ENTREPRENEURS, FLUCTUATIONS AND DYNAMICS WITHIN THE WEALTH DISTRIBUTION

THOMAS MICHAEL PUGH

May 28, 2019

University College London (UCL)

Submitted towards the degree of PhD in Economics

Supervised by Dr. Vincent Sterk and Prof. Mariacristina De Nardi

I, Thomas Michael Pugh, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Thanks to my supervisors, Vincent Sterk and Mariacristina De Nardi for their dedicated support, to Rory McGee and Gonzalo Paz Pardo for their greatly appreciated comments and conversations and to Antonio Guarino for his guidance. I would also like to thank my family for endless availability and care.

This thesis includes elements from my papers Pugh [2018c], Pugh [2018a] and Pugh [2018b].

Abstract

There has been substantial interest in inequality and the distribution of wealth for centuries. After changes beginning in the 1980's, rising income and wealth inequality at the top has become an important item on the policy agenda and in public discussion. This thesis develops the study of wealth inequality in two key directions - firstly, the study and use of mobility in wealth amongst households to understand and discriminate between the mechanisms and theories purporting to explain the highly concentrated distribution of wealth in the upper tail. Secondly, the study of interactions between entrepreneurship, which is prominent amongst the wealthy, and aggregate shocks to the economy.

In Chapter One, I investigate wealth data in the UK find that there is substantial mobility amongst the wealthy and large changes in wealth. In Chapter Two, I use these findings to estimate a model incorporating multiple theories of the upper wealth distribution and identify that heterogeneity in returns has the best fit to the data.

In Chapter Three I turn to entrepreneurship, examining the impact of entrepreneurial constraints on the economy when responding to aggregate shocks. I find evidence of increased dispersion amongst businesses during recessions and use this to calibrate uncertainty- and mobility- increasing 'turbulence shocks amongst entrepreneurs. I find that entrepreneurial behaviour amplifies TFP-style shocks and symmetric turbulence shocks have rich effects, changing the distribution of wealth and delivering medium term decreased output and a spell of longer term increased output through slow capital reallocation.

I conclude that the study of the top of the wealth distribution and entrepreneurship has significant implications, in terms of understanding the mechanisms behind inequality and the mechanisms that drive patterns in the business cycle.

IMPACT STATEMENT

This work impacts a large range of literatures within economics and in wider social science. My investigation of mobility, inequality and entrepreneurship throws light on key questions that are prominent in public discourse and policy discussions as well as academia. How likely are those in the top 1% to be there in 2, 5 or 10 years and why? And how likely are those not at the top to move up to it? Why do those at the top hold so much wealth? Does entrepreneurship, inequality and individual shocks affect the economy over the business cycle? My research answers these questions and provides a framework for future research - for example, when considering government policies to increase welfare for those across the distribution, it is important to understand mobility in that distribution, and what effects changes on the individual level have on the economy, including small businesses which in aggregate provide significant employment and output.

The research demonstrates the importance of considering mobility when explaining the high inequality at the top of the wealth distribution, and shows the role that entrepreneurial capital constraints can have in amplifying and propagating different types of aggregate shocks associated with recessions.This impacts upon policy making with regards to fiscal policy concerning small businesses, prudential lending regulation and monetary policy, all areas where the businesses cycle and the contributions of entrepreneurial behaviour to fluctuations is of great importance.

Outside of academia or policy making, there is a large and varied discussion regarding inequality, inclusion and opportunity. By providing facts on mobility at the very top, especially on the potentially short tenure of some of the very wealthy, this thesis informs the debate. Interpreting and inferring from those facts is also important, and the work understanding the mechanisms driving these patterns in the data is significant. It shows that claims of wealth inequality being sourced from different preferences for future saving or from sudden 'superstar' earnings do not successfully match the data, whilst heterogeneity in returns from wealth and investment does and is key to understanding the downwards movements in the distribution.

Contents

List of Tables

List of Figures

Chapter 1

The Wealth and Assets Survey and Wealth Dynamics at the top

In this chapter, I investigate the UK Wealth and Assets Survey (Office for National Statistics and UK Data Service [2018]). I benchmark inequality and the representation of the wealthy in the survey versus Estate data, other surveys and available rich lists for the top 1000, showing the WAS represents inequality at the top at least as well as other sources. I argue for the use and imputation of business wealth from the survey and study that wealth.

There is a large concentration of wealth in the UK, with the top 1% holding approximately 20% of assets, with a threshold for entry of over £3 million. I document substantial wealth mobility amongst the top, with one third of the top 1% leaving this category in 2 years and half of the top 1% leaving in six years. There is also evidence of greater mobility amongst the very top during the 07-09 recession years.

1.1 Introduction

The very wealthiest hold a large fraction of wealth in most developed economies and there has been substantial interest in documenting shares of wealth held by the very wealthiest - for example, the works of Piketty [2014] and Saez and Zucman [2014] studying the long run sequences of these wealth shares and other features of the wealth distribution. I use a wealth survey dataset for the UK, the Wealth and Assets Survey (WAS) published by the Office for National Statistics and UK Data Service [2018], to contribute to this literature, providing and analysing similar cross-sectional moments to the literature, validating the data against other sources and then contributing new facts through studying longitudinal transitions of the wealthy.

In the UK, Alvaredo et al. [2017] provide a cross-sectional analysis using Estate data, one of four main sources for study of the (upper) wealth distribution - estate data, asset income data, rich lists of the wealthy and wealth surveys. The UK does not yet have an available dataset for adequately recreating wealth from asset income, but has the Sunday Times Rich List covering the wealthiest 1000 individuals or households and an array of surveys. Amongst these, the WAS is unique due to a combination of size, oversampling of the wealthy and depth of wealth, income and portfolio questions. Importantly, amongst its international wealth survey peers, the WAS is rare for including longitudinal features.

The first contribution of this work is to demonstrate that the WAS has a better representation of the top of the wealth distribution than previously thought, on a variety of dimensions, especially in later waves. There has been criticism of the WAS in Crossley et al. [2016] for poor representation of the top 1%. Yet, I show the wealthy in the WAS are richer than in estate data and compare favourably to the rich list, particularly focusing on the impact of business wealth and pensions.

The mobility of the wealthy is an important economic fact - a world with long, stable dynasties may have very different implications to one with rapid rises and falls of different households, something I study with structural macroeconomic models in other chapters. These dynamics of the wealthy tail are less studied than the cross-section, largely due to a relative lack of data. To my knowledge, there are few alternative longitudinal wealth datasets to examine wealth dynamics at the top of the distribution. Thus, longitudinal data from the WAS can shed light on this topic.

Thus the second contribution of this chapter is that I use the WAS to study the relatively unknown distribution of changes in wealth faced by (top) households and their wealth mobility patterns. There is substantial wealth and income mobility at the top. To my knowledge, this study is the first to extensively analyse these distributions of panel changes in wealth including the very wealthy. Over a third of the wealthiest 1% exit this group biennially and are unlikely to return. After six years, only half of the wealthiest 1% remain in the same wealth category.

Their dynamics show rich history dependence and indicate more than a simple Markov-style process. Newer entrants to wealthy groups such as the top 1% are much more likely to leave again in two years (60% exit) versus those already in the group (20% exit). There are also high likelihoods of dramatic changes amongst the wealthy - for example, amongst the wealthiest 5%, one quarter lose over 25% of their wealth and 10% lose over half their wealth in two years. Quantile regression implies that serial dependencies over wealth do not substantially weaken over longer time horizons, unlike the same dependencies for German and American incomes seen in Trede [1998]. In addition, I find moments of the change in log wealth distribution over quantiles of wealth to be similar to the U-shaped skew and variance curves found in Guvenen et al. [2015] for earnings. To my knowledge, this study is the first to extensively analyse these distributions of survey panel changes in wealth including the very wealthy, and I find similar patterns in the Survey of Consumer Finances (SCF), Panel Study of Income Dynamics (PSID) and English Longitudinal Study of Ageing (ELSA) though these sources are less suited to studying the top of the wealth distribution..

1.2 Literature Review

There is a well-established economic literature documenting and studying the cross-sectional distribution of income and wealth over time, notably Atkinson and Piketty [2014] covering income inequality over the 20^{th} century and the study of wealth inequality in Piketty [2014]. This is accompanied by a literature evaluating data quality to establish empirical facts, to which I contribute. Alvaredo et al. [2017] study UK wealth using estate data, which is a useful comparator for the WAS. They provide data on the UK wealth distribution drawn from over a century of Estate data in the UK. This data demonstrates a substantial fall in inequality from the 1920's until the 1980's, with a rise in wealth inequality thereafter, much like the U.S.A.

Vermeulen [2016] studies the effectiveness of wealth surveys in capturing the wealthy, and is a valuable reference for evaluating the WAS results. He finds that many wealth surveys without substantial oversampling fail to match the wealth distribution at the top when comparing to the Power Law found in the very extreme tail of the Forbes list billionaires.

The approach of Bricker et al. [2015] to reconcile the SCF with capitalised income data in Saez and Zucman [2014] informs the attempt to compare survey and estate data in this study. In the case of the SCF versus capitalised incomes, the data can be reconciled with the use of the careful use of the correct definitions and aggregate wealth measures, which aligns to my findings with respect to the WAS and Estate data.

Whilst detailed work has been completed on top income transitions, mobility and distributions - for example, in the US, Auten et al. [2013]; Kopczuk et al. [2007] and Guvenen et al. [2014a] - there is a relative lack of equivalents for wealth dynamics. Substantial European alternatives include longitudinal Nordic and Scandinavian administrative wealth datasets, for example, Fagereng et al. [2016] and the panel subsample of the Italian Survey of Household Income and Wealth (SHIW) discussed by Jappelli and Pistaferri [2000] and Jappelli [1999]. The work by Fagereng et al. [2016] is of particular note as they discuss returns to wealth, which is very close to the focus on wealth transitions and fluctuations later in this study. Their findings of a significant individual component to wealth returns is important when considering the mechanisms driving the wealth distribution in this dissertation.

In the U.S. there are only 2 small one-off transitional datasets from the SCF - a 1989 re-interview of the 1983 wave (Kennickell and Starr-McCluer [1997]) and the same for 2007 and 2009 (Bricker et al. [2011]). There are therefore only two data points for US SCF wealth transitions, separated by 20 years. The Panel Study of Income Dynamics

(PSID) is relatively much longer (1968-present) but does not represent the richest via oversampling like the SCF or WAS, and so misses the wealthiest 1%. Quadrini [2000] and Hurst et al. [1998]) both study wealth in the PSID, and a more recent discussion with reference to wealth mobility can be found in Carroll et al. [2017b]. There are also long-term inter-generational studies of income and wealth persistence, but these are a very different concept of mobility to that embodied in the biennial movements I consider.

The UK has a number of survey datasets which also include wealth information. The closest to the WAS is ELSA, the English Longitudinal Study of Ageing (Banks et al. [2019]. ELSA has detailed information on wealth throughout its 7-wave[1] history of biennial waves from 2000 to the present, but focuses only on those over 50 at the time of the first wave. Meanwhile the British Household Panel Survey (BHPS) is designed to cover the whole population and runs annually, but does not oversample the wealthy and only includes substantial wealth information for 3 special waves (every five years), as this is not a major focus for the survey (of Essex. Institute for Social and Research. [2018]). Both ELSA and BHPS are substantially smaller than the WAS. The Bank of England/NMG Survey, discussed in Anderson et al. [2016] is another alternative, running yearly over a similar period to the WAS and including some wealth variables. But, the wide use of banding and top coding for these variables, together with small survey size and potential sample selection issues, limits its effectiveness.

[1] The 8th wave is being released at the time of writing.

1.3 Top Wealth shares in the UK

In this section I analyse the WAS data and appropriate wealth measures, particularly discussing the wealth shares of the richest $x\%$. I also compare the WAS to estate data, explicitly considering the two components in the wealth share figures, the numerator of in-group wealth and the denominator of total population wealth, to identify the source of differences. By using the careful comparisons and investigation of the WAS, I argue the WAS effectively represents the top of the distribution.

The WAS is a biennial panel survey dataset covering wealth, income and demographics for UK households. It is large versus the U.S. SCF or average country in the EU Household Finance and Consumption Survey (HFCS)[2], with 20,000 or more households in each wave, and new samples added from wave 3 onwards to maintain its size. The WAS contains 4 biennial survey waves, beginning in July 2006 - June 2008 for wave 1. Wealthy households are also oversampled to account for lower response rates, much like other wealth surveys (such as the SCF).[3]

For each household, WAS interviewers ask respondents for information on wealth, income and various demographic features. They catalogue valuations and amounts of different assets, as well as recording surrounding information such as date of purchases for properties or personal opinions towards leaving a bequest. Like other well-designed surveys[4], interviewers endeavour to probe answers and ask respondents to use financial statements and records as aids in their answers. The data provider also performs some imputation for missing responses and I analyse the impact of additional multiple imputation for non-answerers to business wealth questions

I consider various definitions and levels of wealth based on the main categories in the WAS: private business values; financial assets (cash, shares, bonds, investment funds, savings products, deposits minus debts [informal and formal] and credit cards); property (value minus mortgage debt); pensions and physical wealth (vehicles, jewellery, collectibles, household contents). In each case, I consider the net measure of wealth, i.e. including relevant debts as negative wealth. Different longitudinal weighting schemes do not affect the cross-sectional and longitudinal results, particularly the ones used in later chapters.[5] The definition of wealth used by the ONS in the original release is the sum of financial wealth; property wealth; pension wealth and physical wealth, i.e. excluding business wealth. I discuss their exclusion of business wealth and propose that the better measure is the sum of business wealth; property wealth; financial wealth; and physical wealth, i.e.

[2]As examples, the SCF contains 6,000 families whilst the HFCS has 80,000, but contains 20 EU countries - averaging 4,000 per country.

[3]Although the WAS has a lower oversampling rate, at 2x-3x versus 6x for the SCF, it has a larger sample (approximately, WAS_n=20-30,000 households versus $SCF_n < 5000$.), so still maintains a sizeable responding sample for the top quantiles - around 500 observations for the top 1% and 1500 for the top 5%.

[4]The benchmark examples being the U.S. Survey of Consumer Finances and Panel Survey of Income Dynamics as two of the most frequent sources for wealth data in economic research.

[5]See Appendix (A.0.5)

excluding pensions and including businesses.

Pensions are a large component of aggregate wealth and their WAS modelling valuations drive several, from this work's perspective, erroneous, patterns which are important to analyse and address. Pensions are not necessarily consumable assets at the time of survey and are thus complex to model, thus some studies exclude available pensions from wealth, for example Hurd [1989] and Hurd [2002] and some surveys have pension contributions and scheme features but do not offer pension wealth as an item (for example, ELSA).

The WAS creates valuations for defined contribution (DC) plan holders by directly asking for respondents' account value, with probing questions about available financial statements. Defined benefits (DB) pension plans are valued using models, utilising plan details combined with discounting assumptions. The Office of National Statistics (ONS) changes DB modelling assumptions over waves - as stated in the ONS wave 3 report [2014], newer estimates of previous waves' aggregate pension wealth were 20-30% lower and median individual occupational defined benefit plan wealth fell by 45%.

Changes to individual households' modelled pensions over waves are also substantial and varied - amongst the top 5% of households, over a third of households have falls or rises of more than 50% of value.[6] These percentage changes represent large monetary fluctuations and are not supported by corresponding changes in contributions.

In Figure 1.1, wealth excluding pensions shows increases in inequality over time, due to a combination of concentrating non-pension wealth, removal of pension wealth changes not beneficial to the wealthy and sample improvements discussed further on. Given the discussion above, the different changes in inequality over time from the inclusion of pension wealth are likely due to changes in the measurement and valuation of pension wealth rather than changes in holdings or behaviour. The value of pension assets broadly declines in the WAS, due to the modelling changes which produce large individual fluctuations as well as large aggregate changes. It should be noted that the source of these fluctuations and changes is DB pensions - DC pension holders do not experience these changes and there is little difference in wealth concentration statistics or patterns over time when including or excluding DC pensions from the measure of total wealth.

Having considered pensions, let us examine business wealth. Business wealth is important amongst the wealthy - amongst the wealthiest 1%, half hold some form of private business wealth and over a third run their own business or are self employed.[7], similar to US data reported in Cagetti and Nardi [2006], Quadrini [1999] and Quadrini [2000].

The Office of National Statistics [2014] states in their WAS documentation accompanying the dataset that business wealth is excluded from total wealth due to poor response

[6]Note this is not driven by small pension values - less than a tenth of top 5% households have pension wealth below £50,000 and only a fifth below £100,000.

[7]Robust to various definitions of 'self-employed'.

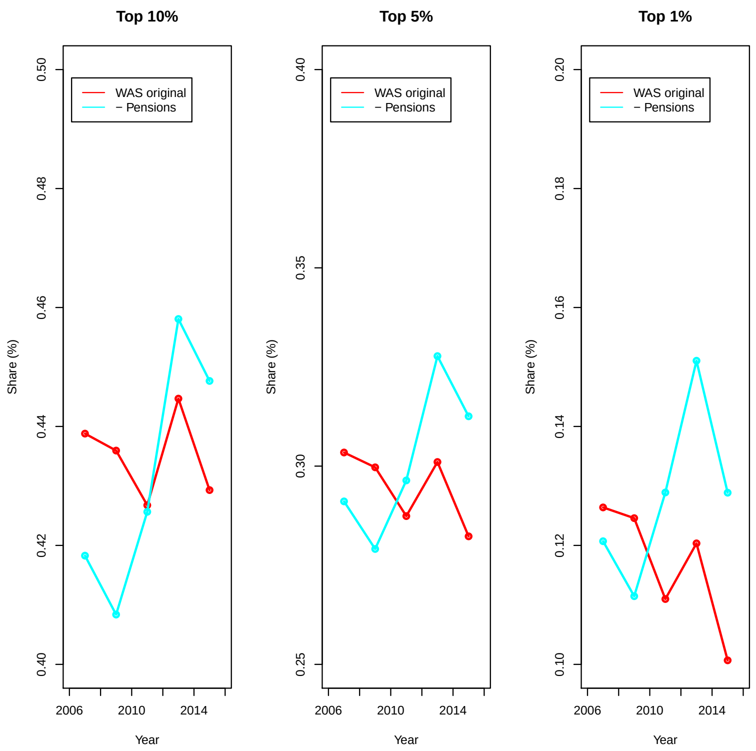

Figure 1.1: Shares of wealth held by top $x\%$ of households - top 10% (left), top 5% (middle), top 1% (right). Original ONS wealth measure (red) and after removing pensions (teal).

rates in initial waves for business valuation, causing concern about representative answers therein. I argue that complete exclusion of the business data by the ONS (implicitly valuing all businesses at zero, which is by definition biased downwards) is a poor answer to the problem when data is available.

I propose using a similar approach to the multiple imputation in the SCF for those respondents not providing numerical business wealth valuation in the WAS cross-sectional estimates. The ONS uses a similar single imputation procedure to impute data gathered with error and for complete non-response, prior to data release. The proposed procedure is focused on the next level, imputing for those who technically 'respond' to the question, but only to state they cannot or will not answer with a value.

Firstly, the WAS improves over the five waves - by wave 4, business value non-response is near 5%. To exclude this later data is unreasonable. The improvement is predominantly due to many recorded non-responses truly being valuations of zero, which can be identified by other survey questions. The response rate is actually 74% and 76% in waves 1 and 2, as can be seen in Table 1.1. This adjustment improves the response rate and identifies the unknown value observations where imputation should be used. The figures are similar to the SCF - over 30% of the 2013 SCF self-employed had zero business wealth, compared to 34-40% in the WAS.

The remaining business data without a valuation ('Bus. Val. unknown' in table 1.1) is

Wave	Bus. Val. >0	Bus. Val. = 0	Bus. Val. Inferred 0	Bus. Val. unknown
2006-08	0.33	0.07	0.34	0.25
2008-10	0.30	0.06	0.40	0.24
2010-12	0.57	0.36	-	0.07
2012-14	0.55	0.40	-	0.05
2014-16	0.54	0.38	-	0.08

Table 1.1: Proportion of self-employed with positive, zero and unknown business value.

a suitable candidate for imputation. Like the SCF, I use multiple imputation by chained equations (MICE). I apply the procedure to impute the specified business wealth, as well as some excluded physical wealth in wave 1 in order to use the full sample. An imputation in this manner requires a good understanding of those not responding to business questions in order to evaluate if the missing data is conditionally-missing-at-random, i.e. that given available covariates and predictive variables for whether the observation is missing, the data is missing at random.

The procedure uses predictive variables to select pools of potential 'matches' for observations with missing values and then draws randomly from this pool. As such, there is a set of random imputee candidates for the missing data. Multiple imputation does not just pick one value, but uses the full array of imputed datasets formed by the procedure, pooling the results from each new dataset in order to construct desired estimates.[8] These results have smaller errors and biases than a single imputation or deleting observations. To see more of the underlying mechanics, arguments and statistical results surrounding MI, consult Rubin [1987] or the SCF Federal Reserve paper concerning multiple imputation, Kennickell [1998].

Business owners that do not provide a valuation are asked why in wave 1 and 2. Some state reasons which directly imply a valuation of (or very near to) zero - the business having no financial assets or no market value. The remainder state they are unwilling to say, unable to say or that they have no records. Some also provide 'other' reasons not available in the end user version of the database. With the variables available in the WAS covering demographic information, income components, asset portfolios and business information, the MICE procedure can control for many features of respondents, increasing the likelihood that the residual data is then conditionally missing at random.

Table 1.2 shows the reasons given by respondents for not providing a value, broken down by a division within the self-employed definition. Considering all self-employed not answering (the full population asked the question, marked row "All SE"), a large number of respondents give a non-response reason implying a direct and clear estimate of zero value, either "the business has no market value" or "the business has no financial assets".[9]. Given

[8]This study uses Van Buuren's MICE package in R, van Buuren and Groothuis-Oudshoorn [2011], with the classification and regression tree options with 5 imputed datasets.

[9]The table covers wave 1 only, but wave 2 is similar

that these can be reasonably inferred to be an observation of 0, a large proportion of data is no longer missing and response rates become vastly improved.

Population	No business financial assets	no market value	unwilling to say	unable to say	no records	other	N
All SE	40%	18%	5%	26%	4%	8%	2479
BSE	20%	15%	9%	44%	3%	8%	942
SEO	52%	20%	2%	14%	4%	7%	1537

Table 1.2: Reasons for non-response ("don't know") to business valuation, Wave 1, proportion of observations.

The type of response is associated with a key distinction in the self-employment categories shown in table 1.2, "self-employed in another way" (SEO) are much likely to state an effective business value of zero under further questioning versus those who are 'sole directors', 'directors' or 'partners', a group I call "Business Self Employed" (BSE).

The BSE/SEO distinction is an example of a predictive variable used in the MICE procedure to ensure any differences between respondents that could cause the data not to be missing-at-random is accounted for. The data shown in this paper uses BSE status; financial wealth; business features (debt, employees, year of creation); education; sample weight and age. Different predictor choices have been tested to establish robustness of results and also, encouragingly, there are no patterns to suggest that characteristics of members of different non-value answer categories ('unwilling', 'unable', 'no records' and 'other') are extremely different.

The distribution of educational qualifications amongst different subsets of business owners is a particularly interesting variable and shown in table 1.3. Breaking respondents down into those with degrees ('degree+'), other qualifications ('other qual') and no qualifications ('no qual'), the respondents who do not answer the valuation question ("No answer") appear to be less educated and mathematically skilled than those with positive business valuations ('Bus. Val.>0'), but about the same as all self-employed ("All SE").

	Bus. Val.>0	BSE	All SE	Bus. Val.=0	No answer	Unwilling to say	Unable to say	No record	other
degree	0.31	0.35	0.29	0.27	0.26	0.22	0.25	0.14	0.35
other qual	0.57	0.55	0.58	0.60	0.61	0.66	0.63	0.66	0.51
no qual	0.12	0.11	0.12	0.13	0.13	0.12	0.12	0.20	0.14

Table 1.3: Education levels across different business ownership status and non-response reasons.

In the BSE subset and amongst those stating 'other' reasons, there is more likelihood of a degree. Interestingly, those 'unable' to say a business value are more likely to hold a degree than those 'unwilling', suggesting the respondents are not likely to be lacking educational skills if they claim not to be able to evaluate the business. Those without business records are substantially less educated, but this difference does not seem to follow through into a similar difference in self-reported mathematical ability in table 1.4. Table 1.4 shows respondents by self-assessed mathematical ability. Those 'unwilling' to provide a valuation claim to be the most mathematically able on average - alongside entrepreneurs with positive business value and the BSE.

	Bus. Val.>0	BSE	All SE	Bus. Val.=0	No answer	Unwilling to say	Unable to say	No record	other
excellent	0.37	0.37	0.31	0.27	0.28	0.38	0.28	0.23	0.25
good	0.46	0.47	0.48	0.48	0.51	0.45	0.50	0.61	0.55
moderate	0.15	0.14	0.18	0.21	0.18	0.13	0.20	0.15	0.17
poor	0.03	0.02	0.03	0.04	0.02	0.02	0.02	0.01	0.02

Table 1.4: Self-reported mathematical ability levels across different business ownership status and non-response reasons.

Overall, there appears to be no straightforward relationship between the non-valuation reason and ability proxies. Those who are in the BSE subset are better educated and claim to be more mathematically able, as do those with positive business valuations (and there is substantial overlap between the two).

Including business wealth in the measure of total wealth after removing pensions raises the household top 1% share of wealth by several percentage points in all waves, as shown in Figure 1.2. The effect of the business value imputation is to further raise the share, except in wave 1, where the effect of an accompanying physical wealth imputation and the resulting larger usable dataset ends in a roughly equal top 1% share.[10] By the fourth wave, wealth shares of the top 1% including business wealth imputation and excluding pensions exceeds 20%.

As indicated above, I also perform multiple imputation for a second variable in wave 1. Physical wealth (vehicles, household contents, collectibles) was only collected for half of households in wave 1, so the ONS excludes the other half, leaving only 50% of the sample. This is a dramatic loss of data, especially for the necessarily small subsample of the top percentiles, and physical wealth is a small part of total wealth (less than 10% of aggregate wealth). I therefore perform multiple imputation for household physical wealth in wave 1 to use of the full wave 1 sample both in the general analysis and in business imputation, leading to a robust dataset. There are higher WAS top wealth shares when excluding physical wealth as, like the SCF findings of Wolff [1987], physical wealth is mainly important at the bottom. In total, removing physical wealth raises the top 1% share by approximately 2 percentage points.

Later waves of the WAS include refreshment samples which have a greater rate of oversampling of the wealthy (3x versus 2.5x) and developments to the process of identifying the wealthy. Figure 1.3 shows a calculation for the share of the top 1% using only the initial sample over the 4 waves, using the original sample plus those recruited in wave 3

[10]Both imputations are shown together, so all observations for business wealth can be used in imputation.

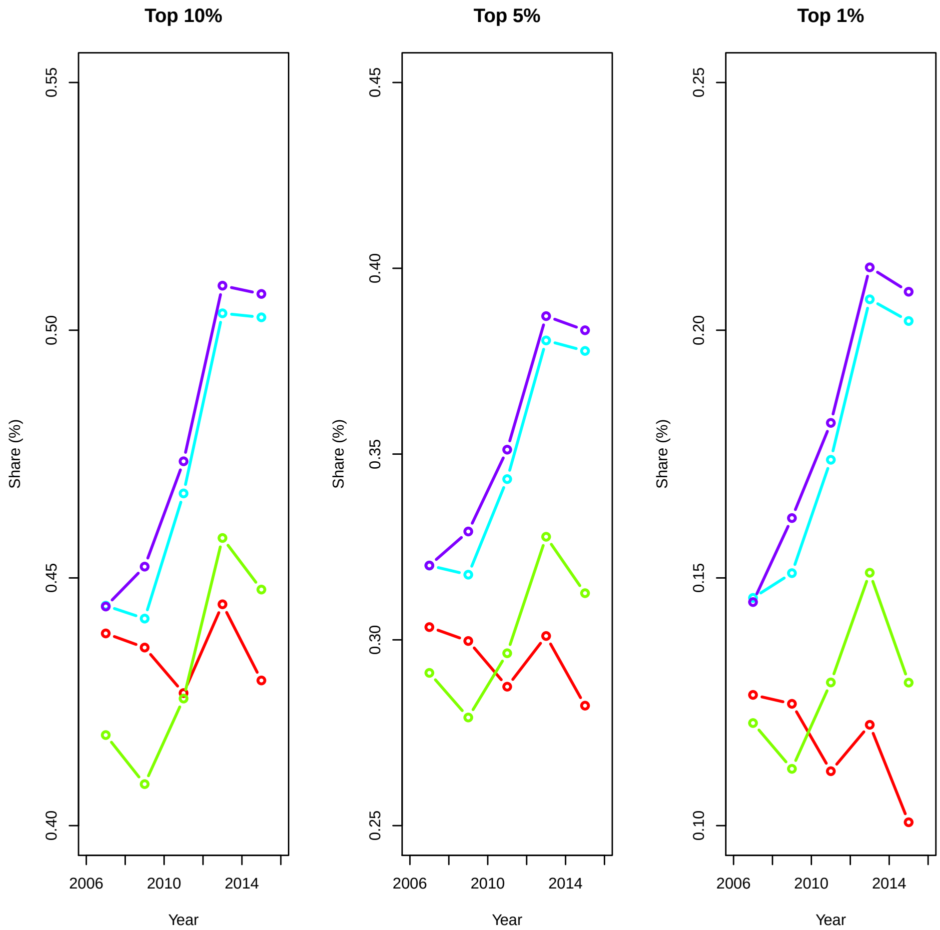

Figure 1.2: Shares of wealth held by top x% of households - top 10% (left), top 5% (middle), top 1% (right). Original ONS wealth measure (red) then, cumulatively, removing pensions (green), including business wealth (teal) and, finally, imputation of business wealth (purple).

and then the entire available sample in wave 4. Excluding all new entrants (i.e. without sample changes), the share of the top 1% experienced a 2007-2009 rise, then a slight fall in 2011-2013. Adding the new wave 3 sample jumps the share, but thereafter the combined group falls much like the original sample. The new wave 4 sample also increases the share by their entry, to 22%. The new households from the refreshment samples are wealthier and increase the wealth share of the very wealthy, accounting for the majority of the rise in those wealth shares since 2009 when using the entire sample (as opposed to calculations using the original sample survivors).

The changes in the WAS with newer refreshment samples suggest validation of the data is a useful exercise. UK estate data contains the officially valued wealth of all deceased persons over a wealth threshold[11]. It covers the top 40% of the population and so can be used, with further assumptions, for calculating the wealth of the top 10, 5, 1 and 0.1 percent groups. For the period WAS data covers, the estate data generate top 1% shares of 18-20% in Alvaredo et al. [2017]. It is therefore a very useful comparator to consider.

[11] approximately £5000 for modern data

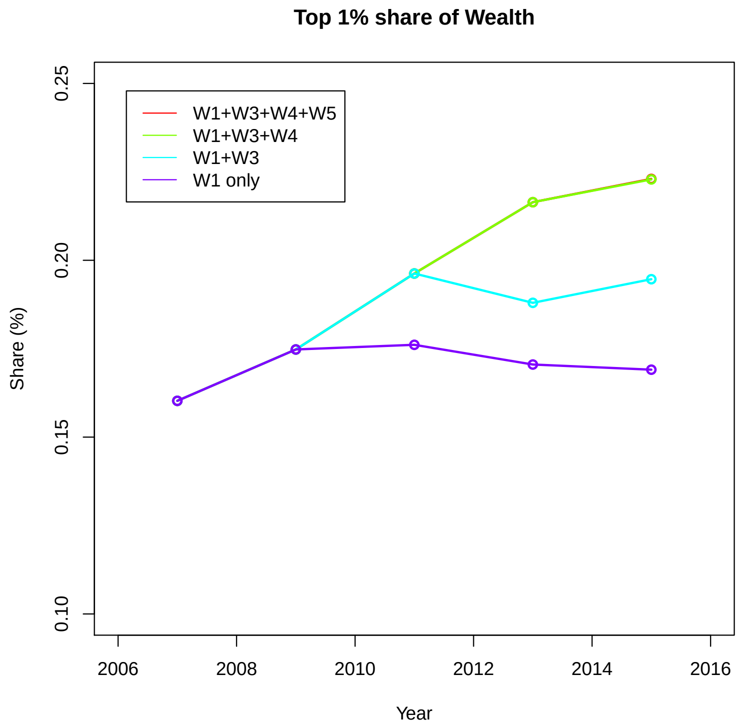

Figure 1.3: Share of wealth held by top 1% individuals, using groups starting in wave 1 (W1 only), wave 1 or wave 3 (W1 + W3), waves 1, 3 or 4 (W1 + W3 + W4) or any wave (W1 + W2 + W3 + W4 + W5).

There are several explanations for the estate data - WAS discrepancy noted in Crossley et al. [2016], which used the ONS original measure of total wealth. As noted above, the published WAS top 1% wealth share under that measure is only 12%. Firstly, there are differences with the measure. WAS reported total wealth excluded business wealth and included pension wealth, whereas estate data does the opposite. Secondly, the WAS total wealth measure is at household level, not at person level as per estate data. Therefore, in my comparison, I adjust WAS data to be comparable to estate data as far as possible to meet both these points.[12]

There may also be different coverage of the wealthy and their assets - there are universal legal valuation requirements at death for estate data but widespread avoidance and evasion. WAS relies on personal interview valuations instead. The WAS has the benefit of including data for lower parts of the wealth distribution and for items such as jointly-held main residences, whereas the estate data includes neither, so must rely on national accounts and assumptions. Lastly, WAS avoids the selection inherent in valuing only at point of death, which Alvaredo et al. [2017] attempt to control for using 'mortality multipliers'.[13]

The wealth shares are constructed by summing wealth in a given group and dividing

[12]See the appendix for details.
[13]These factors weight the assets of the deceased by an inverse probability of death to recreate a representative sample of the living population.

by the sum of all wealth so I now consider wealth shares before examining the wealth levels that contribute to the numerator and denominator of those wealth share figures.

The ONS WAS top wealth shares in Figure 1.4 are below those for estate data in all years, whilst the adjusted sequence I create from the WAS to be comparable to estate data shows substantially higher inequality than the originals and rises to become approximately as unequal as estate data over the waves.

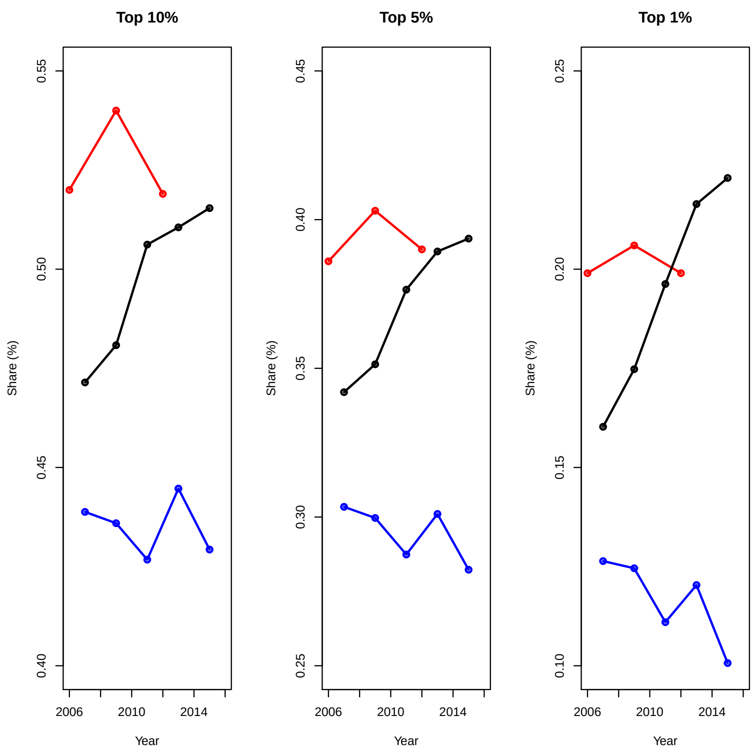

Figure 1.4: Shares of wealth held by top $x\%$ of households - top 10% (left), top 5% (middle), top 1% (right). Estate data from Alvaredo et al Alvaredo et al. [2017] (red), WAS original base measure (blue), adjusted sequence (black).

