FACULTY OF ECONOMICS AND BUSINESS ADMINISTRATION

Wealth Inequality and Wealth Mobility in the United States

Christophe Van Langenhove

Supervisors: Prof. Dr. Freddy Heylen, Prof. Dr. Gert Peersman

A dissertation submitted to Ghent University in partial fulfilment of the requirements for the degree of Doctor of Economics.

Academic year: 2024 – 2025

GHENT UNIVERSITY

Doctoral Examination Board

Prof. dr. Gerdie Everaert
Chair, Ghent University

Prof. dr. Freddy Heylen
Supervisor, Ghent University

Prof. dr. Gert Peersman
Supervisor, Ghent University

Prof. dr. Paula Gobbi
Université Libre de Bruxelles (ECARES)

Prof. dr. Yasin Kürşat Önder
Ghent University

Prof. dr. Alberto Russo
Università Politecnica delle Marche

Prof. dr. Dirk Van de gaer
Ghent University

Acknowledgments

Dit dankwoord markeert het einde van mijn doctoraatstraject, en tegelijkertijd het slot van een fundamenteel hoofdstuk in mijn leven. Dat ik dit traject überhaupt heb mogen ondernemen, voelt als een immens voorrecht, en daar ben ik bijzonder dankbaar voor. Mocht ik opnieuw voor de keuze staan, zou ik er zonder aarzelen en met veel enthousiasme opnieuw aan beginnen.

Het succesvol afronden van dit doctoraat was evenwel onmogelijk geweest zonder de steun van vele mensen en instellingen. Hen één voor één bedanken zou leiden tot een ellenlange opsomming die onwenselijk is – jullie hebben immers nog iets meer dan 200 pagina's leeswerk voor de boeg. Ik gebruik dit dankwoord dan ook graag om een aantal van hen expliciet te bedanken.

Allereerst wil ik mijn promotor, professor Freddy Heylen, bedanken voor zijn onvoorwaardelijke steun gedurende de voorbije jaren. Freddy, vanaf het prille begin van mijn doctoraatstraject voelde ik jouw rotsvast vertrouwen in mijn potentieel als onderzoeker. Nooit heb ik enige twijfel daarrond ervaren, zelfs niet toen ik wellicht iets te ruim mijn tijd nam om me op een definitief onderzoeksthema vast te pinnen. Je gunde me echter alle vrijheid in deze zoektocht, een vrijheid die doorslaggevend was in mijn ontwikkeling tot zelfstandige onderzoeker. Ik heb doorheen de jaren ook bijzonder veel geleerd uit jouw werkethiek, je toewijding aan onderwijs en onderzoek, en jouw uitzonderlijke oog voor detail. Deze professionele kwaliteiten combineer je bovendien met een zeer hartelijke persoonlijkheid. Verder heb je talloze uren besteed aan het nauwgezet nalezen van mijn onderzoeksvoorstellen, papers en presentaties, zelfs wanneer je eigen agenda overvol zat. Heel erg bedankt voor al deze generositeit! Ik had me geen betere promotor kunnen wensen.

I would also like to thank my co-supervisor, professor Gert Peersman, and the other members of my doctoral examination board. Your detailed feedback was extremely helpful in preparing my PhD chapters for journal submission. Moreover, the internal defence took place in a very constructive setting – made even warmer, quite literally, by the heatwave that day. Thank you, Gert Peersman, Paula Gobbi, Yasin Kürşat Önder, Alberto Russo, and Dirk Van de gaer.

Daarenboven prijs ik me uiterst gelukkig dat ik mijn doctoraat heb kunnen schrijven in een aangename en stimulerende werkomgeving. In de eerste plaats wil ik hiervoor Arthur en Marthe bedanken. Dankzij de onberispelijke werkethiek op de bovenste verdieping van Sint-Pietersplein 6 brachten we in sommige weken een behoorlijk absurd aantal uren samen door in *de office*. Ook buiten de muffe kantooromgeving spendeerden we menige uren in elkaars gezelschap: al lopend tussen tractoren in uithoeken van Vlaanderen, al slurpend van belabberde cocktails, en al professioneel ontmijnend van explosieve smoothies in niet nader gespecifieerde ijskasten. Ik koester deze momenten zeer, hoop dat er nog vele mogen volgen, en had me geen fijnere collega's kunnen voorstellen om al deze levensuren mee door te brengen! Arthur, dank je voor jouw ongeëvenaarde en ogenschijnlijk grenzeloze toewijding aan de vriendschap, en voor de buitengewone fijnzinnigheid die jou kenmerkt. Het ontbreekt me aan literair talent om mijn dankbaarheid daarvoor in passende termen uit te drukken, maar we kunnen niet allemaal even quasi-geniaal zijn, uiteraard. Marthe, dank je om het kantoor steeds opnieuw te vullen met jouw gevatheid, voor je aanmoediging in de lancering van mijn ongetwijfeld roemrijke loopcarrière, en uiteraard voor de memorabele academische samenwerking aan de start van ons doctoraat. En ter info: de kombuchaflesjes aan de kantoorschoorsteen laat ik met uiterst veel genoegen nog eventjes stof vergaren.

Daarnaast heb ik ook met Pieter, Vinanda en Yasmine vele uren op kantoor mogen doorbrengen. Dank jullie alle drie voor de ontelbare fijne gesprekken over de meest uiteenlopende onderwerpen, en voor de aangename lunches in de vakgroeptuin en op verplaatsing. Jullie aanwezigheid en enthousiasme maakten zelfs de zwaarste doctoraatsdagen een pak lichter. Verder vielen er ook buiten de immer tv-soap-achtige wereld van Sint-Pietersplein zeer mooie momenten te beleven. Hoewel ik de voorbije periode eerder uitblonk in afwezigheid, heb ik doorheen de jaren sterk genoten van vele Friday Drinks, een PhD-weekend en andere activiteiten samen met collega-doctorandi. Hartelijk dank hiervoor aan Arthur, Bakou, Eva, Femke, Tom en talrijke anderen! Tot slot ben ik ook Melina en Sandy veel dank verschuldigd om steeds klaar te staan om op kundige wijze de praktische zaken van mijn doctoraatstraject in goede banen te leiden.

Mijn dankbaarheid reikt evenwel verder dan de vakgroep economie: ook daarbuiten zijn er meerdere mensen die mijn leven de voorbije jaren aanzienlijk hebben verrijkt. Eerst en vooral wil ik Sara bedanken. Toen we elkaar bijna vier jaar geleden ontmoetten, zat ik in het tweede jaar van mijn doctoraat, en speelde mijn werk zich grotendeels overdag en tijdens de week af. Gaandeweg zijn die werkuren gestaag uitgebreid, en het afgelopen jaar was er uiteindelijk amper nog vrije tijd naast mijn doctoraat over. Sara, dank je om dit telkens opnieuw te verdragen, en dank je om een onvoorwaardelijke steun en toeverlaat te zijn. Dank je ook voor de vele mooie belevenissen samen, gaande van een weekend tussen koeien in Friesland, roadtrips vol schnitzels en *endabrechnungen* door het Zwarte Woud, tot en met hoogstaande padelwedstrijden in Balen en rustigere momenten in bont gezelschap te Ledeberg. Weet dat ik deze momenten enorm waardeer. En hoewel het bijzonder onwaarschijnlijk is dat ze dit ooit zal lezen, wil ik ook Gaston bedanken voor haar uitstekend gezelschap tijdens het thuiswerken, en meer algemeen om een buitengewoon fijne en aanhankelijke kat te willen zijn.

Terugkerend naar het mensendom, wil ik nog graag Michiel bedanken voor de onuitputtelijke stroom van koffietjes en kombucha op zondag. De standvastigheid van deze zondagse momenten samen met Arthur bracht een erg gewaardeerde stabiliteit in mijn weekends de afgelopen jaren. Ook de gezapige citytripjes met ons drieën vormden een waar genot, en ik hoop dat er in de toekomst nog veel van zulke gezapigheid mag volgen. Daarnaast ook bedankt opnieuw aan jullie, Arthur en Michiel, en verder aan Ilja, Jozefien, Lotte, Matilde, Oscar en talrijke anderen voor de vele bijzonder vermakelijke avonden tijdens de Gentse Feesten, op oudejaar en daar tussenin.

Tot slot een speciale vermelding voor mijn familie. Aan mijn ouders, dank jullie voor jullie steun doorheen de voorbije kwarteeuw: te beginnen met het tot stand brengen van mijn geboorte, een ontegensprekelijke conditio sine qua non voor het schrijven van dit doctoraat. Maar ook daarna hebben jullie vele opofferingen gedaan die ik vermoedelijk nooit ten volle zal kunnen doorgronden. Zonder deze opofferingen was ik nooit tot dit doctoraat en het schrijven van dit dankwoord gekomen, en daarvoor ben ik zeer dankbaar. Ook was het de laatste jaren erg fijn om bij gelegenheid te kunnen terugkeren naar de ongetwijfeld meest idyllische gemeente in het Vlaamse grondgebied. Het opsnuiven van de gezonde fabriekslucht aldaar kwam mijn concentratie en welzijn ten zeerste ten goede, waarvoor dank.

Contents

List of Figures

Chapter 1

Chapter 3

List of Tables

Nederlandstalige samenvatting

Dit doctoraatsonderzoek bestudeert de vermogensongelijkheid en vermogensmobiliteit in de Verenigde Staten. Het onderzoek bestaat uit drie hoofdstukken.

In hoofdstuk 1 gebruik ik data uit de Panel Study of Income Dynamics (PSID) om intergenerationele en intragenerationele vermogensmobiliteit in de Verenigde Staten te analyseren. Methodologisch vergelijk ik de vermogensdynamiek in de PSID met die in de Survey of Consumer Finances (SCF). De PSID onderschat de ongelijkheid aan de top van de vermogensverdeling vergeleken met andere cross-sectionele datasets, maar is geschikt voor het bestuderen van vermogensmobiliteit over de volledige verdeling. Op basis van de gevalideerde data behandel ik vier onderzoeksvragen. Ten eerste: wat is het niveau van relatieve intergenerationele vermogensmobiliteit in de VS? Ten tweede: hoe groot is de intragenerationele vermogensmobiliteit? Ten derde: bestaat er interdependentie in vermogensrangen over generaties, i.e. vertonen de veranderingen in de vermogenspositie van individuen binnen een cohort gelijkenissen met die van hun ouders over dezelfde periode? En ten vierde: wat zijn de drijfveren achter intragenerationele vermogensmobiliteit? Deze analyses leveren een rijk geheel van empirische vermogensmobiliteitsmomenten op die relevant zijn voor de calibratie van heterogene agentenmodellen. Bovendien genereren ze meerdere nieuwe inzichten. Ten eerste ontwikkel ik een gradient boosting machine learning-model om vermogensrangen van gezinnen tot 1969 te benaderen. Deze benadering presteert aanzienlijk beter dan de gangbare proxies gebaseerd op woningbezit. Ik stel vast dat de gelijkenis in vermogensrang tussen (groot)ouders en (klein)kinderen toeneemt met leeftijd, dat intergenerationele mobiliteit in de tijd is afgenomen, en dat de Verenigde Staten lagere mobiliteit kennen dan de meeste andere landen waarvoor vergelijkbare data beschikbaar zijn. Ten tweede blijkt intragenerationele mobiliteit zich vooral voor te doen tussen de leeftijden van 30 en 39, en te zijn afgenomen aan de top van de verdeling. De mobiliteit is bovendien aanzienlijk lager dan in de Scandinavische landen. Verschillen in mobiliteit houden verband met schenkingen of erfenissen, ondernemerschap, arbeidsinkomen, gezondheid en niet-hypothecaire schulden. Ten derde vind ik positieve interdependentie tussen de vermogensrangpaden van individuen en die van hun ouders over dezelfde historische periode.

In hoofdstuk 2 gebruik ik opnieuw de PSID om het spaargedrag van Amerikaanse gezinnen over de vermogensverdeling te analyseren. Voor het schatten van de spaarquote maak ik gebruik van twee complementaire schattingsmethode: de cross-sectionele methode en de geaggregeerde methode. Ik identificeer vier empirische patronen. Ten eerste stijgen de totale spaarquota uit arbeidsinkomen en nieuwe middelen (flow-based saving rates) met de vermogensrang, terwijl de spaarquota uit vermogen en total middelen (stock-based saving rates) relatief stabiel blijven of slechts gematigd toenemen. Ten tweede heeft vermogensmobiliteit een belangrijke invloed op deze patronen: hoewel de impact van mobiliteit op spaarquota positief is volgens de cross-sectionele methode, is ze voor de meeste delen van de verdeling negatief binnen de geaggregeerde schattingsmethode. Dit verschil hangt samen met de manier waarop beide methodes omgaan met mobiliteit: terwijl de cross-sectionele methode alle huishoudens binnen een vermogensdeciel gelijkaardig weegt, kent de geaggregeerde methode meer gewicht toe aan huishoudens die dalen in vermogensrang. Ten derde stel ik vast dat de synthetische methode (die vaak in de literatuur wordt gebruikt) de spaarquota tot het 80e percentiel systematisch overschat, maar deze voor de top 20% onderschat. Ten vierde blijkt dat gezinnen hoger in de vermogensverdeling in toenemende mate sparen door het aanhouden van activa die in waarde stijgen. Passief sparen via intergenerationele overdrachten komt frequenter voor bij rijkere gezinnen, maar blijft beperkt in omvang. De empirische spaargedragsmomenten in hoofdstuk 2 zijn relevant voor heterogene agentenmodellen die de Amerikaanse vermogensverdeling willen repliceren.

In hoofdstuk 3 bestudeer ik de relatie tussen vermogensongelijkheid en vermogensmobiliteit, en het onderscheid tussen type dependence en scale dependence. Ten eerste ontwikkel ik een theoretisch denkkader dat beide concepten formeel definieert. Het model bevat de belangrijkste bronnen van vermogensongelijkheid die in de literatuur worden benadrukt: heterogeneit in arbeidsinkomen, spaarheterogeniteit, kapitaalinkomensrisico en het verband tussen rendement en vermogen. Ten tweede toon ik via vereenvoudigde heterogene agentenmodellen aan dat het onderscheid tussen type en scale dependence cruciaal is voor het verklaren van vermogensmobiliteit: bij eenzelfde graad van ongelijkheid genereren modellen met type dependence een hogere vermogensmobiliteit dan modellen met enkel scale dependence. Daarnaast blijkt dat het verband tussen mobiliteit en deze parameters wordt gekenmerkt door nietlineariteiten. Ten derde bouw ik een Aiyagari-Bewley-Huggett economie waarin beide vormen van dependence zijn geïntegreerd. Voor de schattig van het model ontwikkel ik een nieuwe methode die een theoretische scale-dependent functie koppelt aan een empirisch bepaalde type-dependent structuur op basis van PSID-paneldata. Het geschatte model weet zowel de vermogensongelijkheid als de vermogensmobiliteit in de Verenigde Staten in 2021 goed te reproduceren. Ten vierde voer ik verschillende counterfactual analyses uit. Die tonen aan dat realistische type dependence in spaargedrag essentieel is om vermogensmobiliteit in het model in lijn te brengen met de empirische data. Bovendien blijken ongelijkheid in arbeidsinkomen en spaargedrag de belangrijkste factoren achter persistentie in de vermogensverdeling, zowel op korte als op lange termijn. Heterogeniteit in rendementen speelt een kleinere rol. In het algemeen geldt: hoe groter de vermogensongelijkheid, hoe lager de vermogensmobiliteit.

Summary in English

This PhD dissertation studies wealth inequality and wealth mobility in the United States. It consists of three chapters.

In the first chapter, I leverage data from the Panel Study of Income Dynamics (PSID) to analyze inter- and intra-generational wealth mobility in the United States. Methodologically, I compare the wealth dynamics in the PSID to those from the Survey of Consumer Finances (SCF). The PSID underestimates top wealth inequality compared to the SCF, but this does not compromise an analysis of wealth mobility across the entire wealth distribution. I then use the validated dataset to investigate four research questions. First, what is the level of relative inter-generational wealth mobility in the United States? Second, what is the degree of intra-generational wealth mobility? Third, does there exist within-family wealth rank interdependence, i.e. do the changes in individuals' within-cohort wealth ranks relate to the within-cohort wealth rank changes of their parents over the same historical time period? Fourth, what are the sources of intra-generational wealth mobility in the United States? These four analyses provide an extensive set of empirical wealth mobility moments that are useful to the heterogeneous agent literature. Furthermore, they generate several novel findings and contributions. First, from an inter-generational (family-level) perspective, I develop a gradient boosting machine learning model to approximate household wealth ranks back to 1969. This proxy significantly outperforms housing-based proxies commonly used in the literature. Moreover, I find that wealth rank resemblance between (grand)parents and their (grand)children increases with age, that inter-generational wealth mobility has declined over time, and that the United States exhibits lower mobility compared to most other countries with available data. Second, intra-generational (individual-level) wealth mobility is concentrated between ages 30 and 39, has declined at the top of the distribution, and is substantially lower than in the Nordic countries. Diverging wealth rank trajectories are associated with variation in inter-generational transfer receipts, business ownership, labor income, health, and non-mortgage indebtedness. Third, bridging the inter- and intra-generational perspectives, I find positive interdependence between the wealth rank trajectories of individuals and those of their parents over the same historical time period.

In the second chapter, I use household-level data from the Panel Study of Income Dynamics (PSID) to provide evidence on saving behavior across the wealth (rank) distribution in the United States. I estimate saving rates across wealth deciles using two complementary approaches: the cross-sectional method and the aggregate method. I obtain four collections of stylized empirical facts. First, I find that total saving rates out of labor income and new resources rise with wealth ranks (flow-based saving rates). In contrast, total saving rates out of wealth and composite resources are roughly stable and only moderately increasing with wealth ranks (stock-based saving rates). Second, wealth (rank) mobility has a substantial im-

pact on total saving rate patterns across the wealth distribution. However, while the contribution of wealth mobility is strictly positive for the cross-sectional method, it is negative across most of the wealth distribution for the aggregate method. I show that this discrepancy relates to these methods' distinct treatment of wealth (rank) mobility: while the cross-sectional method attaches equal weight to all households in a wealth decile, the aggregate method overweighs households that display downward wealth mobility. Third, I find that the synthetic method (which is commonly used in the absence of panel data) overestimates saving rates up to the 80th percentile, while it underestimates the saving rates of the top 20%. Fourth, I demonstrate that households' reliance on capital gains rises across the wealth rank distribution: the top wealthiest households' total saving consists predominantly of saving by holding appreciating assets. Passive saving out of inter-generational transfers is more common for wealthier households, but relatively unimportant in magnitude. Many of the empirical saving behavior moments across the wealth (rank) distribution reported in Chapter 2 are likely of interest to the heterogeneous agent literature replicating the U.S. wealth distribution.

In the third chapter, I focus on the interplay between the inequality of the wealth distribution and its turnover (i.e. wealth mobility), and on the distinction between type dependence and scale dependence. First, I develop a generalized theoretical framework that provides a formal definition of type dependence and scale dependence. The theoretical framework embeds the core sources of wealth inequality underscored in the theoretical literature (labor income risk, saving rate heterogeneity, capital income risk, link between returns and wealth). Second, using a set of simplified heterogeneous agent models, I show that the type dependence versus scale dependence distinction is critical for matching wealth mobility outcomes: for identical wealth inequality outcomes, type-dependent models generate higher wealth mobility than scale-dependent ones. In addition, the relationship between wealth mobility and the scale- and type-dependent parameters is found to be characterized by non-linearities. Third, I construct an Aiyagari-Bewley-Huggett economy with type dependence and scale dependence. To estimate the type-dependent and scale-dependent parameters, I outline a novel estimation strategy that links a theoretical scale-dependent function to a corresponding, empirically-determined type-dependent structure using panel data from the PSID. The estimated model replicates well the wealth inequality and wealth mobility observed in the United States in 2021. Fourth, I conduct a series of counterfactual analyses on the estimated baseline model. These show that allowing for a realistic degree of saving ratio type dependence is critical in matching wealth mobility in the stationary model state to its empirical counterpart. Moreover, labor income inequality and saving ratio inequality emerge as the key driving forces behind agents' persistence in the wealth (rank) distribution in both the short-run and the long-run. Return heterogeneity is found to be less important. Finally, in general, there is an inverse relationship between wealth inequality and wealth mobility: higher wealth inequality coincides with lower wealth mobility.

Introduction

Since the beginning of the 1980s, wealth inequality in the United States has increased significantly, especially at the upper tail of the distribution. This evolution has triggered renewed interest in wealth inequality in popular and academic writing. It has also raised concerns about social mobility and its future trajectory. Low wealth mobility makes high wealth inequality even more problematic. Some authors have argued that, if left unaddressed, the growing inequality might induce a return to the Gilded Age − a period marked by high wealth inequality and limited wealth mobility. During that era, individuals' parental wealth positions were critical in determining their lifetime economic resources. In this context, it is unsurprising that also debates on wealth and estate taxation have become more prominent in recent academic and public discourse.

In the academic theoretical literature, a subset of the heterogeneous agent literature has focused on investigating the U.S. wealth distribution and its properties. This strand of the literature departs from Aiyagari-Bewley-Huggett economies in which agents make optimal decisions under uninsurable risk. These models have been used to address two types of research questions related to wealth inequality. On the one hand, what explains the existence of wealth inequality at a given point in time? What is the contribution of different agent heterogeneities (labor income risk, capital income risk, saving rate heterogeneity, taxation, etc.) to U.S. wealth inequality outcomes? On the other hand, what are the driving factors behind the rising U.S. wealth inequality since the beginning of the 1980s? And have policy changes contributed to the rise in wealth inequality?

Two shortcomings This theoretical literature has two key shortcomings, however. First, Aiyagari-Bewley-Huggett models tend to focus exclusively on wealth inequality, which measures the degree of dispersion in wealth levels across the population ('shortcoming 1'). Only a handful of models also include relative wealth mobility − the turnover of individuals across the wealth rank distribution − as a target variable. Currently, there therefore exists only limited insight on the interdependence between wealth inequality and wealth mobility. In principle, one expects an inverse relationship: higher wealth inequality implies larger absolute wealth differences between individuals, which renders turnover across the wealth rank distribution less likely. However, does this inverse relationship between wealth inequality and wealth mobility hold by definition? Or does the strength of the relationship interact with the underlying wealth inequality-generating channel? And what do the empirical data tell us?

The absence of wealth mobility in the Aiyagari-Bewley-Huggett models narrows the scope of their policy and societal implications. More precisely, the degree to which wealth inequality is seen as detrimental to a society is linked to the amount of wealth mobility: as mentioned, elevated wealth inequality might be less problematic if it coincides with high wealth mobility. This is because in a high wealth mobility setting, the negative externalities of high wealth inequality − e.g. political capture, social fragmentation and unrest, unequal access to healthcare, underinvestment in human capital − are likely more limited. As a result, there are good reasons to be interested not only in the effects of decision or policy-related variables on wealth inequality, but also their influence on wealth mobility. Yet, given Aiyagari-Bewley-Huggett heterogeneous agent models' almost exclusive focus on wealth inequality, they do not account for wealth mobility.

Second, to explain wealth inequality, Aiyagari-Bewley-Huggett heterogeneous agent models rely on structural agent heterogeneities in saving behavior, portfolio allocation and expected asset returns. This structural heterogeneity is introduced through either type dependence or scale dependence. On the one hand, type dependence implies that agents are ex-ante different: for instance, some agents might be more future-oriented or less risk averse than others, leading to diverging saving rates between these agents even when they have identical wealth levels. On the other hand, scale dependence means that agents are structurally heterogeneous due to differences in wealth levels: the difference between the agents arises ex-post. For example, when preferences are non-homothetic, wealthier individuals will optimally choose higher saving rates than poorer ones.

However, there currently does not even exist a formal definition of type dependence and scale dependence in the literature, and the implications of heterogeneous agent models' reliance on type dependence versus scale dependence are poorly understood ('shortcoming 2'). More precisely, while it has been demonstrated that the degree of type dependence and scale dependence in these models affects optimal wealth taxation, this has not been extended to other research questions. In addition, relating back to shortcoming 1, the degree of type dependence versus scale dependence might affect wealth mobility outcomes: insofar as agents can switch between types, type dependence introduces an additional source of randomness in Aiyagari-Bewley-Huggett models that is not present in purely scale-dependent models. Is accounting for a realistic degree of type dependence versus scale dependence then critical in producing realistic wealth mobility? And how do type and scale dependence parameters influence model outcomes for wealth inequality and wealth mobility?

Addressing these two shortcomings of Aiyagari-Bewley-Huggett models is not straightforward: there are four main challenges. First, no extensive empirical data on wealth mobility is available for the United States. As a result, the few theoretical studies for the U.S. that have integrated wealth mobility into their models have resorted to matching Nordic countries' wealth mobility outcomes, broader social mobility metrics or wealth mobility among a very small subset of the U.S. population − the Forbes 400 families. Second, for their calibration, Aiyagari-Bewley-Huggett models are highly reliant on cross-sectional moments computed across the wealth (rank) distribution using micro datasets. For the United States, it is primarily the Survey of Consumer Finances (SCF) that is being used. However, this survey does not contain a panel dimension. This makes it impossible to conduct an unbiased estimation of saving behavior at the household level, as such unbiased estimation requires as input the first difference of wealth. This is especially unfortunate as saving rate heterogeneity is found to be a critical driver of wealth inequality. Third, the absence of a formalized definition of type dependence and scale dependence in the context of Aiyagari-Bewley-Huggett models makes it challenging to investigate the impact of the type- and scale-dependent assumptions on model outcomes. Fourth, there currently exists no model-based estimation strategy for the degree of type dependence versus scale dependence.

Overview PhD The present PhD project aims to address these four challenges and hence take important steps forward in resolving the two shortcomings of Aiyagari-Bewley-Huggett heterogeneous agent models. It also contributes to answering six broader research questions that have occupied researchers in recent years. First, do wealth inequality and wealth mobility display an inverse relationship, both empirically and theoretically? Second, what are the sources of wealth inequality and wealth mobility in general, and in the U.S. in particular? Third, how high is wealth mobility in the U.S., also compared to other countries, and how did it evolve during the last decades? Fourth, how do type dependence and scale dependence parameters affect wealth inequality and wealth mobility outcomes, and how important is this distinction for matching U.S. wealth mobility outcomes? Fifth, what is the importance of type dependence and scale dependence in households' saving and portfolio allocation behavior? Sixth, does there exist a positive relationship between saving behavior and a household's position in the wealth distribution? I provide an answer to these six research questions across the three chapters of this PhD dissertation.

In a first chapter, I leverage data from the Panel Study of Income Dynamics (PSID) to analyze inter- and intra-generational wealth mobility in the United States. From a methodological perspective, I compare the wealth dynamics in the PSID to those from the Survey of Consumer Finances (SCF). The PSID underestimates top wealth inequality compared to the SCF, but this does not compromise an analysis of wealth mobility across the entire wealth distribution. I then use the validated dataset to investigate four research questions. First, what is the level of relative inter- and intra-generational wealth mobility in the United States? Second, how have both types of wealth mobility indicators evolved over time, and how do they compare to other countries with available data? Third, does there exist within-family wealth rank interdependence, i.e. do the changes in individuals' wealth ranks relate to the wealth rank changes of their parents over the same historical time period? Fourth, what are the sources of intra-generational wealth mobility in the United States? These analyses provide an extensive set of empirical wealth mobility moments that are useful to the heterogeneous agent literature on the U.S. wealth distribution. Furthermore, they generate several novel findings and contributions. First, from an inter-generational (family-level) perspective, I develop a gradient boosting machine learning model to approximate household wealth ranks back to 1969. This proxy significantly outperforms housing-based proxies commonly used in the literature. I find that wealth rank resemblance between (grand)parents and their (grand)children increases with age, that inter-generational wealth mobility has declined over time, and that the United States exhibits lower mobility compared to most other countries with available data. Second, intra-generational (individual-level) wealth mobility is concentrated between ages 30 and 39, has declined at the top of the distribution, and is substantially lower than in the Nordic countries. Diverging wealth rank trajectories are associated with variation in inter-generational transfer receipts, business ownership, labor income, health, and non-mortgage indebtedness. Third, bridging the inter- and intra-generational perspectives, I find positive interdependence between the wealth rank trajectories of individuals and those of their parents over the same historical time period.

In a second chapter, I use household-level data from the Panel Study of Income Dynamics (PSID) to provide evidence on saving behavior across the wealth (rank) distribution in the United States. I estimate saving rates across wealth deciles using two complementary approaches: the cross-sectional method and the aggregate method. I obtain four collections of stylized empirical facts. First, I find that total saving rates out of labor income and new resources rise with wealth ranks (flow-based saving rates). In contrast, total saving rates out of wealth and composite resources are roughly stable or moderately increasing with wealth ranks (stock-based saving rates). Second, wealth (rank) mobility has a substantial impact on total saving rate patterns across the wealth distribution. However, while the contribution of wealth mobility is strictly positive for the cross-sectional method, it is negative across most of the wealth distribution for the aggregate method. I show that this discrepancy relates to these methods' distinct treatment of wealth (rank) mobility: while the cross-sectional method attaches equal weight to all households in a wealth decile, the aggregate method overweighs households that display downward wealth mobility. Third, I find that the synthetic method (which is commonly used in the absence of panel data) overestimates saving rates up to the 80th percentile, while it underestimates the saving rates of the top 20%. Fourth, I demonstrate that households' reliance on capital gains rises across the wealth rank distribution: the top wealthiest households' total saving consists predominantly of saving by holding appreciating assets. Passive saving out of inter-generational transfers is more common for wealthier households, but relatively unimportant in magnitude. Many of the empirical saving behavior moments across the wealth (rank) distribution reported in Chapter 2 are likely of interest to the heterogeneous agent literature replicating the U.S. wealth distribution.

In a third chapter, I use heterogeneous agent models incorporating both type dependence and scale dependence to jointly study wealth inequality and wealth mobility in the United States. The chapter makes four core contributions to the literature. First, I outline a generalized theoretical framework that provides a formal definition of type dependence and scale dependence. The framework embeds the core sources of wealth inequality underscored in the theoretical literature (labor income risk, saving rate heterogeneity, capital income risk, link between returns and wealth). Second, using a set of simplified heterogeneous agent models, I demonstrate that the type dependence versus scale dependence distinction is critical for matching wealth mobility outcomes: for identical wealth inequality outcomes, type-dependent models generate higher wealth mobility than scale-dependent ones. Third, I construct an Aiyagari-Bewley-Huggett economy populated by households and entrepreneurs, and with both type dependence and scale dependence in parameters and decision variables. To estimate the type-dependent and scale-dependent parameters, I outline a novel estimation strategy that links a theoretical scale-dependent function to a corresponding, empirically-determined type-dependent structure using panel data from the PSID. The estimated model replicates well the wealth inequality and wealth mobility observed in the United States in 2021. Fourth, I conduct a series of counterfactual analyses on the estimated baseline model. These show that allowing for a realistic degree of saving ratio type dependence is critical in matching wealth mobility in the stationary model state to its empirical counterpart. Moreover, labor income inequality and saving ratio inequality emerge as the key driving forces behind agents' persistence in the wealth (rank) distribution in both the short-run and the long-run. Return heterogeneity is found to be less important, although this may relate to specific model assumptions. Finally, in general, there exists an inverse relationship between wealth inequality and wealth mobility: higher wealth inequality coincides with lower wealth mobility.

Data and model limitations/choices This dissertation makes various data and modeling choices that warrant some reflection. First, throughout the dissertation, I define wealth mobility as relative wealth mobility, i.e. changes in families' or individuals' positions in the wealth distribution. In line with this choice, I compute outcome metrics such as rank-rank coefficients and transition probabilities. An alternative would be to investigate absolute wealth mobility, i.e. whether families or individuals accumulate higher or lower real wealth over time. However, Chapters 1 and 2 primarily aim to provide calibration inputs to the heterogeneous agent literature. Models in this literature usually derive a stationary state where aggregate wealth growth is zero or constant. In such a setting, it makes more sense to consider relative as opposed to absolute wealth mobility.

Second, each of the three chapters of this PhD dissertation are highly reliant on the Panel Study of Income Dynamics (PSID). The main motivation for using the PSID is that it represents the only panel dataset of U.S. households over a time-span of multiple decades and covering three family generations. Unfortunately, as I outline in Chapter 1, the PSID does not capture the tail of the U.S. wealth distribution well. This means that for instance the top 10% wealth share in the PSID is lower than in the Survey of Consumer Finances (SCF) or Distributional National Accounts (DINA). For the study of relative wealth mobility − which uses the number of individuals in various wealth brackets as input − this is unproblematic. However, the under-representation of the tail does imply that one can only investigate relatively broad top wealth categories − e.g. the top 10% wealthiest. Making claims about or deriving policy implications on a finer group of top wealthiest (such as the top 1% and beyond) is impossible as these households are underrepresented or not represented in the PSID.

Third, Chapters 1 and 2 use micro data to compute empirical moments that are of interest to the Aiyagari-Bewley-Huggett heterogeneous agent literature. Furthermore, the second part of Chapter 3 constructs a heterogeneous agent model that jointly matches U.S. wealth inequality and wealth mobility moments. However, in principle, one could also use other modeling strategies to investigate the properties of the U.S. wealth distribution. A paramount alternative are agent-based (stock-flow consistent) models. Unlike heterogeneous agent models, such agent-based models depart from behavioral heuristics as opposed to intertemporal optimization processes. However, the gap between the present dissertation and the agent-based models is relatively small. That is, the empirical moments generated in Chapters 1 and 2 are also useful for agent-based models of the U.S. wealth distribution: they are definitely not exclusive to the Aiyagari-Bewley-Huggett literature. Moreover, the first part of Chapter 3 provides a generalized type versus scale dependence framework that is also relevant to agent-based models: the generalized framework does not impose any optimization procedure on the scale-dependent functions and can therefore be readily extended to an agent-based modeling context.

Chapter 1

Wealth Mobility in the United States: Empirical Evidence from the PSID [1]

This paper leverages data from the Panel Study of Income Dynamics (PSID) to analyze inter- and intra-generational wealth mobility in the United States. It provides a rich set of empirical moments that are likely to be of interest to the heterogeneous agent literature on the U.S. wealth distribution. The analysis yields several novel contributions and findings. First, from an inter-generational (family-level) perspective, I develop a gradient boosting machine learning model to approximate household wealth ranks back to 1969. This proxy significantly outperforms the housing-based proxies commonly used in the literature. I find that wealth rank resemblance between (grand)parents and their (grand)children increases with age, that inter-generational wealth mobility has declined over time, and that the United States exhibits lower mobility compared to most other countries with available data. Second, intra-generational (individual-level) wealth mobility is concentrated between ages 30 and 39, has declined at the top of the distribution, and is substantially lower than in the Nordic countries. Diverging wealth rank trajectories are associated with variation in inter-generational transfer receipts, business ownership, labor income, health, and non-mortgage indebtedness. Third, bridging the inter- and intra-generational perspectives, I find positive interdependence between the wealth rank trajectories of individuals and those of their parents over the same historical time period. Fourth, I find that the PSID underestimates top wealth inequality compared to other cross-sectional datasets. However, I demonstrate that this does not compromise an analysis of wealth mobility across the entire wealth distribution.

[1]Ghent University (email: christophe.vanlangenhove@ugent.be). I acknowledge the financial support from the Research Foundation Flanders (FWO, project number: 11I5324N). I would like to thank Arthur Apostel, Paula Gobbi, Freddy Heylen, Selma Malmberg, Marthe Mareels, Ive Marx, Yasin Kürşat Önder, Gert Peersman, Morten O. Ravn, Alberto Russo and Dirk Van de gaer for their valuable comments on this working paper. I would also like to thank the participants at the internal seminar of the Department of Economics (UGent), the ESPANet Early Career Researchers Day (ESPANet/KULeuven), the 20th PSE Doctorissimes Conference (Paris School of Economics), the Inequalities & Opportunities Conference (University of Bari) and the 11th ECINEQ Conference (ECINEQ/World Bank) for their discussion and feedback.

1 Introduction

Over the past decade, empirical research on wealth inequality has expanded considerably, both for the United States and at an international level (e.g. Saez & Zucman, 2016; Smith et al., 2023; Zucman, 2019). In contrast, studies investigating relative inter- or intra-generational wealth mobility − changes in families' or individuals' wealth ranks across and within generations − remain hard to come by. This is unfortunate for two main reasons.

First, insights into inter- and intra-generational wealth mobility may inform academic and popular debates on estate taxation, wealth taxation and the economics of opportunity. For example, high wealth inequality may be viewed as less detrimental to society if it coincides with high turnover at the top of the wealth distribution (and vice versa). This is because in such a high wealth mobility setting, the negative externalities of high wealth inequality − e.g. political capture and weakening of political institutions, social fragmentation and unrest, unequal access to healthcare, underinvestment in human capital − are likely more limited. In a similar spirit, wealth mobility outcomes may serve as a key input in shaping and challenging cultural narratives on the American dream and the United States as land of opportunity.

Second, a theoretical literature on heterogeneous agent macro models uses wealth inequality as key outcome variable, while disregarding wealth mobility (e.g. De Nardi & Fella, 2017; Hubmer et al., 2021; Xavier, 2021). Such setting does not allow to take an explicit stance on the importance of type versus scale dependence: an unequal stationary wealth distribution could be generated by ex-ante differences in discount factors or risk aversion across agents (type dependence), or by ex-post heterogeneity in response to different wealth levels (scale dependence). Calibrating these models jointly to inequality and mobility moments could be a starting point in solving this type versus scale dependence puzzle (e.g. Van Langenhove, 2025). While there exists some theoretical work that incorporates wealth mobility outcomes (Atkeson & Irie, 2022; Benhabib et al., 2019; Fernholz, 2016; Gomez, 2023), this literature is constrained by the unavailability of wealth mobility data for the United States.

Research questions To address the scarcity of wealth mobility data over recent decades, this paper leverages the Panel Study of Income Dynamics (PSID) to provide evidence on inter- and intra-generational wealth mobility outcomes in the United States. Many of the generated empirical moments are likely of interest to the heterogeneous agent macro literature on the U.S. wealth distribution. Three research questions are addressed.

First, from an inter-generational (family-level) perspective, I investigate how the within-cohort wealth ranks of individuals compare to the within-cohort wealth ranks of their parents (at identical points in their lifecycles) and grandparents (at different points in their lifecycles, due to data limitations). Such static comparison of wealth ranks across generations is the approach commonly taken in the literature (e.g. Adermon et al., 2018; Boserup et al., 2017; Pfeffer & Killewald, 2018; Siminsky & Yu, 2022).

Second, from an intra-generational (individual-level) perspective, the paper analyzes the within-cohort wealth rank changes of individuals over their lifecycle. For example, given one's within-cohort wealth rank at the age of 30 or 55, what is the probability of this individual moving upward or downward the wealth rank distribution as it progresses through working life or older age? And how do the observed wealth rank trajectories relate to individuals' inter-generational transfer receipts and socio-economic characteristics? I investigate these questions for both working life (ages 30-54) and older age (ages 55-74).

Third, bridging the inter-generational (family-level) and intra-generational (individual-level) perspectives, this paper is the first to investigate the inter-dependence between individuals' wealth rank trajectories and those of family members (i.e. within-family inter-dependence in intra-generational wealth mobility). That is, does there exist covariance between the changes in individuals' wealth ranks and those of their parents over the same historical time period? I find that there does exist covariance. This suggests the presence of altruism across generations and the exposure to identical sources of idiosyncratic risk across family members.

Related literature & contributions This paper contributes to four strands of the literature. I discuss each of these strands in what follows.

Inter-generational (family-level) wealth mobility I add to the literature studying inter-generational wealth mobility in the United States (Charles & Hurst, 2003; Conley & Glauber, 2008; Menchik, 1979; Pfeffer & Killewald, 2018; Siminsky & Yu, 2022). Among these, only Pfeffer & Killewald (2018) extend their analysis to mobility across three generations (grandparents-grandchildren). For the Nordic countries, Adermon et al. (2018), Black et al. (2020), Boserup et al. (2017) and Fagereng et al. (2021) investigate inter-generational wealth mobility for Norway, Denmark and Sweden respectively. Finally, Gregg & Kanabar (2023) and Levell & Sturrock (2023) produce evidence on inter-generational wealth mobility for the United Kingdom, while Siminsky & Yu (2022) do so for Australia. I make three contributions to this inter-generational wealth mobility literature.

First, from a methodological perspective, I develop a novel method to approximate wealth ranks whenever direct wealth data is unavailable. While targeted at the PSID, this method is likely generalizable to other datasets. In the PSID, questions on asset holdings and debt levels date back only to 1984. However, data on main housing values and rental payments are available as early as 1969. A common strategy is then to assume that renters have zero wealth and to approximate total household wealth by main housing values (Pfeffer & Killewald, 2018, but also e.g. Chetty et al., 2020). Instead, I develop a gradient-boosting (GB) machine learning model trained on post-1984 data. This ML-model incorporates additional socio-economic variables from the PSID as input variables, and significantly outperforms the housing proxies in predicting household wealth levels out-of-sample. However, these proxies underestimate the actual degree of intra-generational wealth mobility during working life, as well as the actual degree of inter-generational wealth mobility.

Second, existing studies generally include an age control in the standard rank-rank regressions, hence imposing a functional form on the relationship between age and wealth persistence. On the contrary, I investigate the impact of lifecycle bias on estimated wealth mobility outcomes without imposing a functional form by computing wealth mobility moments across distinct lifecycle stages. I find that age matters for inter-generational (family-level) wealth mobility outcomes: wealth rank resemblance between parents and their children rises significantly with parents' and children's age (parent-child lifecycle bias). In addition, wealth rank resemblance between grandparents and their grandchildren is higher when grandchildren are older than 35 years (grandchild lifecycle bias). Finally, I show that grandparent-grandchild mobility (three generations) exceeds parent-child mobility (two generations). The effect is non-linear over the wealth distribution, however: while mobility at the top is significantly higher across three versus two generations, the difference in mobility at the bottom is comparatively weaker.

Third, this paper is the first to investigate the time trend and cross-country differences in inter-generational wealth mobility in the United States across the entire wealth distribution. That is, has the increase over time in overall wealth inequality (e.g. Saez & Zucman, 2016; Smith et al., 2023; Zucman, 2019) coincided with changes in inter-generational wealth mobility? And how does wealth mobility in the United States compare to other developed countries for which data is available? On the one hand, I find that inter-generational wealth mobility has declined over the past decades. This complements evidence on the decline in wealth mobility among the top 400 wealthiest families in the United States (Fernholz & Hagler, 2023). It also aligns with the decline in inter-generational wealth mobility established for Sweden (Adermon et al., 2018) and the United Kingdom (Gregg & Kanabar, 2023; Levell & Sturrock, 2023). On the other hand, I find that wealth mobility in the United States is lower compared to most other countries with available data.

Intra-generational (individual-level) wealth mobility This paper also contributes to the literature on intra-generational wealth mobility. For the United States, this intra-generational literature is currently limited to Conley & Glauber (2008), Klevmarken et al. (2003) and Shiro et al. (2022)[2]. For the Nordic countries, Audoly et al. (2024) and Hubmer et al. (2024) analyze intra-generational mobility for Norway. I contribute to the intra-generational wealth mobility literature along three lines.

First, in line with the inter-generational mobility analysis, I explicitly investigate the relationship between age and intra-generational wealth mobility outcomes. More precisely, I am the first to investigate intra-generational mobility over roughly the entire lifecycle, from ages 30-34 to ages 70-74. I find that wealth mobility during working life (ages 30-54) exceeds the mobility observed during older age (ages 55-74). Moreover, timing effects indicate that the majority

[2] In addition, Kuhn et al. (2020) and Kalsi & Ward (2025) conduct limited intra-generational wealth mobility analyses using the PSID. These analyses serve as a robustness to their baseline results.

of intra-generational (individual-level) wealth mobility occurs early in working life, between ages 30 and 39.

Second, similar to the inter-generational analysis, I investigate the time trend in intra-generational wealth mobility and compare the wealth mobility outcomes to other countries with available data. On the one hand, this paper is the first to show that the increase in within-cohort life-cycle wealth inequality over the past decades has coincided with a decline in overall intra-generational wealth mobility. This decline in wealth mobility masks opposing effects at the bottom and at the top of the wealth distribution, however: while intra-generational wealth mobility at the top has declined strongly, mobility at the bottom has in fact risen slightly. On the other hand, intra-generational wealth mobility in the United States is substantially lower compared to Norway.

Third, this paper explores the sources of intra-generational wealth mobility in the United States. The analysis indicates that consolidation at the top (bottom) of the wealth distribution is associated with the most substantial (an absence of) inter-vivos transfer and inheritance receipts. However, even for the wealthiest, these receipts make up only a limited fraction of their lifetime resources. Furthermore, business ownership is linked with consolidation at the top and downward wealth mobility, while its association with upward wealth mobility is inconclusive. Last, consolidation at the bottom and downward mobility to the bottom are associated with low labor income, poor and deteriorating health, elevated non-mortgage indebtedness and modest asset ownership. Instead, at the top, labor income and asset ownership are relatively high.

Within-family wealth rank inter-dependence This paper additionally opens a novel literature by bridging the inter- and intra-generational wealth mobility perspectives: I investigate the inter-dependence between individuals' wealth rank trajectories and those of their parents over the same historical time period. I find that individuals who experience upward or downward mobility within their cohort tend to have parents who followed similar wealth rank trajectories within their own cohort. At the same time, individuals that consolidate their position at the top are the most common to have wealthy parents. Such wealth rank inter-dependence within families could reflect two channels. On the one hand, it suggests the presence of altruism across generations. On the other hand, it may be that parents and their children face exposure to identical sources of idiosyncratic risk (e.g. business risk, occupational specializations, housing areas).

PSID-validation Last but not least, this paper adds to a literature spanning Cooper et al. (2019), Insolera et al. (2021) and Pfeffer et al. (2016) by investigating the appropriateness of the Panel Study of Income Dynamics (PSID) for studying U.S. wealth inequality and wealth mobility. More precisely, I harmonize the PSID-data and validate the data by contrasting aggregate wealth and wealth inequality outcomes in the PSID to the outcomes in the top-wealth-adjusted Survey of Consumer Finances (SCF). Compared to existing studies, I validate these outcomes over time rather than for a specific year. Two key findings persist. On the one hand, the PSID underestimates most aggregate wealth components relative to the SCF, but accurately captures their time trends. On the other hand, wealth share trajectories in the PSID closely align with those from the SCF, notwithstanding an underestimation of the top 10% wealth share by moderately over 10%-points in the PSID. Regardless of this top-wealth bias, I argue that the PSID can be effectively used to study wealth-related questions. This is particularly true for the study of wealth mobility (compared to wealth inequality) given that wealth mobility metrics employ the number of households across the wealth distribution as calculation inputs (rather than their wealth levels).

Roadmap Section 2 introduces a theoretical framework to understand the driving forces behind inter- and intra-generational wealth mobility. Section 3 summarizes the data and empirical methods used, building on the detailed exposition provided in Appendices A to F. Section 4 compares the wealth rank outcomes of individuals to those of their parents at the same lifecycle stage and to those of their grandparents at different lifecycle stages. Section 5 presents the results of the intra-generational wealth mobility analyses during working life and older age. Section 6 investigates the inter-dependence between the within-cohort wealth rank trajectories of individuals and those of their parents. Section 7 reports composition statistics for groups and clusters of individuals with distinct wealth rank trajectories, shedding exploratory light on potential channels of wealth mobility. Section 8 concludes.

