distribution at age 31, and who rise quickly as the cohort ages. This reveals that group 6 is not only composed by individuals who are always at the top of the wealth distribution, also including individuals who become wealthier early in life. This is similar to the trajectory of the ex-post wealthiest individuals studied in [Halvorsen, Hubmer, Ozkan, and Salgado \(2023\)](#), who focus on even wealthier individuals than our high rank group. Figure B.2 in Appendix B presents the age-profile of the 10th, and 90th percentiles.

**FIGURE 5. Life Cycle Dynamics of Wealth Mobility:
50th, 25th, and 75th Percentiles of Rank within Group**



Notes: The figures plot the 50th, 25th, and 75th percentiles of the distribution of wealth ranks in each clustered group against the cohort's average age. The continuous line corresponds to the 50th percentile. The dashed lines correspond to the 25th and 75th percentiles. All individuals belong to the 1960-1964 birth cohort. The clusters are constructed from the balanced sample using hierarchical agglomerative clustering and Ward's method with a dissimilarity measure (5).

4.3. Group characteristics

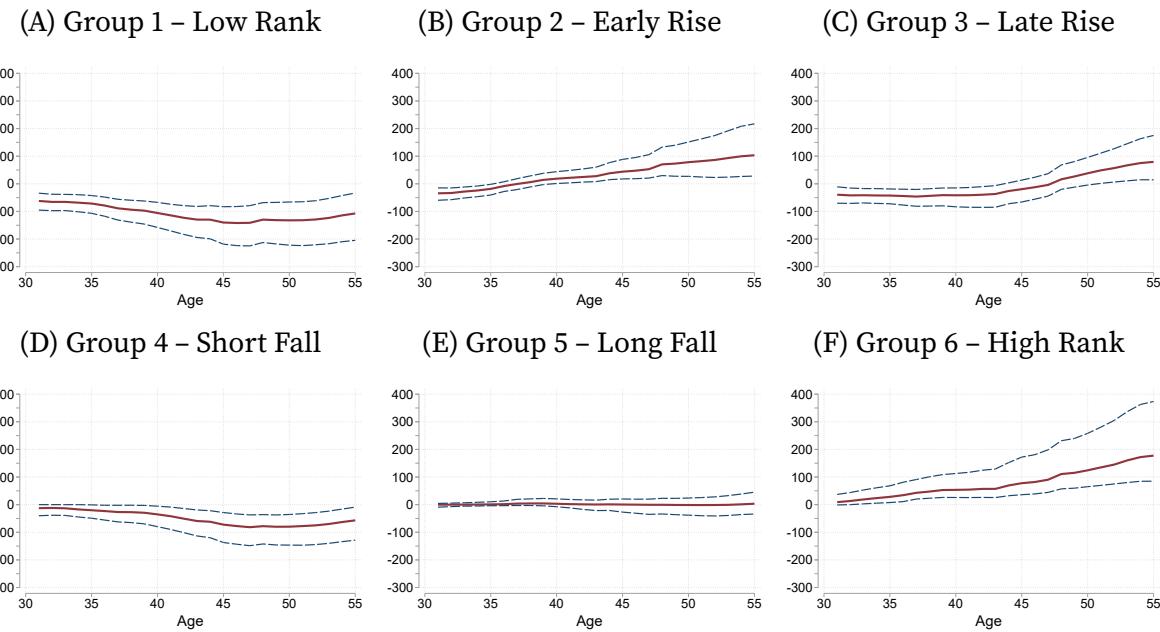
Below we explore the characteristics of the different groups. We look at the distribution of the level of wealth (net worth), assets, and debt for each group, as well as at differences in civil status and education.

In Figure 6 below, we plot the evolution of net worth within groups. One important point these graphs show is that falling down the wealth distribution does not necessarily imply individuals are dissaving and deaccumulating wealth. Instead, because ranks are comparative statements, individuals can fall down the distribution if they save, but do so at a slower rate than the rest of their cohort. For example, the net worth of individuals of group 5 does not change on average, while their rank declines. Similarly, maintaining one's rank in the distribution entails decreasing net worth in group 1 and increasing in group 6. These changes reveal that there is heterogeneity in the saving rate implied by rank dynamics across the wealth distribution.

Figure 7 shows how our measure of average net worth is decomposed into assets and debt. Figures B.5 and B.6 in Appendix B present details of the distribution of assets and debt, respectively, while Figure B.7 presents the aggregate leverage of each group. Zooming in on these broad components of net worth highlights that the underlying asset dynamics of these groups are more similar than net worth shows. The low rank (group 1) and falling individuals (groups 4 and 5) accumulate assets at much the same rate, but the key difference driving their overall net worth profiles is the accumulation of debt. In contrast, among the remaining three groups of early and late risers as well as the persistently high rank individuals (groups 2, 3 and 6, respectively) we see that markedly different accumulation of assets drives their rising net worth.

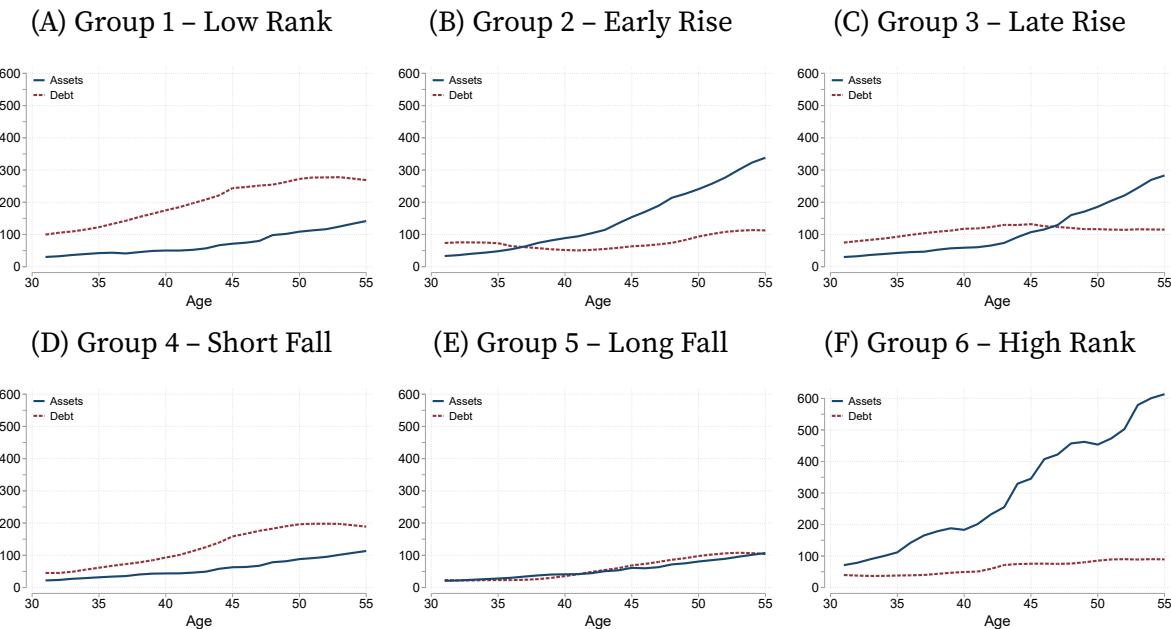
Given the importance differences in the asset and debt holdings of individuals we document above, we now turn to documenting important features of their portfolio in more detail. In Table 2 we describe important features of the portfolio composition of

FIGURE 6. Wealth by Group: 50th, 25th, and 75th Percentiles



Notes: The figures plot the 50th, 25th, and 75th percentiles of the distribution of wealth levels (net worth) in each clustered group against the cohort's average age. The continuous line corresponds to the 50th percentile. The dashed lines correspond to the 25th and 75th percentiles. All numbers are in thousands of 2019 U.S. dollars. All individuals belong to the 1960-1964 birth cohort. The clusters are constructed from the balanced sample using hierarchical agglomerative clustering and Ward's method with a dissimilarity measure (5).

FIGURE 7. Assets and Debt by Group



Notes: The figures plot the mean assets and debt in each clustered group against the cohort's average age. The continuous line corresponds to mean assets. The dashed line corresponds to mean debt. All numbers are in thousands of 2019 U.S. dollars. All individuals belong to the 1960-1964 birth cohort. The clusters are constructed from the balanced sample using hierarchical agglomerative clustering and Ward's method with a dissimilarity measure (5).

TABLE 2. Portfolio Differences Across Clusters

Group	Low Rank (1)	Early Rise (2)	Late Rise (3)	Short Fall (4)	Long Fall (5)	High Rank (6)
<i>A. home-ownership Rates (%)</i>						
Age 35	57.4	49.3	51.1	34.7	20.3	36.1
Age 45	67.4	67.4	63.0	52.6	40.8	61.3
Age 55	72.0	75.3	72.0	60.8	51.5	71.0
<i>B. Portfolios (Age 31)</i>						
Safe Assets	9.4	18.3	17.4	16.7	25.3	27.4
Property	49.9	41.3	40.1	37.6	27.2	16.0
Private Equity	12.5	14.5	13.2	16.0	21.0	31.9
<i>C. Portfolios (Age 35)</i>						
Safe Assets	9.1	16.8	14.7	13.5	21.4	22.0
Property	49.7	38.6	42.5	39.3	26.6	14.1
Private Equity	12.0	16.6	13.1	15.7	20.6	32.1
<i>D. Portfolios (Age 45)</i>						
Safe Assets	14.7	20.7	18.5	16.3	19.5	18.0
Property	44.85	23.9	30.4	40.0	31.4	11.1
Private Equity	12.5	26.2	20.1	14.7	21.7	42.2
<i>E. Portfolios (Age 55)</i>						
Safe Assets	10.2	15.3	13.8	12.4	16.7	15.6
Property	58.0	33.6	40.0	55.0	48.8	21.1
Private Equity	12.4	33.2	27.9	10.7	12.5	46.7

Notes: The table reports key features of household portfolios for each clustered group at different ages. We construct the gross portfolio shares (excluding debt) in panels B through E.

individuals in each cluster, as well as their home-ownership rates at different ages. To minimise the role of outliers, we report the aggregate gross asset share for each group.

Panel A of Table 2 shows the home-ownership rate at key ages for each of our groups. We define home-ownership on the basis of primary residence as the Norwegian tax code allows us to identify secondary and leisure properties separately from the primary residence. We focus first on home-ownership rate as it is typically the largest part of household portfolios and is typically identified as a marker of the financial independence of individuals. For all of our groups, home-ownership rates exhibit a strong life-cycle profile — rising with age. However, the rise in the aggregate rate masks important heterogeneity across these groups. At age 35, persistently low rank individuals (group 1) have the largest home-ownership rates, over 10% higher than either of the groups of rising individuals (groups 2 and 3). While the fallers (groups 4 and 5) have the lowest overall home-ownership rates, the high rank individuals (group 6) are closest to this group at age 35. Their home-ownership rate, however, almost doubles over the next 20 years as they catch up to the risers and the low rank group. The majority of these early differences have been eliminated by age 45, but the falling individuals (groups 4 and 5) never achieve the home-ownership rates of their peers. They too dramatically increase their home-ownership rates, but are still 10-20 years behind the rest of their cohort on average.

Turning next to Panels B through E of Table 2, we report the average portfolio shares of Safe Assets, Property and Private Equity (the largest and, ex-post, most interesting features of individual portfolios) for each of our clusters. In Panel B, we report these at age 31 (the earliest age we are able to calculate them), in Panel C at age 35, Panel D at age 45 and, finally, at age 55 in Panel E (the last age we are able to calculate them). We find important differences in both the initial composition of portfolios, but also their ex-post dynamics as they age. First, consistent with the evidence above on the extensive margin of home-ownership we see that property is the largest part of the

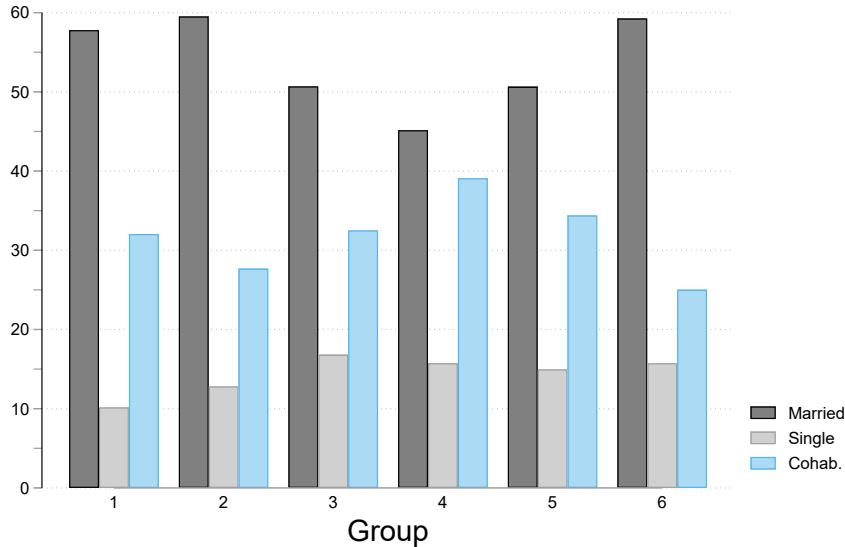
individual portfolio of the persistently low rank individuals (group 1) even as they age. Similarly, we see the persistently high rank individuals (group 6) also hold very different portfolios at age 31. They hold less wealth in housing and a much larger share of their wealth in private equity. These patterns also persist as they age.²⁰

Focusing on groups 2-5, the risers and fallers, we find smaller differences in portfolios at age 31. In particular, the average portfolio composition of both rising groups and the short fall group are very close. The long fall group (group 5) hold portfolios in between the short fallers and the persistently high rank. As these groups age, however, we see markedly different dynamics. First, while initial differences between the fallers persist into their 40s, by age 55 their portfolios have converged and resembles the persistently low rank (group 1). Second, and in stark contrast, despite their initial similarities at age 31 we see the portfolios of early risers (group 2) beginning to look different by age 35 and very different by age 45. For their late rising peers (group 3) these differences do not show up until their forties and are cemented by age 55 when the portfolio of both rising groups has stabilized. Yet most striking is the reason for this divergence and convergence: the role of private equity. The initially and permanently wealthy have far larger portfolio shares in private equity. Importantly, we also provide new evidence that this is true of those who rise through the wealth distribution. Both their eventual position and the timing of their rise is correlated with the increasing portfolio share of private equity. This is an important and new fact, distinct from the role private equity plays in driving both wealth concentration and the existence of the super-wealthy in the top tenth- and hundredth-of-a-percentile in the distribution. The super-wealth are, in fact, charting the same portfolio course in microcosm as the moderately well off.