Despite the lower wealth shares for the adjusted sequence in early waves, the *levels* of average wealth for the comparable wealthy in the WAS have always been equal or higher than estate data. Figure 1.5 displays the average wealth of different top quantile groups for the estate series detailed in Alvaredo et al. and the new WAS equivalent series.[14]

The average wealth held by the WAS top wealth groups exceeds the comparable equivalents for the estate data throughout - i.e. the rich are richer in the WAS. In wave 1, the WAS average is similar to that of the estate data for the top 1%, at £2.6m as opposed to the estate data figure of £2.5m. The refreshment samples are responsible for the dramatic rise in levels of wealth in wave 3 onwards, but the original wave 1 sample top x% is still wealthier than the top x% in the estate data throughout. The WAS does not under-represent wealth at the top versus estate data, despite lower top wealth shares.

Given these higher average levels of wealth amongst the wealthy, the initially lower

[14] Average wealth in top groups is a rescaled numerator for the equation generating the wealth share statistic.

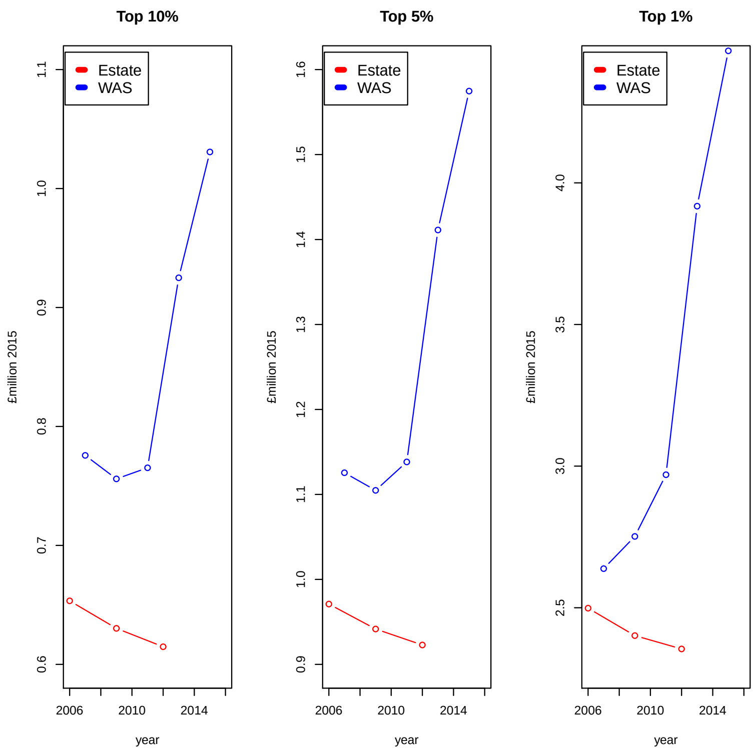

Figure 1.5: Average personal wealth for the top 10% (left), 5% (middle) and 1%(right) in WAS (blue) and estate data (red)

WAS shares in figure 1.4 must be due to the denominator - aggregate wealth. As shown in figure 1.6, the total wealth of the population calculated from the WAS is higher than estate data and closer to the marketable wealth total created from estate data by tax authority HMRC, "Series C". The marketable wealth series is not necessarily 'better' than the estate total as, although attempted corrections in Series C are designed to include items the estate data may not take into account, it contains some assumptions and corrections sourced from brief and outdated data (Crossley et al. [2016]).

It is not easy to identify exactly where in the distribution the WAS finds extra wealth to contribute to this larger total in figure 1.6 - estate statistics do not include the lowest parts of the distribution for direct comparison. The assumptions made about those least wealthy excluded from the estate statistics do not appear to account for lower aggregate wealth versus the WAS - on a simple comparison, such persons are actually poorer in the WAS than in the estate data[15]. Incomplete recording of jointly held housing from estates is one potential explanation for these differences in the wealth distribution.

In the estate data analysed by Alvaredo et al. [2017], housing wealth is around 30% of the assets of the top 1% in estate data over the period the WAS covers. In contrast, the WAS figure is approximately 45% in wave one, which (monotonically) falls to 30% by the fourth wave. As shown in figure 1.7, WAS wealth further excluding housing (i.e. leaving

[15]This is approximately calculated by comparing 'excluded wealth' from the estate statistics with wealth of the lowest x% in the WAS, where x is the share of people not covered by the estate data, i.e. assuming the poorest in the WAS are also at the bottom for probate recording.

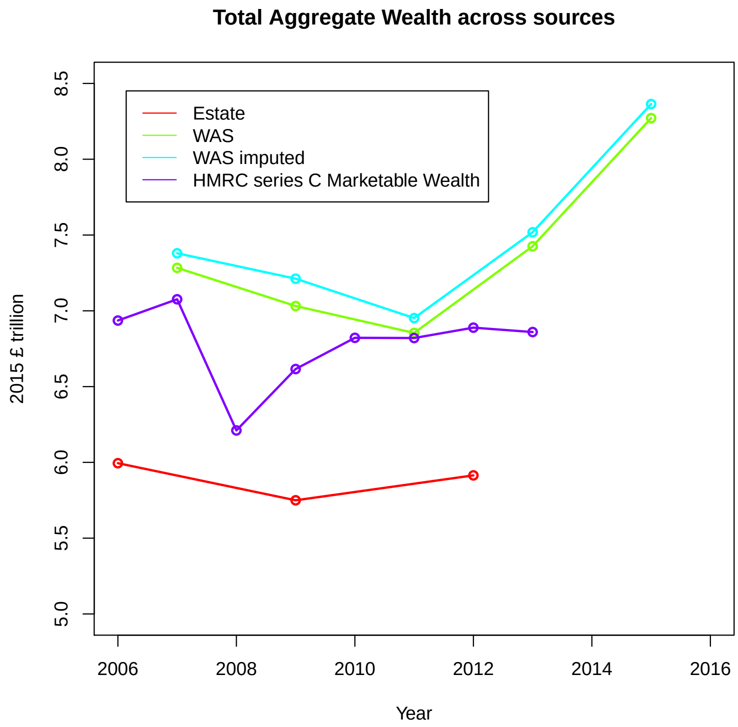

Figure 1.6: Aggregate Real Wealth Totals from Alvaredo et al Alvaredo et al. [2017], HMRC Series C Marketable Wealth and WAS.

only business, financial and physical wealth) has a top 1% share of around 30% and is a similar level to estate data in later waves. Again, the WAS equals or exceeds the estate data in levels of wealth for the top quantiles. The wealthy are still wealthier in the WAS when excluding housing, despite lower shares of wealth for top groups.

A second, brief, validation is also of interest - income data in the WAS. Compared to the World Income Database (WID) UK figures, there are similar top shares of total income among adults, with top 1% shares of around 12% before tax and 9% after tax. When restricting total income to taxpayers only, the WAS broadly matches the UK Survey of Personal Income administrative data (SPI), except for near to the top 1%, where the WAS quantile is around 10% lower.

For investment income amongst taxpayers, the top 10% consistently have 85-90% of investment income in the WAS, whilst the SPI in Alvaredo et al lists approximately 75%.[16]. The WAS investment income and rent total is broadly the same, when adjusting for the exclusion of 'non-profits serving households' (NPISH) in the WAS data.[17]

The identification and processing of 'taxpayers' in the WAS is likely to be a major source of differences as, in end-user data, one needs to infer/model tax payments from different income items, some of which are gathered pre-tax and others post-tax. Further

[16]except in years of tax reform

[17]The WAS total is £50,000m in wave 3, whereas the comparable figure from national accounts/SPI data is £75,000m. NPISH consistently receive 20-30% of the total investment income when statistics are available. Thus a reasonable benchmark to total WAS investment income is £55,000m. Around 5% of responders did not know their investment income, so the WAS figure is close.

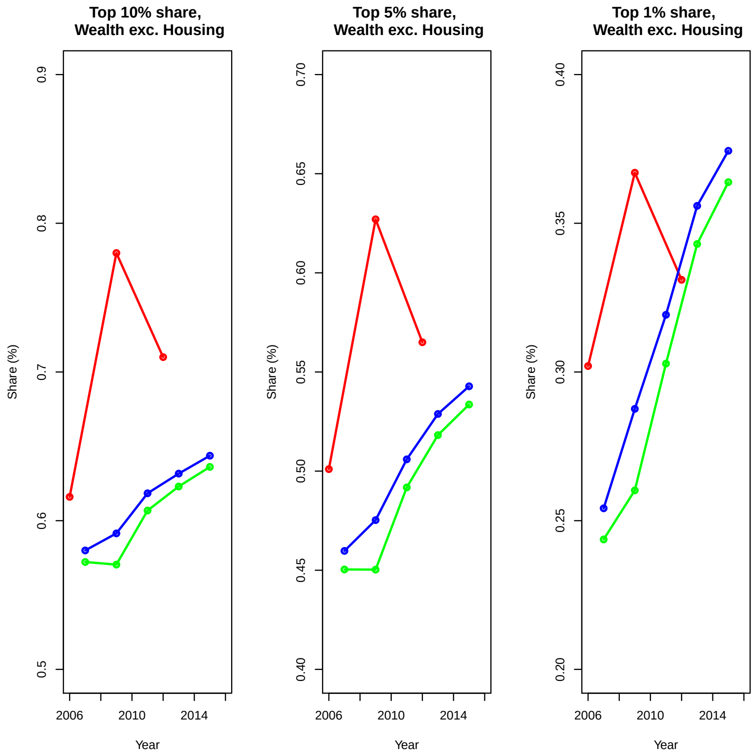

Figure 1.7: Wealth shares of Top 10, 5 and 1%, excluding housing. Estate data from Alvaredo et al. [2017] (red), WAS person level including imputation(blue) and without (green).

discussion of this procedure and explanation for why these figures are likely to slightly overtax the wealthy can be found in the Appendix.

As a further comparison, the Sunday Times Rich List[18] aims to identify the richest 1000 persons or households in the U.K.[19] and catalogue their wealth by a combination of interviews, compiling publicly available information and their own investigations. The vast majority of this wealth is privately held businesses, public stocks or property. The team excludes wealth in private bank accounts and other sources they cannot access. They endeavour to value the private businesses and track changes in ownership. The wealth in this list is substantially above that of the survey - the top 1000 is less than 0.002% of the population and the entry threshold for the 2018 group is over £110 million. In comparison, the top survey member has less than £50 million in wealth. Alvaredo et al. [2017] catalogue estimated Pareto coefficients based upon this list and from estate data, each of which we compare to the survey data below. The differing scale makes the rich list and survey difficult to compare, except by Pareto coefficients. The Pareto, or 'Power Law', distribution describes a fat tail above a threshold c, with the following (scaled) cumulative distribution function in that tail,

[18] https://www.thesundaytimes.co.uk/richlist
[19] U.K. resident or with a significant presence there

$$P(X \leq x) = \left(\frac{c}{x}\right)^{\alpha}$$

with the coefficient α. α is an important parameter defining the distribution and its relation to top wealth shares is one motivating reason for the focus on top shares of wealth in the literature. α can be estimated in a number of ways - graphical log-log plots and regressions; maximum likelihood fitting and shares within shares. To be easily comparable, we follow Alvaredo et al. [2017] and use the shares-within-shares method for the WAS to find the inverse coefficient $\beta = \alpha/(\alpha - 1)$. β is positively correlated with the fatness of the tail and inequality in terms of higher wealth shares.[20] The Pareto distribution assumes that β is constant at all points within the tail - something that I find, like Alvaredo et al. [2017], does not hold. In Figure 1.8 we see that the WAS shows an equal or higher β than the rich list data[21], meaning there is equal or higher inequality amongst the wealthy in the survey versus the super-rich in the rich list. We also see in Figure 1.9 that for the available estate data the WAS has similar Pareto-Lorenz coefficients[22]. The WAS data has a higher β for data in higher parts of the distribution as time passes in line with the arguments earlier concerning better sampling and an increase in wealth inequality amongst the very wealthy.

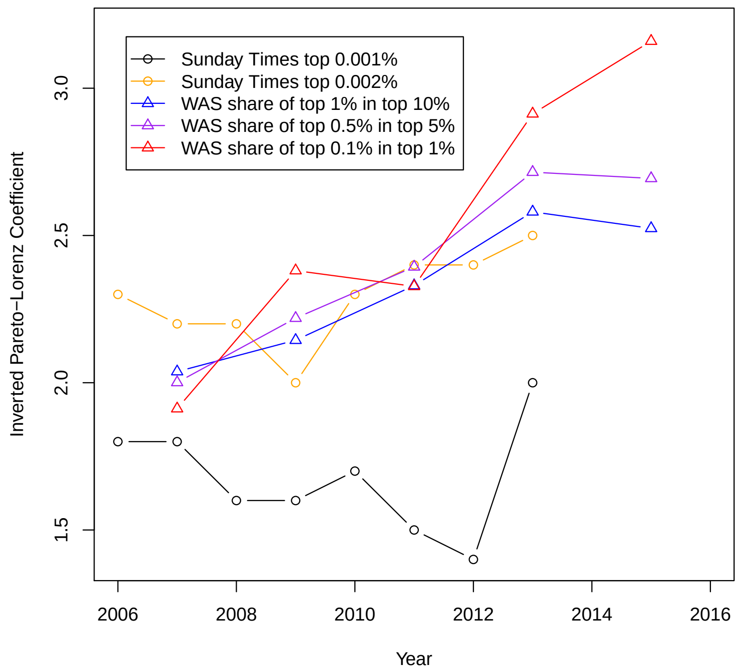

Figure 1.8: Inverted Pareto-Lorenz coefficients for the WAS and the Sunday Times Rich List.

[20] Whereas α is inversely correlated with the wealth share.
[21] Pareto coefficients reproduced from Alvaredo et al. [2017]
[22] Pareto coefficients reproduced from Alvaredo et al. [2017]

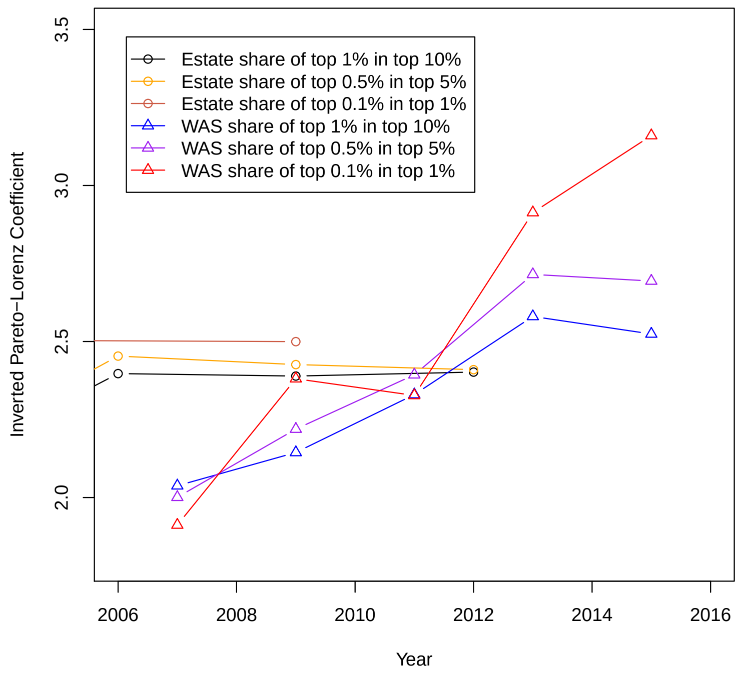

Figure 1.9: Inverted Pareto-Lorenz coefficients for the WAS and Estate Data.

Finally, one important cross-sectional inequality comparison is the work of Vermeulen [2016], who investigates the ability of wealth surveys to represent the wealthiest by combining the survey data with wealth information from rich lists of billionaires to estimate power law parameters. For the WAS, he estimates an original wealth share for the top 1% in wave 2 of 13% and 14-18% when including the rich list data. My analysis has shown that the WAS top 1% wealth share fits into this range after adding business wealth even with no other changes.

To conclude this section, stripping out refreshment sample effects, there has been a small rise in top wealth shares over 2007-2015. Rising top wealth shares over the period is otherwise attributable to better sampling of the wealthy in later waves. I claim the WAS data is suitable for the study of top of distribution inequality in all waves by validation against estate data and available rich list data. The WAS represents the wealthy at least as well as the estate data in levels of wealth, as lower top shares in early years are due to higher aggregate wealth in the survey.

1.4 Transitions and Mobility

1.4.1 Wealth Mobility

In this section, we study transitions and changes in wealth for households. Unless otherwise specified, we use the measure defined in the previous section of financial, business, property and physical wealth. The following results are robust to different wealth definitions and longitudinal weighting schemes.

Table 1.5 presents various transition probabilities for different groups of wealthy households commonly studied in the literature[23]. More than one third of the top percentile exit in two years, whilst the 8 year exit rate is 50%.[24] Membership in higher percentile groups (going right across table 1.5) is generally more unstable and there is greater probability of transitioning downwards, despite larger gaps between the higher thresholds.

Years	Top 10%	Top 5%	Top 1%	Top 0.1%
07-09	0.72	0.68	0.58	0.41
09-11	0.77	0.73	0.64	0.5
11-13	0.79	0.74	0.67	0.44
13-15	0.79	0.75	0.71	0.49
07-15	0.65	0.64	0.52	0.5

Table 1.5: Proportion of households staying in top wealth quantile groups across waves

I find similar patterns in both the U.K. ELSA and both the 07/09 and 1983/89 SCF (latter from Kennickell and Starr-McCluer [1997]) as shown in Table 1.6 and 1.7 The SCF 07-09 transitions are similar to the WAS 07-09, but show less mobility than the WAS, while the 83/89 SCF is more mobile than the WAS 6-year transitions (though this may be due to the different eras). I also note that in Hurst et al. [1998], the PSID has a proportion staying in the top 10% over 5-years of 64%-69%, which is similar to the WAS 6-year staying rate of 65%-72%. ELSA's staying rates are somewhat lower than those in the WAS, but the ELSA figures are otherwise similar, despite being drawn from a smaller sample which does not use oversampling to represent the very top. The ELSA sample for the top 0.1% is very small and thus highly variable, a problem which the WAS also has above the top 0.1%, though to a lesser degree.

Source	Top 10%	Top 5%	Top 1%	Top 0.1%
SCF 07-09	0.78	0.81	0.66	0.56
WAS 07-09	0.72	0.68	0.58	0.41
SCF 83-89	0.41	0.52	0.59	
WAS 07-13	0.65	0.64	0.52	0.5
WAS 09-15	0.72	0.65	0.57	0.42

Table 1.6: Proportion of households staying in top wealth quantile groups across waves,WAS and SCF

Studying exit from the top categories over different horizons in Table 1.8, we see that the proportion remaining in the category decreases as the time horizon expands, so there is a lower probability of staying for a longer time. However, after the first transition (of 2 years), further transitions have a much lower impact on the proportion remaining. Considering the 2007 start, whilst there is a 42% chance of exiting in the first two years,

[23]Note that for this dynamic analysis, the multiple imputation for business value from above cannot be used, as it is only valid cross-sectionally.

[24]much above the (Markov) compound of individual wave-to-wave staying probabilities

	top 10%	top 5%	top 1%	top 0.1%
1-2	0.71	0.65	0.52	0.10
2-3	0.73	0.70	0.45	0.51
3-4	0.73	0.66	0.46	0.18
4-5	0.74	0.64	0.49	0.00
5-6	0.73	0.65	0.44	0.00
6-7	0.74	0.70	0.59	0.00

Table 1.7: ELSA: Staying rates in top wealth groups over waves.

the chance for exiting in 8 years is 52%. The transition over the financial crisis of 2007-2009 has higher exit rates and mobility, but the 2007 cohort retains a similar pattern as the time window expands over multiple waves. We also find the same pattern in ELSA's statistics, although not at the level of the top 1%.

Years	Top 10%	Top 5%	Top 1%
07-09	0.72	0.68	0.58
07-11	0.71	0.68	0.59
07-13	0.68	0.63	0.55
07-15	0.65	0.64	0.52
09-11	0.77	0.73	0.64
09-13	0.76	0.7	0.6
09-15	0.72	0.65	0.57
11-13	0.79	0.74	0.67
11-15	0.77	0.72	0.61

Table 1.8: Proportion of households staying in top wealth quantile groups over time

As an illustration of the substantial wealth fluctuations involved in these transitions, I show the quantiles of the percentage change distribution for the top 5% in table 1.9 and the top 1% in table 1.10. I note the very substantial losses indicated by the lower quartile and lowest decile - for Decile 1 ($Q(0.1)$), 45-60% of wealth lost, for reference this loss is around £600-800,000. There is an even wider distribution of losses amongst the top 1%, with a quarter losing over 35% of their wealth. For both groups, the distribution of changes is shifted downwards during the 07-09 transition which coincides with the financial crisis, particularly affecting the top 1%.

Years	Q(0.1)	Q(0.25)	Q(0.5)	Q(0.75)	Q(0.9)
07-09	-0.6	-0.34	-0.09	0.14	0.42
09-11	-0.46	-0.23	-0.02	0.19	0.54
11-13	-0.49	-0.24	-0.01	0.18	0.53
13-15	-0.48	-0.2	0.03	0.24	0.54

Table 1.9: Quantiles of Proportional Changes in Wealth for Top 5%

Table 1.11 displays before and after statistics for those in the top 5% who experience a fall of 25% or more in their wealth between two waves versus the remainder of the top 5%[25]. The self-employed are over-represented in those with large falls and a substantial

[25] I use the fourth and fifth wave, though other waves are similar, as are the averages.

Years	Q(0.1)	Q(0.25)	Q(0.5)	Q(0.75)	Q(0.9)
07-09	-0.78	-0.55	-0.23	0.12	0.56
09-11	-0.65	-0.35	-0.03	0.23	0.72
11-13	-0.69	-0.39	-0.07	0.17	0.52
13-15	-0.77	-0.38	-0.03	0.22	0.65

Table 1.10: Quantiles of Proportional Changes in Wealth for Top 1%

proportion of these exit self-employment after their fall. I show the median ($Q_{0.5}$) and top quartile ($Q_{0.75}$) of the proportion of total wealth held as business wealth amongst these self-employed in the second and third row. On the left, the 'before' figures show big fallers have a larger proportion of their wealth in their business (versus the other self-employed in the top 5%) before their fall. After their fall, their wealth in their business is substantially reduced.

Those big fallers with large proportions of financial wealth (75th percentile and above in terms of proportion of wealth held in financial wealth) before the transition experience a large reduction in that proportion, roughly halving the size of their financial portfolio versus their other remaining assets. Big fallers have a lower allocation towards property wealth, the proportion of which rises after their fall, indicating their non-property assets are having greater reductions than property assets[26].

	Fall >-25%		Others	
	Before	After	Before	After
Proportion Self-employed	0.34	0.21	0.27	0.25
Self-employed $Q_{0.5}$ % bus. wealth	0.44	0	0.03	0.06
Self-employed $Q_{0.75}$ % bus. wealth	0.76	0.27	0.41	0.42
Median % financial wealth	0.18	0.13	0.24	0.21
$Q_{0.75}$ % financial wealth	0.58	0.27	0.42	0.4
Median % housing wealth	0.35	0.66	0.56	0.59

Table 1.11: Statistics for subsets of Top 5%, before & after transitions.

Examining income, we find a similar pattern for top incomes in table 1.12 when compared to the top wealth transitions in table 1.5. The figures are broadly similar to rates of exit from top income groups in the US found by Guvenen et al. [2014a], Auten et al. [2013] and Kopczuk et al. [2007].[27]

One of the full transition matrices generating the staying rates can be seen in table 1.13, depicting agent's current status by row, and future status given current status in columns. Whilst there is concentration around the diagonal - i.e. that larger moves across wealth categories are less likely than smaller moves, there is significant likelihood of falling very far down the wealth ladder. Amongst the 48% that leave the top 1% over 8 years in Table 1.14, 20% fall out below the top 5%, and this represents a loss of at least several

[26]Note WAS cannot distinguish between asset sales for consumption and intrinsic losses. Thus the tendency to sell other assets before illiquid property may be showing here.

[27]One should note that the top x% in wealth and top x% in income are not all the same people when interpreting these patterns. About half of these respective top 1%'s overlap.

Years	Top 10%	Top 5%	Top 1%	Top 0.1%
07-09	0.62	0.54	0.27	0.28
09-11	0.61	0.55	0.44	0.42
11-13	0.61	0.57	0.6	0.48
13-15	0.62	0.57	0.5	0.59
07-15	0.46	0.4	0.25	0.42

Table 1.12: Proportion of households staying in top gross income quantile groups across waves

million pounds. In short, wealth can be very volatile, even for the wealthy. This also aligns with SCF 07/09 panel findings from Bricker et al. [2011] and the 1980's results from Kennickell and Starr-McCluer [1997]. Other years in the WAS show similar patterns.

from/to	<top 10%	top 5-10%	top 1-5%	top 1%
<top 10%	0.97	0.02	0.01	0.00
top 5-10%	0.37	0.47	0.15	0.00
top 1-5%	0.10	0.23	0.60	0.07
top 1%	0.02	0.01	0.36	0.61

Table 1.13: Transitional Probabilities for HH WAS wealth categories 09-11.

from/to	<top 10%	top 10%	top 5%	top 1%
<top 10%	0.96	0.03	0.01	0.00
top 5-10%	0.53	0.31	0.16	0.00
top 1-5%	0.20	0.22	0.49	0.09
top 1%	0.12	0.02	0.37	0.5

Table 1.14: Transitional Probabilities for HH WAS wealth categories across waves 1-5 (07-15), household wealth.

There is a strong persistence in continued membership of top wealth categories despite the relatively high group exit rates from wave to wave. Table 2.2 shows the probability of staying conditional on history of membership. Those with longer past membership have a much higher probability of remaining in the group, whereas new entrants have a very high chance of exit.[28] Again, ELSA data contains similar findings which are shown in the Appendix.

	top 10%	top 5%	top 1%
$P(T_4 \mid F_1 F_2 T_3)$	0.48	0.39	0.30
$P(T_4 \mid F_1 T_2 T_3)$	0.75	0.68	0.66
$P(T_4 \mid T_1 T_2 T_3)$	0.91	0.88	0.87

Table 1.15: Probability of remaining in top wealth groups given different histories. 'T_t' indicates 'True' for belonging to the group in wave t and 'F_t' indicates 'False' for the same.

We can study more of the distribution of individual wealth changes using a non-parametric quantile regression and plots of resulting quantiles, similar to Trede [1998].

[28]The top 0.1% must be excluded from this conditional analysis as there are too few observations for some categories.

	top 10%	top 5%	top 1%
$P(T_4 \mid F_1F_2T_3)$	0.34	0.29	0.29
$P(T_4 \mid F_1T_2T_3)$	0.69	0.70	0.45
$P(T_4 \mid T_1T_2T_3)$	0.87	0.82	0.67

Table 1.16: ELSA: Conditional staying rates in top wealth groups (3-stage). Notation as per previous table.

The different quantile levels at each x-axis point show the distribution of outcomes at that point. Thus figure 1.10 shows the deciles of wave 3 wealth at each level of wave 2 wealth, much like a series of localised box plots. As an example, households at 4 times median wealth (x-axis=4) in 2009 have a wide range of outcomes - the top 10% (violet, D9) of those households have 5x median wealth in 2011, whilst the lowest 10% (red, D1) have approximately 2.5x median wealth. Considering the whole figure, the range of wealth changes increases as wealth increases.

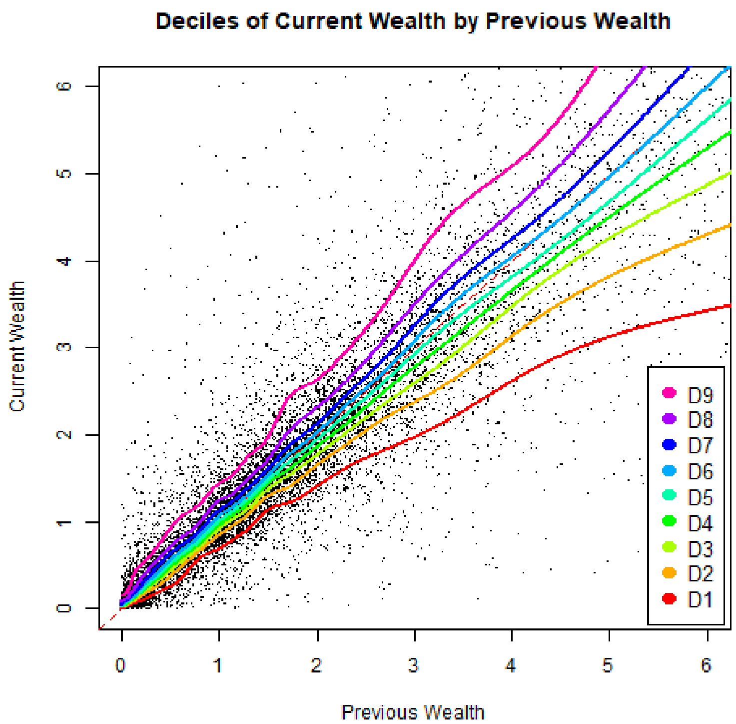

Figure 1.10: Non-linear Quantile Regression for relative-to-median wealth in 2011 vs 2009. Deciles (D1-9) of W_t for a given W_{t-1}.

The patterns in Figure 1.10 are representative of results from other waves and time horizons, as all are very similar. Further diagrams can be found in the Appendix.

Both the slope and the spread of the quantiles indicate mobility features. A lower slope would imply greater mobility, due to weaker local-linear dependence on previous wealth[29]. Similarly, a greater distance between different quantiles at a point would also

[29] An illustrative example is completely horizontal quantile lines, which implies full independence of current wealth from previous wealth, as all across the x-axis face the same outcome quantiles and probability distribution

indicate greater mobility through variance for a household at that point.

Under this analysis, as wealth data shows neither substantially lower slopes nor greater spread over longer time horizons, it indicates that mobility is not substantially higher over the longer horizon of 8, 6 or 4 years versus 2 years.

The wealth quantile dependencies show in Figure 1.10 are well described by a linear relationship (as shown in Appendix). Linear quantile regression coefficients can be easily compared over time horizons to describe mobility differences. Using a quantile regression for deciles of current wealth given wealth in 2007, we observe that both the intercepts and slopes for the 2007 wealth variable are very similar, as shown in Figure 1.11. This also indicates that mobility does not greatly increase over 8 years as opposed to 2. However, there is an increased spread of slope coefficients for the 6 and 8 year regressions, and so there is some evidence of greater mobility in wealth over longer periods of time.

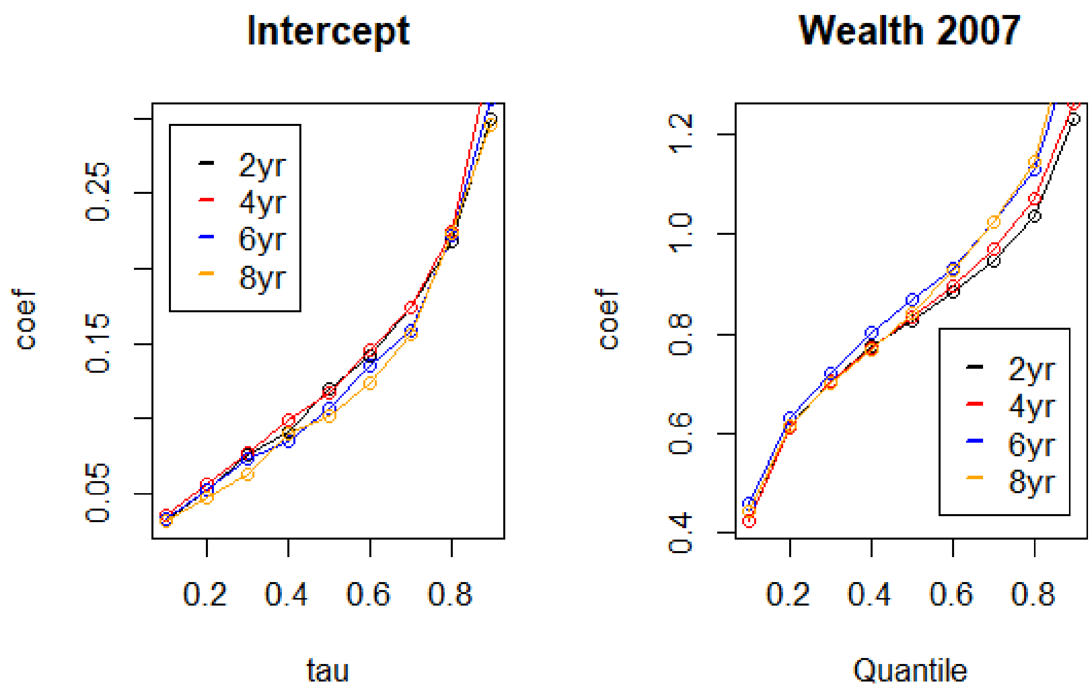

Figure 1.11: Linear Quantile Regression Coefficients for Relative Wealth on Relative Wealth 2007.

Figure 1.12 shows changes in log wealth versus the quantile of previous wealth in the distribution. The distribution of changes is quite substantial over the whole distribution of wealth (from the lowest percentile to highest), with many households gaining or losing 0.25 or 0.5 log points of wealth. Of particular importance, the very wealthiest have a much wider, and slightly lower, $\Delta log(w)$ distribution, whilst the poor below the 4th Decile have a wide but much more positive distribution of proportional wealth change outcomes. For households from the 4th Decile to the 9th Decile, the distribution of log wealth changes faced is broadly the same. Of particular note is the very substantially larger negative tail for the top 2%.

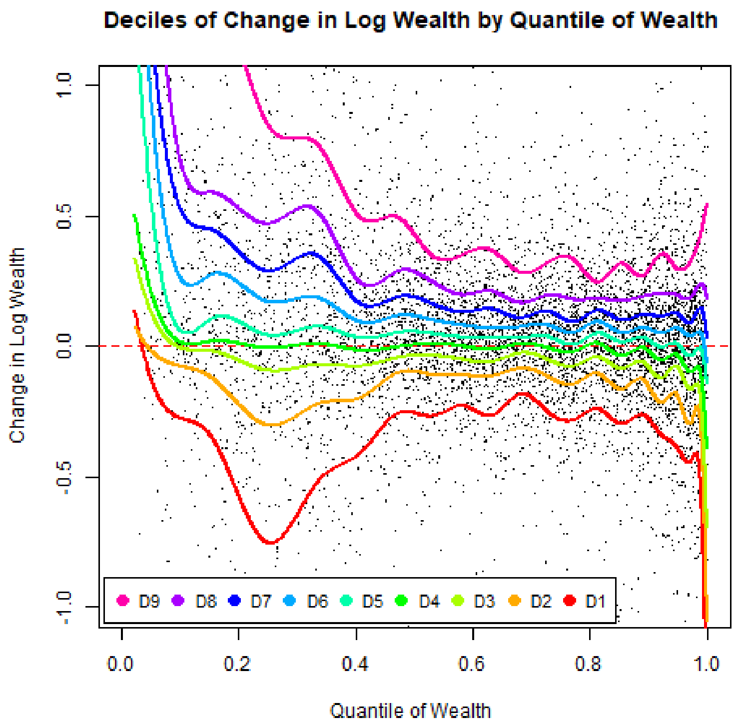

Figure 1.12: Non-Linear Quantile Regression for Changes in Log Wealth vs Quantile of Wealth.

Figure 1.13 shows the first four moments of changes in log wealth, conditional on wealth quantile (using kernel methods). Visually, readers can note the remarkable general similarity to moments of change in log income distributions found in Guvenen, Karahan, Ozkan and Song's study of SSA earnings data (Guvenen et al. [2015]) - variance and skew both U-shaped with the latter negative, whilst kurtosis is substantial and hump-shaped. [30] The results for the mean and variance particularly align with those from the previous quantile diagrams, whilst the skew and kurtosis are less clearly intuitable. We do not observe the increase in negative skew that is observed in the American earnings data over the recessionary transition.

[30] Despite this being in different countries, for wealth rather than income and for households rather than tax units.