2 Wealth inequality & mobility: framework & channels

2.1 Framework

To define wealth mobility outcomes and their sources, let us consider the following simplified budget constraint for an individual j:

$$w_j(t+1) = \left[1 - \theta_j(t)\right] \left[w_j(t)(1 + \alpha_j^a(t)r^a(t) + \alpha_j^i(t)r_j^i(t)) + y_j(t) + m_j(t) + \mu_j(t)\right] \tag{1}$$

where w_j denotes the individual's wealth level, θ_j its consumption rate out of total resources available, α_j^a and r^a the allocation to and return on aggregate investment risk, α_j^i and r_j^i the allocation to and return on idiosyncratic investment risk, y_j labor income, m_j net receipts of inter-vivos transfers and inheritances, and μ_j a residual variable that captures household formation effects. I assume that the return on the riskless asset equals zero. In addition, I abstract from taxation for simplicity. Furthermore, an individual j is assumed to belong to a family, which consists of individuals across multiple generations. In this paper, an individual's family is equaled to its parents and grandparents (so that siblings and great-grandparents are excluded).

θ_j, α_j^a and α_j^i constitute the behavioral (or policy) variables of the individual. The level of any $z_j \in \{\theta_j, \alpha_j^a, \alpha_j^i\}$ is assumed to be determined by an interplay of type- and scale-dependence. Formally:

$$z_j(t) = \bar{z}\left[\kappa_j(t)\right] + \epsilon_j(t) \tag{2}$$

where $\bar{z}\left[\kappa_j\right]$ denotes the level of parameter z specific to wealth rank κ, and ϵ_j the individual-specific variation around $\bar{z}$. ϵ_j is defined as type dependence, whereas $\bar{z}\left[\kappa_j\right]$ represents scale dependence[3]. Specifically:

- Type dependence captures structural parameter heterogeneity across individuals, or − equivalently − ex-ante heterogeneity. For example, despite having near-identical wealth ranks, individual a may display higher saving rates or higher aggregate or idiosyncratic investment risk allocations compared to individual b. This could follow from structural heterogeneity in preferences, cultural attitudes or social norms. If long-lasting, these favorable characteristics of individual a are expected to generate higher wealth accumulation over time for individual a relative to individual b.
- Scale dependence captures the change in parameter z in response to variation in an individual's wealth rank κ_j − or ex-post heterogeneity. Suppose individuals c and d are initially identical in terms of wealth levels, labor income and type-dependent parameter levels. However, individual c experiences a positive idiosyncratic shock, e.g. an increase in its labor income or the receipt of an inter-vivos transfer. As the wealth level of individual c rises, its aggregate risk allocation or saving rate may increase as a result of behavioral (non-homothetic preferences) or institutional determinants (higher expected returns thanks to superior investment fund access).

2.2 Wealth mobility channels

The budget constraint in Equation 1 allows to differentiate between five channels of inter-generational wealth transmission, as well as four channels of intra-generational wealth mobility. While this paper does not quantify the importance of these channels, they aid the interpretation of the reported wealth mobility outcomes later in the paper.

Inter-generational channels There are five channels of inter-generational (family-level) wealth transmission. First, an individual may receive inter-vivos transfers or inheritances from its parents or grandparents. This introduces a positive association between an individual's wealth rank posterior to the transfer receipt and the wealth ranks of the parents or grandparents prior to their transfer or death. Moreover, wealthy parents may be more likely to finance consumption expenditures of their children (inter-vivos transfers in kind). Second, there exists strong

[3]There may also be a lifecycle bias underlying the parameter z_j. In this paper, I abstract from this bias for simplicity.

evidence that parental wealth positively affects labor market outcomes as a result of genetic, social, education and network effects (e.g. Holmberg et al., 2024; Karagiannaki, 2017; Pfeffer, 2018; Staiger, 2023). As high labor income is associated with higher wealth accumulation over the lifecycle, this creates a positive association between wealth ranks across generations. Third, investment in high-return assets (such as housing or business) may require substantial upfront expenditures, meaning that individuals might experience borrowing constraints (e.g. Lee et al., 2020). Access to parental or grandparental wealth could provide the required collateral to circumvent these constraints and allow for higher wealth accumulation over the lifecycle. Fourth, the type-dependent level of an individual's parameters may be influenced by the type-dependent levels of its parents or grandparents. For example, children could inherit saving and risk-taking behavior from their parents with a non-random probability as a result of genetic or social effects (e.g. Black et al., 2020; Fagereng et al., 2021; Lindquist et al., 2015). Fifth, wealth levels may play a critical role in social network formation. If children have access to the social networks of their parents or grandparents, individuals from high-wealth families might be more likely to create a household with individuals from similar-wealth families (e.g. Charles et al., 2013; Wagner et al., 2020; Fagereng et al., 2022).

Intra-generational channels In addition to the sources of inter-generational wealth transmission, I distinguish between four channels of intra-generational (individual-level) wealth mobility. First, diverging idiosyncratic risk realizations may generate individual-level wealth mobility over time. In the framework of Section 2.1, there exist two sources of idiosyncratic risk: labor income and investment idiosyncratic risk (which may include the business-specific risk in a non-Markovian portfolio or the idiosyncratic risk to housing). Second, individuals are type-dependent in behavioral parameters. Insofar as an individual's wealth-rank neighbors have dissimilar type-dependent levels, the individual is expected to experience downward or upward mobility over time even when facing identical aggregate and idiosyncratic risk realizations to its wealth rank neighbors. Third, an individual may experience wealth mobility as a result of inter-generational transfer receipts that diverge from those received by its wealth rank neighbors. Fourth, an individual can move up or down the wealth distribution through its own and its wealth rank neighbors' choices of relationship or marriage partners (household formation).

Three remarks are in place. First, the presence of scale dependence widens absolute differences in wealth levels over time and hence generates a more unequal stationary wealth distribution. However, it does not trigger changes in individuals' wealth ranks, and therefore does not constitute a distinct source of wealth mobility. Second, there may exist type and scale dependence also in individuals' non-behavioral variables (such as rates of return). Third, type and scale dependence might be present in individuals' non-financial variables affecting the idiosyncratic risk realizations and behavioral variables from Equation 1. A prime example of such non-financial variable is health (e.g. De Nardi et al., 2024; Mahler & Yum, 2024). Specifically, an individual's health may affect its labor income outcomes, as well as its saving rates or risky asset allocations. At given wealth ranks, some individuals face better health than others due to genetics or health habits over the lifecycle (a type dependence). At the same time, health may be directly linked with wealth due to the access to healthcare facilities that wealth buffers enable (a scale dependence).

3 Data & methods

3.1 Data

This paper uses data from the Panel Study of Income Dynamics (PSID), which was conducted annually between 1968 and 1997 and bi-annually from 1999 to 2021. All survey waves include data on family units' gross main housing value, gross main housing mortgage debt and rental payments. The waves in 1984, 1989, 1994 and 1999-2021 add questions about other assets and debts, which allows to define wealth as the total of all asset categories minus the total of all debt categories. In the remainder of this paper, I refer to a full sample Ω (spanning years 1969 to 2021) and a reduced sample Ψ (which contains only the years where wealth-related questions were inquired, from 1984 onwards). A detailed description of the dataset is provided in Appendix A.

3.2 Methodological contributions

The paper has two methodological contributions, which are placed in Appendices. They are discussed in detail in the Appendices. Here, I highlight the main elements.

First, I harmonize and validate the PSID-dataset for wealth mobility research (Appendices A and B). Appendix A provides a detailed description and validation of the PSID-data, while Appendix B harmonizes the wealth variables and reports variable-specific outliers. A key concern related to the PSID involves its inaccurate representation of the top wealthiest. The validation exercise in Appendix A underscores this concern: the PSID underestimates the top 10% wealth share by slightly over 10%-points compared to the top-wealth-adjusted Survey of Consumer Finances (SCF). This relative error becomes larger the smaller the group of top wealthiest households under consideration. However, there are two reasons why the PSID can effectively be used to study wealth-related topics, and wealth mobility in particular. On the one hand, wealth mobility metrics use the number of households across the wealth distribution as calculation inputs. If one defines top wealth broadly (e.g. the top 10%), excluding a small number of high-wealth households therefore has a much more limited impact on these wealth mobility measures than for wealth inequality metrics (which instead rely on total wealth owned by households). On the other hand, despite its underestimation of top wealth, the PSID does accurately capture the trends in wealth inequality and accumulation, so that the underestimation bias is time-invariant (see Appendix A).

Second, data on wealth w is available over the reduced sample Ψ, which begins only in 1984. A common approach in the literature has therefore been to approximate wealth levels prior to 1984 based on main housing values and rental payments ('housing proxies'), which are available over the full sample Ω (e.g. Chetty et al., 2020; Pfeffer & Killewald, 2018). In Appendix C, I instead construct a gradient boosting (GB) ML-model to proxy wealth levels and ranks prior to 1984. It uses additional household-level socio-economic data (e.g. labor income, capital income, household size, household status, age, business ownership, health) available in the PSID. I demonstrate that such an ML-model displays superior performance compared to the housing proxies in predicting household wealth levels in a testing set from the post-1984 sample. This outperformance is robust to different performance metrics, to varying time periods and holds both at the bottom and at the top of the wealth distribution. However, despite the outperformance to housing proxies, the ML-proxy still misallocates a significant fraction of households: in a given year, approximately 8% of households have their wealth ranks misallocated by over 25 rank units.

3.3 Empirical strategy

The empirical strategy can be described in three steps. A detailed explanation of these steps is provided in Appendix D. First, I convert household-level to individual-level data based on the individual's household status (single, relationship, marriage)[4]. Second, individuals are allocated to birth cohorts (defined over ten-year intervals) and observations for all variables are summarized by taking the median per lifecycle stage (spanning ages 30-34 to ages 75+). This aggregation over multiple years is a common approach in the mobility literature (e.g. Boserup et al., 2017; Gregg & Kanabar, 2023). It has several advantages: it smooths out remaining transitory measurement errors and survey non-response, minimizes noise from household transitions, and circumvents the non-uniform timing of PSID survey waves. Third, I define individual-level within-cohort wealth ranks (with maximum ranks normalized to 100), which constitute the principal inputs in the wealth mobility analyses in this paper. The usage of ranks (as opposed to for instance log wealth) has the advantage of dealing with zero and negative observations appropriately and being robust to data transformations (e.g. Boserup et al., 2017).

Two within-cohort wealth ranks series are defined. On the one hand, κ^{Ψ} is computed from actual wealth data in the reduced sample (Ψ). On the other hand, $\hat{\kappa}^{\Omega}$ is based on proxied wealth data in the full sample (Ω). Mobility outcomes based on these two benchmark series are reported in the main text. As a robustness, I have additionally computed mobility outcomes using $\hat{\kappa}^{\Psi}$, which is calculated from proxied wealth in the reduced sample. Across all wealth mobility analyses in this paper, $\hat{\kappa}^{\Psi}$ yields mobility outcomes that align very closely with those based on $\hat{\kappa}^{\Omega}$. Consequently, differences in outcomes between κ^{Ψ} and $\hat{\kappa}^{\Omega}$ are due to the usage

[4]Given that individuals may switch households over time, we are ultimately interested in individuals' wealth rank trajectories rather than those of households.

of a different measure (κ versus $\hat{\kappa}$) rather than differences in underlying samples (Ψ versus Ω). Throughout this paper, it will become clear that the proxy wealth series $\hat{\kappa}$ underestimate the actual degree of inter- and intra-generational wealth mobility (based on κ).

3.4 Outcome metrics

The inter- and intra-generational analyses rely on a comprehensive set of inequality and mobility metrics, defined in detail in Appendix E. These measures allow to study overall inter- and intra-generational wealth inequality and mobility, as well as mobility at the bottom and top of the wealth distribution. In what follows, I provide an overview of the mobility and inequality metrics.

To study overall wealth mobility (across the entire wealth distribution), I compute rank-rank coefficients β. These regress within-cohort wealth ranks at some final lifecycle stage on within-cohort wealth ranks at some initial lifecycle stage using Ordinary Least Squares (OLS). This is a common approach in the mobility literature (e.g. Deutscher & Mazumder, 2021; Mogstad & Torsvik, 2023). As a robustness, I have also computed overall mobility outcomes based on a squared mobility metric that attaches greater weight to large wealth rank changes (see Appendix E). It produces the same conclusions as the rank-rank coefficients, so that I do not report this squared mobility metric in the main text.

To investigate mobility at both the bottom and top of the wealth distribution, I primarily use transition probabilities. These measure the ex-ante probability of moving to a specific wealth bin from a given starting point, as well as the ex-post probability of originating from a specific bin given a final position. In addition, I categorize families or individuals into discretionary groups and hierarchical clusters. More precisely:

- Discretionary groups: families or individuals with distinct wealth rank combinations or trajectories are allocated to a discretionary group. At the bottom, (i) the steady poor include the families or individuals that start and end in the bottom 20%, (ii) the past poor those that display upward wealth mobility to the top 50% originating from the bottom 20%, and (iii) the new poor start off in the top 50% but experience downward mobility to the bottom 20%. At the top, (iv) the steady wealthy start and end in the top 10%, (v) the past wealthy begin in the top 10% but display downward mobility to the bottom 70%, and (vi) the new wealthy experience upward mobility to the top 10% after starting off in the bottom 70%.
- Hierarchical clusters[5]: individuals are grouped into clusters based on their wealth rank trajectories over the lifecycle, in line with Audoly et al. (2024). These provide complementary evidence to the discretionary groups: while the discretionary groups capture

[5]This hierarchical clustering procedure is applied only in the intra-generational analysis as it requires wealth rank trajectories (rather than combinations) as input.

only the subset of individuals with the most extreme wealth rank trajectories, the clusters group every single individual in the sample into a distinct cluster. The clusters therefore provide insight into how broad-based the overall wealth mobility is. A mathematical derivation of the clustering algorithm is provided in the Online Supplement.

In addition, for the intra-generational analysis, I define variables that capture within-cohort wealth inequality and accumulation. These encompass within-cohort wealth shares, wealth to average labor income ratios, and the proportion of low- and high wealth individuals across the lifecycle. The latter are defined as individuals with wealth levels below annual average labor income (low wealth) and in excess of twenty times annual average labor income (high wealth).

4 Inter-generational family-level mobility

This section investigates wealth mobility within families from a static perspective. Wealth rank outcomes of individuals are compared to those of their parents at identical lifecycle stages and to those of their grandparents at different lifecycle stages (since data is unavailable at the same lifecycle stages). Section 4.1 provides the outcomes across two generations (parent-child), while Section 4.2 produces the results across three generations (grandparent-grandchild).

I restrict the sample to (grand)children's birth cohorts that have at a minimum 750 observations in at least one (grand)parent-(grand)child lifecycle stage combination[6]. The analyses below report rank-rank coefficients, as well as transition probabilities across two and three generations. Crucially, the rank-rank regressions do not include age controls. Instead, they are computed across different (grand)parent-(grand)child lifecycle combinations to more distinctly quantify the impact of (grand)parent and (grand)child age on rank-rank coefficient estimates.

Previewing the results, parent-children wealth rank resemblance is found to increase with parent-child age (parent-child lifecycle bias), while wealth rank resemblance between grandparents and their grandchildren is higher when grandchildren are older than 35 years (grandchild lifecycle bias). In addition to these timing effects, two-generational wealth mobility has declined over time (specifically between ages 35-44), and three-generational wealth mobility exceeds two-generational wealth mobility. The latter effect is non-linear, however: mobility at the top is significantly higher across three generations than across two, while the difference in mobility at the bottom is comparatively weaker.

[6] Letting PC and GC denote parent-child and grandparental-grandchild linkages, the (grand)children's birth cohorts that fulfill the minimum observation criterion include:

$$Y^{\mathrm{PC}} = \{P^{\mathrm{PC}}, 1936\text{–}45, 1946\text{–}55,\ 1956\text{–}65,\ 1966\text{–}75,\ 1976\text{–}85\}$$
$$Y^{\mathrm{GC}} = \{P^{\mathrm{GC}}, 1956\text{–}65,\ 1966\text{–}75,\ 1976\text{–}85\}$$

Here, P^{PC} and P^{GC} denote the pooled dataset in the two- and three-generational samples. These contain the observations across all other selected birth cohorts.

4.1 Inter-generational mobility across two generations

Section 4.1 evaluates inter-generational parent-child mobility (across two generations). Using rank-rank coefficient estimates β (Figure 1), I quantify the degree of mobility across the entire wealth distribution (overall mobility) as well as the impact of the parent-child lifecycle bias on the estimations. I subsequently investigate the mobility at the bottom and top of the wealth distribution using ex-ante and ex-post transition matrices $T_{\mathrm{EA}}(a)$ and $T_{\mathrm{EP}}(a)$ (Figures 2 and 3).

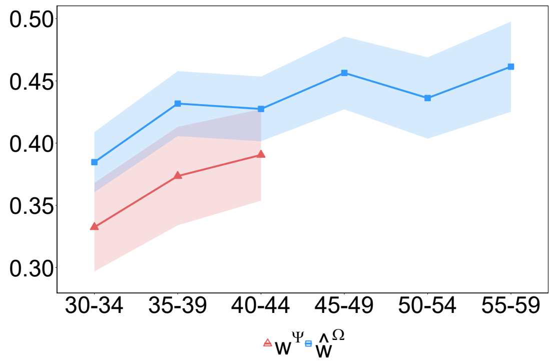
Overall mobility The analysis of overall mobility across two generations generates three key findings (Figure 1). First, the estimated parent-child rank-rank coefficients β range from 0.33 to 0.39 (based on actual wealth) and from 0.39 to 0.46 (based on proxy wealth). Second, the resemblance between parents and their children in terms of within-cohort wealth ranks is significantly higher at ages 35-39 compared to ages 30-34. At later ages, the two-generational resemblance increases further, peaking between ages 55 and 59 (with no data available for later stages). This follows from the upward-sloping profile of the β-values, and indicates the presence of a parent-child lifecycle bias in two-generational wealth mobility outcomes. Third, although they accurately capture age dynamics, the proxy wealth ranks underestimate the degree of two-generational wealth mobility, contrary to the claim in Pfeffer & Killewald (2018).

The increased parent-child resemblance with age (parent-child lifecycle bias) may be attributed to two mechanisms. First, the fraction of individuals that receives an inter-vivos transfer or inheritance increases strongly during working life, from around 10% at ages 30-34 to close to 40% by ages 50-54 (Appendix G). This is likely to generate greater alignment between parent and child within-cohort wealth ranks as children's working life progresses (channel 1 in Section 2.3). Second, individuals may have inherited labor market outcomes or type-dependent parameter levels from their parents, or could have married household partners with similar parental wealth. These channels (channels 2-5 in Section 2.3) increasingly affect individuals' wealth levels as their lifecycle progresses and are therefore expected to generate greater parent-child wealth rank resemblance after some time.

Literature comparison How do the β-estimates in Figure 1 compare to those reported in the literature? In what follows, I compare the findings to existing estimates for the United States, the Nordic countries, Australia and the United Kingdom.

For the United States, my estimated values for actual wealth (0.33 to 0.39) are slightly below those of Pfeffer & Killewald (2018): these authors find a two-generational rank-rank coefficient of 0.39 using PSID-data until 2015. Moreover, based on a PSID-sample until 2017, Siminsky & Yu (2022) produce a β-estimate of 0.34, which is similar to the estimates produced in this paper. Both Pfeffer & Killewald (2018) and Siminsky & Yu (2022) use actual wealth series in a regression where parents' and children's ages are included as control variables. Finally, using log-log regressions, Charles & Hurst (2003) and Conley & Glauber (2008) find wealth rank

Figure 1: Two-generational rank-rank coefficients β for parents and children at identical lifecycle stages for the pooled dataset.



Note: this figure reports rank-rank coefficients β computed from parents' and children's within-cohort wealth ranks. These are compared at identical lifecycle stages (shown on the x-axis). Coefficients are reported based on actual wealth if available (from w^{Ψ}) and proxy wealth (from $\hat{w}^{\Psi}$). In the rank-rank regressions, children's wealth ranks are the dependent variable. The usage of the pooled dataset indicates that individuals across all selected birth cohorts are included in the sample. The shaded areas display the 95% confidence intervals.

coefficient estimates of 0.37 and 0.28 respectively based on two-generational PSID-samples that include relatively young children.

Two-generational wealth mobility in the United States is lower than in Norway, Denmark and Australia, but similar to the United Kingdom and Sweden. Specifically, Boserup et al. (2017) report a wealth rank coefficient of 0.27 (at age 45) for Denmark, while Fagereng et al. (2021) and Audoly et al. (2024) respectively find a rank-rank coefficient of 0.17 (regression with age controls) and a rank-rank coefficient of 0.25 (at parent-child age 55) for Norway. Moreover, Siminsky & Yu (2022) produce a β-estimate of 0.25 (regression with age controls) for Australia. These estimates lay significantly below my estimates for the United States (0.33 to 0.39). By contrast, for the United Kingdom, Gregg & Kanabar (2023) and Levell & Sturrock (2023) produce rank-rank coefficients of 0.30 and 0.36 respectively (regressions with age controls). In addition, for Sweden, Adermon et al. (2018) observe β-estimates between 0.30 and 0.39 (regression with age controls), while Black et al. (2020) produce a coefficient of 0.35 (regression with age controls). All these studies rely on actual wealth data as opposed to housing or machine learning wealth proxies.

The parent-child lifecycle bias in two-generational mobility is well established in the literature. For the United States, Pfeffer & Killewald (2018) find that two-generational wealth rank resemblance increases with parent-child age: their estimated two-generational rank-rank coefficient rises from 0.33 at ages 25-34 to 0.44 at ages 55-64. Moreover, regressing child wealth ranks between ages 20 and 45 on parent wealth ranks at age 45 for Denmark, Boserup et al. (2017) find a U-shaped pattern that bottoms at children's mid-twenties. Likewise, Audoly et al. (2024) regress child wealth ranks from ages 30 to 55 on parent wealth ranks at age 55 (on average)

Table 1: Two-generational rank-rank coefficients β across children's birth cohorts $\in \Upsilon^{\mathbf{PC}}$ for parents and children at identical lifecycle stages.

Variable	Stage	1946–55	1956–65	1966–75	1976–85	1986–95	Pooled
κ^{Ψ}	30–34	-	-	-	0.35	0.32	0.33
					(0.02)	(0.04)	(0.02)
	35–39	-	-	0.33	0.40	-	0.37
				(0.03)	(0.03)		(0.02)
	40–44	-	-	0.35	0.46	-	0.39
				(0.03)	(0.03)		(0.02)
$\hat{\kappa}^{\Omega}$	30–34	-	0.39	0.39	0.38	0.40	0.39
			(0.03)	(0.02)	(0.02)	(0.03)	(0.01)
	35–39	-	0.38	0.44	0.45	-	0.43
			(0.03)	(0.02)	(0.02)		(0.01)
	40–44	0.43	0.37	0.43	0.51	-	0.43
		(0.04)	(0.02)	(0.02)	(0.03)		(0.01)
	45–49	0.48	0.44	0.47	-	-	0.46
		(0.03)	(0.02)	(0.03)			(0.02)
	50–54	0.42	0.41	-	-	-	0.44
		(0.03)	(0.02)				(0.02)
	55–59	0.48	0.46	-	-	-	0.46
		(0.03)	(0.03)				(0.02)

Note: This table reports rank-rank coefficients for parents' and children's within-cohort wealth ranks across all children's birth cohorts. These are compared at identical lifecycle stages (ranging from 30–34 to 55–59) based on actual wealth ranks κ^{Ψ} and proxy wealth ranks $\hat{\kappa}^{\Omega}$. In the rank-rank regression, children's wealth ranks are the dependent variable. The rank-rank coefficients are calculated only when an birth cohort has at the minimum 750 observations for the respective variable (as specified in the introduction to Section 4). Standard errors are shown in parentheses below the respective rank-rank coefficient estimate.

and report a positive linear relationship between child age and their β-estimates for Norway. Finally, Adermon et al. (2018) and Siminsky & Yu (2022) find evidence on two-generational age effects for Sweden and Australia respectively.

Cross-cohort differences Next to cross-country heterogeneity, we are interested in the evolution of two-generational wealth mobility over time. To this end, I compare rank-rank coefficient estimates across children's birth cohorts (Table 1).

Inter-generational wealth mobility across two generations in the United States is found to have declined over time. The declining mobility can be established only between ages 35 and 44, however: for the 35-39 and 40-44 lifecycle stages, two-generational β-estimates are significantly higher the more recent the child birth cohort. For the other lifecycle stages, no definite conclusions can be drawn[7]. The increase in β is particularly strong for the 40-44 stage, with β rising from 0.37 (1956-65 cohort) to 0.51 (1976-85 cohort) based on proxy wealth. The decline in wealth mobility across two generations is due to both stronger persistence at the bottom and at the top, as shown in the Online Supplement.

How does this relate to existing literature? This paper is the first to investigate the time trend in overall two-generational wealth mobility. As a result, it is also the first to demonstrate the decline in overall two-generational wealth mobility for the United States. However, complementary evidence is provided by Fernholz & Hagler (2023), who report a decline in intergenerational wealth mobility for the top 400 wealthiest American families since 1985 (using Forbes 400 data). Furthermore, the decline in overall inter-generational wealth mobility aligns with evidence from Blanden et al. (2023) and Gregg & Kanabar (2023) for the United Kingdom, as well as with findings from Adermon et al. (2018) for Sweden.

Mobility at the bottom & top The rank-rank coefficients provide insight into wealth mobility outcomes across the entire wealth distribution. Instead, the ex-ante and ex-post transition matrices ($T_{\text{EA}}(a)$ and $T_{\text{EP}}(a)$, Figures 2 and 3) provide more detail on mobility at the bottom and top of the wealth distribution. Complementary evidence is provided by the discretionary groups, which together contain approximately 25% of parent-child pairs (Appendix H). In what follows, I report the baseline transition probabilities derived from κ^{Ψ} as a benchmark, and provide the results based on ML-proxy $\hat{\kappa}^{\Omega}$ in parentheses.

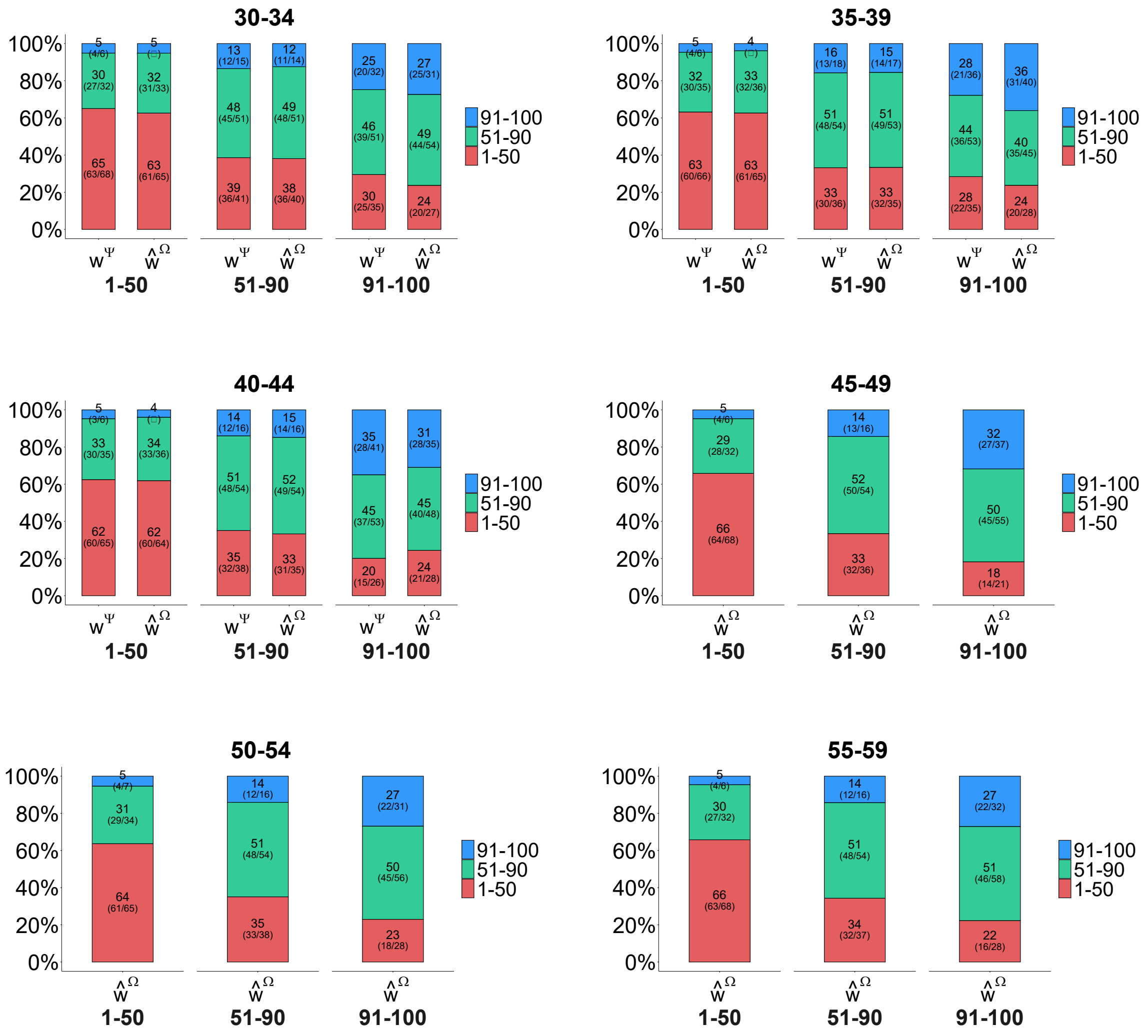
The pooled T_{EA} and T_{EP} show that:

- At the bottom: of parents in the bottom 50% at some lifecycle stage, 62%-65% (63%-66%) of their children end up in the bottom 50% at the same stage (Figure 2). Moreover, 5% (4%-5%) of children from parents in the bottom 50% at some stage ascend to the top 10% (Figure 2). Vice versa, of children in the bottom 50% at a given lifecycle stage, 3%-5% (4%-5%) originate from parents in the top 10% at that stage (Figure 3).
- At the top: 25%-35% (27%-36%) of the children from parents in the top 10% at some stage end up in the top 10% at the same stage (Figure 2). Furthermore, 20%-30% (18%-24%) of the children from parents in the top 10% at some stage drop to bottom 50% (Figure 2). Last, 23%–27% (17%–25%) of children who end up in the top 10% at some stage originate from parents in the bottom 50% (Figure 3).

These results have three implications. First, overall mobility across two generations is driven by both mobility at the bottom and at the top. Second, the parent-child lifecycle bias in two-generational samples is stronger at the top than at the bottom: the probability of families con-

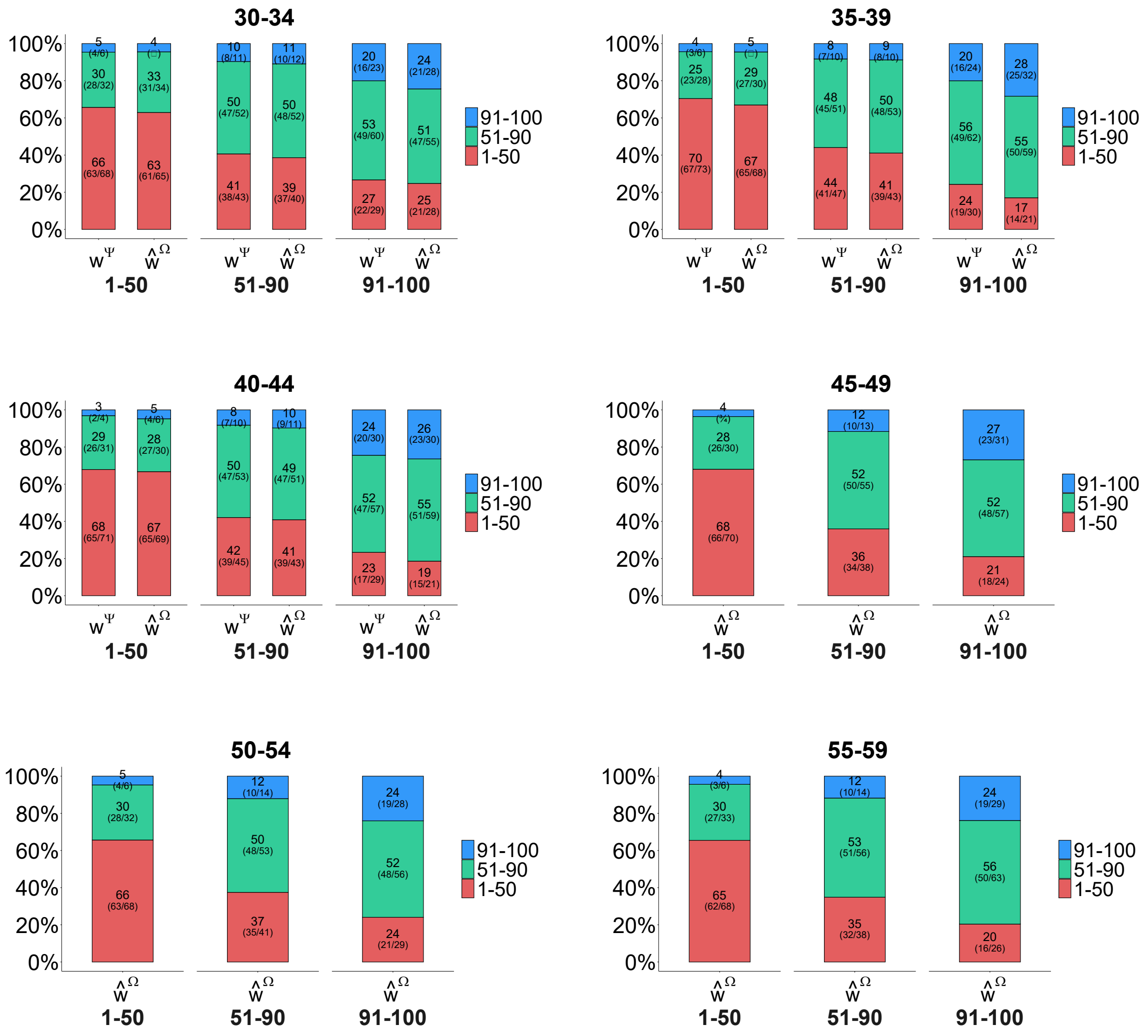
[7] For the 30-34 stage, the β-estimates are available only for the most recent child birth cohorts and do not display a clear trend. For the 45-49, 50-54 and 55-59 lifecycle stages, differences between the β-estimates for the 1946-55 and 1956-65 or 1966-75 cohorts are also limited and do not display a clear trend. These stages do not have sufficient data available for the more recent child birth cohorts.

Figure 2: Ex-ante transition matrices $T_{\mathbf{EA}}(a)$ between parental and children wealth ranks at identical lifecycle stages for the pooled dataset.



Note: these transition matrices compare parents' and children's within-cohort wealth ranks at identical lifecycle stages (ranging from stages 30-34 to 55-59). The transition probabilities are reported both for actual wealth w^{Ψ} (if available) and proxy wealth $\hat{w}^{\Omega}$. Given that the matrices are computed ex-ante, the x-axis represents parental wealth ranks. The y-axis displays children's wealth ranks given the wealth ranks of their parents at the same lifecycle stage. The numbers in parentheses display the 95% confidence intervals for the respective transition probability.

Figure 3: Ex-post transition matrices $T_{EP}(a)$ between parental and children wealth ranks at identical lifecycle stages for the pooled dataset.



Note: these transition matrices compare parents' and children's within-cohort wealth ranks at identical lifecycle stages (ranging from stages 30-34 to 55-59). The transition probabilities are reported both for actual wealth w^{Ψ} (if available) and proxy wealth $\hat{w}^{\Omega}$. Given that the matrices are computed ex-post, the x-axis represents children's wealth ranks. The y-axis displays parental wealth rank outcomes given their children's wealth ranks at the same lifecycle stage. The numbers in parentheses display the 95% confidence intervals for the respective transition probability.

solidating in the top 10% rises strongly with (parent-child) age considered, following the age pattern for overall mobility reported in Figure 1. By contrast, the link between parent-child age and persistence at the bottom is comparatively weaker. Third, in line with the overall mobility analysis, the proxy wealth ranks generally underestimate actual inter-generational wealth mobility.

4.2 Inter-generational mobility across three generations

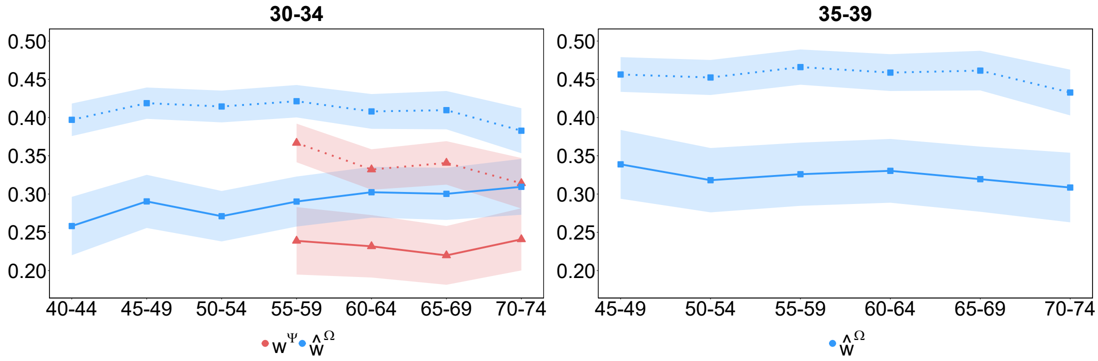
Having discussed inter-generational wealth mobility across two generations, I now discuss wealth mobility for grandparent-grandchild combinations (across three generations) for the pooled dataset. I report overall wealth mobility and age effects based on rank-rank coefficients (Figure 4), and provide more detail through the transition probabilities (Figure 5). A cross-cohort comparison is not feasible as the majority of grandchildren is concentrated in the same birth cohort (1976-85).

Grandparental wealth ranks are observed from age 40 or 45 onwards, while grandchild wealth ranks are recorded only between ages 30 and 39. As a result, in contrast to the two-generational mobility analysis, a comparison of the within-cohort wealth ranks of grandparents and grandchildren at identical lifecycle stages is infeasible. This mismatch is a common issue in the literature on inter-generational wealth mobility across three generations (Boserup et al., 2014; Pfeffer & Killewald, 2018). To allow for a direct comparison with the degree of three-generational wealth mobility, I consistently report the two-generational outcomes over the same stage combinations (as dotted lines).

Overall mobility Grandparent-grandchild inter-generational wealth mobility is higher than parent-child mobility. This is evidenced by the lower three-generational rank-rank coefficient estimates (solid lines) compared to the two-generational estimates (dotted lines) in Figure 4. The three-generational rank-rank coefficient β-estimates vary in function of the grandparent-grandchild stage combination considered. For grandchildren aged between 30 and 34, rank-rank coefficients range between $0.22-0.25$ based on actual wealth and between $0.26-0.29$ based on proxy wealth. For grandchildren aged between 35 and 39, the β-range based on proxy wealth increases to $0.30-0.34$ (with no estimate available for actual wealth).

Wealth rank resemblance between grandparents and grandchildren is stronger for grandchildren aged 35-39 compared to ages 30-34: the rank-rank coefficient estimates based on the proxy wealth ranks are approximately 4 to 6 points higher for grandchildren in stage 35-39 relative to grandchildren in stage 30-34 (regardless of grandparents' ages). This follows from the higher β-levels observed on the right-hand relative to the left-hand side plot in Figure 4, and indicates the presence of a grandchild lifecycle bias. Although unsurprising in light of the parent-child lifecycle bias (see Section 4.1.1), this grandchild lifecycle bias in β-estimates constitutes a novel result in the three-generational wealth mobility literature.

Figure 4: Rank-rank coefficients β for grandparents and grandchildren (solid lines) and parents and children (dotted lines) when (grand)children are aged 30-34 and 35-39.



Note: this figure produces rank-rank coefficients β computed from the within-cohort wealth ranks of grandparents-grandchildren (solid lines) and parents-children (dotted lines). These are calculated at different lifecycle stage combinations. Specifically, I compare (grand)child wealth ranks at ages 30-34 (left-hand side) and ages 35-39 (right-hand side) to (grand)parental wealth ranks across the lifecycle stages reported on the x-axis. The coefficients are computed based on actual wealth if available (from w^{Ψ}) and proxy wealth (from $\hat{w}^{\Psi}$). In the rank-rank regression, (grand)child wealth ranks are the dependent variable. The pooled dataset is used. The shaded areas display the 95% confidence intervals.

The grandchild lifecycle bias has two key implications. First, the three-generational β-ranges for stage 30-34 (0.22−0.25 for actual wealth) likely overestimate the degree of grandparent-grandchild inter-generational wealth mobility during midlife in the United States: rank-rank coefficients when grandchildren reach midlife are likely to be significantly higher. Given that Pfeffer & Killewald (2018) use a PSID-sample with young grandchildren, their rank-rank coefficient estimates likely suffer from the same downward bias (see below). Second, the grandchild age effect appears to be unique to the United States: Boserup et al. (2014) do not find a link between their benchmark rank-rank coefficients and average grandchild age in their Danish sample. However, more research on international three-generational wealth mobility would be needed to validate this conclusion.

Literature comparison How do these findings compare to existing literature? For the United States, Pfeffer & Killewald (2018) report three-generational rank-rank coefficient estimates of 0.23 (using actual wealth) and 0.21 (using proxy wealth). They obtain their rank-rank coefficients through a regression that includes grandparental and grandchild age controls. Moreover, their proxy wealth series uses main housing values and rental payments and display inferior performance compared to the ML-proxy used in this paper (see Appendix C). While my actual wealth rank regressions (0.22−0.25 at grandchild ages 30-34) yield similar rank-rank coefficients to theirs (0.23), the proxy wealth rank-rank coefficients reported in this paper (0.26 to 0.29 at grandchild ages 30-34) are substantially higher than the one in Pfeffer & Killewald (2018) (0.21).

What explains this large discrepancy in proxy wealth rank β-estimates? The three-generational samples in Pfeffer & Killewald (2018) rely on grandchild observations for 2013 and 2015. As these observations follow closely after the 2008-2009 real estate bust, a housing proxy using solely main housing values may not accurately approximate individual wealth. This is shown in Appendix C: the measurement error of the housing proxy used by Pfeffer & Killewald (2018) is particularly higher during the post-crisis years. Instead, the extended sample (until 2021) and superior ML-proxy used in this paper are likely to have generated a more accurate rank-rank estimate. In line with the two-generational analysis (see Section 4.1.1), these results also imply that $-$ contrary to the claim in Pfeffer & Killewald (2018) $-$ rank-rank coefficients based on proxy wealth underestimate actual three-generational wealth mobility.

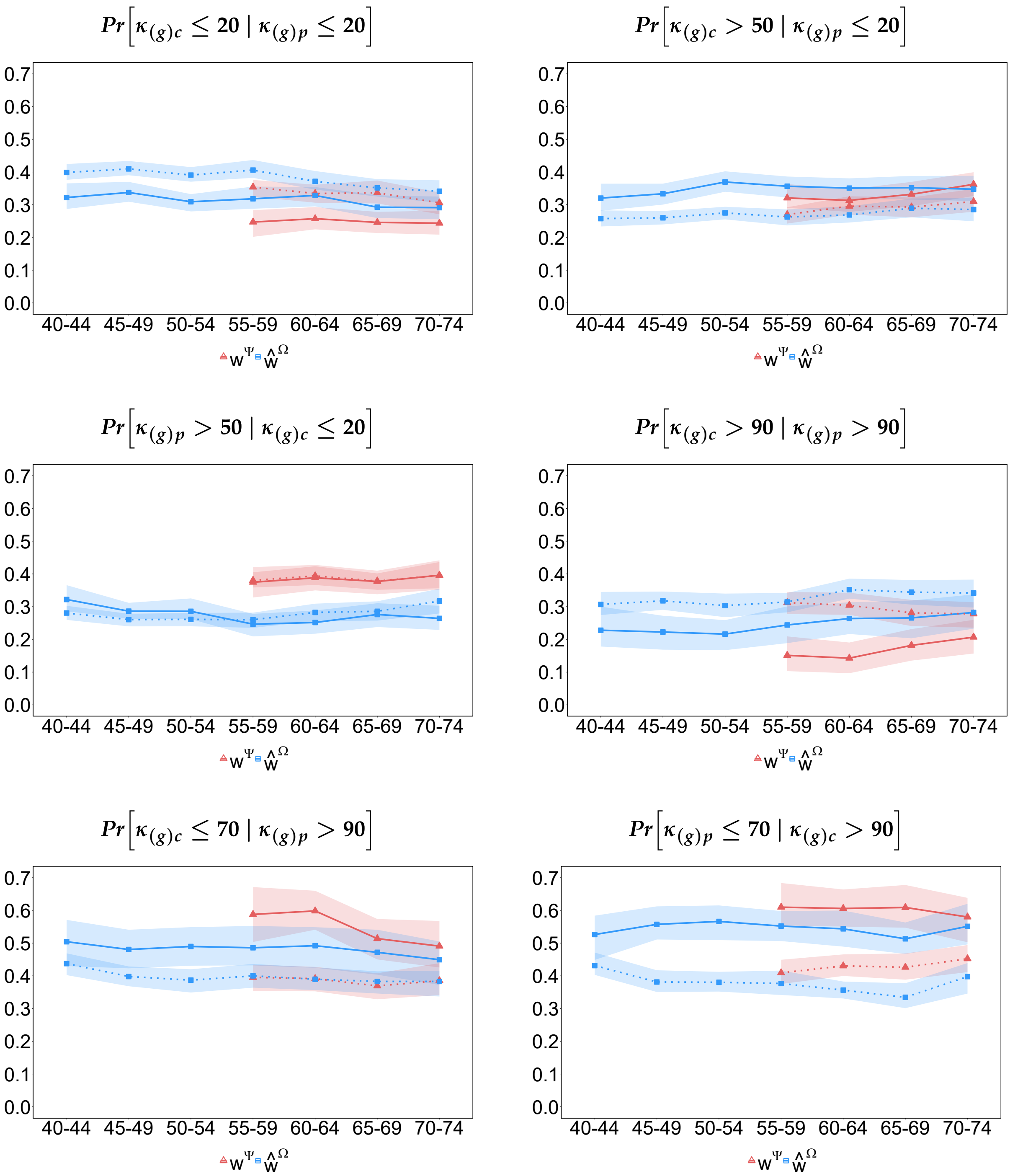
Three-generational wealth mobility in the United States is lower compared to Denmark (in line with the two-generational analysis) and lower than in Sweden (contrary to the two-generational analysis): using actual wealth, Boserup et al. (2014) report a three-generational benchmark β-estimate of 0.16 for Denmark, while Adermon et al. (2018) produce rank-rank coefficients of 0.14 to 0.17 for Sweden. The estimates in both papers are robust to the grandchild lifecycle bias. They are lower than the $0.22-0.25$ values (actual wealth, grandchild ages 30-34) established for the United States in this paper. If sufficient actual wealth data were available for later grandchild stages in the PSID, the grandchild lifecycle bias implies that the gap in three-generational mobility between the United States and Nordic countries would likely be even more pronounced.