We also track the civil status of individuals in each group. We use civil status, which accounts for cohabitation identified by the tax authority in addition to marriage due

²⁰While we show these at a snapshot of ages and a subset of financial instruments, in Figure B.9 in Appendix B we report the full portfolio composition at each age for each group.

FIGURE 8. Marriage and Cohabitation at Age 55



Notes: The figure shows the composition of each clustered group in terms of the civil status of their individuals at age 55. For each group, the figure shows the share of individuals who are married (dark gray), the share who are single (light gray), and the share who are cohabiting (light blue), in that order. All individuals belong to the 1960-1964 birth cohort. The clusters are constructed from the balanced sample using hierarchical agglomerative clustering and Ward's method with a dissimilarity measure (5).

to the high prevalence of the former among the cohort we consider Figure 8 shows that, by age 55, the marriage rate is U-shaped across groups, with groups 1, 2, and 6 having higher marriage rates and groups 3, 4, and 5 having lower ones. This pattern is reversed for the cohabitation rate. Figure B.8 (in Appendix B) the life-cycle of the marriage, cohabitation, and single rates. Notably, we see larger declines in the share married for groups 4 and 5, the falling individuals, than other groups. Unfortunately, however, we cannot cleanly attribute causality in either direction. One possibility is that individual net worth falls as divorces increase due to the division of assets or payment of spousal support, on the other hand it may be that falls through the wealth distribution and increasing dissolution of marriage are both the result of economic distress.

5. Ex-ante Analysis: Predicting patterns of wealth mobility

The groups of individuals we recover from our analysis above differ in their ex-post trajectories across the wealth distribution. Despite these differences, we showed that some of these groups have similar average ranks and portfolio composition at age 31. We now dig deeper into individuals' ex-ante characteristics to determine whether this ex-post heterogeneity across groups is predicted by observed factors early in life.

Formally, we estimate group-assignments using a multinomial logit specification controlling for parental wealth, sex, education level and place of birth:²¹

$$Pr(g_i = j) = F\left(\beta_0^j + \alpha_{q(i)}^j + \delta_{\text{male}(i)}^j + \lambda_{\text{bcounty}(i)}^j + \gamma_{\text{educ}(i)}^j + \mu_{\text{subj}(i)}^j\right), \quad (6)$$

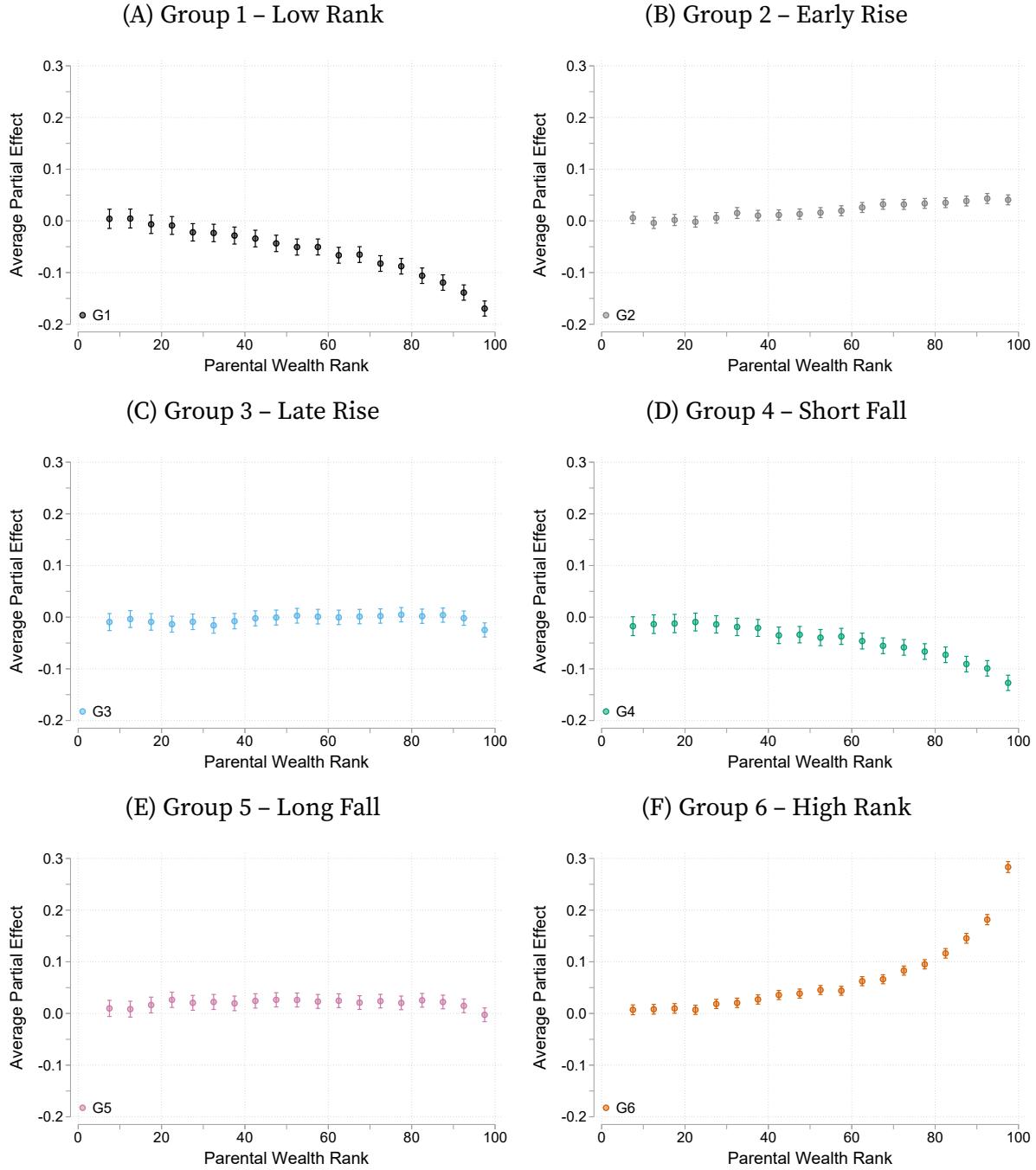
where $F(\cdot)$ denotes the logit transformation. We include parental background by incorporating ventile of parental wealth specific fixed effects, $\alpha_{q(i)}^j$. We measure parental wealth at the beginning of our sample (1993). We also include a gender specific effect, $\delta_{\text{male}(i)}^j$, and a birth county fixed effect, $\lambda_{\text{bcounty}(i)}^j$, that captures potential place-based differences between the Oslo Metropolitan Area, Other major cities, and rural regions. Finally, we include education level fixed effects (post-compulsory high school, technical college, undergraduate, postgraduate and PhD), $\gamma_{\text{educ}(i)}^j$, and subject specific fixed effects, $\mu_{\text{subj}(i)}^j$, for those with undergraduate or above education.²² We focus on these characteristics due to longstanding interest in, and recent evidence on, the roles of parental background, gender, education, and place-based effects.

Figure 9 reports the estimated average partial effects of parent's wealth rank on predicted group assignment and their 95% confidence interval. This shows that the

²¹Note that our ordering of groups is arbitrary. While we adopt a convenient ordering based on average rank at age 31, we show that these groups switch position as they age. We therefore do not impose an ordinal ranking across groups.

²²We aggregate degree specific education codes into six categories: arts and humanities; business, economics and agricultural management; computer science and engineering; natural sciences; health; education specialists.

FIGURE 9. Parental Wealth Rank and the Probability of Group Assignment



Notes: The figures plot the average partial effect of Parental Wealth (measured in 1993) relative to being born to parents in the bottom ventile of the distribution. We construct the average partial effect by integrating over the empirical joint distribution of other covariates. We report point estimates separately for each outcome, the probability of being assigned to each of our six groups, along with their 95% confidence intervals.

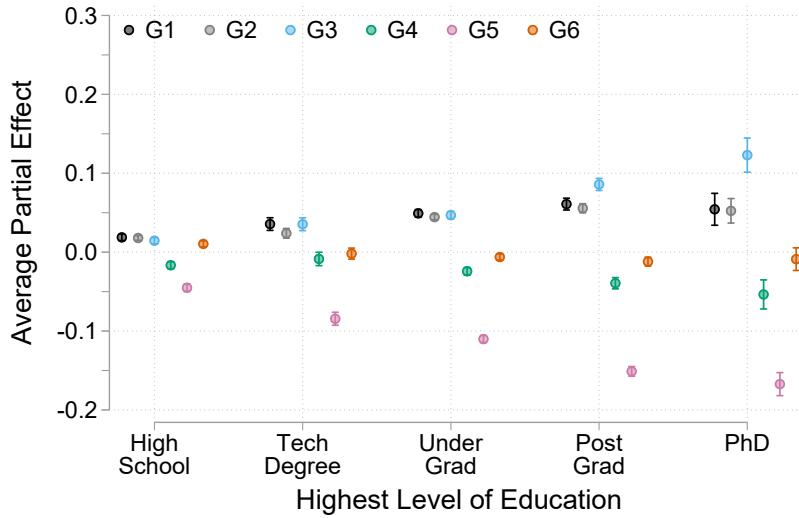
role of parental background is highly non-linear. Comparing an individual with parent's in the bottom ventile of their within cohort wealth distribution with somebody who's parents are in the top ventile sees an almost 20 percentage point decrease in their probability of belonging to group 1 (Panel A) – the persistently poor. In contrast, those with parent's at the median have less than a 6 percentage point decrease in their probability of belonging to group 1. Similarly to our within generation findings, the inter-generational effect of parental background helps insure those at the top from large falls.

While we see no statistically significant effect of parental background on belonging to the late risers (Group 3, Panel C), we see the probability of rising earlier in the life cycle (Group 2, Panel B) is increasing in parental background and statistically significant. These effects are smaller in magnitude than the effect on belonging to the persistently poor, but given the large differences in lifetime outcomes they still show an economically significant relationship. Turning to the fallers (Groups 4 and 5, Panels D and E) we find small effects on the long fallers, but statistically significant decreases in the probability of belonging to the short fall group as we increase parental background. These results are consistent with an important initial role for parents in placing their children in the distribution by age 31.

The largest effects, however, are for belonging to the persistently wealth group 6 (Panel F). An individual with parent's in the top ventile of the wealth distribution is almost 30 percentage points more likely to belong to this group than an individual with parents in the bottom quartile of the wealth distribution. Again, the effects are highly non-linear as we move up the parental wealth distribution.

Taken together, these results (which are consistent with the evidence on inter-generational wealth mobility measured by rank-rank persistence at fixed ages) highlight that parental background plays an important role in shaping the life cycle dynamics of individual wealth mobility. Not only does it predict that an individual with

FIGURE 10. Educational Attainment and the Probability of Group Assignment



Notes: The figures plot the average partial effect of each education level relative to compulsory schooling age. We construct the average partial effect by integrating over the empirical joint distribution of other covariates. We report point estimates separately for each outcome, the probability of being assigned to each of our six groups, along with their 95% confidence intervals. G1 corresponds to Low Rank, G2 corresponds to Early Rise, G3 corresponds to Late Rise, G4 corresponds to Short Fall, G5 corresponds to Long Fall, and G6 corresponds to High Rank.

rich (poor) parents is on average richer (poorer) at any given age, but it is also able to determine the persistently rich (poor) individuals over the life cycle. This suggests that single-age snapshots of inter-generational wealth persistence understate inequality of opportunity. However, while they are important at the extremes of the lifetime wealth distribution, parental background is less able to disentangle the churning middle classes.

We now turn to understanding the role of education, which we report in Figure 10. Increasing education level is associated with three broad patterns. First, those with higher education are less likely to belong to either falling group. For example, an individual with a PhD is 5 (17) percentage points less likely to belong to group 4 (group 5) than an individual who only has compulsory schooling level, which may reflect the segregated job-market for these individuals or the delayed accumulation of wealth as a

graduate student.