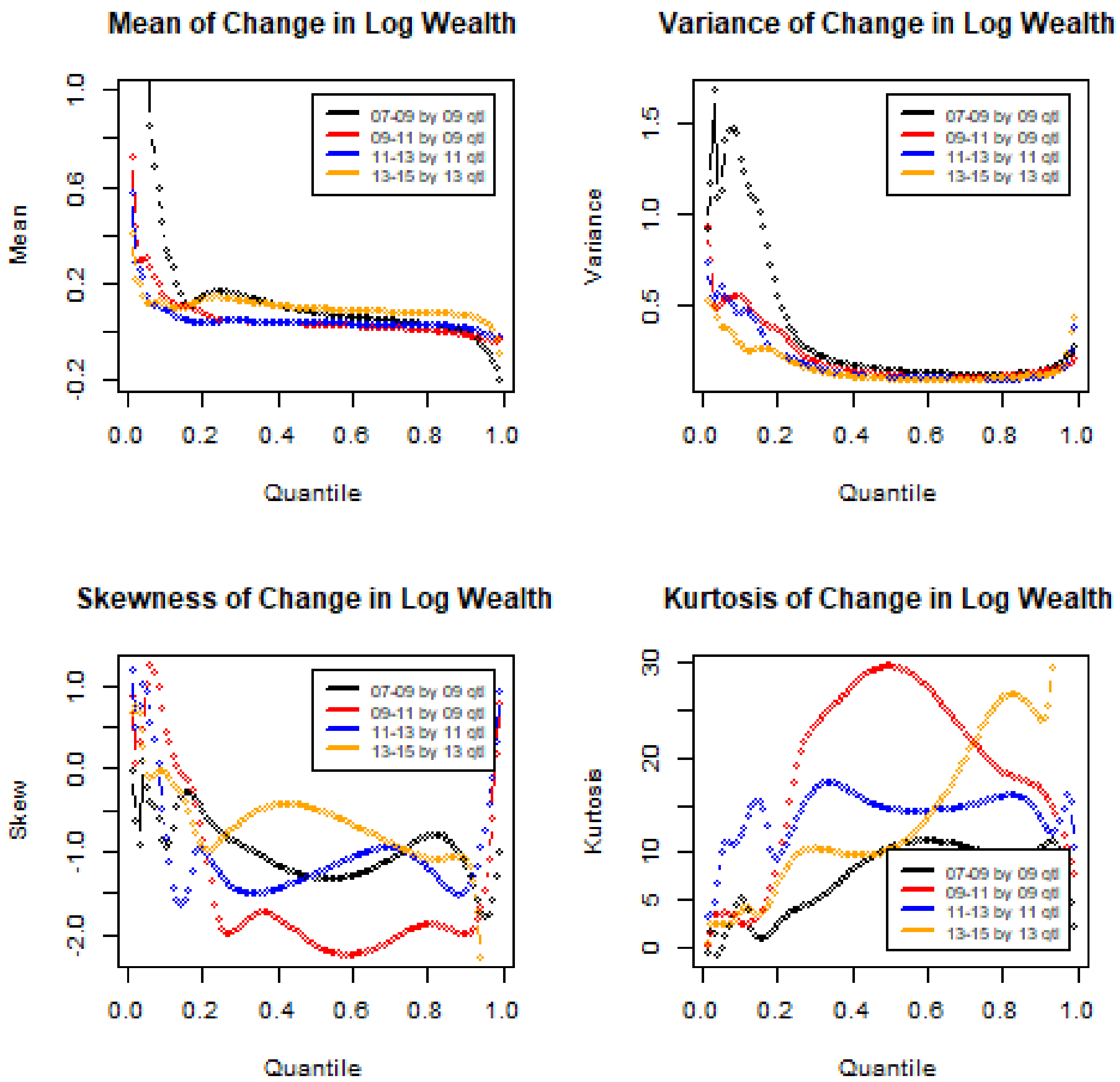

Figure 1.13: Moments for the Change in Log Wealth distribution over quantiles of previous wealth: $\Delta log(W_t)$ by $\tau = F_{W_{t-1}}(W_{t-1})$

Since there are 5 waves, one can also consider the distribution of changes in log wealth conditional on previous changes in log wealth in Figure 1.14. There is some reversion, shown by the generally negative slope of the quantile functions, but there is also a spread of quantiles further from the x-origin in both directions. This can be interpreted as those households experiencing large changes then continuing to experience large changes, regardless of direction[31]. This dependence weakens over a longer horizon when one compares the 11-13 vs 07-09 diagram versus the 09-11 vs 07-09 diagram displayed. The longer horizon plot is both flatter and relatively smaller in spread, appearing closer to independence[32].

[31] Although the bottom 20% and top 10% in wealth are overweighted for $\Delta log(W) > 0.5$ and $\Delta log(W) < -0.5$ respectively, removing these high and low wealth observations does not change the findings.
[32] As mentioned earlier, independence shown in this visualisation would be a series of horizontal lines.

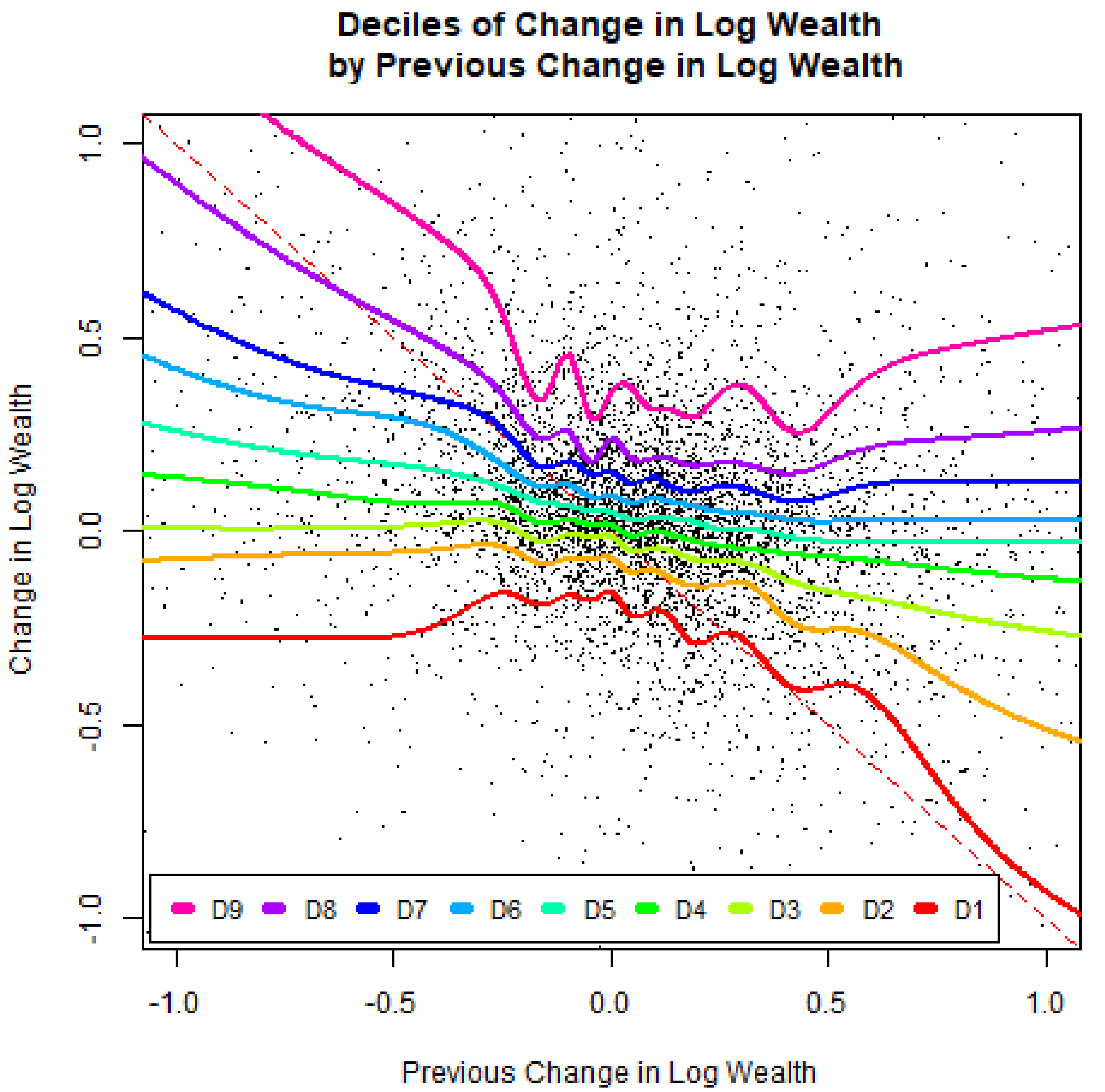

Figure 1.14: Non-Linear Quantile Regression showing Deciles of Differenced Log Wealth 2011-2009 vs Differenced Log Wealth 2009-2007.

1.4.2 Predicting Wealth Changes and Measurement Error

In wealth survey data, it is good practise to consider the possibility and strength of measurement error in the data. Before considering measurement error, I briefly examine whether the above wealth changes are easily predictable. I use the top 5% (n=1500) for a logistic regression model to predict whether a household continues to be a member of the top 5% in the next wave [33]. Although it does not very successfully predict the binary of whether a household leaves or not, it does separate the data into two groups - one predicted as very unlikely to leave ($\hat{P_{stay}} > 0.85$) versus a group forecast to be more likely to leave ($\hat{P_{stay}} \approx 0.55$). The highly likely stayers are wealthier, income richer, not self-employed and have proportionally less business/financial wealth[34] as well as less likely to

[33] as in Table 1.5

[34] Therefore, inversely, relatively more housing and physical wealth. However, portfolios are much less explanatory than income and wealth.

have extreme negative income changes. I also fit continuous regression models to changes in log wealth and find an out-of-sample R^2 of 0.15-0.24.[35] Income variables alone have an out of sample R^2 not above 0.05.

In the first wave of the survey, participants are asked to predict their "financial situation" in two years. Examining a basic breakdown of wealth changes conditional on participants' view of their future circumstances in Table 1.17, those predicting worse personal circumstances are actually slightly more likely to gain wealth over 07-09 and their predictions are insignificant when used in modelling.

Situational prediction (07 to 09)	Probability better off in 2009
not asked	0.46
better off	0.36
worse off	0.41
same	0.35
don't know	0.31

Table 1.17: Breakdown of proportion of top 5% respondents gaining wealth ($w_2 - w_1 > 0$) by those respondents' predictions of their future financial situation.

To confidently use survey data to identify wealth dynamics, measurement error is an important issue. Mechanically, simple noise in log wealth would reduce the appearance of persistence and can cause bias in a variety of estimates. Neri and Ranalli [2012] and Neri and Monteduro [2013] tackle measurement error in the SHIW using information from connected bank details and find significant underestimation of assets amongst the more wealthy. They do not identify time varying measurement error, but do find connections between education, wealth and income and the underestimation or overestimation of assets.

To account for measurement error in the dynamic moments that I have presented, I utilise the panel element of the data to identify this. Lee et al. [2017] studies the Korean Labour and Income Panel Study (KLIPS) and finds, using a set of instruments together with a panel estimation, an approximately equal split between measurement error and 'true' residual variation in individual household consumption and income and I use a similar approach.

As WAS is a dynamic (longitudinal) panel where fixed effects ensure any dynamic estimation would be inconsistent without the use of differencing and instruments, I use the strategies of Holtz-Eakin et al. [1988], Arellano and Bond [1991] and Anderson and Hsiao [1982]. These all use previous lagged values of the dynamic variable in question as instruments for estimation.

In this case, observed wealth $w_{i,t}$ is the dynamic variable of interest. Throughout, all variables are in logs, and I assume that there is classical i.i.d. measurement error (which is thus be multiplicative for actual wealth) with some variance σ_v^2. Hence, the estimating

[35] Please see appendix for more details.

equation is,

$$w_{i,t} = \rho w_{i,t-1} + \beta X_{i,t} + \alpha_i + \epsilon_{i,t}$$

Except,

$$w_{i,t} = w^*_{i,t} + v_{i,t}$$

Where $w^*_{i,t}$ is 'true' wealth. For estimation, the equation is differenced to remove α_i (fixed effects) and the methodology would normally use $w_{i,t-2}$ (and further back) as instruments to estimate ρ. But, in a world with measurement error, $w_{i,t-2}$ is no longer a valid instrument as it contains a link between differenced measurement error $\Delta v_{i,t-1}$ and $\Delta w_{i,t-1}$, the differenced right-hand-side regressor. Yet $w_{i,t-3}$ and beyond remain valid instruments.

If measurement error is restricted to be zero mean, i.i.d. and with homogeneous variance (as above) then the residuals from the differenced equation u_t (a function of $\hat{\rho}$) can be used to identify the variance of the measurement error and equation error,

$$\mathrm{E}(u_t u_t) = 2\sigma^2_\epsilon + 2(1 + \rho + \rho^2)\sigma^2_v$$

$$\mathrm{E}(u_t u_{t-1}) = -\sigma^2_\epsilon - (1 + 2\rho + \rho^2)\sigma^2_v$$

which can be solved for σ^2_ϵ and σ^2_v. Further lags on u can be used to create more restrictions (which can be used for ELSA, but WAS is too short with only 5 periods).

With the added assumption of normality, the distribution is fully defined and can then be used to generate simulated output.

I use a bootstrap to find the distribution of estimates, much like Lee et al. [2017], running dynamic panel regressions using difference GMM, as per Arellano-Bond.[36] Below are shown estimated results from WAS at both household- and individual-level for σ_v, σ_ϵ and ρ, as well as from ELSA.

Measurement error standard deviation is approximately half of true residual error standard deviation in both individual and household WAS, somewhat lower than Lee et al's results of an approximately equally sized σ_v and σ_ϵ for income and consumption in KLIPS. Persistence ρ is not extremely high, though it should be noted that this is after fixed effects and co-regressor effects. The persistence confidence interval is smaller for WAS when dealing with individuals. Whilst ELSA only includes adults aged over 50, it provides a useful benchmark and results are similar, though with somewhat lower ρ and a 1:1 ratio of σ_v:σ_ϵ.

[36]Other variables included are lags and polynomials of self-employment flag, business ownership, years in current job, degree holding, age, and income (including investment income), detailed in the appendix. Negative variance results are excluded throughout.

Data	Feature	Mean	$Q_{0.05}$	$Q_{0.5}$	$Q_{0.95}$	Std. Dev.
WAS Household	σ_v	0.104	0.041	0.106	0.158	0.035
	σ_ϵ	0.206	0.149	0.201	0.279	0.045
	ρ	0.453	0.011	0.446	0.930	0.293
WAS Individual	σ_v	0.111	0.038	0.113	0.176	0.041
	σ_ϵ	0.307	0.270	0.306	0.346	0.025
	ρ	0.533	0.275	0.518	0.857	0.179
ELSA	σ_v	0.186	0.077	0.198	0.253	0.054
	σ_ϵ	0.203	0.089	0.209	0.287	0.060
	ρ	0.307	0.147	0.295	0.503	0.112
	σ_v + OIDR	0.182	0.000	0.204	0.277	0.079
	σ_ϵ + OIDR	0.181	0.000	0.199	0.285	0.086

Table 1.18: Bootstrap Measurement error results. "OIDR" refers to use of overidentifying restrictions.

1.5 Conclusions

In this chapter, I investigated the WAS dataset and the evidence on the wealth at the top and dynamics of wealthy households. I showed that the WAS information on the wealthy includes greater wealth and inequality than estate data and that lower shares of wealth versus estate data are due to a higher aggregate wealth figure. I discussed business wealth data in the WAS and showed that later waves have little missing data and are safe to use, whilst multiple imputation can be used for earlier waves. I found that the wealthiest 1% of households in the UK hold 20% of total wealth, and inclusion of business wealth is important to that figure. The WAS sampling and representation of the wealthy improves over waves, yet defined benefit pension wealth is difficult to use and generates spurious patterns in inequality. Comparing the WAS dataset to both estate data and the rich list of the top 1000 in the UK, the WAS data does not under-represent the wealth of the wealthy and shows similar inequality features at the top to other sources.

I then discuss the new moments and facts this data contributes to the literature - transitions in wealth at the top. I find rich wealth dynamics amongst the wealthy, including high probabilities of exiting the richest wealth categories, with one third exiting every two years and half every six years. I compare these transitions to available data from other surveys - including the SCF and PSID from the U.S. and ELSA in the U.K., finding similar patterns and results. Wealth transitions have significantly negative skew and high kurtosis. Kurtosis is particularly high at the very top, and variance decreases with wealth until the top 5% whereafter wealth and variance increase together for those above the 95th quantile. Quantile regressions indicate that the very wealthy suffer great variability in wealth. Those at the very top, above the 95th quantile, experience a wider distribution of changes in log wealth than those below, and the distribution widens as one moves further up the tail of the wealth distribution. I investigate those at the top who experience particularly large falls in wealth and find that they are more likely to be self-employed or a business owner and hold proportionally more business and financial wealth. Importantly, business wealth of those who experience large falls in wealth is the element of their wealth which suffers the especially large falls.

I note the existence of a pattern whereby those in a top wealth group, such as the top 1%, have a higher probability of staying in that group versus newer entrants to the group and investigate the possibility of measurement error. Using an AR1 dynamic panel estimation strategy, I identify the variance of time-varying i.i.d. measurement error and 'true' residual error, finding that measurement error variance is approximately half the size of 'true' error variance in the WAS (a 1:2 ratio). Noting the same pattern occurs in ELSA, I also examine that dataset with the same methodology and find a 1:1 ratio.

I therefore conclude that the WAS dataset offers new and interesting insights, especially at the top of the wealth distribution. The results I have presented do not encompass the full depth of rich information in the dataset, which includes risk and saving attitudes, breakdown of debts and inheritance information, amongst a host of other variables. There is great scope for future work and insights for this data.

Chapter 2

Wealth and Mobility: Superstars, Returns Heterogeneity and Discount Factors

The wealthy hold a large fraction of total wealth but to what extent do they stay wealthy over time? What theory explains both cross-sectional inequality and the dynamics of wealthy households? This chapter uses the longitudinal UK Wealth and Assets Survey (WAS) to answer these questions. I examine three main theories for the highly concentrated distribution of wealth against the data - heterogeneous returns to wealth, temporary high earnings and discount factor heterogeneity. I identify heterogeneous returns to wealth as the theory that best explains the inequality and mobility data and I corroborate my findings with a model which combines all three mechanisms. This result occurs because poor heterogeneous wealth returns realisations simultaneously reduce stocks of wealth and discourage future saving through expected persistence in wealth returns. This generates very large downwards mobility. My estimated model matches both wealth inequality and mobility moments and can show that, structurally, 12% of the top 1% leave this category within two years and 25% leave within six years.

2.1 Introduction

Inequality, the behaviour of the wealthy and the distribution of wealth have long been topics of discussion for economists. Recently, inequality has become more prominent in policy and academic questions and the implications of heterogeneous wealth distributions to economic and policy questions is still being widely explored.[1] The very wealthiest hold a large fraction of wealth in most developed economies, so much so that the rich right tail of the empirical cross-sectional wealth distribution often follows a fat-tailed Pareto

[1]For example, the recent announcement of a wide-ranging Institute of Fiscal Studies review on inequality headed by Angus Deaton.

distribution. In this chapter, I focus on the mobility of the wealthy in that tail. I use data on both inequality and mobility to evaluate quantitative theories of inequality. The incomplete markets Aiyagari-Hugget-Bewley framework, often used by macroeconomists to generate a non-trivial distribution of wealth through self-insurance buffer stock savings against earnings shocks, cannot create the thick right tail and concentration found in the data. Hence, three main theories of tail wealth accumulation have been proposed - heterogeneous returns to wealth; temporary 'superstar' high earnings state(s) and discount factor heterogeneity. Using the data, I estimate a structural model to identify which mechanisms are driving inequality and mobility, and the parameters governing those mechanisms.

Understanding the drivers of wealth inequality is key to the implications of many heterogeneous agent macroeconomic models. For example, Kindermann and Krueger [2014] find optimal tax on top earners to be over 90% with an exogenous 'superstar' earnings process whilst the entrepreneurial model used by Cagetti and Nardi [2004] shows that reducing estate tax and raising income tax is welfare decreasing. Ocampo et al. [2017] find efficiency through improved capital allocation under wealth taxation and Carroll et al. [2017a] argue that wealth differences resulting from preference heterogeneity is important to household consumption responses.

Motivated by the need to distinguish the driving force behind wealth inequality, I utilise the UK Wealth and Assets Survey (WAS) panel dataset. As described earlier, this wealth survey is significantly larger than its peers, is longitudinal and it oversamples the wealthy to capture them accurately. This allows us to study wealth transitions amongst those at the top and to use that data to evaluate different explanations for top wealth inequality. I therefore apply moments from the data to a Bewley-Huggett-Aiyagari incomplete markets framework[2] with the three additional explanations to generate realistic inequality.

De Nardi [2015] and De Nardi and Fella [2017] examine major hypotheses about extensive wealth accumulation: earnings and income risks; idiosyncratic returns and wealth risk; heterogeneous saving/risk preferences; bequests, human capital and altruism towards descendants; medical expenses and, lastly, entrepreneurship. I choose to focus on the first three in this chapter, though the model also incorporates some stylistic features of bequests and inheritance.

Very high 'superstar' earnings states (Castaneda, Diaz-Gimenez and Rios-Rull [2003]) that last a limited period of time have been found to generate very high wealth inequality. Superstardom is temporary such that households save most of their earnings due to knowledge that they will eventually lose superstar status and will want to use these savings to smooth their consumption over time. Due to the extreme level of the earnings state, these wealth stocks can be very large, generating the high inequality found in the data.

Benhabib, Bisin and Zhu [2014] and Benhabib, Bisin and Luo [2015]) offer an alter-

[2] The key papers for this literature being Aiyagari [1994], Huggett [1996] and Bewley [1983]

native explanation in the form of exogenous heterogeneous returns to wealth. They show that a distribution of returns can replicate cross-sectional wealth inequality and has simple implications for mobility. In this theory, wealthy agents are those who experience a series of excessive returns - as they become richer the impact of greater returns increases, leading to a process that generates a fat tail of a few wealthy agents who control very large asset holdings. Non-perfect persistence of the returns process (including birth and death) ensures that wealth does not excessively concentrate, leading to a Pareto distribution.

Discount factor heterogeneity, as used by Krusell and Smith [1998], Hendricks [2004] and Carroll et al [2017a] explains wealth heterogeneity by different weightings on future consumption, often labelled as 'patience' or a desire to smooth consumption. Explanations from this theory are rarely targeted at the very wealthy tail, as Hendricks notes, and relies on more patient households accumulating greater asset holdings due to greater desire to save for the future and to keep their consumption stream smooth.

Understanding the dynamics of the wealthy and how they come to be wealthy is important in and of itself - is there a dominant perpetual 'rentier' class who live from their income? Or are the wealthy better characterised as the lucky tail of portfolio risk? Are they recipients of sudden rewards for extraordinary skills or gradual wealth builders? Whilst we can identify the cross-sectional features of the wealthy - more likely to be entrepreneurs, hold more stocks, be slightly older - we need longitudinal data to understand their dynamics, and to discipline mechanisms that claim to represent and drive the distribution of wealth.

This chapter use the features of those at the top of the wealth distribution, including the changes in wealth faced by top households, which I extensively analysed in the earlier chapter concerning the WAS data. To summarise, there are substantial wealth and income mobility at the top in the raw data, where around a third of the wealthiest 1% exit this group biennially and are unlikely to return. After six years, half of the wealthiest 1% have exited that category.

Using the WAS data moments, my main finding is that returns heterogeneity is the mechanism that best explains the data. This is because it has the ability to generate larger and faster downward mobility than other mechanisms. It can do so because it has two effects, one directly affecting the agent's budget constraint and one behavioural effect through expected future returns. Agents with particularly poor realisations of returns will experience falls in their wealth stock. This can force rapid changes in wealth, depending on the persistence of returns and degree of variance. Poor returns also feeds through into an incentive not to hold wealth if one expects poor returns to continue in future, causing further de-accumulation. In contrast, superstars de-accumulate slowly after losing their very high earnings as there is no downward pressure on their wealth except gradual consumption-smoothing pressures. Discount factor shocks only operate through the behavioural channel of expected value of future wealth, not affecting the agent's budget constraint or resources.

The estimated returns heterogeneity has a positive yearly autocorrelation of approximately 0.5 and standard deviation of 0.1 in the joint model with all three theories of inequality present. This volatility is in the region of direct wealth return heterogeneity estimates by Fagereng, Guiso, Malacrino and Pistaferri [2016] using Norwegian administrative wealth tax datasets. As a benchmark, the unconditional yearly wealth returns standard deviation is 0.16 versus Campbell's 0.5-0.6 for a single U.S. public stock (Campbell [2001]). The other two mechanisms do not substantially contribute to explaining inequality in the joint estimation.

I correct for time-varying measurement error, as this can play a quantitatively important role in wealth survey data[3]. I still find substantial mobility after the correction, with around 12% leaving the top 1% every two years and 25% every six years. Without this correction attributing some variation to measurement error, returns heterogeneity would be even more prominent as the most successful mechanism since it is the only one that can accommodate rapid and large wealth changes and thus greater variation in wealth favours it.

2.2 Data: The Wealth and Assets Survey

In this section I describe the WAS data used and the wealthy within it, building a picture of their relevant characteristics. As stated earlier, the WAS is a biennial panel survey dataset covering wealth, income and demographics for UK households and thus supplies useful moments with which to estimate theories of wealth inequality. In my chapter on the WAS data, I examine the WAS in detail, comparing its cross-sectional implications versus estate data, rich lists and other survey and administrative datasets. I find it effectively represents the top of the distribution and here, I provide a short summary of relevant cross-sectional findings from the WAS concerning the wealthy and a brief recall of key mobility features used in my modelling and estimation.

Throughout this chapter, the benchmark definition of 'wealth' is that used earlier - the sum of private business values; financial assets (cash, shares, bonds, investment funds, savings products, deposits minus debts and credit cards); property (value minus mortgage debt) and physical wealth (vehicles, jewellery, collectibles, household contents), minus any other liabilities.

Table 2.1 shows statistics for the whole population and from wealthy groups[4]. The

[3] An example could be Biancotti, D'Alessio and Neri's [2008] study of the Italian Survey of Household Income and Wealth

[4] Income is before taxes and without social benefits, other income categories are investments, rental properties, pensions and other (including irregular items). Earnings includes self-employed or business earnings paid as wages. Age[5], self-employed and business ownership (amongst the self-employed) refer to the Household Reference person, whilst all other rows are for the entire household. The 'wealthy' groups

Group	All	top 10%	top 5%	top1%
Age	54	62	61	60
Income	38732	91138	120262	228355
Earnings	31883	60866	79532	146140
Self-employed	0.09	0.21	0.27	0.4
Business owner	0.05	0.15	0.2	0.34
Wealth (total)	317572	1596584	2396994	6355747
Property / Total	0.55	0.48	0.44	0.31
Financial (net) / Total	0.18	0.22	0.22	0.19
Physical / Total	0.15	0.07	0.06	0.03
Business / Total	0.12	0.22	0.28	0.46

Table 2.1: Means for top groups and population

heads of households ('household reference person') in top wealth groups are a little older than those of the general population[6]. Unsurprisingly, the wealthy have much higher gross incomes than the population, and a lower proportion of income from earnings (and thus proportionately higher income from investments and assets). They are much more likely to be headed by an entrepreneur or business owner and whilst they still concentrate a large proportion of their wealth in housing, the prominence of business wealth and financial wealth is much greater amongst the very wealthy.

The 'average' wealthy household is quite varied - some households are dominated by business wealth, others by property. There is great variation in their incomes versus their wealth and the sources of their incomes. As expressed in the previous chapter, there is significant mobility between different wealth groups, with large proportions exiting categories of the wealthiest. It is also important to recall two other features - the large variation in wealth changes, even (especially) amongst those at the top, and the strong persistence in continued membership of top wealth categories, despite the relatively high group exit rates from wave to wave, or 'stayers stay' pattern. In Table 2.2 (which repeats earlier tables for convenience) the probability of staying is highly conditional on history of membership. Those with longer past membership appear to have a much higher probability of remaining in the group, whereas new entrants have a very high chance of exit - 'stayers stay'[7].

are defined by the wealth variable, which is as described in the text. The proportions are dividing one average by another.

[6]The age of Head of Households is structurally higher than that of the population of individuals. The WAS distribution of individual ages matches other demographic data perfectly.

[7]As mentioned, ELSA data contains similar findings and the high wealth mobility is not dissimilar from the SCF or PSID.

	top 10%	top 5%	top 1%
$P(T_4\|F_1F_2T_3)$	0.48	0.39	0.30
$P(T_4\|F_1T_2T_3)$	0.75	0.68	0.66
$P(T_4\|T_1T_2T_3)$	0.91	0.88	0.87

Table 2.2: Probability of remaining in top wealth groups given different histories. 'T_t' indicates 'True' for belonging to the group in wave t and 'F_t' indicates 'False' for the same.

2.3 Model

I now consider incomplete markets explanations for the highly skewed wealth distribution versus the dynamic facts in the WAS data.

The basic structure for the following is an Aiyagari model containing a distribution of agents deciding to save or consume a simple, liquid asset and facing labour earnings shocks. It is well known that this model cannot replicate the substantial cross-sectional wealth inequality in the data, hence I add the different inequality generating mechanisms discussed in the Introduction.

Households have CRRA utility,

$$u(c) = \frac{c^{1-\gamma}}{1-\gamma}$$

In the model, a household can be young or old, with probabilistic ageing and probabilistic death for the old (who are then reborn as young, subject to estate taxes). The probabilities are selected to replicate actuarial population statistics. I denote the age status as O and its transitions as Π_O.

They also have (discretised) earnings ability z, which follows a transition matrix Π_z and returns ability R which follows transitions Π_z. Similarly, discount factors β are stochastic and follow transitions Π_β. The age, discount factor, earnings and returns transition matrices are exogenous. They choose to save or consume c in an asset a, creating a state vector of $\{(a_t, z_t, R_t, O_t, \beta_t\}$ describing an agent in a given period. The agent aims to maximise their sum of expected discounted utility, forming the following Bellman equation,

$$V(a_t, z_t, R_t, O_t, \beta_t) =$$

$$\max_{c_t, a_{t+1}} \{u(c_t) + \beta_t \mathbb{E}_t(V(a_{t+1}, z_{t+1}, R_{t+1}, O_{t+1}, \beta_{t+1}))\}$$

The budget constraint for a young agent ($O_t = young$) is

$$c_t + a_{t+1} = wz_t + R_t(1+r)a_t$$

They choose to save or consume out of their earnings income wz_t, where w is the equilibrium wage, and wealth a_t subject to interest and capital gains earnings $R_t r a_t$, r

being the equilibrium interest rate on the asset.

For the old ($O_t = old$) the budget constraint is

$$c_t + a_{t+1} = p_t + R_t(1+r)a_t$$

This is the same as young agents, except for receipt of a fixed pension p rather than earnings, which the government pays for using income and consumption taxes.[8] I do not show the taxes in this exposition for clarity and brevity.[9]

For agents who die and are replaced by a young descendent ($O_t = born$), the equation is equivalent to the young but their assets are subject to estate tax τ_{estate},

$$c_t + a_{t+1} = wz_t + R_t(1+r)a_t(1 - \tau_{estate}(a_t))$$

Later in this paper, I characterise and use the processes R, z and β in order to match wealth inequality. I allow for the possibility of stochastic inheritance of discount factors after death transitions, but impose full inheritance of R and a redraw of earnings from the stationary distribution.

To close the model, I have a production sector with a representative firm who produces a consumable output good using capital and labour. The firm is Cobb-Douglas with capital share $\alpha = 0.33$ and pays depreciation of $\delta = 0.07$ on capital. The firm pays r to rent capital and w to pay workers. I find an equilibrium r which matches capital demand and holdings among agents.

As the agents are receiving different returns for assets, I create a zero-cost, risk-neutral and perfectly competitive representative financial intermediary who holds household assets on their behalf. The intermediary rents the capital to firms, receives the rental income and return of the capital and then pays a stochastic return to each household on their units of capital, which is such that on average the intermediary makes zero profit. I assume this return is $(1+r)R$ where R stochastic and can be viewed as random efficiency of the intermediary for each individual household. Households have to hold this asset or consume. Effectively, this intermediary amalgamates the capital stock for the firm and then distributes the total returns so that households receive different returns. In reality, we may prefer to think of this as household 'ability' rather than financial intermediary efficiency/success.

This is a stationary rational expectations equilibrium, with prices and policies:

- HH policy function $a_{t+1}(a_t, z_t, R_t, O_t, \beta_t)$ from solving value function problem above

[8]the tax revenue always exceeds these payments. I assume the remainder is spent on non-utility-enhancing projects rather than rebated to households for a balanced budget.

[9]The estate tax is calibrated in the style of Cagetti and Nardi [2006] and Cagetti and Nardi [2004] by matching proportion of deceased paying (3.5%) and generating a flat effective tax rate by matching revenue (0.18% of GDP) due to widespread avoidance and tax relief versus headline rates. I use a Gouveia and Strauss [1994] income tax function estimated for UK taxes, and a UK consumption tax of 17.5%. Simplified state pension payments follow the ratio of state pensions to earnings in the WAS data.

given w and r

- the competitive intermediary makes zero profit
- Firm maximises profit $K^\alpha N^{1-\alpha} - wN - (r+\delta)K$ with factor prices $r = MPK - \delta$ and $w = MPN$
- markets clear when firm capital demand equals household supply, weighted by their returns $K = \int R_i a_i di$
- labour market clears, $\int n_i di = \int z_i \mathrm{I}(O_i = 0) di = N$

To operationalise this model, throughout I use a log AR1 distribution of earnings y for agents[10], calibrated to the UK earnings Gini and the Shorrocks Index for Quintiles. I use WAS figures, as administrative earnings data reported by De Nardi, Fella and Paz Pardo [2018] has very similar results. Other parameters take well-known values - unless otherwise specified, there is a discount factor of $\beta = 0.95$ and CRRA preferences with parameter $\gamma = 2$ for all agents.

The next step is to add the three wealth inequality generating mechanisms - superstar earnings, returns heterogeneity and discount factor heterogeneity.

Superstar earnings are in the form of an extra z earnings state with a level Y, which can be entered into equally from any earnings state ($P_{Y,in}$) and exits equally into any earnings state ($P_{Y,out}$). This is a modified version of the Castaneda et al-style super-high ability level $\bar{y}$ used to generate wealth inequality ("CDR model").

Individual returns R are characterised as a discretised log-normal AR1 process, with parameters of autocorrelation ρ_r and standard deviation σ_r and a mean of 1. I use this process to nest the ideas in Benhabib, Bisin & Luo [2015] ("BBL model") and Benhabib, Bisin and Zhu [2014] ("BBZ model") that heterogeneous returns with different persistences (BBL is lifelong R whilst BBZ has zero autocorrelation) can generate tail wealth inequality in line with the data - one of my aims is to shed light on the appropriate persistence.

Discount factors β follow the literature[11] in assuming a discrete state symmetric process. I use two states β_l, β_h and probability of transition P_β. I assume that earnings ability is not inherited and is redrawn from the stationary distribution after death, whilst returns status is fully inherited[12] and I allow stochastic inheritance of β, so there is a parameter $P_{\beta,d}$ which governs the probabilistic inheritance of β.

2.3.1 Estimation

With the model complete, I now turn to the estimation procedure for recovering parameters, understanding the mechanisms and comparing to the data. After calculating an

[10] I am mostly concerned with the upper tail which, as De Nardi, Fella and Paz Pardo [2016] note, even realistic non-parametric earnings processes do not match, so I keep earnings simple.

[11] Examples include Krusell and Smith [1998] , Hendricks [2004] and Carroll et al. [2017a].

[12] As the portfolio, its managers and so on would be inherited, etc.

equilibrium I simulate 100,000 agents. In summary, I calculate the same moments, transition matrices and quantile regressions from the model as the WAS data shown earlier and compare the two using a Simulated Methods of Moments structure. To estimate and generate the distribution of parameters in the structural model I construct an objective function based on the model's moments equally weighted, normalised deviations from a set of equivalent data moments. I then use a methodology based upon Chernozhukov and Hong [2003]. I use a Monte Carlo Markov Chain approach, starting from a point in the parameter space, iteratively applying an innovation to a set of parameters, simulating the model for a given set of parameters, considering the new value of the objective function, deciding to accept or reject the new position based on a probabilistic rule and then drawing a new innovation to create new parameters to add to that position and so on. I use the Metropolis-Hastings algorithm, and calibrate such that the acceptance rate is approximately 20%. I discard initial points as a burn-in before calculating the distribution of parameters.