Mobility at the bottom & top The transition probabilities affirm the conclusion of higher wealth mobility over three generations (grandparent-grandchild) compared to two generations (parent-child). This finding holds both at the top and bottom of the wealth distribution. There exists a non-linearity, however: while mobility at the top is significantly higher over three relative to two generations, the divergence in wealth mobility at the bottom is more limited.

In what follows, I quantify this non-linearity based on Figure 5. It compares the transition probabilities across three generations (solid lines) versus two generations (dotted lines) when grandchildren are aged between 30-34. Data for this grandchild stage is available for both actual wealth (reported as benchmark) and proxy wealth (reported in parentheses):

- At the bottom: 22%-23% (31%-34%) of grandchildren with grandparents in the bottom 20% during their lifecycle end up in the bottom 20% during stage 30-34, compared to 31%-33% (37%-41%) of children from bottom 20% parents (steady poor). Moreover, 31%-37% (33%-37%) of grandchildren with grandparents in the bottom 20% during their lifecycle end up in the top 50% at ages 30-34, while this number equals 29%-32% (27%-29%) for children from bottom 20% parents (past poor). Conversely, 39%-40% (29%-32%) of grandchildren belonging to the bottom 20% at ages 30-34 originate from grandparents

Figure 5: Transition probabilities for grandparents and grandchildren (solid lines) and parents and children (dotted lines) when (grand)children are aged 30-34.



Note: these plots produce transition probabilities over specific wealth bin combinations. These are defined in line with the discretionary groups (see Section 3.4 and Appendix E). In the notation above, $\kappa_{(g)p}$ denotes the within-cohort wealth ranks of (grand)parents and $\kappa_{(g)c}$ the within-cohort wealth ranks of (grand)children. The transition probabilities are computed at different lifecycle stage combinations: child wealth ranks at ages 30-34 are compared to (grand)parental wealth ranks at the stages between ages 40-44 and 70-74 (plotted on the x-axis). As an example, the values produced for the right-hand plot on the top row indicate the probability of (grand)children belonging to the top 50% at stage 30-34 given that their (grand)parents belonged to the bottom 20% at any of the x-axis stages. I report the outcomes for child lifecycle stage 35-39 in Appendix H. The shaded areas display the 95% confidence intervals, which have been determined through bootstrapping.

belonging to the top 50% over their lifecycle, relative to 40%-43% (29%-31%) of children from top 50% parents (new poor).

- At the top: 12%-20% (23%-27%) of grandchildren with grandparents in the top 10% during their lifecycle end up in the top 10% during stage 30-34, compared to 28%-30% (30%-33%) of children from top 10% parents (steady wealthy). Moreover, 52%-58% (49%-52%) of grandchildren with grandparents in the top 10% during their lifecycle end up in the bottom 70% at ages 30-34, while this number equals 38%-39% (40%-42%) for children of top 10% parents (past wealthy). Finally, 61%-68% (53%-58%) of grandchildren belonging to the top 10% at ages 30-34 originate from grandparents belonging to the bottom 70% over their lifecycle, relative to 41%-48% (37%-44%) of children from bottom 70% parents (new wealthy).

These results show that the relative differences between grandparent-grandchild and parent-child transition probabilities are significantly higher for wealthy discretionary families (steady wealthy, past wealthy and new wealthy) compared to the poor discretionary families (steady poor, past poor and new poor). The same conclusion persists when considering grandchild lifecycle stage 35-39 (Figure 21, Appendix H). As a result, mobility at the top is significantly higher over three relative to two generations, while the difference in mobility at the bottom is comparatively more limited.

What are tentative theoretical mechanisms behind this non-linearity? I focus on channels that explain the relatively low persistence at the top, versus those that center on the relatively strong persistence at the bottom. At the top, families pass along significant wealth across generations through inter-vivos transfers and inheritances. In theory, this should generate strong persistence at the top. However, even for steady wealthy parent-child pairs, the inter-generational transfers make up only 4% of their lifetime resources on average during working life[8]. This suggests that their impact on inter-generational wealth transmission may be limited. In addition, business ownership is passed along wealthier families (either through type or scale dependence), as evidenced by high business ownership rates among steady and past wealthy families at the start of the children's working life (see Online Supplement). Given the high idiosyncratic risk involved in business ownership, this is expected to lead to downward mobility for a significant fraction of wealthy families over longer time-frames[9]. At the bottom, a non-negligible number of families are stuck in multi-generational spirals of low labor income and asset ownership, little or no saving, poor health and − as a result of their low wealth levels − minimal inter-generational transfer receipts (see Online Supplement).

[8]Towards the end of working life, this fraction is higher for the most wealthy individuals (around 11%-16%). Nonetheless, in line with evidence from Black et al. (2022) for Norway, inter-vivos transfers and inheritances still make up a relatively limited fraction of lifetime resources even for individuals from wealthy families.

[9]This coincides with the argument made in Kalsi & Ward (2025), who find that persistence among the elite wealthiest during the Gilded Age period in the United States was relatively low.

5 Intra-generational wealth inequality & individual-level wealth mobility

In this section, I investigate intra-generational (individual-level) within-cohort wealth accumulation, inequality and mobility over the lifecycle. The lifecycle is split into working life (ages 30-54) and older age (ages 55-74). Section 5.1 presents within-cohort wealth shares and wealth-to-income ratios. Section 5.2 elaborates on the determinants of the within-cohort wealth distribution at the start of the lifecycle. Given their initial wealth ranks at ages 30-34, individuals' wealth rank trajectories during working life and older age are investigated in Section 5.3 and 5.4 respectively. Section 5.5 compares wealth mobility outcomes across birth cohorts, while Section 5.6 explores the timing of intra-generational wealth mobility.

Two sample restrictions are applied. First, I limit the working life and older age samples to individuals with wealth rank observations in both the initial (30-34 or 55-59) and final stage (50-54 or 70-74) of the respective lifecycle phase. To ensure a balanced panel, I recalculate the within-cohort wealth ranks for these restricted samples. Second, the sample is further limited to individuals in birth cohorts with a minimum of 250 observations for either κ_j^{Ψ} or $\hat{\kappa}_j^{\Omega}$ after the first restriction is applied[10]. The Ψ- and Ω-samples for the pooled data respectively contain 2524 and 3750 observations for working life, and 1534 and 2046 observations for the older age phase. Note that the samples of individuals used for working life and older age are distinct, with no overlap between the individuals of the two samples. Outcomes across the two lifecycle phases are therefore only indirectly comparable.

Previewing the results, within-cohort wealth inequality is found to be stable over the lifecycle. Wealth-to-income ratios rise around fivefold over working life, while wealth decumulation during older age occurs only after age 65. The initial wealth distribution at ages 30-34 overlaps significantly with the distribution of family wealth and distribution of cumulative inter-vivos transfers and inheritances received at that point. Next, I report rank-rank coefficients and transition probabilities. Intra-generational wealth mobility during older age is significantly lower than during working life, and most intra-generational wealth mobility is found to occur between ages 30 and 39. Finally, the data shows a negative correlation between within-cohort wealth inequality (which has increased over time) and wealth mobility at the top 10% (which has significantly declined over time).

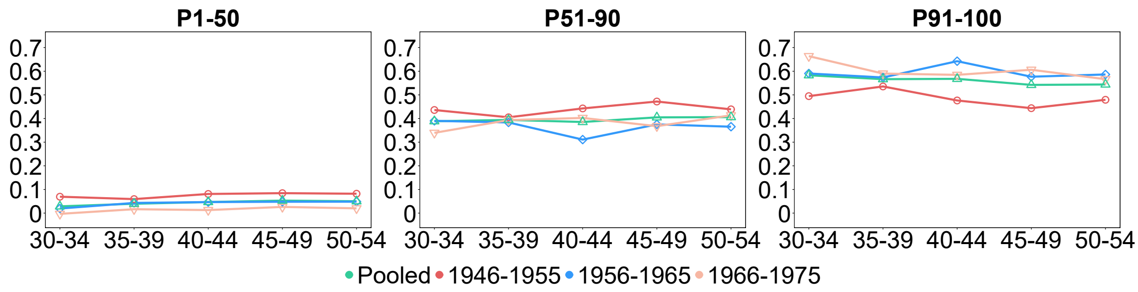
[10] The cohorts that fulfill the minimum observation criterion include:

$$\Upsilon^{\mathrm{WL}} = \{P^{\mathrm{WL}},\ 1936\text{–}45,\ 1946\text{–}55,\ 1956\text{–}65,\ 1966\text{–}75\}$$
$$\Upsilon^{\mathrm{OA}} = \{P^{\mathrm{OA}},\ 1916\text{–}25,\ 1926\text{–}35,\ 1936\text{–}45,\ 1946\text{–}55\}$$

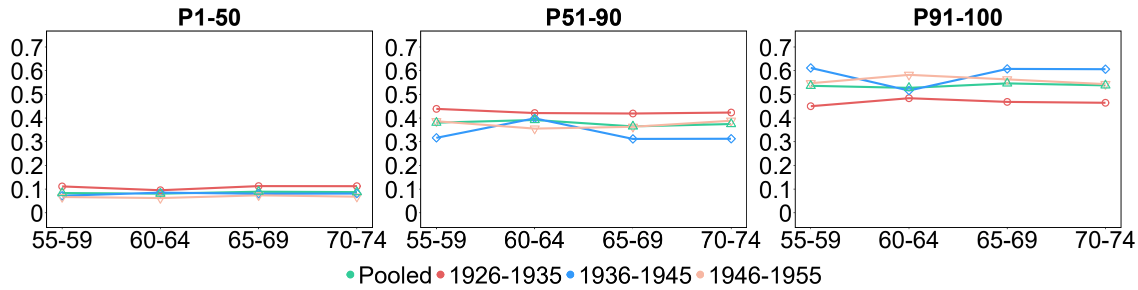
Here, P^{WL} and P^{OA} denote the pooled datasets for working life and older age respectively. These pooled datasets contain all observations from the other cohorts.

Figure 6: Wealth shares λ_b across lifecycle stages for birth cohorts $\in \Upsilon^{WL}$ and $\in \Upsilon^{OA}$ based on actual wealth levels w^{Ψ}.

Working life: ages 30-54



Older age: ages 55-74



Note: these plots show the within-cohort wealth shares for the bottom 50%, middle 50%-90% and top 10% wealthiest at each lifecycle stage per birth cohort. The shares are calculated using working life and older age samples for actual wealth levels w^{Ψ}. Given that the working life and older age samples contain different individuals, the wealth shares are not directly comparable across the upper and lower panel. The pooled wealth share is computed as the average of the wealth shares across the birth cohorts per lifecycle stage.

5.1 Wealth inequality & accumulation

Within-cohort wealth inequality remains roughly stable throughout the lifecycle, as shown by the relatively flat profile of the pooled wealth shares in Figure 6. During working life, the top 10% wealth share fluctuates between 55% and 60%, while the bottom 50% own 0% to 5% of total wealth. The top 10% wealth shares track closely the SCF-estimates of Bauluz & Meyer (2024), although I do not find higher wealth inequality during the early stages of the working lifecycle (ages 30-34), except for the 1966-75 cohort. During older age, pooled top 10% wealth shares equal approximately 53% between ages 55 and 74, while the bottom 50% owns around 8% of total within-cohort wealth (Figure 6). This leaves an approximate wealth share for the middle 50%-90% of 39%. The observed stability in within-cohort wealth inequality during older age is also consistent with the results of Bauluz & Meyer (2024).

The stability of within-cohort wealth inequality implies that wealth growth rates are similar across the wealth distribution. During working life, wealth-to-income ratios increase around fivefold for the bottom 50%, middle 50-90% and top 10% brackets (Figure 7). This substantial accumulation of wealth over the working lifecycle leads to an increase in the fraction of high wealth individuals from roughly 1% to 7%, and a decline in the proportion of low-wealth individuals from approximately 58% to 33% (Figure 22, Appendix H). During older age, all wealth brackets exhibit additional wealth accumulation between ages 55 and 64, followed by wealth decumulation between ages 65-74 (Figure 7). The 1946-55 cohort stands out to the others by notably higher wealth to income ratios, and a higher fraction of high-wealth individuals (Figure 22 in Appendix H). This is likely related to the extreme asset price trajectories (Dotcom bubble and Great Financial Crisis) experienced by this cohort at the end of its working life and beginning of its older age.

Within-cohort wealth inequality has increased over time, in line with the SCF-estimates from Bauluz & Meyer (2024). This follows from a cross-cohort comparison of wealth inequality outcomes (Figures 6). For working life, two findings stand out. First, the 1966-75 cohort displayed significantly higher wealth inequality at the start of the working lifecycle compared to the two earlier cohorts (1946-55 and 1956-65), with wealth shares above 70%. In addition, the 1966-75 and 1956-65 cohorts experienced higher wealth inequality from ages 40 to 54 compared to the 1946-55 cohort. Second, for the two most recent cohorts (1956-65 and 1966-75), wealth shares for the bottom 50% were significantly closer to zero compared to the 1946-55 cohort. This likely follows from increased non-mortgage indebtedness in recent decades (e.g. Bartscher et al., 2024). For older age, the two most recent cohorts (1936-45 and 1946-55) experienced higher within-cohort wealth inequality than the 1926-35 cohort: the top 10% wealth share in the most recent cohorts was at least 10%-points higher (at 63% and 56% for the 1936-45 and 1946-55 cohorts compared to 46% for the 1926-35 one). Accordingly, bottom 50% wealth shares in the most recent cohorts lay substantially below those of the 1926-35 one (around 7% versus 11%).

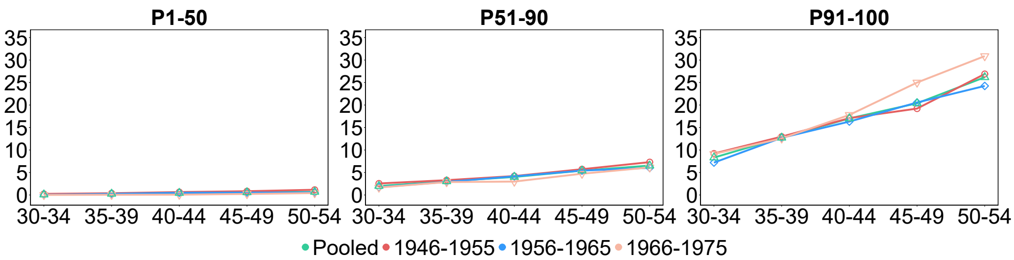
5.2 Wealth distribution: ages 30-34

Around 60% of individuals at ages 30-34 have wealth levels lower than the annual average labor income (Figure 22, Appendix H). Only around 1% of individuals display wealth levels in excess of twenty times labor income. This suggests that approximately 39% of individuals start off working life with wealth levels between one and twenty times average labor income. This begs the question: where does this wealth come from?

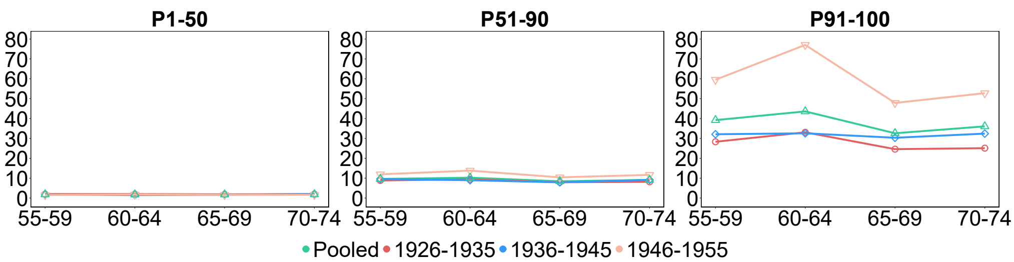
While the structure of the data does not allow for a comprehensive accounting decomposition, it demonstrates that family wealth plays a critical role in determining the within-cohort wealth distribution at ages 30-34: wealthy individuals at the start of working life tend to belong to wealthy families and are the most likely to have received an inter-vivos transfer or inheritance. This overlap in individuals' wealth ranks at ages 30-34 with family wealth and

Figure 7: Wealth-to-income ratios θ_b across lifecycle stages for birth cohorts $\in \Upsilon^{WL}$ and $\in \Upsilon^{OA}$ based on actual wealth levels w^{Ψ}.

Working life: ages 30-54



Older age: ages 55-74



Note: these plots show the within-cohort wealth-to-income ratios for the bottom 50%, middle 50%-90% and top 10% wealthiest at each lifecycle stage per birth cohort. The ratios are calculated using working life and older age samples for actual wealth ranks. Income is computed as average annual labor income. Given that the working life and older age samples contain different individuals, the ratios are not directly comparable across the upper and lower panel. The pooled wealth-to-income ratio is computed as the average of the wealth-to-income ratios across the birth cohorts per lifecycle stage.

inter-generational transfer receipts aligns with evidence from Boserup et al. (2018): these authors find a similar overlap in Denmark, albeit for much younger individuals (at age 18). In what follows, I quantify these findings in more detail.

Of the individuals in the within-cohort top 10% at ages 30-34, 55% have parents that belong to the top 30% of their own cohort at that time. Furthermore, close to 30% of individuals in the top 10% have already received an inter-vivos transfer or inheritance. This is higher than for the middle 50%-90% (15%) and bottom 50% (8%). Total transfer receipts of the top 10% by ages 30-34 make up around 50% of the total cumulative transfers received by individuals at that stage. Instead, of the individuals in the within-cohort bottom 20% at ages 30-34, only 15% have parents that belong to the top 30% of their own cohort at that time. Moreover, only 6% of individuals in the within-cohort bottom 20% at the start of working life have received an inter-vivos transfer or inheritance at that point.

5.3 Wealth mobility during working life

While wealth growth rates over the working lifecycle (ages 30-54) are broadly similar across wealth brackets (Section 5.1), this conceals intra-generational mobility of individuals across the within-cohort wealth distribution. That is, the within-cohort bottom 50%, middle 50%-90% and top 10% are not fixed groups: significant turnover takes place over the lifecycle. In what follows, I quantify the degree of wealth mobility during working life at the individual level.

Overall mobility The rank-rank coefficient (based on κ^{Ψ}) in the pooled dataset between ages 30-34 and 50-54 equals 0.57 (Table 2). This finding is in line with Shiro et al. (2022), who obtain a rank-rank estimate of 0.59 for the United States using a PSID-sample over the same age span (30-54). The minor difference to my estimate likely follows from sample differences: while I use the SRC-subsample in the PSID (as detailed in Appendix A), Shiro et al. (2022) use this SRC-subsample in combination with the SEO-subsample and two immigrant subsamples. Furthermore, Conley & Glauber (2008) produce a log-log estimate of 0.47 using a PSID sample that spans twenty years. However, these authors' sample constitutes of individuals across a broad spectrum of initial age levels, and is therefore not directly comparable to mine.

Two conclusions can be drawn. First, intra-generational wealth mobility in the United States is significantly lower compared to the Nordic countries. Specifically, over the same age span as this paper (30-54), Audoly et al. (2024) find a rank-rank coefficient slightly in excess of 0.20 for Norway. Moreover, Boserup et al. (2018) obtain a β-estimate of 0.22 for Denmark in a study where individuals' wealth ranks at age 45 are regressed on those at age 18. Second, in line with the inter-generational analysis, rank-rank coefficients based on proxy wealth ranks underestimate the actual degree of intra-generational wealth mobility: the β-estimate based on $\hat{\kappa}^{\Omega}$ equals 0.66 (compared to the actual value of 0.57).

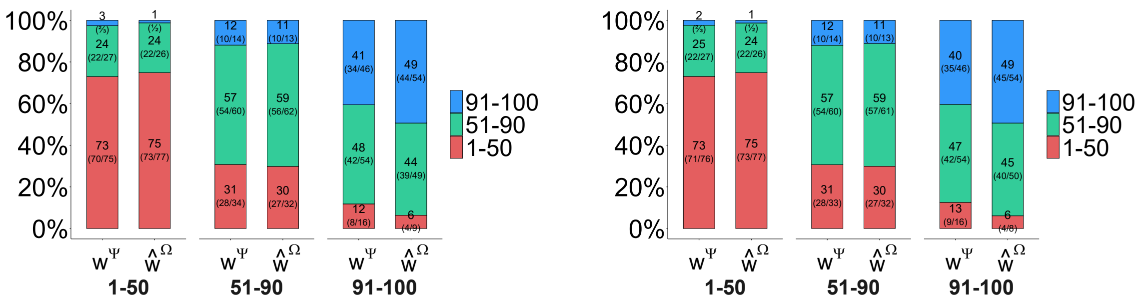
Table 2: Fraction of individuals belonging to each of the discretionary groups (in %) and rank-rank coefficients β across cohorts $\in \Upsilon^{WL}$ based on actual wealth ranks κ^{Ψ}.

Cohort	Poor Groups (%) Steady	Past	New	Wealthy Groups (%) Steady	Past	New	β
Pooled	9.2	3.8	3.5	4.1	2.9	2.3	0.57
	(8.3, 10.1)	(2.6, 5.0)	(2.4, 4.7)	(3.4, 4.7)	(2.0, 4.0)	(1.4, 3.3)	(0.02)
1946–55	9.3	4.2	3.4	3.0	3.4	3.3	0.54
	(7.9, 10.6)	(2.3, 6.5)	(1.8, 5.4)	(2.0, 3.9)	(2.0, 5.3)	(1.6, 5.4)	(0.03)
1956–65	9.5	3.2	3.6	4.3	2.8	1.9	0.57
	(8.4, 10.7)	(1.8, 4.7)	(2.0, 5.5)	(3.4, 5.3)	(1.4, 4.4)	(0.8, 3.2)	(0.02)
1966–75	8.1	4.7	3.3	5.5	2.1	1.4	0.60
	(6.1, 10.3)	(2.0, 8.4)	(1.3, 6.2)	(4.1, 6.9)	(0.9, 4.5)	(0.7, 3.8)	(0.04)

Note: this table reports the fraction of individuals in the sample belonging to each of the discretionary groups (in %). Moreover, it reports rank-rank coefficients β. These metrics are calculated with within-cohort wealth ranks at stage 50–54 as dependent variable. The values are shown with confidence intervals for the discretionary group shares and standard errors for the estimated rank-rank coefficients (in parentheses).

Figure 8: Ex-ante and ex-post transition matrices during working life (ages 30-54) for the pooled dataset.

(a) ***Ex-ante:*** $T_{EA}(a)$ **(b)** ***Ex-post:*** $T_{EP}(a)$



Note: these transition matrices compare the within-cohort wealth ranks of individuals in the working life sample at ages 30-34 and ages 50-54. The ex-ante matrix shows individuals' wealth ranks at ages 50-54 given their initial wealth rank at ages 30-34 (shown on the x-axis). Instead, the ex-post matrix shows individuals' initial wealth ranks at ages 30-34 given their final wealth rank at ages 50-54 (shown on the x-axis). The usage of the pooled dataset indicates that individuals across all selected birth cohorts are included in the sample. The numbers in parentheses display the 95% confidence intervals for the respective transition probability.

Mobility at the bottom & top Rank-rank coefficients provide insight into intra-generational wealth mobility across the entire wealth distribution, but do not show how broad-based mobility over the lifecycle is. Next, I therefore report transition probabilities and hierarchical clustering outcomes.

The pooled ex-ante transition matrix $T_{\mathrm{EA}}(a)$, ex-post $T_{\mathrm{EP}}(a)$ transition matrix (Figure 8) based on actual wealth ranks κ^{Ψ} reveal that[11]:

- At the bottom: 73% (75%) of individuals in the bottom 50% of their cohort at ages 30-34 still belong to the bottom 50% at ages 50-54. Conversely, 27% (25%) of the individuals in the bottom 50% at age 30-34 displayed upward mobility during their working life, with 3% (1%) migrating to the top 10% of the distribution. Finally, 2% (1%) of the individuals that end the working lifecycle in the bottom 50% originate from the within-cohort top 10% at ages 30-34.
- At the top: 41% (49%) of individuals in the top 10% at ages 30-34 have remained in this wealth bin by ages 50-54. Conversely, 59% (51%) of the top 10% wealthiest at ages 30-34 exhibit downward wealth mobility, with 12% (9%) falling to the bottom 50%. Last, among those individuals in the top 10% at ages 50-54, 60% (51%) started working life in the bottom 70%. 13% (6%) of the top 10% individuals at ages 50-54 began their working life in the bottom 50% wealthiest.

These results have two key implications. First, overall mobility during working life is induced by both wealth mobility at the bottom and at the top of the wealth distribution. Second, the proxy wealth series' bias in estimating wealth mobility relates both to the bottom and to the top: the proxy ranks overestimate the persistence at the bottom (75% versus 73%), as well as the persistence at the top (49% versus 41%).

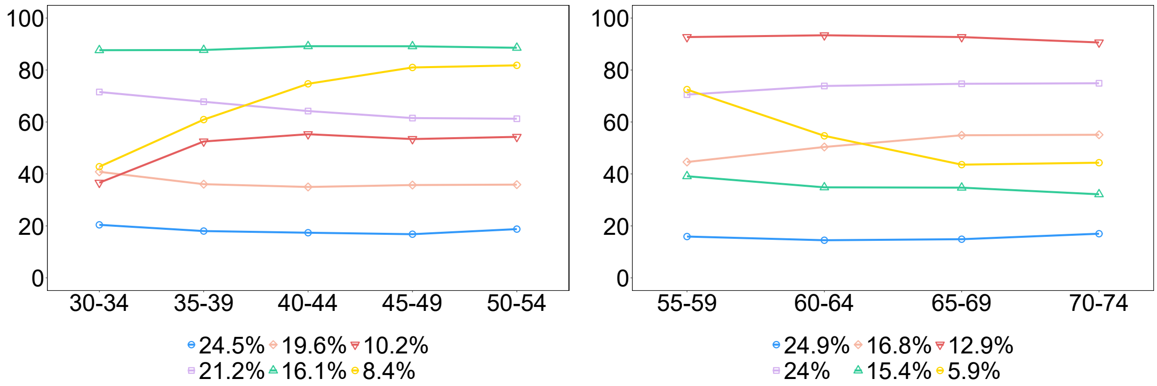
Complementary evidence to the transition matrices and discretionary groups is provided by the hierarchical clustering algorithm. Its application to the actual wealth series κ^{Ψ} for working life is presented in Figure 9 (panel a). Unlike for the discretionary groups, all individuals in sample Ψ have been categorized into one of the six benchmark clusters. I report the proportion of individuals in each cluster in parentheses. That is:

- Two immobile clusters at the bottom (45%): akin to the steady poor group, the steady bottom cluster (25%) contains individuals that spend their entire working life in the vicinity of the 20th wealth percentile. Instead, the steady supra-bottom cluster (20%) includes individuals that display fluctuating wealth ranks between the 30th and 40th wealth percentile.
- Two mobile clusters (18%): akin to the new wealthy group, the average individual in the strong risers cluster (8%) starts off around the 40th wealth percentile, exhibits a drastic

[11]The transition probabilities based on proxy wealth $\hat{\kappa}^{\Psi}$ are reported in parentheses.

Figure 9: Hierarchical clustering wealth rank trajectories for working life and older age for the pooled dataset based on actual wealth ranks κ^{Ψ}.

(a) ***Working life (ages 30-54)*** **(b)** ***Older age (ages 55-74)***



Note: the plots report the average within-cohort wealth rank trajectories of the individuals in the six hierarchical clusters. The clusters have been computed through the hierarchical clustering algorithm described in the Online Supplement using actual wealth ranks κ^{Ψ} as input. I report cluster outcomes based on proxy wealth ranks in the Online Supplement.

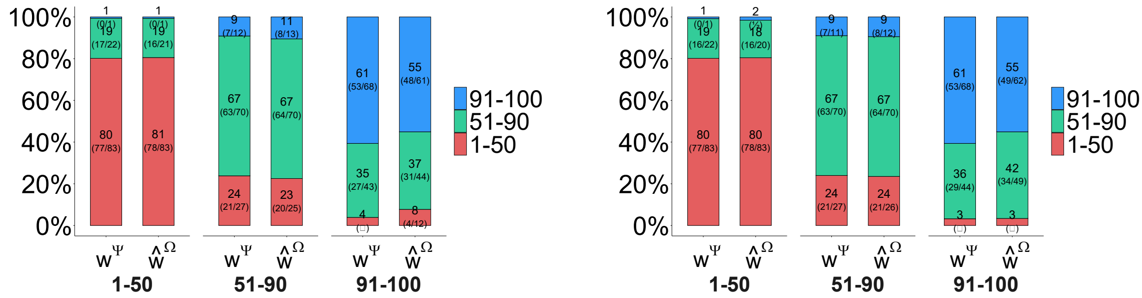
rise to the 70th wealth percentile by ages 40-44, and a slight further increase to above the 80th wealth percentile thereafter. Instead, individuals in the moderate risers cluster (10%) experience a more limited rise in their wealth ranks from below the 40th to around the 60th wealth percentile.

- Two immobile clusters in the upper half of the distribution (37%): the steady middle cluster (21%) contains individuals that spend their entire working lifecycle between the 60th and 70th wealth percentiles. Instead, akin to the steady wealthy group, the individuals in the steady top cluster (16%) maintain a stable wealth rank around the 90th wealth percentile throughout their working lifecycle.

Only a relatively small fraction of individuals (18%) displays significant wealth mobility over working life. The remainder of individuals in the sample (82%) is relatively immobile. This fraction of mobile individuals in the United States (18%) is lower than in Norway: Audoly et al. (2024) find that 36% of the individuals in their sample display substantial upward or downward mobility between ages 30 and 54. As a result, intra-generational wealth mobility in the United States is lower than in Norway, which aligns with the conclusion based on rank-rank coefficients.

Figure 10: Ex-ante and ex-post transition matrices during older age (ages 55-74) for the pooled dataset.

(a) ***Ex-ante:*** $T_{EA}(a)$ **(b)** ***Ex-post:*** $T_{EP}(a)$



Note: these transition matrices compare the within-cohort wealth ranks of individuals in the old age sample at ages 55-59 and ages 70-74. The ex-ante matrix shows individuals' wealth ranks at ages 70-74 given their initial wealth rank at ages 55-59 (shown on the x-axis). Instead, the ex-post matrix shows individuals' initial wealth ranks at ages 55-59 given their final wealth rank at ages 70-74 (shown on the x-axis). The usage of the pooled dataset indicates that individuals across all selected birth cohorts are included in the sample. The numbers in parentheses display the 95% confidence intervals for the respective transition probability.

5.4 Wealth mobility during older age

Having discussed wealth mobility during working life, I now move to the discussion of intra-generational wealth mobility during older age (ages 55-74). This paper is the first to explicitly study wealth mobility during this lifecycle phase.

Intra-generational wealth mobility during older age is found to be lower than intra-generational wealth mobility during working life: the rank-rank coefficient estimate in the pooled dataset equals 0.77 between ages 55 and 74 (Table 2), compared to 0.57 for working life (Section 5.3). This finding holds also when accounting for the disparity in lifecycle time span: the estimated rank-rank coefficients for a 20-year working lifecycle span equal 0.60 (for ages 30-49) and 0.68 (for ages 35-54), which are still significantly lower than the estimate for older age (0.76).

These findings should be approached with caution, however: the sample restrictions (see introduction to Section 5) imply that older age wealth mobility moments are computed based on a sample of individuals that are still alive by ages 70-74. This may introduce a selection bias: individuals that have poor health and die prematurely (and are thus not included in the sample) could face downward wealth mobility due to high healthcare expenditures or as a result of voluntary bequests. As this downward mobility will not be captured, the estimated rank-rank coefficient (0.77) may underestimate the actual degree of wealth mobility during older age.

The lower wealth mobility during older age compared to working life holds both at the bottom and top of the wealth distribution. More precisely, the ex-ante $T_{\text{EA}}(a)$ and ex-post $T_{\text{EP}}(a)$ transition matrices (Figure 10) reveal that[12]:

- At the bottom: 80% (81%) of the individuals in the bottom 50% at ages 55-59 still belong to this bin by age 70-74. Of those displaying upward mobility, only 1% (1%) migrated to the top 10% wealthiest. Roughly 1% (2%) of the individuals in the bottom 50% at age 70-74 started the older age phase in the top 10%.
- At the top: 61% (55%) of the individuals in the top 10% at the start of older age still belong to this bin by age 70-74. 4% (8%) of the individuals starting at the top drop to the bottom 50% of the wealth distribution. Finally, 3% (3%) of the individuals ending older age in the top 10% started off in the bottom 50%.

The hierarchical clustering procedure underscores that wealth mobility during older age is lower than during working life (Figure 9, panels a and b): while the older age cluster types overlap with those from working life, the strong risers cluster is replaced by a strong droppers cluster whose wealth ranks decline from right below the 80th to slightly above the 40th wealth percentile. In addition, the steady middle cluster now lies closer to the top, with stable wealth ranks around the 75th percentile. The relative occurrence of the cluster types also differs: in older age compared to working life, the steady bottom make up 25% (versus 25%), the steady supra-bottom 15% (versus 16%), the middle risers 17% (versus 10%), the strong droppers 6% (versus 8% of stronger risers), steady middle 24% (versus 21%), and the steady top 13% (versus 10%).

Finally, the proxy wealth series approximates actual wealth mobility during older age more accurately than during working life and across generations: the rank-rank coefficient estimate based on proxy wealth (0.78) lies very close to the one based on actual wealth (0.77). Moreover, the degree of persistence at the bottom and top align a lot more closely (81% versus 80% and 61% versus 55% respectively) than in previous sections of the paper. This better approximation during older age may relate to the lower importance of hard-to-capture variables such as business returns and non-mortgage indebtedness during older age.

5.5 Timing effects

In this section, I investigate timing effects in intra-generational wealth mobility: is within-cohort wealth rank mobility stronger at specific points of the lifecycle? The analysis relies primarily on Figure 11, which presents rank-rank coefficients from a rolling window analysis.

Two key findings persist. First, individuals' wealth rank position at ages 30-34 is increasingly less predictive of their current wealth rank as individuals progress through their lifecycle. Second, the majority of wealth mobility over the lifecycle occurs between ages 30 and 39: the

[12]Results based on the proxy wealth data are reported in parentheses.

rolling analysis based on w^{Ψ} reports a β-estimate of around 0.72 for the transition from stage 30-34 to 35-39. This is significantly lower than the $0.80-0.85$ estimates for the other transitions during working life and $0.85-0.90$ for the transitions during older age. The timing effect is corroborated by Figure 9 (panel a): the moderate risers and strong risers clusters for working life exhibit the majority of their mobility during the earlier stages of the working lifecycle, particularly between ages 30 and 39. Additionally, the moderate risers and strong droppers clusters for older age display gradual rather than abrupt shifts in wealth rank trajectories (Figure 9, panel b). The timing effect of intra-generational wealth mobility aligns with evidence for Norway in Audoly et al. (2024).

The higher intra-generational wealth mobility between ages 30 and 39 holds both at the bottom and top of the wealth distribution (Figure 23, Appendix H)[13]. At the bottom, an individual in the bottom 20% at stage 30-34 has a 56% probability of remaining in the bottom 20% by ages 35-39 (steady poor). This probability increases to above 60% in the subsequent stages. Moreover, the likelihood of moving from the bottom 20% to the top 50% between stages 30-34 and 35-39 equals 13% (past poor). It drops to below 10% for later transitions. Finally, the probability of dropping to the bottom 20% when starting from the top 50% remains relatively stable at 8%-9% throughout the working lifecycle (new poor). At the top, an individual in the top 10% at stage 30-34 has a 53% probability of still belonging to the top 10% by ages 35-39 (steady wealthy). For later transitions, this probability consistently exceeds 60%. Furthermore, an individual belonging to the top 10% at ages 30-34 has a 15% probability of dropping to the bottom 70% by ages 35-39, which declines to below 8% for later transitions (past wealthy). Last, the probability of rising from the bottom 70% to the top 10% declines from around 6% to 4% (new wealthy).

What are potential theoretical mechanisms underlying the timing effects in intra-generational wealth mobility? First, absolute differences in wealth levels between individuals are significantly smaller at the start of working life (Section 5.1). This implies that given additive shocks (labor income, inter-generational transfers, household formation) are expected to generate more substantial wealth mobility early in working life. Second, idiosyncratic investment risk-taking is slightly stronger between ages 30 and 39, which follows from the peak observed for conditional business portfolio shares at these ages (see Online Supplement). Third, equity market participation is at its lowest at the start of the working lifecycle, which makes it more likely that an individual will have heterogeneous aggregate investment risk exposures relative to its wealth rank neighbors (Figure 18, Appendix G).

5.6 Cross-cohort differences

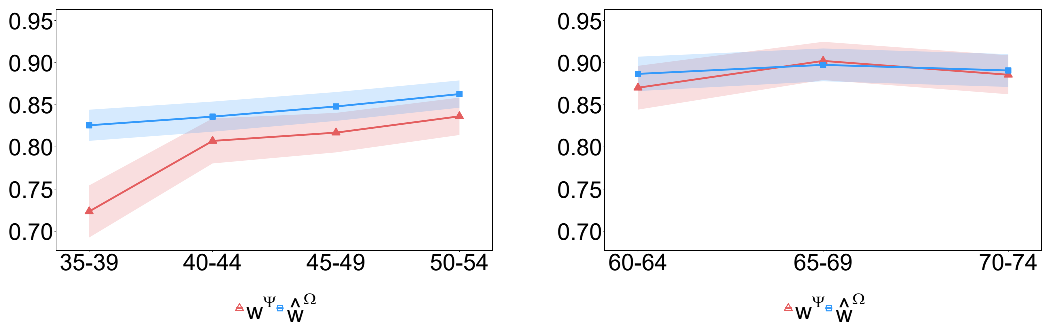
To investigate the time trend in intra-generational wealth mobility, I compare rank-rank coefficients β across birth cohorts based on actual wealth ranks κ^{Ψ} (Table 2). Given that most wealth

[13]The patterns shown here are even more pronounced when using age group 25-29 as starting point, as shown in the Online Supplement.

Figure 11: Rolling window analysis for rank-rank coefficient β.

(a) ***Working life: ages 30-54.***

(b) ***Older age: ages 55-74***



Note: the rank-rank coefficient β is computed with Ξ_{k-1} as initial stage and Ξ_k as final stage, where Ξ denotes working lifecycle stages and $k \in \{1, 2, 3, 4\}$. The reported data for stage k gives an indication of wealth mobility outcomes between this stage k and the previous stage $k-1$. For example, when $k = 3$, the cross-section of individuals' within-cohort wealth ranks at ages 45-49 is regressed on the cross-section at ages 40-44. The shaded areas display the 95% confidence intervals.

mobility occurs during working life and because of the lower sample size for the older age phase, I present only the results for the working life phase.

Overall intra-generational wealth mobility has dropped over time, which reflects a substantial decline in intra-generational wealth mobility at the top. Specifically, β-estimates are higher in the 1966-75 birth cohort compared to earlier cohorts (0.60 versus 0.54), but this conceals contrasting dynamics at the bottom and top of the wealth distribution. At the top, wealth consolidation during working life has increased significantly: the fraction of steady wealthy has risen from 3% in the 1946-55 cohort to close to 6% in the 1966-75 cohort. This has coincided with a strong drop in downward mobility from the top (the fraction of past wealthy has declined from over 3% to close to 2%) and a decrease in upward mobility to the top (the fraction of new wealthy has dropped from over 3% to below 2%). Instead, wealth consolidation at the bottom during working life has declined, as evidenced by the declining fraction of steady poor from roughly 9% to approximately 8%. This was accompanied by an increase in upward mobility from the bottom (the fraction of past poor has risen from over 3% to close to 5%), and a roughly stable degree of downward mobility to the bottom (the fraction of new poor has fluctuated around 3.5%).

Together with the findings from Section 5.1, the data therefore points towards a negative correlation between within-cohort wealth inequality and mobility at the top of the wealth distribution: the higher within-cohort wealth inequality in recent cohorts (Section 5.1) has coincided with stronger wealth consolidation (and thus weaker mobility) among the top 10% wealthiest. Instead, at the bottom, intra-generational persistence is found to have slightly declined over time. The net effect on overall wealth mobility was negative as well.

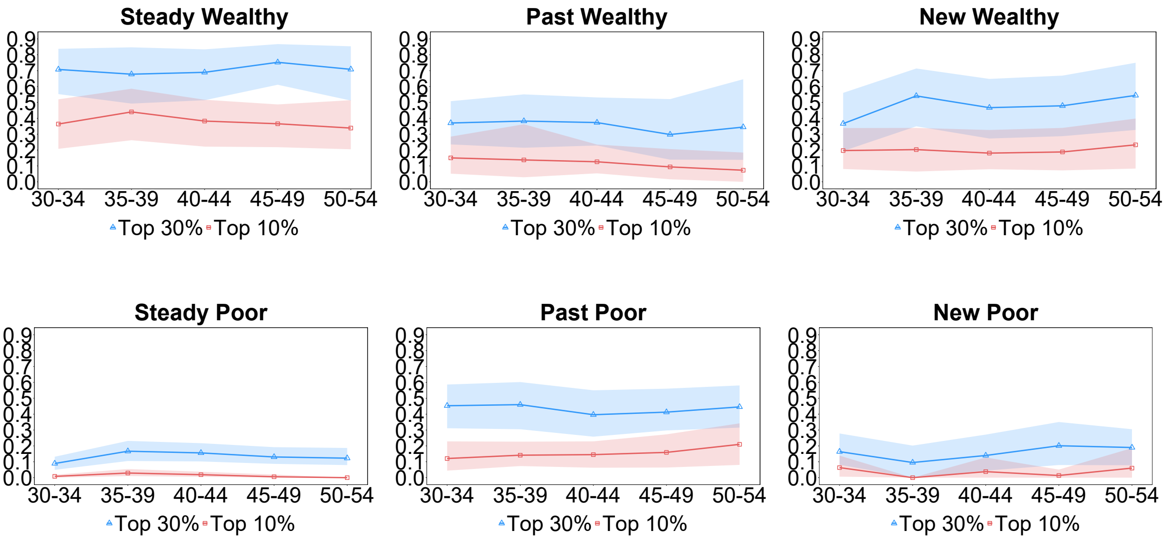
6 Within-family inter-dependence in intra-generational wealth mobility

I have established that there exists significant inter-generational persistence in within-cohort wealth ranks of parents and their children at identical points in their lifecycles (Section 4.1). Moreover, individuals' within-cohort wealth ranks at the start of working life overlap significantly with their parents' within-cohort wealth ranks at that time: the wealthiest individuals at age 30-34 are the most likely to have wealthy parents (Section 5.2). In this section, I build on these findings to investigate the third research question of this paper: does there exist inter-dependence between the within-cohort wealth rank trajectories of individuals and those of their parents (conditional on these being alive) as these individuals progress through working life?

In short, the answer is yes: there seems to be inter-dependence between individuals' wealth rank trajectories and those of their parents. Individuals who experience upward mobility from the bottom and to the top in their cohort (past poor, new wealthy) are likely to have parents that experience upward mobility in their own cohort as well. Furthermore, individuals that experience downward mobility from the top (past wealthy) often have parents that encounter downward wealth mobility also. Last, individuals that consolidate their position at the top of the wealth distribution are the most likely to have wealthy parents throughout their entire working life. These findings are quantified in Figure 12:

- Steady poor: throughout their working life, individuals that start and end working life in the bottom 20% (the steady poor) face a slightly increasing probability (from 9% to 12%) of having parents that belong to the top 30% of their within-cohort wealth distribution. On the contrary, the probability of having top 10% parents remains close to 0%.
- Past poor: individuals that display upward mobility from the bottom 20% to the top 50% of the within-cohort wealth distribution (the past poor) face a rising likelihood of having very wealthy parents: while the likelihood of having top 30% parents remains stable around 45%, the probability of having top 10% parents increases from 12% to 21%.
- New poor: for individuals that drop from the within-cohort top 50% to the bottom 20%, there is little inter-dependence with parental wealth ranks: the likelihood of having top 30% parents remains relatively stable around 17%, and the probability of having top 10% parents fluctuates between 0% and 5%.
- Steady wealthy: throughout their working life, individuals that start and end working life in the top 10% (the steady wealthy) have a 68% to 75% probability of having parents in the top 30% of their own within-cohort wealth distribution. Instead, the probability of having top 10% parents remains roughly constant around 36% for these individuals.

Figure 12: Inter-dependence between individuals' and their parents' wealth rank trajectories based on actual wealth ranks κ^{Ψ}.



Note: this plot uses the individuals from the working life sample defined in Section 5 of the paper. For a given discretionary group, it computes (for each lifecycle stage) the fraction of individuals in that group that have parents belonging to the top 10% and top 30% of their within-cohort wealth distribution at that historical point in time. Individuals that have no parents are excluded from the sample. The shaded areas display the 95% confidence intervals, which have been determined through bootstrapping.

- Past wealthy: individuals starting working life in the top 10% but dropping to the bottom 70% (the past wealthy) encounter a declining likelihood of having wealthy parents: the probability of having top 30% parents drops from 37% to 33% throughout working life. Furthermore, the likelihood of having top 10% parents declines from 15% to 7% by the end of working life.
- New wealthy: individuals that display upward mobility from the bottom 70% to the top 10% over working life (the new wealthy) face a strongly rising likelihood of having top 30% parents as these individuals' working life progresses: the likelihood of having top 30% parents rises from 36% at ages 30-34 to 54% at ages 50-54. The probability of having top 10% parents displays a relatively flat lifecycle profile between 19% and 23%.

Of course, this analysis does not take a stance on causality. For example, it may be that new wealthy individuals share part of their newly accumulated wealth with their parents via inter-vivos transfers (channel 1 in Section 2.3). On the contrary, the strong wealth accumulation of new wealthy individuals may relate to a reversal in their parents' fortunes that is transmitted through inter-vivos transfers or other channels (see Section 2.3). Moreover, the inter-dependence could be driven by exposures to identical sources of idiosyncratic risk, for instance a family business or a similar portfolio of stocks. A causal decomposition of the importance of these effects is left to future research.

7 Sources of mobility: composition analysis

As a final step in this paper, I conduct a composition analysis to investigate the inter-generational transfer receipts and socio-economic characteristics of individuals across the discretionary groups and clusters. The composition analysis should be interpreted as an exploratory exercise: it does not disentangle the causal driving forces behind wealth mobility dynamics, nor does it draw conclusions regarding the quantitative importance of the variables under consideration. The infeasibility of causal and quantitative identification follows from the presence of type and scale dependencies in families' and individuals' behavioral parameters. These create endogeneity between wealth accumulation and individuals' socio-economic characteristics (see Section 2.3). As an example, suppose one observes in the data high business ownership rates among the wealthiest individuals. This could relate to the easier access to business financing that wealth enables (a scale dependence). However, it could also reflect that only a subset of individuals hold valuable entrepreneurial ideas, generating higher wealth positions for these individuals over time (a type dependence).

The analysis is conducted on the intra-generational working life sample and therefore focuses on the sources of intra-generational (individual-level) wealth mobility over the working lifecycle. However, the Online Supplement demonstrates that the intra-generational findings extend to two-generational (family-level) mobility: the sources of inter-generational wealth mobility are found to be equivalent to those of intra-generational wealth mobility. This observation makes sense intuitively: as individual wealth ranks at ages 30-34 overlap to a large extent with family wealth ranks (Section 5.2), reversals in individuals' fortunes over the lifecycle generate similar reversals from an inter-generational perspective (at the family level). Furthermore, the Online Supplement shows that families consolidating their position at the bottom or top over two generations exhibit highly similar socio-economic characteristics between parents and children at the same age. Only for families with high (upward or downward) inter-generational wealth mobility do children's composition metrics diverge from those of their parents. This aligns with a literature documenting inter-generational persistence in socio-economic characteristics (e.g. Adermon et al., 2021; Charles & Hurst, 2003; Fagereng et al., 2021; Lindquist et al., 2015)[14].