Second, we find increased education is associated with an increased probability of belonging to a rising group (groups 2 and 3) and the persistently poor group (group 1). The effect of belonging to the persistently poor or early risers stabilizes at around 5 percentage points with an undergraduate degree and is approximately constant for higher levels of education. In contrast, additional education beyond the undergraduate level has an economically and statistically significant increase in the probability of belonging to the late rising group — a 4 percentage point increase for each level of advanced degree held.

Finally, we find no economically significant effect of education on belonging to the persistently wealthy group. While we do find some statistically significant relationships (due to our large sample size) these are on the order of magnitude of a single percentage point. Even for our smallest education group, the PhDs, we are able to rule out effects larger than 2 percentage points with 95% confidence.

These results highlight an important role for parental background and education, however, they do not provide information on the joint explanatory power of the ex-ante characteristics we include. To do this we explore how each set of ex-ante covariates helps to explain the variation across groups. We quantify this using two measures. Our preferred measure is a Distance Weighted Classification Rate $\in [0, 1]$ given by:

$$1 - \frac{\sum_{i=1}^N \sum_{k=1}^G \widehat{Pr}(g = k | X_i) d(g(i), k)}{\sum_{i=1}^N \sum_{k=1}^G \widehat{Pr}(g = k) d(g(i), k)}, \quad (7)$$

where $d(g, g')$ corresponds to Ward's distance metric in equation (5). This measure, which is bounded between 0 and 1, corresponds to the average implied distance between an individual's true group and their predicted groups weighted against a naive predictor that uses a homogeneous random assignment. As all of the distances between disjoint groups are positive, the numerator of the fraction in equation (7) can be interpreted

similarly to the residual sum of squares in the coefficient of determination. Similarly, the denominator takes on an interpretation as the total sum of squares. Consequently, a value of unity implies perfect classification while a value of zero implies the covariates contain no information. Relative to a correct classification rate (which we discuss in detail below), this measure is able to capture the intuitive idea that those who start at the bottom of the distribution and rise in their 30s (group 2) are more different from those who fall (groups 4 and 5) than they are from those who start at a similar level, but rise later (group 3).

Table 3 reports the total, and individual, contribution from our four groups of ex-ante regressors. To decompose the total contribution into individual contributions We use a Shapley-Owen decomposition (following [Shorrocks 2013](#)) which we report under the partial contribution heading.²³

Consistent with the average partial effects we report above, we find that parental background and sex explain the majority of the variation we can explain. Although we found statistically significant effects of increased education on some group membership probabilities, we find education and birth place effects explain only a small fraction of the variation – less than 1% each. We interpret this quantification as evidence that observable ex-ante characteristics are important. However, this also shows that there is substantial variation in outcomes over the life cycle which cannot be captured by our set of ex-ante observables.

Although our average partial effects suggest an important role for family background, there is substantial variation in the full history of lifetime outcomes that it does not capture. This does not rule out that individuals in our sample are ex-ante heterogeneous along unobservable margins. However, the results in section above show that at age 31 differences among risers and fallers are small.

²³We describe the Shapley-Owen-Shorrocks decomposition in more detail in Appendix C. We select this approach as it allows us to calculate a single value per covariate category that is both permutation invariant and additively decomposable in our multivariate setting.

TABLE 3. Share of Distance Variation Explained by Variable (pp)

Total Contribution	Partial Contribution			
	Parent	Sex	Education	Birth Place
6.44	2.32	3.38	0.65	0.08

Notes: The table reports the distance weighted probability of belonging to an individual's true group (corresponding to the measure in equation 7 along with the contribution of this prediction that is attributable to the covariates in our estimated multinomial logit classifier. We describe the estimated model fully in equation (6). we use a Shapley-Owen decomposition which averages across permutations of decompositions and sums to the total contribution. Appendix C provides more details on the implementation.

This interpretation is consistent with R-squared values reported in inter-generational estimates of the rank correlation in wealth²⁴. However, it also emphasises the value in our clustering approach which is able to exploit the full life cycle history of individuals. In doing so, it reveals a low dimensional representation that is broadly orthogonal to ex-ante characteristics. It thus captures important features of the data that we could not have known ex-ante.

To better understand the predictive content of these variables, we now turn to analysing the Correct Classification Rate, $\in [0, 1]$, which we formally define as:

$$\frac{1}{N} \sum_{i=1}^N \sum_{k=1}^G \widehat{Pr}(g = k | X_i) \mathbb{1}[g(i) = k] \quad (8)$$

Table 4 reports the same decomposition as Table 3 using this alternative outcome metric. Additionally, it shows how each group of covariates contributes to the correct classification of individuals conditional on their true group.

Relative to the results we report above, this reveals two important additional patterns. First, while on average the discriminating power of Birth Place and Education are lower,

²⁴Specifically for Norway reported in (Fagereng et al. 2020) and Denmark (Boserup, Kopczuk, and Kreiner 2018). Value for the US in Charles and Hurst (2003) are slightly higher

TABLE 4. Share of Individuals Correctly Classified

Group	Full Model	Total Contribution ^a	Partial Contribution			
			Parent	Sex	Education	Birth Place
1	25.82	5.87	1.13	4.42	0.26	0.06
2	10.34	0.93	0.27	0.09	0.52	0.05
3	18.71	0.86	0.05	0.28	0.50	0.04
4	20.46	0.88	0.56	0.02	0.22	0.08
5	26.18	6.18	0.14	4.44	1.51	0.09
6	17.57	4.34	3.87	0.24	0.21	0.02
All	21.02	3.39	0.90	1.86	0.56	0.06

Notes: The table reports the predicted probability of belonging to an individual's true group (corresponding to the measure in equation 8 along with the contribution of this prediction that is attributable to the covariates in our estimated multinomial logit classifier. We report values for the overall population contribution and conditional on their true group. We describe the estimated model fully in equation (6). We use a Shapley-Owen decomposition which averages across permutations of decompositions and sums to the total contribution. Appendix C provides more details on the implementation.

^aThe share of individuals correctly classified by random classification is 17.63%.

their ability to classify individuals is much more consistently spread across groups. Second, parental background is most effective at correctly classifying those at the extremes of the distribution. As figure 9 shows, having high wealth parents has a large and statistically significant decrease in the probability an individual belongs to groups 1 and 4 (the persistently poor and short-fallers, respectively). In contrast, those individuals are more likely to belong to group 6 –the persistently wealthy – an effect that is both economically and statistically significant. However, despite the importance of parental wealth for determining positions at the tails of the distribution, it only has limited informational content for predicting those who rise or fall through the churn in the middle of the distribution. We view this as highlighting an important notion of equality of opportunity: extreme comparisons point to *inequality* of opportunity, but there is more *equality* of opportunity in the middle of the distribution.

6. Directions for Future Research

In Progress

References

- Adermon, Adrian, Mikael Lindahl, and Mårten Palme. 2021. “Dynastic Human Capital, Inequality, and Intergenerational Mobility.” *American Economic Review* 111 (5):1523–48. URL <https://doi.org/10.1257/aer.20190553>.
- Ahn, Hie Joo, Bart Hobijn, and Ayşegül Şahin. 2023. “The Dual U.S. Labor Market Uncovered.” URL https://www.dropbox.com/s/7vosrziqoq9yzdg/AhnHobijnSahin_EFG.pdf?raw=1. Working Paper.
- Arellano, Manuel, Richard Blundell, and Stéphane Bonhomme. 2017. “Earnings and consumption dynamics: a nonlinear panel data framework.” *Econometrica* 85 (3):693–734. URL <https://doi.org/10.3982/ECTA13795>.
- Bhuller, Manudeep, Magne Mogstad, and Kjell G Salvanes. 2017. “Life-cycle earnings, education premiums, and internal rates of return.” *Journal of Labor Economics* 35 (4):993–1030.
- Black, Sandra E, Paul J Devereux, Fanny Landaud, and Kjell G Salvanes. 2020. “Where Does Wealth Come From?” Working Paper 28239, National Bureau of Economic Research. URL <http://doi.org/10.3386/w28239>.
- Black, Sandra E, Paul J Devereux, and Kjell G Salvanes. 2005. “Why the apple doesn’t fall far: Understanding intergenerational transmission of human capital.” *American economic review* 95 (1):437–449.
- Bolt, Uta, Eric French, Jamie Hentall MacCuish, and Cormac O’Dea. 2021. “The intergenerational elasticity of earnings: exploring the mechanisms.” Tech. Rep. Discussion Paper 15975, CEPR. URL <https://cepr.org/publications/dp15975>.
- Bonhomme, Stéphane, Thibaut Lamadon, and Elena Manresa. 2022. “Discretizing Unobserved Heterogeneity.” *Econometrica* 90 (2):625–643. URL <https://doi.org/10.3982/ECTA15238>.
- Bonhomme, Stéphane and Elena Manresa. 2015. “Grouped patterns of heterogeneity in panel data.” *Econometrica* 83 (3):1147–1184. URL <https://doi.org/10.3982/ECTA11319>.
- Boserup, Simon Halphen, Wojciech Kopczuk, and Claus Thustrup Kreiner. 2018. “Born with a Silver Spoon? Danish Evidence on Wealth Inequality in Childhood.” *The Economic Journal* 128 (612):F514–F544. URL <https://doi.org/10.1111/eco.12496>.
- Charles, Kerwin Kofi and Erik Hurst. 2003. “The correlation of wealth across generations.” *Journal of political Economy* 111 (6):1155–1182.
- Chetty, Raj, David Grusky, Nathaniel Hendren, Maximilian Hell, Robert Manduca, and Jimmy Narang. 2017. “The Fading American Dream: Trends in Absolute Income Mobility Since 1940.” *Science* URL <https://doi.org/10.1126/science.aal4617>.
- Chetty, Raj, Nathaniel Hendren, Patrick Kline, Emmanuel Saez, and Nicholas Turner. 2014. “Is the United States Still a Land of Opportunity? Recent Trends in Intergenerational Mobility.” *American Economic Review* 104 (5):141–47. URL <https://doi.org/10.1257/aer.104.5.141>.
- Crow, Edwin L. and M. M. Siddiqui. 1967. “Robust Estimation of Location.” *Journal of the American Statistical Association* 62 (318):353–389. URL <https://doi.org/10.1080/01621459.1967.10482914>.
- Davies, James B. and Anthony F. Shorrocks. 2000. “The distribution of wealth.” In *Handbook of Income Distribution*, vol. 1, edited by Anthony B. Atkinson and François Bourguignon, chap. 11. Elsevier, 605–675. URL [https://doi.org/10.1016/S1574-0056\(00\)80014-7](https://doi.org/10.1016/S1574-0056(00)80014-7).

- De Nardi, Mariacristina, Giulio Fella, Marike Knoef, Gonzalo Paz Pardo, and Raun Van Ooijen. 2021. “Family and government insurance: Wage, earnings, and income risks in the Netherlands and the US.” *Journal of Public Economics* 193:104327. URL <https://doi.org/10.1016/j.jpubeco.2020.104327>.
- De Nardi, Mariacristina, Giulio Fella, and Gonzalo Paz Pardo. 2019. “Nonlinear Household Earnings Dynamics, Self-Insurance, and Welfare.” *Journal of the European Economic Association* 18 (2):890–926. URL <https://doi.org/10.1093/jeea/jvz010>.
- Dijkstra, Wil and Toon Taris. 1995. “Measuring the Agreement between Sequences.” *Sociological Methods & Research* 24 (2):214–231. URL <https://doi.org/10.1177/0049124195024002004>.
- Dlouhy, Katja and Torsten Biemann. 2015. “Optimal matching analysis in career research: A review and some best-practice recommendations.” *Journal of Vocational Behavior* 90:163–173. URL <https://doi.org/10.1016/j.jvb.2015.04.005>.
- Fagereng, Andreas, Luigi Guiso, Davide Malacrino, and Luigi Pistaferri. 2020. “Heterogeneity and Persistence in Returns to Wealth.” *Econometrica* 88 (1):115–170. URL <https://doi.org/10.3982/ECTA14835>.
- Fagereng, Andreas, Martin B. Holm, and Gisle J. Natvik. 2021. “MPC Heterogeneity and Household Balance Sheets.” *American Economic Journal: Macroeconomics* 13 (4):1–54. URL <https://doi.org/10.1257/mac.20190211>.
- Fagereng, Andreas, Martin Blomhoff Holm, and Kjersti Næss Torstensen. 2020. “Housing wealth in Norway, 1993–2015.” *Journal of Economic and Social Measurement* 45 (1):65–81. URL <https://doi.org/10.3233/JEM-200471>.
- Fagereng, Andreas, Magne Mogstad, and Marte Rønning. 2021. “Why Do Wealthy Parents Have Wealthy Children?” *Journal of Political Economy* 129 (3):703–756. URL <https://doi.org/10.1086/712446>.
- Gale, William G., Hilary Gelfond, Jason J. Fichtner, and Benjamin H. Harris. 2021. “The Wealth of Generations, With Special Attention to the Millennials.” In *Measuring Distribution and Mobility of Income and Wealth*, edited by Raj Chetty, John N. Friedman, Janet C. Gornick, Barry Johnson, and Arthur Kennickell. University of Chicago Press, 145–174. URL <http://www.nber.org/chapters/c14445>.
- Gregory, Victoria, Guido Menzio, and David G Wiczer. 2021. “The Alpha Beta Gamma of the Labor Market.” Working Paper 28663, National Bureau of Economic Research. URL <http://www.nber.org/papers/w28663>.
- Guvenen, Fatih, Greg Kaplan, Jae Song, and Justin Weidner. 2022. “Lifetime Earnings in the United States over Six Decades.” *American Economic Journal: Applied Economics* 14 (4):446–79. URL <https://doi.org/10.1257/app.20190489>.
- Guvenen, Fatih, Fatih Karahan, Serdar Ozkan, and Jae Song. 2021. “What Do Data on Millions of U.S. Workers Reveal About Lifecycle Earnings Dynamics?” *Econometrica* 89 (5):2303–2339. URL <https://doi.org/10.3982/ECTA14603>.
- Guvenen, Fatih, Luigi Pistaferri, and Giovanni L. Violante. 2022. “Global trends in income inequality and income dynamics: New insights from GRID.” *Quantitative Economics* 13 (4):1321–1360. URL <https://doi.org/10.3982/QE2260>.
- Haider, Steven and Gary Solon. 2006. “Life-Cycle Variation in the Association between Current