Throughout, I include time-varying i.i.d. measurement error standard deviation as a parameter in the estimation. I view this inclusion as best practise in using survey data and a straightforward correction for which I consider robustness checks. The measurement error is in logs and identification of measurement error versus returns heterogeneity centres on the use of conditional transition probabilities as targets. The log specification and i.i.d. draws create negative autocorrelation which influences the 'stayers stay' pattern and so identifies the variance parameter, together with pressure from the moment of total variance of observed changes in log wealth.

The data moments, or targets, are:

- top 1, 5 and 10% wealth shares
- 2, 4, 6 and 8 year top wealth staying rates for top 10%, 5% and 1%
- 2-stage (e.g. $T|FT$ in my notation) and 3-stage (e.g. $T|FFT$) conditional staying rates for top 1 and 5%
- standard deviation of changes in log wealth above median wealth (0.32)
- UK Capital-Income ratio (2.5)

There are 23 targets in total, shown in Table 2.3 and I also provide the full list of targets and their data values when discussing and comparing versus results in Table 2.16. The total parameter count from the above is 10, leaving 13 degrees of freedom for the joint estimation and thus being overidentified.

The data targets are estimated from the WAS, in the manner described earlier in this work and using the same notation. Thus $P(T|FT)$ refers to the probability that someone will be a member of a category, given that they have been a member of the category (T)

Moment Definition	Targeted Value
Share of wealth held by Top 1%	0.206
Share of wealth held by Top 5%	0.385
Share of wealth held by Top 10%	0.478
Probability of staying in top 1% (2yr)	0.73
Probability of staying in top 5% (2yr)	0.67
Top 1% $P(T\|FT)$	0.37
Top 1% $P(T\|TT)$	0.81
Top 5% $P(T\|FT)$	0.44
Top 5% $P(T\|TT)$	0.87
Top 1% $P(T\|FFT)$	0.3
Top 5% $P(T\|FFT)$	0.39
Top 1% $P(T\|TTT)$	0.87
Top 5% $P(T\|TTT)$	0.88
Probability of staying in top 1% (4yr)	0.59
Probability of staying in top 1% (6yr)	0.55
Probability of staying in top 1% (8yr)	0.51
Probability of staying in top 5% (4yr)	0.68
Probability of staying in top 5% (6yr)	0.63
Probability of staying in top 5% (8yr)	0.61
Probability of staying in top 10% (4yr)	0.71
Probability of staying in top 10% (6yr)	0.68
Probability of staying in top 10% (8yr)	0.63
$\sigma_{\Delta log(wealth)}$ (above median only)	0.34
Capital:Income Ratio	2.5

Table 2.3: Estimation Moments.

only for one period, before which they were not in the category (F). I use both two-stage and three-stage conditional probabilities in this estimation, though I exclude $P(T|FTT)$ given that the other two- and three-stage moments together with the overall probability of staying make this predictable and thus a possible source of collinearity. I only include those above median wealth in the standard deviation moment, as those at the bottom are dominated by the (simple AR1) earnings process. The lower end of the wealth distribution is not my focus and this model does not aim to explain it with great accuracy, so I use those above the median. The capital income ratio for the UK is somewhat lower than the US at 2.5, although it varies over the relevant period (a decade or so) between 2.4 and 2.6, so I take the average.

In the most general model I use, which incorporates all three mechanisms, the 10 parameters from the model are displayed in Table 2.4.

Definition	Parameter
R autocorrelation	ρ_r
R standard deviation	σ_r
superstar level	Y
superstar entry probability	$P_{Y,in}$
superstar exit probability	$P_{Y,in}$
probability of staying in β state	P_β
β inheritance probability	$P_{\beta,d}$
first β state	β_l
second β state	β_h
measurement error standard deviation	σ_v

Table 2.4: Parameters for estimation.

2.4 Results

The main result is that heterogeneous returns to wealth fits the data best amongst the three mechanisms. I consider estimations using each theory on its own and then a joint estimation with mechanisms from all three theories in Table 2.5. The sum of squared errors (SSE) from the data moments finds R shocks superior to the other two mechanisms on this fit index and quite close to the errors of the unconstrained estimation involving all three explanations. This method of comparison mirrors the equally-weighted GMM objective function used for all the estimations.[13] The parameters for the R process are very similar between the estimation using R alone and the multiple explanation version - an autocorrelation of 0.5 and standard deviation of 0.1. The multiple explanation estimation has superstars with very low earnings versus the canonical extraordinary levels used (only 4x median earnings) and limited β heterogeneity, suggesting returns heterogeneity remains the driver behind inequality even when other mechanisms are allowed.

Measure	Model	Min.	Median	Mean
Sum Squared Error	All mechanisms	0.09	0.14	0.16
	R only	0.12	0.26	0.38
	Superstars	0.42	0.94	0.89
	β only	0.7	0.9	0.89

Table 2.5: Fit of estimations.

In table 2.5, the minimum sum of squared errors represents the best fit of the particular model, which is particularly close between R only and all mechanisms. The mean and median are draw from the distribution of results for each model. In terms of minimum, mean or median SSE, Superstars are a much poorer fit than R shocks or the unconstrained estimation, and are very similar to β only. As would be expected, the unconstrained mechanism does improve over R shocks alone, but by significantly less.

The reason for the identification of returns heterogeneity as the best theory to explain the data comes from the tension between inequality and mobility across the different the-

[13]'moment condition' is used interchangeably with 'target' throughout this chapter.

ories. With the exception of wealth returns variance, the mechanisms to create inequality rely on incentivising persistent above average saving over time in a subset of the population and thus generating a wealthy group. But this (almost necessarily) generates stasis in wealth. The mobility moments force the model to generate wealthy households who lose wealth rapidly enough to exit wealthy groups at the correct rates and in the right time-frame, providing tension against allowing this stasis. Whilst time-varying measurement error can increase mobility, it is particularly restricted on the upside by the need to match the standard deviation of wealth.

Wealth returns heterogeneity can cause the rapid changes in wealth found in the data due to both directly affecting the stock of wealth and changing incentives to save in the future. It can do so whilst also creating inequality at realistic levels. This is particularly important for matching downward changes in wealth, as Superstars lose high income but only consume their wealth stock gradually to smooth consumption, whilst β shocks focus on savings incentives alone and are very persistent to generate inequality. Whilst both of the two other mechanisms can attempt to match mobility, they do so either with the aid of excessive measurement error volatility, which causes the model to overshoot wealth variability ($\sigma_{\Delta log(w)}$) or, as realistic inequality would cause a failure to match mobility, they choose parameters which generate too little inequality. For example, the top 1% wealth share for the model with superstars only is 13%, versus 20% in the data. These large deviations are then punished in the fit index.

As a demonstration of the mechanism by which returns heterogeneity generates rapid downward changes in Figure 2.1 I examine a a wealthy household at the 99.5 percentile suffering a series of the worst shocks under each theory. Low heterogeneous returns realisations are in black, a loss of superstar ability in red and a lower discount factor in orange. I show the points when the agent reaches key quantiles such at the 99th (top 1%) in text alongside each curve. The very unlucky agent in black continually experiences the very lowest state of heterogeneous returns R in the discretised AR1 process (-26%) and has constant median earnings wz. He rapidly falls to below the median wealth in less than a decade.[14] In blue, I show the same agent path, but compensated for the direct losses of wealth and changes in his budget constraint. This disentangles the mechanisms of direct changes to wealth from R and changes to savings incentives - discovering the incentive effect by compensating the agent for the direct loss of wealth but having the same R state and expectations. The blue agent deaccumulates much more slowly, showing a large proportion of mobility from R shocks comes from the direct changes, nonetheless, there is still a significant fall from R expectations alone.

The red superstar agent deaccumulates slowly and from a significantly higher wealth

[14] Note that this unlucky agent is indeed unlucky given the medium persistence of the R process ($\rho_r = 0.47$) and is illustrative. Yet annual falls of -26% are not uncommon in the wealth data earlier or in asset markets.

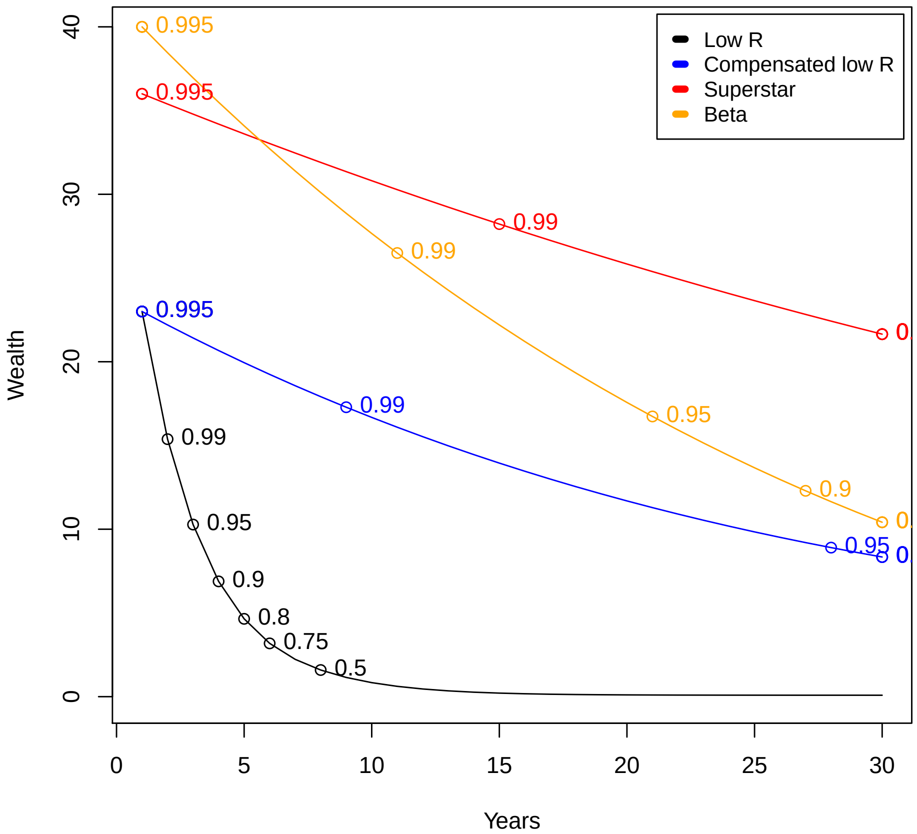

Figure 2.1: Simulation of agents wealth over time, starting at the 99.5th percentile and experiencing very bad shocks in different models. Point at which agent passes key percentile of wealth shown with text of that percentile next to curves.

position, visually depicting the greater wealth immobility resulting from superstars. This can be seen in the percentiles the agents pass through - the unlucky R agent is below the 95th percentile in 3 years, the compensated low R agent reaches the 95th percentile in 30 years and the superstar agent remains above this. One can thereby see the need for higher measurement error amongst superstars to create mobility and the constraint on the superstar mechanism from high mobility leading to inability to match inequality. The estimated R model depicted in the figure has realistic inequality, yet the estimated superstar process that would replicate inequality would have even higher wealth.

An agent from the β model is shown in orange. This agent is at the 99.5th percentile and is given the lower β, in this case 0.935 versus a high β of 0.975. The β model has a very high persistence (with an estimated average state duration of 2000 years), so having an agent with such high wealth without a high β is exceedingly rare and not typical of transitions in the β model, which are mostly attributable to measurement error in the estimation. The agent deaccumulates quite quickly in absolute terms, but still remains above the 90th percentile after 25 years. The high persistence of the lower state means the agent expects to have a low value for future savings for a very long time and so the impact of the lower discount factor is magnified by the long future expectation, resulting in fast wealth stock consumption. The compensated R agent has a gentler slope than the β agent due to the lower expected persistence of their R state and thus a smaller impact on their future expected returns and value of savings.

I now turn to the results of estimating each explanation in turn, before covering the joint estimation of all three mechanisms and robustness checks.

2.4.1 Superstar earnings

Giving a small number of households incredibly high earnings with a significant chance of losing those earnings generates substantial inequality. These lucky agents are aware of their eventual superstar-less future and save a substantial proportion of their income to insure against this, as agents save most of a temporary income shock. When they do lose their superstar ability, they then dis-save gradually, smoothing their consumption over time according to their discounting preferences.

Although typically the population of superstars used is very small and with extraordinary income (for example, 0.01% in Kindermann and Krueger [2014] earn over 1000 times median earnings) I allow the entry and exit probabilities for superstars to be estimated such that different populations with different longevity are possible, as described above. The superstar earnings process has three parameters: a level Y, a superstar entry probability $P_{Y,in}$ and exit probability $P_{Y,out}$. In addition, there is the standard deviation of time-varying measurement error σ_v.

Parameter	Mean	$Q_{0.05}$	$Q_{0.5}$	$Q_{0.95}$	s.e.
Y	15.409	8.71	14.14	25.43	5.411
$P_{y,in}$	0.002	0.00112	0.00195	0.00406	0.001
$P_{y,out}$	0.324	0.13	0.338	0.478	0.096
σ_v	0.3	0.238	0.296	0.373	0.038

Table 2.6: Estimated parameters

I find the superstar estimates to be have much lower earnings than is usual for such models, only 10-20 times median, and around 0.6% of the population are superstars. The model then struggles to match tail inequality with these weak superstars. The estimation procedure prefers to minimise the earnings of superstars in order to attempt to match mobility. In table 2.7 the match to conditional mobility moments and staying rates is good, but the wealth share of the top 1% is significantly too low, as is their staying rate. This is likely due to the large estimated measurement error volatility. At 0.3, this is almost as large as total data volatility of wealth ($\sigma_{\Delta log(w)}$ targeted moment, 0.34) and causes the model's $\sigma_{\Delta log(w)}$ to significantly overshoot the target.

Separately calibrating the model to match cross-sectional inequality moments alone,

Moment	Data	Mean	$Q_{0.05}$	$Q_{0.5}$	$Q_{0.95}$	s.d.
Top 1% share	0.206	0.13	0.09	0.13	0.17	0.02
Top 5% share	0.385	0.34	0.28	0.34	0.42	0.031
Top 10% share	0.478	0.48	0.41	0.48	0.56	0.031
Top 5% stay	0.73	0.74	0.69	0.74	0.78	0.013
Top 1% stay	0.67	0.61	0.55	0.61	0.68	0.027
Top 1% $P(T\|FT)$	0.37	0.39	0.35	0.39	0.44	0.015
Top 1% $P(T\|TT)$	0.81	0.75	0.7	0.75	0.8	0.02
Top 5% $P(T\|FT)$	0.44	0.44	0.37	0.44	0.5	0.026
Top 5% $P(T\|TT)$	0.87	0.84	0.82	0.85	0.88	0.011
$\sigma_{\Delta log(w)}$	0.34	0.44	0.36	0.43	0.53	0.049

Table 2.7: Selected moments from data and estimation.

I find that to match the top 1% wealth share the model needs earnings of around 50 times the median - and this is very different to the estimation including mobility targets, or to top earners in the administrative earnings and survey data, who are significantly lower.[15] This aligns with criticisms from Benhabib et al. [2015] that superstar models have to use earnings far above that found in surveys or administrative data when matching inequality[16]. These findings show that the high earnings and resultant inequality disappear when confronted with mobility.

If the model is forced to focus solely on cross-sectional inequality, as mentioned above, wealth shares can be matched, but only by greater immobility - for example, a biennial staying rate of 80% for the top 1%. This is because the earnings level needed to match wealth inequality is so high that agents take a very long time to fall to another category. In the case of imposing realistic inequality, measurement error would have to be even greater to match mobility and would further overshoot $\sigma_{\Delta log(w)}$. In the estimation, using measurement error to match mobility is constrained by targeting variance of log changes in wealth and the 'stayers stay' pattern, leaving the superstars mechanism to choose between matching mobility or inequality.

[15]'Real' superstars' probabilities of entry and exit for the top earnings 0.1% from the WAS are yearly equivalents of 0.0002-0.0005 and 0.3-0.4, with similar figures for the top 0.5% and top 1%. They earn an average of 30 times median household earnings. Results from De Nardi et al. [2018] using the UK administrative earnings survey dataset are very similar.

[16]Though the debate on effective capturing of high earners in tax data is still open.

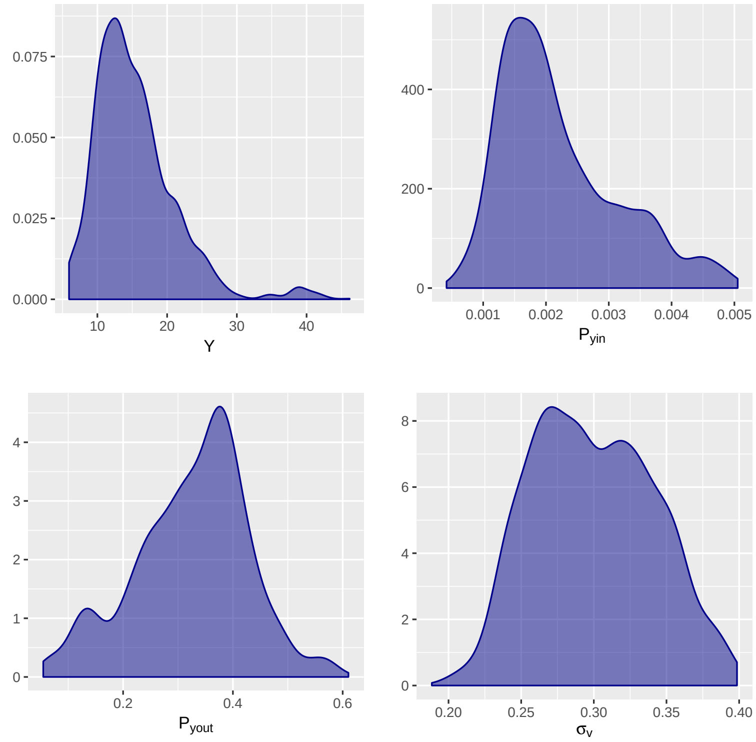

Figure 2.2: Estimated parameter density distribution.

2.4.2 Discount Factor Heterogeneity

It is difficult to use symmetric preference heterogeneity to generate inequality that matches the right tail of the wealth distribution, as noted by Hendricks [2004]. I estimate the persistence of discount factors both within lives (P_β for staying in a β state) and through inheritance ($P_{\beta,d}$ to keep β state). The two discount factors β_l and β_h are parameters estimated within the unit interval.

The estimation results reflect the difficulty of replicating inequality at the very top with discount factors alone, ending with point-densities at corner solutions where $P_\beta \to 1$. As $P_{\beta,d}$ is also very close to 1, the agents have very long preferences - they keep their β almost certainly for their entire life and only have a one in 40 chance their children will not have the same discount factor state. Given the expected working life and estimated probabilities, the average household will stay in the same state for over 2000 years. Despite the immense longevity and opportuntiy for large differentiation between discount factors, this only results in a top 1% wealth share of less than 15% and top 5% share of 30%. Because there are only two symmetric states, too much longevity or differentiation could decrease tail inequality as the different populations are too big to cause the concentrated accumulation by a very small group that occurs in the empirical Pareto distribution.

Nonetheless, due to the allowance for measurement error, the longevity of the preference dynasties does not result in surface level secular stasis. However, the staying rates are not well matched, as can be seen in Table 2.9. The pattern of the conditional staying rates

Parameter	Mean	$Q_{0.05}$	$Q_{0.5}$	$Q_{0.95}$	s.e.
β_1	0.936	0.932	0.937	0.938	0.002
β_2	0.976	0.963	0.979	0.984	0.007
P_β	0.999	0.9993	0.9998	0.9999	0.002
$P_{\beta,d}$	0.949	0.931	0.952	0.955	0.007
σ_v	0.218	0.2	0.221	0.234	0.011

Table 2.8: Estimated parameters

is relatively close to the data for the top 1%, but at the cost of not matching staying rates at different horizons or moments at the top 5% and 10%. However, the poorest match is that this long discount factor heterogeneity results in a capital income ratio far in excess of the target (and in excess of other models). Whilst this target could be matched by lowering one or both β's it appears the pressure to match other moments (such as inequality) prevents this from occurring.

Moment	Data	Mean	$Q_{0.05}$	$Q_{0.5}$	$Q_{0.95}$	s.d.
Top 1% share	0.206	0.13	0.1	0.14	0.15	0.018
Top 1% stay 2yr	0.67	0.74	0.67	0.76	0.79	0.05
Top 1% stay 4yr	0.59	0.74	0.68	0.76	0.79	0.048
Top 1% stay 6yr	0.55	0.74	0.65	0.75	0.78	0.051
Top 1% stay 8yr	0.51	0.73	0.66	0.75	0.78	0.053
Top 1% $P(T\|FFT)$	0.3	0.27	0.24	0.27	0.3	0.019
Top 5% $P(T\|FFT)$	0.39	0.32	0.3	0.31	0.34	0.01
$\sigma_{\Delta log(w)}$	0.34	0.32	0.3	0.33	0.35	0.016
K:Y ratio	2.5	3.36	2.87	3.48	3.65	0.293

Table 2.9: Selected moments from data and estimation.

2.4.3 Returns heterogeneity

Returns heterogeneity can generate significant wealth inequality, either through high persistence of different returns and gradual accumulation or through high variance and sudden exogenous gains of wealth. It also has the advantage of being able to destroy or limit a stock of wealth through negative returns, something the other mechanisms lack. This can, for example, aid a speedy descent for some of the wealthy to help match mobility data as discussed earlier.

Parameter	Mean	$Q_{0.05}$	$Q_{0.5}$	$Q_{0.95}$	s.e.
ρ_r	0.328	0.119	0.328	0.535	0.088
σ_r	0.131	0.096	0.129	0.174	0.019
σ_v	0.207	0.172	0.204	0.247	2.022

Table 2.10: Estimated parameters

I estimate (annual) positive autocorrelation of approximately 0.33 and standard deviation of 0.13 for R. There is a trade off between autocorrelation and standard deviation, as agents need greater variance to gain enough wealth to match inequality when persistence of wealth returns is low, as seen in Table 2.11. This leads to negative correlation between ρ_r and σ_r. Unsurprising, in the correlation of parameters, ρ_r is positively correlated with measurement error volatility, as higher wealth returns persistence decreases mobility, leading to a need for measurement error σ_v to increase variation and mobility to that found in the data.

	ρ_r	σ_r	σ_v
ρ_r	1.00	-0.27	0.21
σ_r	-0.27	1.00	0.23
σ_v	0.21	0.23	1.00

Table 2.11: Correlation of parameters from estimation.

Top 1% (and below) wealth shares are accurately captured, as are conditional mobility moments. In table 2.12 there is a qualitative match to the data overall in terms of decreasing staying rates in top categories with greater time horizons, though the top 1% differentiation over time is not as large in the model as in the data.

One moment not used in the estimation is the general equilibrium interest rate r. This can be high in these estimations, ranging from 5% up to 10% with some R parameter sets. Given the significant variance in the single wealth asset it is not surprising that r is above the usual range that the risk-free market-clearing interest rates in general equilibrium models lie within.

Moment	Data	Mean	$Q_{0.05}$	$Q_{0.5}$	$Q_{0.95}$	s.d.
Top 1% share	0.206	0.2	0.11	0.2	0.27	0.041
Top 1% stay 2yr	0.67	0.7	0.64	0.7	0.75	0.026
Top 1% stay 4yr	0.59	0.66	0.6	0.66	0.73	0.025
Top 1% stay 6yr	0.55	0.61	0.55	0.61	0.71	0.026
Top 1% stay 8yr	0.51	0.58	0.5	0.58	0.68	0.029
Top 5% stay 2yr	0.73	0.69	0.65	0.69	0.73	0.02
Top 5% stay 4yr	0.68	0.64	0.61	0.64	0.68	0.016
Top 5% stay 6yr	0.63	0.6	0.56	0.6	0.64	0.014
Top 5% stay 8yr	0.61	0.56	0.52	0.56	0.6	0.014
Top 1% $P(T\|FFT)$	0.3	0.36	0.28	0.35	0.44	0.029
$\sigma_{\Delta log(w)}$	0.34	0.36	0.28	0.36	0.46	0.042

Table 2.12: Selected moments from data and estimation.

The 'true' fluctuations in wealth can be observed by studying simulations without the measurement error input. Examining the staying probabilities for agents with different histories in Table 2.13, there is a higher staying rate in the underlying structural model, with around 85% staying. In Table 2.14 the underlying model still demonstrates some of the 'stayers stay' pattern (more so than other estimated models), but is not as mobile as the previous results and the data.

As explained above, the effects of returns heterogeneity can be broken down into two major effects: returns affect both income today and saving incentives for tomorrow by

Source	top 10%	top 5%	top 1%
Data	0.76	0.72	0.65
with ME	0.72	0.69	0.68
underlying	0.86	0.84	0.85

Table 2.13: Probability of remaining in top wealth groups for data and estimated models.

Source	History	top 10%	top 5%	top 1%
Data	$T_3 \vert F_1 T_2$	0.51	0.37	0.4
Data	$T_3 \vert T_1 T_2$	0.88	0.83	0.79
with ME	$T_3 \vert F_1 T_2$	0.49	0.45	0.4
with ME	$T_3 \vert T_1 T_2$	0.81	0.8	0.82
w/out ME	$T_3 \vert F_1 T_2$	0.73	0.73	0.68
w/out ME	$T_3 \vert T_1 T_2$	0.88	0.87	0.89

Table 2.14: Probability of remaining in top wealth groups, given different histories for data and models. 'T_t' indicates 'True' for belonging to the group in wave t and 'F_t' indicates 'False' for the same.

realising gains or losses on the stock of wealth and by giving different expectations of future returns. In the case of exactly zero returns persistence, there is no difference in expected returns, but for the case of positive autocorrelation, there is an incentive to make savings decisions correlated with today's returns, to take advantage of future high returns by investing or to spend now to avoid the poor returns in the future. Of course, this ignores the counter-balance of wealth effects - there is a further effect that an agent who expects to be poorer from a negative wealth change is incentivised to keep saving in expectation of that potential poverty even though it is the low returns to wealth which would cause that poverty.[17].

These effects are very different to those generated by superstars. Superstar ability only directly changes the flow of wealth, not the stock. Not only this, but they do not have a negative flow aspect, and thus find it difficult to create mobility. In contrast, R shocks scale with wealth, ensuring the wealthy are equally vulnerable, and can result in negative income. The incentive effects under persistent returns shocks are similar to discount factor shocks as β changes in future wealth value can be mapped to different future returns, but the discount factor variation does not include direct changes in the stock of wealth.

[17]This effect is small for the rich right tail focused upon

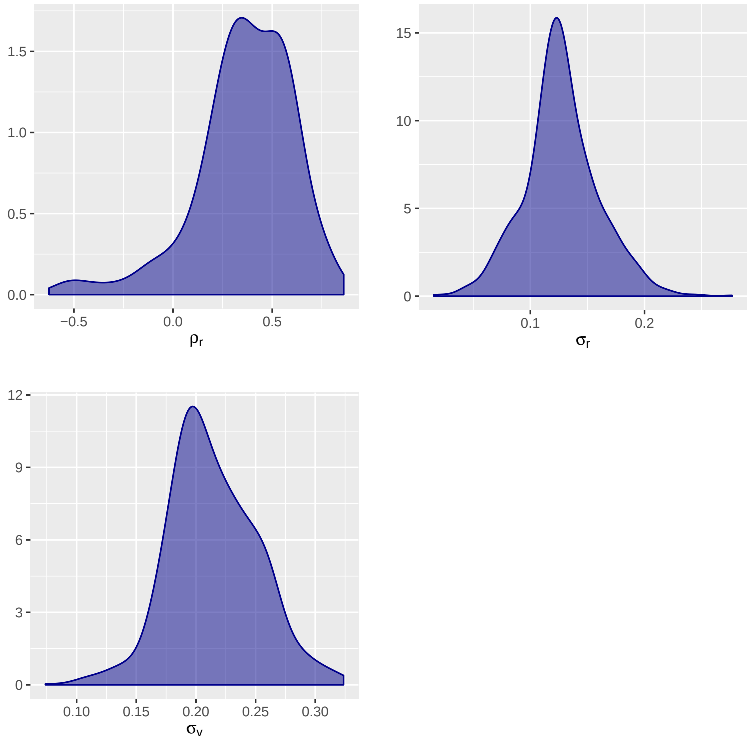

Figure 2.3: Estimated parameter density distribution for R shocks.

2.4.4 Joint Estimation

The parameters in the joint estimation of all three theoretical mechanisms are similar to the estimation restricted to R heterogeneity alone, with positive autocorrelation in R of 0.5 and standard deviation of 0.1. Superstars are not very super, with an average estimate of only 4 times median earnings for a superstar population of the top 0.6%, as opposed to the approximate 50 time median earnings for the top 0.1% needed to match inequality solely using superstars. The two levels of discount factors have some deviations but are extremely short-lived versus the 50 year average duration in Krusell & Smith or the expected 2000 years in the β only estimation, with agents staying in a state for an average of 3 years and inheriting the same ability with a roughly 50% chance[18]. Measurement error volatility also displays a similar level to that with R shocks alone, with σ_v close to 0.2.

The model fits key targets, including both wealth shares and staying probabilities - I show the full estimation results for the joint model and the individual mechanism models in table 2.16. It is unsurprising that the fit to many targets for the joint estimation is very similar to that with R shocks alone, given the similarity of parameters.

In table 2.17 I compare the data, model results and underlying fluctuations for staying rates. Wealth mobility is somewhat lower than wealth surveys, but still very much present. Similarly, there is still a pattern that new entrants are less likely to stay.

[18]The same as the symmetric stationary distribution probabilities.

Parameter	Mean	$Q_{0.05}$	$Q_{0.5}$	$Q_{0.95}$	s.e.
ρ_r	0.502	0.308	0.533	0.625	0.088
σ_r	0.1	0.071	0.097	0.134	0.019
Y	4.341	2.225	3.747	9.781	2.022
$P_{Y,enter}$	0.002	0	0.001	0.005	0.002
$P_{Y,exit}$	0.312	0.03	0.366	0.582	0.191
$P_{\beta,d}$	0.475	0.122	0.627	0.829	0.261
P_β	0.678	0.412	0.687	0.957	0.153
β_l	0.949	0.894	0.957	0.984	0.029
β_h	0.949	0.925	0.942	0.985	0.02
σ_v	0.232	0.206	0.225	0.275	0.02

Table 2.15: Estimated parameters

Moment	Target	Joint	R	β	Superstars
Top 1% wealth share	0.21	0.2	0.2	0.13	0.13
Top 5% wealth share	0.38	0.36	0.35	0.31	0.34
Top 10% wealth share	0.48	0.48	0.46	0.46	0.48
Prob. stay top 5%, 2yr	0.73	0.7	0.69	0.68	0.74
Prob. stay in top 1%, 2yr	0.67	0.69	0.7	0.74	0.61
Top 1% $P(T\|FT)$	0.37	0.39	0.41	0.36	0.39
Top 1% $P(T\|TT)$	0.81	0.82	0.82	0.87	0.75
Top 5% $P(T\|FT)$	0.44	0.45	0.45	0.4	0.44
Top 5% $P(T\|TT)$	0.87	0.81	0.8	0.81	0.84
Top 1% $P(T\|FFT)$	0.3	0.33	0.36	0.27	0.35
Top 5% $P(T\|FFT)$	0.39	0.4	0.41	0.32	0.39
Top 1% $P(T\|TTT)$	0.87	0.87	0.87	0.91	0.8
Top 5% $P(T\|TTT)$	0.88	0.85	0.84	0.87	0.88
Prob. stay in top 1%, 4yr	0.59	0.65	0.66	0.74	0.59
Prob. stay in top 1%, 6yr	0.55	0.62	0.61	0.74	0.56
Prob. stay in top 1%, 8yr	0.51	0.58	0.58	0.73	0.53
Prob. stay in top 5%, 4yr	0.68	0.66	0.64	0.68	0.72
Prob. stay in top 5%, 6yr	0.63	0.62	0.6	0.67	0.7
Prob. stay in top 5%, 8yr	0.61	0.58	0.56	0.66	0.67
Prob. stay in top 10%, 4yr	0.71	0.7	0.68	0.75	0.72
Prob. stay in top 10%, 6yr	0.68	0.66	0.63	0.74	0.7
Prob. stay in top 10%, 8yr	0.63	0.62	0.59	0.73	0.69
$\sigma_{\Delta log(wealth)}$ $(> Q_2)$	0.34	0.38	2.46	0.32	0.44
K:Y Ratio	2.5	2.45	0.36	3.36	2.67

Table 2.16: Mean Estimation Moments.

Source	top 10%	top 5%	top 1%
Data	0.77	0.73	0.67
with ME	0.72	0.69	0.69
underlying	0.88	0.87	0.88

Table 2.17: Probability of remaining in top wealth groups for data and estimated models.

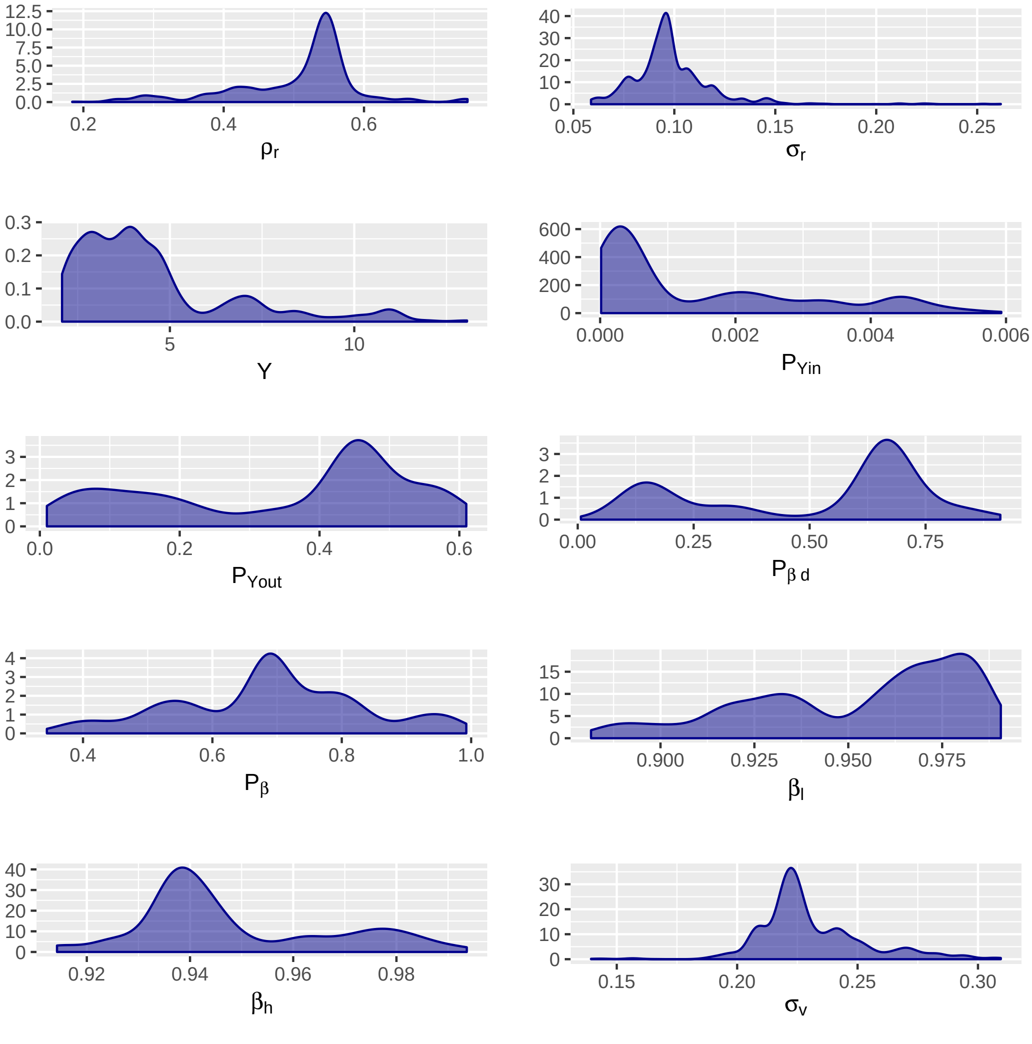

Figure 2.4: Estimated parameter density distribution for joint estimation.

2.5 Robustness

In this section, I check robustness of these results with two examples: firstly, implementing 'real superstars' - taking high earnings from the data and using their earnings levels and dynamics to fit the superstar earnings process whilst estimating the other parameters. Secondly, restricting measurement error to be a ratio to variation in wealth, based on findings from a measurement error identification exercise in the previous chapter.