In what follows, I first define the individual-level composition metrics used, and subsequently present the key findings with respect to inter-generational transfers and socio-economic characteristics. The latter are visualized in Appendix G. Given the top wealth bias of the PSID (see Section 3.2), the findings in this section relate to the entire wealth distribution and have little

[14] Some of the papers in this literature also disentangle the role of pre- versus post-birth factors using data for adoptees. In the PSID, such strategy is hard to implement due to the lack of extensive adoptee data.

say on the wealth accumulation dynamics of the very top wealthiest (top 1% and beyond). The composition of these very top wealthiest are the focus in for instance König et al. (2023) for Germany and Hubmer et al. (2024) for Norway.

Composition metrics The composition analysis is executed based on various individual-level metrics, detailed in Appendix F. The metrics are organized into four categories. First, a labor income and saving category calculates within-cohort labor income ranks, gross saving rates, non-mortgage debt participation and non-mortgage debt-to-income ratios (conditional on holding non-mortgage debt). Second, an asset ownership and allocation category computes homeownership, equity ownership, unincorporated business ownership, incorporated business ownership and mortgage participation rates. In addition, it calculates housing, equity, business and mortgage allocations relative to total assets (conditional on participation in the respective asset or debt market). Third, the health and household status category calculates whether an individual belongs to a household where at least one member has poor health and whether the individual is single, in a relationship or married. Fourth, the inter-vivos transfers and inheritances category assesses whether the individual has received an inter-generational transfer at any point in its lifecycle, and computes the ratio of its cumulative (capitalized) transfer receipts to its lifetime resources, in line with Black et al. (2022). Lifetime resources are defined as the cumulative sum of (capitalized) labor income[15]. These individual-level measures are summarized over the set of individuals in a discretionary group or cluster, as outlined in Appendix F.

Inter-generational transfers Inter-vivos transfers and inheritances are associated with wealth persistence at the top during working life. At ages 30-34, the wealth and cumulative transfer distributions overlap: top 10% individuals have received substantial transfers already, while the bottom 50% have hardly received any (see Section 5.2). Over their lifecycle, individuals consolidating their position at the top (steady wealthy, steady top) receive additional transfers: the proportion of recipients among these individuals rises to 60%-65% by ages 50-54, and their receipts make up around 11%-16% of lifetime resources at these ages. On the contrary, among the individuals stuck at the bottom (steady poor, steady bottom), the proportion of recipients rises to at most 20% by ages 50-54, and their receipts constitute a mere 4%-6% of lifetime resources.

The association between inter-generational transfers and upward wealth mobility is weaker. In comparison to the median American, past poor and new wealthy individuals are more likely to receive transfers (40%-45% by ages 50-54), which additionally comprise a more significant fraction of their lifetime resources (13%-14% by ages 50-54). It is therefore possible that these two groups include individuals that belong to wealthier families, but received inter-generational

[15]Unlike in Black et al. (2022), I do not include government transfers as part of lifetime resources. This assumption may induce an upward bias in the inter-generational transfers to lifetime resources variable for poorer individuals.

transfers only later in life relative to the steady money and steady top. Alternatively, these individuals' parents may have experienced favorable reversals in their fortunes only later in their lifecycle. On the contrary, the strong risers cluster is not linked with unusually high transfer receipts (approximately 6% of resources by ages 50-54).

Of course, these arguments do not by definition imply that these inter-vivos transfers and inheritances (their relative absence) are critical in consolidating individuals' position at the top (at the bottom). In fact, the inverse conclusion prevails: even for the wealthiest individuals their cumulative receipts constitute a limited fraction of lifetime resources (at most 16%). However, the comparatively high inter-generational transfer receipts of the consistently wealthy do indicate that these individuals are more likely to belong to wealthier families (in line with the finding in Section 6). They may therefore have benefited from their parental wealth through other channels (channels 2-5 in section 2.3). The apparent minimal importance of inter-vivos transfers and inheritances in generating inter-generational wealth persistence (the first channel in Section 2.3) aligns with other evidence for the United States (Charles & Hurst, 2003; Pfeffer & Killewald, 2018) and with results for Norway (Audoly et al., 2024). However, it contradicts findings for Sweden (Adermon et al., 2018).

Nonetheless, this conclusion regarding the minimal importance of inter-generational transfers warrants caution. There are two reasons for this. First, the PSID contains survey data, and is therefore prone to under-reporting of inter-generational transfers. On top of that, the transfer variable in the PSID likely suffers from a significant downward bias due to the irregular structure of the PSID survey waves (see Appendix F for a detailed explanation). Second, the timing of inter-generational transfers matters if there exist scale dependencies in individuals' behavior. For example, an early receipt of transfers may enable individuals to allocate higher fractions of their assets to high-return assets such as housing or businesses (e.g. Lee et al., 2020). In expectation, such early receipt of inter-generational transfers may therefore generate higher wealth accumulation over these individuals' lifecycle. Whether there actually exists heterogeneity in the timing of transfer receipts across individuals in the United States and whether such timing affects individuals' wealth rank trajectories over the lifecycle are questions that I leave to future research.

Socio-economic characteristics Persistence at the top (steady wealthy, steady top, steady subtop) is linked to high labor income, with individuals in these groups and clusters consistently belonging to the top 40% highest labor income earners over working life. High labor income is also associated with upward wealth mobility (past poor, new wealthy), although the evidence does not extend to the strong risers cluster. Instead, persistence at the bottom (steady poor, steady bottom, steady supra-bottom) and downward mobility to the bottom (new poor) are linked with low and declining ranks in the labor income distribution throughout working life. These results relate to Charles & Hurst (2003), who find that inter-generational income persistence explains half of the inter-generational persistence in wealth in the United States.

Their results are in line with those for the United Kingdom (Davenport et al., 2021; Levell & Sturrock, 2023). For the Nordic countries, Audoly et al. (2024) find human capital to be the main predictor of individuals' falling and rising over the wealth distribution (for Norway). Instead, Adermon et al. (2018) obtain that earnings and education only account for a quarter of two-generational wealth persistence (for Sweden).

Business ownership is linked to consolidation at the top (steady wealthy, steady top) and to significant downward mobility (past wealthy, middle decline, new poor). This suggests that business ventures can sustain or break individuals' and families' positions in the wealth distribution. The association between business ownership and upward wealth mobility is inconclusive, however: business ownership is clearly linked with the new wealthy, but not particularly with past poor or strong risers individuals. Matching evidence is provided for Norway: even though Audoly et al. (2024) do not find a marked role for business ownership in generating upward wealth mobility during working life, their results do show a clear correlation with consolidation at the top and downward wealth mobility. For the United States, both Charles & Hurst (2003) and Pfeffer & Killewald (2018) establish that business ownership has a non-negligible impact on inter-generational wealth persistence.

Individuals that are wealthy over the lifecycle (steady wealthy, steady top, steady subtop) or rise to the top (new wealthy, strong risers) display higher equity ownership rates compared to poor individuals. This relates to Charles & Hurst (2003), who find that equity ownership contributes significantly to inter-generational wealth persistence. Moreover, while homeownership and wealth ranks are positively correlated, wealthier individuals display lower conditional housing allocations. The disparity between the wealthy and poor is less pronounced for conditional equity allocations. Finally, persistence at the bottom (steady poor, steady bottom) and downward wealth mobility (new poor, past wealthy) are associated with poor and deteriorating health, a high likelihood of belonging to single households, and elevated and increasing non-mortgage indebtedness over the lifecycle.

8 Conclusion

Even though there exists an extensive body of research on social and income mobility, research on wealth mobility over the past decades is very limited. In this paper, I fill this gap for the United States by studying inter- and intra-generational wealth mobility using data from the Panel Study of Income Dynamics (PSID). The paper starts by providing two methodological contributions. First, I harmonize and validate the PSID-dataset. I argue that the PSID can be effectively used to study wealth-related questions, in particular those that relate to wealth mobility. Second, I construct a proxy wealth rank series using a gradient-boosting machine learning model that improves the housing proxies used in the literature. Throughout the paper, it is demonstrated that these proxies provide a useful tool for extending wealth mobility analyses across generations, although they underestimate the actual degree of wealth mobility.

Building on these two methodological contributions, I then formulate and provide insight into three research questions.

First, I study inter-generational (family-level) wealth mobility from a static perspective, comparing individuals' within-cohort wealth ranks to those of their parents and grandparents at specific lifecycle stages. In addition to providing a rich set of empirical moments and contrasting these to existing studies, I show that two-generational wealth mobility has declined over time and that inter-generational wealth mobility in the U.S. is lower than in most other countries with available data. Moreover, wealth mobility across three generations exceeds the mobility across two generations, although this effect is significantly stronger for mobility at the top than for mobility at the bottom. Finally, wealth rank resemblance between parents and their children increases with age (parent-child lifecycle bias), while wealth rank resemblance between grandparents and their grandchildren is higher when grandchildren are older than 35 years (grandchild lifecycle bias).

Second, this paper investigates intra-generational (individual-level) wealth inequality and mobility given the initial wealth rank distribution at ages 30-34. Within-cohort wealth inequality is found to be roughly stable over the lifecycle. Next, having provided a broad set of empirical moments, I show that intra-generational wealth mobility at the top has declined over time, and that the majority of wealth mobility occurs between ages 30 and 39. In addition, intra-generational wealth mobility is lower than in other countries with available data (in this case, the Nordic countries). Moreover, the composition analysis shows that persistence at the top is associated with the most substantial inter-vivos transfers and inheritances receipts. Individuals that are stuck at the bottom stand out by an overall absence of inter-generational transfers. Business ownership is linked with persistence at the top and downward mobility, while its association with upward mobility is inconclusive.

Third, this paper is the first to show that there exists inter-dependence between the within-cohort wealth rank trajectories of individuals and those of their parents (conditional on these being alive) over the same historical time period as individuals progress through working life. Specifically, individuals that face upward mobility from the bottom and to the top in their cohort are likely to have parents that encounter upward mobility in their own cohort as well. Vice versa, individuals that experience downward mobility from the top are likely to have parents facing downward wealth mobility also. Last, individuals that consolidate their position at the top are the most likely to have wealthy parents. These findings suggest the presence of altruism across generations and the exposure of parents and their children to identical sources of idiosyncratic risk.

9 References

Adermon, A., Lindahl, M., Waldenström, D. (2018). Intergenerational Wealth Mobility and the Role of Inheritance: Evidence from Multiple Generations. The Economic Journal 128, p. 482-513.

Adermon, A., Lindahl, M., Pälme, M. (2021). Dynastic Human Capital, Inequality, and Intergenerational Mobility. The American Economic Review 11 (5), p. 1523-1548.

Ager, P., Boustan, L., Eriksson, K. (2021). The Intergenerational Effects of a Large Wealth Shock: White Southerners After the Civil War. American Economic Review 111 (11), p. 3767–3794.

Atkeson, A., Irie, M. (2022). Rapid Dynamics of Top Wealth Shares and Self-Made Fortunes: What Is the Role of Family Firms? AER: Insights 4 (4), p. 409–424.

Audoly, R., McGee, R., Ocampo, S., Paz-Pardo, G. (2024). The Life-Cycle Dynamics of Wealth Mobility. Federal Reserve Bank of New York Staff Reports, no. 1097.

Bartscher, A., Kuhn, M., Schularick, M. (2024). The Distribution of Household Debt in the United States, 1950-2022. Available at SSRN: https://ssrn.com/abstract=4931920.

Bauluz, L., Meyer, T. (2024). The Wealth of Generations. Working Paper, CUNEF Universidad.

Black, S., Devereux, P., Lundborg, P., Majlesi, K. (2020). Poor Little Rich Kids? The Role of Nature versus Nurture in Wealth and Other Economic Outcomes and behaviors. Review of Economic Studies 87, p. 1683-1725.

Black, S., Devereux, P., Landaud, F., Salvanes, K. (2022). The (un)Importance of Inheritances. NBER Working Paper No. 29693.

Benhabib, J., Bisin, A., Luo, M. (2019). Wealth Distribution and Social Mobility in the U.S.: A Quantitative Approach. American Economic Review 109 (5), p. 1623-1647.

Blanden, J., Eyles, A., Machin, S. (2023). Intergenerational Home Ownership. The Journal of Economic Inequality 21, p. 251-275.

Borella, M., Bullano, F., De Nardi, M., Krueger, B., Manresa, E. (2024). Health Inequality and Health Types. NBER Working Paper No. 32799.

Boserup, S., Kopczuk, W., Kreiner, C. (2014). Stability and Persistence of Intergenerational Wealth Formation: Evidence from Danish wealth Records of Three Generations. Unpublished.

Boserup, S., Kopczuk, W., Kreiner, C. (2017). Intergenerational Wealth Formation over the Life Cycle: Evidence from Danish Wealth Records 1984-2013. Unpublished.

Boserup, S., Kopczuk, W., Kreiner, C. (2018). Born With A Silver Spoon? Danish Evidence on Wealth Inequality in Childhood. The Economic Journal 128, p. 514-544.

Charles, K., Hurst, E. (2003). The Correlation of Wealth across Generations. Journal of Political Economy 111 (6), p. 1555-1582.

Charles, K., Hurst, E., Killewald, A. (2013). Marital Sorting and Parental Wealth. Demography 50, p. 51-70.

Chetty, R., Friedman, J.N., Saez, E., Turner, N., Yagan, D. (2020). Income Segregation and Intergenerational Mobility Across Colleges in the United States. The Quarterly Journal of Economics 135(3), p. 1567–1633.

Conley, D., Glauber, R. (2008) Wealth Mobility and Volatility in Black and White. Washington, DC: Center for American Progress. Available from https://www.americanprogress.org.

Cooper, D., Dynan, K., Rhodenhiser, H. (2019). Measuring Household Wealth in the Panel Study of Income Dynamics: The Role of Retirement Assets. Working Papers, No. 19-6, Federal Reserve Bank of Boston, Boston, MA.

Davenport, A., Levell, P., Sturrock, D. (2021). Why Do Wealthy Parents Have Wealthy Children? The Institute for Fiscal Studies, London.

De Nardi, M., Fella, G. (2017). Saving and Wealth Inequality. Review of Economic Dynamics 26, p. 280-300.

De Nardi, M., Pashchenko, S., Porapakkarm, P. (2024). The Lifetime Costs of Bad Health. Review of Economic Studies 00, p. 1-40.

Deutscher, N., Mazumber, B. (2023). Measuring Intergenerational Income Mobility: A Synthesis of Approaches. Journal of Economic Literature 61 (3), p. 988-1036.

Dupont, B., Rosenbloom, J.L. (2022). Wealth mobility in the United States: 1860–1870. Social Science History 46(4), p. 801–829.

Fagereng, A., Mogstad, M., Ronning, M. (2021). Why Do Wealthy Parents Have Wealthy Children? Journal of Political Economy 129 (3), p. 703-756.

Fagereng, A., Guiso, L., Pistaferri, L. (2022). Assortative Mating and Wealth Inequality. National Bureau of Economic Research Working Paper No. 29903.

Fernholz, R. (2016). A Model of Economic Mobility and The Distribution of Wealth. Journal of Macroeconomics 50, p. 168-192.

Fernholz, R., Hagler, K. (2023). Rising Inequality and Declining Mobility in the Forbes 400. Economics Letters 230 111235.

Gomez, M. (2023). Decomposing the Growth of Top Wealth Shares. Econometrica 91 (3), p. 979-1024.

Gregg, P., Kanabar, R. (2023). Intergenerational Wealth Transmission in Great Britain. Review of Income and Wealth 63, p. 807-837.

Holmberg, J., Simmons, M., Trapeznikova, I. (2024). Parental Wealth and Early Labor Market Outcomes. Unpublished.

Hubmer, J., Krusell, P., Smith, A. (2021). Sources of U.S. Wealth Inequality: Past, Present and Future. NBER Macroeconomics Annual 35 (1), p. 391-455.

Hubmer, J., Halvorsen, E., Salgado, S., Ozkan, S., (2024). Why Are the Wealthiest So Wealthy? New Longitudinal Empirical Evidence and Implications for Theories of Wealth Inequality. Federal Reserve Bank of St. Louis Working Paper 2024-013.

Heathcote, J., Storesletten, K., Violante, G. (2010). The Macroeconomic Implications of Rising Wage Inequality in the United States. Journal of Political Economy 118 (4), p. 681-722.

Insolera, N., Simmers, B., Johnson, D. (2021). An Overview of Data Comparisons Between PSID and Other U.S. Household Surveys. Technical Series Paper, 21-02.

Jordà, O., Knoll, K., Kuvshinov, D., Schularick, M., Taylor, A. (2019). The Rate of Return on Everything, 1870–2015. The Quarterly Journal of Economics 134 (3), p. 1225–1298.

Kalsi, P., Ward, Z. (2025). The Gilded Age and Beyond: The Persistence of Elite Wealth in American History. NBER Working Paper No. 33355.

Kaplan, G., Violante, G., Weidner, J. (2014). The Wealthy Hand-to-Mouth. NBER Working Paper No. 20073.

Karagiannaki, E. (2017). The Effect of Parental Wealth on Children's Outcomes in Early Adulthood. Journal of Economic Inequality 15, p. 217-243.

Klevmarken, N.A., Lupton, J., Stafford, F. (2003). Wealth Dynamics in the 1980s and 1990s: Sweden and the United States. The Journal of Human Resources 38 (2), p. 322-353.

König, J., Schluter, C., Schröder, C. (2023). Routes To The Top. DIW Discussion Papers, No. 2066, Deutsches Institut für Wirtschaftsforschung (DIW), Berlin.

Kuhn, M. Schularick, M., Steins, U. (2020). Income and Wealth Inequality in America, 1949-2016. Journal of Political Economy 128 (9), p. 3469-3519.

Lee, H., Myers, D., Painter, G., Thunell, J., Zissimopoulos, J. (2020). The Role of Parental Financial Assistance in the Transition to Homeownership by Young Adults. Journal of Housing Economics 47.

Levell, P., Sturrock, D. (2023). Using Understanding Society to Study Intergenerational Wealth Mobility in the UK. Fiscal Studies 44, p. 417-432.

Lindquist, M., Sol, J., Van Praag, M. (2015). Why Do Entrepreneurial Parents Have Entrepreneurial Children? Journal of Labor Economics 33 (2), p. 269-296.

Mahler, L., Yum, M. (2024). Lifestyle Behaviors and Wealth-Health Gaps in Germany. Econometrica 92 (5), p. 1697-1733.

Menchik, P. (1979). Inter-generational Transmission of Inequality: An Empirical Study of Wealth Mobility. Economica 46, p. 349-362.

Mogstad, M., Torsvik, G. (2023). Family Background, Neighbourhoods and Intergenerational Mobility. NBER Working Paper No. 28874.

Pfeffer, F. (2018). Growing Wealth Gaps in Education. Demography 55, p. 1033-1068.

Pfeffer, F., Schoeni, R., Kennickel, A., Andreski, P. (2016). Measuring wealth and wealth inequality: Comparing two U.S. surveys. Journal of Economic and Social Measurement 41, p. 103-120.

Pfeffer, F., Killewald, A. (2018). Generations of Advantage. Multigenerational Correlations in Family Wealth. Social Forces 96(4), p. 1411-1442.

Saez, E., Zucman, G. (2016). Wealth inequality in the United States since 1913: Evidence from capitalized income tax data. Quarterly Journal of Economics 131 (2), p. 519-578.

Shiro, A., Pulliam, C., Sabelhaus, J., Smith, E. (2022). Stuck On The Ladder: Intragenerational Wealth Mobility in the United States. Working paper, Brookings Institution.

Siminsky, P., Yu, S.H. (2022). The Correlation of Wealth Between Parents and Children. The Australian Economic Review 55 (2), p. 195-214.

Smith, M., Zidar, O., Zwick, E. (2023). Top Wealth in America: New Estimates Under Heterogeneous Returns. Quarterly Journal of Economics 138 (1), p. 515-573.

Staiger, M. (2023). The Intergenerational Transmission of Employers and the Earnings of Young Workers. Unpublished manuscript, Opportunity Insights Harvard University.

Straub, L. (2019). Consumption, Savings, and The Distribution of Permanent Income. Unpublished Manuscript, Harvard University.

Van Langenhove, C. (2025). Wealth Inequality and Wealth Mobility in the United States: Type Dependence versus Scale Dependence. Forthcoming.

Wagner, S., Boertin, D., Gørtz, M. (2020). The Wealth of Parents: Trends Over Time in Assortative Mating Based on Parental Wealth. Demography 57, p. 1809-1831.

Xavier, I. (2021). Wealth Inequality in the U.S.: the Role of Heterogeneous Returns. Available at SSRN 3915439.

Zucman, G. (2019). Global Wealth Inequality. Annual Review of Economics 11, p. 109-138.

A Data

A.1 Waves & samples

This paper uses data from the Panel Study of Income Dynamics (PSID), which was conducted annually between 1968 and 1997, and bi-annually from 1999 until 2021. All waves infer about households' gross main housing value, gross main housing mortgage debt and rents paid. The waves in 1984, 1989, 1994 and 1999-2021 add to this questions about other assets and debts, allowing to define households' wealth.

The original 1968 PSID-sample consists of two independently drawn subsamples: (1) the SRC-subsample (Survey Research Center): a nationally representative sample of households, and (2) the SEO-sample (Survey of Economic Opportunities): an over-sample of low-income families. In 1990, a Latino subsample was added to the PSID, but this sample was dropped again from 1995 onwards. In 1997 and 2017, the PSID was permanently augmented with two representative immigrant subsamples to reflect the changing composition of the U.S. population. For each of these four subsamples, the PSID tracks over time the individuals belonging to the original set of households. In addition, it tracks individuals that descended from these original individuals, as well as non-sample individuals that entered the PSID through their connection to the former (e.g. a relationship or marriage).

The default in economic research using PSID-data is to focus on the SRC-subsample (e.g. Cooper et al., 2019; Heathcote et al., 2010; Kaplan et al., 2014; Straub, 2019). The SEO-subsample and the immigrant sub-samples are thus typically excluded from the analysis. In this paper, I follow this approach. As a robustness, I include the two representative immigrant samples from 1997 and 2017 onwards. The results for this alternative sample are presented in the Online Supplement. It shows that the conclusions of this paper are robust to the inclusion of these two immigrant samples.

Let us define the two core samples used in this paper. Denote N as the total number of households that to have responded to the PSID-questionnaires in at least one year between 1969 and 2021. Moreover, let us denote a specific household by subscript i. We have:

$$\mathcal{T}_\Omega = \{1969, 1970, \ldots, 1997, 1999, 2001, \ldots, 2021\} \tag{3}$$

$$\Omega = \{\mathbf{I}_i^\Omega(t) \mid i = 1, 2, \ldots, N, \; t \in \mathcal{T}_\Omega\} \tag{4}$$

where $\mathcal{T}_\Omega$ is the set of years corresponding to full sample Ω and $\mathbf{I}^\Omega$ denotes the vector of PSID variables that are available over $\mathcal{T}_\Omega$. 1968 is excluded from the sample due to the high number of outliers for this year. In addition, we can write:

$$\mathcal{T}_\Psi = \{1984, 1989, 1994, 1999, 2001, \ldots, 2021\} \tag{5}$$

$$\Psi = \{\mathbf{I}_i^\Psi(t) \mid i = 1, 2, \ldots, N, \; t \in \mathcal{T}_\Psi\} \tag{6}$$

where $\mathcal{T}_\Psi$ is the set of years corresponding to reduced sample Ψ and $\mathbf{I}_i^\Psi$ the vector of PSID variables that are available over $\mathcal{T}_\Psi$. It holds that $\mathbf{I}^\Psi = [\mathbf{I}^\Omega, \mathbf{I}^\Phi]$, where $\mathbf{I}^\Phi$ is defined as the vector of additional variables exclusive to sample Ψ.

A.2 Definitions

Unit of analysis The unit of analysis in the PSID-questionnaires is the family unit. Pfeffer et al. (2016) argue that the family unit may not always be equivalent to the household unit. For example, when an adult child that previously lived outside of the parental home moves back into the parental home, it will still be considered as a separate family unit even though its financial decisions, financial flows and wealth levels may be intertwined with the parents' ones. Still, this is more often than not a temporary situation, and it seems likely that at least some independence in financial decision-making, flows and wealth levels is maintained. For that reason, I do equate family units to households.

Wealth & wealth ranks In full sample Ω, the wealth categories are limited to gross main housing (h) and main housing mortgages outstanding (m). Non-homeowners are asked to report their rental payments (r). In contrast, reduced sample Ψ extends the PSID by including questions about a broader range of assets and debts. Beyond gross main housing, the asset categories encompass business holdings, equity holdings, fixed-income holdings, pension wealth, and gross other housing. On the liabilities side, besides main housing mortgages outstanding, households report the value of other housing debt and non-mortgage debt. I define household wealth w as the total of all asset categories minus the total of all debt categories. For a household i at time t, wealth is computed only if values are reported for every asset and debt category. If any category is missing, the household is considered a non-respondent at t.

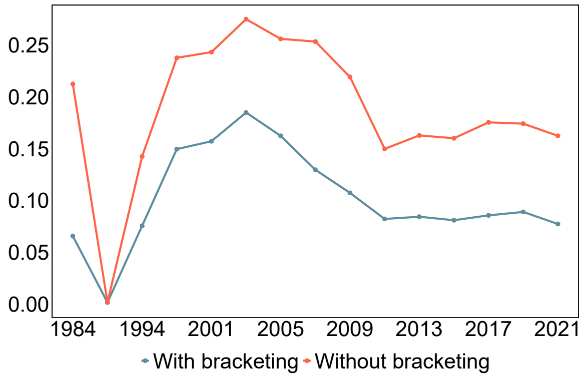
To study wealth mobility, the ultimate interest lays in households' wealth ranks, denoted as κ. Let $N(t)$ represent the total number of responsive households at time t, with their wealth levels given by $w_1(t), w_2(t), \ldots, w_{N(t)}(t)$. I define the wealth rank $\kappa_i(t)$ for household i at time t as:

$$\kappa_i(t) = \left\lceil \frac{100 \times \left(1 + \sum_{k=1}^{N_t} \mathbf{1}(w_k(t) < w_i(t))\right)}{N_t} \right\rceil \tag{7}$$

where $\mathbf{1}(w_k(t) > w_i(t))$ is an indicator function equal to 1 if $w_k(t) > w_i(t)$ and 0 otherwise. $\lceil \cdot \rceil$ denotes a ceiling function, which ensures the rank is placed into an integer bin from 1 to 100. Since the wealth variable w is defined exclusively in sample Ψ ($w \in \mathbf{I}^\Phi$) the same applies to κ: $\kappa_i \in \mathbf{I}^\Phi$.

Outliers & non-response The wealth-related sections of the PSID face two primary challenges. First, there is significant non-response, as shown in Figure 13. This issue is partially mitigated through the use of bracketing. Given that such bracketing effectively reduces non-

Figure 13: Fraction of non-respondent households with and without bracketing applied.



Note: this plot displays the fraction of households in each year that responded to the PSID-survey but displayed non-response for at least one wealth category. As a result, their total wealth for that year is undefined. The fraction of non-responsive households is shown for the dataset without bracketing applied and the dataset with bracketing applied. In 1989, non-response was close to zero.

response (Figure 13), I apply it whenever available. The details of the bracketing procedure are provided in Appendix B. Second, asset- and debt-related variables in Ψ are not harmonized over time, and both the reduced sample Ψ as the full sample Ω exhibit measurement errors. To address this, I have carefully aligned wealth categories across time periods, as discussed in Appendix B. Moreover, I have applied different outlier-correction procedures. These include on the one hand variable-specific outliers (see Appendix B) and on the other hand general outlier-correction procedures (see Online Supplement).

A.3 PSID-validation

In what follows, I validate the PSID by comparing its time trajectories for aggregate wealth (and its underlying components) to the trajectories for these variables in the top-wealth-adjusted Survey of Consumer Finances (SCF) (Figure 14). Furthermore, I set side by side the wealth shares observed in the PSID to those seen in the SCF (Figure 15).

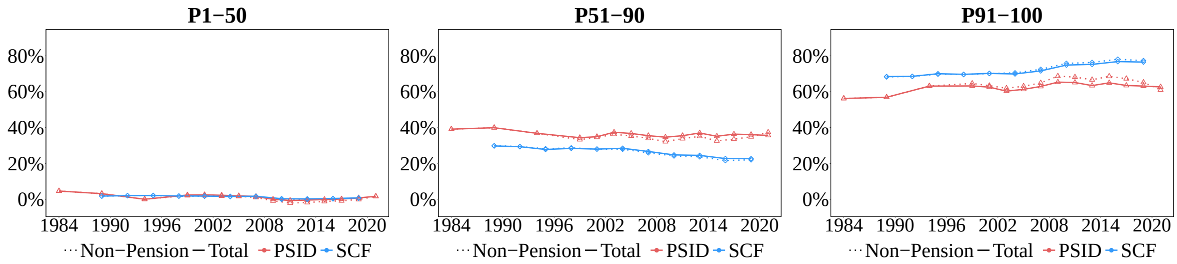
With respect to aggregate wealth, the PSID systematically underestimates all wealth categories compared to the SCF (Figure 14). This is consistent with previous findings (e.g., Pfeffer et al., 2016; Insolera et al., 2021). The underestimation is particularly strong for net business holdings. However, crucially given this paper's focus on wealth mobility, the PSID does accurately capture the time evolution of wealth and its underlying categories.

In addition, the evolution of wealth shares in the PSID aligns closely with those of the SCF: both databases indicate a slight increase in overall wealth inequality since the early 1980s (Figure 15). However, as noted in previous studies (e.g., Pfeffer et al., 2016; Cooper et al., 2019), the PSID underestimates top wealth inequality. For instance, in 2019, the top 10% wealth share

Figure 14: Average wealth levels per household (for total wealth and its underlying categories).

Note: these plots report the aggregate holdings of wealth and its different underlying categories, averaged across households. The outcomes are compared across the PSID and SCF databases over time.

Figure 15: Wealth shares (in %) in the PSID and SCF databases.



Note: these plots report the share of three commonly used wealth brackets (the bottom 50%, middle 50%-90% and top 10%) in aggregate wealth over time. I report outcomes both for non-pension wealth and for total wealth (consisting of non-pension and pension wealth). The wealth shares in the PSID are set side to side to those in the SCF.

(including pension wealth) equaled 62% in the PSID, compared to 77% in the SCF. The same top-wealth bias is observed when comparing the fraction of low- and high-wealth households in the PSID versus the SCF (see Online Supplement).

What explains this discrepancy between the PSID and SCF? While the SCF adjusts its nationally representative sample by oversampling at the top of the wealth distribution, the PSID does not. To address this, the PSID could in principle be supplemented with data from the Forbes 400 to better approximate top wealth (as is done for the distributional national accounts of Saez & Zucman (2016)). However, there are two key reasons against this approach. First, the composition of the Forbes 400 changes annually. Incorporating rich-list data into a wealth mobility study across the entire wealth distribution would therefore require making assumptions about the households that entered or exited the Forbes 400 during the period under consideration. This would introduce significant uncertainty into the wealth mobility analysis. Second, the primary focus of this paper is wealth mobility rather than wealth inequality. For wealth mobility measures, the number of households across the wealth distribution serves as the key calculation input. Excluding a small number of high-wealth households has a minimal impact on these outcomes. In contrast, wealth inequality metrics rely on the total wealth owned by households as the main calculation input. In such setting, excluding a small number of high-wealth households disproportionately skews the results downward. Therefore, correcting for top wealth is less critical in the context of this paper's focus on wealth mobility.

B Data definitions, outliers & non-response

B.1 Bracketing

Responding families in the PSID are occasionally unaware of the exact value of their wealth variables. In that case, for some years and variables, bracketing questions are provided. As an example, let x be the variable of interest, and let x_1, x_2 denote the thresholds, where it holds that $x_1 < x_2$. Using the answers to the bracketing question, I allocate x to one of the following three intervals: $[0, x_1[$, $[x_1, x_2[$ and $[x_2, +\infty[$. For the first two brackets, the actual x-value is estimated using the average of the lower and upper bound. For the last bracket, the estimate is calculated as $x_2 + \frac{1}{2}x_2$. When available, I apply this bracketing procedure for missing observations. In the Online Supplement, I have verified that the findings of this paper are robust to whether or not the bracketing procedure is applied.

B.2 Variable-specific definitions & outliers

B.2.1 Housing-related wealth categories

Main housing & rent Main housing mortgages outstanding are not reported in the years 1973-1975 and 1982. These are interpolated as follows. First, for all non-missing years in Ω, I compute the mortgage ratio as $\frac{h_i(t)}{m_i(t)}$. Second, a distance-weighted interpolation procedure (specified in the Online Supplement) is applied to the mortgage ratio over the missing years in Ω. Third, given the observed $h_i(t)$, the interpolated value for $\frac{h_i(t)}{m_i(t)}$ is used to trace out $m_i(t)$ for the missing years in Ω.

For the period 1969-1992, rental payments by renter households are reported on an annual basis. Instead, for the period 1993-2021, they are disclosed on a monthly basis. I define rents r on an annual basis, and therefore annualize the values for the latter period. In addition, rental payments are not provided for the years 1988-1989. These missing values are interpolated using a distance-weighted linear interpolation (specified in the Online Supplement). Furthermore, in 1970, reported rents for a select subset of homeowners takes on the value '768'. These outliers are set to their correct value of zero.

Other housing For the period 1984-2011, other housing is reported net of mortgage debt. Instead, for the period 2013-2021, gross other housing and mortgage debt on other housing are reported separately. To compute portfolio allocations (in Appendix F), our interest lays in the gross representation. I therefore calculate the average mortgage ratio on other housing conditional on ownership using data from the Survey of Consumer Finances (SCF) from 1989 to 2019. This mortgage ratio is found to equal 52 percent. I then use this ratio to trace out approximations for gross other housing and mortgages outstanding on other housing for the period 1984-2011.

B.2.2 Non-housing wealth categories

Business & equity For the period 1984-2011, business holdings are reported net of business debts. Instead, for the period 2013-2021, gross business assets and debts are reported separately. Between 2013 and 2021, I therefore compute the net measure. Additionally, there exist a handful of observations for net business holdings that take on unrealistically large negative values (for one survey wave only). These outliers are set equal to zero.

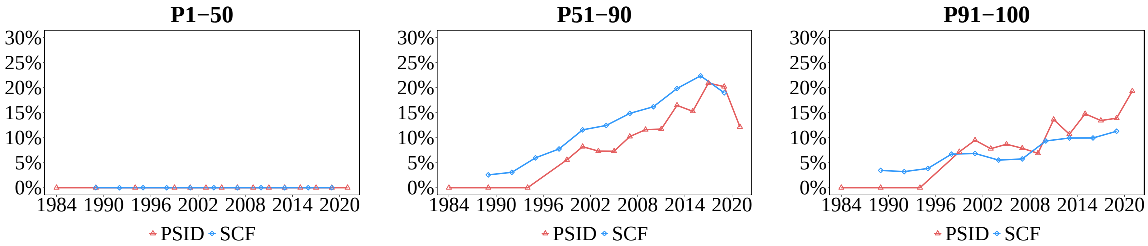
Equity holdings are defined as the cumulative value of stocks in publicly-traded corporations, stock market mutual funds or investment trusts. However, in the period 1984-1997, this variable also includes holdings of stocks in IRAs. Similar to business holdings, there are a handful of observations that take on unrealistically large negative values for one survey wave only. These values are corrected to zero.

Fixed income For the period 1984-1997, fixed income is computed as the sum of two survey questions. In a first question, labeled as 'baseline fixed income' in the variable codes in the Online Supplement, households report the cumulative value of their checking accounts, saving accounts, money market fund holdings, certificates of deposits, government savings bonds and Treasury bills, including those held in IRAs. In a second question, labeled as 'other' in the Online Supplement, the household is asked about the cumulative value of any other assets, including bond funds and cash value of lifecycle insurance values. For the period 1999-2017, there exists a minor difference: the questions are the same as for the 1984-1997 period, but fixed income IRAs are now inferred about in a separate question and are therefore excluded from the fixed income variable. This shows up as a minor trend-change for the fixed income variable in 1999 (Figure 14). Finally, for the period 2019-2021, the 'baseline fixed income' question is split up into two separate questions: on the one hand a question on checking accounts, saving accounts and money market funds, and on the other hand a question on certificates of deposits, government bonds and treasury bills. The 'other' question remains unchanged. For the period 2019-2021, fixed income is then computed as the sum of the reported values over the three questions (two baseline fixed income questions and the other question).

B.2.3 Pension wealth

Pension wealth is calculated as the sum of (1) defined contribution plans, and (2) IRAs and private annuities. On the one hand, defined contribution account values are reported only from 1999 onwards and equal the sum of the reported values in defined contribution accounts held by the reference person and by the partner. These comprise not only the plans held with the current employer, but also those held with the two previous employers from both individuals. On the other hand, as noted earlier, IRA-and private annuity wealth are inferred about in a separate question from 1999 onwards. This structure implies that pension wealth prior to 1999 will equal zero (Figure 14). Also for pension wealth, in some years, there are one-off outliers that take on a negative value of multiple billions. These are set equal to zero.

Figure 16: Ratio of pension to total wealth across the non-pension wealth distribution.



Note: these plots display the ratio of pension wealth to total wealth across three wealth brackets: the bottom 50%, middle 50%-90% and top 10%. Households have been allocated to the one of the three brackets based on their rank in the non-pension wealth distribution. For the PSID, pensions equal zero prior to 1999 given that pensions are not inquired about in the PSID-questionnaire for these years.

There exist two difficulties related to the measurement of pension wealth. First, IRA wealth is included in equity and fixed income questions prior to 1999, and inferred about in a separate question from 1999 onwards. There does not exist a straightforward method of separating the IRA-proportion of the equity and fixed income questions prior to 1999, nor a reliable method of allocating IRA wealth to equity and fixed income afterwards. Therefore, I keep IRA wealth as part of the equity and fixed income variables prior to 1999, and include it in pension wealth from 1999 onwards. As this discrepancy affects only the portfolio share calculations and not households' total wealth levels or ranks, its impact on the findings in this paper is minimal. Second, from 1999 onwards, defined contribution plan wealth is included in the calculation of w, while it is not in the years prior. As our ultimate interest lays in wealth ranks κ, this shift may be problematic insofar as there exists heterogeneity in pension wealth across the non-pension wealth distribution. Figure 16 shows that this heterogeneity is relatively limited: the share of pension to total wealth displays roughly similar levels and time-trajectories across the wealth bins. Nevertheless, as a robustness, in the Online Supplement I show that the main conclusions of the paper continue to hold when restricting the wealth variable to non-pension wealth.

B.2.4 Non-mortgage debt

In the period 1984-2009, the PSID captures non-mortgage debt through a variable 'other debt'. In 2011, the 'other debt' variable is subdivided into credit card debt, student loan debt, medical debt and debt to relatives. I then calculate non-mortgage debt as the sum of these four categories. In the period 2013-2021, a residual debt category is added to the four categories from the 2011-wave. I include it as part of non-mortgage debt. This shift in definitions implies that the underlying non-mortgage debt variable may be slightly different across the three periods. In particular, the absence of a residual category in 2011 might imply a minor underestimation of non-mortgage debt compared to the other years. However, the impact of this exclusion is marginal, as evidenced by the absence of a major trend-shift for non-mortgage debt in 2011 (Figure 14).

C ML-proxies over the full sample Ω

C.1 Framework

Data on wealth w is available over the reduced sample $\mathcal{T}_\Psi$, which begins only in 1984. This leads to two limitations. First, it restricts a comparison of wealth mobility outcomes across birth cohorts. Second, it limits the feasibility of an inter-generational wealth mobility analysis, particularly in examining grandparent-grandchild wealth linkages (across three generations). However, gross main housing value h (for homeowner households) and rental payments r (for renter households) are available over the full period $\mathcal{T}_\Omega$. To approximate wealth over the entire period $\mathcal{T}_\Omega$, it is therefore common to estimate wealth based on h or r:

$$\hat{w}_i^\Omega(t) = \begin{cases} \hat{f}_h(\mathbf{x_h})h_i^\Omega(t) & \text{if } h_i(t) > 0 \\ \hat{f}_r(\mathbf{x_r})r_i^\Omega(t) & \text{if } h_i(t) = 0 \end{cases} \tag{8}$$

where $\hat{w}_i^\Omega(t) = \hat{w}_i(t) \mid t \in \mathcal{T}_\Omega$ represents the predicted wealth level over $\mathcal{T}_\Omega$. For homeowners, wealth is approximated by multiplying the observed main housing value $h_i^\Omega(t)$ (available for $t \in \mathcal{T}_\Omega$) by a scaling factor $\hat{f}_h$. For renters, wealth is approximated in parallel using observed rental payments $r_i^\Omega(t)$ and a scaling factor $\hat{f}_r$.

While $h_i^\Omega(t)$ and $r_i^\Omega(t)$ constitute variables that are directly observable, the scaling factors $\hat{f}_h$ and $\hat{f}_r$ need to be estimated as a function of some vector of input variables available over $\mathcal{T}_\Omega$. Let us define these input vectors as $\mathbf{x_h} = \mathbf{x_{h,i}^\Omega(t)} = \mathbf{x_{h,i}(t)} \mid t \in \mathcal{T}_\Omega$ for homeowners and $\mathbf{x_r} = \mathbf{x_{r,i}^\Omega(t)} = \mathbf{x_{r,i}(t)} \mid t \in \mathcal{T}_\Omega$ for renters.

C.2 Common assumptions

Existing literature (e.g. Chetty et al., 2020; Pfeffer & Killewald, 2018) makes the assumption that $\hat{f}_h = C$ and $\hat{f}_r = 0$, where C is some fixed number such as the average or median wealth-to-gross main housing value ratio in the sample. Mathematically:

$$\hat{w}_i^\Omega(t) = \begin{cases} Ch_i^\Omega(t) & \text{if } h_i(t) > 0 \\ 0 & \text{if } h_i(t) = 0 \end{cases} \tag{9}$$

implying that total wealth is approximated as main housing value for homeowners, while renters are assumed to have zero wealth. When studying wealth mobility $-$ where the interest lays in wealth ranks rather than absolute wealth levels $-$ the correctness of this approach hinges on the assumptions of (1) main housing values being positively correlated with wealth levels, (2) this relationship being stable over time, and (3) renters having zero wealth.

For the first assumption, the Pearson correlation coefficient in the PSID between gross main housing value and wealth over Ψ equals 0.66. However, there exists substantial heterogeneity in the homeowner scaling factors across households: the standard deviation of this variable equals 2.23. This suggests that the constant C constitutes a strong simplification. For the second assumption, while the median renter scaling factor indeed equals 0, a non-negligible proportion of renter households reports positive wealth levels. Furthermore, there exists significant heterogeneity in wealth levels among renters: the standard deviation for the renter scaling factor equals 10.03.

As a result, the proxy in Equation 9 can be improved by accounting for household heterogeneity in homeowner and renter scaling factors. To address this, in Section C.3, I estimate two machine learning (ML-models) that incorporate additional household-level information available in the PSID-dataset in full sample Ω. In Section C.4, I define four housing proxies, which represent variations to Equation 9. These serve as benchmarks against which the performance of the ML-models can be compared in Section C.5. The results demonstrate that the ML models significantly outperform the housing proxies.

C.3 ML-models

In what follows, I construct and estimate a gradient-boosting (GB)-model to predict scaling factors $\hat{f}_h$ and $\hat{f}_r$. Additionally, in the Online Supplement, I develop an alternative ML-model $-$ a multi-layer perceptron (MLP) model $-$ to which the performance of the GB-model can be compared. Both ML-models are trained and tested on observable sample Ψ, and estimated for homeowners and renters separately. Their inputs $\mathbf{x_h}$ and $\mathbf{x_r}$ consist of household-level variables available over the full sample period $\mathcal{T}_\Omega$. The models can then be used to make predictions over $\mathcal{T}_\Omega$.

The construction of the ML-models proceeds in three steps. First, I define the inputs $\mathbf{x_h}$ and $\mathbf{x_r}$ used by the models. Second, I outline the equations of the (homeowner and renter) GB-model, with a detailed derivation provided in the Online Supplement. Cross-validation is employed to determine the optimal hyperparameters. Thereafter, the GB-model is estimated. The full development of the (homeowner and renter) MLP-model is presented in the Online Supplement also. Third, I perform a series of diagnostic tests on ML-model outcomes, with the procedures and results again described in the Online Supplement.

Input variables The selection of the input variables $\mathbf{x_h}$ and $\mathbf{x_r}$ occurs according to two criteria. A first criterion is availability: the variable should be available over the full period $\mathcal{T}_\Omega$, or equivalently $x \in \mathbf{I}^\Omega(t)$. Due to the limited number of variables in $\mathbf{I}^\Omega(t)$, this criterion imposes a relatively strong restriction. A second criterion is relevance: the variable should contribute to the predictive performance of the ML-models. Based on these two criteria, and defining $A, B, C \in \mathbb{R},\ A, B, C < \infty$, the following input variables are selected for the homeowner and renters ML-models:

1. Labor income $\frac{y_i(t)}{\bar{y}(t)}$: the household's labor income $y_i(t)$ normalized by the average labor income across all households $\bar{y}(t)$.
2. Capital income $\frac{\gamma_i(t)}{\bar{\gamma}(t)}$: the household's capital income $\gamma_i(t)$ relative to the average capital income income across all households $\bar{\gamma}(t)$.
3. Household size $h_i^n \in [1, A]$: the number of individuals living in household i, comprising the reference person, partner, and children.
4. Household status $h_i^s \in \{0, 1, 2\}$: indicates whether the reference person is single (0), in a relationship with the partner (1), or married to the partner (2).
5. Age $h_i^a \in [1, B]$: the age of the oldest individual in the household (between the reference person or the partner).
6. Business ownership $n_i^b \in \{0, 1, 2\}$: indicates whether the household does not own a business (0), owns an unincorporated business (1), or owns an incorporated business (2).
7. Health status $h_i^h \in [0, 1]$: the proportion over the past four years in which at least one of the core household members was unable to work due to poor health.
8. Cars per adult $\frac{h_i^c}{\bar{h}(t)}$, with $h_i^c \in [1, C]$: the number of cars per adult owned by the household, normalized by the sample median at time t.

Different outlier correction procedures are applied to these input variables. For cars per adult, data is missing for years 1973-1974 and 1987-1997. To address this, I apply to this variable the distance-weighted interpolation procedure outlined in the Online Supplement. The PSID-questionnaire codes for the input variables are provided in the Online Supplement as well.

In addition to variables (1)-(8), the inputs for the homeowner ML models ($\mathbf{x_h}$) include normalized gross main housing values $h_i(t)/\bar{h}(t)$ and the mortgage ratio $m_i(t)/h_i(t)$. For the renter models, the inputs ($\mathbf{x_r}$) also include normalized rental payments $r_i(t)/\bar{r}(t)$. These additional variables allow for scale dependence between the scaling factors ($\hat{f}_h$ and $\hat{f}_r$) and the value of the household's residence. For instance, as households accumulate more wealth, they may transition to more valuable houses, but the value of the house or corresponding rental payments might constitute a declining proportion of their total wealth over time.

Estimation & cross-validation The homeowner and renter GB-model is estimated over sample $\mathcal{T}_\Psi$. Observations over $\mathcal{T}_\Psi$ are divided into a randomly generated training set and testing set. I use a mean squared error (MSE) loss function, which is the benchmark in the literature.