- and Lifetime Earnings.” *American Economic Review* 96 (4):1308–1320. URL <https://doi.org/10.1257/aer.96.4.1308>.
- Halvorsen, Elin, Joachim Hubmer, Serdar Ozkan, and Sergio Salgado. 2023. “Why are the wealthiest so wealthy?” URL https://www.dropbox.com/s/dl/pvpz8t49o8r87fp/OHSH_Wealth_2023.pdf. Working Paper.
- Halvorsen, Elin, Serdar Ozkan, and Sergio Salgado. 2022. “Earnings dynamics and its intergenerational transmission: Evidence from Norway.” *Quantitative Economics* 13 (4):1707–1746. URL <https://doi.org/10.3982/QE1849>.
- Hastie, Trevor, Robert Tibshirani, Jerome H Friedman, and Jerome H Friedman. 2009. *The elements of statistical learning: data mining, inference, and prediction*, vol. 2. Springer.
- Humphries, John Eric. 2021. “The Causes and Consequences of Self-Employment over the Life Cycle.” URL https://johnericumphries.com/Humphries_SelfEmployment.pdf. Working Paper.
- Karahan, Fatih and Serdar Ozkan. 2013. “On the persistence of income shocks over the life cycle: Evidence, theory, and implications.” *Review of Economic Dynamics* 16 (3):452–476. URL <https://doi.org/10.1016/j.red.2012.08.001>.
- Kelley, Truman Lee. 1947. *Fundamentals of statistics*. Harvard University Press.
- Lentz, Rasmus, Suphanit Piyapromdee, and Jean-Marc Robin. 2022. “The Anatomy of Sorting – Evidence from Danish Data.” URL https://www.dropbox.com/s/l6uplfj99y226nl/LPR_MobilityWages.pdf?raw=1. Working Paper.
- McKinney, Kevin L., John M. Abowd, and Hubert P. Janicki. 2022. “U.S. long-term earnings outcomes by sex, race, ethnicity, and place of birth.” *Quantitative Economics* 13 (4):1879–1945. URL <https://doi.org/10.3982/QE1908>.
- McVicar, Duncan and Michael Anyadike-Danes. 2002. “Predicting successful and unsuccessful transitions from school to work by using sequence methods.” *Journal of the Royal Statistical Society: Series A (Statistics in Society)* 165 (2):317–334. URL <https://doi.org/10.1111/1467-985X.00641>.
- Paz Pardo, Gonzalo. 2022. “Homeownership and portfolio choice over the generations.” URL https://gonzalopazpardo.github.io/jm/PazPardo_Homeownership.pdf. Working Paper.
- Shorrocks, Anthony F. 2013. “Decomposition procedures for distributional analysis: a unified framework based on the Shapley value.” *Journal of Economic Inequality* 11 (1):99.
- Solon, Gary. 1992. “Intergenerational income mobility in the United States.” *The American Economic Review* 82 (3):393–408. URL <http://www.jstor.org/stable/2117312>.
- Vermeulen, Philip. 2016. “Estimating the Top Tail of the Wealth Distribution.” *American Economic Review* 106 (5):646–50. URL <https://doi.org/10.1257/aer.p20161021>.

Appendix A. Additional Results on the Dynamics of Ranks

Table A.1 presents the levels of wealth percentiles for individuals in our cohort of interest for the years 1993, 1997, 2007, and 2017. These years correspond to the individuals having an average age of 31, 35, 45, and 55.

TABLE A.1. Wealth Percentiles by Age for Main Cohort

Age	p10	p25	p50	p75	p90
31	-\$86k	-\$48k	-\$11k	\$2 k	\$19k
35	-\$99k	-\$53k	-\$9k	\$6k	\$36k
45	-\$232k	-\$106k	-\$4k	\$49k	\$132k
55	-\$58k	-\$79k	\$42k	\$195k	\$459k

Notes: The table presents selected percentiles of wealth by age for the 1960-1964 birth cohort. All numbers are in thousands of 2019 U.S. Dollars.

A.1. Persistence in wealth ranks

Figure A.1 presents the persistence in wealth rank over the life cycle measured by their h^{th} -order autoregressive coefficient for our selected sample that focuses on the 1960-1964 cohort and compares it with the results presented in Figure 1 for all available cohorts. The figure confirms that the age-profile of the autoregressive coefficient for our cohort of interest is comparable to that of the population at large for the ages in which we observe them.

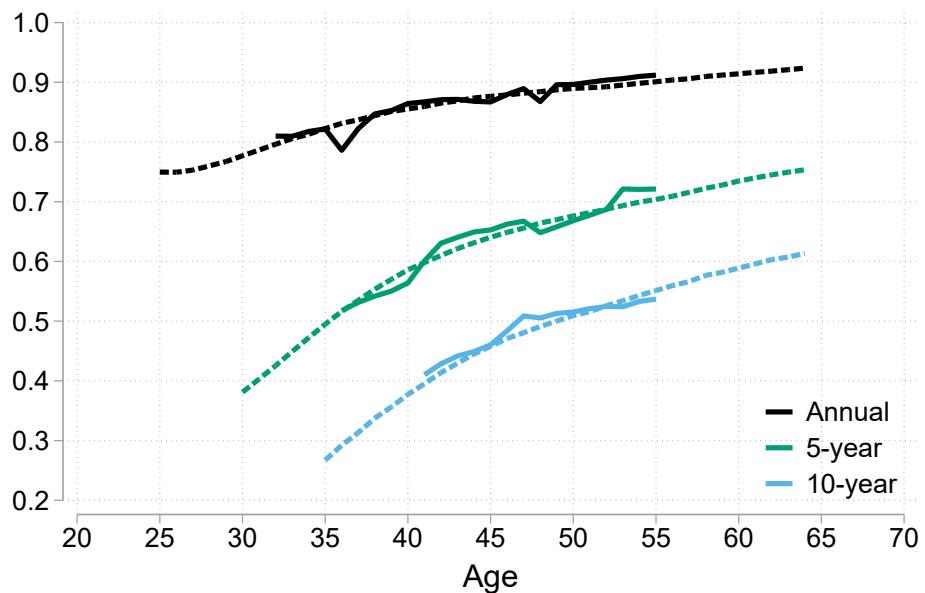
A.2. The distribution of change in wealth ranks

Figure A.2 complements figure 3 by presenting the distribution of rank changes for a ten year horizon. The distribution of changes is wider than for shorter horizons but close to the distribution at a five year horizon.

Figure A.3 shows the average change in ranks conditional on previous rank. A flat line would represent no change in ranks and, absent risk, perfect persistence in wealth positions. The less persistent wealth ranks are, the more negative will be the slope of this curve; positive slopes are ruled out by the definition of rank. Looking first at the dark annual lines, we observe a relatively flat pattern, consistent with very little changes year-on-year in wealth ranks. However, as the horizon grows we see more evidence of mean reversion. At age 31, for example, those in the bottom decile see their rank increase by 7 percentiles on average over a single year and by almost 20 percentiles over half a decade. This still places them well below the median, which suggests that there is some (upwards) wealth mobility, but it is not very strong.

As we move up the wealth distribution, the expected annual growth in their rank falls to zero, before becoming negative in the top half of the distribution. Comparing

FIGURE A.1. Persistence in Wealth Rank



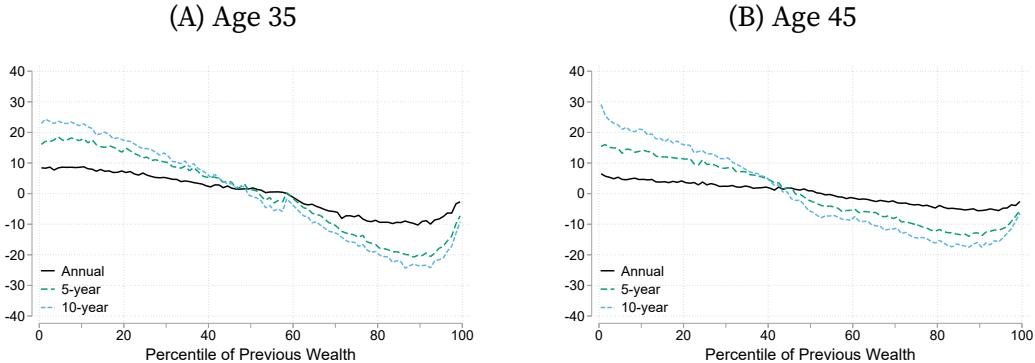
Notes: The figure plots the persistence measure (equation 2) against cohort's average age. The dashed lines reproduce Figure 1 and uses information of all birth cohorts included in our data. The continuous lines use only data from the 1960-1964 birth cohort and is therefore only possible to compute the persistence measure for a narrower window of ages. The dark line corresponds to a one year horizon ($h = 1$). The green line corresponds to a five year horizon ($h = 5$). The blue line corresponds to a ten year horizon ($h = 10$).

FIGURE A.2. Distribution of 10 Year Rank Changes



Notes: The figures plot the 10th, 25th, 50th, 75th, and p90th percentiles of distribution of 10 year wealth rank changes of individuals in the 1960-1964 birth cohort conditional on their rank at ages 35, left panel, and 45, right panel.

FIGURE A.3. Average Rank Changes h Years Ahead



Notes: The figures plot the average change in wealth rank of individuals in the 1960-1964 birth cohort conditional on their rank at ages 35, panel 2A, and 45, panel 2B. The dark-continuous line corresponds to a one year horizon ($h = 1$). The dashed line corresponds to a five year horizon ($h = 5$). The short-dashed line corresponds to a ten year horizon ($h = 10$).

across panels, the extent of this mean reversion declines as individuals age. While the relationship is monotone across the bottom 90 percent of the wealth distribution, the pattern reverses in the top decile. While those at the top are still more likely to fall to lower ranks (in part the role of truncation), the average fall is much smaller than for individuals in the next decile. This striking feature of the data is robust to examining different ages and is more extreme for longer horizons. This suggests that there is persistent heterogeneity across wealth groups, which we can capture with our clustering approach.

However, mean changes can have very different implications depending on their dispersion and in general their higher order moments. In Figure A.5 we plot the standard deviation, skewness, and kurtosis of the change in ranks. We measure skewness using [Kelley \(1947\)](#)'s measure of asymmetry,

$$S_{\mathcal{K}} = \frac{(p90 - p50) - (p50 - p10)}{p90 - p10}, \quad (\text{A.1})$$

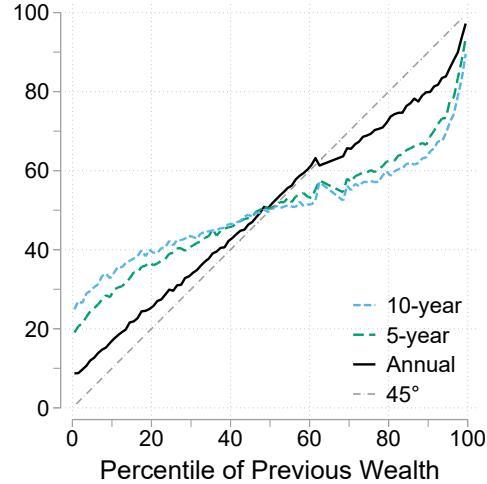
and kurtosis using [Crow and Siddiqui \(1967\)](#)'s measure,

$$\mathcal{K}_{\text{CS}} = \frac{p97.5 - p2.5}{p75 - p25}. \quad (\text{A.2})$$

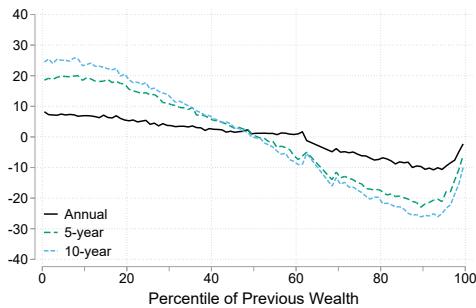
These quantile-based estimators of skewness and kurtosis are easy to interpret and robust to outliers, they have been commonly used in the literature (see, e.g., [Guvenen, Pistaferri, and Violante 2022](#)). The Kelley skewness is positive (right skewness) if the probability mass between the median and the top decile exceeds the probability mass between the median and the bottom decile, while the Crow-Siddiqui kurtosis, if large, denotes heavy tails, as reflected by the range $p97.5 - p2.5$ being large relative to the range