2.5.1 Real Superstars

One simple way to test the robustness of the estimation is to consider changing the earnings process - high earners can be identified in the WAS dataset and in administrative data, as mentioned earlier, so information can be used to implement realistic superstar earnings. From this, I can examine whether my results from the main estimation continue to hold, or does the prominence of returns heterogeneity wither when faced with high earnings from the data?

I implement superstars based on earnings of the top 0.1% and re-estimate the remaining discount factor heterogeneity and wealth returns parameters. Using the WAS and the administrative earnings data (De Nardi et al. [2016]), the top 0.1% of earners have a yearly transition probability of 0.0004 into this category and 0.4 out of it, with an average earnings of about 30 times the median (which, as mentioned earlier, is approximately half the level needed to match cross-sectional inequality). I note the earnings transition probabilities of the top 0.1% are similar to the top 1%, 0.5% and 0.01%.

Parameter	Mean	$Q_{0.05}$	$Q_{0.5}$	$Q_{0.95}$	s.e.
ρ_r	0.478	0.241	0.482	0.702	0.131
σ_r	0.129	0.084	0.127	0.182	0.024
P_β	0.225	0.083	0.247	0.351	0.079
$P_{\beta,d}$	0.565	0.452	0.546	0.783	0.093
β_l	0.97	0.949	0.971	0.988	0.011
β_h	0.969	0.944	0.97	0.987	0.012
σ_v	0.277	0.236	0.275	0.332	0.027

Table 2.18: Estimated parameters

I find similar results to the earlier joint estimation, though the variation of wealth returns is higher to compensate for the lower mobility the data-superstars cause[19]. In line with this reasoning, σ_r is somewhat higher. Discount factor persistence is very low, with an average duration of less than 2 years. There are some differences between the two β's despite similar mean levels. This short duration β variation is also likely to stem from

[19] As the joint estimation is allowed to have very low earning superstars, there is less pressure towards immobility.

pressure to mitigate immobility caused by superstars. Superstar-sourced immobility can also explains the need for higher measurement error variation.

Moment	Data	Mean	$Q_{0.05}$	$Q_{0.5}$	$Q_{0.95}$	s.d.
Top 1% share	0.206	0.2	0.165	0.196	0.242	0.019
Top 5% share	0.385	0.371	0.327	0.366	0.429	0.024
Top 10% share	0.478	0.491	0.446	0.488	0.549	0.024
Top 5% stay	0.73	0.691	0.652	0.689	0.74	0.021
Top 1% stay	0.67	0.724	0.681	0.723	0.767	0.021
Top 1% $P(T\|FT)$	0.37	0.449	0.396	0.451	0.5	0.017
Top 1% $P(T\|TT)$	0.81	0.827	0.796	0.827	0.861	0.013
Top 5% $P(T\|FT)$	0.44	0.435	0.387	0.434	0.479	0.018
Top 5% $P(T\|TT)$	0.87	0.807	0.781	0.807	0.832	0.013
$\sigma_{\Delta log(w)}$	0.34	0.444	0.385	0.439	0.517	0.036
K:Y ratio	2.5	2.772	2.462	2.762	3.066	0.136

Table 2.19: Selected moments from data and estimation.

The fit to the wealth and mobility targets is similar, as would be expected. However, wealth variance and K:Y ratio are too large (rather like the results from superstars alone). The pattern of higher staying rates at the top 1% versus the top 5% is in conflict with the data. Otherwise, the overall conclusion is that the qualitative and major quantitative results from earlier parts are not largely affected by direct use of earnings data for superstars.

2.5.2 Proportionally Restricted Measurement Error

As an alternative benchmark to directly fitting measurement error, I utilise the procedure of Lee et al. [2017] to identify the size of i.i.d. time-varying measurement error variance in the WAS in the earlier chapter. Using an AR1 dynamic panel instrumental variable GMM estimation in the style of Arellano and Bond [1991], I found the measurement error standard deviation to be half that of 'true' equation error standard deviation, suggesting it has a quantitatively significant presence, but does not dominate. I now use this ratio of estimated measurement error standard deviation to total standard deviation of changes in log wealth to generate the size of measurement error for a given model output, i.e. using the model-generated wealth volatility to anchor a proportional measurement error variance ("every unit of wealth variance has x units of measurement error variance"). Under this restriction, I add a proportionally fixed amount of measurement error to the model output each time, rather than allowing σ_v to fluctuate and using a target of wealth variation and other dynamics to identify it.

I use a minimiser in each estimation iteration to find a σ_v that creates an output wealth process with a 1:2 ratio of $\sigma_v : \sigma_{\Delta log(w)}$.

Comparing to the joint estimation, positive wealth returns autocorrelation is stronger, near to 0.8 rather than 0.5 and standard deviation is correspondingly lower (as it has to decrease with higher autocorrelation to have similar inequality). Superstar earnings are no longer extremely low and instead around 17x median earnings, which is close to the data level for the top 0.5%, though with higher exit. Discount factor heterogeneity is larger, but similarly (im)persistent.

Parameter	Mean	$Q_{0.05}$	$Q_{0.5}$	$Q_{0.95}$	s.d.
ρ_r	0.722	0.4	0.756	0.936	0.149
σ_r	0.07	0.036	0.067	0.118	0.021
Y	13.074	2.745	15.806	20.39	6.221
$P_{Y,enter}$	0.002	0	0.002	0.004	0.001
$P_{Y,exit}$	0.728	0.557	0.756	0.848	0.086
$P_{\beta,d}$	0.772	0.634	0.775	0.92	0.081
P_{β}	0.566	0.368	0.564	0.753	0.121
β_l	0.952	0.918	0.951	0.98	0.016
β_h	0.961	0.918	0.965	0.993	0.023

Table 2.20: Estimated parameters

I show the match to the data for proportional measurement error and real superstars versus the main joint estimation and the data in Table 2.21. The inequality and mobility moments are better matched under proportional measurement error, at the cost of excessive wealth variation at 0.47. In particular, the probabilities of staying in different groups over different horizons are very well matched. Without variance of changes in log wealth as a target, the generated value of σ_v causes excessive variance of changes in log wealth. With greater measurement error, the wealth-inequality-generating theories have less pressure to generate mobility. I do not target the wealth variance in this exercise as that would push σ_v to take a specific value like the other estimations rather than simply respond proportionally to the variation in wealth generated by the mechanisms.

Moment	Target	Joint	Restricted M.E.	Real Superstars
Top 1% wealth share	0.21	0.2	0.21	0.2
Top 5% wealth share	0.38	0.36	0.4	0.37
Top 10% wealth share	0.48	0.48	0.53	0.49
Prob. stay top 5%, 2yr	0.73	0.7	0.69	0.69
Prob. stay in top 1% (2yr)	0.67	0.69	0.65	0.72
Top 1% $P(T\|FT)$	0.37	0.39	0.37	0.45
Top 1% $P(T\|TT)$	0.81	0.82	0.8	0.83
Top 5% $P(T\|FT)$	0.44	0.45	0.43	0.43
Top 5% $P(T\|TT)$	0.87	0.81	0.81	0.81
Top 1% $P(T\|FFT)$	0.3	0.33	0.32	0.42
Top 5% $P(T\|FFT)$	0.39	0.4	0.39	0.4
Top 1% $P(T\|TTT)$	0.87	0.87	0.85	0.86
Top 5% $P(T\|TTT)$	0.88	0.85	0.85	0.85
Prob. stay in top 1%, 4yr	0.59	0.65	0.62	0.67
Prob. stay in top 1%, 6yr	0.55	0.62	0.58	0.62
Prob. stay in top 1%, 8yr	0.51	0.58	0.55	0.58
Prob. stay in top 5%, 4yr	0.68	0.66	0.65	0.64
Prob. stay in top 5%, 6yr	0.63	0.62	0.62	0.6
Prob. stay in top 5%, 8yr	0.61	0.58	0.58	0.56
Prob. stay in top 10%, 4yr	0.71	0.7	0.69	0.66
Prob. stay in top 10%, 6yr	0.68	0.66	0.65	0.62
Prob. stay in top 10%, 8yr	0.63	0.62	0.61	0.58
$\sigma_{\Delta log(wealth)}$ $(> Q_2)$	0.34	0.38	2.6	0.44
K:Y Ratio	2.5	2.45	0.47	2.77

Table 2.21: Mean Estimation Moments.

2.6 Conclusions

My conclusion is that by using transitions in top wealth groups I can identify exogenous wealth returns heterogeneity as the wealth accumulation mechanism that best explains the inequality and mobility data. I find that discount factor heterogeneity and superstar earnings cannot match inequality and mobility simultaneously on their own. When the three theories are combined in a joint estimation, I find returns heterogeneity dominates. I explain these results through the ability of returns heterogeneity to account for higher mobility due to affecting wealth via two mechanisms - direct changes to the stock of wealth/budget constraints and changes to savings incentives via different expected future returns. This can create the fast wealth losses we see in the data.

I use a number of facts about fluctuations in wealth amongst the wealthy from the longitudinal and representative WAS wealth dataset in an estimation of theories generating wealth inequality. My modelling matches these patterns and demonstrates how such data is useful for identifying different processes behind wealth inequality. By identifying the mechanisms generating wealth inequality and mobility in a clear methodology and explaining why they fit the data, I hope to contribute to better modelling of the real processes governing the wealth distribution. The results make clear that any process hoping to be realistic and match mobility must have a direct impact on both the budget constraint and change savings incentives to generate the rapid (downwards) changes in wealth in the data.

This chapter suggests that when considering the wealth distribution, study into how and why these differential returns come about and their impact is of greater importance that studying earnings. For development, these models do not explicitly consider entrepreneurship, nor portfolios or risk preferences which would be natural routes to follow, given the importance of wealth returns I find and this data has the potential to be informative about this.

Chapter 3

Entrepreneurs, Turbulence and Inequality Dynamics - Who Has Wealth Matters.

This chapter quantitatively studies the recessionary effects of firm-level productivity dispersion on credit-constrained entrepreneurs through capital misallocation. It links business fluctuations and turbulence during recessions to dynamics of personal wealth and inequality through a heterogeneous model of entrepreneurship and aggregate shocks. In firm-level data, there is greater mobility and a wider distribution of turnover changes during recessions, so I calibrate aggregate shocks to entrepreneurial productivity transitions and add to a heterogeneous Cagetti-De Nardi model. The increases in turbulence cause quantitatively substantive and persistent negative responses of capital, consumption and output. Propagation occurs through increases in capital misallocation under turbulence. This is due to credit constraints with endogenously greater impact on less productive or smaller entrepreneurial firms: newly productive firms are unable to fully utilise upward productivity gains whilst previously productive large firms remain holding inefficiently large capital stocks. The transmission of these effects are somewhat counter-balanced by the response of unconstrained corporate firms. Negative shocks to productivity levels for all producers cause a large reaction and slow recovery from the entrepreneurial sector which dominates the economy's response, amplifying the initial shock.

3.1 Introduction

This chapter makes three main contributions to the literature: developing a heterogeneous general equilibrium model with aggregate shocks that matches and explains features of wealth inequality, business trends and business cycles; utilising a new type of distributional entrepreneurial shock that affects aggregates through micro-changes and discussing the role of entrepreneurial constraints over the business cycle.

Entrepreneurs form a large part of the wealthiest in advanced economies (40%+ of the U.S. or U.K. top 1%[1]) and hold a large part of their wealth in their businesses. The framework of Quadrini [2000], Cagetti and Nardi [2006] and Cagetti and Nardi [2006] uses credit-constrained entrepreneurs to generate realistic wealth inequality, through a reliance on personal wealth to collateralise business borrowing. This is justified by the empirical dominance of business owners, entrepreneurs and the self-employed at upper wealth quantiles and their relatively undiversified asset portfolios. The modelling of entrepreneurs has implications for tax policy in Cagetti and Nardi [2004] and Kitao [2008] whilst Bassetto et al. [2015] examine the impact of rising financial intermediation costs. Further examples are bankrupcty in Meh and Terajima [2008] and incorporation decisions in Short and Glover [2011].

I incorporate aggregate shocks into this framework and solve with a Krusell-Smith methodology (Krusell and Smith Jr. [1998]), examining total factor productivity (TFP) shocks for constrained entrepreneurial firms as well as uncertainty shocks to stochastic, idiosyncratic entrepreneurial productivity transitions, which I call 'turbulence shocks'. These shocks increase the probability of changes in productivity for entrepreneurs and are similar to those in the work of Bloom and others, for example Bloom et al. [2016], Bloom et al. [2018] and Bloom [2014]. These works use time varying higher moments of stochastic shock processes, typically for a heterogeneous distribution of firms and study the effects of changes in uncertainty. In this case, aggregate shocks increase the dispersion of entrepreneurial productivity transitions and are motivated by patterns of greater recessionary dispersion observed in longitudinal firm data. There is also supporting evidence in observations from the Wealth and Assets Survey that the probability of transition out of top income and wealth groups increases in recessions, as also found in Guvenen et al. [2014b], Guvenen et al. [2014a] and Auten et al. [2013].

Entrepreneurs and workers are aware of the likelihood of state changes, aggregate shocks and resultant turbulence, which they incorporate into decision making. The transmission mechanism from a turbulent shock is the unequal effects of productivity shocks on different entrepreneurs. Some rich and productive entrepreneurs fall to a relatively unproductive firm status (which also has a higher chance of exit). This imposes a tighter borrowing constraint on them, as borrowing is positively dependent on the creditor-seizable output of the firm, which is correlated with productivity. Thus, these previously highly productive entrepreneurs experience a drop in income which prevents their building of wealth to expand their firm borrowing constraint. Further, if they exit, their capital becomes very unproductive, earning only the risk-free rate for lenders. Those benefiting from a positive shock are constrained by their available capital and so must increase their

[1]Based on the U.S. Survey of Consumer Finances and U.K. Wealth and Assets Survey

wealth and personal firm size slowly. Simply put, constraints ensure there is no easy way to redistribute capital from previously productive entrepreneurs to newly productive ones. hence, greater mobiltiy in productivity creates more allocation problems.

This leads to lower output (as entrepreneurs borrow, employ and produce less), lower consumption and also lower inequality, after a very brief initial period of increased precautionary saving. Inequality in income and wealth reduces as entrepreneurs dominate the top of the distribution and are, as a group, poorer. In aggregate, there is a move of both capital and labour from more to less productive usage, whilst individuals simultaneously reduce savings through normal self-insurance motives.

However, if the turbulence shock increases variance symmetrically, there is a post-dip boom as the higher density of entrepreneurs with improved productivity from the shock begin to build capital and expand their borrowing constraints over time. This mechanism can also be considered alongside the concept of a 'financial accelerator' in Bernanke et al. [1999]. They propose that positive productivity shocks can incite a virtuous cycle of greater profits, expansion of entrepreneur-level capital constraints due to increased entrepreneurial assets to leverage upon and further profits. I find a similar effect (although I focus on the negative, inverse side of the mechanism), whereby there is a persistent response of the economy to changes in individual entrepreneurial productivity.

In many ways, this approach is closest to that of Khan and Thomas [2013], who study the effect of TFP and credit shocks on credit-constrained firms. Their framework is motivated by aggregate expansion and contraction of firm borrowing for (partly irreversible) investment. They find the distribution of capital amongst heterogeneous firms interacts with borrowing shocks to cause contractions in output and other aggregates. Like Khan and Thomas, I examine the implications of non-optimal capital allocation after shocks and the inability of agents to redistribute effectively. However, my innovation is the combination of the heterogeneous entrepreneurial household framework with endogenous borrowing constraints and aggregate shocks that also change the probability distribution of entrepreneurial productivity.

The findings are also close to Moll [2014] regarding the persistence of entrepreneurial ability and the ability of entrepreneurs to self-finance. Moll finds persistence of entrepreneurial ability engenders less steady state capital misallocation due to a greater ability to self-finance entrepreneurial operations. However, it also causes slower transitions between states such that studying steady states alone can be misleading. This result shows the value of fully implementing aggregate shocks in models of entrepreneurship and capital constraints. In my work, high turbulence reduces persistence and stability of entrepreneurial ability and thus, as expected, I find similar results of reduced inequality and reduced aggregate output, through a similar mechanism. Achdou et al. [2014] use a continuous time model incorporating standard aggregate TFP shocks and entrepreneurs, finding that changes in the wealth distribution are extremely slow. Although the general framework is similar to this study, turbulence shocks, entrepreneurial heterogeneity and endogeneity of the borrowing constraint are not.

To examine the distribution of firm-level changes, I use an administrative dataset containing all active U.K. firms (including small firms and single-person businesses). This covers 1997 to 2015, and within the series there are recession-associated rises in dispersion of changes in turnover, as well as a drop in the mean, indicating that the distribution of outcomes for entrepreneurs widens whilst becoming more negative. I discuss these changes, and incorporate them into the model.

U.S. data for entrepreneurship and business income has been studied in DeBacker et al. [2012], who find the distribution of business income has greater spread and much greater likelihood of extreme values, propelling people rapidly through the income distribution versus labour income. They also find business income falls as a proportion of overall income over the Great Recession. Evans and Jovanovic [1989] attempt to structurally identify the capital constraints for entrepreneurs, finding a value of 1.5x wealth, whilst Moskowitz and Vissing-Jørgensen [2002] provide a study of the relative returns to private equity and business ownership with surprisingly low private equity returns.

This chapter first describes the model before covering the data for calibration. The administrative dataset containing all U.K. firms is combined with information from wealth surveys and time series of macroeconomic aggregates to provide targets for fitting the model. Then, the impact of different aggregate shocks is discussed and the results are examined in a series of Impulse Response Functions and simulations before concluding.

3.2 Model Description

I use a general equilibrium Bewley [1977] model where agents face individual shocks and can partially insure themselves through asset holdings, largely following Cagetti and Nardi [2006] and Bassetto et al. [2015]. There is a continuum of households of measure one. These households can use their resources to save in capital or to consume an output/consumption good. They also make an occupational choice to be a worker or an entrepreneur during each period, facing both idiosyncratic and aggregate shocks. If they become a worker, they inelastically supply an amount of labour, determined by their idiosyncratic (and stochastic) worker productivity, also known as 'effective labour'. This strategy (common in the macroeconomic literature) enables the representation of a distribution of wages with minimal computational impact. If they have entrepreneurial ability, they may choose to become an entrepreneur and then select levels of capital and labour to employ in their firm to maximise their profit. There is also a representative corporate firm, which follows a standard Cobb-Douglas production function and is designed to reflect the existence of large, publicly owned firms not facing the same constraints as entrepreneurs and privately held businesses. Both types of firm (entrepreneurial and corporate) operate in the same input/output markets and both are price takers, so pay the same interest rates and wages and produce the same output good, which households purchase. There is an aggregate shock which, remaining ambivalent about its exact nature, is included as a state z that can affect levels of productivity and transitions.

3.2.1 Households

There is a measure one continuum of households. They are infinitely lived and discount the future, with discount factor β. They aim to maximise utility and have constant relative risk aversion, so their expected utility function is

$$\mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \frac{c_t^{1-\sigma}}{1-\sigma}.$$

$\mathbb{E}_t$ denotes expected value in period t and c_t is consumption in t.

Households retire and die probabilistically, following the approach of Cagetti and Nardi [2006] and Cagetti and Nardi [2004], with an average working life of 45 years, and retirement of 11 years (calibrated to UK statistics). They also pay taxes on their income, and estate taxes on their wealth when necessary. Retirees cannot work, but can still run businesses if they have the relevant ability. If not running a business, retirees will be given a fixed social security income p instead of earnings wy. Throughout the following exposition, these additional life-cycle elements are ignored to simplify the explanation of entrepreneurial and worker behaviour.

3.2.2 Technologies

Each agent has some ability y as worker, which reflects the unit of effective labour he can supply, and some ability θ as entrepreneur.

Entrepreneurial output is given by

$$f(k, n, \theta) = \theta(z)(k^{\gamma}(1+n)^{(1-\gamma)})^{\nu}.$$

where k is the entrepreneur's working capital, n is workers' labour employed by the entrepreneur and z is the aggregate state/

Entrepreneurs face decreasing returns from investment ($0 < \nu < 1$) due to the difficulty of stretching managerial skills across larger projects ("span of control"). Hence, while entrepreneurial ability is exogenous, an entrepreneur's return from investing in capital is an endogenous function of his project's size.

The entrepreneur uses all of his labour to run the entrepreneurial technology and n is labour employed in addition to the one unit of the entrepreneur. If the optimal amount of labour is smaller than 1, the entrepreneur only employs his own unit of labour.

The real economy is made up of both small entrepreneurial firms and larger corporate firms, which are unlikely to face entrepreneurial financing restrictions. To represent this, a corporate production sector is included, which does not face borrowing constraints and is represented by a standard Cobb-Douglas production function:

$$F(K_c, L_c) = \theta^c(z)K_c^{\alpha}L_c^{1-\alpha},$$

where K_c, L_c are, respectively, aggregate capital and aggregate labour employed by the corporate sector. θ^c is the productivity of the corporate sector, which may be affected by z. In both sectors capital depreciates at the same rate, δ.

The corporate sector maximises profit at the representative firm, setting prices equal to marginal products through first order conditions. Since the entrepreneurs and corporates are competitive, the equilibrium interest rate and wage for period t is given by the corresponding marginal products in the representative corporate sector, which are

$$r_t = \alpha\theta^c(z)\left(\frac{K_{ct}}{L_{ct}}\right)^{\alpha-1}$$

$$w_t = (1-\alpha)\theta^c(z)\left(\frac{K_{ct}}{L_{ct}}\right)^{\alpha}$$

One can rearrange such that w_t is a function of r_t,

$$w_t = (1-\alpha)\theta^c(z)\left(\frac{r_t+\delta}{\alpha}\right)^{\frac{\alpha}{\alpha-1}}$$

This represents a no-arbitrage restriction - entrepreneurs wishing to use capital saved by households must offer to pay the same interest rate as the corporate firms. In equilibrium, the result is that the corporate sector is somewhat passive, absorbing labour and capital supply above the level demanded by entrepreneurial firms. All capital pays r to savers and all effective labour units are paid w.

While entrepreneurial firms direct their profits to the owner, corporate firms have constant returns to scale and make zero profit so their ownership and size is not important.

3.2.3 Credit markets

Working capital $k = (a + b)$ in an entrepreneurial business equals the entrepreneur's own assets (a) plus or minus assets that are borrowed or lent (b). The method of borrowing is direct lending of capital, where the entrepreneur agrees to pay the prevailing (corporate) interest rate on the capital he borrows. This lending is carried out post-shock realisation, so there is no uncertainty on the returns to these capital loans.

Assume that if the entrepreneur is borrowing and defaults, the creditor seizes a fraction $0 < \lambda < 1$ of output and undepreciated capital after wages are paid. It follows that if the entrepreneur wishes to borrow, the amount that the lender is willing to lend is such that debt and interest on debt are repaid in case of default

$$(1+r)(k-a) \leq \lambda\Big(f(k,n,\theta) + (1-\delta)k - wn\Big),$$

which can be written as

$$k \leq \lambda\frac{(f(k,n,\theta) + (1-\delta)k - wn)}{(1+r)} + a.$$

The entrepreneur can thus invest an amount k that satisfies the above equation. If the entrepreneur is not borrowing constrained, the optimal firm size will be implemented and leftover assets will be invested at the equilibrium interest rate.

Notice that the permitted borrowing, b depends on one's assets (a), entrepreneurial ability (θ) and the equilibrium prices (which depend on the aggregate shock.

3.2.4 The basic household decision problem

At the beginning of each period, current ability levels and prices are fully known, whilst future ones are unknown. Each individual starts a period with assets a, entrepreneurial ability $\theta(z)$ and worker ability y, and chooses whether to be an entrepreneur or a worker in the current period (a binary indicator, e) and then how much to save for next periods assets (a').

Both workers and entrepreneurs supply all their labour (1 unit) inelastically. A worker's unit becomes y units in the labour market and the entrepreneur supplies 1 to his project. Households, whether entrepreneurs or workers, borrow and lend at rate r.

The state variables for the household's problem are given by the household's assets (a), ability levels (y, θ), the aggregate state z and prices r, w (all actors are price takers). These are the variables the agent takes as given when making his choices. a is endogenous to households decisions, but y, θ and z are all exogenous and follow transition matrices Π_y, $\Pi_\theta(z)$ and Π_z. The functional form of $\theta(z)$ is defined later - the below applies for multiple entrepreneurial abilities with any form of underlying transition $\Pi_\theta(z)$, which may be aggregate state dependent.

The infinitely-lived household's problem can thus be written recursively as:

$$V(a,y,\theta;z,w,r) = \max\{V_e(a,y,\theta;z,w,r), V_w(a,y,\theta;z,w,r)\}, \tag{3.1}$$

$$V_e(a,y,\theta;z,w,r) = \max_{c,k,n,a'}\{u(c) + \beta\mathbb{E}_t(V(a',y',\theta';z',w',r'))\}. \tag{3.2}$$

where V_e is the entrepreneur's value (e=1) and V_w is the worker's value (e=0). The expectation of the future value function is taken with respect to $(y', \theta'; w', r')$, conditional on $(y, \theta; z, w, r)$. The maximization process is subject to the following constraints:

$$f(k, n, \theta, z) = \theta(z)(k^{\gamma}(1+n)^{(1-\gamma)})^{\nu}. \tag{3.3}$$

$$c + a' = f(k, n, \theta, z) + (1-\delta)k - wn - (1+r)(k-a), \tag{3.4}$$

$$k \leq \lambda \frac{(f(k, n, \theta, z) + (1-\delta)k - wn)}{(1+r)} + a \tag{3.5}$$

$$a \geq 0, \tag{3.6}$$

$$k \geq 0. \tag{3.7}$$

3.2.5 The worker's problem

The worker solves the following problem

$$V_w(a, y, \theta; z, r, w) = \max_{c, a'} \{u(c) + \beta \mathbb{E}_t V(a', y', \theta'; z', r', w')\} \tag{3.8}$$

subject to equation (3.6) and

$$c + a' = (1+r)a + w\, y, \tag{3.9}$$

where w is the given wage in the state.

3.2.6 The entrepreneur's problem

If the household has decided to become an entrepreneur, he selects how much capital and labour to use, as well as how much of the output to save to maximise his value function, subject to his capital borrowing constraint, current aggregate productivity and prices.

The first order condition for n from the production function and budget constraint imply:

$$1 + n^*(w, z) = \left[\frac{w}{(1-\gamma)\nu} \quad \frac{1}{\theta z^{\eta} k^{\gamma\nu}} \right]^{\frac{1}{(1-\gamma)\nu - 1}}.$$

Let

$$\mathrm{t}_1 = \frac{-\gamma\nu}{(1-\gamma)\nu - 1}; \quad \mathrm{f}_1 = \left[\frac{w}{(1-\gamma)\nu\theta} \right]^{\frac{1}{(1-\gamma)\nu - 1}}$$

so,

$$1 + n = \mathrm{f}_1\, k^{\mathrm{t}_1} \tag{3.10}$$

This is the optimal n the entrepreneur will choose, given k.

The first order condition for k implies:

$$k = \left[\frac{r + \delta}{\gamma\nu\theta(1+n)^{(1-\gamma)\nu}} \right]^{\frac{1}{\gamma\nu - 1}}$$

Notice that for $0 < \gamma < 1$ and $0 < \nu < 1$, n is increasing in k, and k is increasing in n.

Either the entrepreneur wishes to hire $n > 0$, or he only uses his own labour $n = 0$. If the capital borrowing constraint binds, the two relevant equations to use are the capital borrowing constraint and the first order condition for n. Otherwise, the optimal firm size applies.

First, the first order condition for n is used to the find the level of k for which n is equal to zero. I call this k level k^{break}

$$k^{\text{break}} = \Big(\frac{w}{(1-\gamma)\nu\theta z^{\eta}}\Big)^{\frac{1}{\gamma\nu}}$$

Assuming that the entrepreneur is in the situation where $n > 0$, then substituting optimal n as a function of k in the non-linear equation implied by the borrowing constraint:

$$0 = a - k + \frac{\lambda}{1+r}\Big[w + (1-\delta)k - wf_1 k^{\mathrm{t}_1} + \theta \mathrm{f}_1{}^{(1-\gamma)\nu} k^{\gamma\nu + \mathrm{t}_1(1-\gamma)\nu}\Big]$$

Let us also define an additional f_2, f_3 and t_2,

$$\mathrm{f}_2 = \lambda\theta \mathrm{f}_1{}^{(1-\gamma)\nu}$$

$$\mathrm{t}_2 = \gamma\nu + \mathrm{t}_1\,(1-\gamma)\nu$$

$$\mathrm{f}_3 = -\lambda w\,\mathrm{f}_1$$

Thus, the (comparatively) simple equation is derived,

$$0 = a(1+r) + \lambda w + (\lambda(1-\delta) - (1+r))k + \mathrm{f}_3 k^{\mathrm{t}_3} + \mathrm{f}_2 k^{\mathrm{t}_2} \tag{3.11}$$

which is a non-linear equation in k, given r and w. One can then compare the solution for k from this equation to k^{break}. If k is larger than k^{break} then this case is a feasible solution for k, n. If not, then the case where $n = 0$ and the entrepreneur only uses his own labour is considered instead. Then, $n = 0$ and k solves

$$0 = a(1+r) + (\lambda(1-\delta) - (1+r))k + \lambda\theta k^{\gamma\nu} \tag{3.12}$$

These equations form the two feasible constrained cases - either $n > 0$ and solve for k from constraint 3.11 and then use equation 3.10, or if this is infeasible choose $n = 0$ and solve for k from the borrowing constraint, given $n = 0$. Checking if the optimal firm size is achievable and solving the equations above as so if not gives a policy of $\{k(a,\theta,w,r,z), n(a,\theta,w,r,z)\}$ for the entrepreneur and a budget constraint. One can compare this to the worker value, take the preferred career option, then use value function iteration to find the savings policy.

3.2.7 Equilibrium and Algorithm

Equilibrium in this model follows the Krusell-Smith (K-S) definition used in their equity premium paper (Krusell and Smith Jr. [1997]) and the calculation of r is similar to the calculation of bond prices in their paper. The recursive definition of equilibrium in this model is given in a similar[2] manner:

First, define Γ as the aggregate state, consisting of the distribution over entrepreneurial ability, worker productivity, assets and more generally any variable (e.g. prior corporate capital usage) that influences current or future behaviour. H is the law of motion for the aggregate state, using transition matrix Π_z for productivity transitions,

$$\Gamma' = H(\Gamma, z, z')$$

Next, define policy functions over savings (a'), entrepreneurial capital (k) and labour (n) and choice to become entrepreneur (e),

$$a'(a, \theta, y, z, \Gamma)$$

$$k(a, \theta, y, z, \Gamma)$$

$$n(a, \theta, y, z, \Gamma)$$

$$e(a, \theta, y, z, \Gamma), e\epsilon\{0, 1\}$$

Note that $k(.)$ and $n(.)$ are null if $\theta = 0$ since entrepreneurial policy is meaningless when one has no entrepreneurial ability. Similarly, $e = 0$ if $\theta = 0$.

Pricing functions r, w are defined as in section 3.2.2.

Together, the law of motion H, the functions a', k, n, e define an allocation.

Agents (workers, entrepreneurs and firms) optimise according to their problems as defined in sections 3.2.4, 3.2.5 and 3.2.6 and markets clear, obeying no arbitrage between different sectors (entrepreneurial and corporate). For labour and capital, market clearing imposes that individual labour and capital supply sums to aggregate supply and equals aggregate demand, which itself is summed from individual demand:

$$\int_i (1 - e_i) y_i di = L = L^c + L^e = L^c + \int_i e_i n_i di$$

$$\int_i a'_i di = K = K^c + K^e = K^c + \int_i e_i k_i di$$

The set of allocation, policy functions and pricing functions, together with market clearing and solution of agents' problems, defines an equilibrium.

As per K-S, I attempt to find an approximate equilibrium, by suitable definition of $\hat{\Gamma}$

[2] see section 2.2 of Krusell and Smith Jr. [1997], p395.

which approximates true Γ in the dimensions relevant to the agent's decisions.

First, one must define a computationally feasible aggregate state $\hat{\Gamma}$ and laws of motion $\hat{H}$. In the original K-S (Krusell and Smith Jr. [1998]) there is only one intertemporal decision, savings in capital, which requires knowledge of future prices. Today's average capital and the aggregate shock state are (combined) a good predictor of average capital tomorrow, and therefore of tomorrow's prices as there is no production heterogeneity or labour hours margin. K-S use:

$$ln(K_{t+1}) = I\{z = g\}(a_0 + a_1 ln(K_t)) + I\{z = b\}(b_0 + b_1 ln(K_t))$$

In this framework there are two sectors, and aggregate capital is invested in the two sectors. The equilibrium interest rate is given by the marginal product of capital invested in the corporate sector only, not by the marginal product of total capital. As stated above, there is a no arbitrage relationship between users of capital - all corporate and entrepreneurial capital renters accept the same market rate. Notice that to split capital between entrepreneurs and corporates today agents need to know, at a minimum, prices today, (meaning the interest rate and the wage) which are endogenous equilibrium objects that depend on people's occupational decisions. To approach this in the K-S limited rationality methodology, recall the equilibrium interest rate and wage for period t is given by the corresponding marginal products in the corporate sector, $r_t = MPK_{c_t}$ and $w_t = MPL_{c_t}$, and that w_t is a function of r_t and quantities affected by the aggregate state z_t, as per section 3.2.2. As agents know z_t, this leaves r_t as the only unknown, a function of the (unknown) corporate K-L ratio. For the household to make savings and entrepreneurial decisions, they must forecast the interest rate, the wage and the future capital stock. Thus, if they use an accurate polynomial approximation to r_t and K_{t+1}, this is sufficient for equilibrium, following Krusell and Smith's equity pricing paper (Krusell and Smith Jr. [1997]). To find the equilibrium, these forecasts can be compared to the interest rate and capital stock resulting from household's and firm's behaviour based on the forecasts, and when the two are sufficiently similar, this is a numerical equilibrium.

The numerical algorithm for solving the entrepreneurial problem, performing value function iteration, simulating and updating forecast rules has to be extremely robust. This model has a number of non linearities - occupational choice, hiring decisions and ability transitions - that make it difficult to solve (quickly).

3.3 Shock Process, Data and Calibration

3.3.1 Shock Process

The shock process is a distributional shock, affecting the idiosyncratic transitions of entrepreneurs, changing Π_θ as well as traditional level shocks to productivity $\theta(z)$ and

$\theta^c(z)$. There are two usual states of z, normal turbulence and high turbulence, which correspond to non-recession and recessions. There is an exogenous transfer between these states, for which agents know probabilities, just as they know the structure of Π_θ.

Entrepreneurial ability, θ is distributed across the population with most having ability zero and a constant 'entry' probability into a positive process for θ_i. For entrepreneurs with positive ability, each period they receive a shock η_i to their ability,

$$log(\theta_{i,t+1}) = log(\theta_{i,t}) + \eta_{i,t}$$

This forms a process similar to the entrepreneurial 'ladder' ability process used by Quadrini [2000] and Kitao [2008], and other processes used by Luttmer [2010] and Luttmer [2007]. To make the process stationary and usable, the distribution of θ is bounded at the top and bottom, with the bounds part of the calibration. The lower bound is an absorbing boundary whilst the upper bound is reflective. Whilst these assumptions can be changed, it seems unreasonable that firms would 'bounce' out of low productivity with a very high probability in turbulent recessionary times, and having a reflective upper boundary ensures there is a tail distribution of productivity rather than a mass at the upper limit.[3] I discretise the process by simulating the bounded unit root process to recover transition probabilities between equally-spaced θ states.

The exit probabilities for entrepreneurial ability (i.e. transitioning from $\theta > 0$ to $\theta = 0$) are calibrated to match business survival rates, as in the model of Kitao [2008] with multiple entrepreneurial ability levels and similar to Sedlacek and Sterk [2014]. The exit function is,

$$P(\theta = 0|\theta_i) = \epsilon_1 e^{-\epsilon_2 \theta_i}$$

Higher ability entrepreneurs face a lower exit probability, but these probabilities do not vary with the aggregate state - exit only endogenously increases from events shifting greater numbers of previously high ability entrepreneurs into lower ability categories. Thus the lowest state intuitively captures a tail of relatively unproductive firms with high exit likelihood, and the increase in probability of transitioning there is a part of the turbulence mechanism.