The predictions for the scaling factors over full sample $\mathcal{T}_\Omega$ of the GB-model are given by:

$$\hat{f}_h^{\mathrm{GB}} = \hat{f}_h^{M_h^*}(\mathbf{x_h}) = \hat{f}_h^{(0)} + \sum_{m=1}^{M_h^*} \lambda_h^* g_h^{(m)}(\mathbf{x_h}) \tag{10}$$

$$\hat{f}_r^{\mathrm{GB}} = \hat{f}_r^{M_r^*}(\mathbf{x_r}) = \hat{f}_r^{(0)} + \sum_{m=1}^{M_r^*} \lambda_r^* g_r^{(m)}(\mathbf{x_r}) \tag{11}$$

where a detailed derivation is provided in the Online Supplement. $\hat{f}^{(0)}$ denotes the initial guesses and $g^{(m)}$ is the weak learner at iteration m. The hyperparameters include the optimal number of boosting rounds M^*, the optimal learning rate λ^*, and the optimal maximum depth of a tree d^*. Predictions for $\hat{w}_i^\Omega(t; \mathcal{M}_{\mathrm{GB}})$ are obtained by substituting $\hat{f}_h^{\mathrm{GB}}$ and $\hat{f}_r^{\mathrm{GB}}$ into Equation 8.

To optimize the hyperparameters, a k-fold cross-validation is performed separately for the homeowner and renter GB-model. Using the MSE loss function $\mathcal{L}_{\mathrm{CV}}$, the average cross-validation losses are defined as:

$$\mathcal{L}_{\mathrm{CV}}(M_h, d_h, \lambda_h) = \frac{1}{k} \sum_{j=1}^{k} \mathcal{L}^{(j)}(M_h, d_h, \lambda_h) \tag{12}$$

$$\mathcal{L}_{\mathrm{CV}}(M_r, d_r, \lambda_r) = \frac{1}{k} \sum_{j=1}^{k} \mathcal{L}^{(j)}(M_r, d_r, \lambda_r) \tag{13}$$

where I set $k = 10$, consistent with standard practices. The optimal hyperparameters are obtained by minimizing the cross-validation loss:

$$(M_h^*, d_h^*, \lambda_h^*) = \arg\min \mathcal{L}_{\mathrm{CV}}(M_h, d_h, \lambda_h) \tag{14}$$

$$(M_r^*, d_r^*, \lambda_r^*) = \arg\min \mathcal{L}_{\mathrm{CV}}(M_r, d_r, \lambda_r) \tag{15}$$

For the homeowner GB-model, the resulting optimal hyperparameters equal $M^* = 140$, $d^* = 9$ and $\lambda^* = 0.045$. Those for the renter GB-model are given by $M^* = 90$, $d^* = 6$ and $\lambda^* = 0.06$. In the Online Supplement, I compute summary metrics of the SHAP-values for the homeowner and renter GB-model across all observations.

C.4 Housing measures

I aim to evaluate whether the predictions of the GB and MLP models outperform proxies that neither rely on an optimization procedure nor utilize all available information in the sample Ω, as is typically the case in the existing literature (e.g. Chetty et al., 2020; Pfeffer & Killewald, 2018). In the following, I define four such proxies, referred to as housing proxies.

A first housing proxy, denoted as $\hat{w}_i^{\Omega}(t; \mathcal{M}_{\text{NP1}})$, is defined by Equation 9. It assumes that homeowners' wealth equals C-times their main housing value, while renters' wealth is zero. Since we are ultimately interested in wealth rankings, the value of C is irrelevant as long as $C > 0$. The second housing proxy $\hat{w}_i^{\Omega}(t; \mathcal{M}_{\text{NP2}})$, third housing proxy $\hat{w}_i^{\Omega}(t; \mathcal{M}_{\text{NP3}})$ and fourth housing proxy $\hat{w}_i^{\Omega}(t; \mathcal{M}_{\text{NP4}})$ attempt to refine the estimation of renters' wealth.

The second housing proxy is defined as:

$$\hat{w}_i^{\Omega}(t; \mathcal{M}_{\text{NP2}}) = \begin{cases} C h_i^{\Omega}(t) & \text{if } h_i(t) > 0 \\ C \frac{r_i^{\Omega}(t)}{v(t)} & \text{if } h_i(t) = 0 \end{cases} \tag{16}$$

where C is again a fixed number, and $v(t)$ the rental yield, which is taken from Jordà et al. (2019). This proxy (1) assumes that rental yields are uniform across houses, (2) approximates the value of renters' residence as the inverse of the rental yield, and (3) assumes that renters' wealth corresponds to the value of the house they occupy. However, given that the median wealth of renters equals zero, the latter assumption seems particularly strong. To address this, a third housing proxy is introduced:

$$\hat{w}_i^{\Omega}(t; \mathcal{M}_{\text{NP3}}) = \begin{cases} \bar{C}_h h_i^{\Omega}(t) & \text{if } h_i(t) > 0 \\ \bar{C}_r \frac{r_i^{\Omega}(t)}{v(t)} & \text{if } h_i(t) = 0 \end{cases} \tag{17}$$

where $\bar{C}_h$ is the average scaling factor to gross main housing value for homeowners, and $\bar{C}_r$ is the average scaling factor for renters' estimated housing values. Both are calculated over the sample $\mathcal{T}_{\Psi}$. That is:

$$\bar{C}_h = \frac{1}{n} \sum_{i=1}^{n} \frac{w_i^{\Psi}(t)}{h_i^{\Psi}(t)} \qquad \text{if } h_i(t) > 0 \tag{18}$$

$$\bar{C}_r = \frac{1}{n} \sum_{i=1}^{n} \frac{w_i^{\Psi}(t)}{r_i^{\Psi}(t)} v(t) \qquad \text{if } h_c(t) = 0 \tag{19}$$

Finally, to mitigate the influence of outliers in scaling factors, a fourth housing proxy is defined as:

$$\hat{w}_i^{\Omega}(t; \mathcal{M}_{\text{NP4}}) = \begin{cases} \tilde{C}_h h_i^{\Omega}(t) & \text{if } h_i(t) > 0 \\ \tilde{C}_r \frac{r_i^{\Omega}(t)}{v(t)} & \text{if } h_i(t) = 0 \end{cases} \tag{20}$$

where $\tilde{C}_h$ and $\tilde{C}_r$ are defined analogously to Equations 18 and 19, but calculate the medians instead of the averages.

C.5 Performance comparison

Given the wealth predictions from the optimal ML models and the four housing proxies, the approximated wealth rank series can be determined. These ranks represent the ultimate objects of interest and are defined as:

$$\hat{\kappa}_i^{\Omega}(t;\chi) = \left\lceil \frac{100 \times \left(1 + \sum_{k=1}^{N_t} \mathbf{1}(\hat{w}_k^{\Omega}(t;\chi) < \hat{w}_i^{\Omega}(t;\chi))\right)}{N_t} \right\rceil \tag{21}$$

where $\chi = \{\mathcal{M}_{\text{GB}}, \mathcal{M}_{\text{MLP}}, \mathcal{M}_{\text{NP1}}, \mathcal{M}_{\text{NP2}}, \mathcal{M}_{\text{NP3}}, \mathcal{M}_{\text{NP4}}\}$.

To evaluate the performance of the two ML models and the four housing measures, I compare the proxy wealth ranks ($\hat{\kappa}$) to the actual ones (κ) over the testing set. Performance metrics include the mean squared error (MSE), mean absolute error (MAE), and the proportion of wealth rank predictions that deviate by more than 25 and 50 ranks to actual ones. These metrics are summarized using two approaches. In a first approach, the performance metric $\mathcal{M}$ is calculated for each year and averaged across years:

$$\mathcal{M}_t = \frac{1}{N_t} \sum_{i=1}^{N_t} m(a_{i,t}, p_{i,t}), \quad \mathcal{M} = \frac{1}{T} \sum_{t=1}^{T} \mathcal{M}_t \tag{22}$$

In a second approach, the performance metric $\mathcal{M}$ is computed for each household and averaged across all households:

$$\mathcal{M}_i = \frac{1}{T_i} \sum_{t \in \mathcal{V}_i} m(a_{i,t}, p_{i,t}), \quad \mathcal{M} = \frac{1}{I} \sum_{i=1}^{I} \mathcal{M}_i \tag{23}$$

where $\mathcal{V}_i$ denotes the set of valid time points for individual i. $m(a, p)$ represents the specific calculation for the chosen metric, such as $(a - p)^2$ for MSE or $|a - p|$ for MAE.

The performance results are displayed in Table 3. Two key findings persist. First, across the housing proxies, the third housing proxy consistently displays superior performance. Second, the housing proxies' performance does not come close to those of the ML-models. Moreover, between the ML-models, it is the GB-model that outperforms the MLP-model. Therefore, in the following sections, I use the GB-model predictions as a proxy for wealth and wealth ranks over the sample $\mathcal{T}_{\Omega}$. To simplify notation, I define:

$$\hat{w}_i(t) = \hat{w}_i^{\Omega}(t; \mathcal{M}_{\text{GB}}), \quad \hat{\kappa}_i(t) = \hat{\kappa}_i^{\Omega}(t; \mathcal{M}_{\text{GB}}) \tag{24}$$

Despite the superior performance of the GB model, a significant number of predictions remains inaccurate (Table 3). On average, 9% of wealth rank predictions deviate by more than 25 ranks from their actual values in any given year, while approximately 1% of the predictions diverge by more than 50 ranks. Additionally, 21% of households in the sample experience a wealth rank misallocation of at least 25 ranks at some point during their lifecycle. When the misallocation threshold is raised to 50 ranks, this proportion drops to 3%.

Table 3: Model performance for housing and machine learning wealth proxies.

Across years

Proxy	MSE	MAE	≥ 25	≥ 50
NP1	453.88	15.58	0.20	0.04
NP2	909.32	23.57	0.41	0.10
NP3	429.09	15.42	0.19	0.03
NP4	528.38	17.53	0.26	0.04
MLP	238.51	11.02	0.10	0.01
GB	195.67	10.00	0.08	0.01

Across households

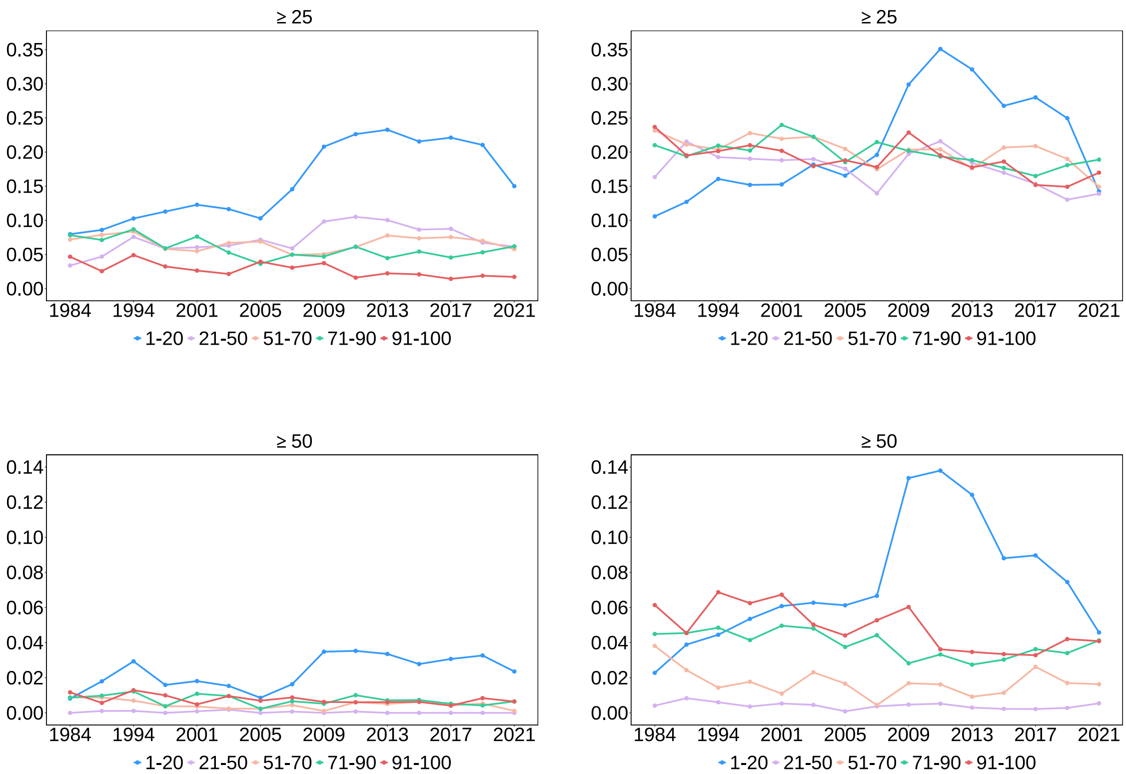
Proxy	MSE	MAE	≥ 25	≥ 50
NP1	438.17	15.10	0.37	0.09
NP2	925.15	23.58	0.61	0.21
NP3	410.61	14.87	0.36	0.08
NP4	510.56	17.00	0.44	0.09
MLP	235.76	10.68	0.23	0.03
GB	196.15	9.80	0.19	0.02

Note: panel (a) computes the performance metrics per year and averages across time. For example, in the average year, 8% of households have their wealth ranks misallocated by at least 25 units based on the GB-proxy. Instead, in panel (b), the performance metrics are calculated per household and are averaged across households. For example, 19% of households have their wealth rank misallocated by at least 25 wealth rank units at some point in this household's existence based on the GB-proxy.

Figure 17 highlights the timing of misallocations and its distribution over actual wealth. Three key observations emerge. First, the GB-proxy series outperforms the housing proxy series in all time periods and across all actual wealth levels. Second, for both the GB- and housing proxy, misallocations are more common among poor households (those belonging to the bottom 20%). This makes sense: wealth levels near the bottom of the distribution are closer to zero, so that small errors in estimated scaling factors disproportionately affect wealth ranks. Third, for the poor households, there exists time variation in the likelihood of misallocation: the degree of misallocation was significantly higher during and in the aftermath of the global financial crisis of 2008. This effect holds for both the GB-proxy and housing proxy, but is significantly stronger for the latter.

Figure 17: Proportion of misallocated households per wealth bin (according to actual wealth) for the GB-proxy and the third housing proxy.

GB-proxy (left panel) *NP3-proxy (right panel)*



Note: this plot reports the fraction of households that is misallocated by at least 25 wealth rank units (upper panel) or 50 wealth rank units (lower panel) for each year. The left panels report the outcomes for the GB-proxy series (GB), while the right panel does so for the third housing proxy (NP3). Households are allocated to wealth bins according to their actual wealth levels.

D Empirical strategy

D.1 Individual-level

Notation & eligibility Ultimately, our focus is on the mobility of individuals, rather than households. This requires taking into account that individuals may switch households over time. For instance, an individual living alone may begin cohabiting with a partner or get married, causing the original household (e.g. $i = 1$) to dissolve and a new household with different characteristics (e.g. $i = 2$) to be formed. Such transitions might influence the individual's wealth positively or negatively. Let us write variable z of an individual j belonging to household i at t as $z_j(t, i)$, where i may vary over time.

I restrict the analysis to individuals that have at least some control over their finances, and $-$ consequently $-$ influence the decisions of the household to which they belong. Therefore, I limit the PSID-sample to individuals identified as either the reference person or the partner within their household i. I designate an individual as partner if its relationship to the reference person is classified in the individual-level PSID file as legal spouse, partner, uncooperative legal spouse, or other non-relatives (which primarily includes same-sex partners).

Wealth levels & rankings A key question regarding individual-household linkages is how to allocate household-level wealth categories and total wealth w_i to the individual level. This allocation is performed using the household status variable h_i^s, which was defined in Section C.3. I use the following allocation rules:

1. Single individual ($h_i^s = 1$): when the household consists of a single financially-independent individual, the entire household-level wealth $w(i)$ is allocated to this individual: $w_j(i) = w_i$.
2. Non-married couple ($h_i^s = 2$): when the household comprises a non-married couple, the household-level wealth level $w(i)$ is allocated in proportion to each individual j's contribution (averaged over the past three survey waves) to the household's labor income: $w_j(i) = \frac{y_j(i)}{y(i)} w(i)$.
3. Married couple ($h_i^s = 3$): when the household consists of a married couple, the household-level wealth level $w(i)$ is divided equally between both individuals: $w_j(i) = \frac{1}{2}\, w(i)$.

Once w_j is defined for j in $1, 2, \ldots, N_j$ $-$ with N_j defined as the number of eligible individuals $-$ I compute both the individual-level actual wealth ranks $\kappa_j(t, i)$ and proxy wealth ranks $\hat{\kappa}_j(t, i)$. The individual-level wealth ranks are calculated as in Equation 7, with the household subscript

i replaced by the individual subscript j:

$$\kappa_j(t,i) = \left\lceil \frac{100 \times \left(1 + \sum_{k=1}^{N_t^j} \mathbf{1}(w_k(t,i) < w_j(t,i))\right)}{N_t^j} \right\rceil \tag{25}$$

$$\hat{\kappa}_j(t,i) = \left\lceil \frac{100 \times \left(1 + \sum_{k=1}^{N_t^j} \mathbf{1}(\hat{w}_k(t,i) < \hat{w}_j(t,i))\right)}{N_t^j} \right\rceil \tag{26}$$

where N_t^j represents the number of eligible individuals at time t, w is the actual wealth level, and $\hat{w}$ is the estimated wealth level using the GB-model from Appendix C.

D.2 Cohorts & lifecycle stages

Definitions To structure the analysis, each individual j is assigned to a time-invariant birth cohort a and time-varying lifecycle stage s. A variable z at time t of individual j belonging to household i, birth cohort a and lifecycle stage s is then defined as $z_j(t,i,s;a)$. Here, i and s vary with t, while a remains time-invariant. birth cohorts Υ are defined over ten-year intervals, beginning with 1866-1975 up until 2006-2015. Lifecycle stages Ξ are based on age brackets and defined as 0-24, 25-29, 30-34, 40-44, 45-49, 50-54, 55-59, 60-64, 65-69, 70-74 and 75+, determined by the individual's age $a_j(t)$.

Within-cohort wealth ranks In the literature, wealth rank outcomes are typically calculated across the entire population. However, since older individuals tend to have accumulated more wealth, they naturally occupy higher positions in the overall wealth distribution. To address this, I define individual-level within-cohort wealth ranks, using the previously introduced birth cohorts. These ranks are calculated for both actual wealth (using observed values) and proxy wealth (using GB-model estimates). The within-cohort ranks are derived by applying the ranking formula to the subset of individuals belonging to a specific birth cohort a:

$$\kappa_j(t,i,s;a) = \left\lceil \frac{100 \times \left(1 + \sum_{k=1}^{N_t^a} \mathbf{1}(w_k(t,i,s;a) < w_j(t,i,s;a))\right)}{N_t^a} \right\rceil \tag{27}$$

$$\hat{\kappa}_j(t,i,s;a) = \left\lceil \frac{100 \times \left(1 + \sum_{k=1}^{N_t^a} \mathbf{1}(\hat{w}_k(t,i,s;a) < \hat{w}_j(t,i,s;a))\right)}{N_t^a} \right\rceil \tag{28}$$

where N_t^a is the number of eligible individuals in cohort a at time t, w denotes the actual wealth level, and $\hat{w}$ represents the wealth level predicted by the GB-model. These within-cohort wealth ranks serve as the primary input of the wealth mobility analyses conducted in this paper.

Summary across stages Finally, I summarize each variable for an individual j over their lifecycle stages. For a given variable z of individual i during lifecycle stage s, the summarized value, denoted as $z_j(s;a)$, is defined as the median of all observations of z for individual j across the years t within the lifecycle stage s:

$$z_j(s;a) = \tilde{z}_j(t,i,s;a) \quad \forall t \in \mathcal{T}_s \tag{29}$$

where $\tilde{z}_j(i,s;a)$ represents the median value of z for individual i, belonging to birth cohort a, during the lifecycle stage s. The set $\mathcal{T}_s$ includes all years t that correspond to the lifecycle stage s for individual j. This approach allows us to drop the time indicator t. The key objects $\kappa_j(s;a)$ and $\hat{\kappa}_j(s;a)$ are then defined as the median actual and proxy wealth ranks of individual j over their lifecycle stage s, with $s \in \Xi$.

This summary over multiple observations per lifecycle stage offers four key advantages. First, any remaining transitory measurement errors − even after the application of outlier correction procedures − are likely to be smoothed out. Second, the formulation reduces the impact of occasional non-response, helping to preserve sample size in wealth mobility analyses. Third, aggregating data by lifecycle stage helps minimizing noise arising from household transitions, such as marriage or divorce. These might otherwise distort wealth mobility estimates. Fourth, it circumvents the non-uniform timing of PSID survey waves, in particular for the reduced sample Ψ.

D.3 Proxy wealth & wealth ranks over Ψ

In earlier appendices and sections, I have defined actual wealth $w_j(s;a) = w_j^{\Psi}(s;a)$ and within-cohort wealth ranks $\kappa_j(s;a) = \kappa_j^{\Psi}(s;a)$. Additionally, I introduced proxy wealth $\hat{w}_j(s;a) = \hat{w}_j^{\Omega}(s;a)$ and proxy within-cohort wealth ranks $\hat{\kappa}_j(s;a) = \hat{\kappa}_j^{\Omega}(s;a)$. These take values over reduced sample $\mathcal{T}_\Psi$ and full sample $\mathcal{T}_\Omega$ respectively.

Let us now define proxy wealth and within-cohort wealth ranks, summarized per lifecycle stage s, restricted to the reduced sample Ψ:

$$\hat{w}_j^{\Psi}(s;a) = \hat{w}_j^{\Omega}(s;a)\Big|_{\mathcal{T}_\Psi}, \quad \hat{\kappa}_j^{\Psi}(s;a) = \hat{\kappa}_j^{\Omega}(s;a)\Big|_{\mathcal{T}_\Psi} \tag{30}$$

where $|_{\mathcal{T}_\Psi}$ indicates that the values are restricted to the time frame $\mathcal{T}_\Psi$.

Equation 30 covers the same time-frame and individuals as actual wealth $w_j^{\Psi}(s;a)$ and wealth ranks $\kappa_j^{\Psi}(s;a)$. Therefore, a comparison of the outcomes of $\hat{w}_j^{\Psi}(s;a)$ to those of $w_j^{\Psi}(s;a)$ provides insight into the validity of the GB-model predictions and the accurateness of $\hat{w}_j^{\Omega}(s;a)$

and $\hat{\kappa}_j^{\Omega}(s;a)$. As argued in Section 3.3, throughout the mobility analyses, the outcomes based on $\hat{w}_j^{\Psi}(s;a)$ align more closely to those based on $\hat{w}_j^{\Omega}(s;a)$ than those based on $w_j^{\Psi}(s;a)$. This indicates that differences in the results between $w_j^{\Psi}(s;a)$ and $\hat{w}_j^{\Omega}(s;a)$ relate to the usage of the proxy ($\hat{\kappa}$ versus κ) rather than sample differences (Ω versus Ψ).

For future reference, I define the relevant sets of actual and proxy within-cohort wealth and wealth ranks as:

$$W = \{w_j^{\Psi}(s;a),\ \hat{w}_j^{\Omega}(s;a),\ \hat{w}_j^{\Psi}(s;a)\} \tag{31}$$

$$K = \{\kappa_j^{\Psi}(s;a),\ \hat{\kappa}_j^{\Omega}(s;a),\ \hat{\kappa}_j^{\Psi}(s;a)\} \tag{32}$$

D.4 Inter-generational linkages

The PSID enables the construction of family trees, allowing individuals to be linked to their parents and grandparents. I focus on biological and adoptive parents, excluding step-parenting. An individual can thus have at most two parents and four grandparents. Parent indices are denoted as $p_1(j)$ and $p_2(j)$, so that $p(j) = \{p_1(j), p_2(j)\}$. The set of grandparent indices is then defined as:

$$g(p(j)) = \{g_1(p_1(j)), g_2(p_1(j)), g_1(p_2(j)), g_2(p_2(j))\} \tag{33}$$

while a variable z associated with the k-th parent of individual j is expressed as $z_{p_k(j)}$. Similarly, a variable z of the first grandparent of the k-th parent of individual j is denoted as $z_{g_1(p_k(j))}$.

D.5 Intra-generational lifecycle phases

For the intra-generational analyses, individuals' wealth rank trajectories are investigated over two lifecycle phases: working life (ages 30-54) and older age (55-74). For completeness, I define the lifecycle stages relevant to working life and older age as Ξ^{WL} and Ξ^{OA} respectively. The distinction in two lifecycle phases offers two main advantages. First, not a single individual has data points spanning the entire lifecycle. By separating the analysis into two phases, it becomes possible to examine intra-generational mobility across the entire lifecycle, albeit using data from different birth cohorts. Second, this approach aligns with both theoretical and empirical literature, which frequently differentiates between models of wealth dynamics during working life and during older age.

E Inequality & mobility metrics

In this section, I define the outcome measures used in the inter- and intra-generational wealth mobility analyses. These include (i) metrics related to wealth inequality and accumulation over the lifecycle, (ii) rank-rank coefficients, (iii) a squared mobility metric, (iv) transition probabilities, (v) discretionary groups and (vi) hierarchical clustering. For intra-generational analyses, all six measures are calculated and reported. The inter-generational analyses are restricted to measures (ii), (iii), (iv) and (v).

Outcome metrics (ii) to (vi) compare two cross-sections of wealth ranks. Specifically, for the inter-generational analyses, different individuals (parent-child or grandparent-grandchild) of the same family are compared at the same lifecycle stage (if available) or different lifecycle stages (otherwise). In the intra-generational analyses, the same individuals are evaluated at an initial and a final lifecycle stage.

E.1 Wealth dynamics over the lifecycle

For each wealth bin b, I calculate their wealth shares and wealth-to-average labor income ratios across the lifecycle stages $s \in \Xi^{\mathrm{WL}}$ or $s \in \Xi^{\mathrm{OA}}$:

$$\lambda_b(s;a) = \frac{\sum_{j \in b} w}{\sum_j w}, \quad \theta_b(s;a) = \frac{\sum_{j \in b} w}{|b| \cdot \bar{y}(t)} \tag{34}$$

where $|b|$ denotes the number of individuals in wealth bin b, $w \in W$ represents wealth, and $\bar{y}(t)$ is the average labor income across all individuals at time t. Depending on the lifecycle phase under consideration, $a \in \mathrm{Y}^{\mathrm{WL}}$ or $a \in \mathrm{Y}^{\mathrm{OA}}$, and $s \in \Xi^{\mathrm{WL}}$ or $s \in \Xi^{\mathrm{OA}}$. In addition to these measures, I compute the proportion of low-wealth and high-wealth individuals for each lifecycle stage s. These groups are defined as individuals with wealth levels below $\bar{y}(s;a)$ and in excess of twenty times $\bar{y}(s;a)$ respectively:

$$\vartheta^l(s;a) = \frac{1}{|a|} \sum_j^{j \in a} w < \bar{y}(s;a), \quad \vartheta^h(s;a) = \frac{1}{|a|} \sum_j^{j \in a} w > 20 \cdot \bar{y}(s;a) \tag{35}$$

where $|a|$ denotes the number of individuals in birth cohort a, and $w \in W$.

E.2 Overall mobility

Rank-rank coefficients I calculate a rank-rank coefficient β, obtained by regressing wealth ranks in a final stage ($s = f$) on wealth ranks in the initial stage ($s = i$) using Ordinary Least Squares (OLS). It is defined as:

$$\kappa_k(s = f) = \alpha + \beta \kappa_k(s = i) + \epsilon_k, \tag{36}$$

where α denotes the intercept, β the regression coefficient capturing the degree of wealth persistence, and ϵ_k the error term for an individual, parent-child pair or grandparent-grandchild pair k.

Squared mobility To attach higher weight to large wealth rank fluctuations, I define a squared mobility measure η as:

$$\eta(a) = \frac{\sum_k \left[\kappa_k(\mathrm{s = f}) - \kappa_k(\mathrm{s = i})\right]^2}{|a|} \tag{37}$$

where $|a|$ denotes the number of individuals in birth cohort a, and i and f denote the initial and final lifecycle stages under consideration. Across all analyses, the squared mobility metric yields identical findings to the rank-rank coefficient β. I therefore do not report this squared mobility metric in the main text.

E.3 Mobility at the bottom and top

Transition matrices Transition matrices summarize the probability of individuals moving from a wealth bin b_i to a wealth bin b_f from an initial stage $s = i$ to a final stage $s = f$.

After categorizing individuals into wealth bins b based on $\kappa \in K$, the transition probability from b_i to b_f is calculated for a given cohort a as:

$$P(b_\mathrm{i} \to b_\mathrm{f})(a) = \frac{n_a(b_\mathrm{i}, b_\mathrm{f})}{\sum_{b_\mathrm{f}} n_a(b_\mathrm{i}, b_\mathrm{f})} \tag{38}$$

where $n_a(b_\mathrm{i}, b_\mathrm{f})$ represents the number of individuals in cohort a transitioning from bin b_i to bin b_f. The total number of individuals in the initial bin b_i is given by $\sum_{b_\mathrm{f}} n_a(b_\mathrm{i}, b_\mathrm{f})$. The ex-ante and ex-post transition matrices $T_{\mathrm{EA}}(a)$ and $T_{\mathrm{EP}}(a)$ for cohort a are then defined as:

$$T_{\mathrm{EA}}(a) = \left[P(b_\mathrm{i} \to b_\mathrm{f})(a)\right]_{b_\mathrm{i}, b_\mathrm{f}}, \quad T_{\mathrm{EP}}(a) = \left[P(b_\mathrm{i} \to b_\mathrm{f})(a)\right]_{b_\mathrm{f}, b_\mathrm{i}} \tag{39}$$

where each element of $T_{\mathrm{EA}}(a)$ and $T_{\mathrm{EP}}(a)$ represents the probability of transitioning between two wealth bins b for birth cohort a. While the underlying calculations are identical, the interpretation of the columns differs between the two matrices. In the ex-ante matrix $T_{\mathrm{EA}}(a)$, a column represents the probability of moving to wealth bins b given the initial wealth bin b_i. In the ex-post matrix, a column represents the probability of originating from wealth bins b given the final wealth bin b_f.

Discretionary groups Using Equation 38, I calculate the relative occurrence of six discretionary groups that focus on wealth mobility at the bottom 20% and top 10% of the wealth distribution. The groups include the steady poor (SP), past poor (PP), new poor (NP), steady wealthy (SW), past wealthy (PW) and new wealthy (NW). At the bottom, (i) the steady poor include those families or individuals that start and end in the bottom 20%, (ii) the past poor the families or individuals that display upward wealth mobility to the top 50% originating from the bottom 20%, and (iii) the new poor start off in the top 50% but experience downward mobility to the bottom 20%. At the top, (iv) the steady wealthy start and end in the top 10%, (v) the past wealthy begin in the top 10% but display downward mobility to the bottom 70%, and (vi) the new wealthy experience upward mobility to the top 10% after starting off in the bottom 70%.

Hierarchical clusters The transition matrices and discretionary groups have the advantage of being intuitive and easily interpretable. They also facilitate cross-cohort comparisons. However, these methods may be considered somewhat ad hoc, as they require defining thresholds for wealth bins and discretionary groups prior to the analysis. Additionally, only a fraction of the sample is allocated to one of the six discretionary groups.

To complement these approaches, I employ hierarchical clustering, a method from the machine learning literature. This technique groups individuals' wealth rank trajectories into clusters, providing an alternative perspective to the discretionary groups. The four-step procedure used for this clustering is adapted from Audoly et al. (2024) and is detailed in the Online Supplement. As it requires wealth rank trajectories as input, it is used only for the intra-generational analyses.

The clustering process ultimately results in a set of k clusters, where each cluster c contains the wealth rank trajectories of the individuals assigned to it. All individuals in the sample are allocated to a specific cluster. Each cluster c is summarized by its average wealth rank trajectories $\bar{\kappa}_c(s)$, with $s \in \Xi^{\mathrm{WL}}$ or $\in \Xi^{\mathrm{OA}}$. Denoting $|C_{\mathrm{c}}|$ as the number of individuals in a specific cluster, we have:

$$\bar{\kappa}_{\mathrm{c}}(s) = \frac{1}{|C_{\mathrm{c}}|} \sum_{i \in C_{\mathrm{c}}} \kappa_i(s) \tag{40}$$

F Composition metrics

In Appendix E, I have defined different outcome measures to assess the degree of inter -and intra-generational wealth mobility. As part of this, I defined six discretionary groups (steady poor, past poor, new poor, steady wealthy, past wealthy and new wealthy), as well as a set of hierarchical clusters. In this Appendix, I define a set of variables that can be used to compare the composition of the individuals within a sample or across different discretionary groups or hierarchical clusters.

F.1 Labor income, saving rates & non-mortgage indebtedness

To assess heterogeneity in labor incomes, I define the within-cohort labor income rank $\delta_j(t,s,i;a)$, which is computed by applying the ceiling function to labor income $y_j(t,s,i;a)$ for an birth cohort a:

$$\delta_j(t,i,s;a) = \left\lceil \frac{100 \times \left(1 + \sum_{k=1}^{N_t^a} \mathbf{1}(y_k(t,i,s;a) < y_j(t,i,s;a))\right)}{N_t^a} \right\rceil \tag{41}$$

where N_t^a is the number of individuals in birth cohort a. $\delta_j(t,s,i;a)$ is then summarized into lifecycle stages as $\delta_j(s;a)$ according to the procedure described in Appendix D. In addition, I calculate the non-mortgage debt-to-income ratio $v_j(t,s,i;a)$ – summarized over s as $v_j(s;a)$ – as the ratio of non-mortgage debt to total household income. It is equated to the level of the household i individual j is linked with. These variables are aggregated over a sample, group or cluster g by taking the median observation across the relevant individuals.

F.2 Asset ownership & allocation

With respect to asset ownership and allocation, I formalize two types of measures. First, I define a homeownership dummy variable $d_j^h(s;a)$ which equals one whenever individual j belonged more often than not to a household i owning at least one house during lifecycle stage s. Additionally, two dummy variables $d_j^{bu}(s;a)$ and $d_j^{bi}(s;a)$ equal one whenever individual j was linked to a household i that respectively owned an unincorporated or incorporated business more often than not throughout lifecycle stage s. These variables are aggregated across the sample, group or cluster by calculating the fraction of individuals with dummies equal to one. Second, I define the conditional equity, housing and mortgage portfolio shares at the individual level as $\alpha_j^e(s;a)$, $\alpha_j^h(s;a)$ and $\alpha_j^d(s;a)$. These are equated to their household-level counterparts. They are aggregated across the sample, group or cluster by computing the median.

F.3 Inter-vivos transfers & inheritances

The PSID contains two variables that could possibly capture inter-vivos transfers and inheritances. For the first variable $-$ available over sample Ω $-$ households are asked how much they have received in lumpsum payments (comprising inheritances and payouts from insurance) since the previous survey wave. Prior to 1982, this lumpsum-question provides a bracketing response only. Summary statistics for this lumpsum variable provide highly non-robust outcomes, however. For example, the lumpsum variable suggests that the cumulative proportion of individuals having received a payment remains more or less constant over working life, which strongly contradicts empirical evidence (e.g. Black et al., 2022). For that reason, I do not proceed with this variable. For the second variable $-$ defined over sample Ψ $-$ households are asked how much they have received in gifts or inheritances since the previous survey wave. For $\mathcal{T}_{\Psi}[1] = 1984$, the gifts or inheritances inferred about are those that have been received overall prior to 1984. For the 1984-1989 survey waves, the respondent can provide two separate inheritances or gifts, while this number was raised to three from 1994 onwards. I apply bracketing to the responses if necessary and available, link the household-level responses to individual-level ones based on the procedure described in Appendix D, and define the received gifts or inheritances at t as $l_j(t, s, i; a)$.

For an individual j, I then compute at each t the cumulative value of the inter-vivos transfers and inheritances it has received up until that that point in time. This allows to define two composition metrics. First, I calculate a dummy variable $\iota_j^d(t, s, i; a)$ which indicates whether the individual j has received any transfer in its lifetime up until t. It is summarized per lifecycle stage s to obtain $\iota_j^d(s; a)$, and aggregated by calculating the fraction of individuals in the sample, group or cluster that has received a transfer. Second, I define the individual's cumulative transfer receipts to its lifetime resources (Black et al., 2022), defined as $\iota_j(t, s, i; a)$. Lifetime resources are computed as the cumulative sum of capitalized labor income. $\iota_j(t, s, i; a)$ is then summarized over lifecycle stage s to obtain $\iota_j^s(s; a)$. Finally, $\iota_j^s(s; a)$ is aggregated across the individuals in the group or cluster by taking the mean.

There are three remaining issues with $\iota_j^s(t, s, i; a)$. First, given that it depends on the cumulative sum, the accurateness of $\iota_j^i(a)$ for a given individual j is strongly affected by the timing of non-response or the timing of an individual's entry into the dataset. For example, the 1984-question infers about inheritances and gifts ever received prior to 1984. If a household displays non-response specifically for 1984, or enters the dataset only after 1984, $\iota_j^i(a)$ may strongly underestimate actual transfers received. However, as long as the non-response is random across the different discretionary groups or clusters, it should not affect the observed relative differences between these groups or clusters. Second, as noted, gifts and inheritances are allocated from the household to the individual level according to the rules described in Appendix D. For gifts and inheritances, which are mostly intertwined with the family of a specific individual in the household, these allocation rules may be suboptimal. Nevertheless, given the data, there is no straightforward option to execute the linkages more appropriately. Third, the gift and inheritance questions are self-reported, and may thus suffer from a downward bias. This holds specifically at the top of the inter-generational transfer distribution.

F.4 Health & household composition

Regarding health and household composition, I delineate two measures[16]. First, to assess an individual's health level, I define a dummy health variable $d_j^h(t, s, i; a)$, which is summarized per lifecycle stage s as $d_j^h(s; a)$. The dummy variable uses a question in the household PSID-dataset which categorizes the household's reference person's and partner's health. Whenever an individual j belongs to a household where at least one of the two core members is stated to have poor health, variable $d_j^g(t, s, i; a)$ is set to one. It is aggregated as the fraction of individuals in a sample, group or cluster that are part of a household with a poor health member. Second, the individual's household status variable, $s_j^h(t, s, i; a)$ and $s_j^h(s; a)$, is equated to the status variable of the household it belongs to (see Appendix C). It is aggregated across a sample, group or cluster by computing the fraction of individuals that is co-habiting with a partner or married (i.e. is non-single).

[16]In addition, I have considered the number of children in the household and the integration of the household. The results indicate that these two variables are not clearly associated with wealth rank combinations or trajectories. I therefore do not report their outcomes in this paper.

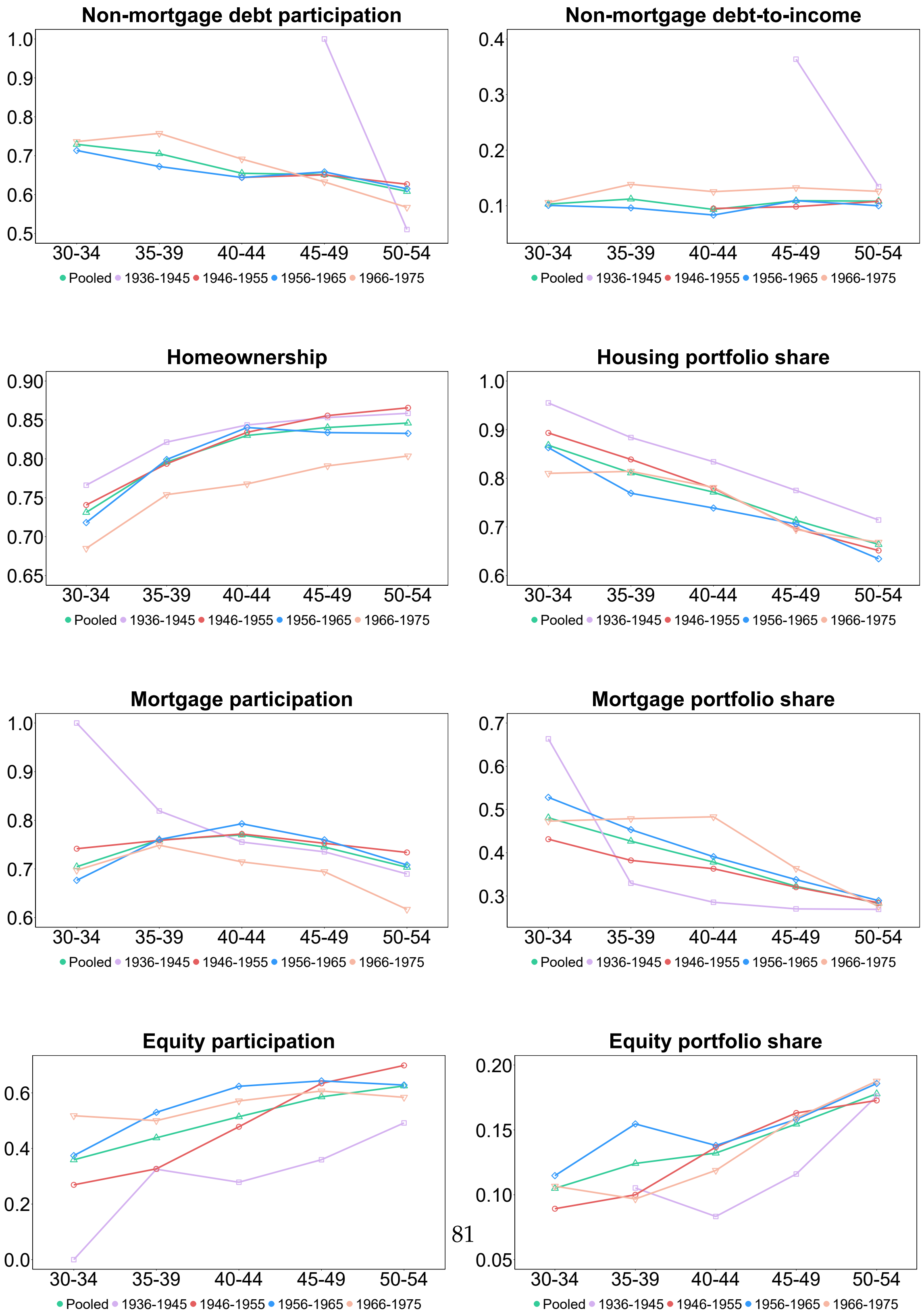
G Composition analysis: intra-generational wealth mobility during working life

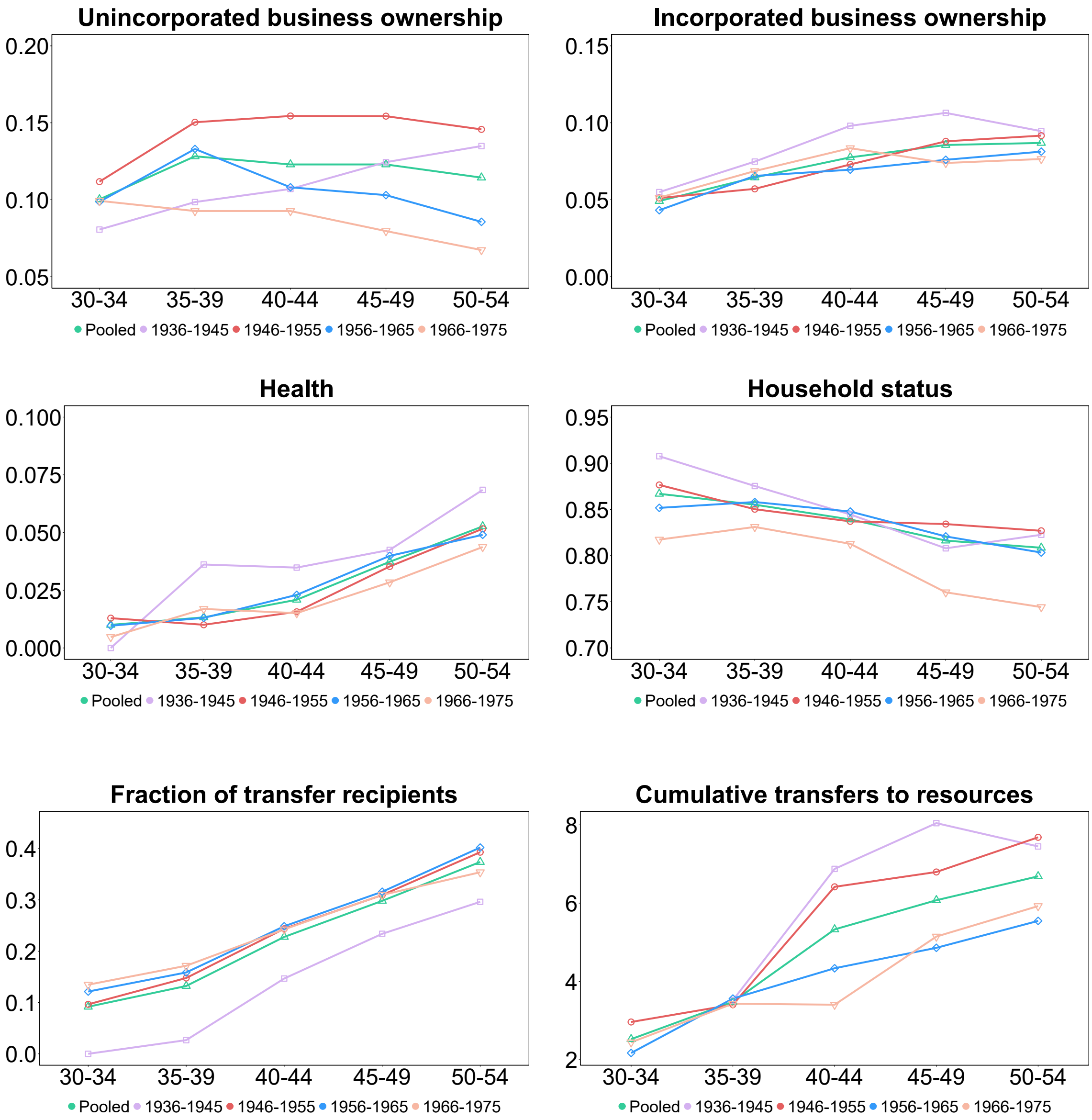
In this Appendix, I present the results of the composition analysis for working life. Specifically, Section G.1 provides the composition outcomes for the entire working life sample per lifecycle stage. Instead, Sections G.2 and G.3 compute the outcome metrics for the individuals in each of the discretionary groups and clusters. The composition metrics reported in these sections have been defined in Appendix F. Moreover, in the same Appendix, I have discussed how the individual-level metrics are aggregated across the sample or across the individuals in a specific group or cluster.