FIGURE A.4. Distribution of Rank Changes – Age 31

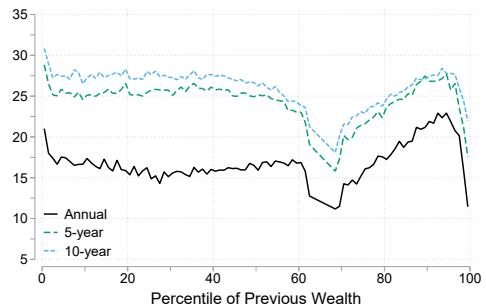
(A) Average Rank h Years Ahead



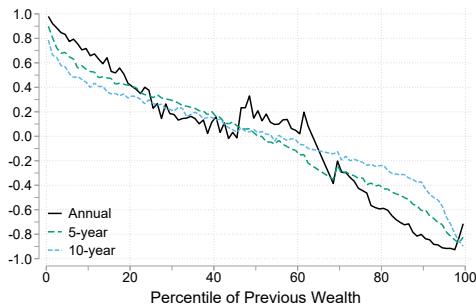
(B) Mean Change



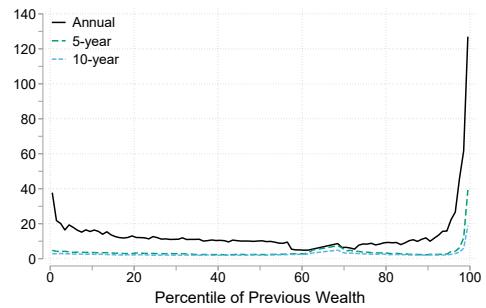
(C) S.D. of Changes



(D) Kelley's Skewness



(E) Crow-Siddiqui Kurtosis



Notes: The figures plot moments of the change in wealth ranks of individuals in the 1960-1964 birth cohort conditional on their rank at age 31. Panels correspond to the average rank level h years ahead, and the average, standard deviation, skewness, and kurtosis of the change in ranks h periods ahead. We use [Kelley \(1947\)](#)'s skewness measure, as in (A.1), and [Crow and Siddiqui \(1967\)](#)'s kurtosis measure, as in (A.2). The dark-continuous line corresponds to a one year horizon ($h = 1$). The dashed line corresponds to a five year horizon ($h = 5$). The short-dashed line corresponds to a ten year horizon ($h = 10$).

between $p75 - p25$. For a normal distribution, Kelley skewness equals zero and Crow-Siddiqui kurtosis equals 2.91. As above, we consider changes 1, 5, and 10 years ahead for 35 and 45 year olds and provide results for 31 year olds in Figure A.4 of Appendix A.

We plot the standard deviation of rank changes in panel A.5A. Consistent with Figure 2, we observe that the dispersion of wealth rank changes increases with the length of the horizon we consider. As with average changes, the very top of the distribution appears to follow a different process—not only do they suffer smaller falls, they are also exposed to much lower risk in the size of the fall. For the rest of the distribution the pattern is relatively flat, with the exception of individuals between the 50th and 70th percentiles, who experience lower variability in the changes of their wealth rank.

Turning to the asymmetry of the distribution of rank changes (panel A.5B), we observe, as expected, that individuals in the bottom of the wealth distribution experience mostly higher-than-average rank changes, as reflected by their positive skewness. This is in line with the mobility results of Figure 2. By contrast, rank changes of individuals in the upper 40 percent of the wealth distribution are mostly on the downside. Skewness displays roughly similar patterns across ages and horizons, and suggests that innovations to wealth ranks strongly deviate from normality.

Finally, we observe that most individuals experience relatively little changes in wealth ranks across years, with only a few individuals experiencing relatively large changes. This is evidenced by the high values for the kurtosis of wealth rank changes, panel A.5C. This pattern is, as expected, stronger for annual changes, and for the lowest and highest earners. The average kurtosis over 1- and 10-year horizons are above 15 and 6, respectively, while a normal distribution would imply a Crow-Siddiqui kurtosis of around 3.

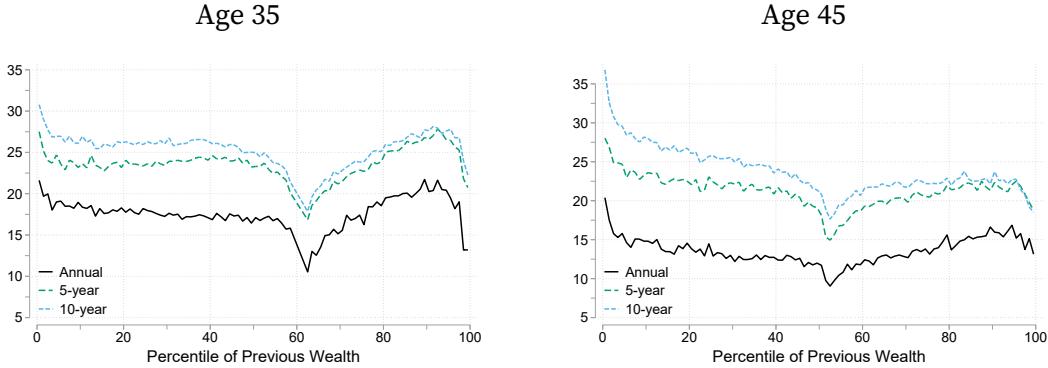
A.3. The distribution of change in income ranks

Figure A.6 replicates the exercise of Figure 2 for income using data from the Global Repository of Income Dynamics database (Guvenen, Pistaferri, and Violante 2022), specifically from Norwegian and U.S. data (Halvorsen, Ozkan, and Salgado 2022; McKinney, Abowd, and Janicki 2022). The data can be accessed at <https://www.grid-database.org>. The figure plots the average future income rank of individuals conditional on their present income rank for a 1 and a 5 year horizon. We report results for a sample of 35-44 year-old individuals that approximates the sample we use in our study, and for a larger sample of prime-aged adults, aged 25-55. The results are not largely affected by this sample selection.

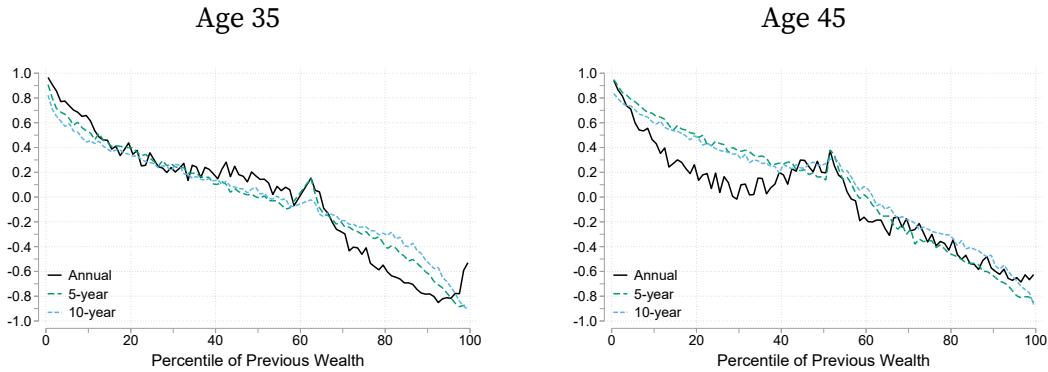
The persistence of income ranks is higher than that of wealth ranks, reported in Figure 1. The one year persistence as measured by $\rho(1)$, defined in equation (2), is 0.97, and the five year persistence is 0.73. Our results for wealth show increasing persistence of wealth ranks, but the persistence is at most 0.9 for a one year horizon and 0.75 for a five year horizon.

FIGURE A.5. Distribution of Rank Changes

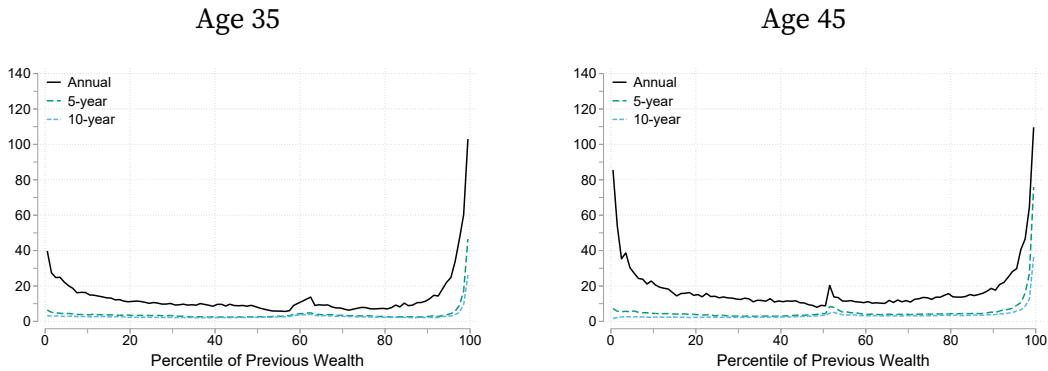
(A) Standard Deviation of Rank Changes



(B) Kelley's Skewness of Rank Changes

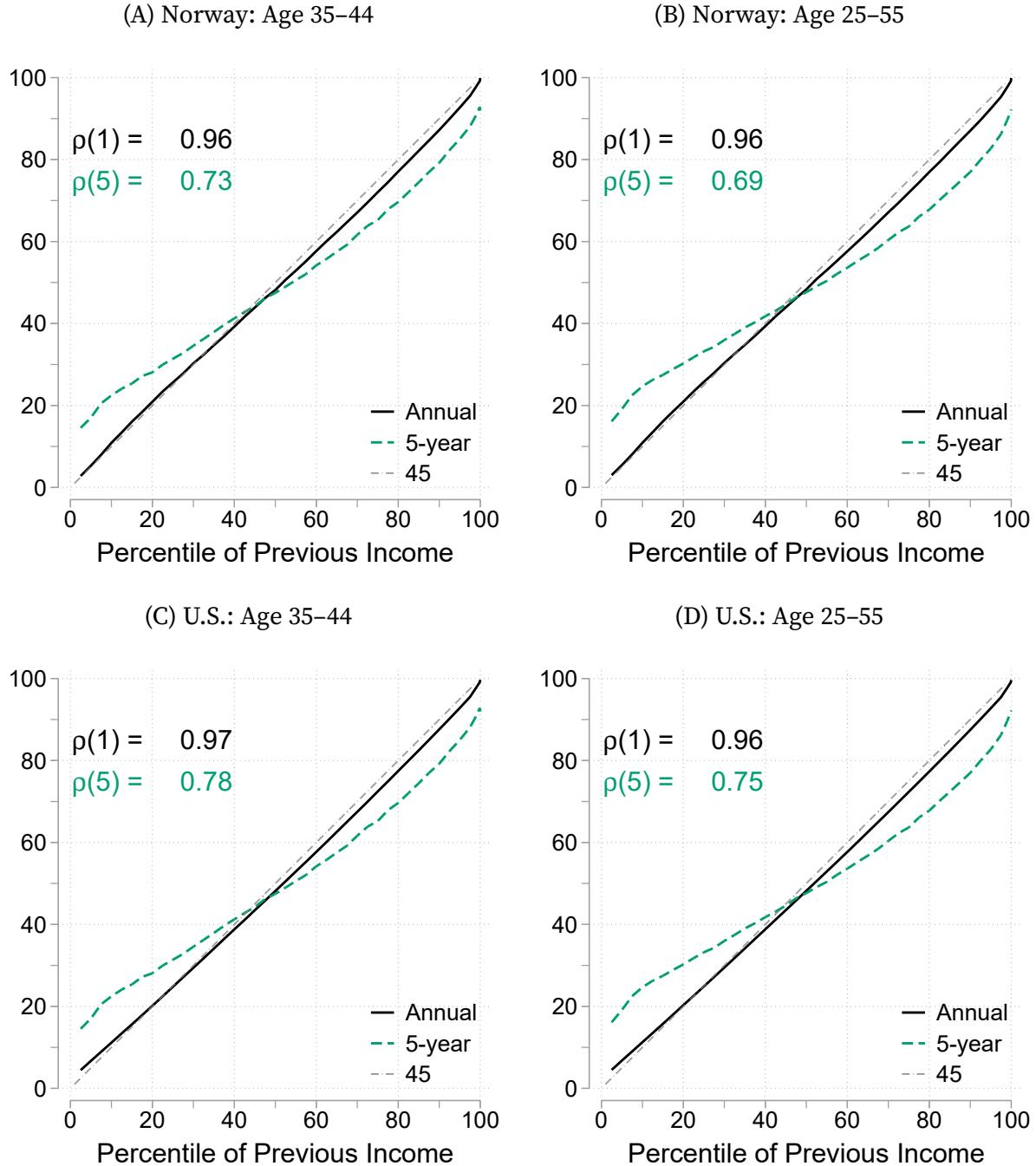


(C) Crow-Siddiqui Kurtosis of Rank Changes



Notes: The figures plot higher order moments of the change in wealth ranks of individuals in the 1960-1964 birth cohort conditional on their rank at ages 35, left panels, and 45, right panels. Panel A.5A plots the standard deviation. Panel A.5B plots Kelley (1947)'s skewness measure, as in (A.1). Panel A.5C plots Crow and Siddiqui (1967)'s kurtosis measure, as in (A.2). The dark-continuous line corresponds to a one year horizon ($h = 1$). The dashed line corresponds to a five year horizon ($h = 5$). The short-dashed line corresponds to a ten year horizon ($h = 10$).