Being as general as possible, the aggregate shock can change θ levels or change the distribution of η. Raising the variance σ^2_η would be akin to Bloom-style dispersion shocks and changing the mean μ_η would be very similar to reducing the state vector θ with the exception that agents moving downwards may encounter different exit probabilities. To accommodate the data and/or findings of Bloom et al. [2016] that dispersion increases are mostly due to skew changes from expansion in the lower tail of productivity shocks, it is

[3] Regardless, experiments with different bounds do not change the conclusions.

possible to change the distribution for η in a variety of ways. To remain simple, I focus on increasing the variance of the shock η, which creates a greater amount of turbulence and movement across the ability spectrum. I also show changes to the θ vector as a simple comparison to canonical TFP level shocks.

The above defines the aggregate-state-dependent transition matrices for entrepreneurs $\Pi_\theta(z)$ and ability vector $\theta(z)$.

I note in advance of the results that alternative characterisations of the process provide similar results as long as there is the important feature that there are multiple entrepreneurial productivity levels and increased transition probabilities between levels under the turbulent state.

3.3.2 Data

There are two main sources of data used for calibration of the model - the U.K. Wealth and Assets Survey ('WAS') and Business Structure Database ('BSD'). I choose to base the model upon the U.K. due to favourable qualities of these datasets. The WAS is used here for moments regarding entrepreneurial transitions, the distribution of wealth and the wealth of entrepreneurs. Importantly, the WAS shows greater mobility amongst the wealthy during the 07-09 transition and an increase in entrepreneurial exit and losses of business wealth.

	07-09	09-11	11-13	13-15
% staying in wealthiest 1%	0.54	0.64	0.67	0.71
% entrepreneurs in wealthiest 1%	0.35	0.41	0.43	0.38
% leaving entrepreneurship	0.48	0.35	0.34	0.39

Table 3.1: Wealth and Assets Survey entrepreneurial statistics by years.

The definition of "entrepreneur" used for the data is those that are self-employed and say that they are one of 'sole director', 'director' or 'partner', or are 'self-employed in another way' and state a business value of over £1000. This excludes a large group of self-employed who do not own a business and earn relatively little.

The BSD contains employment and turnover information for all firms in the UK for 1997-2015 and provides the link between entrepreneurial household features and the outcomes for those entrepreneurial enterprises. The moments I use are sourced from my short note on the database, Pugh [2018a]. There are approximately 2 million active businesses in a given year and these businesses can be tracked over time. This dataset is different to many used in the macro literature as it includes and tracks small firms as well as large and is at enterprise level, rather than establishment level. An enterprise is much closer to the concept of an entrepreneurial firm than an establishment. There is also some less complete data at the 'enterprise group' level in the BSD, but the vast majority of enterprises are the sole member of their enterprise group. The main findings are that during a large recession businesses face a wider and more negative distribution of outcomes than in non-recessionary times, in addition to increases in business exits and falls in entry.

As shown in Figure 3.1, births decrease whilst exits rise during the recessionary period around 2008 before births recover to their former level while deaths remain elevated. The rise in entry in 2002-2005 may be partly due to a change in tax incentives encouraging the self-employed to register as business owners.

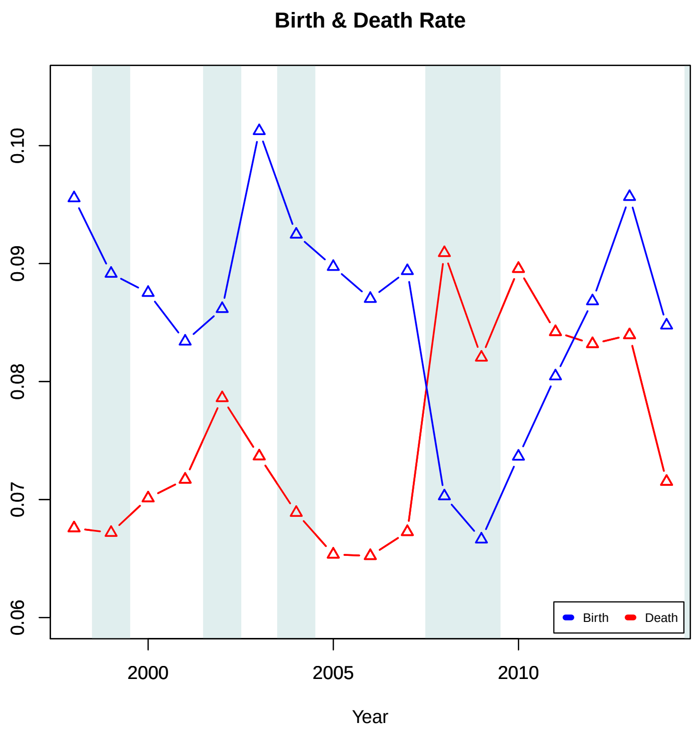

Figure 3.1: Birth and death rates for enterprises. Years with more than 6 months of OECD recession indicator shaded.

Figure 3.2 depicts the moments of the distribution for equal weighted percentage changes in turnover[4]. Equal weighted changes are well-suited due to the high number of zero values and extreme movements which would make standard percentage changes or log changes difficult to use. These show a clear change over the 2009 crisis and other recessions. The mean falls and variance rises though, surprisingly, skewness increases whilst kurtosis falls. The rise in skewness is not matched by an examination of the quantiles of the distribution. There, in Figure 3.3 the lower tail of outcomes falls quite significantly, more so than the middle and top of the distribution. In words, the distribution of poor turnover outcomes faced by businesses is getting more negative and more extreme around recessionary periods, which are shaded in Figure 3.2.

[4]Defined as 2*(New-Old)/(New+Old), putting equal weight on both elements of the change in the denominator.

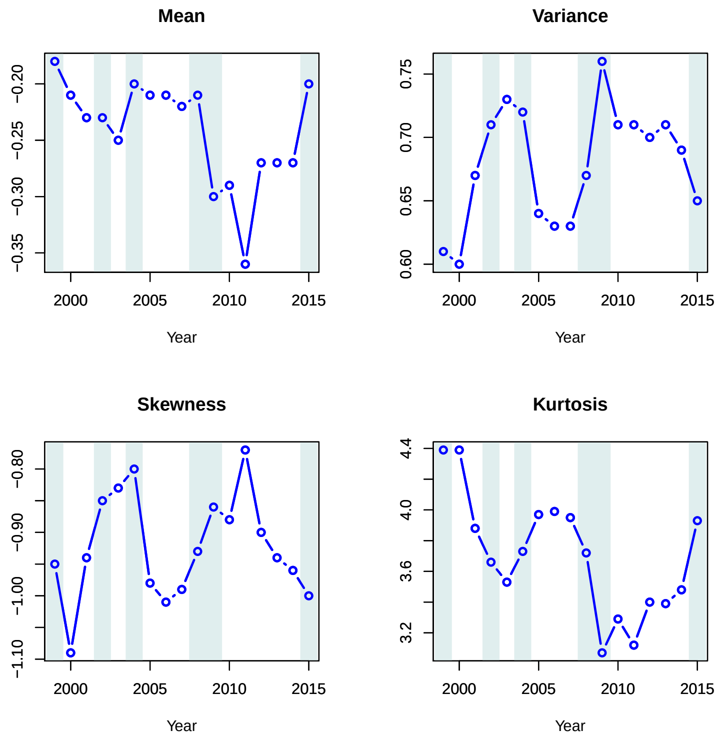

Figure 3.2: Moments for equal-weighted % change in turnover by year. Years with more than 6 months of OECD recession indicator shaded.

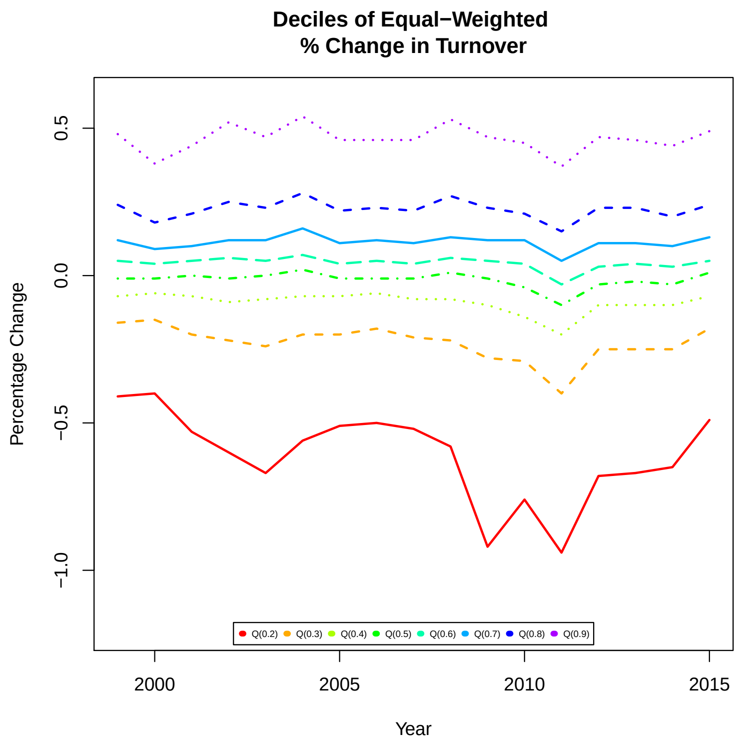

Figure 3.3: Deciles of equal-weighted % change in turnover by year, period including Great Recession.

The turnover figures are preferable for model calibration compared to employment moments as small firms or lone self-employed entrepreneurs have a very discrete level of employment and thus large and lumpy changes if and when they change their hiring. Whilst the distribution of employment change quantiles shows a similar pattern to turnover, it is harder to interpret due to a strong presence of lumpy changes from the very many small firms [5]. The model is not designed for representing this discrete hiring behaviour accurately, as employment beyond an entrepreneur's own labour is continuous.

3.3.3 Calibration

Income taxes are characterised by the functional form in Gouveia and Strauss [1994], whilst estate taxes use the format and calibration strategy of Cagetti and Nardi [2004], which studies the impact of estate tax changes upon entrepreneurs and the wealth distribution. Pensions p are paid at the ratio of UK state pensions to mean earnings, similar to the ratio from Kotlikoff et al. [1999] in Cagetti and Nardi [2006].

The worker ability process y and Π_y is a discretised AR1 in logs, calibrated to the earnings Gini and the Shorrocks Index in the WAS data and UK earnings administrative data examined in De Nardi et al. [2018]. The simplicity of the worker process is due to the focus on entrepreneurial dynamics and choices. De Nardi et al. [2016] use very rich worker earnings dynamics (which exclude entrepreneurs) and find a better match to lower tail inequality and savings behaviour of the poor but they find little improvement of fit for upper wealth quantiles, where entrepreneurial dynamics and our targets are mainly located.

The parameters are calibrated to a stationary model with permanently low turbulence / non-recessionary state, with the target ranges from the data and output from the model shown below in Table 3.3, and parameters and values shown in Table 3.2. The parameters use the same notation as in the model exposition[6] in Section 3.2.

Parameter definition	label	best fit values
Discount factor	β	0.932
Borrowing limit	λ	0.35
Entr. capital share	α_e	0.56
Entr. DRS	ν	0.74
Entr. ability upper limit	θ_h	3.25
Entr. ability lower limit	θ_l	1.35
Entr. ability entry	$P(\theta_1\|\theta_0)$	0.012
Entr. ability exit level	ϵ_1	1.6
Entr. ability exit curvature	ϵ_2	1.25
Entr. ability variance	σ_η^2	0.22

Table 3.2: Parameters

The targets and their ranges are taken from the WAS and the BSD. These targets are similar to those used in Cagetti and Nardi [2006] and related papers. As the WAS is

[5] The UK firm size distribution is power-law distributed, much like other firm size distributions as noted by Luttmer [2007], Luttmer [2010] and Gabaix [2009] amongst others

[6] 'Entr.' refers to 'Entrepreneur'.

biennial, the targets for entrepreneurial household transitions are also biennial and two year transitions are calculated from the model to match the data. The model is itself annual, like the BSD data. The overall match to targets is good, and the model captures the data (stylised) facts regarding the position of entrepreneurs, mobility and the wealth distribution in the UK. Unlike previous work, the use of the WAS in calibration means both the facts about the top of the wealth distribution and the longitudinal facts come from the same database (and those not used in the calibration can become testable predictions).

Most of the target definitions are self-explanatory, except for "IQR of turnover changes" and "Median Entr : Median Worker wealth", which refer respectively to the inter-quartile range of equal-weighted percentage changes in turnover from the BSD and the ratio of median wealth of entrepreneurs versus median wealth of workers.

Target definition	data range	model value
% entrepreneurs	3.6-5.5	4.8
% entrepreneurs exit in 2 years	37-48	40
% entrepreneurs enter in 2 years	1.8-2.3	2
% entrepreneurs do not hire	41-59	48
Median Entr : Median Worker wealth	2.9-3.5	3.3
Top 1% wealth share	16-21%	23
% entrepreneur in Top 1%	34-39	37
Entrepreneur wealth share %	17-20	18.7
Capital:Income ratio (K:Y)	2.4-2.6	2.45
IQR turnover changes	0.32-0.34	0.33

Table 3.3: Targets

Although each parameter does not have a unique link to a target, there are some clear economic intuitions. The borrowing limit λ raises entrepreneurial borrowing limits, inequality and entrepreneurial income, though it also reduces pressure on poorer entrepreneurs to gather assets and overcome their constraints. The capital share α_e and DRS ν of the entrepreneurial production function are strongly linked to the proportion of entrepreneurs in the top 1%, proportion of entrepreneurs hiring and wealth of entrepreneurs as they provide incentives for entrepreneurs to gather capital (rather than use labour). The K:Y ratio constrains the amount of capital accumulated by entrepreneurs and is strongly related to the discount factor β. The limits for the entrepreneurial ability process θ_l and θ_h affect inequality very strongly and are constrained by targets for the wealth of entrepreneurs. The entry and exit of entrepreneurs is closely determined by the entry and exit process for ability, as would be expected. Variance of the ability shock is almost solely determined by the match to the inter-quartile range of changes in log turnover from the BSD.

There is tension between the calibration having enough entrepreneurial wealth (% wealth held by entrepreneurs) whilst maintaining low enough inequality (top 1% share of wealth), sufficient non-hiring and reasonable median wealth ratios. The more wealth entrepreneurs have, the more significant they are to the economy. A Low DRS parameter ν and a big spread of entrepreneurial ability $\theta_h - \theta_l$ generates a high proportion of entrepreneurs in the top 1%, the top 1% share and matches median entrepreneur:worker wealth ratio. The entry and exit parameters determine the entrepreneurial population, and lower exit or higher entry does raise inequality/entrepreneurial wealth targets as would be expected. Entrepreneurs are generally rich, and become richer on average when they have longer expected tenure with a positive entrepreneurial ability and the resultant high income.

3.4 Aggregate Equilibria and Krusell-Smith Aggregation

I will explore two sets of shocks. Firstly, a 'standard' reduction in TFP during a recession (much like Krusell and Smith Jr. [1998] without changes in earnings transitions). Secondly, an increase in variance of entrepreneurial shocks as noted in the turnover data (and following the literature on dispersion increases during recessions). My objective is to study the impact of credit-constrained entrepreneurs.

The aggregate state is a two-state process (normal, recession) with transition probabilities based upon the average lengths of recessions from UK aggregate data, taken from the Federal Reserve bank of St Louis FRED database. When the recession is characterised as a mean reduction in TFP, it is calibrated by the difference between the Bank of England's TFP estimates in the two states, which is -0.8%[7]. Turbulence shocks are instead calibrated using the increase in inter-quartile range from Figure 3.3 that occurs in 2009-11 versus the earlier period, which is 0.54 as opposed to 0.36. When this is calibrated to the model, η has a standard deviation of 0.22 and 0.36 in the normal and recessionary states respectively.

This model does not follow the Krusell-Smith (K-S) mean-only aggregation result, whereby the economy and its required forecasts can be predicted with great accuracy using only the current aggregate state and mean level of capital. Usually, models with aggregate shocks and heterogeneity rely on a simple forecast rule of $K' = \alpha + \beta K$ conditional on z. Here, the forecast rules need to use the lagged shock value and lagged interest rate for prediction and forecasting dynamics. This indicates it 'matters' to forecasts whether the economy is in a high or a low turbulence state previously and this information is not fully transmitted through the total capital stock. One can infer that the order of shocks and their combination influences dynamics significantly. The reasons behind this become clear when studying the Impulse Response Functions (IRFs) in the next section - each aggregate shock changes the allocation of capital amongst individuals who may have a lot of capital, which influences aggregate dynamics when entrepreneurs are involved.

[7] Bank of England, Total Factor Productivity Growth in the United Kingdom [TFPGUKA], retrieved from FRED, Federal Reserve Bank of St. Louis; https://fred.stlouisfed.org/series/TFPGUKA.

Table 3.4: Equilibrium Accuracy tests (MAE in %)

Model	$R^2_{K'}$	$MAE_{K'}$	R^2_r	MAE_r
Level z shocks	0.9993	0.60	0.9988	0.11
Turbulence shocks	0.987	0.08	0.981	0.27
K-S equity model	0.99999	-	0.99999	-

Throughout, the two accuracy measures used to evaluate the accuracy of the equilibrium are the simulation R^2 and a version of the test from Haan [2010], which uses the maximum absolute error (MAE) over a simulation period, both shown in Table 3.4. Den Haan's test is significantly stricter and detects numerical aggregation errors that the R^2 measure may not. It has a simple intuition which is easy to compare over models - in every simulation used as part of convergence iterations or in the test, this is the maximum deviation between a sequence simulated using the forecast rules alone and a separate sequence from the full model using the same shocks.

Once extra predictors are included, the model with TFP-mean-only shocks has good forecast rules according to the R^2 measure and the maximum absolute error test of Haan [2010] to evaluate. In the turbulence shock model, there is a lower R^2 but not large maximum absolute deviations over a simulation.

The source for the lack of aggregation is the difficulty of inferring entrepreneurial behaviour from aggregates - entrepreneurs are very much affected by their individual heterogeneity and personal history and, particularly under turbulence shocks, this affects the economy.

With level shocks alone, the model performs better on these tests, with high R^2 similar to K-S models that follow the mean-only aggregation result and maximum absolute errors of less than 1% over the entire simulation (10000 periods). Turbulence shocks have lower R^2 and, as well as the greater impact of entrepreneurial heterogeneity, this is due to the compensation of the corporate sector in response to shocks to the entrepreneurial sector. This results in much smaller K and r variation and thus the R^2 is lower but the maximum absolute error (the stricter test in terms of deviations from forecasts) is still very low.

3.5 Aggregate and Inequality Dynamics and Mechanisms - IRFs and TFP

To explain the model's shock dynamics, I now present impulse response functions (IRFs). These are the mean response of a variable to a change in state for normal to recessionary conditions. There are two things to note - firstly, the extensive response of the entrepreneurial sector to a shock, which is longer than a non-entrepreneurial model without constraints (which is similar to the response of the unconstrained corporate sector). Secondly, by contrast to a more usual mean-changing shock, I note the novel effects of the turbulence shock with entrepreneurs, particularly in persistence of capital changes and the strong sectoral differences. Mean-only shocks affecting all entrepreneurs and corporates equally also show the amplifying power of the entrepreneurial sector.

3.5.1 Effects of level shocks

On the aggregate level, the response of the economy to a downwards TFP shock affecting both corporates and entrepreneurs in figure 3.4 looks very much like a response to a level shock to economy-wide TFP in many standard models. The economy begins in state 1 (the normal state) in period 1 at the start of the graph and moves to state 2 in period 2. This impulse response function is formed from an average of simulations. Since the economy begins in state 1, which has higher TFP, these variables begin above the average levels for a simulation. Output, investment and consumption all fall following the shock, as does the capital stock. The responses follow the well-recognised relation that investment is more volatile than output and consumption less (and by reasonable magnitudes). Then, the economy, having fallen below the average level recovers towards the simulated mean. After 20 years, the output and investment level are at the mean, but the capital and consumption levels are still below (though they do return to the mean within 50 years).

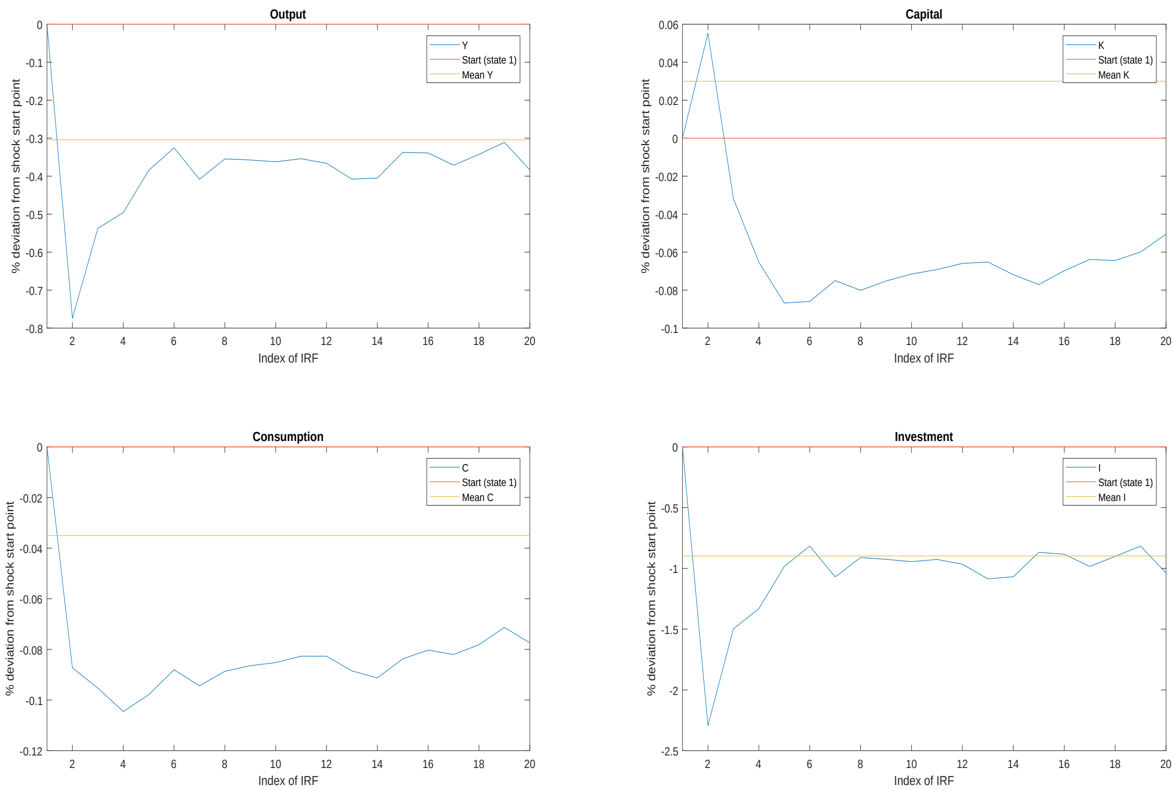

Figure 3.4: Average IRFs for aggregate measures, shock from normal to recession state in period 1 to 2.

If one examines the breakdown of responses into corporate and entrepreneurial sectors in figure 3.5, entrepreneurial production and resource usage (both labour and capital) fall much more dramatically than the corporate equivalents or the overall economic response. There is an amplifying effect in the entrepreneurial output reduction versus the shock size, as the entrepreneurial output has an initial fall of -0.97%, an extra 0.17% versus the shock amount and therefore over 20% larger than the size of the original shock.

Entrepreneurial capital has a very large response to the shock relative to the aggregate response - approximately three times larger at the fullest extent of both responses, 4 years after the shock (in period 5). The corporate capital stock initially rises, as there is a substitution of capital from the entrepreneurial to the corporate sector. This model has no frictions on the movement of capital between different sectors in the lending market, thus the corporate sector dampens the overall effect on the capital stock, as well as providing lenders the ability to move their capital away from the entrepreneurial sector easily.

It takes a particularly long time for the entrepreneurial capital stock to return to the average level, given the severity of the fall. In labour usage, the narrative is similar, with the exception that the overall stock of labour does not change in this model, thus there is only a substitution effect. The entrepreneurial sector employs significantly less labour, which is absorbed by the representative corporate firm. The slow entrepreneurial capital recovery and large effects upon entrepreneurs speaks to an inverse version of the 'financial accelerator' mechanism in Bernanke et al. [1999], Kiyotaki and Moore [1997] and Bernanke and Gertler [1989] - entrepreneurs need to save (invest) their own assets to build their businesses and this process is slowed by reduced productivity, reduced borrowing and reduced profits to invest, generating a propagation mechanism. In this case, a compensating corporate sector dampens the aggregate response.

In this framework, there is a non-trivial wealth distribution which supports the explanation above and can lend insight into the mechanisms at work. Figure 3.6 shows the wealth changes experienced by different groups in the wealth distribution. The wealthier subsets experience an initially softer fall in wealth (as a group), because of their increased ability to self-insure against shocks. But, seven years after the impact, the proportional wealth decreases for the wealthier begin to exceed groups including less wealthy members such that the maximum negative impact on wealth of the top 1% is after 15 years, versus 6 years for the top 50%. Further, this negative impact is larger and takes longer to resolve for the richer. This aligns with the larger proportion of entrepreneurs in the very top wealth groups and the longer, deeper impact of the shock upon entrepreneurial capital and production.

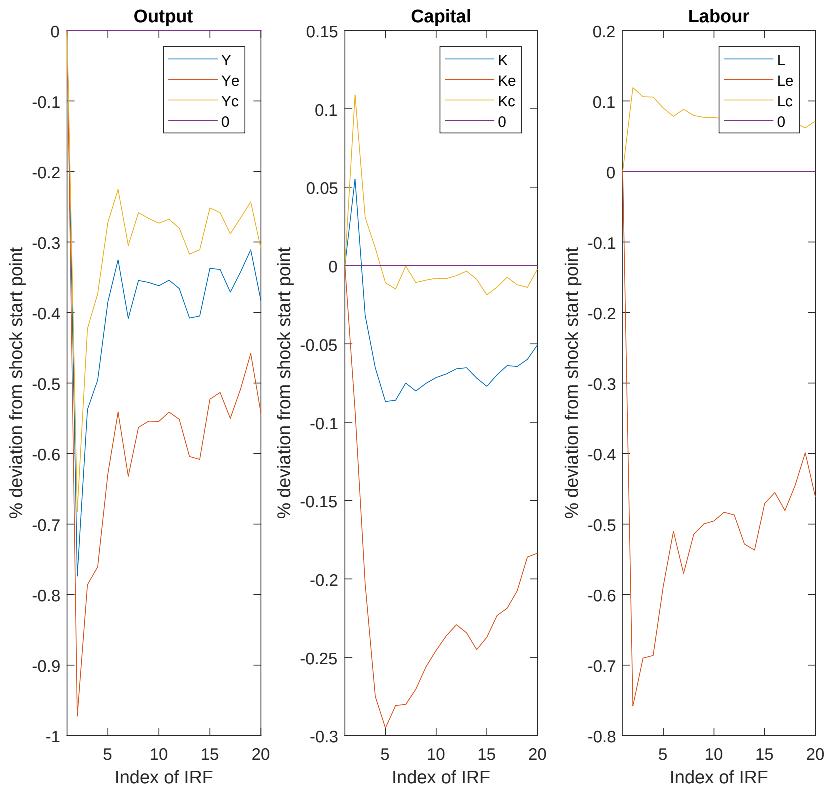

Figure 3.5: IRFs for entrepreneurial and corporate sectoral quantities, shock from normal to recession state in period 1 to 2.

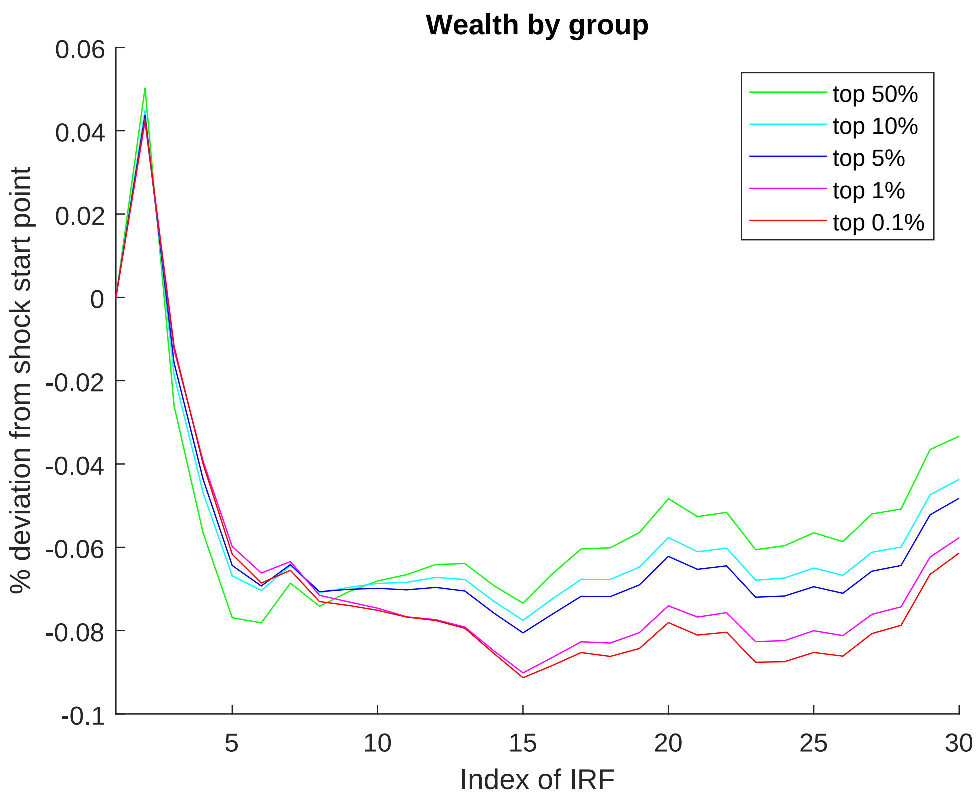

Figure 3.6: IRF for the wealth held by different groups, level shocks.

3.5.2 Effects of turbulence shocks

Examining turbulence shocks, we also see large changes in the entrepreneurial sector, which are also significantly compensated by the corporate sector. Note that the levels of the productivity states are unchanged as the turbulence only affects entrepreneurial productivity transition probabilities. In the model, the transition probabilities are realised at the end of the period, so there is an initial rise in saving due to precautionary motives from the full knowledge of a turbulent state affecting the end of the first period. This causes an initial rise in investment, output and capital, and a negative response of consumption. Quickly, this is replaced with an inversion of these changes, with investment, output and capital all falling. Consumption remains elevated for some time, from a combination of a reduced interest rate and smoothing behaviour.

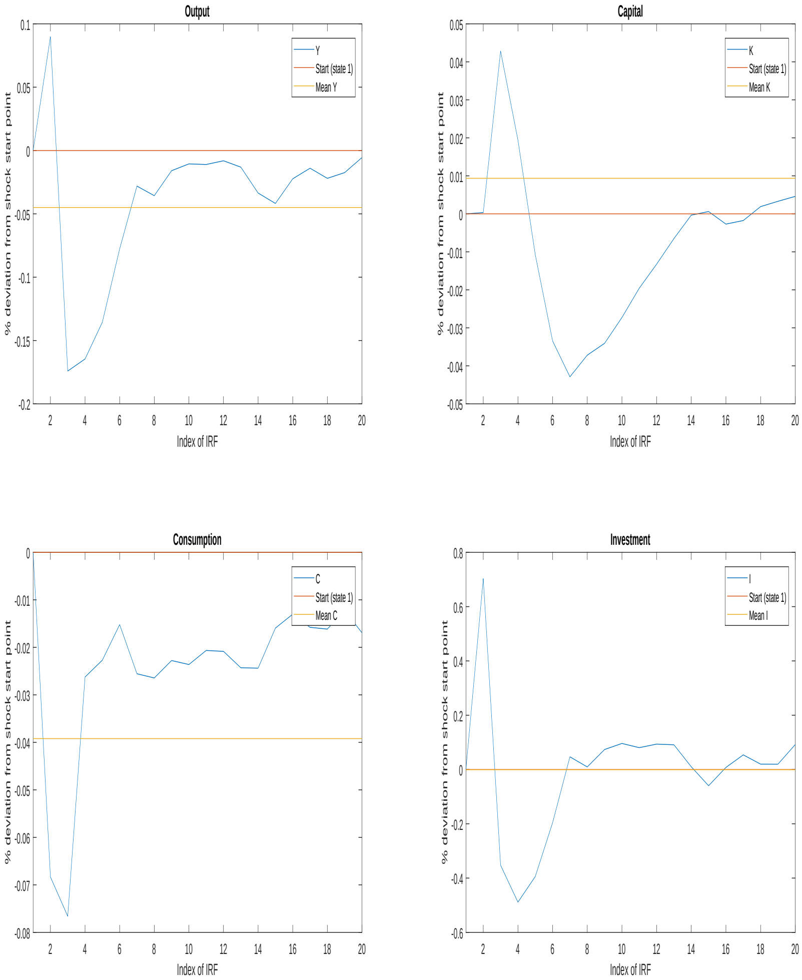

Figure 3.7: Average IRFs for aggregate measures, shock from normal to recession state in period 1 to 2.

Moving on from the aggregate dynamics, once the shocks are realised in the second period of the simulation, there is a large and rapid drop in entrepreneurial capital and output. This change is very much larger than that of the TFP-style shocks discussed earlier, with a fall over 1.5% in output and over 2% in capital. The corporate sector's ability to compensate for the changes explains the comparatively muted aggregate response from the economy, though there is an aggregate drop, as discussed above.

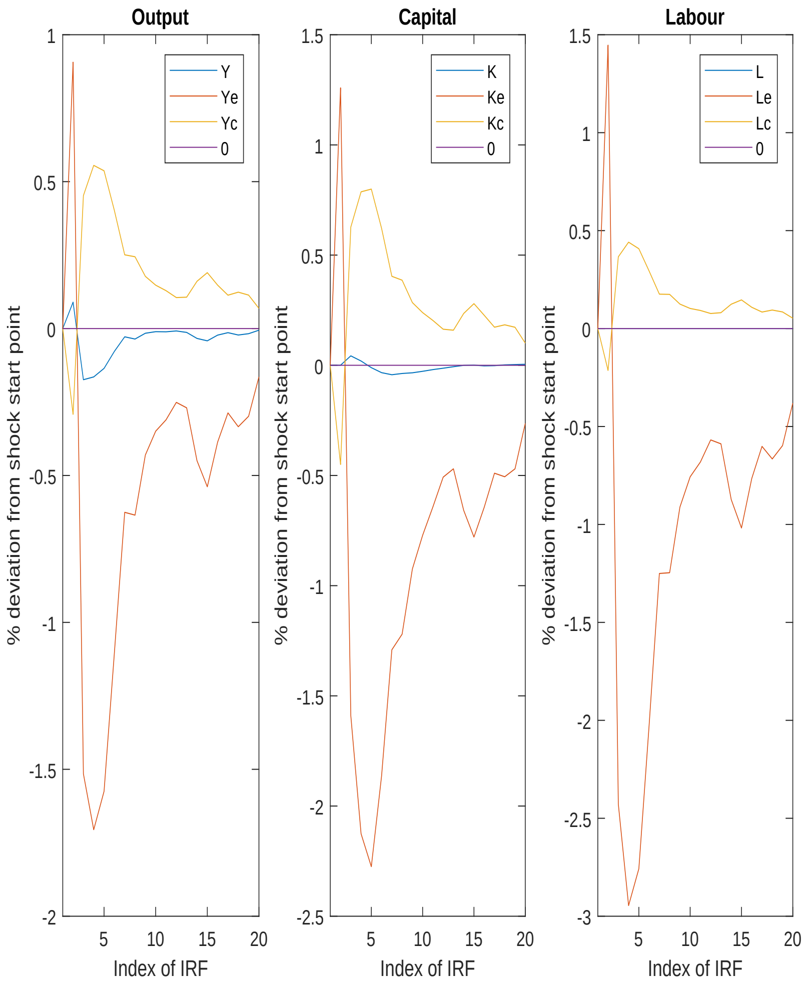

Figure 3.8: IRFs for entrepreneurial and corporate sectoral quantities, shock from normal to recession state in period 1 to 2.