G.1 Composition across the entire sample

Figure 18 presents the composition metrics per birth cohort for all individuals in the working life sample. Four findings persist. First, non-mortgage debt participation and non-mortgage debt-to-income ratios are relatively stable over working life. Overall, non-mortgage indebtedness has increased over time: more recent cohorts have higher participation rates and higher non-mortgage debt-to-income ratios. Second, homeownership rises over the working lifecycle, while the conditional share of housing in individuals' portfolios follows a downward trajectory. Homeownership is lower in the most recent (1966-75) cohort, while the conditional housing share was higher in the oldest (1936-45) cohort. Mortgage participation displays an inverse U-shaped pattern (peaking at ages 40-44), while conditional mortgage-to-total assets ratios decline over the working lifecycle. Moreover, equity market participation and the conditional equity portfolio share rise with age, and were significantly lower for the 1936-45 cohort compared to more recent cohorts. Instead, business ownership rates are roughly stable over the working lifecycle and across cohorts. Third, individuals are more likely to belong to a poor health household and more likely to be part of a single household as the working lifecycle progresses. The fraction of single individuals is higher in the most recent (1966-75) cohort. Fourth, the fraction of inter-generational transfer recipients and the size of their cumulative receipts increases strongly over working life. Figure 18 suggests that the fraction of inter-generational transfer recipients lays significantly lower for the 1936-45 cohort. However, this is likely related to a measurement error: the 1936-45 birth cohort is likely to have received significant transfers prior to 1984, which may not be accurately captured in the PSID-data (see Appendix F for a detailed explanation).

Figure 18: Socio-economic characteristics and inter-generational transfers of individuals in the working life sample.

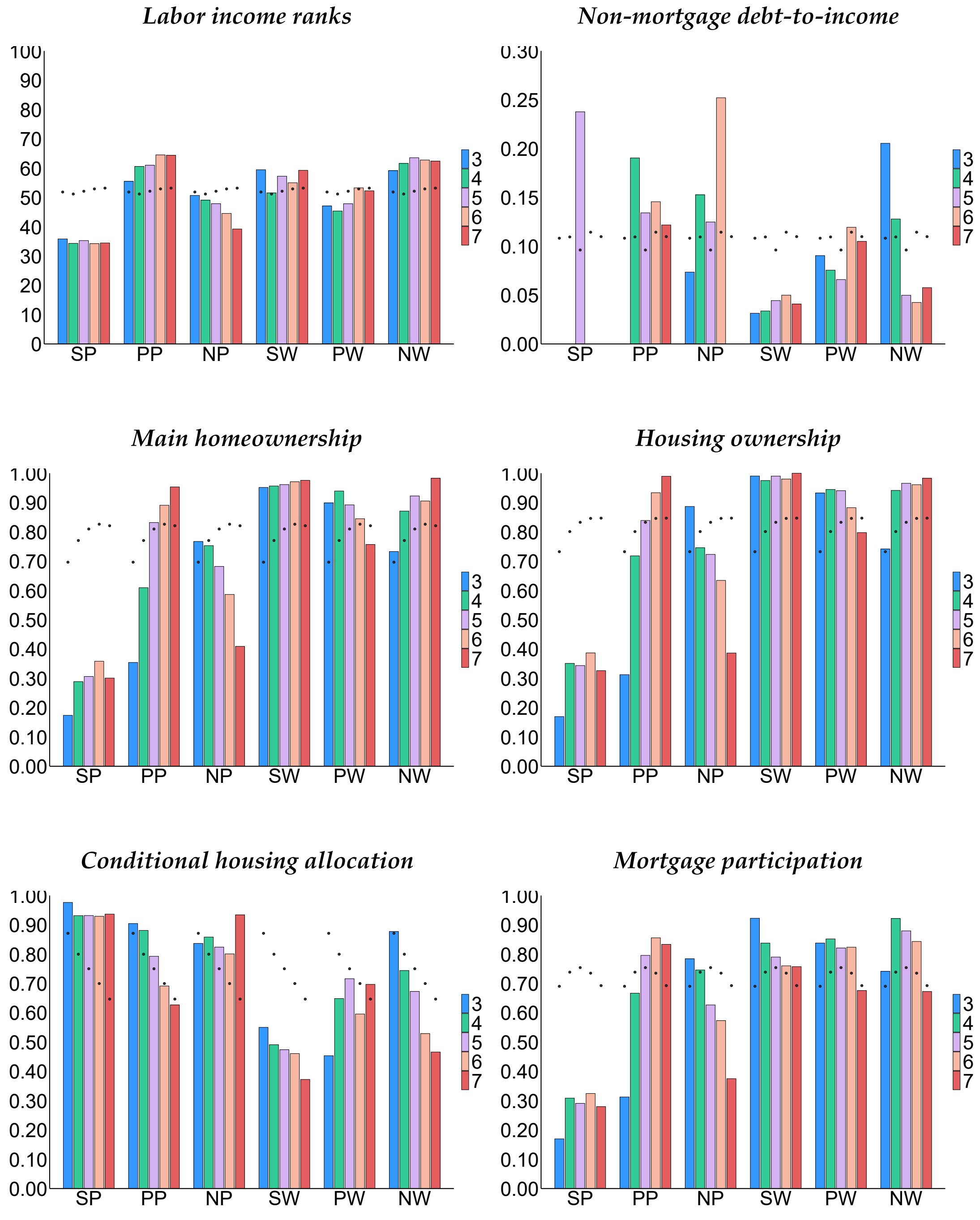


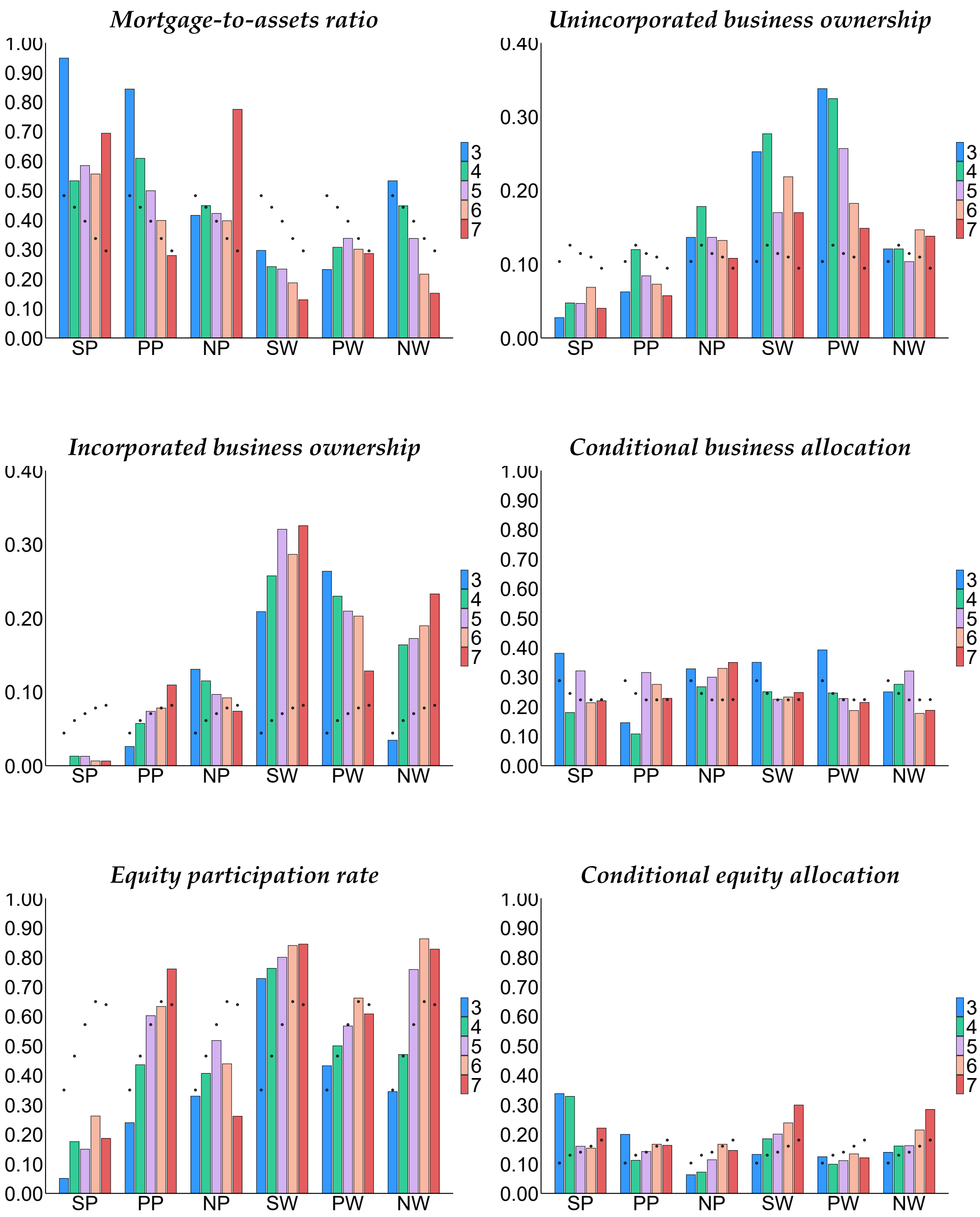


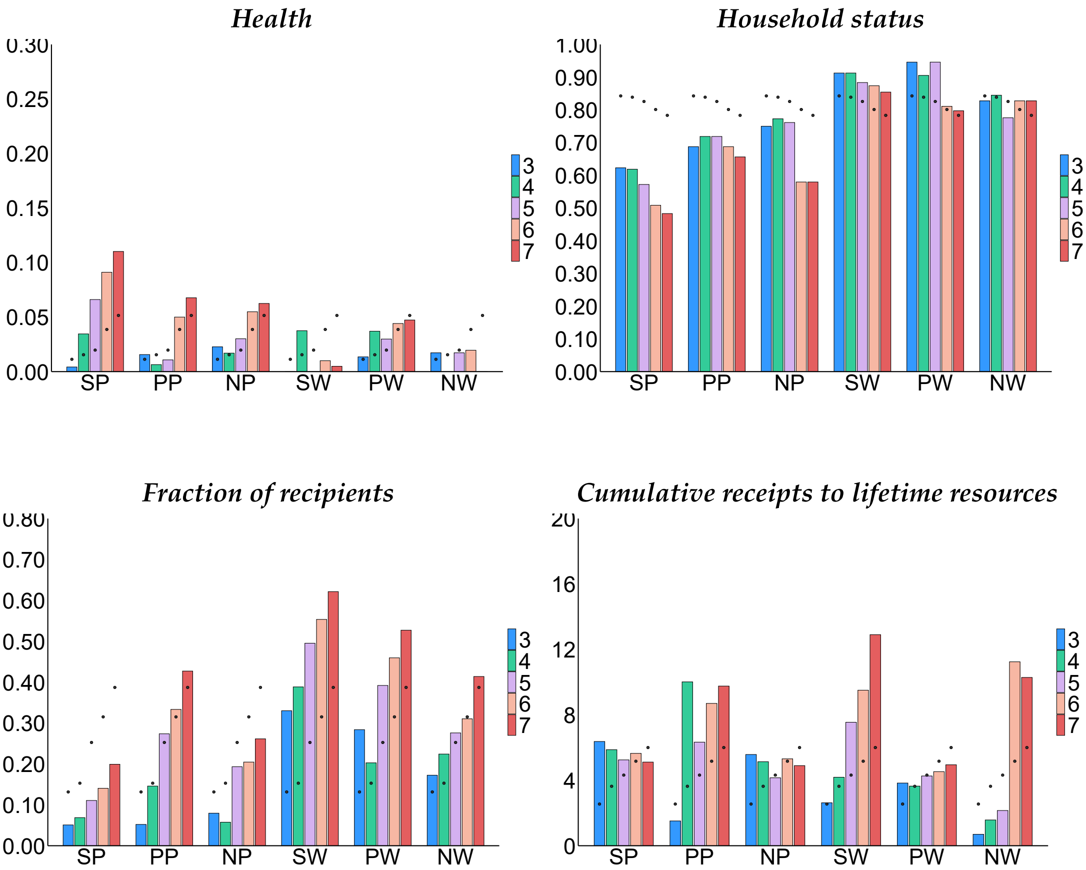
Note: this figure summarizes the key composition metrics across all individuals in the working life sample per lifecycle stage. The composition metrics and their aggregation method are defined in Appendix F.

G.2 Composition discretionary groups

Figure 19: Composition metrics for the individuals per discretionary group across the working lifecycle stages.



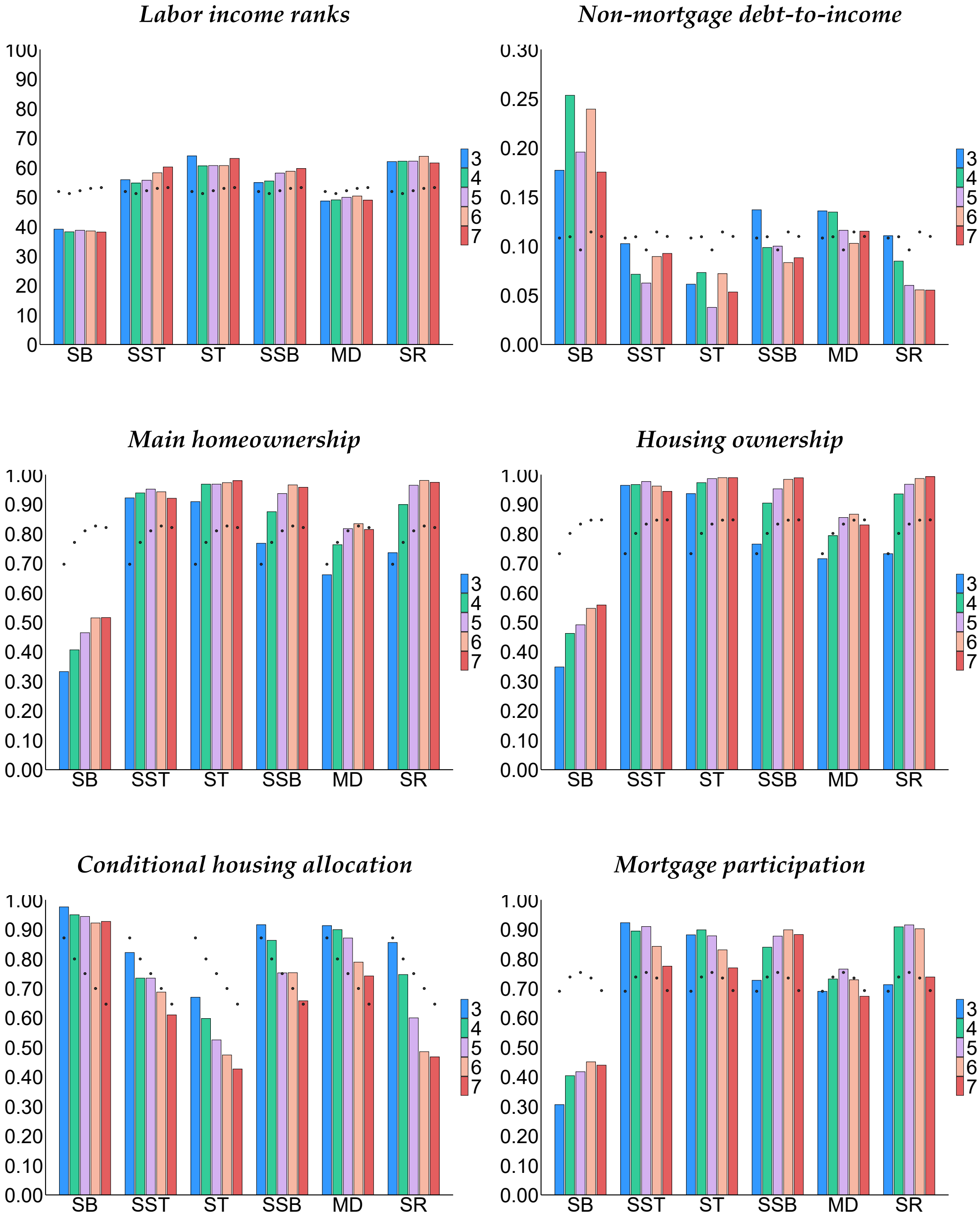

Mortgage-to-assets ratio
Unincorporated business ownership
Incorporated business ownership
Conditional business allocation
Equity participation rate
Conditional equity allocation
SP
PP
NP
SW
PW
NW
3
4
5
6
7
0.00
0.10
0.20
0.30
0.40
0.50
0.60
0.70
0.80
0.90
1.00

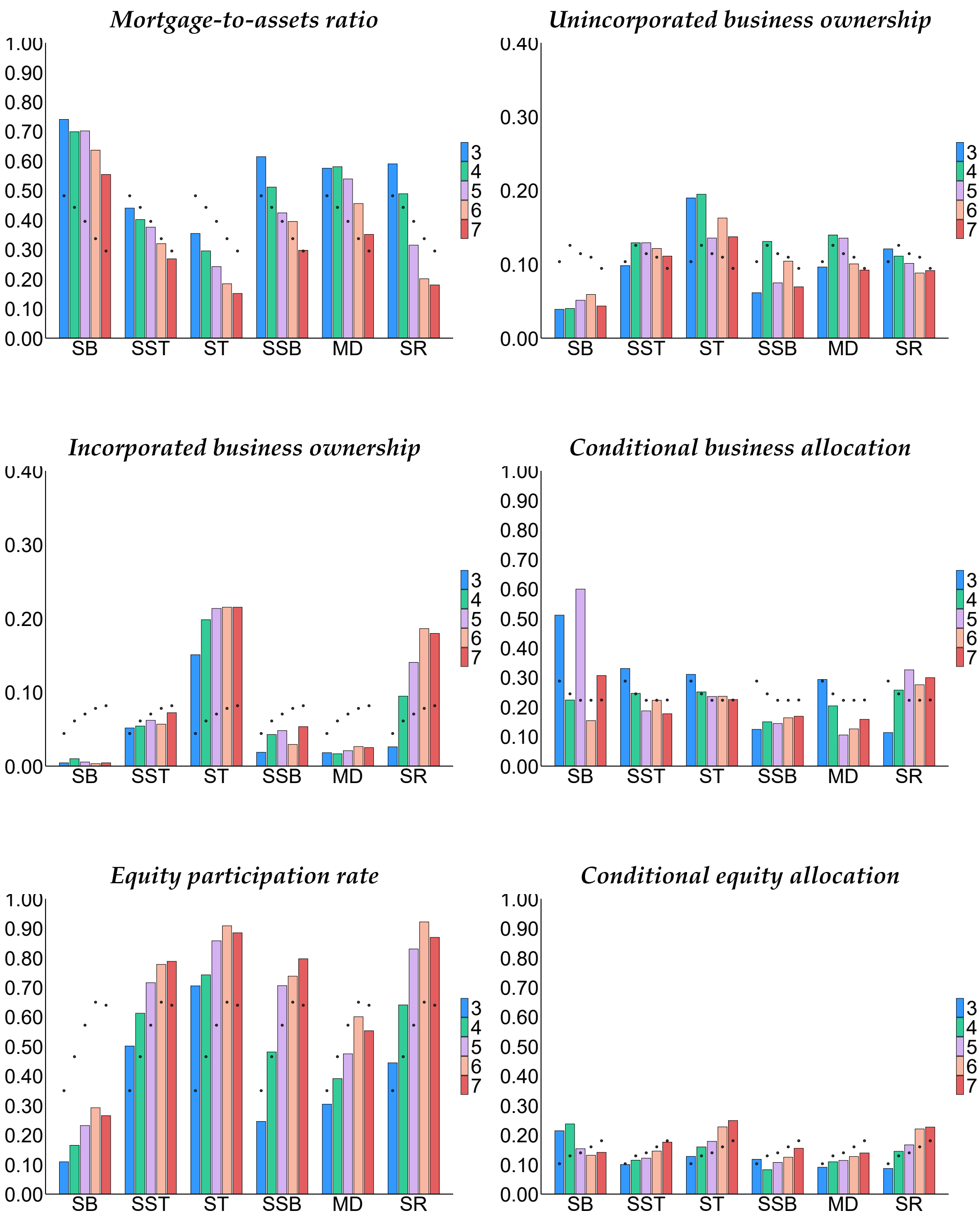


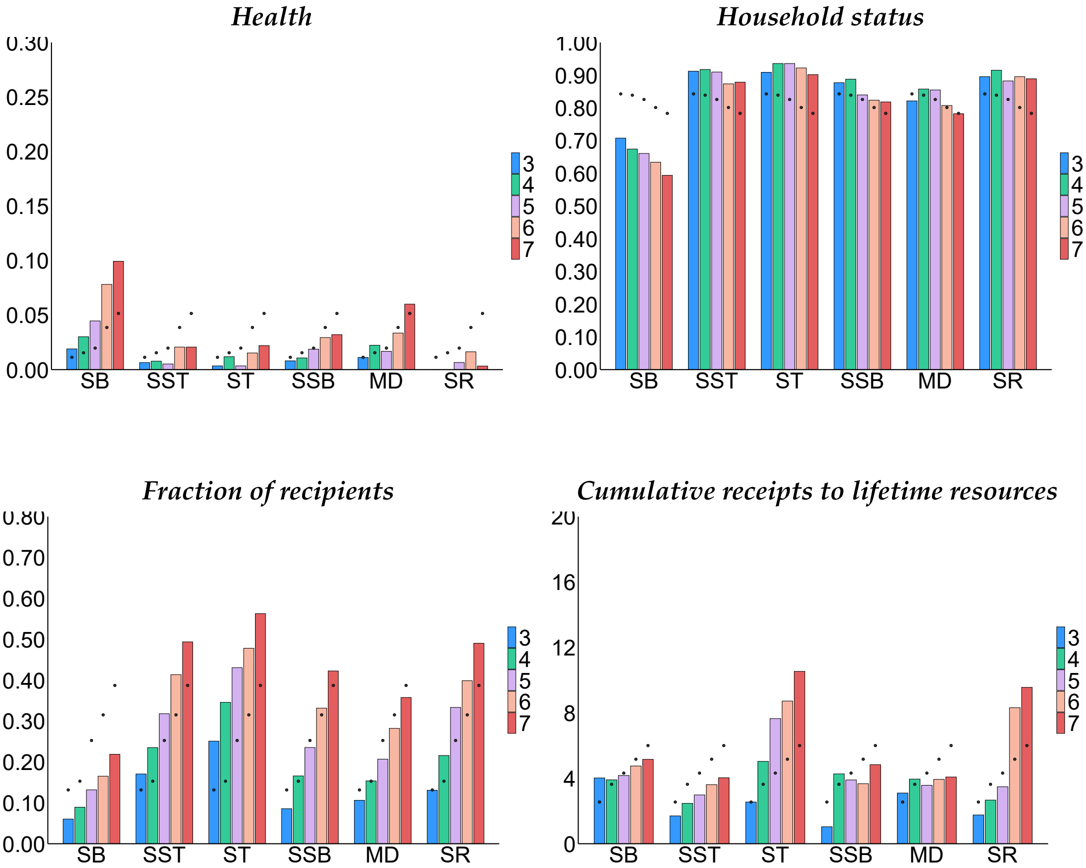
Note: this figure summarizes the key composition metrics across all individuals in each of the discretionary groups per lifecycle stage. The composition metrics and their aggregation method are defined in Appendix F.

G.3 Composition hierarchical clusters

Figure 20: Composition metrics for the individuals per hierarchical cluster across the working lifecycle stages.




Mortgage-to-assets ratio
Unincorporated business ownership
Incorporated business ownership
Conditional business allocation
Equity participation rate
Conditional equity allocation
SB
SST
ST
SSB
MD
SR
3
4
5
6
7



Note: this figure summarizes the key composition metrics across all individuals in each of the discretionary clusters per lifecycle stage. The composition metrics and their aggregation method are defined in Appendix F.

H Additional visualizations

In this Appendix, I report additional visualizations related to inter- and intra-generational wealth mobility outcomes. The structure of the appendix follows the chronology of the main text: I first provide additional visualizations for three- and two-generational wealth mobility, and then move to intra-generational wealth mobility outcomes.

Table 4: Share of families (in %) consolidating in the bottom 20% over two generations, computed across children's birth cohorts $\in \Upsilon^{PC}$ for parents and children at identical lifecycle stages.

Variable	Stage	1946–55	1956–65	1966–75	1976–85	1986–95	Pooled
$\hat{\kappa}^{\Omega}$	30–34	-	7.2	7.6	6.8	6.6	7.0
			(5.8, 8.9)	(6.4, 8.8)	(6.0, 7.6)	(5.4, 7.9)	(6.6, 7.5)
	35–39	-	6.1	7.2	7.8	-	7.3
			(5.2, 7.3)	(6.0, 8.2)	(7.1, 8.8)		(6.8, 7.9)
	40–44	10.9	7.2	7.7	8.6	-	8.2
		(9.5, 12.3)	(6.0, 8.5)	(6.8, 8.5)	(7.2, 9.8)		(7.7, 8.9)
	45–49	9.0	8.3	6.7	-	-	8.0
		(7.6, 10.5)	(7.2, 9.3)	(5.7, 7.7)			(7.2, 8.6)
	50–54	7.4	7.9	-	-	-	7.8
		(6.0, 8.7)	(6.8, 9.0)				(7.3, 8.5)
	55–59	9.2	8.9	-	-	-	9.0
		(8.0, 10.8)	(7.5, 10.3)				(8.1, 9.8)

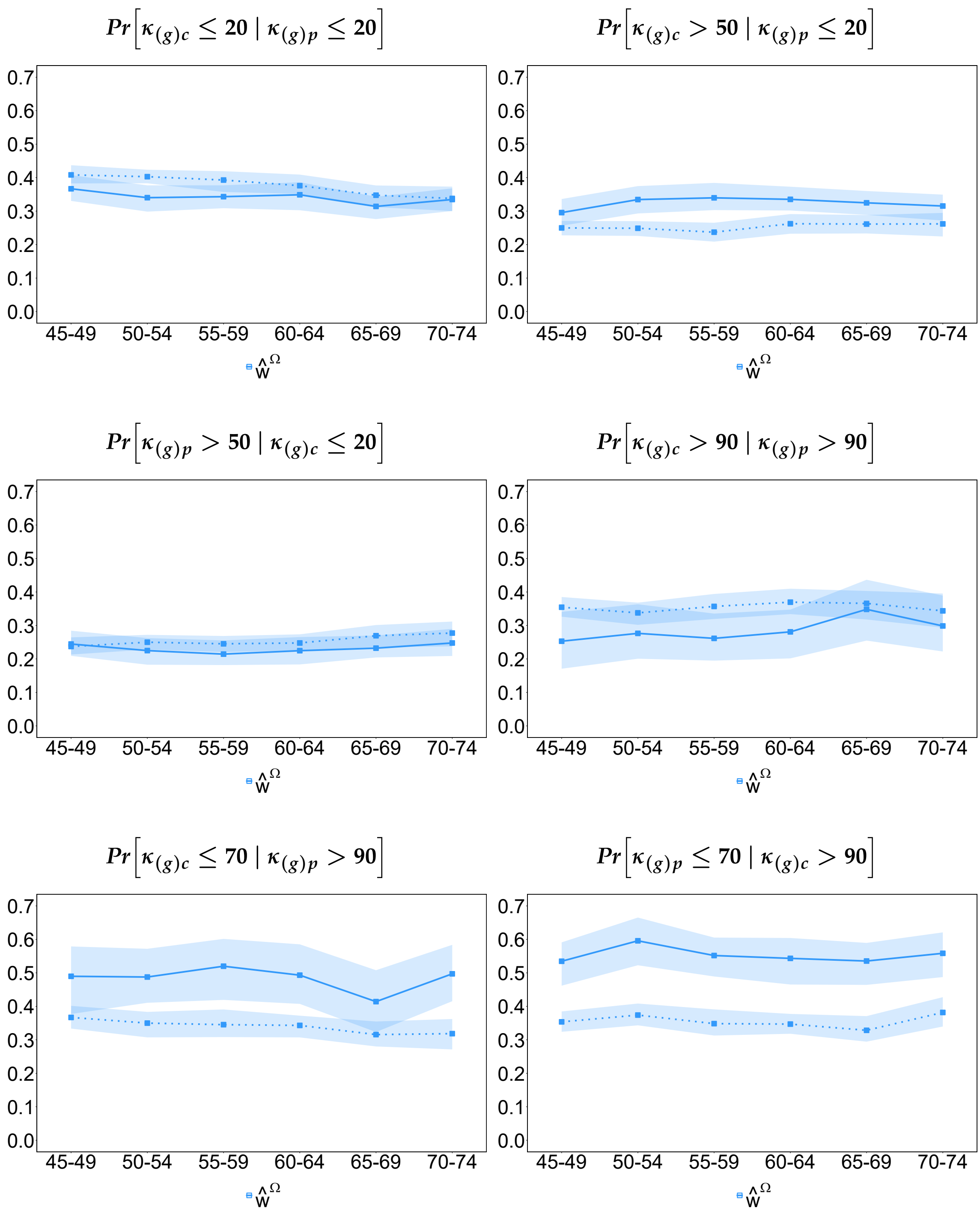
Note: the share of families belonging to each discretionary group are calculated based on parents' and children's within-cohort wealth ranks at identical lifecycle stages (ranging from 30–34 to 55–59) using the proxy wealth measure $\hat{\kappa}^{\Omega}$. The 95% confidence intervals are reported in parentheses. Estimates are included only for cohort-stage combinations with at least 750 observations.

Table 5: Share of families (in %) consolidating in the top 10% over two generations, computed across children's birth cohorts $\in \Upsilon^{PC}$ for parents and children at identical lifecycle stages.

Variable	Stage	1946–55	1956–65	1966–75	1976–85	1986–95	Pooled
	30–34	-	2.2	3.1	2.5	3.3	2.7
			(1.3, 3.3)	(2.4, 3.8)	(2.0, 3.0)	(2.3, 4.5)	(2.4, 3.1)
$\hat{\kappa}^{\Omega}$	35–39	-	3.1	3.6	4.0	-	3.6
			(2.4, 3.9)	(2.9, 4.4)	(3.5, 4.5)		(3.2, 4.0)
	40–44	2.1	2.9	3.8	3.0	-	3.1
		(1.2, 3.0)	(2.2, 3.5)	(3.1, 4.5)	(2.0, 3.9)		(2.6, 3.5)
	45–49	2.2	3.3	3.8	-	-	3.2
		(1.5, 3.0)	(2.5, 3.8)	(2.9, 4.7)			(2.8, 3.7)
	50–54	1.9	2.6	-	-	-	2.7
		(1.2, 2.7)	(2.0, 3.2)				(2.2, 3.1)
	55–59	3.1	2.4	-	-	-	2.7
		(2.0, 4.2)	(1.7, 3.0)				(2.1, 3.3)

Note: the share of families belonging to each discretionary group are calculated based on parents' and children's within-cohort wealth ranks at identical lifecycle stages (ranging from 30–34 to 55–59) using the proxy wealth measure $\hat{\kappa}^{\Omega}$. The 95% confidence intervals are reported in parentheses. Estimates are included only for cohort-stage combinations with at least 750 observations.

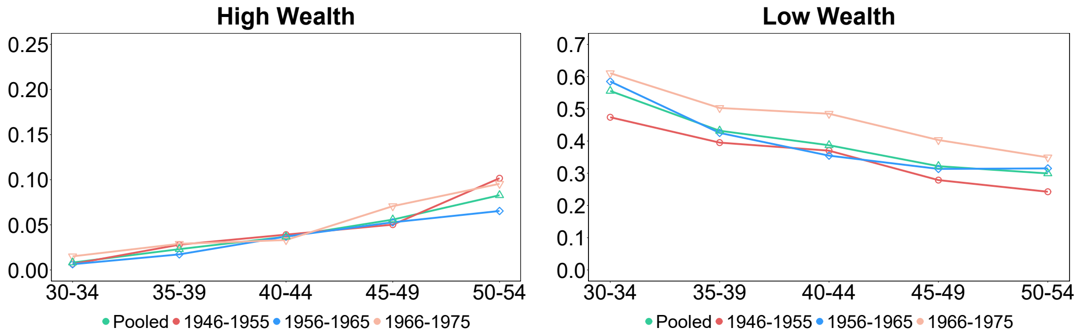
Figure 21: Transition probabilities for grandparents and grandchildren (solid lines) and parents and children (dotted lines) when (grand)children are aged 35-39.



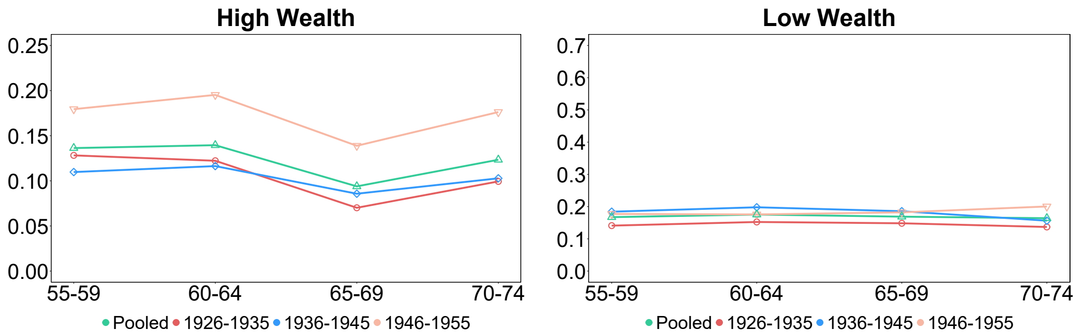
Note: these plots produce transition probabilities across specific wealth bins. These are defined in line with the discretionary groups (see Section 3.4 and Appendix E). In the notation above, $\kappa_{(g)p}$ denotes the within-cohort wealth ranks of (grand)parents, and $\kappa_{(g)c}$ the within-cohort wealth ranks of (grand)children. The transition probabilities are computed at different lifecycle stage combinations: child wealth ranks at ages 35-39 are compared to (grand)parental wealth ranks at stages between 45-49 and 70-74 (plotted on the x-axis). As an example, the values produced for the right-hand plot on the top row denote the probability of children belonging to the top 50% at stage 35-39 given that their parents belonged to the bottom 20% at any of the x-axis stages. The pooled dataset is used. The shaded areas display the 95% confidence intervals, which have been determined through bootstrapping.

Figure 22: Proportion of high-and low-wealth individuals for birth cohorts $\in \Upsilon^{WL}$ and $\in \Upsilon^{OA}$ based on actual wealth levels w^{Ψ}.

Working life: ages 30-54

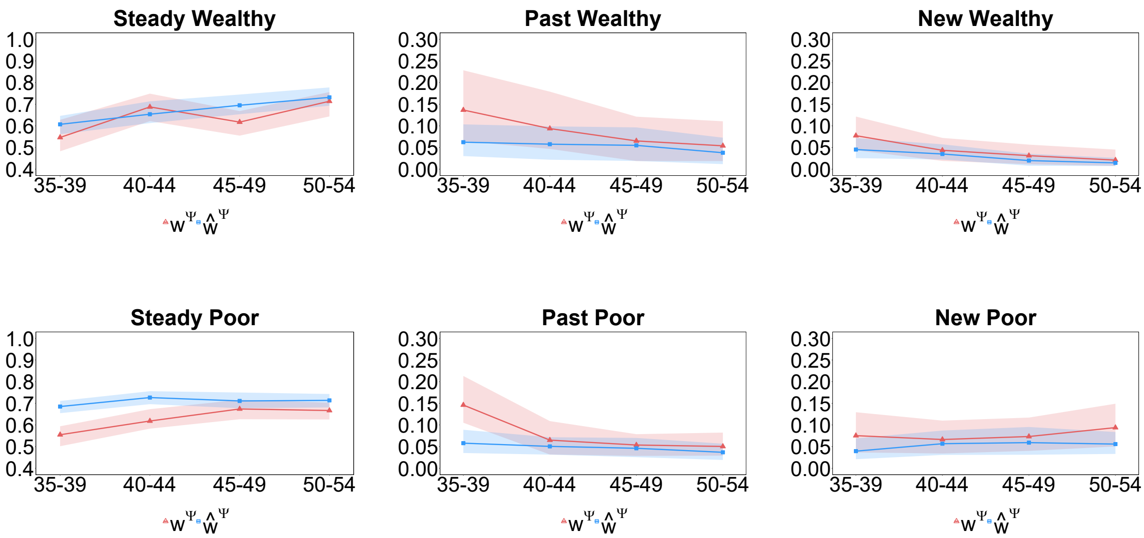


Older age: ages 55-74



Note: these plots show the fraction of high- and low-wealth individuals at each lifecycle stage per birth cohort. These fractions are computed based on actual wealth levels w^{Ψ}. Given that the working life and older age samples contain different individuals, the proportions of high-and low-wealth individuals are not directly comparable across the upper and lower panels.

Figure 23: Rolling window analysis for the discretionary groups.



Note: this plot shows the probability of shifting from one wealth bin to another. The combination of bins considered relates to the definitions of the discretionary groups (see Section 3.4 and Appendix E). For instance, the past wealthy plot displays the probability of an individual moving from the top 10% to the bottom 70% between two lifecycle stages $k-1$ and k. The results are reported for the actual wealth w^{Ψ} and proxy wealth $\hat{w}^{\Omega}$ series. The shaded areas display the 95% confidence intervals, which have been determined through bootstrapping.

Chapter 2

Saving Rate Heterogeneity across the Wealth (Rank) Distribution in the United States [1]

This paper leverages household-level data from the Panel Study of Income Dynamics (PSID) to investigate saving rate heterogeneity across the wealth distribution in the United States. I estimate saving rates across wealth deciles using two complementary approaches: the cross-sectional method and the aggregate method. Four empirical facts emerge. First, saving rates out of labor income and new resources rise with wealth ranks (flow-based saving rates), whereas saving rates out of wealth and composite resources are roughly stable or only modestly increasing (stock-based saving rates). Second, wealth (rank) mobility contributes substantially to saving rate heterogeneity. However, the direction of its effect differs between cross-sectional and aggregate methods due to their distinct treatment of wealth (rank) mobility. Third, the synthetic method (commonly used in the absence of panel data) overestimates saving rates below the 80th percentile, and underestimates them for the top 20%. Fourth, saving increasingly consists of saving out of capital gains as households are richer: wealthier households save for the most part by holding appreciating assets. The paper provides several empirical moments that are of interest to the heterogeneous agent macro literature.

[1]Ghent University (email: christophe.vanlangenhove@ugent.be). I acknowledge the financial support from the Research Foundation Flanders (FWO, project number: 11I5324N). I would like to thank Arthur Apostel, Paula Gobbi, Freddy Heylen, Yasin Kürşat Önder, Gert Peersman, Alberto Russo and Dirk Van de gaer for their valuable feedback on this paper.

1 Introduction

Saving rate heterogeneity is one of the fundamental drivers of wealth inequality (e.g. De Nardi et al., 2017; Hubmer et al., 2021; Van Langenhove, 2025b). Nonetheless, empirical evidence on saving behavior across the wealth (rank) distribution is limited to studies from Norway (Fagereng et al., 2021) and Sweden (Bach et al., 2018). For the United States, the only available evidence comes from Saez & Zucman (2016) and Bauluz & Meyer (2024). However, the saving rate estimates across the wealth distribution in these studies rely on the synthetic method, which only partially incorporates wealth mobility. Therefore, this synthetic method is biased. As a result, heterogeneous agent models of the U.S. wealth distribution have resorted to using Nordic data (e.g. Fernandez-Villaverde & Levintal, 2024) or empirical data across the income (rank) distribution (e.g. Hubmer et al., 2021) as saving behavior calibration targets.

In this paper, I fill this research gap by providing empirical evidence on saving rate heterogeneity across the wealth (rank) distribution in the United States based on household-level data from the Panel Study of Income Dynamics (PSID). I provide an answer to four research questions. First, what is the relationship between total saving rates and the wealth (rank) distribution in the United States? Second, what is the contribution of wealth (rank) mobility to the observed total saving rate patterns across the wealth (rank) distribution? Third, how significant is the bias of the synthetic method compared to unbiased estimation methods? Fourth, does the composition of total saving into active saving and passive saving vary across the wealth (rank) distribution? To address these four research questions, the present paper proceeds in two stages.

In a first stage, I introduce various saving concepts and saving rates, and outline two strategies to estimate saving rates across the wealth (rank) distribution. On the one hand, the saving concepts distinguish between total saving (the change in wealth), active saving (disposable income minus consumption expenditures) and passive saving (capital gains and inter-generational transfers). I define the saving rates out of labor income and new resources (flow-based saving rates) and the saving rates out of wealth and composite resources (stock-based saving rates). On the other hand, I show that the saving rate of a wealth bracket can be estimated using the cross-sectional method (which computes a summary metric over the cross-section of saving rates) and the aggregate method (which estimates saving rates using aggregated variables over households sets). In addition, I propose approaches to quantify the impact of wealth (rank) mobility on observed saving rates patterns across the wealth distribution. I also outline the synthetic estimation method that is used in the literature on U.S. saving behavior and propose an approach to quantify its bias.

In a second stage, I conduct empirical analyses that address the four research questions of this paper. First, I find that total saving rates out of labor income and new resources rise with wealth ranks (flow-based saving rates). In contrast, total saving rates out of wealth and composite resources are roughly stable and only moderately increasing with wealth ranks (stock-based saving rates). Second, wealth (rank) mobility has a significant impact on total saving rate patterns across the wealth distribution. However, while the contribution of wealth mobility is strictly positive for the cross-sectional method, it is negative across most of the wealth distribution for the aggregate method. I show that this discrepancy relates predominantly to these methods' distinct treatment of wealth (rank) mobility: while the cross-sectional method attaches equal weight to all households in a wealth decile, the aggregate method overweighs households that display downward wealth mobility. Third, I find that the synthetic method overestimates saving rates up to the 80th percentile, while it underestimates the saving rates of the top 20%. Fourth, I demonstrate that households' reliance on capital gains rises across the wealth rank distribution: the top wealthiest households' total saving consists for the most part of saving by holding appreciating assets. Passive saving out of inter-generational transfers is more prevalent for wealthier households, but relatively unimportant in magnitude.

Related literature & contributions This paper makes contributions to three strands of the literature.

First, I contribute to the empirical literature studying saving rate heterogeneity across the wealth (rank) distribution. For the Nordic countries, using panel datasets, Bach et al. (2018) provide empirical evidence for Sweden, while Fagereng et al. (2021) do so for Norway. In addition, Saez & Zucman (2016) investigate saving behavior across the wealth distribution for the United States using cross-sectional data without a panel dimension. Because of this data restriction, their estimation relies on the synthetic method, which is biased. The present paper is therefore the first to provide unbiased empirical evidence on the relationship between saving rates and wealth (ranks) for the United States[2]. It is also the first to study the importance of active saving versus passive saving across the U.S. wealth (rank) distribution.

Second, I quantify the contribution of wealth (rank) mobility to the observed saving rate heterogeneity across the wealth (rank) distribution. I also show that estimation differences between the cross-sectional and aggregate method relate almost entirely to their distinct treatment of wealth (rank) mobility. This paper is one of the first to explicitly study the contribution of wealth mobility to saving rate patterns across the entire wealth distribution. One exception is Gomez (2023), which uses an accounting decomposition to analyze the impact of wealth (rank) mobility for Forbes 400 households. However, Gomez (2023) does not analyze the entire wealth distribution, and uses a narrower saving rate definition than is the case in the present paper.

Third, when panel data is absent, some studies have resorted to the synthetic method to estimate saving rate heterogeneity across the wealth (rank) distribution (e.g. Saez & Zucman, 2016; Bauluz & Meyer, 2024). However, the synthetic method is biased as it compares different

[2]There does exist a literature that quantifies saving behavior across the income (rank) or lifetime income (rank) distribution (e.g. Dynan et al., 2004).

sets of households over consecutive time periods. This paper is the first to quantify the bias faced by the synthetic method compared to an unbiased aggregate estimation method.

Roadmap This paper proceeds as follows. Section 2 defines a budget constraint, saving concepts, and two flow-based and two stock-based saving rates. Section 3 elaborates on the cross-sectional and aggregate methods, proposes approaches to quantify the contribution of wealth (rank) mobility, and discusses the synthetic method and its bias. Section 4 outlines the samples taken from the Panel Study of Income Dynamics (PSID) and the measurement of the different variables used for this study. Section 5 provides the empirical results on total saving rate patterns across the wealth (rank) distribution, the contribution of wealth (rank) mobility, and the bias of the synthetic method. Section 6 decomposes total saving into active saving and passive saving. Section 7 concludes.

2 Theoretical framework

In this Section, I outline the theoretical concepts and definitions used to study saving behavior across the wealth distribution. I start by presenting a budget constraint. Based on the budget constraint equation, I then define three saving concepts (total saving, active saving and passive saving) and four saving rates (saving rate out of labor income, new available resources, wealth and composite available resources).

2.1 A budget constraint

To outline theoretical saving behavior concepts, it is worthwhile to define a household budget constraint. Let us consider for a household i the wealth accumulation equation:

$$\Delta w_i(t+1) = y_i(t) + g_i(t) + \left[r_i^i(t) + r_i^c(t)\right] w_i(t) - \tau_i(t) - c_i(t) + m_i(t) + \eta_i(t) \tag{1}$$

where w denotes wealth, y labor income, g government transfer receipts, r^i the return on wealth from dividends and net interest, r^c the return on wealth from realized and unrealized capital gains, τ household tax payments, c consumption expenditures, m net inter-generational transfer receipts, and η a residual category capturing household formation dynamics[3]. It holds that $E\left[\eta_i(t)\right] = 0$.

For later reference, the budget constraint can be used to define two forms of household available resources. On the one hand, a household's *new* resources denote a household's total in-

[3]In the data, this primarily involves children moving out or into the household. Such composition change generates an inflow or outflow of wealth that is unrelated to any of the other budget constraint variables. I correct for this term in the definition of total saving (Equation 4).

flows net of taxes. Algebraically:

$$\Lambda_i^N(t) = y_i(t) + \left[r_i^i(t) + r_i^c(t)\right] w_i(t) + g_i(t) - \tau_i(t) + m_i(t) \tag{2}$$

On the other hand, a household's *composite* resources add to their *new* resources its wealth level at the start of time period t. Specifically:

$$\Lambda_i^C(t) = w_i(t) + \Lambda_i^N(t) \tag{3}$$

2.2 Three saving concepts

I define three saving concepts: total saving, active saving and passive saving. Total saving equals the change in a household's wealth Δw_i over period t corrected for the residual η_i:

$$s_i^T(t) = \Delta w_i(t+1) - \eta_i(t) \tag{4}$$

and reflects the portion of new resources Λ^N that the household has not used for consumption c. This can be seen from Equations 1 and 2.

Total saving s_i^T represents the sum of active saving s_i^A and passive saving s_i^P (see Equation 1). That is:

$$\underbrace{s_i^T}_{\text{Total}} = \underbrace{s_i^A}_{\text{Active}} + \underbrace{s_i^P}_{\text{Passive}} \tag{5}$$

Active saving s^A equals the remaining household's resources after using its received labor income, capital income and government transfers net of taxes (or equivalently, its disposable income) to finance consumption expenditures. Mathematically:

$$s_i^A(t) = y_i(t) + g_i(t) + r_i^i(t)w_i(t) - \tau_i(t) - c_i(t) \tag{6}$$

Passive saving s^P equals the sum of (realized and unrealized) capital gains and net intergenerational transfer receipts. Algebraically:

$$s_i^P(t) = r_i^c(t)w_i(t) + m_i(t) \tag{7}$$

There exists no consensus on the definitions of active and passive saving in the literature. In particular, Fagereng et al. (2021) interpret capital income $r_i^i(t)w_i(t)$ as a component of active saving. On the contrary, Bach et al. (2018) attribute it to passive saving. The definitions used in this paper are in line with those from Fagereng et al. (2021).

2.3 Flow-and stock-based saving rates

To study saving behavior, the existing literature defines different saving rates. I distinguish between two flow-based saving rates and two stock-based saving rates. The former normalize household's saving flows with another flow variable (labor income or new resources), while the latter normalize the saving flow based on a stock variable (wealth) or a variable derived from a stock variable (composite resources). In what follows, I outline these four saving rates, using total saving s^T as an illustration. I also define saving ratios.