FIGURE A.6. Average Income Rank h Years Ahead

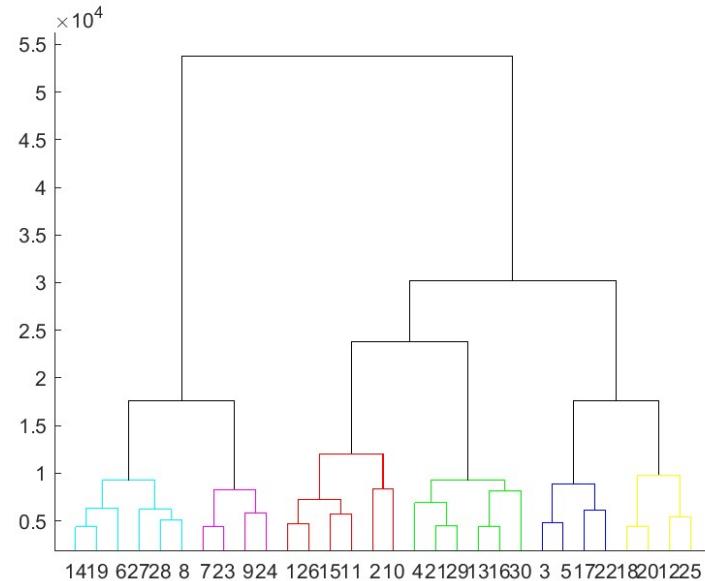


Notes: The figures plot the average future income rank of individuals in Norway and in the U.S. for selected samples of the population. Data is from the Global Repository of Income Dynamics database: <https://www.grid-database.org>. The dark-continuous line corresponds to a one year horizon ($h = 1$). The dashed line corresponds to a five year horizon ($h = 5$). The light-line corresponds to the 45 degree line.

Appendix B. Additional Results on Typical Trajectories

B.1. Agglomerative Hierarchical Clustering

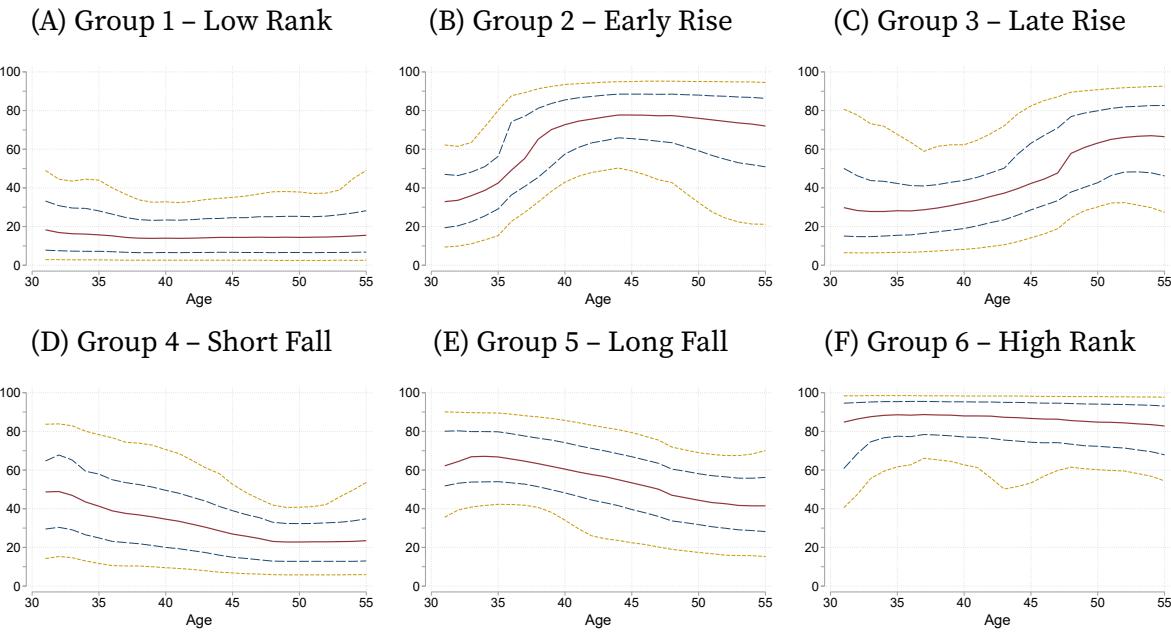
FIGURE B.1. Agglomerative Hierarchical Clustering: Dendogram



Notes: The figure presents the Dendrogram of the hierarchical agglomerative clustering procedure as executed on the balanced sample for the 1960-1964 birth cohort. The Dendrogram shows the tree of clusters up to a hierarchy of $G = 30$ groups. The tree shows how groups are merged as the clustering procedure recursively reduces the number of groups.

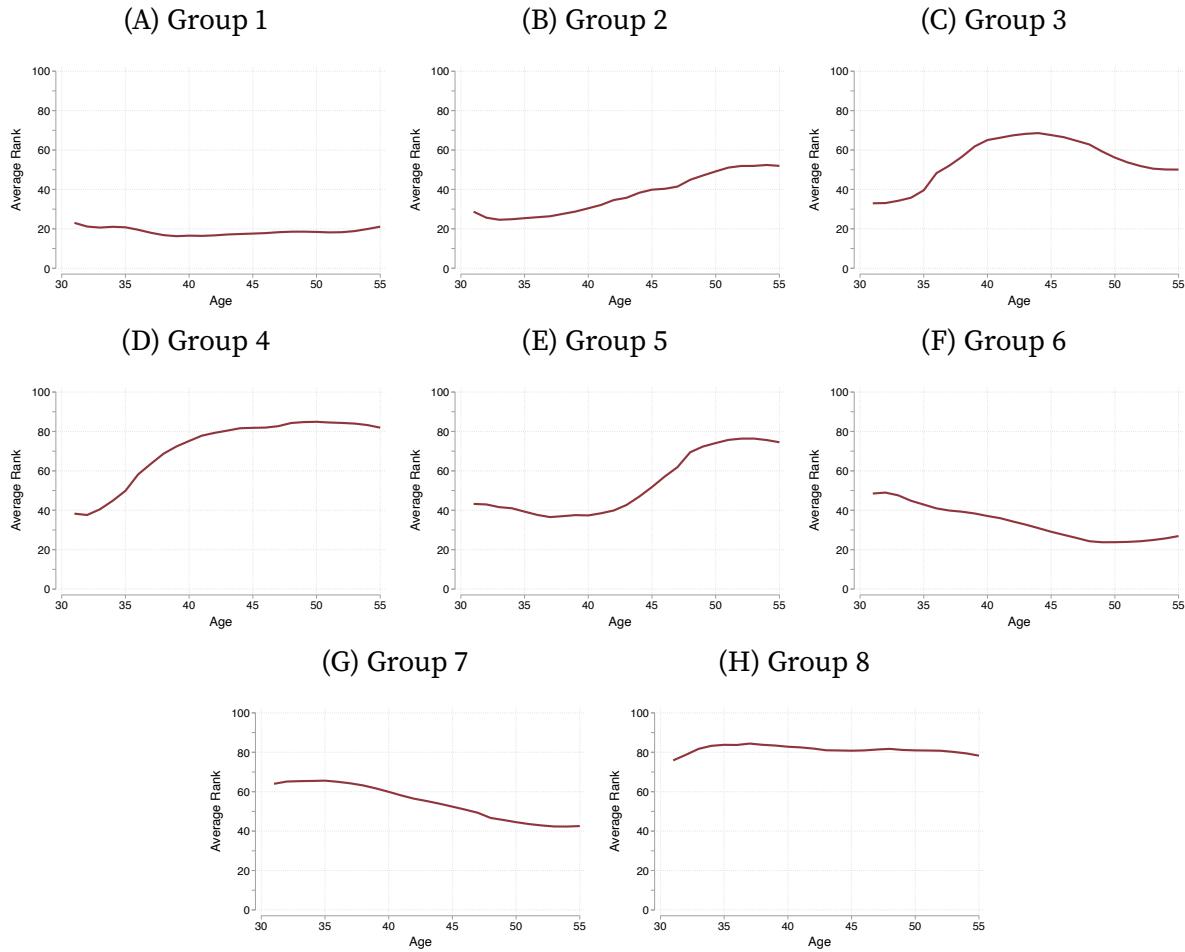
B.2. Distribution of typical trajectories

**FIGURE B.2. Life Cycle Dynamics of Wealth Mobility:
10th, 25th, 50th, 75th, and 90th Percentiles of Rank within Group**



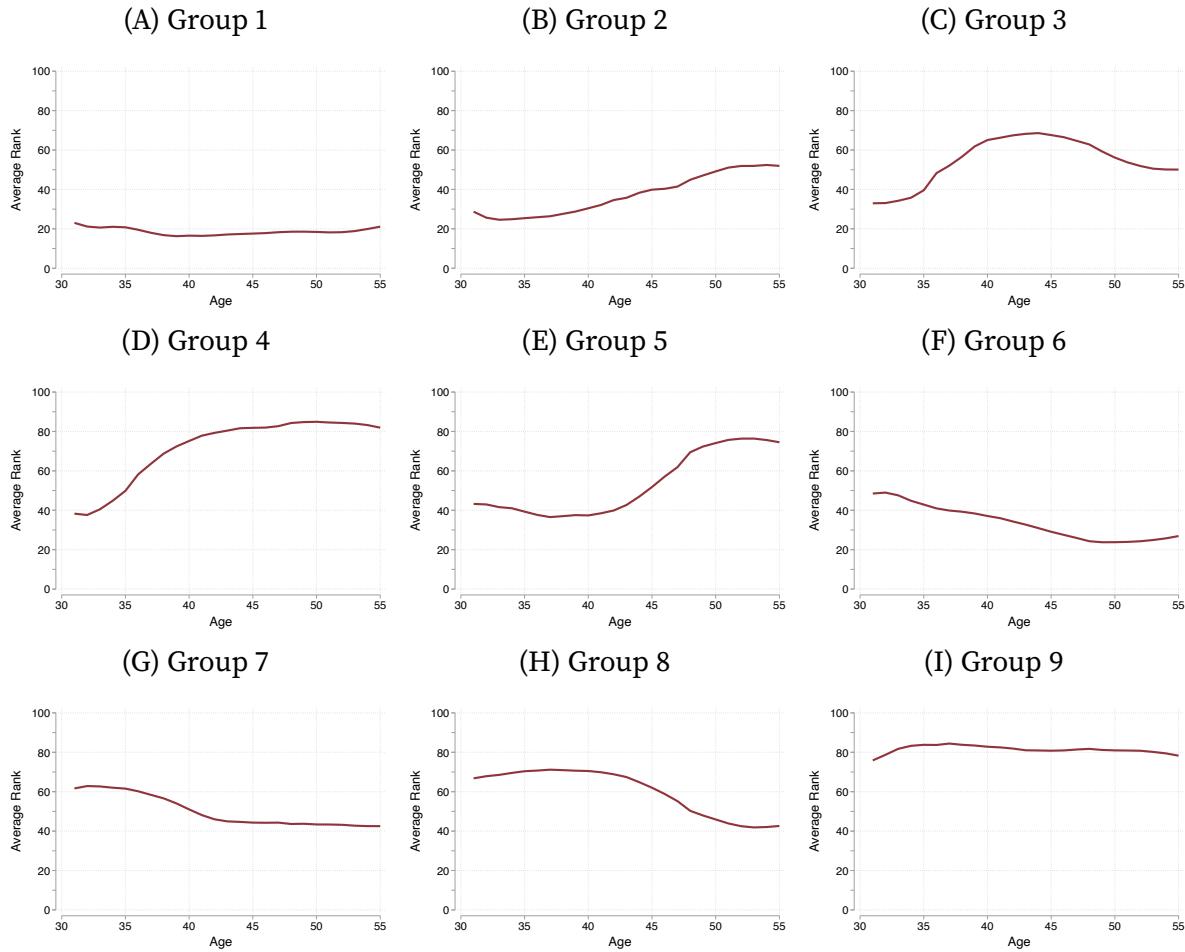
Notes: The figures plot the 10th, 25th, 50th, 75th, and 90th percentiles of the distribution of wealth ranks in each clustered group against the cohort's average age. The continuous line corresponds to the 50th percentile. The inner-dashed lines correspond to the 25th and 75th percentiles. The outward-dotted lines correspond to the 10th and 90th percentiles. All individuals belong to the 1960-1964 birth cohort. The clusters are constructed from the balanced sample using hierarchical agglomerative clustering and Ward's method with a dissimilarity measure (5).