After the initial effects driven by precautionary motives, previously productive and relatively unconstrained entrepreneurs fall down the ability ladder, have tighter borrowing constraints (here, the more complex borrowing constraint is important - a simple multiple of assets rule would weaken this mechanism) and so are less able to produce and have lower input demand. Their capacity to produce is not replaced by new/smaller entrepreneurs due to the stronger impact of these borrowing constraints, and the corporate sector is a less efficient user of this newly 'leftover' capital. There is thus a shift of resources into the corporate sector due to the financial frictions. Through the introduction of new entrepreneurs and low-turbulence productivity transitions, the economy fully returns to its average state over approximately 30 years.

Even though the shock is symmetric, and despite the greater density of entrepreneurs at the lowest (entry) level of ability meaning more entrepreneurs will be able to move up versus down, the downward effect dominates over the medium term 1-5 years after the shock, and longer for capital.

This mechanism fits well with that examined in Moll [2014] concerning self-financing and borrowing constraints. In that work, which is far more theoretically focused than this quantitative exercise, increasing the persistence of entrepreneurial ability increases the efficiency of the allocation of capital, as more able entrepreneurs build wealth faster and thus acquire more assets, improving aggregate productivity.

Here, the same effect applies - when turbulence is reduced, entrepreneurial ability is more persistent and output is indeed higher, whereas an increase in turbulence increases the relative misallocation of capital and strengthens the role of borrowing constraints. This is especially stark if one considers a rich entrepreneur at the top of the ability scale, who falls to the bottom and exits through the high attrition of entrepreneurs with low ability. Now, their entire capital stock is passively invested in the general capital market, earning r and only given to constrained entrepreneurs who can borrow it or the corporate sector.

Relatively, this usage is less efficient than a highly productive entrepreneur using the capital for their own production and to collateralise their own borrowing (the average entrepreneur earning marginal returns at around 25%). In essence, financing constraints ensure inefficiency by preventing the most productive entrepreneurs acquiring capital rapidly and turbulence limits the alternative of self-financing by building a personal asset base.

This corresponds to the effects upon inequality observed in figure 3.9. The wealth of the richest shrinks by more than those lower in the wealth distribution, reducing inequality under turbulence. The richest groups have a much greater density of entrepreneurs, and thus are more affected by the turbulence, as well as having a greater income from capital if not entrepreneurs, thus suffering more from the (small) interest rate reduction accompanying the turbulence increase. The greater initial precautionary saving and much slower impact on the wealth of the richest, seen particularly for the top 1% groups and above, is very likely to be a result of the slow recovery in entrepreneurial capital and output described above. These differentials are quite large, as the top 50% has a maximum reduction in assets of significantly less than 0.1% seven years later whilst the top 0.1% has a reduction approximately ten times larger, at 0.85%, twenty years later. The large assets of the top 1 and 0.1% (who hold, respectively, 20% and 5% of aggregate wealth) indicates that the absolute amounts of capital are still large, and are especially large compared to the effects upon the wealth distribution observed in figure 3.6 with TFP-level shocks.

Comparing between the different shock processes, we can evaluate the relative effects on the length of the aggregate output recovery in figure 3.10. The diagram plots the relative recovery of output from the IRF's trough to the simulation mean, scaled proportionately from 0 (trough) to 1 (mean). As a benchmark, a version of the model with a zero density of entrepreneurs is included in the figure, with the same size TFP shocks as in Section 3.5.1. The inclusion of entrepreneurs in the model, whilst keeping the same size shocks (and same sequences to generate the IRF), there is a reduction in the rate of recovery from

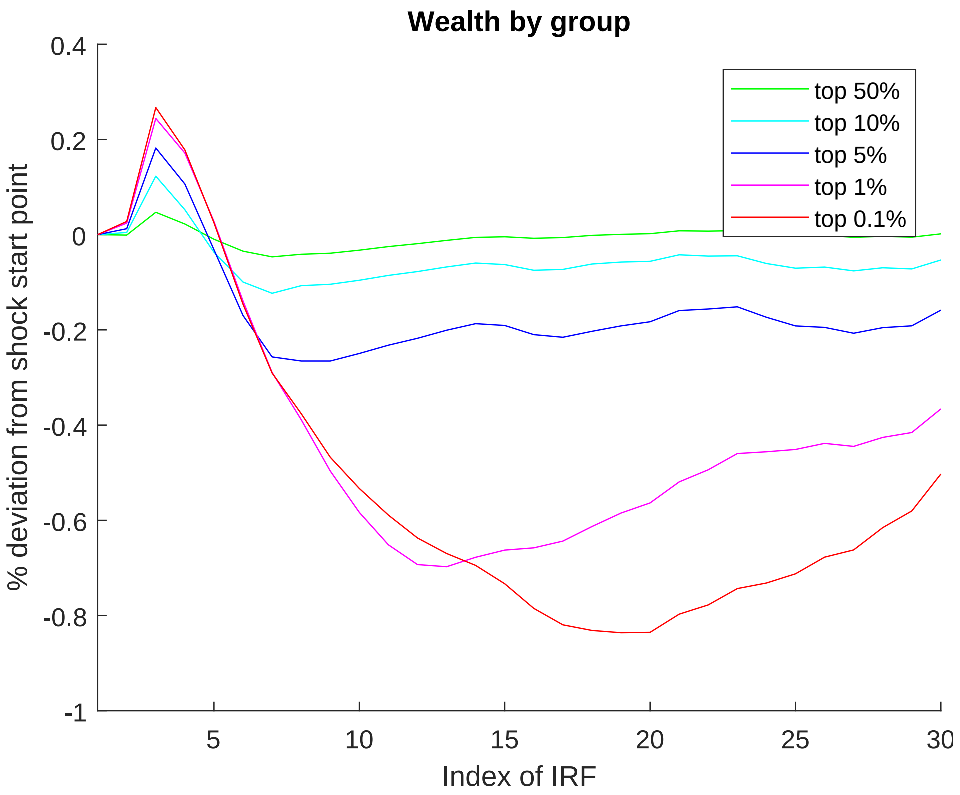

Figure 3.9: IRF for the wealth held by different groups, turbulence shocks.

the trough point, with the benchmark economy achieving a 'full' recovery to the mean output in 4.5 years, as opposed to 4.8 when including entrepreneurs. As noted above, this ignores the very large difference between sectors and does not account for not capturing the cross-sectional wealth distribution when not including entrepreneurs.

Turbulence shocks are substantially different. The full recovery occurs at approximately the same time as the other models, but the initial response is much slower, before gathering higher speed and overshooting the mean output. The speed-up and eventual outperformance is due to the symmetric upside of the turbulence shock. Over a long enough time, the greater density of entrepreneurs at higher entrepreneurial abilities gather increasing resources and expand their constraints, creating output growth which can gather pace as they grow, before the eventual stochastic return to the average entrepreneurial distribution.

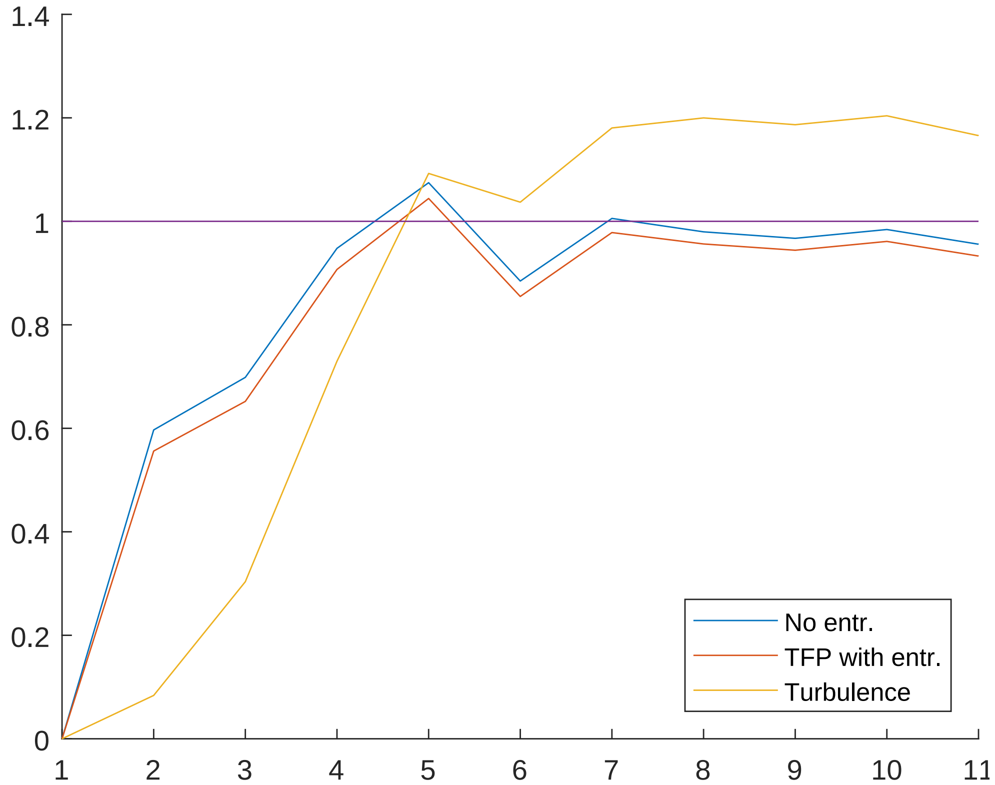

Figure 3.10: Rescaled Output IRF, recovery of output from trough of different shocks and models. Trough is 0, average simulated output is 1, representing a full recovery.

3.6 Conclusion and Developments

I have presented a model for understanding the role of aggregate shocks and entrepreneurs, small businesses and personal wealth fluctuations. This model demonstrates that fluctuations that are within the bounds of firm-level data have large and highly persistent effects on the aggregate economy through entrepreneurial households. The existence of entrepreneurial constraints and turbulence together creates capital misallocation and transmits micro-level changes into aggregate fluctuations - as claimed at the very start, 'who has wealth matters'. I consider the data on entrepreneurship and the dynamics of enterprises in the UK and offer a quantitatively evaluated heterogeneous version of the 'financial accelerator' channel, showing that non-directional dispersion in entrepreneurial productivity can have large and persistent effects, as well as entrepreneurial behaviour's amplifying effects in responses to common TFP-style shock processes.

The entrepreneurial modelling provides rich insights and I find that the effects of shocks upon the entrepreneurial production sector are much larger than aggregate effects due to the compensating role of the unconstrained corporate sector in this model. The role of frictions between the two sectors or within the corporate sector limiting the dampening effect and exacerbating fluctuations or potentially providing different responses is of substantial interest to those interested in the performance of the economy and business constraints. In the UK, entrepreneurs are only a small part of the population and not as rich as their

U.S. equivalents, so we would expect larger effects if calibrated to the U.S.A.

This model can be developed further. One key dimension is to include features such as personal loss of capital from business collapse, negative income or 'disaster shocks'. The current framework does not create very large changes in the distribution of wealth between groups, without very large or persistent turbulence shocks, likely because it does not include risk and the destruction of capital when an entrepreneur/owner closes or loses a firm. Extending the model of entrepreneurial businesses to include more complex borrowing, ownership and choice of investment in the firm is something that would likely yield interesting results, but richer entrepreneurial/firm data would be needed. Similarly, the turbulence itself could be represented with other parameterisations accounting for different moments of the distribution of firm dynamics and household mobility.

I also discuss inequality, showing how calibrated recessionary turbulence in entrepreneurial ability can reduce inequality in the short run. The turbulent entrepreneurship mechanism has the possibility of explaining longer term transitions and developments in inequality than the business cycle environment depicted here. For example, the decreasing downwards mobility amongst top income groups seen in Guvenen et al. [2014a] since the 1980's and increasing inequality align with the results from this work. Further, the process for workers in this model is very simple, and turbulence-style earnings and capital income shocks affecting non-entrepreneurs and the poor may also be important to explanations for changing inequality and mobility in earnings, income and wealth.

In sum, this chapter demonstrates the importance of capital-constrained entrepreneurs in heterogeneous models not just for replicating the cross-sectional distribution of wealth or public finance questions but also as a transmission mechanism for shocks to the economy.

Chapter 4

Conclusions

This thesis investigates the wealth distribution in the UK and the related role of entrepreneurship in the economy. Each chapter has its own conclusions and suggestions for development, but there is one constant thread - understanding the mobility and behaviour of the wealthy is important to understanding the economy. Whilst there is (rightly) great focus on poverty in studies of inequality, the large holdings of the wealthy and their involvement in entrepreneurship and investment make them an important consideration for public policy and academic enquiry. Whilst not tackling the longer term trends of rising wealth and income of the very wealthiest groups versus the remainder of the population, this thesis demonstrates top wealth inequality and mobility should be a matter of interest for both those thinking about about equity, those optimising welfare and those considering the stability of the economy.

I find that wealth is not stable, with large movements amongst those at the very top. Business wealth and entrepreneurship is a key element in these movements, especially the largest changes. The importance of heterogeneity in returns gives a clear message to the literature that variation in earnings alone is not driving the patterns in mobility, and that individual exposure to different returns on wealth is key. The importance of downwards movements in wealth is a demonstration that destruction and loss of wealth is an important mechanism affecting the wealthy, which deserves further investigation.

Chapter one provides a careful examination of wealth data in the UK, where I conclude that the WAS dataset is a good representation of the wealthy and provides interesting and important facts about the cross-sectional distribution of wealth and novel implications about the mobility of the wealthiest in the UK. Chapter two builds upon this work and implements structural estimation of different theories of wealth inequality, and discerns that returns heterogeneity is important to match both the static empirical inequality present and the high mobility amongst the wealthy. The estimations provide evidence against theories of wealth accumulation based on extraordinary earnings or different preferences, due to their reliance on immobility to generate inequality. Chapter three develops the facts about entrepreneurship from chapter one and combines with data on UK firms to work with a macroeconomic model of entrepreneurship, aggregate shocks and the wealth distribution. It finds that constrained entrepreneurs exacerbate the effect of TFP shocks often used to drive fluctuations in the Real Business Cycle literature, and are thus important to understanding business cycle variability. I also use the business data for 'turbulence shocks', which increase the variability and uncertainty of entrepreneurial productivity. These shocks have interesting effects, and suggest that this turbulence can contribute to explanations of changing wealth inequality and medium-term (5 years-20 years) economic fluctuations.

I look forward to developing the themes in this work further, to the impact that this work has and am pleased at any assistance or insight it provides to researchers or other interested parties in the future.

Bibliography

Yves Achdou, Jean-Michel Lasry, Pierre-Louis Lions, and Benjamin Moll. Wealth distribution and the business cycle: The role of private firms. *mimeo*, July 2014.

S. Rao Aiyagari. Uninsured idiosyncratic risk and aggregate saving. *Quarterly Journal of Economics*, 109(3):659–684, August 1994.

Facundo Alvaredo, Salvatore Morelli, and Anthony B. Atkinson. Top wealth shares in the uk over more than a century. Working Papers 2017:01, Department of Economics, University of Venice "Ca' Foscari", 2017.

gareth Anderson, Philip Bunn, Alice Pugh, and Arzu Uluc. The Bank of England/NMG Survey of Household Finances. *Fiscal Studies*, 37(1):131–152, 2016.

T. W. Anderson and Cheng Hsiao. Formulation and estimation of dynamic models using panel data. *Journal of Econometrics*, 18(1):47–82, January 1982.

Manuel Arellano and Stephen Bond. Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations. *Review of Economic Studies*, 58(2):277–297, 1991.

A. B. Atkinson and Thomas Piketty, editors. *Top Incomes: A Global Perspective.* Number 9780198727743 in OUP Catalogue. Oxford University Press, May 2014.

Gerald Auten, Geoffrey Gee, and Nicholas Turner. Income Inequality, Mobility, and Turnover at the Top in the US, 1987-2010. *American Economic Review*, 103(3):168–72, May 2013.

J. Banks, GD. Batty, K. Coughlin, K. Deepchand, M. Marmot, J. Nazroo, Z. Oldfield, N. Steel, M. Steptoe, Wood, and P. A Zaninotto. *English Longitudinal Study of Ageing: Waves 0-8, 1998-2017. 29th Edition.* UK Data Service. SN: 5050, 2019. doi: http://doi.org/10.5255/UKDA-SN-5050-16.

Marco Bassetto, Marco Cagetti, and Mariacristina De Nardi. Credit Crunches and Credit Allocation in a Model of Entrepreneurship. *Review of Economic Dynamics*, 18(1), January 2015.

Jess Benhabib, Alberto Bisin, and Shenghao Zhu. The Wealth Distribution in Bewley Models with Investment Risk. NBER Working Papers 20157, National Bureau of Economic Research, Inc, May 2014.

Jess Benhabib, Alberto Bisin, and Mi Luo. Wealth Distribution and Social Mobility in the US: A Quantitative Approach. NBER Working Papers 21721, National Bureau of Economic Research, Inc, November 2015.

Ben Bernanke and Mark Gertler. Agency costs, net worth, and business fluctuations. *American Economic Review*, 79(1):14–31, March 1989.

Ben S. Bernanke, Mark Gertler, and Simon Gilchrist. The financial accelerator in a quantitative business cycle framework. In J. B. Taylor and M. Woodford, editors, *Handbook of Macroeconomics*, volume 1 of *Handbook of Macroeconomics*, chapter 21, pages 1341–1393. Elsevier, 1999.

Truman Bewley. A Difficulty with the Optimum Quantity of Money. *Econometrica*, 51 (5):1485–1504, September 1983.

Truman F. Bewley. The permanent income hypothesis: A theoretical formulation. *Journal of Economic Theory*, 16(2):252–292, December 1977.

Claudia Biancotti, Giovanni D'Alessio, and Andrea Neri. Measurement Error In The Bank Of Italy'S Survey Of Household Income And Wealth. *Review of Income and Wealth*, 54 (3):466–493, September 2008.

Nicholas Bloom. Fluctuations in Uncertainty. *Journal of Economic Perspectives*, 28(2): 153–176, Spring 2014.

Nicholas Bloom, Fatih Guvenen, and Sergio Salgado. Skewed Business Cycles. 2016 Meeting Papers 1621, Society for Economic Dynamics, 2016.

Nicholas Bloom, Max Floetotto, Nir Jaimovich, Itay Saporta Eksten, and Stephen J. Terry. Really Uncertain Business Cycles. *Econometrica*, 86(3):1031–1065, May 2018.

Jesse Bricker, Brian K. Bucks, Arthur B. Kennickell, Traci L. Mach, and Kevin B. Moore. Surveying the aftermath of the storm: changes in family finances from 2007 to 2009. Finance and Economics Discussion Series 2011-17, Board of Governors of the Federal Reserve System (U.S.), 2011.

Jesse Bricker, Alice M. Henriques, Jacob Krimmel, and John Sabelhaus. Measuring Income and Wealth at the Top Using Administrative and Survey Data. Finance and Economics Discussion Series 2015-30, Board of Governors of the Federal Reserve System (U.S.), April 2015.

Marco Cagetti and Mariacristina De Nardi. Taxation, entrepreneurship and wealth. Federal Reserve Bank of Minneapolis Staff Report 340, July 2004.

Marco Cagetti and Mariacristina De Nardi. Entrepreneurship, Frictions, and Wealth. *Journal of Political Economy*, 114(5):835–870, October 2006.

John Y. Campbell. Have Individual Stocks Become More Volatile? An Empirical Exploration of Idiosyncratic Risk. *Journal of Finance*, 56(1):1–43, February 2001.

Christopher Carroll, Jiri Slacalek, Kiichi Tokuoka, and Matthew N. White. The distribution of wealth and the marginal propensity to consume. *Quantitative Economics*, 8(3): 977–1020, November 2017a.

Daniel R. Carroll, Nicholas Hoffman, and Eric R. Young. Mobility. Working Paper 1634, Federal Reserve Bank of Cleveland, December 2017b.

Ana Castañeda, Javier Díaz-Giménez, and José-Victor Ríos-Rull. Accounting for the U.S. earnings and wealth inequality. *Journal of Political Economy*, 111(4):818–857, August 2003.

Victor Chernozhukov and Han Hong. An MCMC approach to classical estimation. *Journal of Econometrics*, 115(2):293–346, August 2003.

Thomas F. Crossley, Cormac O'Dea, Facundo Alvaredo, Anthony B. Atkinson, and Salvatore Morelli. The Challenge of Measuring UK Wealth Inequality in the 2000s. *Fiscal Studies*, 37:13–33, 03 2016.

Mariacristina De Nardi and Giulio Fella. Saving and Wealth Inequality. *Review of Economic Dynamics*, 26:280–300, October 2017.

Mariacristina De Nardi, Giulio Fella, and Gonzalo Paz Pardo. The Implications of Richer Earnings Dynamics for Consumption and Wealth. NBER Working Papers 21917, National Bureau of Economic Research, Inc, January 2016.

Mariacristina De Nardi, Giulio Fella, and Gonzalo Paz Pardo. Earnings mobility in the UK: evidence from the NESPD. 2018.

Jason DeBacker, Bradley Heim, Vasia Panousi, Shanthi Ramnath, and Ivan Vidangos. The properties of income risk in privately held businesses. Finance and Economics Discussion Series 2012-69, Board of Governors of the Federal Reserve System (U.S.), 2012.

David S. Evans and Boyan Jovanovic. An estimated model of entrepreneurial choice under liquidity constraints. *Journal of Political Economy*, 97(4):808–827, August 1989.

Andreas Fagereng, Luigi Guiso, Davide Malacrino, and Luigi Pistaferri. Heterogeneity and Persistence in Returns to Wealth. NBER Working Papers 22822, National Bureau of Economic Research, Inc, November 2016.

Xavier Gabaix. Power Laws in Economics and Finance. *Annual Review of Economics*, 1 (1):255–294, May 2009.

Miguel Gouveia and Robert P. Strauss. Effective federal individual income tax functions: An exploratory empirical analysis. *National Tax Journal*, 47(2):317–339, June 1994.

Fatih Guvenen, Greg Kaplan, and Jae Song. The Glass Ceiling and The Paper Floor: Gender Differences among Top Earners, 1981-2012. NBER Working Papers 20560, National Bureau of Economic Research, Inc, October 2014a.

Fatih Guvenen, Serdar Ozkan, and Jae Song. The Nature of Countercyclical Income Risk. *Journal of Political Economy*, 122(3):621 – 660, 2014b.

Fatih Guvenen, Fatih Karahan, Serdar Ozkan, and Jae Song. What Do Data on Millions of U.S. Workers Reveal about Life-Cycle Earnings Risk? NBER Working Papers 20913, National Bureau of Economic Research, Inc, January 2015.

Wouter J. Den Haan. Assessing the accuracy of the aggregate law of motion in models with heterogeneous agents. *Journal of Economic Dynamics and Control*, 34(1):79–99, January 2010.

Lutz Hendricks. How important is preference heterogeneity for wealth inequality? Mimeo. Iowa State University, 2004.

Douglas Holtz-Eakin, Whitney Newey, and Harvey S Rosen. Estimating Vector Autoregressions with Panel Data. *Econometrica*, 56(6):1371–1395, November 1988.

Mark Huggett. Wealth distribution in life-cycle economies. *Journal of Monetary Economics*, 38(3):469–494, December 1996.

Michael D. Hurd. Mortality risk and bequests. *Econometrica*, 57(4):779–813, July 1989.

Michael D. Hurd. Portfolios of the elderly. In Luigi Guiso, Michael Haliassos, and Tullio Jappelli, editors, *Household Portfolios*, pages 431–472. MIT Press, 2002.

Erik Hurst, Ming Ching Luoh, and Frank P. Stafford. Wealth dynamics of American families, 1984-94. *Brookings Papers on Economic Activity*, (1):267–337, 1998.

Tullio Jappelli. The age-wealth profile and the life-cycle hypothesis: A cohort analysis with a time series of cross-sections of italian households. *Review of Income and Wealth*, 45(1):57–75, March 1999.

Tullio Jappelli and Luigi Pistaferri. The dynamics of household wealth accumulation in italy. *Fiscal Studies*, 21:269–295, Jun 2000.

Arthur B. Kennickell. Multiple imputation in the survey of consumer finances. Technical report, Board of Governors of the Federal Reserve System, 1998. Mimeo.

Arthur B Kennickell and Martha Starr-McCluer. Household Saving and Portfolio Change: Evidence from the 1983-89 SCF Panel. *Review of Income and Wealth*, 43(4):381–99, December 1997.

Aubhik Khan and Julia K. Thomas. Credit Shocks and Aggregate Fluctuations in an Economy with Production Heterogeneity. *Journal of Political Economy*, 121(6):1055–1107, 2013.

Fabian Kindermann and Dirk Krueger. High Marginal Tax Rates on the Top 1%? Lessons from a Life Cycle Model with Idiosyncratic Income Risk. NBER Working Papers 20601, National Bureau of Economic Research, Inc, October 2014.

Sagiri Kitao. Entrepreneurship, taxation and capital investment. *Review of Economic Dynamics*, 11(1):44–69, January 2008.

Nobuhiro Kiyotaki and John Moore. Credit Cycles. *Journal of Political Economy*, 105(2): 211–48, April 1997.

Wojciech Kopczuk, Emmanuel Saez, and Jae Song. Uncovering the American Dream: Inequality and Mobility in Social Security Earnings Data since 1937. NBER Working Papers 13345, National Bureau of Economic Research, Inc, August 2007.

Laurence J. Kotlikoff, Kent A. Smetters, and Jan Walliser. Privatizing Social Security in the United States: Comparing the options. *Review of Economic Dynamics*, 2(3): 532–574, July 1999.

Per Krusell and Anthony A. Smith Jr. Income And Wealth Heterogeneity, Portfolio Choice, And Equilibrium Asset Returns. *Macroeconomic Dynamics*, 1(02):387–422, June 1997.

Per Krusell and Anthony A. Smith Jr. Income and Wealth Heterogeneity in the Macroeconomy. *Journal of Political Economy*, 106(5):867–896, October 1998.

Nayoung Lee, Geert Ridder, and John Strauss. Estimation of Poverty Transition Matrices with Noisy Data. *Journal of Applied Econometrics*, 32(1):37–55, January 2017.

Erzo G. J. Luttmer. Selection, Growth, and the Size Distribution of Firms. *The Quarterly Journal of Economics*, 122(3):1103–1144, 2007.

Erzo G.J. Luttmer. Models of Growth and Firm Heterogeneity. *Annual Review of Economics*, 2(1):547–576, September 2010.

Cesaire A. Meh and Yaz Terajima. Unsecured debt, consumer bankruptcy, and small business. Working Papers 08-5, Bank of Canada, February 2008.

Benjamin Moll. Productivity Losses from Financial Frictions: Can Self-Financing Undo Capital Misallocation? *American Economic Review*, 104(10):3186–3221, October 2014.

Tobias J. Moskowitz and Annette Vissing-Jørgensen. The returns to entrepreneurial investment: A private equity premium puzzle? *American Economic Review*, 92(4):745–778, 2002.

Mariacristina De Nardi. Quantitative Models of Wealth Inequality: A Survey. NBER Working Papers 21106, National Bureau of Economic Research, Inc, April 2015.

Andrea Neri and Maria Teresa Monteduro. The housing wealth of Italian households: a comparison of administrative and survey data. Questioni di Economia e Finanza (Occasional Papers) 146, Bank of Italy, Economic Research and International Relations Area, January 2013.

Andrea Neri and Maria Giovanna Ranalli. To misreport or not to report? The measurement of household financial wealth. Temi di discussione (Economic working papers) 870, Bank of Italy, Economic Research and International Relations Area, July 2012.

Sergio Ocampo, Gueorgui Kambourov, Daphne Chen, Burhanettin Kuruscu, and Fatih Guvenen. Use It or Lose It: Efficiency Gains from Wealth Taxation. 2017 Meeting Papers 913, Society for Economic Dynamics, 2017.

University of Essex. Institute for Social and Economic Research. *British Household Panel Survey: Waves 1-18, 1991-2009, 8th Edition.* UK Data Service. SN:5151, 2018. doi: http://doi.org/10.5255/UKDA-SN-5151-2.

Office of National Statistics. Wealth and assets survey, waves 1-3, 2006-2012: Report. Technical report, Newport, UK, May 2014.

Social Survey Division Office for National Statistics and (UKDS) UK Data Service. Wealth and assets survey, waves 1-5, 2006-2016, 2018.

Thomas Piketty. *Capital in the twenty-first century.* Harvard University Press, Cambridge, MA., 2014.

Thomas Michael Pugh. Uk business dynamics. Mimeo, September 2018a.

Thomas Michael Pugh. Wealth and mobility: superstars, returns heterogeneity and discount factors. Mimeo, September 2018b.

Thomas Michael Pugh. The wealth and assets survey and wealth dynamics at the top. Mimeo, September 2018c.

Vincenzo Quadrini. The importance of entrepreneurship for wealth concentration and mobility. *Review of Income and Wealth*, 45(1):1–19, March 1999.

Vincenzo Quadrini. Entrepreneurship, saving, and social mobility. *Review of Economic Dynamics*, 3(1):1–40, January 2000.

Donald B. Rubin. *Multiple Imputation for Nonresponse in Surveys.* John Wiley and Sons, 1987.

Emmanuel Saez and Gabriel Zucman. Wealth Inequality in the United States since 1913: Evidence from Capitalized Income Tax Data. NBER Working Papers 20625, National Bureau of Economic Research, Inc, October 2014.

Petr Sedlacek and Vincent Sterk. The Growth Potential of Startups over the Business Cycle. Discussion Papers 1403, Centre for Macroeconomics (CFM), February 2014.

Jacob M. Short and Andy Glover. Bankruptcy, Incorporation and the Nature of Entrepreneurial Risk. 2011 Meeting Papers 836, Society for Economic Dynamics, 2011.

Mark Trede. Making mobility visible: a graphical device. *Economics Letters*, 59(1):77–82, April 1998.

Stef van Buuren and Karin Groothuis-Oudshoorn. mice: Multivariate imputation by chained equations in r. *Journal of Statistical Software*, 45(3):1–67, 2011.

Philip Vermeulen. Estimating the Top Tail of the Wealth Distribution. *American Economic Review*, 106(5):646–650, May 2016.

Gianluca Violante, Greg Kaplan, and Justin Weidner. The Wealthy Hand-to-Mouth. 2014 Meeting Papers 192, Society for Economic Dynamics, 2014.

Edward N. Wolff. Estimates of wealth inequality in the U.S., 1962-1983. *Review of Income and Wealth*, 33:231–256, 1987.

Appendix A

Appendix for supporting data and further details

A.0.1 Supporting transitions data

ELSA data has similar patterns to the WAS in terms of top wealth transitions, although it has a smaller sample of the top 1 and 0.1%, so a number of conditional moments are not calculable (or are extremely lumpy). We see the 'stayers stay' pattern in top groups and around a third of the top 5% exit that group between every biennial wave. The gradual decrease in the number staying in the group over time is present at the top 10%, but not clearly demonstrated above this (unlike the WAS, which has the pattern up to the top 0.1%).

	top 10%	top 5%	top 1%	top 0.1%
1-2	0.71	0.65	0.52	0.10
1-3	0.67	0.57	0.51	0.00
1-4	0.65	0.55	0.38	0.29
1-5	0.65	0.58	0.44	0.00
1-6	0.62	0.56	0.38	0.00
1-7	0.62	0.59	0.52	0.41

Table A.1: ELSA: Staying rates in top wealth groups over waves.

A.0.2 Person level wealth

The value of main residences and some physical wealth is collected at the household level and must be allocated to individuals to create person-level total wealth. The main residence and household physical wealth[1] are allocated equally to the head of household[2] and their partner, if one is present, and otherwise to the head of household entirely.

Alternatives include allocating entirely to the head of household, the partner-splitting approach and an equal split amongst adults in the household. As most households have

[1]Household contents, vehicles and collectibles.
[2]Also known as household reference person.

one or two adults, the third assumption makes little difference. Unsurprisingly, the head of household-only assumption leads to greater concentration of wealth but by less than 2 percentage points for the top 1% share. Later, exploration of wealth excluding housing obviously bypasses this issue.

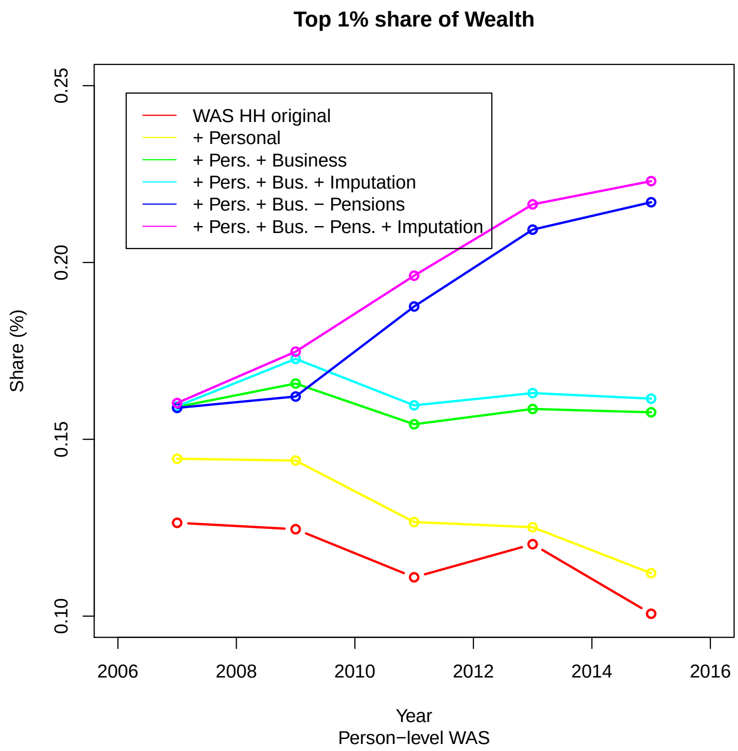

Figure A.1: Wealth shares of Top 1% for Person level WAS.

The movement from household to person level in figure A.1 increases wealth shares a small amount, but clearly the same patterns are present. The changes are also within predicted bounds from Crossley et al. [2016] for person- versus household-level measurements.

A.0.3 WAS estate-comparable dataset

There are a number of less important changes are needed to create an estate-comparable WAS dataset. Estate data only includes those over 18. For WAS waves 1, 2 and 4 excluding individuals below 18 is difficult as age is in banded categories which do not overlap with that limit. I explored ways to mitigate this problem, but found no difference to results because 16-20 year olds generally have very little wealth. As a result, in waves 1,2 and 4 the estate-comparable database also exclude 19 year olds, as the most direct option. For wave 3 and 5, age is directly provided and thus cleanly comparable to estate data.

A.0.4 Calculating total income and tax paying in the WAS

The WAS does not directly have a total income variable, either in net or gross value in the end-user data. Income in the WAS is recorded under the following broad categories:

- (labour or self-employed) earnings
- pension income
- social benefit income
- investment income
- rental income
- other income (royalties, gambling, irregular income, etc)

Some of these items are collected both pre- and post-tax, others vary. Some items, such as income from bonuses allow respondents to choose to offer either a pre- or post-tax value but not both. It is therefore difficult to construct total income directly from the variables in the dataset, as UK income taxes and national insurance are calculated based on total income from all sources rather than item by item.

In order to represent incomes as accurately as possible, I first construct tax functions for income tax and national insurance for 2006-2014. I then create a weighted average function of these for each wave period. For example, as the first wave covers 06-08 and respondents are interviewed at different times, the taxes in force over this period are weighted in line with their respective length of enforcement.

Incrementally, I apply/invert[3] these tax functions to each item of an individual's income - so first inferring net and gross earnings income, then adding pension income to that total and apply the next set of incremental taxes and so on. In effect, this replicates the procedure used by tax collecting agencies. There are subtleties within this, such as different limits at different ages, different tax of pension income from overseas, etc, which are accounted for. This is broadly the procedure and ordering used by the tax authority.

However, this cannot account for some key rebates. Firstly, large dividend income is taxed at a lower rate than other capital income (10% lower than the maximum rate for other income). The WAS asks for investment income from all sources as a single figure, so one cannot isolate income from shares in order to apply this. The procedure is therefore likely to over-tax the wealthy, who hold more shares. It also cannot account for tax rebates on rental income. In the UK, landlords are able to reclaim tax based on mortgage payments for the rented property. Landlords can potentially pay near-zero tax on rental income on this basis. The procedure simply applies full tax to rental income.

[3] A minimising algorithm infers the gross income where there is only have net income, as the functions are not explicitly invertible, but are one-to-one.