Flow-based saving rates I distinguish between two flow-based saving rates. First, some studies consider the saving rate out of labor income (e.g. Fagereng et al., 2021). Using the budget constraint (Equation 1), this saving rate is defined as:

$$\zeta_i^T(t) = \frac{s_i^T(t)}{y_i(t) + g_i(t)} \tag{8}$$

where it should be noted that the denominator includes not only labor income, but also replacement income and other government transfers. For simplicity, however, I refer to the saving rate in Equation 8 as the saving rate out of labor income.

The saving rate out of labor income discriminates between the sources of a household's new resources. An example can illustrate this. Suppose households A and B have identical new resources Λ^N and identical total saving s^T: $\Lambda_A^N = \Lambda_B^N$ and $s_A^T = s_B^T$. However, household A faces higher labor income (and lower capital income) than household B: $y_A > y_B$. As a result, despite identical Λ and s^T, household A reports a lower total saving rate out of labor income ($\zeta_A^T < \zeta_B^T$) relative to household B.

Second, as a response to this, one can define the saving rate out of new resources. Algebraically:

$$\phi_i^T(t) = \frac{s_i^T(t)}{\Lambda_i^N(t)} \tag{9}$$

which is invariant to the composition of a household's new resources. Specifically, in the previous illustrative example, households A and B display identical saving rates out of new resources ($\phi_A^T = \phi_B^T$) even though they derive new resources from different sources (mostly labor income versus mostly capital income). The saving rate out of new resources is conceptually similar to the saving rate out of total income.

Stock-based saving rates The flow-based saving rates normalize a household's savings flow based on the financial flows a household has received from its human capital (labor income), accumulated wealth (capital income and gains) and other sources (government and family). However, in addition to its inflows at t, a household could draw down its wealth stock to finance consumption expenditures. Stock-based saving rates take this option into account.

As a third saving rate, a strand of literature (e.g. Bach et al., 2018) considers the saving rate out of wealth. For total saving, the saving rate out of wealth is defined as:

$$\mu_i^T(t) = \frac{s_i^T(t)}{w_i(t)} \tag{10}$$

which is equivalent to the growth rate of the household's wealth (using the time subscript notation from Equation 1). The saving rate out of wealth abstracts from the income, capital gains and transfer flows a household may have obtained in period t.

Fourth, as a response to this, I define the saving rate out of composite resources. It is set equal to:

$$\theta_i^T(t) = \frac{s_i^T(t)}{\Lambda_i^C(t)} \tag{11}$$

which shows how much a household adds to its wealth relative to the total resources it has available for consumption (or alternatively, its maximum possible consumption at t[4]).

Saving ratios The flow-based and stock-based saving rates consider in the numerator a saving variable (total saving, active saving or passive saving). These saving rates have been defined using a budget constraint for the change in total wealth Δw (Equation 1). Alternatively, some heterogeneous agent models work with the saving ratio ξ, which is defined as:

$$\xi(t) = \frac{w_i(t)}{\Lambda_i^C(t)} \tag{12}$$

and has wealth w in the numerator, rather than its first difference. This has as a drawback that a decomposition into active and passive components is infeasible. For that reason, I do not report saving ratio outcomes in the main text of this paper. However, Appendix F provides empirical evidence on saving ratios ξ across the wealth (rank) distribution for the United States.

3 Estimating saving rates across the wealth distribution

Our aim is to investigate saving rate heterogeneity across the wealth distribution. For practical purposes, this comes down to studying saving rates across the wealth rank distribution, for instance across wealth deciles. Such a transition from the wealth distribution to the wealth decile distribution raises two key questions. First, what method should be used to estimate the saving rate of a wealth decile from the available data? Second, how does wealth (rank) mobility affect the saving rate estimates per wealth decile?

[4]This statement is true only when there exists a borrowing constraint at zero: $w \geq 0$. Insofar as the borrowing constraints lies below zero (consumer or other credit), composite resources underestimate households' total consumption capabilities.

In this Section, I introduce some additional definitions, and outline two complementary methods to estimate saving rates across the wealth rank distribution: the cross-sectional method and the aggregate method. While the former computes some summary metric over the cross-section of household-level saving rates per wealth decile, the latter relies on aggregates per wealth decile. I also propose approaches to quantify the contribution of wealth (rank) mobility to saving rate estimates. Finally, I compare the cross-sectional and aggregate method to the so-called synthetic method that has been used to estimate saving rate patterns for the U.S. when no household-level saving rate or panel data is available (e.g. Saez & Zucman, 2016; Bauluz & Meyer, 2024).

3.1 Composition of a wealth decile

Let us define as P_t^d the set of households belonging to wealth decile d at time period t. The composition of a wealth decile d is likely to change over time: due to household-level heterogeneity in the total saving rate out of wealth, there may exist turnover of households across wealth deciles. If such wealth rank mobility takes place, it holds that $P_t^d \neq P_{t-1}^d$. In addition, the existence of wealth mobility implies the presence of exiting and entrant households for a decile d.

On the one hand, let O_t^d and D_t^d denote two sets of households that exit wealth decile d at t. These exiting households respectively already belonged to the sample at $t-1$ (O^d) and exited the sample at t because of death or non-response (D_t^d). On the other hand, define as I_t^d and B_t^d two sets of households that entered into the wealth decile at t. These entrant households respectively already belonged to the sample at $t-1$ (I^d) and entered the sample at t (B^d).

Using the exiting and entrant households definitions, the set of households belonging to wealth decile d at a time period t can be written algebraically as:

$$P_t^d = P_{t-1}^d \setminus (O_t^d \cup D_t^d) \cup (I_t^d \cup B_t^d) \tag{13}$$

where $P_{t-1}^d \setminus (O_t^d \cup D_t^d)$ represents the immobile households that stayed in the same wealth decile d over two consecutive time periods. Let us define these immobile households as:

$$S_t^d = P_{t-1}^d \setminus (O_t^d \cup D_t^d) \tag{14}$$

so that the set of households in a decile d at any t (P_t^d) can be conveniently re-written as the union of immobile households and entrant households:

$$P_t^d = S_t^d \cup (I_t^d \cup B_t^d) \tag{15}$$

which allows to distinguish between two types of wealth mobility. First, some households may enter a decile d at t because of upward or downward wealth mobility: $I_t^d \neq \emptyset$. I refer to such mobility as endogenous wealth (rank) mobility. Second, the composition of decile d might change because households that previously did not belong to the sample enter the sample in decile d at t: $B_t^d \neq \emptyset$. I label such mobility as sample-related wealth (rank) mobility.

3.2 Two estimation methods

There exist two complementary methods to estimate the saving rate of a wealth decile d: a cross-sectional method and an aggregate method. In what follows, I detail each of the two estimation methods and propose methods to quantify the impact of wealth (rank) mobility on the saving rate estimates. I depart from the total saving rate out of wealth (Equation 10) as an illustrative example. In terms of notation, the saving rate estimate for a wealth decile d is referenced as $\tilde{\mu}^{T,d}$.

3.2.1 Cross-sectional method

The cross-sectional method estimates the total saving rate out of wealth by taking a summary metric over a cross-section of household-level saving rates. Common summary approaches used in this cross-sectional method include taking the mean (e.g. Bach et al., 2018) or taking the median (e.g. Fagereng et al., 2021). Using the median, the saving rate out of wealth estimate is obtained as:

$$\tilde{\mu}^{T,d}(t) = \text{median}\left\{\mu_i^T(t) \,:\, i \in P_t^d\right\} \tag{16}$$

which calculates the median saving rate across all households in decile d at time period t. As the set P_t^d contains both immobile households and all entrant households, it takes into account all wealth (rank) mobility (both endogenous and sample-related).

How to quantify the impact of wealth (rank) mobility on saving rate estimates according to the cross-sectional method? I distinguish between two approaches that relate to the two wealth mobility concepts outlined in Section 3.1. First, in a broad approach, I take the difference between the saving rate estimate from Equation 16 and a counterfactual saving rate estimate computed for the subset of immobile households. Algebraically:

$$\textit{broad:} \quad \text{median}\left\{\mu_i^T(t) \,:\, i \in P_t^d\right\} - \text{median}\left\{\mu_i^T(t) \,:\, i \in S_t^d\right\} \tag{17}$$

which quantifies the joint contribution of endogenous and sample-related wealth mobility to the saving rate estimates across deciles d. However, we are primarily interested in the contribution of endogenous wealth mobility to saving rate patterns. Second, therefore, under the narrow approach I compute:

$$\textit{narrow:} \quad \text{median}\left\{\mu_i^T(t) \,:\, i \in P_t^d \setminus B_t^d\right\} - \text{median}\left\{\mu_i^T(t) \,:\, i \in S_t^d\right\} \tag{18}$$

where the left-most term represents the median saving rates of all households in decile d at t, with the exception of those that entered decile d from out of the sample (B_t^d). As a result, Equation 18 quantifies the contribution of endogenous wealth (rank) mobility to the estimated saving rate from decile d.

3.2.2 *Aggregate method*

The cross-sectional method relies on the cross-section of household-level saving rates in a wealth decile d. On the contrary, the aggregate method estimates saving rates per decile d using aggregated variables for that d. In what follows, I denote as $w_{\mathcal{P}}(t)$ the average wealth of the households in set $\mathcal{P}$ at time period t[5].

The aggregate method estimates the total saving rate out of wealth for a wealth decile d as follows:

$$\tilde{\mu}^{T,d}(t) = \frac{w_{\mathcal{P}_t^d}(t) - w_{\mathcal{P}_t^d}(t-1)}{w_{\mathcal{P}_t^d}(t-1)} \tag{19}$$

which represents the growth rate of the aggregate wealth held by households belonging to wealth decile d at time period t. This approach entirely incorporates wealth (rank) mobility: it allows for exiting and entrant households and computes the growth rate over an identical household set. As the same group of households is traced over two time periods, the aggregate method requires panel data.

How to quantify the contribution of wealth (rank) mobility to saving rates across deciles d estimated using the aggregate method? In line with the cross-sectional method, I distinguish between a broad and narrow approach. First, the broad approach calculates:

$$\textit{broad:} \quad \frac{w_{\mathcal{P}_t^d}(t) - w_{\mathcal{P}_t^d}(t-1)}{w_{\mathcal{P}_t^d}(t-1)} - \frac{w_{S_t^d}(t) - w_{S_t^d}(t-1)}{w_{S_t^d}(t-1)} \tag{20}$$

which equals the difference between the saving rate estimation from Equation 19 and the saving rate computed for households that remained in that decile over two consecutive time periods (i.e. immobile households). Therefore, Equation 21 quantifies the joint contribution of endogenous and sample-related wealth mobility. Second, the narrow approach quantifies the contribution of endogenous wealth mobility only. It therefore calculates:

$$\textit{narrow:} \quad \frac{w_{\mathcal{P}_t^d \setminus B_t^d}(t) - w_{\mathcal{P}_t^d \setminus B_t^d}(t-1)}{w_{\mathcal{P}_t^d \setminus B_t^d}(t-1)} - \frac{w_{S_t^d}(t) - w_{S_t^d}(t-1)}{w_{S_t^d}(t-1)} \tag{21}$$

[5]In other words, the aggregate wealth of all households in the set divided by the number of households in that set.

where the left-most term represents the average wealth growth of all households in decile d at t, with the exception of those that entered decile d from out of the sample.

3.3 The synthetic method

The cross-sectional method from Equation 16 and aggregate method from Equation 19 require household-level saving rate or panel data. Often, however, researchers do not have such datasets available. When this is the case, some studies have resorted to a synthetic form of the aggregate method (e.g. Saez & Zucman, 2016; Bauluz & Meyer, 2024). This synthetic method computes the total saving rate out of wealth for a wealth decile d as:

$$\tilde{\mu}^{T,d}(t) = \frac{w_{\mathcal{P}_t^d}(t) - w_{\mathcal{P}_{t-1}^d}(t-1)}{w_{\mathcal{P}_{t-1}^d}(t-1)} \tag{22}$$

which represents the growth rate of the aggregate wealth held by households belonging to wealth decile d. The synthetic method requires only aggregated data per wealth decile. However, it only partially incorporates wealth (rank) mobility: while it computes aggregate wealth at t over all households belonging to the decile d, aggregate wealth at $t-1$ is calculated across all households belonging to the same decile at $t-1$. Therefore, the method compares a different group of households across the two time periods if wealth (rank) mobility has taken place, i.e. whenever:

$$P_t^d \neq P_{t-1}^d \tag{23}$$

which introduces an estimation bias: Equation 22 is expected to underestimate saving rates at the top of the wealth distribution, and overestimate them at the bottom and the middle of the distribution (e.g. Bach et al., 2018). This bias can be quantified as:

$$\frac{w_{\mathcal{P}_t^d}(t) - w_{\mathcal{P}_t^d}(t-1)}{w_{\mathcal{P}_t^d}(t-1)} - \frac{w_{\mathcal{P}_t^d}(t) - w_{\mathcal{P}_{t-1}^d}(t-1)}{w_{\mathcal{P}_{t-1}^d}(t-1)} \tag{24}$$

which equals the difference between the biased method from Equation 22 relative to the unbiased aggregate method from Equation 19.

4 Data & empirical strategy

In this Section, I describe the dataset and empirical strategy implemented to generate empirical evidence on saving behavior across the U.S. wealth (rank) distribution. I use two samples of the Panel Study of Income Dynamics (PSID), a representative panel dataset of U.S. households. The two samples cover the periods 2001-2021 and 2005-2021 respectively, as explained in more detail below. Given the panel structure of the PSID-dataset, both the cross-sectional and aggregate method can be applied. In addition, it is possible to quantify the contribution of wealth (rank) mobility to the saving rates estimates, and to quantify the bias of the synthetic method.

For the computation of total saving rates out of labor income and out of wealth, it suffices to have a measure of labor income and wealth at the household level. However, more data is needed when considering saving rates out of new resources and composite resources, or when distinguishing between active and passive saving. In this case, one needs data on all variables in the budget constraint in Equation 1. Each of these budget constraint variables can be directly or indirectly imputed from the PSID.

In what follows, I first briefly outline the key properties of the PSID-dataset and the imputation of the budget constraint variables. Thereafter, I discuss the properties of the two samples.

4.1 Dataset

The empirical analyses in this paper leverage household-level data from the Panel Study of Income Dynamics (PSID). Specifically, I use the SRC-subsample, as is common in economic research (e.g. Straub, 2019; Van Langenhove, 2025a). This is a representative sample. However, it underrepresents the top of the wealth distribution (e.g. Pfeffer et al., 2016; Van Langenhove, 2025a). This implies that the PSID cannot be used to investigate saving rate and wealth inequality at the tail. I therefore use the top 10% to represent the wealthiest households. Van Langenhove (2025a) finds that the top wealth bias relative to the Survey of Consumer Finances (SCF) is stable over time.

The PSID contains sufficiently rich micro-level information to impute most of the budget constraint variables from the responses to the household-level questionnaire. There are two exceptions to this, however. First, consumption c is measured accurately only from 2005 onwards. Second, tax payments τ are not reported and therefore need to be imputed. I do this tax imputation using the NBER TAXSIM program (V35). A detailed explanation on the imputation of the budget constraint variables using the PSID and the NBER TAXSIM program is provided in Appendix A.

4.2 Empirical strategy

Saving flows & rates Total, active and passive saving flows for a household i are imputed based on the PSID as:

$$\tilde{s}_i^T(t) = \Delta\tilde{w}_i(t+1) - \tilde{\eta}_i(t) \tag{25}$$

$$\tilde{s}_i^A(t) = \tilde{y}_i(t) + \tilde{g}_i(t) + \tilde{r}_i^i(t)\tilde{w}_i(t) - \tilde{\tau}_i(t) - \tilde{c}_i(t) \tag{26}$$

$$\tilde{s}_i^P(t) = \tilde{r}_i^c(t)\tilde{w}_i(t) + \tilde{m}_i(t) \tag{27}$$

where the $\tilde{x}$-notation reflects the PSID-generated measure of a variable x. As active saving relies on the PSID consumption-estimate $\tilde{c}$, it can be computed only over the 2005-2021 period. On the contrary, total saving and passive saving can be computed in both the 2001-2021 and 2005-2021 periods.

The three saving flows are normalized using labor income, new resources, wealth and composite resources to obtain four saving rates (as explained in Section 2.3). Imputing labor income $\tilde{y}$ and wealth $\tilde{w}$ is relatively straightforward. However, the computation of new resources $\tilde{\Lambda}^N$ and composite resources $\tilde{\Lambda}^C$ relies on close to the full set of budget constraint variables. For the calculation of new and composite resources, I distinguish between an inflow-based and expenditure-based approach.

The inflow-based approach uses the 2001-2021 sample and imputes new and composite resources directly from Equations 2 and 3:

$$\tilde{\Lambda}_i^N(t) = \tilde{y}_i(t) + \tilde{g}_i(t) + \tilde{r}_i^i(t)\tilde{w}_i(t) + \tilde{r}_i^c(t)\tilde{w}_i(t) - \tilde{\tau}_i(t) + \tilde{m}_i(t) \tag{28}$$

$$\tilde{\Lambda}_i^C(t) = \tilde{w}_i(t) + \tilde{\Lambda}_i^N(t) \tag{29}$$

which relies on the imputation of a large number of variables in the PSID. Conversely, the expenditure-based approach uses the 2005-2021 sample to estimate resources from the budget constraint in Equation 1. It relies on the PSID-measure of consumption expenditures $\tilde{c}$. Specifically:

$$\tilde{\Lambda}_i^N(t) = \Delta\tilde{w}_i(t+1) + \tilde{c}_i(t) - \tilde{\eta}_i(t) \tag{30}$$

$$\tilde{\Lambda}_i^C(t) = \tilde{w}_i(t+1) + \tilde{c}_i(t) - \tilde{\eta}_i(t) \tag{31}$$

In the remainder of this paper, I will present the total saving rates out of new resources and out of composite resources calculated based on the inflow-based approach over the 2001-2021 sample. Appendix G compares the baseline outcomes based on the inflow-based approach to those using the expenditure-based approach over the 2005-2021 sample. They yield highly similar results. The sample sizes vary between 30000 and 40000 observations, depending on the saving rate under consideration[6] and after applying the two sample restrictions discussed below.

Treatment of edge cases The PSID-samples contain households that display zero labor income or new resources. Moreover, some households have zero or negative wealth or composite resources. This generates distorted saving rate estimates. To deal with these edge-cases, I impose two restrictions for each saving rate. These restrictions apply both to the cross-sectional

[6] For saving rates out of new resources and composite resources, the sample size is closer to 30000 given that the calculation of resources relies on several budget constraint variables. This raises the probability of household item non-response. The same applies to active saving. I have checked that the results of this paper are robust to sample definitions.

and aggregate method. First, for flow-based saving rates, whenever the flow in the denominator (labor income or new resources) equals zero, the corresponding saving rate is set equal to zero as well. Second, for stock-based saving rates, the denominator (wealth or composite resources) can become negative if the household is indebted. Whenever this is the case, the saving rate is set to 0.05 when the saving flow in the numerator is positive, and to -0.05 when it is negative. When the denominator equals zero, the corresponding saving rate is set to zero also. These edge case restrictions predominantly affect saving rates at the bottom of the wealth distribution (wealth percentile 35 and below), where saving rates are therefore ill-defined throughout this paper.

Two sample restrictions Finally, I impose two additional sample restrictions. First, I restrict the 2001-2021 and 2005-2021 samples to households where the PSID-designated reference person of the household is older than 20. This restriction is in line with for example Bach et al. (2018), and a large literature on wealth inequality measurement. Second, I trim for each year the most extreme 0.5 percent saving rate observations at both the bottom and top of its distribution. This is a common practice in the literature using PSID data (e.g. Gaillard & Wangner, 2023; Straub, 2019). It has been verified that the results of this paper are robust to alternative trimming parameters.

5 Total saving rates across the wealth rank distribution

In this Section, I address three research questions. First, what is the relationship between total saving rates and wealth (ranks) according to the cross-sectional and aggregate method? Second, what is the contribution of wealth (rank) mobility to total saving rate estimates according to both methods? Third, how large is the bias of the synthetic method compared to the (unbiased) aggregate method in estimating total saving rates across the wealth (rank) distribution?

5.1 Total saving rates across the wealth (rank) distribution

How do total flow-based and stock-based saving rates vary with wealth (ranks)? I present the empirical evidence for these total saving rates according to the cross-sectional method and the aggregate method. Results are displayed in Table 1. Three key findings persist.

First, at the extensive margin, the share of households with positive total saving (or equivalently, rising wealth levels) increases across the wealth rank distribution. For the bottom 30%, merely 25% to 40% of households display positive total saving. This fraction rises to above 50% from the 40th wealth percentile onwards, and continues to rise at higher wealth percentiles. The share of positive total savers peaks at a little over 70% for the top 20% wealthiest households (Table 1). Inversely, this implies that still close to 30% of the top wealthiest households displays negative total saving rates.

Table 1: Total saving rates across wealth deciles using the cross-sectional method and aggregate method.

Wealth Bin	**1–10**	**11–20**	**21–30**	**31–40**	**41–50**	**51–60**	**61–70**	**71–80**	**81–90**	**91–100**
Positive Savers (%)	28%	33%	40%	58%	63%	65%	65%	65%	69%	71%
Saving Rate out of Labor Income										
Cross-Sectional	-0.13	-0.01	0.00	0.03	0.09	0.13	0.19	0.24	0.44	0.85
	(-0.14, -0.12)	(-0.01, -0.01)	(0.00, 0.00)	(0.02, 0.03)	(0.08, 0.10)	(0.12, 0.14)	(0.17, 0.20)	(0.21, 0.26)	(0.40, 0.47)	(0.80, 0.93)
Aggregate	-0.05	-0.05	-0.04	-0.04	-0.00	0.04	0.06	0.12	0.29	1.28
	(-0.05, -0.05)	(-0.05, -0.05)	(-0.05, -0.04)	(-0.05, -0.03)	(-0.02, 0.02)	(0.00, 0.07)	(0.01, 0.11)	(0.05, 0.18)	(0.20, 0.40)	(0.80, 1.73)
Saving Rate out of New Resources										
Cross-Sectional	-0.10	0.00	0.00	0.03	0.09	0.13	0.19	0.21	0.30	0.38
	(-0.11, -0.09)	(-0.00, 0.00)	(0.00, 0.00)	(0.03, 0.04)	(0.08, 0.10)	(0.12, 0.14)	(0.17, 0.20)	(0.20, 0.23)	(0.28, 0.32)	(0.36, 0.40)
Aggregate	-0.05	-0.05	-0.04	-0.04	-0.01	0.04	0.06	0.11	0.23	0.53
	(-0.05, -0.05)	(-0.05, -0.05)	(-0.05, -0.03)	(-0.05, -0.02)	(-0.02, 0.01)	(0.00, 0.07)	(0.01, 0.12)	(0.05, 0.17)	(0.15, 0.30)	(0.37, 0.65)
Saving Rate out of Wealth										
Cross-Sectional	-0.05	-0.05	0.00	0.05	0.11	0.10	0.09	0.08	0.09	0.10
	(-0.05, -0.05)	(-0.05, -0.05)	(0.00, 0.00)	(0.05, 0.05)	(0.09, 0.12)	(0.09, 0.12)	(0.09, 0.10)	(0.07, 0.09)	(0.09, 0.10)	(0.09, 0.11)
Aggregate	-0.05	-0.05	-0.05	-0.05	0.00	0.04	0.03	0.03	0.06	0.11
	(-0.05, -0.05)	(-0.05, -0.05)	(-0.05, -0.05)	(-0.05, -0.05)	(-0.02, 0.03)	(-0.00, 0.07)	(-0.00, 0.06)	(0.00, 0.06)	(0.03, 0.08)	(0.07, 0.15)
Saving Rate out of Composite Resources										
Cross-Sectional	-0.05	-0.01	0.00	0.03	0.05	0.06	0.06	0.06	0.07	0.08
	(-0.05, -0.05)	(-0.02, -0.01)	(0.00, 0.00)	(0.02, 0.03)	(0.05, 0.06)	(0.06, 0.07)	(0.06, 0.07)	(0.05, 0.06)	(0.07, 0.08)	(0.08, 0.09)
Aggregate	-0.05	-0.05	-0.04	-0.03	-0.01	0.01	0.02	0.02	0.04	0.07
	(-0.05, -0.05)	(-0.05, -0.05)	(-0.05, -0.03)	(-0.05, -0.01)	(-0.02, 0.01)	(-0.01, 0.04)	(-0.00, 0.04)	(0.00, 0.04)	(0.03, 0.06)	(0.06, 0.10)

Note: this table shows (1) the fraction of positive savers per wealth decile, and (2) total saving rates per wealth decile, computed according to the cross-sectional method (Equation 16) and the aggregate method (Equation 19). The 95% confidence intervals have been determined using bootstrapping and are shown in parentheses. The calculations are executed for the flow-based saving rates (out of labor income and new resources) and stock-based saving rates (out of wealth and composite resources). The 2001–2021 sample is used. Edge cases are dealt with as specified in Section 4.2. The edge cases imply that the saving rates out of wealth and composite resources are ill-defined for the bottom 35% (wealth decile 31-40 and lower).

Second, at the intensive margin, the two total flow-based saving rates (saving rates out of labor income and new resources) increase significantly across the wealth rank distribution. The difference between saving rates at the top and middle of the wealth distribution is the highest for the saving rate out of labor income (Table 1). This follows from varying shares of labor income to total income: on average, the wealthier are more reliant on capital income than households in the middle of the wealth (rank) distribution. On the contrary, for the stock-based saving rates, the saving rate out of wealth is roughly stable (or slightly declining) from wealth decile 41-50 to wealth decile 81-90, but rises again for the top 10% wealthiest. The saving rate out of composite resources is moderately increasing from the middle part of the wealth distribution onwards (Table 1).

Third, the cross-sectional method and aggregate method yield similar saving rate patterns across the wealth (rank) distribution: for both methods, saving rates out of labor income and new resources are strongly increasing with wealth ranks (flow-based saving rates), while saving rates out of wealth and composite resources are respectively increasing only at the top or increasing moderately with wealth ranks (stock-based saving rates). However, there does exist a level difference between both estimation methods: the aggregate method predicts higher saving rates at the top 10% compared to the cross-sectional method, but lower saving rates over the remainder of the wealth distribution (Table 1). In Section 5.3, I show that this discrepancy relates primarily to these methods' distinct treatment of wealth (rank) mobility.

Literature comparison How do the observed total saving rate patterns compare to other research? I compare these patterns to existing research for the United States and the Nordic countries.

For the United States, Saez & Zucman (2016) employ the synthetic method (as outlined in Equation 22) to compute the total saving rate out of income. They find that this saving rate rises with wealth (ranks), both over their full sample (1917-2012) and over the period that overlaps with my sample (2001-2012). This finding is in line with my results for the total saving rate out of new resources (Table 1), which is methodologically closest aligned to the total saving rate out of income used by Saez & Zucman (2016). However, for the 2001-2012 period, Saez & Zucman (2016) obtain a total saving rate out of income of 7% to 15% for the top 10% to 1% wealthiest, and a saving rate of 35% to 38% for the top 1%. On the contrary, my aggregate method estimate for the top 10% based on the PSID yields a total saving rate out of new resources that is significantly higher, at 53%. In Section 5.4, I argue that this divergence most likely follows from the bias in the synthetic method that is used by Saez & Zucman (2016).

Using Swedish administrative data, Bach et al. (2018) establish that the total saving rate out of wealth declines with wealth (ranks): their (median) saving rates out of wealth decline from 11.4% at decile 40-50 to 7.4% (and lower) from the 90th percentile onwards based on the cross-sectional method. On the contrary, based on the same method, I obtain saving rates out of wealth that are higher and only slightly declining from 11% to 9% from percentile 40 onwards. Moreover, the saving rates out of wealth rise to approximately 10% again for the top 10% wealthiest (Table 1). There could be three reasons for this divergence between U.S. and Swedish data. First, Bach et al. (2018) compute saving rates at the individual-level, while I do so at the household-level. Second, I consider a sample over period 2001-2021, while Bach et al. (2018) covers instead the period 2000-2007. Third, the divergence may reflect actual differences in saving rate heterogeneity between Sweden and the United States: it is possible that saving rate inequality in the United States is higher compared to Sweden.

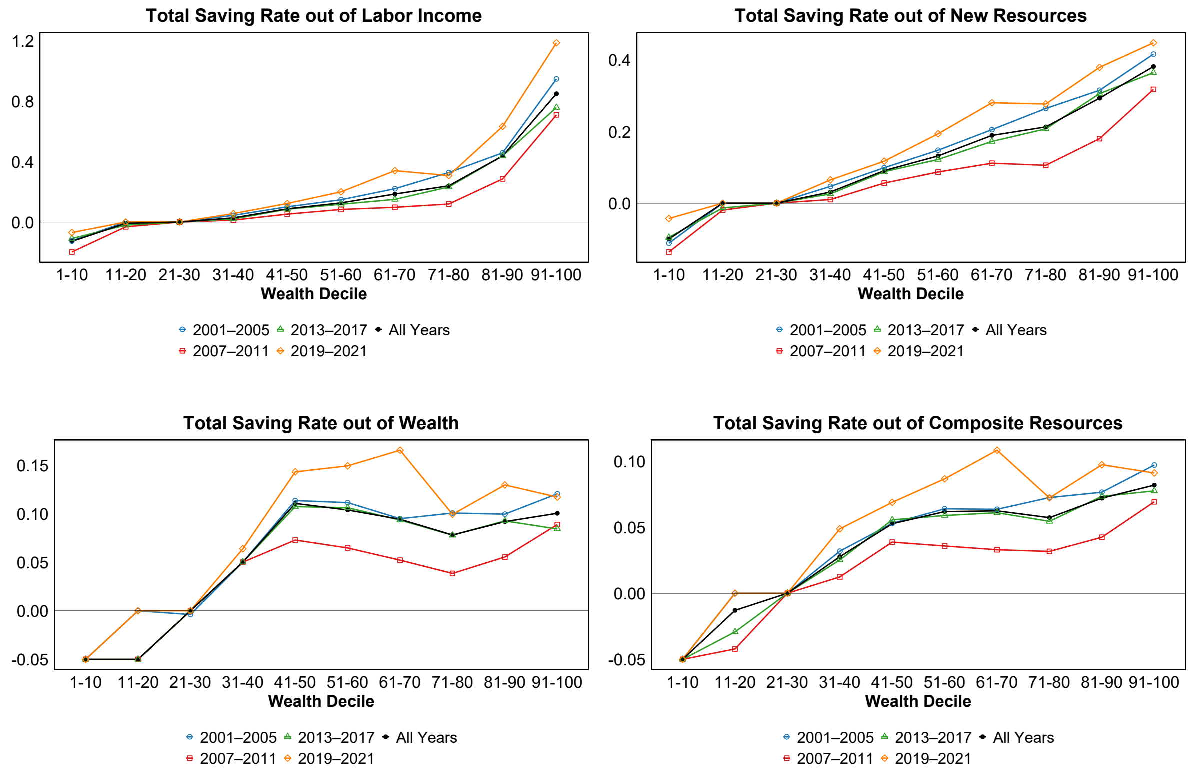
Fagereng et al. (2021) use Norwegian administrative data to study total saving rates out of labor income and wealth at the household level using the cross-sectional method. On the one hand, these authors find that the total saving rate out of labor income rises strongly with wealth (ranks). This is in line with my results for the U.S. (Table 1). However, the functional form of the relationship between saving rates out of labor income and wealth (ranks) differs between my study and theirs: I find a convex relationship from the 41st percentile onwards, while Fagereng et al. (2021) obtain a concave pattern. Moreover, the total saving rate out of labor income of the top 10% is significantly higher in my analysis (over 80%) compared to Fagereng et al. (2021) (at most 50%). On the other hand, in line with Bach et al. (2018), Fagereng et al. (2021) obtain a total saving rate out of wealth that declines strongly with wealth (ranks) and is lower for the top 50% of the wealth distribution than is the case for the U.S.. While the difference may relate in part to diverging sample periods (2001-2021 in this paper versus 2005-2015 in Fagereng et al.), these results suggest greater saving rate inequality in the U.S. than in Norway.

Has the relationship changed over time? Does there exist time variation in the relationship between total saving rates and wealth (ranks), and/or time variation in saving rates levels for any wealth (rank)? I analyze this in Figure 1 for the cross-sectional method and Figure 2 for the aggregate method. Two findings persist.

First, the relationship between total saving rates and wealth (ranks) is stable over time. More precisely, total saving rates out of labor income and new resources rise significantly with wealth ranks (flow-based saving rates). On the contrary, total saving rates out of wealth are roughly stable (or slightly declining) from the middle of the wealth distribution onwards, before rising again at the top 10% (stock-based saving rates). Finally, total saving rates out of composite resources are moderately increasing with wealth ranks (stock-based saving rates).

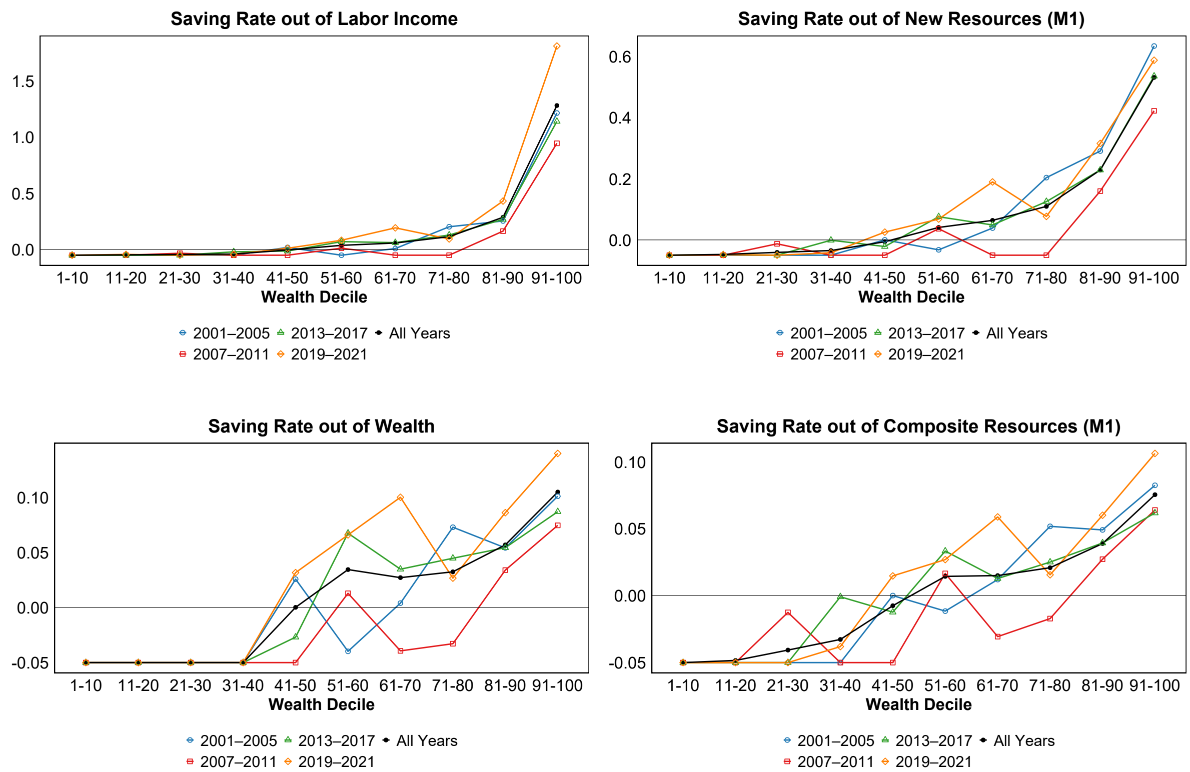
Second, however, there does exist significant time variation in the level of total saving rates across the wealth (rank) distribution: the 2019-2021 period was characterized by higher saving over the entire distribution, while saving for the period 2007-2011 was unusually low. The higher saving in the 2019-2021 period follows from a combination of higher active saving and higher passive saving, while the lower saving in the 2007-2011 period relates primarily to very low passive saving (Figures 5 and 6, Appendix B).

Figure 1: Cross-sectional method – total flow-based and stock-based saving rates across wealth deciles over different survey waves.



Note: this figure plots saving rates per wealth decile, grouped across different PSID survey waves. The saving rates per decile have been estimated using the cross-sectional method (Equation 16). The waves are pooled as: $\{2001, 2003, 2005\}$, $\{2007, 2009, 2011\}$, $\{2013, 2015, 2017\}$ and $\{2019, 2021\}$. The black line shows the baseline results without grouping (from Table 1). The calculations are executed for the two flow-based (total saving rates out of labor income and new resources) and two stock-based saving rates (total saving rates out of wealth and composite resources). The 2001-2021 sample is used. Edge cases are dealt with as specified in Section 4.2. The edge cases imply that the saving rates out of wealth and composite resources are ill-defined for the bottom 35% (wealth decile 31-40 and lower).

Figure 2: Aggregate method – flow-based and stock-based saving rates across wealth deciles over different survey waves according to the aggregate method.



Note: this figure plots saving rates per wealth decile, grouped across different PSID survey waves. The saving rates per decile have been estimated using the aggregate method (Equation 19). The waves are pooled as: $\{2001, 2003, 2005\}, \{2007, 2009, 2011\}, \{2013, 2015, 2017\}$ and $\{2019, 2021\}$. The black line shows the baseline results without grouping (from Table 1). The calculations are executed for the two flow-based (total saving rates out of labor income and new resources) and two stock-based saving rates (total saving rates out of wealth and composite resources). The 2001-2021 sample is used. Edge cases are dealt with as specified in Section 4.2. The edge cases imply that the saving rates out of wealth and composite resources are ill-defined for the bottom 35% (wealth decile 31-40 and lower).

5.2 Robustness of the saving rate patterns

The patterns of total flow-based and stock-based saving rates across the wealth (rank) distribution according to the cross-sectional method are robust to several checks. These robustness checks are presented in Appendix B. The same conclusions pertain when applying the robustness checks to the aggregate method.

First, I report the empirical results when summarizing cross-sectional data for a set of households using the mean instead of the median (Figure 7, Appendix B). This check does not fundamentally alter the empirical conclusions of this paper, with one exception: the total saving rate out of wealth is higher when using the mean instead of the median and declines with wealth (ranks). The other stock-based saving rate − the total saving rate out of composite resources − continues to display a slightly increasing pattern over the wealth distribution, but also takes on somewhat higher values when using the mean rather than the median.

Second, labor income may affect the relationship between total saving rates and wealth (ranks). To investigate this, I group households according to their labor income ranks (where labor income includes government transfer income), and plot the relationship between the saving rates and wealth ranks for each of these groups (Figure 8, Appendix B). I find that the main empirical conclusion outlined above is not affected by labor income: for all labor income groups, flow-based saving rates increase with wealth (ranks), while stock-based saving rates are stable (wealth) or slightly increasing (composite resources). However, the level of the stock-based saving rates (conditional on a household's wealth decile) is slightly higher for household groups with higher labor income[7].

Third, age might constitute a key variable affecting the link between total saving rates and wealth (ranks). Using a similar procedure as for the previous robustness check, I group households according to the age of their reference person, and plot the relationship between the saving rates and wealth (ranks) for each of these groups (Figure 9, Appendix B). Two findings persist. On the one hand, the relationship between total saving rates and wealth (ranks) appears unaffected by age: flow-based saving rates rise with wealth (ranks), while stock-based rates are relatively stable or slightly increasing. On the other hand, the level effect of age is strong: conditional on a household's position in the wealth distribution, its expected saving rate is significantly higher at younger ages. This is in line with a literature on the lifecycle dynamics of wealth accumulation (e.g. Bauluz & Meyer, 2024).

Fourth, total saving rates may be different between households that are business owners (entrepreneurs) and households that are not. Figure 10 (Appendix B) generates two findings. On the one hand, the relationship between saving rates and wealth (ranks) remains consistent with previous results when conditioning on households that are not entrepreneurs. On the

[7] This is by definition not the case for flow-based saving rates, as these are normalized entirely (saving rate out of labor income) or predominantly (saving rate out of new resources) by labor income itself.

other hand, for the wealth deciles where sufficient data on entrepreneurs is available, the saving rates of entrepreneurs significantly exceed the saving rates of non-entrepreneurs across the entire region of the wealth distribution where saving rates are defined.

5.3 Contribution of wealth (rank) mobility

What is the contribution of endogenous wealth (rank) mobility to the total saving rate patterns observed in Section 5.1 (Table 1)? In what follows, I quantify the contribution of wealth (rank) mobility using the narrow approaches for the cross-sectional and aggregate method, as outlined in Section 3.2. I report the empirical results only for the total saving rate out of wealth (in Table 2). The outcomes for the other saving rates are shown in Appendix C and yield similar conclusions. For the sake of briefness, I refer to wealth rank mobility as wealth mobility in what follows.

For the cross-sectional method, the contribution of wealth mobility to total saving rates out of wealth is positive along the entire wealth distribution (in the region where saving rates are defined): wealth mobility accounts for roughly 40% to 60% of the total saving rate out of wealth estimates according to the cross-sectional method. On the contrary, for the aggregate method, the contribution of wealth mobility is negative in the middle part of the wealth distribution (from wealth decile 41-50 to wealth decile 71-80), and positive only for the top 10% wealthiest. Also here the contribution of wealth mobility is substantial: in the absence of wealth mobility, total saving rates out of wealth according to the aggregate method would double in the middle of the wealth distribution. At the top, wealth mobility accounts for close to 40% of the saving rate estimates.

What explains the discrepancy between the wealth mobility contribution outcomes for the cross-sectional and aggregate method? To understand this, recall that the group of entrant households I_t^d for any wealth decile between decile 11-20 to decile 81-90 consists of two groups: (i) households displaying upward mobility from lower wealth deciles ('upward mobility entrants'), and (ii) households experiencing downward mobility from higher wealth deciles ('downward mobility entrants'). In expectation, upward mobility entrants display relatively high saving rates compared to immobile households and downward mobility entrants, whereas downward mobility entrants experience relatively low saving rates. Moreover, as downward mobility entrants originate from higher wealth deciles, their initial wealth levels are higher than those of immobile households S_t^d and upward mobility entrants.

How does this allow to explain the discrepancy in wealth mobility contribution outcomes? On the one hand, the cross-sectional method takes the median of the cross-sectional distribution of saving rates of households in P_t^d. Every household $-$ regardless of whether it is immobile, an upward mobility entrant or a downward mobility entrant $-$ implicitly receives equal weight. In contrast, the aggregate method estimates saving rates using aggregated variables, and implicitly attaches more weight to households with higher initial wealth. Downward mobility

Table 2: Contribution of wealth (rank) mobility to total saving rate out of wealth according to the narrow approach.

Wealth Bin	Cross-Sectional Method				Aggregate Method			
	Baseline	Counterfactual	Contribution		Baseline	Counterfactual	Contribution	
1–10	-	-	-	-	-	-	-	-
11–20	-	-	-	-	-	-	-	-
21–30	-	-	-	-	-	-	-	-
31–40	-	-	-	-	-	-	-	-
41–50	0.11	0.07	0.04	40%	0.00	0.08	-0.08	-1229%
51–60	0.10	0.05	0.06	58%	0.03	0.06	-0.03	-71%
61–70	0.09	0.05	0.05	48%	0.03	0.06	-0.03	-108%
71–80	0.08	0.05	0.03	39%	0.03	0.06	-0.03	-73%
81–90	0.09	0.05	0.04	44%	0.06	0.06	-0.00	-0%
91–100	0.10	0.05	0.05	47%	0.11	0.06	0.04	40%

Note: this table reports the total saving rate out of wealth, estimated using the cross-sectional method (left panel) and the aggregate method (right panel). For both methods, I apply the narrow approach to quantify the contribution of endogenous wealth (rank) mobility, as defined in Equation 18 for the cross-sectional method and Equation 21 for the aggregate method. The 'baseline' column shows the observed saving rate; the 'counterfactual' column shows the rate under a hypothetical scenario without wealth mobility. The 'contribution' columns report (1) the level difference between the baseline and counterfactual, and (2) the percentage deviation from the baseline. These indicate the magnitude and direction of the mobility effect. The analysis is based on the 2001–2021 sample. Edge cases are dealt with as specified in Section 4.2. From decile 1-10 to decile 31-40, saving rates are generally ill-defined. Therefore, the values of the narrow approach are not reported.

entrants are thus given more weight in the estimation compared to immobile households and upward mobility entrants. As downward mobility entrants have lower saving rates in expectation, the contribution of wealth mobility to saving rate estimates in the aggregate method is strongly negative, with the exception of the top 10% (where upward mobility entrants make up all entrants).

Finally, adding to this argument, in Section 5.1 I found that the aggregate method generates lower saving rate estimates across the entire wealth (rank) distribution compared to the cross-sectional method, except at the top 10%. However, without wealth mobility, the total saving rate out of wealth estimates would be almost identical across the cross-sectional method and aggregate method (Table 2, counterfactual columns). In other words, the underestimation of the aggregate method compared to the cross-sectional method until wealth decile 81-90 relates for the most part to these methods' distinct treatment of entrant households' saving rates, and therefore to diverging contributions of wealth (rank) mobility to saving rate estimates.

5.4 Bias of the synthetic method

In the absence of household-level saving rate data or panel data, researchers have resorted to the synthetic method to estimate saving rates across the wealth (rank) distribution (e.g. Saez

Table 3: Bias of the synthetic method versus the unbiased aggregate method for the total saving rates out of labor income and out of wealth.

Wealth Bin	Saving Rate out of Labor Income				Saving Rate out of Wealth			
	Baseline	Counterfactual	Bias		Baseline	Counterfactual	Bias	
1–10	-	-	-	-	-	-	-	-
11–20	-	-	-	-	-	-	-	-
21–30	-	-	-	-	-	-	-	-
31–40	-	-	-	-	-	-	-	-
41–50	-0.00	0.04	0.04	-	0.00	0.06	0.06	-
51–60	0.04	0.06	0.02	48%	0.04	0.05	0.02	50%
61–70	0.06	0.10	0.04	67%	0.03	0.05	0.02	79%
71–80	0.12	0.15	0.03	29%	0.03	0.05	0.01	43%
81–90	0.29	0.25	-0.04	-12%	0.06	0.05	-0.01	-13%
91–100	1.28	0.78	-0.50	-39%	0.11	0.05	-0.06	-53%

Note: this table compares the total saving rate out of labor income (left panel) and out of wealth (right panel), as estimated by the unbiased aggregate method (baseline) and the biased synthetic method (counterfactual). The bias introduced by the synthetic method is calculated in two ways: (1) the level difference between the baseline and counterfactual estimates, and (2) the percentage deviation from the baseline. These calculations follow Equation 21. The analysis uses the 2001–2021 PSID sample. Edge cases are dealt with as specified in Section 4.2. From decile 1-10 to decile 31-40, saving rates are generally ill-defined. Therefore, the bias estimates for these lower deciles are not reported.

& Zucman, 2016; Bauluz & Meyer, 2024). As argued in Section 3.3, the synthetic method is a special type of aggregate estimation method. However, as it traces a different set of households over time, the synthetic method is biased. In what follows, I quantify this bias based on the PSID-sample for the saving rate out of labor income and saving rate out of wealth. The outcomes for the saving rates out of new resources and composite resources are shown in Appendix D. Two findings persist.