FIGURE B.3. Life Cycle Dynamics of Wealth Mobility (8 Clusters)



Notes: The figures plot the average wealth rank of the individuals in each clustered group against the cohort's average age. All individuals belong to the 1960-1964 birth cohort. The clusters are constructed from the balanced sample using hierarchical agglomerative clustering and Ward's method with a dissimilarity measure (5)

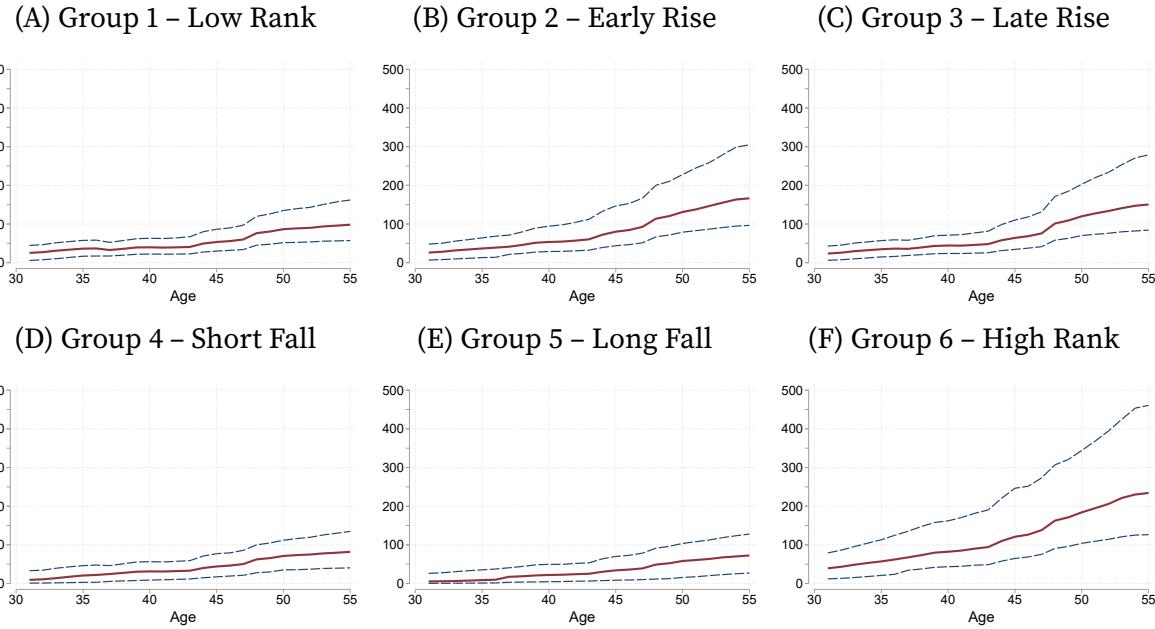
FIGURE B.4. Life Cycle Dynamics of Wealth Mobility (9 Clusters)



Notes: The figures plot the average wealth rank of the individuals in each clustered group against the cohort's average age. All individuals belong to the 1960-1964 birth cohort. The clusters are constructed from the balanced sample using hierarchical agglomerative clustering and Ward's method with a dissimilarity measure (5)

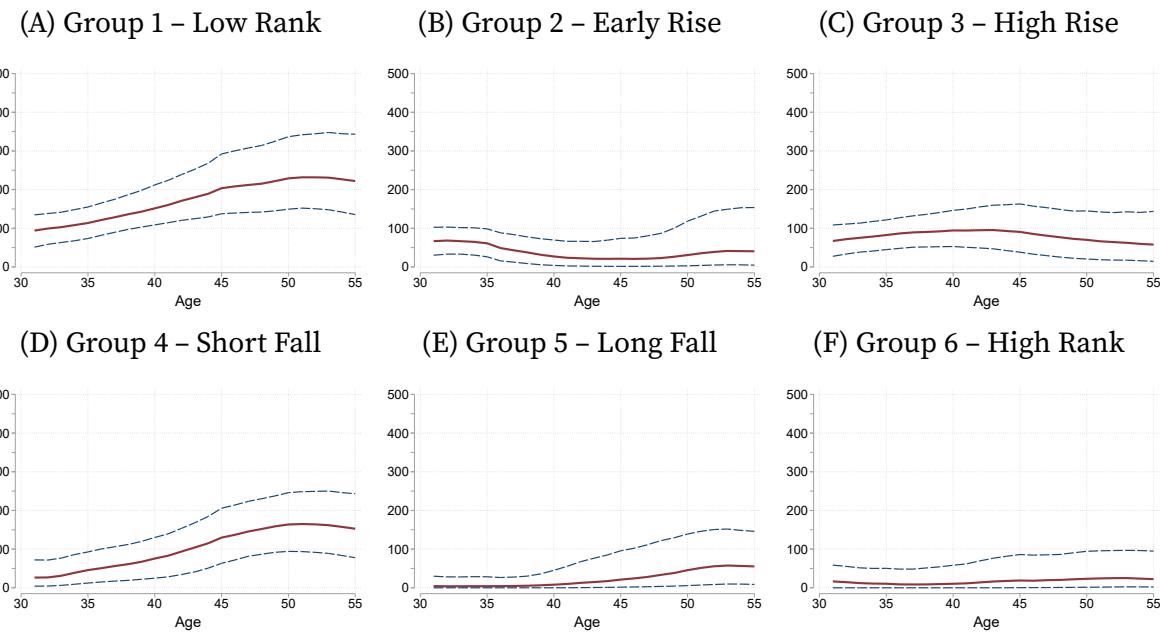
B.3. Further group characteristics

FIGURE B.5. Assets by Group: 50th, 25th, and 75th Percentiles



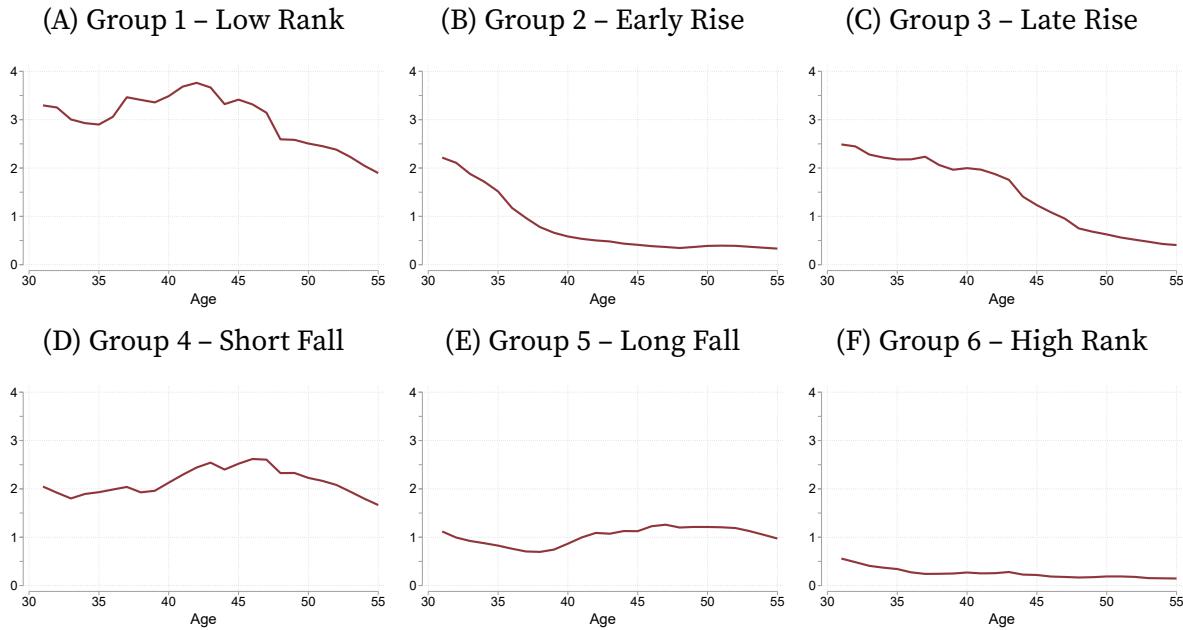
Notes: The figures plot the 50th, 25th, and 75th percentiles of the distribution of assets in each clustered group against the cohort's average age. The continuous line corresponds to the 50th percentile. The dashed lines correspond to the 25th and 75th percentiles. All numbers are in thousands of 2019 U.S. dollars. All individuals belong to the 1960-1964 birth cohort. The clusters are constructed from the balanced sample using hierarchical agglomerative clustering and Ward's method with a dissimilarity measure (5).

FIGURE B.6. Debt by Group: 50th, 25th, and 75th Percentiles



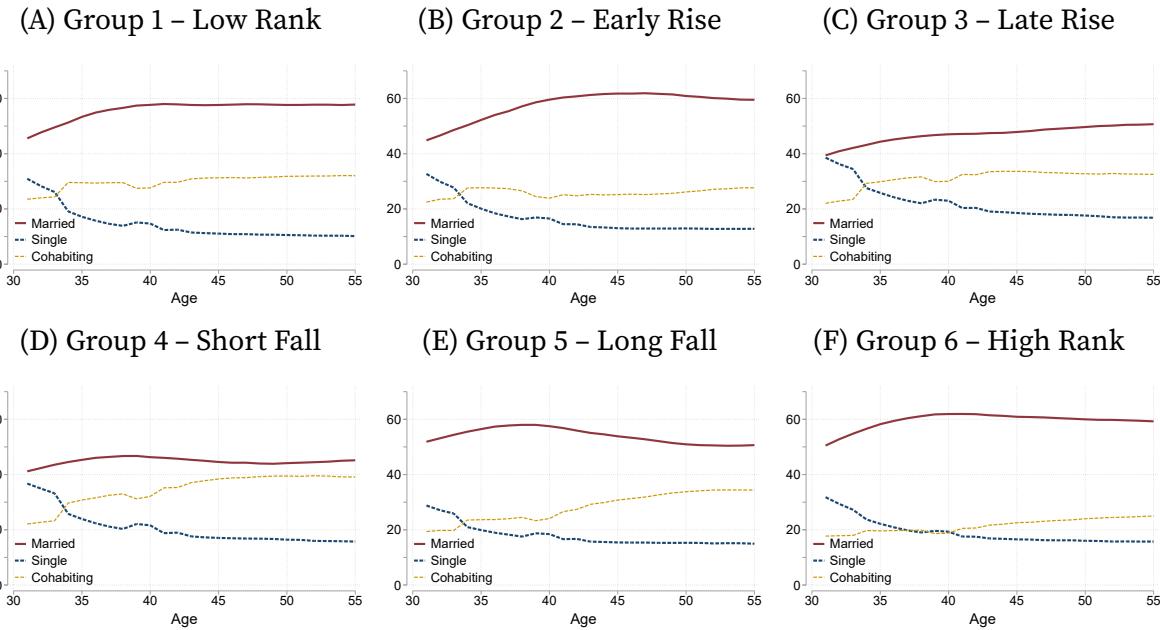
Notes: The figures plot the 50th, 25th, and 75th percentiles of the distribution of debt each clustered group against the cohort's average age. The continuous line corresponds to the 50th percentile. The dashed lines correspond to the 25th and 75th percentiles. All numbers are in thousands of 2019 U.S. dollars. All individuals belong to the 1960-1964 birth cohort. The clusters are constructed from the balanced sample using hierarchical agglomerative clustering and Ward's method with a dissimilarity measure (5).

FIGURE B.7. Leverage by Group



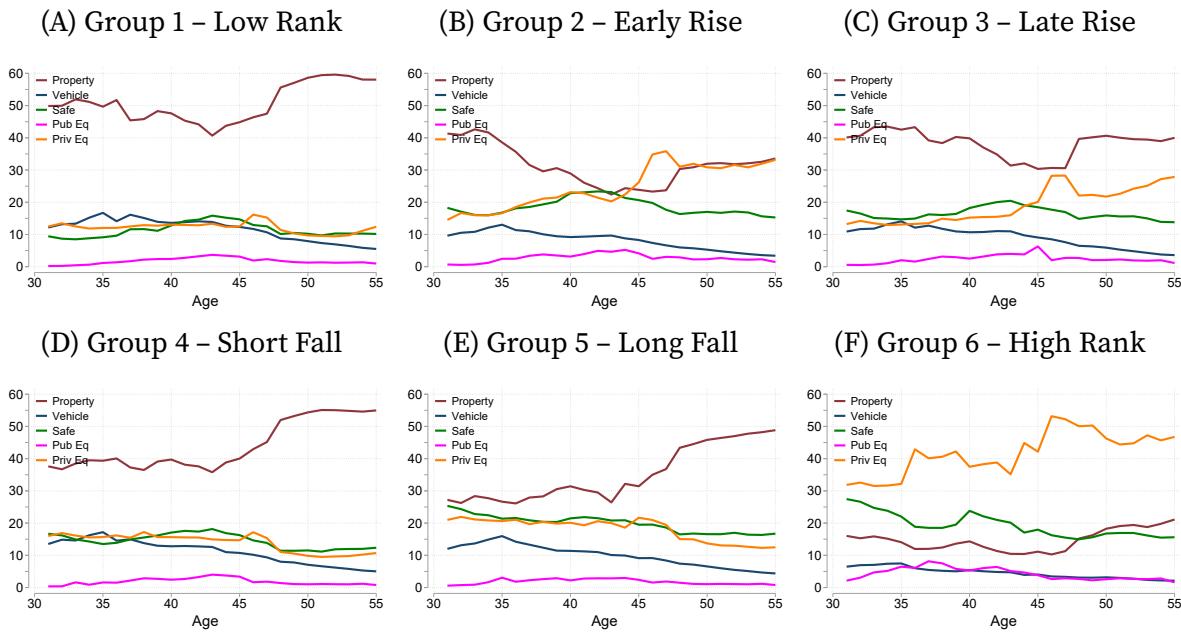
Notes: The figures plot the aggregate leverage in each clustered group against the cohort's average age. Leverage is computed as the ratio of aggregate assets to aggregate debt in each group. All numbers are in thousands of 2019 U.S. dollars. All individuals belong to the 1960-1964 birth cohort. The clusters are constructed from the balanced sample using hierarchical agglomerative clustering and Ward's method with a dissimilarity measure (5).