Again, this is likely to over-tax the wealthy. There are also income tax rebates available to those making investments in small businesses or charitable donations, at substantial rates - the 'EIS' (Enterprise Investment Scheme) delivers an income tax credit of 30% of the value of an investment and the 'Seed' version of the scheme offers 50%. There are also a multitude of subsidy and rebate schemes for business owners. Therefore, again, the procedure is likely to over-tax the wealthy, especially those with business wealth.

As well as those rebates, there is no correction for student loan repayment. Recently, the UK government provides student loans and grants centrally and requires loan repayment via additional tax on income. The WAS does not gather the amount the former student pays off, only the outstanding amount. Tax paid is usually 9% over a threshold, up to the outstanding loan amount, however, it also depends on features of the individual's personal history and loans are cancelled after a very long period of low earnings. This affects a relatively small number of individuals in the WAS, less than 1%, due to the relatively late introduction of widespread student loans to the UK ensuring only a small adult population is affected in the survey. Other UK sources do not include student loan repayment in their net income measures.

Finally, irregular income is excluded. The reasoning for removing this small item is that the taxation of this income is hard to calculate. Inherited items are included in the WAS directly as transfers of wealth, rather than income items. Irregular income in the WAS concerns payouts from insurance schemes, gambling and lump sum redundancy payments. General redundancy payments are included in the income measure and few respondents mention lump sum redundancy payments here, possibly as they include it earlier in the income questioning process.

Overall, the gross and net income imputation is likely to overtax the wealthy, even prior to considering tax avoidance and evasion. Given the good match to the income data in the World Income Database, which uses administrative data akin to the SPI, I am confident that these oversights do not damage the analysis. The lower income in the WAS database at very high income quantiles is likely due to factors such as the above, as well as measurement error and other common causes.

Each wave contains its own minor problems, which are dealt with on a case-by-case basis and details are available upon request, where taxes paid or benefits received are calculated by combining headline rates with eligibility conditions.

A.0.5 Attrition

There are multiple methods to deal with attrition - (multiple) imputation; reweighting using sample exit propensity and simple rescaling/removing.

I argue that, for the household-level WAS, simple rescaling is the best option as exit appears to be at random. Reweighting using inverse estimated sample exit propensity is likely to introduce substantial noise. Multiple imputation of all variables for all waves for all exiting households/individuals over the entire sample would be risky - with around 30% of the sample lost to attrition in each wave, imputed data would be the majority of the dataset by wave 3 or 4.

Much like the ONS' longitudinal weights for individuals, one can use a binary model to predict household sample exit, and then reweight to account for any differential attrition propensity in order to reconstruct a 'correctly' weighted sample. However, when applying this analysis to the household level WAS, the logistic regression normally used for such a procedure is extremely poor at predicting exit.

To demonstrate this, it is compared to a low benchmark - if one were to take the sample attrition rate and allocate exit randomly to respondents at this rate, one would obtain the matrix in table A.3 when comparing this 'model' to the data. By pure chance, one would correctly identify some leavers. One could easily do better, for example by predicting 'all stay'. Comparing the random allocation to the logistic model when engineered to generate the right proportion of exit (and thus be comparable in a simple manner) in Table A.2, the logistic model does somewhat better. But, its level of false positive and false negative are extremely high (the two non-diagonal cells) and close to the benchmark of random allocation. Examining the correct-classification-ratio for all three transitions in table A.4, the logistic model is not far ahead of the random benchmark throughout and is worse than simply predicting none leave. This is concerning for implementing such a model to predict exit and then reweighting with the inverse of its implied probabilities - high predicted exit probabilities of, say, 80% would be more than doubly weighted but given that the attrition model appears to include a lot of noise, reweighting using these noise predictions has the potential to damage rather than improve estimates.

Model/Data	D0	D1
M0	0.41	0.2
M1	0.20	0.2

Table A.2: Distribution of Model binary exit predictions versus data. (Exit=1) Wave 1-2.

	D0	D1
M0	0.37	0.24
M1	0.24	0.16

Table A.3: Distribution of random binary exit predictions versus data. (Exit=1) Wave 1-2.

The poor performance of the logistic exit regression directly suggests that simple random attrition is occurring. Indeed, directly studying attrition rates by wealth and income categories (since these would be areas particularly concerning to this study), there is not any large variation. In the main text, random attrition is used. In this section are some calculations for staying in top wealth categories showing exit when accounting for differ-

Wave	W1-2	W2-3	W3-4
Model	0.60	0.63	0.63
Random	0.52	0.57	0.57
P(Stay)	0.60	0.69	0.68

Table A.4: Correct Classification Ratio for model and random raw probability assignment with raw probability of staying in sample.

ential propensity to leave by wealth category.

from/to	top 0.1%	top 0.1-1%	top 1-5%	top 5-10%	<top 10%	NA
top 0.1%	0.26	0.19	0.10	0.03	0.10	0.33
top 0.1-1%	0.03	0.33	0.17	0.04	0.04	0.39
top 1-5%	0.00	0.05	0.37	0.13	0.08	0.37
top 5-10%	0.00	0.00	0.12	0.29	0.25	0.33
<top 10%	0.00	0.00	0.00	0.02	0.56	0.42

Table A.5: Transitional Probabilities for top wealth groups across WAS waves 1-2 (07-09), household wealth. Includes exit as 'NA'.

	top 0.1%	top 0.1-1%	top 1-5%	top 5-10%	<top 10%	NA
top 0.1%	0.36	0.41	0.11	0.00	0.00	0.12
top 0.1-1%	0.01	0.38	0.23	0.01	0.01	0.36
top 1-5%	0.00	0.05	0.42	0.17	0.08	0.28
top 5-10%	0.00	0.00	0.09	0.36	0.26	0.28
<top 10%	0.00	0.00	0.00	0.01	0.65	0.33

Table A.6: Transitional Probabilities for top wealth groups across WAS waves 2-3 (09-11), household wealth. Includes exit as 'NA'.

	top 0.1%	top 0.1-1%	top 1-5%	top 5-10%	<top 10%	NA
top 0.1%	0.30	0.25	0.08	0.00	0.00	0.37
top 0.1-1%	0.02	0.36	0.18	0.01	0.03	0.40
top 1-5%	0.00	0.04	0.45	0.16	0.07	0.28
top 5-10%	0.00	0.00	0.11	0.39	0.23	0.26
<top 10%	0.00	0.00	0.00	0.01	0.65	0.33

Table A.7: Transitional Probabilities for top wealth groups across WAS waves 3-4 (11-13), household wealth. Includes exit as 'NA'.

from/to	top 0.1%	top 0.1-1%	top 1-5%	top 5-10%	<top 10%
top 0.1%	0.39	0.28	0.15	0.04	0.14
top 0.1-1%	0.05	0.54	0.28	0.06	0.06
top 1-5%	0.00	0.08	0.59	0.20	0.13
top 5-10%	0.00	0.01	0.18	0.43	0.38
<top 10%	0.00	0.00	0.01	0.03	0.96

Table A.8: Transitional Probabilities for top wealth groups across WAS waves 1-2 (07-09), household wealth. Adjusts for sample exit by wealth category.

from/to	top 0.1%	top 0.1-1%	top 1-5%	top 5-10%	<top 10%
top 0.1%	0.40	0.47	0.13	0.00	0.00
top 0.1-1%	0.02	0.59	0.36	0.02	0.02
top 1-5%	0.00	0.06	0.59	0.24	0.11
top 5-10%	0.00	0.00	0.13	0.51	0.36
<top 10%	0.00	0.00	0.01	0.02	0.97

Table A.9: Transitional Probabilities for top wealth groups across WAS waves 2-3 (09-11), household wealth. Adjusts for sample exit by wealth category.

from/to	top 0.1%	top 0.1-1%	top 1-5%	top 5-10%	<top 10%
top 0.1%	0.47	0.40	0.13	0.00	0.00
top 0.1-1%	0.04	0.59	0.30	0.02	0.05
top 1-5%	0.00	0.05	0.62	0.22	0.10
top 5-10%	0.00	0.00	0.14	0.54	0.32
<top 10%	0.00	0.00	0.00	0.02	0.98

Table A.10: Transitional Probabilities for top wealth groups across WAS waves 3-4 (11-13), household wealth. Adjusts for sample exit by wealth category.

	top 10%	top 5%	top 1%	top 0.1%
07-09	0.75	0.71	0.60	0.39
09-11	0.77	0.72	0.64	0.40
11-13	0.79	0.72	0.65	0.47

Table A.11: WAS, proportion staying in top quantile groups across waves, household wealth. Adjusts for sample exit by wealth category.

A.0.6 Additional Transitions

Below are the Markov matrices for all the wave-to-wave transitions not included in the main text.

from/to	<top 10%	top 5-10%	top 1-5%	top 0.1-1%	top 0.1%
<top 10%	0.97	0.02	0.01	0.00	0.00
top 5-10%	0.43	0.41	0.16	0.01	0.00
top 1-5%	0.15	0.22	0.56	0.08	0.00
top 0.1-1%	0.06	0.07	0.30	0.51	0.06
top 0.1%	0.16	0.04	0.13	0.25	0.41

Table A.12: Transitional Probabilities for top HH WAS wealth groups 07-09.

from/to	<top 10%	top 5-10%	top 1-5%	top 0.1-1%	top 0.1%
<top 10%	0.97	0.02	0.01	0.00	0.00
top 5-10%	0.37	0.47	0.15	0.00	0.00
top 1-5%	0.10	0.23	0.60	0.07	0.00
top 0.1-1%	0.02	0.01	0.36	0.58	0.03
top 0.1%	0.00	0.00	0.02	0.48	0.50

Table A.13: Transitional Probabilities for top HH WAS wealth groups 09-11.

from/to	<top 10%	top 5-10%	top 1-5%	top 0.1-1%	top 0.1%
<top 10%	0.98	0.02	0.00	0.00	0.00
top 5-10%	0.33	0.50	0.17	0.00	0.00
top 1-5%	0.10	0.21	0.62	0.06	0.01
top 0.1-1%	0.05	0.01	0.28	0.62	0.03
top 0.1%	0.00	0.00	0.14	0.41	0.44

Table A.14: Transitional Probabilities for top HH WAS wealth groups 11-13.

	<top 10%	top 5-10%	top 1-5%	top 0.1-1%	top 0.1%
<top 10%	0.98	0.02	0.00	0.00	0.00
top 5-10%	0.34	0.50	0.16	0.00	0.00
top 1-5%	0.09	0.20	0.64	0.06	0.00
top 0.1-1%	0.05	0.02	0.22	0.66	0.04
top 0.1%	0.02	0.00	0.21	0.27	0.49

Table A.15: Transitional Probabilities for top HH WAS wealth groups 13-15.

Wealth mobility & transitions including pensions

This section contains transitions including pension wealth. Due to the DB pension modelling changes in the WAS and the resultant fluctuations, mobility is higher. However, the patterns of 'stayers stay', increased mobility in 07-09 and the quantile regression results all remain robust to including pensions. In most cases, the patterns described in the main text become stronger.

	top 0.1%	top 0.1-1%	top 1-5%	top 5-10%	<top 10%
top 0.1%	0.23	0.31	0.22	0.06	0.17
top 0.1-1%	0.05	0.35	0.30	0.14	0.17
top 1-5%	0.00	0.07	0.58	0.20	0.15
top 5-10%	0.00	0.02	0.16	0.43	0.39
<top 10%	0.00	0.00	0.01	0.03	0.96

Table A.16: Transitional Probabilities matrix for 07-09 transition in top quantile groups for total wealth including pensions

	top 0.1%	top 0.1-1%	top 1-5%	top 5-10%	<top 10%
top 0.1%	0.40	0.27	0.25	0.08	0.00
top 0.1-1%	0.01	0.45	0.31	0.10	0.13
top 1-5%	0.00	0.08	0.60	0.21	0.10
top 5-10%	0.00	0.01	0.17	0.47	0.36
<top 10%	0.00	0.00	0.01	0.02	0.97

Table A.17: Transitional Probabilities matrix for 09-11 transitions in top quantile groups for total wealth including pensions

	top 0.1%	top 0.1-1%	top 1-5%	top 5-10%	<top 10%
top 0.1%	0.54	0.32	0.09	0.06	0.00
top 0.1-1%	0.03	0.57	0.32	0.01	0.07
top 1-5%	0.00	0.05	0.61	0.21	0.12
top 5-10%	0.00	0.01	0.17	0.46	0.36
<top 10%	0.00	0.00	0.01	0.02	0.97

Table A.18: Transitional Probabilities matrix for 11-13 transitions in top quantile groups for total wealth including pensions

Extra Conditional Staying Probabilities

Shown here are further conditional probabilities of staying in top wealth groups, as well as the same for top income groups. Table A.19 shows the conditional probabilities for the shorter history of wave 3 transitions (versus the wave 4 transitions and more extensive histories used in the main text) whilst table A.20 shows the same shorter conditional histories for wave 4 transitions. The conclusions are exactly the same as in the main text - those with a shorter history in the group (in this case, new entrants with no further history) are much more likely to leave.

Table A.21 shows the same table as in the main text, but for income. Again, there is the same pattern with similar magnitudes to wealth.

These tables include the top 0.1%, unlike the main text. This is to show the unreliability of conditional transition probabilities at this level. As one can see, sometimes the probability is 1 (or 0 in some tables not shown here) and experiences much more dramatic changes due to the small number of observations available for each specific history when dealing with the top 0.1%.

	top 10%	top 5%	top 1%	top 0.1%
$P(T_3\|F_1T_2)$	0.52	0.44	0.37	0.46
$P(T_3\|T_1T_2)$	0.87	0.87	0.81	0.55

Table A.19: Probability of remaining in top wealth groups given different histories, where 'T_t' indicates membership in wave t and 'F_t' indicates not.

	top 10%	top 5%	top 1%	top 0.1%
$P(T_4\|F_2T_3)$	0.50	0.42	0.47	0.06
$P(T_4\|T_2T_3)$	0.88	0.86	0.83	0.71

Table A.20: Probability of remaining in top wealth groups given different histories, where 'T_t' indicates membership in wave t and 'F_t' indicates not.

	top 10%	top 5%	top 1%	top 0.1%
$P(T_4\|F_1F_2T_3)$	0.43	0.41	0.46	0.35
$P(T_4\|F_1T_2T_3)$	0.70	0.63	0.43	1.00
$P(T_4\|T_1T_2T_3)$	0.80	0.79	0.82	1.00

Table A.21: Probability of remaining in top income groups given different histories, where 'T_t' indicates membership in wave t and 'F_t' indicates not.

A.0.7 Further Quantile regression diagrams and outputs

Log Quantile Regressions

Transforming the data into logs (and excluding negative wealth) in figure 1.12, the majority of the data[4] is close to unit root, with an almost linear relationship and slope similar to 1 for most of the distribution and greater spread for the richest and poorest.

[4] exp(8) to exp(14) covers below the 10th percentile to above the 99th percentile.

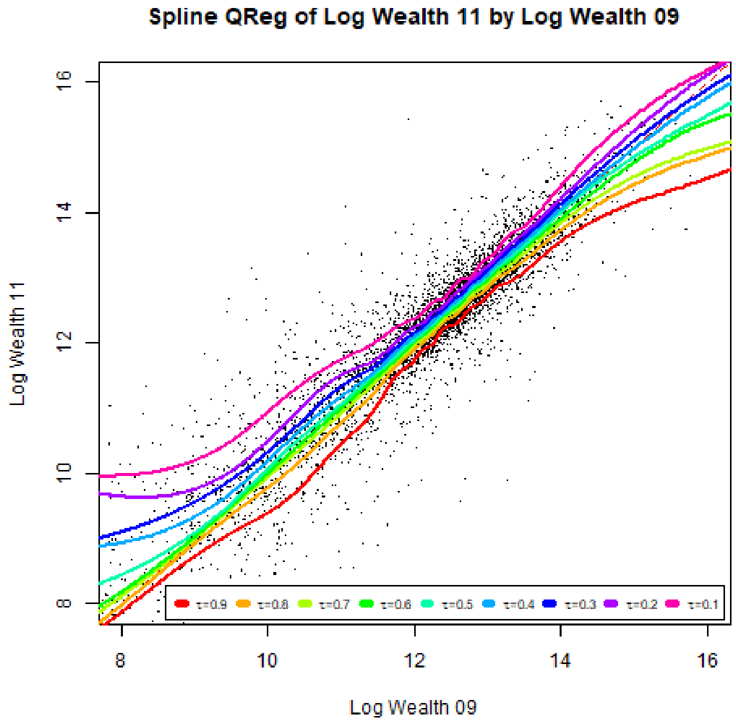

Figure A.2: Non-Linear Quantile Regression for Log Relative Wealth 2011 vs Log Relative Wealth 2009.

Linear Quantile Regressions

The linear quantile regression relationship described in reference to figure 1.11 is shown in figure A.3. The possibility of fixed effects means that quantile regression estimates of lagged linear coefficients are potentially biased. However, the purpose in this particular section is not identifying the parameter, but rather showing that different waves all have the same relationships in terms of mobility. The wave-by-wave linear quantile regression coefficients are shown in figure A.4.

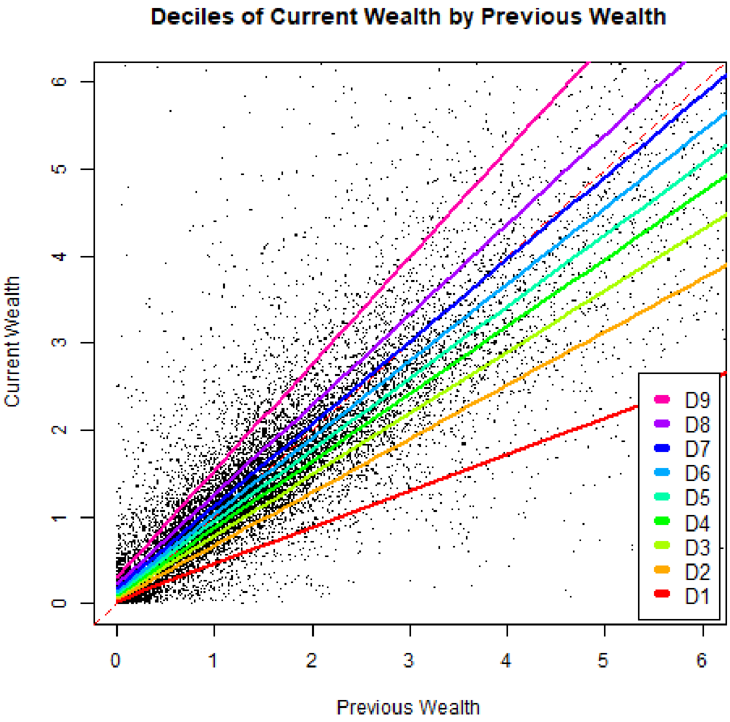

Figure A.3: Linear Quantile Regression for Relative Wealth 2009 vs Relative Wealth 2007.

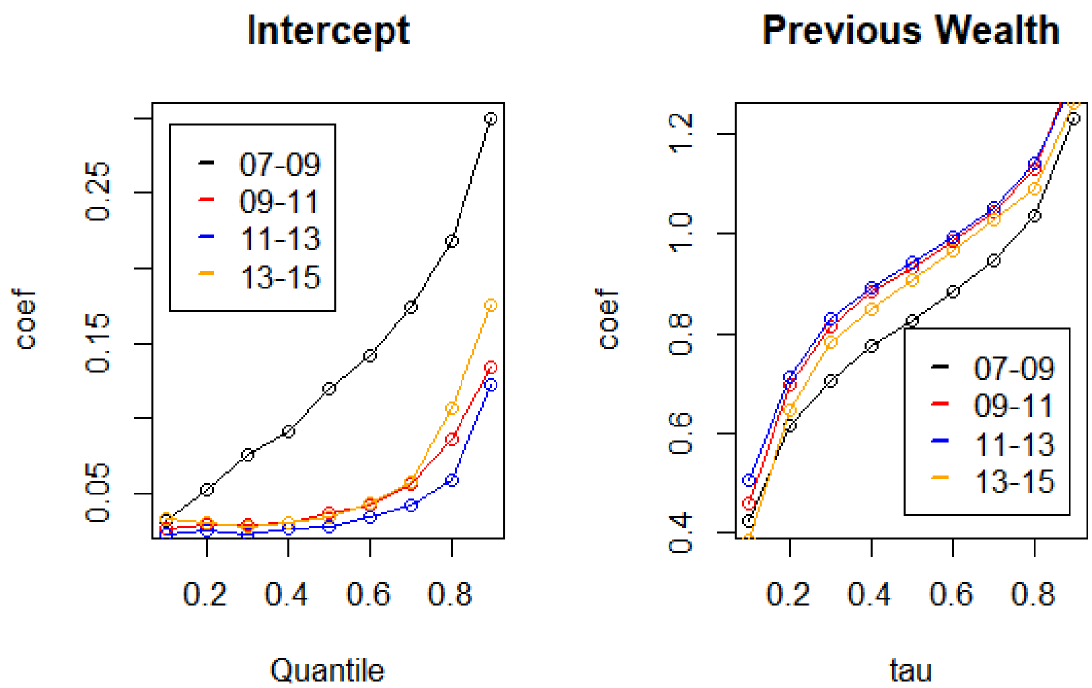

Figure A.4: Linear Quantile Regression Coefficients for Relative Wealth t on Relative Wealth $t-1$ for $t = 2009$, 2011 & 2013.

Quantile Regressions of changes

The distribution of changes in log wealth in 2011-13 versus that in 2009-11 is similar to 2009-11 vs 2007-09, but has a relatively small area of flat quantiles with minimal spread, and the rise in spread further from the origin is relatively more gentle. However, the general pattern is the same. Similarly, the longer horizon plot of log wealth changes in 2011-13 vs 2007-09 shows the widening volatility/spread away from the origin (particularly beyond ± 0.25) but there is more mobility over proportional changes in the sense that the curves are flatter, especially in the central region. For most households, most of the time, biennial proportional changes in wealth are nearly independent with a 4 year horizon, but have some reversion on a 2 year horizon.

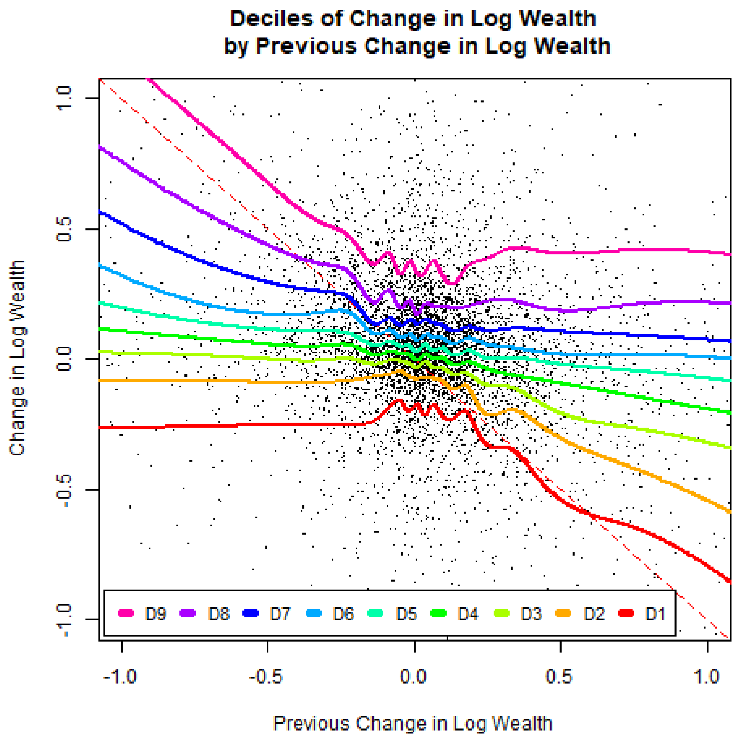

Figure A.5: Non-Linear Quantile Regression for Differences Log Wealth 2013-2011 vs Differences Log Wealth 2011-2009.

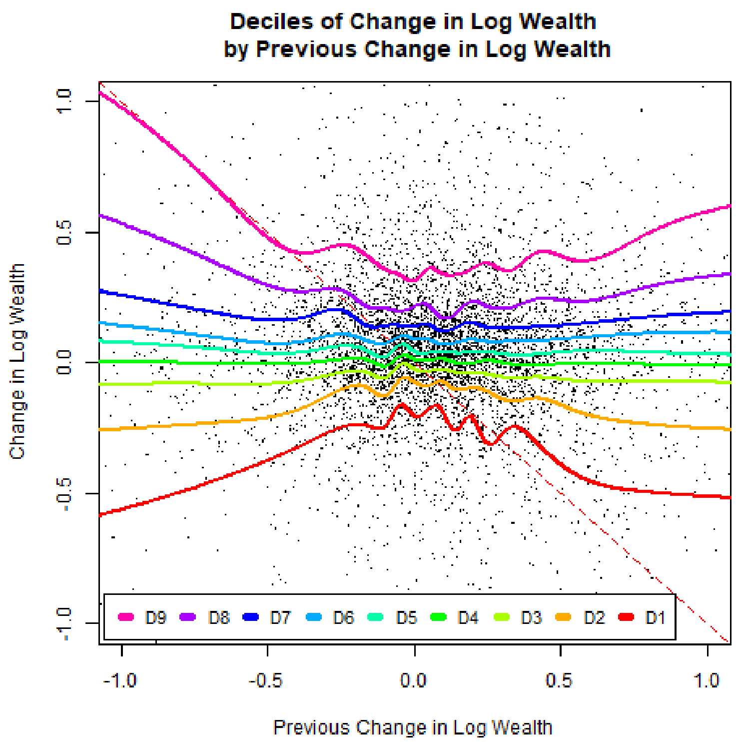

Figure A.6: Non-Linear Quantile Regression for Differences Log Wealth 2009-2011 vs Differences Log Wealth 2007-2009.

A.0.8 Percentage and log change distributions

Percentage change places 100% weight on the earlier period as the denominator, whereas the log difference weights the two periods in the change more equally - with wealth changes of -90% or 200% in the data, the two measures can be radically different. However, there is relatively little difference between the changes in log and percentage change distributions in terms of analysis and conclusions, despite said numerical differences. Here, I present both log and percentage statistics for additional robustness.

Figure A.7 shows the density of changes in log wealth for the top 5%. The distribution contains several outliers and has a substantial appearance of negative skew (substantial density spread to the left of peak versus the right). There is a noticeable difference between the recessionary transitions (07-09, black) and other transitions (09-11, red and 11-13, green). 2007-2009 has greater spread with a substantially greater appearance of negative skew. The appearance of skew can be different to the estimated coefficient - outliers dramatically influence the calculated moments.

Table A.22 presents the Pearson-style empirical moments for the changes in log distribution. Trimming outliers to±100% shows the influence of the extreme observations in table A.23. The most obvious and robust time difference is the reduction in mean wealth change for 07-09 - unsurprisingly, on average, wealth falls in the recession for these households. The other changes in the empirical moments do not clearly match the conclusions from the density graph, nor show other clear patterns.

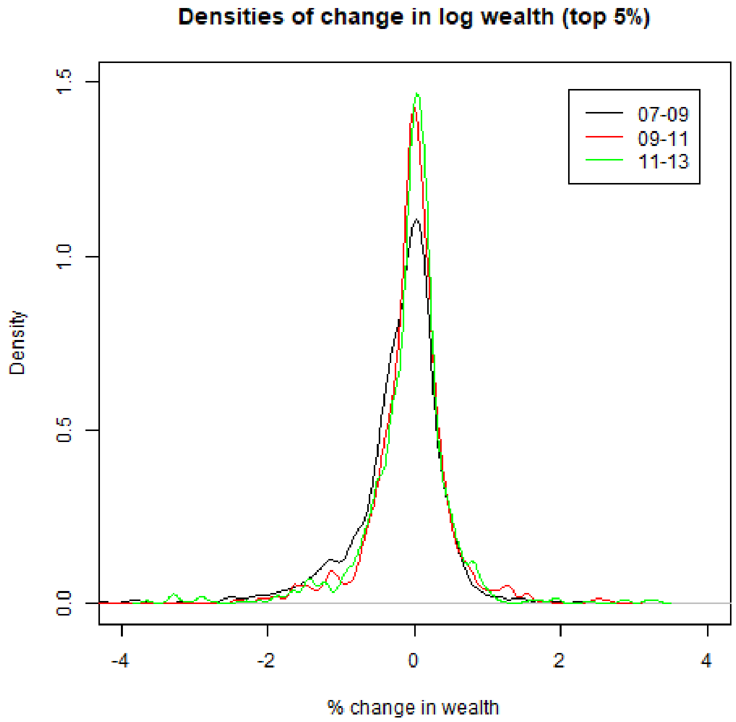

Figure A.7: Distribution of changes in log household wealth for top 5% in WAS

	mean	var	skew	kurt
07-09	-0.24	0.65	-5.31	54.08
09-11	-0.08	0.52	-6.15	83.34
11-13	-0.08	0.39	-0.92	9.65

Table A.22: Moments for change in log household wealth for top 5%

Quantiles of the changes in log wealth for the top 5 and 1% in tables A.24 and A.25 show similar conclusions to the percentage change tables in the main text - a wider distribution at the top and large proportional losses.

Moving onto percentage change statistics, the density figure A.8 shows the pattern of greater left density and heavier outlier presence in 07-09. The moments for percentage change in wealth do show a reduction in skew (i.e. greater negative skew) in 07-09 (similar to Guvenen et al's finding for income in the US Guvenen et al. [2014b]) but also less variance and kurtosis. This is likely due to percentage changes limiting losses in wealth to the [0,1] interval (excluding debt) whilst upward changes are unbounded, so more negative changes will result in less variation contributing to moments.

In table A.26 the % change distribution shows an overall positive mean in 09-11 and 11-13, through this is entirely driven by extreme positive observations and disappears when trimming these extreme observations in table A.27.

In table A.28, the quantiles of changes in wealth are shown for the top 5%. Those losing money lose much more money over the 07-09 transition in terms of absolute pounds (£). The table shows the large amounts that are gained or lost with high probability by

	mean	var	skew	kurt
07-09	-0.16	0.20	-0.18	-0.15
09-11	-0.05	0.18	-0.12	0.62
11-13	-0.05	0.18	-0.26	0.44

Table A.23: 100% trimmed moments for change in log household wealth for top 5%

Quantile	0.1	0.25	0.5	0.75	0.9
07-09	-0.92	-0.42	-0.09	0.13	0.36
09-11	-0.61	-0.26	-0.02	0.17	0.42
11-13	-0.65	-0.27	-0.01	0.17	0.43

Table A.24: Quantiles of changes in log wealth for top 5%.

Quantile	0.1	0.25	0.5	0.75	0.9
07-09	-1.53	-0.82	-0.26	0.10	0.45
09-11	-1.04	-0.44	-0.03	0.21	0.54
11-13	-1.32	-0.51	-0.08	0.14	0.45

Table A.25: Quantiles of changes in log wealth for top 1%.

the wealthy.

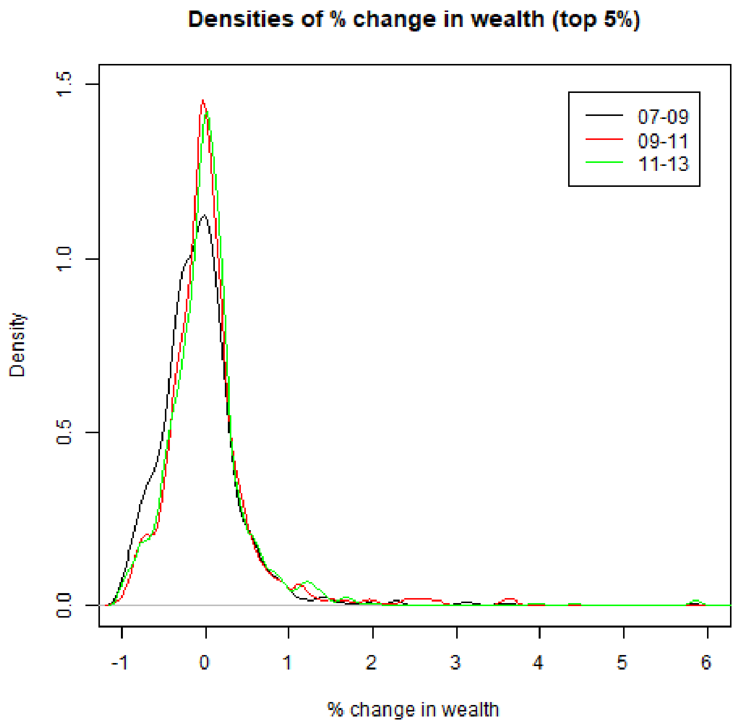

Figure A.8: Distribution of % changes in household wealth for top 5% in WAS

	mean	var	skew	kurt
07-09	-0.05	0.36	5.18	54.22
09-11	0.09	0.78	7.20	71.73
11-13	0.09	0.92	7.46	69.47

Table A.26: Moments for % change in household wealth for top 5%

	mean	var	skew	kurt
07-09	-0.08	0.17	0.40	0.37
09-11	0.01	0.16	0.52	0.66
11-13	0.00	0.16	0.39	0.60

Table A.27: Moments for % change in household wealth for top 5%, data trimmed at 100%

Quantile	0.1	0.25	0.5	0.75	0.9
07-09	-848991	-373489	-92800	140094	473402
09-11	-579200	-247473	-22349	178751	641100
11-13	-707300	-288104	-8603	214595	697396

Table A.28: Quantiles of changes in household wealth for top 5%

A.0.9 Further wealth regression details

To predict continuous wealth changes (rather than group exit), the fit of a flexible machine learning Random Forest algorithm is considered alongside weighted least squares (WLS) in table A.29. This is restricted to changes amongst the top 5%, following the focus on the wealthy tail. Polynomials of previous income; changes in income; wealth and the breakdown of wealth are all significant in various regressions but overall predictive power is low, as shown by the test-set R^2 in table A.29 of less than 0.25. Linear models perform much worse out-of-sample (in the test set of data), due to extreme observations dominating residuals even when trimming data, whereas the more successful random forest has greater flexibility and can avoid this.

Method	WLS	WLS	WLS	WLS	RF	RF
Variable	W	W	$log(W)$	$log(W)$	$log(W)$	$log(W)$
Trim data?		tr		tr		tr
R^2_{train}	0.53	0.10	0.41	0.27	0.88	0.95
R^2_{test}	0.10	0.01	0.03	0.15	0.23	0.24

Table A.29: Fit statistics for predicting changes in wealth (W), log wealth (log(W)), under WLS or Random Forest (RF). 'tr' indicates trimming at ±200% for log(W) and $\pm 10^6$ for W. Training set, 5/8 of data, test set 3/8 of data.

The full WLS predictive regression for changes in log and raw wealth uses orthogonal polynomials (except for factors) of, Total Wealth (as defined in the main text), Property Wealth, Business Wealth, Net Financial Wealth, Credit Card Balances, Income, Income Change, Log Income change, Mortgage value, Age of Household Reference Person (HRP), value of shares, self-employment, Financial Liabilities, (wealthy) hand-to-mouth status[5], Net Income, Mortgage payments, Education level of HRP. Output and changes in log wealth trimmed to absolute values of 200% for purpose of fit statistics when indicated in the main text. The Random forest uses the same variables (without utilising polynomials, since these are unnecessary under the tree methodology).

Least squares regression coefficients for previous income, income changes and log income changes are all highly significant, but do not explain much more than 5% the variation in the R^2 sense. The results for wealth variables tell a similar tale.

[5] As defined by Violante et al. [2014].

Feature	Estimate	Std. Error	t value	Pr(>|t|)
Intercept	-0.1898	0.0176	-10.81	0.0000
GI	-71.4397	24.8536	-2.87	0.0041
GI square	-58.0360	19.2383	-3.02	0.0026
GI cube	-5.3412	2.3840	-2.24	0.0253
GId	-70.1439	24.2393	-2.89	0.0039
GId square	58.6674	19.6330	2.99	0.0029
GId cube	-12.6568	4.4106	-2.87	0.0042
lGId	3.8870	0.8643	4.50	0.0000
lGId square	-1.5528	0.8942	-1.74	0.0828
lGId cube	0.6804	1.0177	0.67	0.5039
			R^2_{train}	0.0608
			R^2_{test}	0.0521

Table A.30: WLS regression of 200%-trimmed changes in log wealth on cubic orthogonal polynomials of income variables - prior gross income (GI), change in income (GId), change in log income (lGId).