FIGURE B.8. Marriage and Cohabitation over the Life Cycle



Notes: The figures plot the composition of each clustered group in terms of the civil status of their individuals against the cohort's average age. The darkest-bottom region corresponds to the share of individuals who are single (neither cohabitating or married). The middle region corresponds to the share of individuals who are cohabitating and unmarried. The top region corresponds to the share of individuals who are married. All individuals belong to the 1960-1964 birth cohort. The clusters are constructed from the balanced sample using hierarchical agglomerative clustering and Ward's method with a dissimilarity measure (5).

FIGURE B.9. Portfolio Shares by Group



Notes: The figures plot the aggregate portfolio shares of gross assets in each clustered group against the cohort's average age. For included categories these correspond to Table 2 in the main text. We describe the construction of portfolio shares in Section 2. All numbers are in percentage points. All individuals belong to the 1960-1964 birth cohort. The clusters are constructed from the balanced sample using hierarchical agglomerative clustering and Ward's method with a dissimilarity measure (5).

Appendix C. The Shapley-Owen-Shorrocks Decomposition

Given an arbitrary function $Y = f(X_1, X_2, \dots, X_n)$, the Shapley-Owen-Shorrocks decomposition is a method to decompose the value of $f(\cdot)$ into each of its arguments X_1, X_2, \dots, X_n . Intuitively, the contribution of each argument if it were to be “removed” from the function. However, because the function can be nonlinear the order in which the arguments are removed matters in general for the decomposition. The function f can be the outcome of a regression, like the predicted values or sum of square residuals, or the output of a structural model, such as a counterfactual value for a variable given a list of model parameters or components, or a transformation of the sample, for example the Gini coefficient.

The Shapley-Owen-Shorrocks decomposition is the unique decomposition satisfying two important properties. First, the decomposition is exact decomposition under addition, letting C_j denote the contribution of argument X_j to the value of the function $f(\cdot)$,

$$\sum_{j=1}^n C_j = f(X_1, X_2, \dots, X_n), \quad (\text{C.1})$$

so that $C_j/f(\cdot)$ can be interpreted as the proportion of $f(\cdot)$ that can be attributed to X_j .²⁵ Second, the decomposition is symmetric with respect to the order of the arguments. That is, the order in which the variable X_j is removed from $f(\cdot)$ does not alter the value of C_j .

The decomposition that satisfies both those properties is

$$C_j = \sum_{k=0}^{n-1} \frac{(n-k-1)!k!}{n!} \left(\sum_{s \subseteq S_k \setminus \{X_j\}: |s|=k} [f(s \cup X_j) - f(s)] \right), \quad (\text{C.2})$$

where n is the total number of arguments in the original function f , $S_k \setminus \{X_j\}$ is the set of all “sub-models” that contain k arguments and exclude argument X_j .²⁶ For example,

$$\begin{aligned} S_{n-1} \setminus X_n &= f(X_1, X_2, \dots, X_{n-1}) \\ S_1 \setminus X_n &= \{f(X_1), f(X_2), \dots, f(X_{n-1})\}. \end{aligned}$$

²⁵The interpretation holds as long as f is non-negative. If f can take negative values, then the interpretation of C_j under the exact additive rule can be misleading as some arguments can have $C_j < 0$.

²⁶We abuse notation here. A sub-model is an evaluation of function f with only some of its arguments. This language is motivated by the function corresponding in practice to the outcome of a regression or structural model. Formally when we write $f(X_1)$ we mean $f(X_1, \emptyset_2, \dots, \emptyset_n)$, where we assume the j -th argument of the function can always take on a null value denoted \emptyset_j . In our regression example below this null value corresponds to a zero valued regressor or parameter. In the case of structural model this null value can correspond to setting some parameters to a predetermined value or excluding certain model components, like the adjustment of prices or a specific shock agents face.

The decomposition in (C.2) accounts for all possible permutations of the decomposition order. Thus, $\frac{(n-k-1)!k!}{n!}$ can be interpreted as the probability that one of the particular sub-model with k variables is randomly selected when all model sizes are all equally likely. For example, if $n = 3$, there are sub-models of size $\{0, 1, 2\}$. In particular, there are 2^2 permutation of models that exclude each variable: $\{(0, 0), (1, 0), (0, 1), (1, 1)\}$.

$$\underbrace{(0, 0)}_{k=0} \quad \underbrace{(1, 0)}_{k=1} \quad \underbrace{(0, 1)}_{k=2} \quad (1, 1)$$

$$\begin{aligned} k = 0 : \frac{(n - k - 1)!k!}{n!} &= \frac{(3 - 0 - 1)!0!}{3!} = \frac{1}{3} \\ k = 1 : \frac{(n - k - 1)!k!}{n!} &= \frac{(3 - 1 - 1)!1!}{3!} = \frac{1}{6} \\ k = 2 : \frac{(n - k - 1)!k!}{n!} &= \frac{(3 - 2 - 1)!2!}{3!} = \frac{1}{3} \end{aligned}$$

Non-linear example

We illustrate the value of this decomposition with a simple non-linear model including $n = 3$ variables:

$$Y = f(X_1, X_2, X_3) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 X_2. \quad (\text{C.3})$$

The objective is to decompose the value of Y into the contribution (or partial effect) of each variable.

Removing X_1

There are 4 possible models that exclude X_1 , one with no variable, 2 with one variable and one with 2 variables

$$\begin{aligned} k = 0 : &\beta_0 \\ k = 1 : &\{\beta_0 + \beta_2 X_2, \beta_0\} \\ k = 2 : &\beta_0 + \beta_2 X_2 + \beta_3 X_3 X_2 \end{aligned}$$

In all 4 models, the partial effect of including X_1 is always $f(s \cup X_1) - f(s) = \beta_1 X_1 \quad \forall s$. This reflects the fact that the order that the order in which variables are included does not matter to construct C_1 :

$$C_1 = \sum_{k=0}^2 \frac{(3 - k - 1)!k!}{3!} \left(\sum_{s \subseteq S_k \setminus \{X_3\}: |s|=k} [f(s \cup X_j) - f(s)] \right) = \beta_1 X_1 \quad (\text{C.4})$$

This would be the same for any argument X_j entering linearly into f an arbitrary number of variables: $Y = f(X_1, X_2, X_3, X_4, \dots, X_n) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 X_2 + \sum_{j=4}^n \beta_j X_j$. The only difference is that the number of sub-models grows exponentially, 2^{n-1} , but the

partial effect of including X_j for some $j \in \{4, \dots, n\}$ is always $C_j = \beta_j X_j$.

Removing X_2

In this case, the partial effect can be decomposed into all the possible ways X_2 can be added into the model, $f(s \cup X_2) - f(s)$, these are

$$\begin{aligned} k=0 (\emptyset_1, \emptyset_3) : \beta_0 + \beta_2 X_2 - \beta_0 &= \beta_2 X_2 \\ k=1 (X_1, \emptyset_3) : \beta_0 + \beta_1 X_1 + \beta_2 X_2 - (\beta_0 + \beta_1 X_1) &= \beta_2 X_2 \\ k=1 (\emptyset_1, X_3) : \beta_0 + \beta_2 X_2 + \beta_3 X_2 X_3 - \beta_0 &= \beta_2 X_2 + \beta_3 X_2 X_3 \\ k=2 (X_1, X_3) : \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_2 X_3 - (\beta_0 + \beta_1 X_1) &= \beta_2 X_2 + \beta_3 X_2 X_3 \end{aligned}$$

Here, the partial effects of adding X_2 are not the same across sub-models because X_2 enters non-linearly into the original model. The symmetric property of the decomposition takes care of this.

$$\begin{aligned} C_2 &= \underbrace{\frac{1}{3} \beta_2 X_2}_{k=0} + \underbrace{\frac{1}{6} (\beta_2 X_2)}_{k=1} + \underbrace{\frac{1}{6} (\beta_2 X_2 + \beta_3 X_2 X_3)}_{k=1} + \underbrace{\frac{1}{3} (\beta_2 X_2 + \beta_3 X_2 X_3)}_{k=2} \\ &= \beta_2 X_2 + \frac{1}{2} \beta_3 X_2 X_3 \end{aligned} \tag{C.5}$$

The result is quite intuitive. $\beta_2 X_2$ appears in all sub-models, hence its probability of appearing in the decomposition is 1. $\beta_3 X_2 X_3$ appears in 2 of the 4 sub-models, hence its probability of appearing is 1/2. Weighting each term by its probability of appearing in the decomposition ensures symmetry.

Removing X_3

We proceed in the same way for X_3 as we did for X_2 . There are 4 sub-models. In 2 of them the effect of adding X_3 is null because X_3 is not in the model. In the 2 remaining sub-models the effect is $\beta_3 X_2 X_3$. Hence,

$$C_3 = \frac{1}{2} \beta_3 X_2 X_3. \tag{C.6}$$

Finally, we verify the decomposition:

$$\begin{aligned} C_1 + C_2 + C_3 &= \beta_1 X_2 + \left(\beta_2 X_2 + \frac{1}{2} \beta_3 X_2 X_3 \right) + \left(\frac{1}{2} \beta_3 X_2 X_3 \right) \\ &= \beta_1 X_2 + \beta_2 X_2 + \beta_3 X_2 X_3 \\ &= f(X_1, X_2, X_3) - \beta_0 \\ &= f(X_1, X_2, X_3) - f(\emptyset_1, \emptyset_2, \emptyset_3). \end{aligned}$$

Note: The decomposition is additive with respect to the reference “null” model where none of the variables are included. This is made apparent in the previous result, where the decomposition does not include the value of β_0 .

R-Squared

Finally, we consider a decomposition of the coefficient of determination in the linear model. Our use of the decomposition applies this for a non-linear model (combining the insights from this and the preceding example).

Consider a linear regression model with n regressors and $i = 1, \dots, M$ observations,

$$y_i = \mathbf{x}'_i \beta + u_i = \beta_0 + \sum_{j=1}^n \beta_j x_{ij} + u_i, \quad (\text{C.7})$$

and define the average value of y as $\bar{y} \equiv \sum_{i=1}^M y_i / M$ and the predicted value

$$\hat{y}_i = \mathbf{x}'_i \hat{\beta} = \hat{\beta}_0 + \sum_{j=1}^n \hat{\beta}_j x_{ij}, \quad (\text{C.8})$$

where we assume that all regressors have zero mean so that $\hat{\beta}_0 = \bar{y}$.

The function of interest is $f(X_1, \dots, X_K) = R^2$, defined as the explained sum of squares SSE over the total sum of squares SST

$$R^2(X_1, X_2, \dots, X_n) = \frac{SSE}{SST} = \frac{\sum_{i=1}^M (\hat{y}_i - \bar{y})^2}{\sum_{i=1}^M (y_i - \bar{y})^2}. \quad (\text{C.9})$$

This makes it clear that the function being decomposed is non-linear even though the model that generates it is itself linear.

Note: The reference value for the R^2 in the Shapley-Owen-Shorrocks decomposition is given by the model without regressors, satisfying

$$R^2(\emptyset) = \frac{\sum_i^M (\hat{\beta}_0 - \bar{y})^2}{\sum_i^M (y_i - \bar{y})^2} = 0, \quad (\text{C.10})$$

so that, in this case, the decomposition recovers the level of the R^2 of the full model (with all variables), unlike the previous example.

Details of the decomposition when $n = 3$ Consistent with the previous example, we show the decomposition for $n = 3$ regressors. As before, we abuse notation by only listing the arguments being included in each sub-model. The contribution of each variable is:

$$\begin{aligned} R_1^2 &= \frac{1}{3} \left[R^2(X_1) - R^2(\emptyset) \right] + \frac{1}{6} \left(\left[R^2(X_1, X_2) - R^2(X_2) \right] + \left[R^2(X_1, X_3) - R^2(X_3) \right] \right) \\ &\quad + \frac{1}{3} \left[R^2(X_1, X_2, X_3) - R^2(X_2, X_3) \right]; \end{aligned} \quad (\text{C.11})$$

$$R_2^2 = \frac{1}{3} [R^2(X_2) - R^2(\emptyset)] + \frac{1}{6} \left([R^2(X_1, X_2) - R^2(X_1)] + [R^2(X_2, X_3) - R^2(X_3)] \right) \\ + \frac{1}{3} [R^2(X_1, X_2, X_3) - R^2(X_1, X_3)]; \quad (\text{C.12})$$

$$R_3^2 = \frac{1}{3} [R^2(X_3) - R^2(\emptyset)] + \frac{1}{6} \left([R^2(X_3, X_2) - R^2(X_2)] + [R^2(X_1, X_3) - R^2(X_1)] \right) \\ + \frac{1}{3} [R^2(X_1, X_2, X_3) - R^2(X_2, X_1)]. \quad (\text{C.13})$$

Summing across all the contributions we obtain back $R^2(X_1, X_2, X_3)$,

$$R_1^2 + R_2^2 + R_3^2 = R^2 = f(X_1, X_2, X_3). \quad (\text{C.14})$$

Note: The value of the contribution differs from the standard definition of partial R-squared. This is because the partial R-squared is an all else equal comparison of excluding regressor X_j from the regression. It does not satisfy the exact decomposition requirement, nor (when applied iteratively) the symmetry requirement.