First, the synthetic method overestimates total saving rates in the middle of the wealth distribution (wealth decile 41-50 to wealth decile 71-80), and underestimates them at the top 20%. As a result, the percentile threshold for the transition between underestimation and overestimation lies around the 80th wealth percentile. Second, the underestimation of the synthetic method is in general the most substantial at wealth decile 61-70, while the overestimation is the most significant for the top 10% wealthiest. This explains to a large extent the discrepancy between the PSID saving rate estimates for the top 10% and the estimates of Saez & Zucman (2016) that I noted in Section 5.1.

6 Active and passive saving − a decomposition

In this Section, I decompose total saving rate patterns across the wealth (rank) distribution into active saving (disposable income minus consumption expenditures) and passive saving

Figure 3: Decomposition of total saving rates into active saving rates and passive saving rates.

Note: this figure plots active saving rates and passive saving rates per wealth decile, computed according to the cross-sectional method (Equation 16). The calculations are executed for the two flow-based (saving rate out of labor income and new resources) and two stock-based saving rates (saving rate out of wealth and composite resources). The 2005-2021 sample is used. Edge cases are dealt with as specified in Section 4.2. The edge cases imply that the saving rates out of wealth and composite resources are ill-defined for the bottom 35% (wealth decile 31-40 and lower). The sum of the median active and passive saving rates may be moderately higher than the total saving rate reported in Table 1. This is because the median total saving rate household will likely be different from the median active saving rate and median passive saving rate household.

(capital gains and inter-generational transfers). Thereafter, I further decompose the passive saving across the wealth (rank) distribution into its two components. The decompositions are conducted using the cross-sectional method. It has been verified that the aggregate method produces similar findings.

6.1 Active versus passive saving

How does the composition of total saving vary across the wealth (rank) distribution? Two conclusions can be drawn.

First, the composition of total saving shifts away from active saving towards passive saving the higher a household's position in the wealth (rank) distribution (Figure 3): as households become wealthier, they become more reliant on capital gains and/or inter-generational transfers. More precisely, households' total saving between wealth percentiles 21 and 40 relies entirely on active saving. From the 41st percentile onwards, total saving turns into a mixture of active and passive saving, with the latter becoming dominant from percentile 51 onwards. For the top 10%, total saving consists for the most part of passive saving.

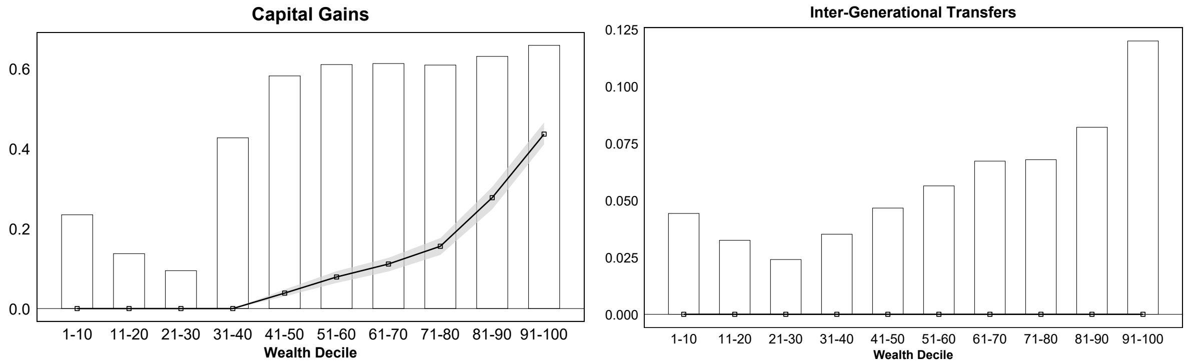
Second, as a result of these composition effects, the observed relationship between total saving rates and wealth (ranks) (in Section 5.1) does not translate directly to active saving rates and passive saving rates. On the one hand, active flow-based saving rates are rising with wealth (ranks), but the increase is significantly less pronounced than for total saving rates. In addition, in contrast to total saving rates, active stock-based saving rates are strongly declining with wealth (ranks). On the other hand, passive flow-based and stock-based saving rates are both strongly increasing in wealth (ranks).

6.2 Composition of passive saving

The previous section demonstrated that the composition of total saving shifts toward passive saving as households become wealthier. At the top 10%, total saving consists for the most part of passive saving. However, does such passive saving consist primarily of capital gains or inter-generational transfers? I analyse this based on Figure 4, which reports the passive saving rate out of new resources for capital gains and net inter-generational transfers separately. The results for the other saving rates (out of labor income, wealth and composite resources) yield similar conclusions.

Passive saving of the wealthier households consists predominantly of capital gains rather than inter-generational transfers. On the one hand, the fraction of households with positive capital gains rises from around 40% at wealth decile 31-40 to a little over 60% at wealth decile 51-60 (extensive margin). These levels remain stable thereafter (Figure 4, left panel). However, capital gains make up an increasing fraction of new resources for higher wealth ranks (intensive margin): the share of capital gains to new resources rises from approximately 10% at the 51st percentile to over 40% for the top 10% wealthiest (Figure 4, left panel). As a result, passive saving through capital gains rises across the wealth (rank) distribution primarily through the intensive margin. On the other hand, the fraction of households with positive inter-generational transfers rises from slightly over 0% for the lowest wealth deciles to close to 15% for the highest ones (Figure 4, right panel). This low extensive margin makes that the unconditional median passive saving rate from inter-generational transfers (relative to new resources) equals zero across the entire wealth (rank) distribution (Figure 4, right panel). The median conditional on transfer receipts is roughly flat across the wealth (rank) distribution for the flow-based saving rates and declining for the stock-based ones (Figure 11, Appendix E). Passive saving through

Figure 4: Decomposition of passive saving into capital gains and inter-generational transfers for the saving rate out of new resources.



Note: this figure plots (1) the fraction of households with positive passive saving per wealth decile (as bars), and (2) passive saving rates out of new resources (as lines). The latter are computed based on the cross-sectional method (Equation 16). The 95% confidence intervals have been determined using bootstrapping. Unlike in Figure 3, passive saving rates are plotted separately for capital gains (left-hand side) and inter-generational transfers (right-hand side). The 2001-2021 sample is used. Edge cases are dealt with as specified in Section 4.2. The edge cases imply that the saving rates out of wealth and composite resources are ill-defined for the bottom 35% (wealth decile 31-40 and lower).

inter-generational transfers thus rises across the wealth (rank) distribution uniquely through the extensive margin.

To conclude, wealthier households save for the most part by holding appreciating assets. Active saving out of (non capital gains) income is of much less importance. Passive saving out of inter-generational transfers is more common for wealthier households, but also relatively unimportant in magnitude[8]. These results are in line with existing findings for the Nordic countries: both Bach et al. (2018) as Fagereng et al. (2021) find that the reliance on saving out of capital gains increases strongly with wealth (ranks).

7 Conclusion

While there exists empirical evidence on saving behavior across the wealth (rank) distribution in Nordic countries, such evidence is largely absent for the United States. This paper uses household-level data from the Panel Study of Income Dynamics (PSID) to fill this gap. I obtain four collections of stylized empirical facts. First, I find that total saving rates out of labor income and new resources rise with wealth ranks (flow-based saving rates). In contrast, total saving rates out of wealth and composite resources are roughly stable and only modestly increasing with wealth ranks (stock-based saving rates). Second, wealth (rank) mobility has

[8]Passive saving out of inter-generational transfers has been computed for a given time period t. A proper analysis on the importance of inter-generational transfers in wealth accumulation would require a cumulative metric. I leave this question to future research.

a substantial impact on total saving rate patterns across the wealth distribution. However, while the contribution of wealth mobility is strictly positive for the cross-sectional method, it is negative across most of the wealth distribution for the aggregate method. I show that this discrepancy relates to these methods' distinct treatment of wealth (rank) mobility: while the cross-sectional method attaches equal weight to all households in a wealth decile, the aggregate method overweighs households that display downward wealth mobility. Third, I find that the synthetic method overestimates saving rates up to the 80th percentile, while it underestimates the saving rates of the top 20%. Fourth, I demonstrate that households' reliance on capital gains rises across the wealth rank distribution: the top wealthiest households' total saving consists predominantly of saving by holding appreciating assets. Passive saving out of inter-generational transfers is more common for wealthier households, but relatively unimportant in magnitude. Many of the empirical saving behavior moments across the wealth (rank) distribution reported in this paper are likely to be of interest to the heterogeneous agent literature replicating the U.S. wealth distribution.

8 References

Andreski, P., Li, G., Samancioglu, M.Z., Schoeni., R. (2014). Estimates of Annual Consumption Expenditures and Its Major Components in the PSID in Comparison to the CE. American Economic Review 104 (5), p. 132-135.

Bach, L., Calvet, L., Sodini, P. (2018). From Saving comes Having? Disentangling the Impact of Saving on Wealth Inequality. Swedish House of Finance Research Paper (June 2018).

Bauluz, L., Meyer, T. (2024). The Wealth of Generations. Working Paper, CUNEF Universidad.

De Nardi, M., Fella, G. (2017). Saving and Wealth Inequality. Review of Economic Dynamics 26, p. 280-300.

Fagereng, A., Holm, M.B., Moll, B., Natvik, G. (2021). Saving Behavior Across the Wealth Distribution: The Importance of Capital Gains. National Bureau of Economic Research w26588.

Fernández-Villaverde, J., Levintal, O. (2024). The Distributional Effects of Asset Returns. NBER Working Paper 32182.

Gaillard, A., Wangner, S. (2023). Wealth, Returns, and Taxation: A Tale of Two Dependencies. Unpublished.

Gomez, M. (2023). Decomposing the Growth of Top Wealth Shares. Econometrica 91 (3), p. 979-1024.

Hubmer, J., Krusell, P., Smith, A. (2021). Sources of U.S. Wealth Inequality: Past, Present and Future. NBER Macroeconomics Annual 35 (1), p. 391-455.

Kimberlin, S., Kim, J. and Shaefer, L. (2015). An Updated Method for Calculating Income and Payroll Taxes from PSID data using the NBER's TAXSIM, for PSID Survey Years 1999 through 2011. University of Michigan Manuscript.

Pfeffer, F., Schoeni, R., Kennickel, A., Andreski, P. (2016). Measuring Wealth and Wealth Inequality: Comparing Two U.S. Surveys. Journal of Economic and Social Measurement 41, p. 103-120.

Saez, E., Zucman, G. (2016). Wealth inequality in the United States since 1913: Evidence from capitalized income tax data. Quarterly Journal of Economics 131 (2), p. 519-578.

Straub, L. (2019). Consumption, Savings, and The Distribution of Permanent Income. Unpublished Manuscript, Harvard University.

Van Langenhove, C. (2025a). Wealth Mobility in the United States: Empirical Evidence from the PSID. Working Papers of Faculty of Economics and Business Administration, WP 25/1104.

Van Langenhove, C. (2025b). Wealth Inequality and Wealth Mobility in the United States: Type Dependence versus Scale Dependence. Forthcoming.

A Budget constraint variables in the PSID

As argued in the main text, computing active saving and passive saving (in the numerator of the ratios), or new resources and composite resources (in the denominator of the ratios) requires panel data on all budget constraint variables. In this subsection, I outline how I compute these budget constraint variables using the Panel Study of Income Dynamics (PSID). I denote PSID-estimated variables with a tilde, e.g. $\tilde{x}$. The PSID includes wealth-related questions only in the 1984, 1989, 1994 waves and bi-annually from 1999 onwards. I therefore restrict the PSID data to the period 1999-2021. For future reference, denote a survey wave as s, where:

$$s \in \{1999, 2001, ..., 2019, 2021\} \tag{32}$$

Two remarks are in place. First, responding families in the PSID are occasionally unaware of the exact value of their income or wealth. In that case, for some years and variables, bracketing questions are provided. I apply the bracketing procedure as discussed in Van Langenhove (2025a). The results of this paper are robust to whether or not this bracketing procedure is implemented. Second, for some variables, we are interested in the first difference. For example, the computation of total saving requires taking the first difference of total wealth (Δw_{t+1}). A difficulty is that PSID survey waves are conducted only bi-annually, which means that the wealth in between two waves needs to be interpolated. I do so by taking the midpoint between two data points (as in e.g. Gaillard & Wangner, 2023). For example, for wealth w:

$$\tilde{w}_t = \begin{cases} \frac{\tilde{w}_{s-1} + \tilde{w}_s}{2} & \text{if } s-1 < t < s \\ \tilde{w}_s & \text{if } t = s \end{cases} \tag{33}$$

As it determines total saving, the most critical budget constraint variable is wealth $\tilde{w}$. It is defined as households' total assets minus its total liabilities. Assets include gross main housing, business holdings, equity holdings, fixed-income holdings, pension wealth and gross other housing. Pension wealth includes defined contribution (DC) plans, IRAs and private annuities. Liabilities comprise main mortgages outstanding, other housing debt and non-mortgage debt. I follow the definitions and harmonization procedures outlined in Appendix A of Van Langenhove (2025a).

A.1 Income variables

Labor income labor income $\tilde{y}$ is readily available from the PSID questionnaire. It is reported separately for the household reference person and spouse. From 2005 onwards, data is available also for other working household members. I compute household labor income by taking the sum across all individuals reported in the household.

Capital income For the computation of the capital income return $\tilde{r}^i$, I largely follow Gaillard & Wangner (2023). This capital income return can be computed straightforwardly from PSID questions: the PSID questionnaire asks households about different sources of capital income. Outcomes are reported separately for the reference person and the spouse. I therefore compute household-level variables by taking the sum over the two individuals. The capital income categories include farm income, business income, rental income, dividend income, interest income and trust and royalty income, which can be linked to their corresponding asset categories.

Three notes are in place. First, unlike Gaillard & Wangner (2023), I attribute farm and business income entirely to capital income. The results in this paper are robust to this assumption. Second, rental income is reported only when the household has a secondary house which it rents out. I do not impute rental income on the main house or impute rental income on occupied secondary housing as these do not reflect actual financial flows. Third, contrary to Gaillard & Wangner (2023), I attribute interest income entirely to fixed income assets. However, as this paper uses only a composite capital income variable as input in the calculation of the saving rates, this assumption does not affect results.

Capital gains Also for the capital gains return $\tilde{r}^g$, I largely follow Gaillard & Wangner (2023). Households are asked about the total value of and total inflows and outflows for most assets. In principle, I define as capital gains the change in the asset's total value that cannot be explained by net inflows. I account for the bi-annual nature of survey waves by dividing capital gains between two survey waves by two, in line with Equation 33. Capital gains are then computed by taking the sum across all wealth categories. There are a number of exceptions and particular computational choices to this general procedure, however. I discuss these choices next.

First, for farm and business holdings, the reported values are net of debt up until 2011. Instead, from 2013 onwards, assets and debts are reported separately. I compute capital gains from the net value throughout the entire period. This is different from Gaillard & Wangner (2023), who instead trace out a gross business holdings series prior to 2011 by assuming that changes in business debt affect business assets one-to-one.

Second, housing consists of main housing and other housing. On the one hand, for main housing, I compute unrealized capital gains as the change in the reported house value between the two survey waves, net of depreciation. Moreover, I calculate realized capital gains as the difference between the housing selling price and the reported housing value in the previous period, again net of depreciation. On the other hand, for other housing, capital gains are computed as the difference in reported net other housing values between two survey waves, corrected for net inflows. Total capital gains on housing are then obtained by summing all three components and subtracting housing improvement expenditures. For both main and other housing, I assume a depreciation rate of 2%, in line with Gaillard & Wangner (2023).

Third, the pension wealth category includes both DC pension plans and IRA (and 401k accounts, Keogh accounts or similar) accounts. However, net inflows are reported only for IRAs. For DC pension plans, I therefore assume that net inflows equal the sum of the employee and employer contributions. These are reported for both the reference person and spouse, and both for current and previous employers. I take the sum across all.

A.2 Government and inter-generational transfers

Government transfer and social security income The PSID contains very detailed questions on government transfer and social security income. These capture all possible government transfer and social security income categories. Prior to 2005, social security income is reported at the family level. From 2005 onwards, it is instead reported separately for the reference person, spouse and other family members. I then take the sum across all three individuals to obtain a household-level variable. Transfer income is reported separately for all three individual types throughout the entire 1999-2021 period.

Inter-generational transfers & lumpsum payments The variable $\tilde{m}$ contains primarily inter-generational transfer receipts. Households are asked to report their received gifts or inheritances since the previous survey wave. Moreover, households are asked to report the amount of financial help they have received from relatives, and the amount of financial help they have provided to others. I assume that inter-generational transfers equal the sum of received gifts and inheritances and net help received from others. Moreover, I augment $\tilde{m}$ with lumpsum payment receipts that cannot be attributed to inheritances. These mainly include payouts from insurance and lottery winnings. However, these lumpsum receipts make up a limited fraction of $\tilde{m}$.

A.3 Consumption expenditures & taxes

Consumption The PSID asks households detailed questions about their consumption expenditures from 1999 onwards. However, for the 1999-2003 waves, PSID questions capture only around 70% of the expenditures from the Consumption Expenditure Survey (CEX) and National Income and Product Accounts (NIPA). From 2005 onwards, it captures almost all categories from the CEX (Andreski et al., 2014). A reliable consumption estimate $\tilde{c}$ in the PSID is therefore available only from 2005 onwards.

Four remarks are in place. First, I include interest payments on first mortgages, second mortgages and consumer debt in the consumption expenditures. While interest payments on first mortgages are reported in the PSID questionnaires, those on second mortgages and consumer debt are not. I therefore estimate these by respectively taking the fixed 30-year mortgage interest rate and personal loan interest rate from the FRED and multiplying these with the household's previous-period debt stock. Second, mortgage principal payments however are excluded from consumption: I consider these as a form of saving as they generate higher wealth (i.e. higher net worth). Third, rental payments of non-house-owning households are included as part of consumption expenditures.

Taxes The PSID does not provide any tax payment data beyond property taxes. Taxes ($\tilde{\tau}$) therefore needs to be estimated. I use the following four-step strategy. First, to estimate income, payroll and capital gains taxes, I use the NBER tax simulator. The details of this procedure are described below and are similar to Kimberlin et al. (2015). Second, I collect tax rates for estates throughout the years and apply these to the inter-generational transfers variable. However, due to the high exemptions on estate taxes, only a handful of observations in the sample are in fact affected by this estate taxation. Third, property taxes and motor vehicle taxes are reported by the household in the PSID. Fourth, I sum the estimated payroll, income and capital gains taxes (NBER simulator), estate taxes and reported property and motor vehicle taxes to obtain a total household-level tax estimate $\tilde{\tau}$.

NBER tax simulator I use NBER tax simulator version v35. The estimation of payroll, income and capital gains taxes requires providing a large set of demography- and income-related input variables to the simulator. All required variables can be computed from the PSID. A number of remarks are in place, however. First, the NBER tax simulator expects a value for short-term and long-term realized capital gains. I have no clear-cut method to distinguish between the short versus long term nature of capital gains, or between realized and unrealized (apart from main housing) gains. I therefore assume that realized capital gains equal 20% of total capital gains, and attribute these entirely to long-term gains. Moreover, I take into account the exemption threshold on housing capital gains over the years. Second, the PSID questionnaire reports household tax deductions on charitable contributions, childcare and medical expenses. I therefore do not include these in the NBER tax simulator and use the PSID-reported value instead.

A.4 Household composition changes

A unique PSID family unit consists of a reference person and possibly a partner. The data structures guarantees that saving rates can be computed for a wave s only when the family unit had the same reference person and partner over wave s and previous wave $s-1$. In principle, $\tilde{\eta}$ therefore equals zero in our data. However, individuals different from the reference person or partner (most commonly children, siblings or elderly) may enter or exit the family unit, and bring assets or debts with them. In case of such event, these new assets and debts are reported in the PSID questionnaire. I define $\tilde{\eta}$ to their net value.

B Robustness: total saving behavior across the wealth (rank) distribution (Sections 5.1 and 5.2)

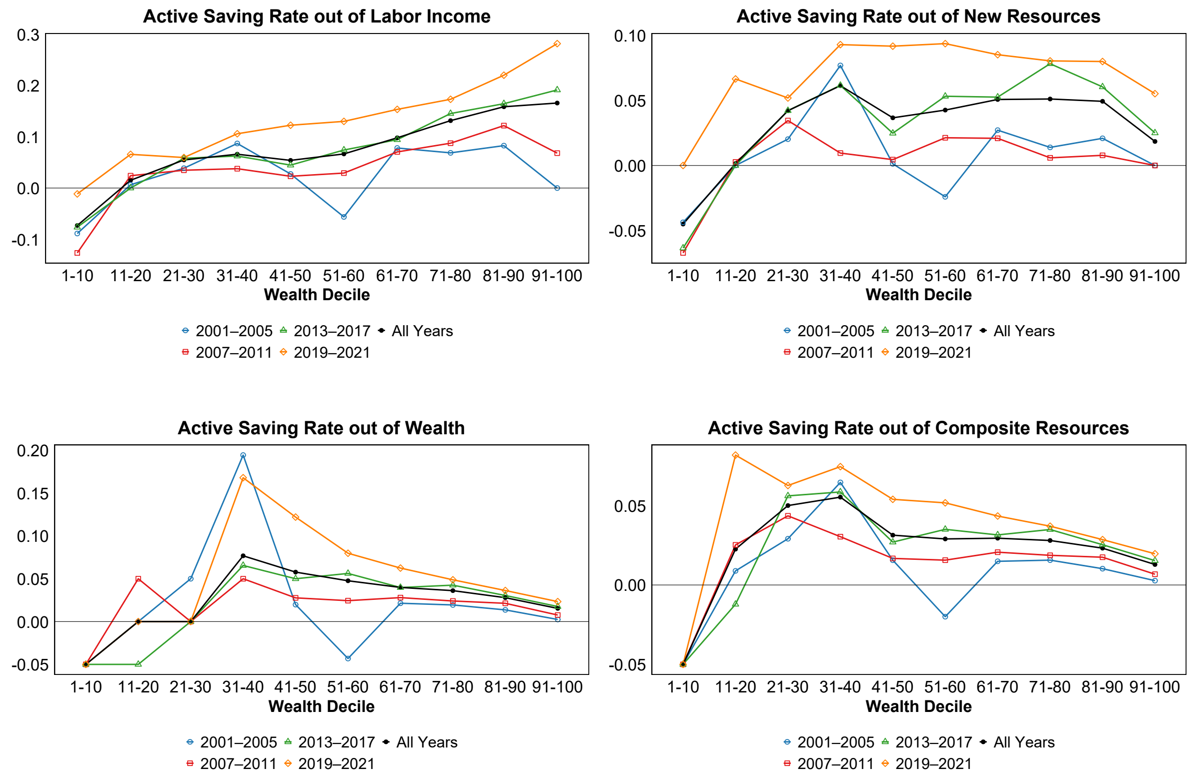
B.1 Saving rates over time

In this subsection, I plot active and passive saving rates across the wealth (rank) distribution for different year groups. In the main text, I demonstrated that total saving rates were higher across the wealth (rank) distribution for the 2019-2021 period, and lower for the 2007-2011 period. Here, I plot active saving rates and passive saving rates (Figures 5 and 6). I find that the high total saving across the wealth (rank) distribution for the 2019-2021 period followed from both higher active and passive saving, whereas the low total saving for the 2007-2011 related primarily to lower passive saving.

B.2 Robustness checks

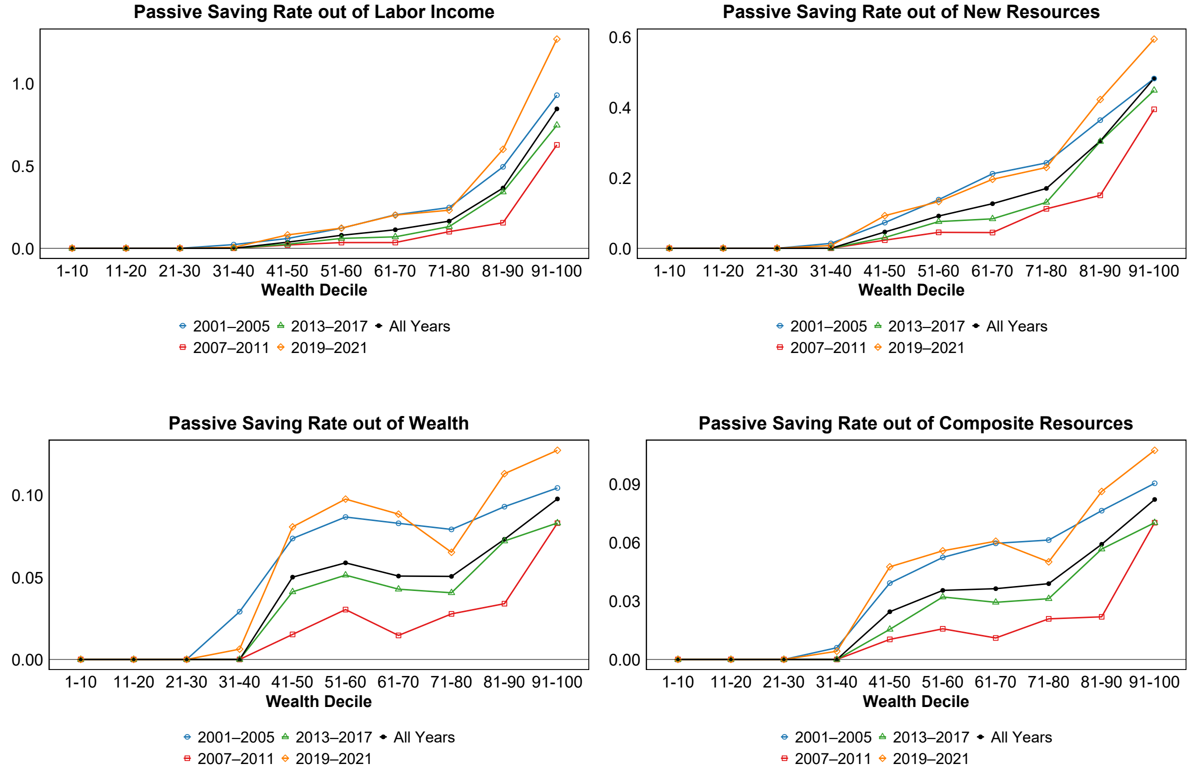
In this subsection, I present the figures for the robustness checks to the relationship between total saving rates and wealth (ranks), as discussed in Section 5.1 of the main text. I discuss four robustness checks. First, I summarize the saving rates per wealth decile d using the mean instead of the median (Figure 7). Second, I plot the relationship between total saving rates and wealth (ranks) across different labor income groups (Figure 8). Third, I do the same for different age groups (Figure 9). Fourth, I compare total saving rates across the wealth (rank) distribution for non-entrepreneurial and entrepreneurial households (Figure 10).

Figure 5: Cross-sectional method– active flow-based and stock-based saving rates across wealth deciles over different survey waves.



Note: this figure plots the active saving rates per wealth decile, grouped across different PSID survey waves. The saving rates per wealth decile have been computed using the cross-sectional method (Equation 16). The waves are pooled as: $\{2001, 2003, 2005\}$, $\{2007, 2009, 2011\}$, $\{2013, 2015, 2017\}$ and $\{2019, 2021\}$. The first wave only contains the year 2005 as the active saving variable is defined only from 2005 in the PSID. The black line shows the baseline results without grouping. The calculations are executed for the two flow-based (active saving rates out of labor income and new resources) and two stock-based saving rates (active saving rates out of wealth and composite resources). The 2001-2021 sample is used. Edge cases are dealt with as specified in Section 4.2. The edge cases imply that the saving rates out of wealth and composite resources are ill-defined for the bottom 35% (wealth decile 31-40 and lower).

Figure 6: Cross-sectional method − passive flow-based and stock-based saving rates across wealth deciles over different survey waves.

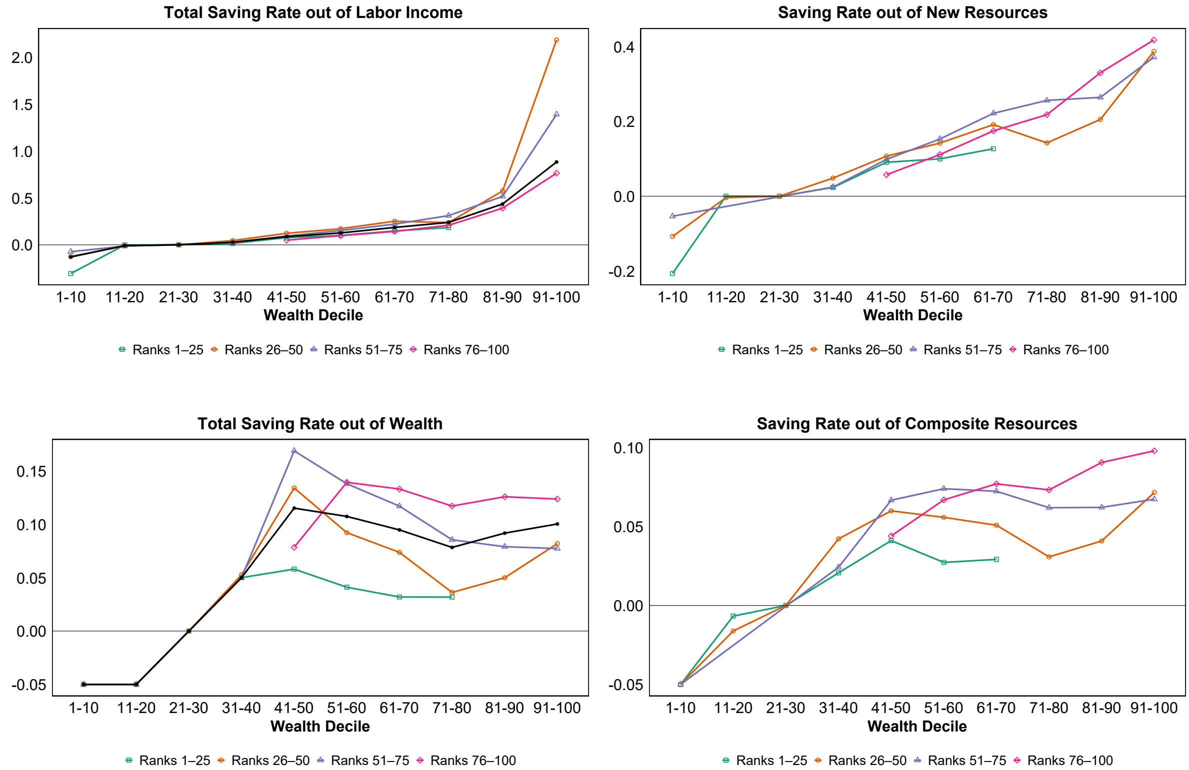


Note: this figure plots the passive saving rates per wealth decile, grouped across different PSID survey waves. The saving rates per wealth decile have been computed using the cross-sectional method (Equation 16). The waves are pooled as: $\{2001, 2003, 2005\}$, $\{2007, 2009, 2011\}$, $\{2013, 2015, 2017\}$ and $\{2019, 2021\}$. The black line shows the baseline results without grouping. The calculations are executed for the two flow-based (passive saving rates out of labor income and new resources) and two stock-based saving rates (passive saving rates out of wealth and composite resources). The 2001-2021 sample is used. Edge cases are dealt with as specified in Section 4.2. The edge cases imply that the saving rates out of wealth and composite resources are ill-defined for the bottom 35% (wealth decile 31-40 and lower).

Figure 7: Cross-sectional method – total flow-based and stock-based saving rates across wealth deciles: mean (robustness, red line) versus median (baseline, black line) as summary metric.

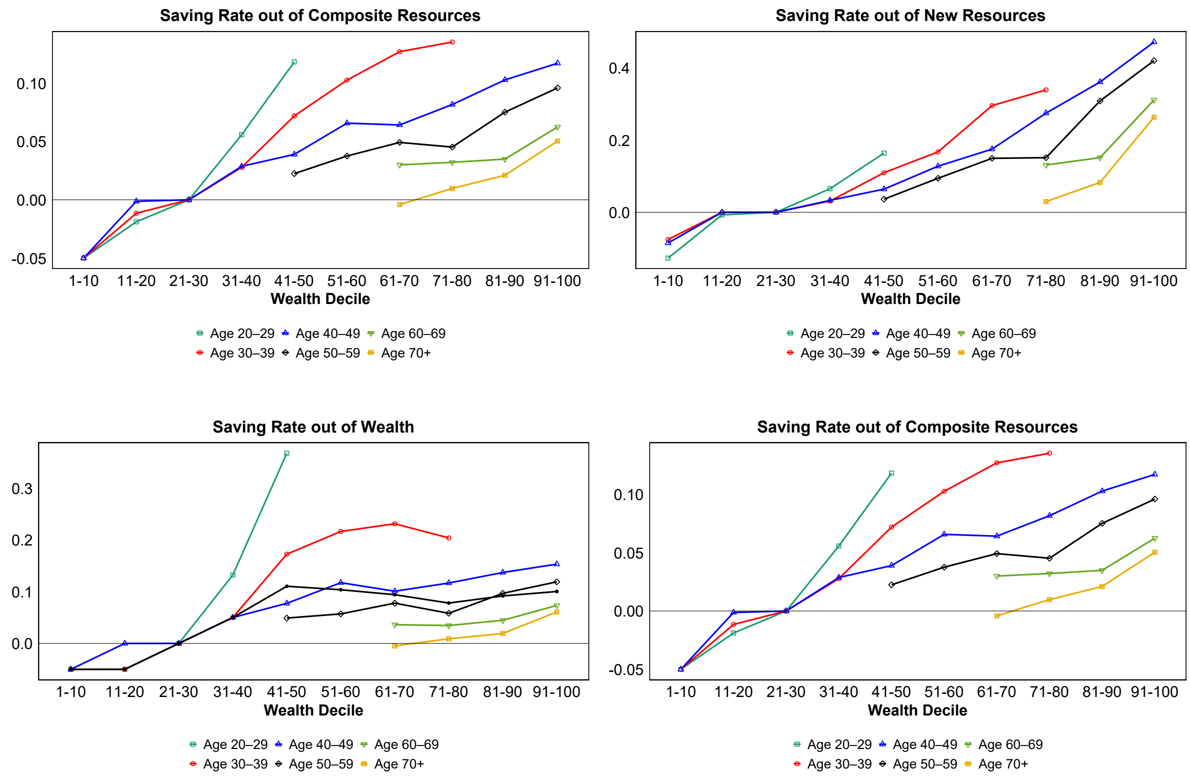
Note: this figure plots the (1) the fraction of households with positive total savings per wealth decile (as bars), and (2) mean and median total saving rates per wealth decile, computed according to the cross-sectional method of Equation 16 (as lines). The median (baseline) is plotted as a black line, the mean (robustness) as a red one. The calculations are executed for the two flow-based (total saving rate out of labor income and new resources) and two stock-based saving rates (total saving rate out of wealth and composite resources). The 2001-2021 sample is used. Edge cases are dealt with as specified in Section 4.2. The edge cases imply that the saving rates out of wealth and composite resources are ill-defined for the bottom 35% (wealth decile 31-40 and lower).

Figure 8: Cross-sectional method – total flow-based and stock-based saving rates across wealth deciles for different labor income (rank) groups.



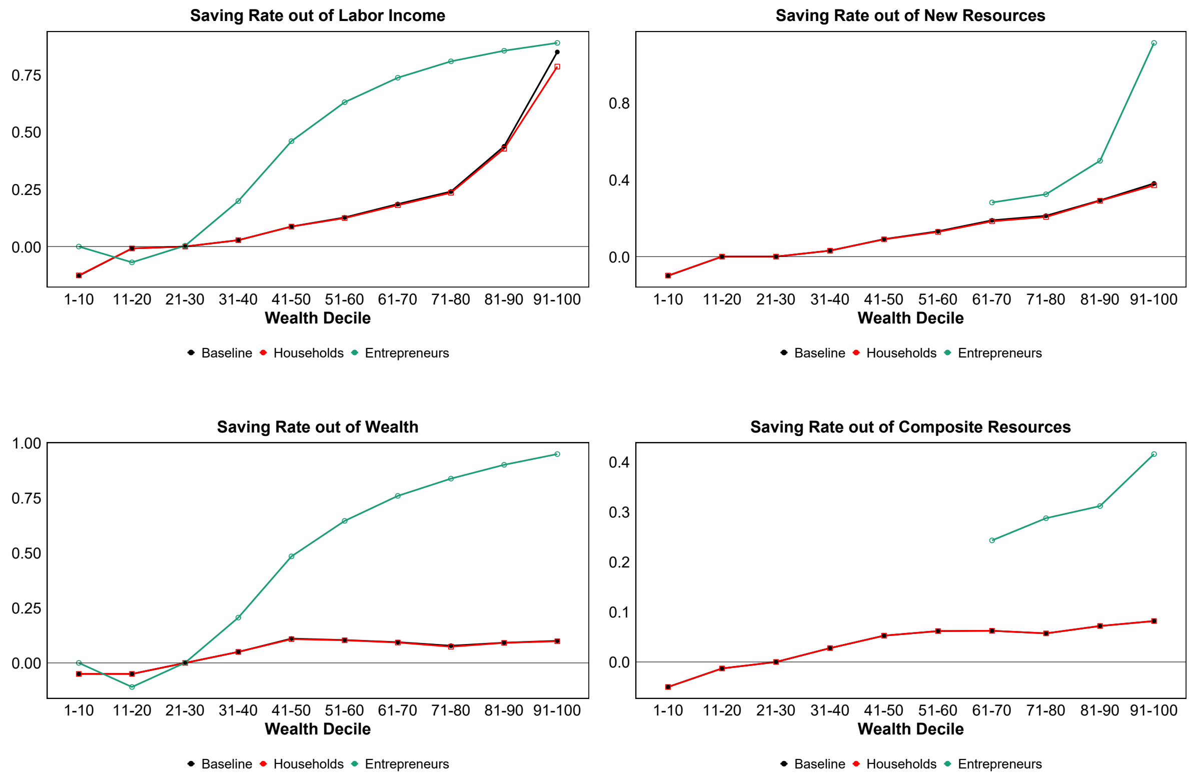
Note: this figure plots total saving rates per wealth decile across different labor income groups. The saving rates per decile have been computed using the cross-sectional method (Equation 16). Households are allocated to a labor income group based on their rank in the labor income distribution. Labor income also includes government transfers (including social security income). The calculations are executed for the two flow-based (total saving rate out of labor income and new resources) and two stock-based saving rates (total saving rate out of wealth and composite resources). The 2001-2021 sample is used. Edge cases are dealt with as specified in Section 4.2. The edge cases imply that the saving rates out of wealth and composite resources are ill-defined for the bottom 35% (wealth decile 31-40 and lower). Values are plotted only when the labor income group–wealth decile combination has a minimum of 250 observations.

Figure 9: Cross-sectional method – total flow-based and stock-based saving rates across wealth deciles for different age groups.



Note: this figure plots total saving rates per wealth decile across different age groups. The saving rates per decile have been computed using the cross-sectional method (Equation 16). Households are allocated to a labor income group based on the age of the reference person of the PSID-household. The calculations are executed for the two flow-based (total saving rate out of labor income and new resources) and two stock-based saving rates (total saving rate out of wealth and composite resources). The 2001-2021 sample is used. Edge cases are dealt with as specified in Section 4.2. The edge cases imply that the saving rates out of wealth and composite resources are ill-defined for the bottom 35% (wealth decile 31-40 and lower). Values are plotted only when the age group–wealth decile combination has a minimum of 250 observations.

Figure 10: Cross-sectional method – total flow-based and stock-based saving rates across wealth deciles: non-entrepreneurial versus entrepreneurial households.



Note: this figure plots the median total saving rates per wealth decile (computed according to the cross-sectional method in Equation 16) for non-entrepreneurial households ('households', in red) and entrepreneurial households ('entrepreneurs', in green) separately. The baseline over all households is added for comparison purposes (in black). Households are designated as entrepreneurs when they report business ownership in the PSID-questionnaires. The calculations are executed for the two flow-based (total saving rate out of labor income and new resources) and two stock-based saving rates (total saving rate out of wealth and composite resources). The 2001-2021 sample is used. Edge cases are dealt with as specified in Section 4.2. The edge cases imply that the saving rates out of wealth and composite resources are ill-defined for the bottom 35% (wealth decile 31-40 and lower).

C Contribution of wealth (rank) mobility (Section 5.3)

In this Appendix, I quantify the contribution of wealth (rank) mobility to observed saving rate patterns across the wealth (rank) distribution for the total saving rates out of labor income, new resources and composite resources.

Table 4: Contribution of wealth (rank) mobility to total saving rate out of labor income according to the narrow approach.

Wealth Bin	**Cross-Sectional Method**				**Aggregate Method**			
	Baseline	Counterfactual	Contribution		Baseline	Counterfactual	Contribution	
1–10	-	-	-	-	-	-	-	-
11–20	-	-	-	-	-	-	-	-
21–30	-	-	-	-	-	-	-	-
31–40	-	-	-	-	-	-	-	-
41–50	0.09	0.04	0.05	53%	-0.00	0.05	-0.05	-1223%
51–60	0.13	0.06	0.07	53%	0.04	0.07	-0.03	-88%
61–70	0.19	0.10	0.08	44%	0.06	0.12	-0.06	-111%
71–80	0.24	0.15	0.09	37%	0.12	0.19	-0.07	-60%
81–90	0.44	0.26	0.18	41%	0.29	0.31	-0.02	-7%
91–100	0.85	0.50	0.35	41%	1.28	0.98	0.30	23%

Note: this table reports the total saving rate out of labor income, estimated using the cross-sectional method (left panel) and the aggregate method (right panel). For both methods, I apply the narrow approach to quantify the contribution of endogenous wealth (rank) mobility, as defined in Equation 18 for the cross-sectional method and Equation 21 for the aggregate method. The 'baseline' column shows the observed saving rate, the 'counterfactual' column shows the rate under a hypothetical scenario without wealth mobility. The 'contribution' columns report (1) the level difference between the baseline and counterfactual, and (2) the percentage deviation from the baseline. These indicate the magnitude and direction of the mobility effect. The analysis is based on the 2001–2021 sample. Edge cases are dealt with as specified in Section 4.2. From decile 1-10 to decile 31-40, saving rates are generally ill-defined. Therefore, the values of the narrow approach are not reported.

Table 5: Contribution of wealth (rank) mobility to total saving rate out of new resources according to the narrow approach.

Wealth Bin	Cross-Sectional Method				Aggregate Method			
	Baseline	Counterfactual	Contribution		Baseline	Counterfactual	Contribution	
1–10	-	-	-	-	-	-	-	-
11–20	-	-	-	-	-	-	-	-
21–30	-	-	-	-	-	-	-	-
31–40	-	-	-	-	-	-	-	-
41–50	0.09	0.05	0.04	49%	-0.01	0.05	-0.06	-1032%
51–60	0.13	0.06	0.07	51%	0.04	0.08	-0.04	-90%
61–70	0.19	0.12	0.07	38%	0.06	0.12	-0.06	-95%
71–80	0.21	0.14	0.07	33%	0.11	0.17	-0.06	-58%
81–90	0.29	0.18	0.12	39%	0.23	0.25	-0.03	-11%
91–100	0.38	0.24	0.14	37%	0.53	0.40	0.13	25%

Note: this table reports the total saving rate out of new resources, estimated using the cross-sectional method (left panel) and the aggregate method (right panel). For both methods, I apply the narrow approach to quantify the contribution of endogenous wealth (rank) mobility, as defined in Equation 18 for the cross-sectional method and Equation 21 for the aggregate method. The 'baseline' column shows the observed saving rate, the 'counterfactual' column shows the rate under a hypothetical scenario without wealth mobility. The 'contribution' columns report (1) the level difference between the baseline and counterfactual, and (2) the percentage deviation from the baseline. These indicate the magnitude and direction of the mobility effect. The analysis is based on the 2001–2021 sample. Edge cases are dealt with as specified in Section 4.2. From decile 1-10 to decile 31-40, saving rates are generally ill-defined. Therefore, the values of the narrow approach are not reported.

Table 6: Contribution of wealth (rank) mobility to total saving rate out of composite resources according to the narrow approach.

Wealth Bin	Cross-Sectional Method				Aggregate Method			
	Baseline	Counterfactual	Contribution		Baseline	Counterfactual	Contribution	
1–10	-	-	-	-	-	-	-	-
11–20	-	-	-	-	-	-	-	-
21–30	-	-	-	-	-	-	-	-
31–40	-	-	-	-	-	-	-	-
41–50	0.05	0.03	0.03	49%	-0.01	0.03	-0.04	-541%
51–60	0.06	0.03	0.03	57%	0.01	0.03	-0.02	-107%
61–70	0.06	0.04	0.02	44%	0.02	0.03	-0.02	-128%
71–80	0.06	0.04	0.02	35%	0.02	0.04	-0.01	-78%
81–90	0.07	0.04	0.03	44%	0.04	0.04	-0.00	-8%
91–100	0.08	0.05	0.03	39%	0.08	0.05	0.03	37%

Note: this table reports the total saving rate out of composite resources, estimated using the cross-sectional method (left panel) and the aggregate method (right panel). For both methods, I apply the narrow approach to quantify the contribution of endogenous wealth (rank) mobility, as defined in Equation 18 for the cross-sectional method and Equation 21 for the aggregate method. The 'baseline' column shows the observed saving rate, the 'counterfactual' column shows the rate under a hypothetical scenario without wealth mobility. The 'contribution' columns report (1) the level difference between the baseline and counterfactual, and (2) the percentage deviation from the baseline. These indicate the magnitude and direction of the mobility effect. The analysis is based on the 2001–2021 sample. Edge cases are dealt with as specified in Section 4.2. From decile 1-10 to decile 31-40, saving rates are generally ill-defined. Therefore, the values of the narrow approach are not reported.

D Bias of the synthetic method (Section 5.4)

Table 7: Bias of the synthetic method relative to the (unbiased) aggregate method for the total saving rates out of new resources and out of composite resources.

Wealth Bin	Saving Rate out of New Resources				Saving Rate out of Composite Resources			
	Baseline	Counterfactual	Bias		Baseline	Counterfactual	Bias	
1–10	-	-	-	-	-	-	-	-
11–20	-	-	-	-	-	-	-	-
21–30	-	-	-	-	-	-	-	-
31–40	-	-	-	-	-	-	-	-
41–50	-0.01	0.03	0.04	621%	-0.01	0.02	0.02	286%
51–60	0.04	0.05	0.01	21%	0.01	0.02	0.01	38%
61–70	0.06	0.08	0.02	24%	0.02	0.02	0.01	62%
71–80	0.11	0.11	0.00	0%	0.02	0.03	0.01	28%
81–90	0.23	0.16	-0.07	-32%	0.04	0.03	-0.01	-13%
91–100	0.53	0.47	-0.07	-12%	0.08	0.04	-0.03	-45%

Note: this table compares the total saving rate out of new resources (left panel) and composite resources (right panel), as estimated by the unbiased aggregate method (baseline) and the biased synthetic method (counterfactual). The bias introduced by the synthetic method is calculated in two ways: (1) the level difference between the baseline and counterfactual estimates, and (2) the percentage deviation from the baseline. These calculations follow Equation 21. The analysis uses the 2001–2021 PSID sample. Edge cases are dealt with as specified in Section 4.2. From decile 1-10 to decile 31-40, saving rates are generally ill-defined. Therefore, the bias estimates for these lower deciles are not reported.

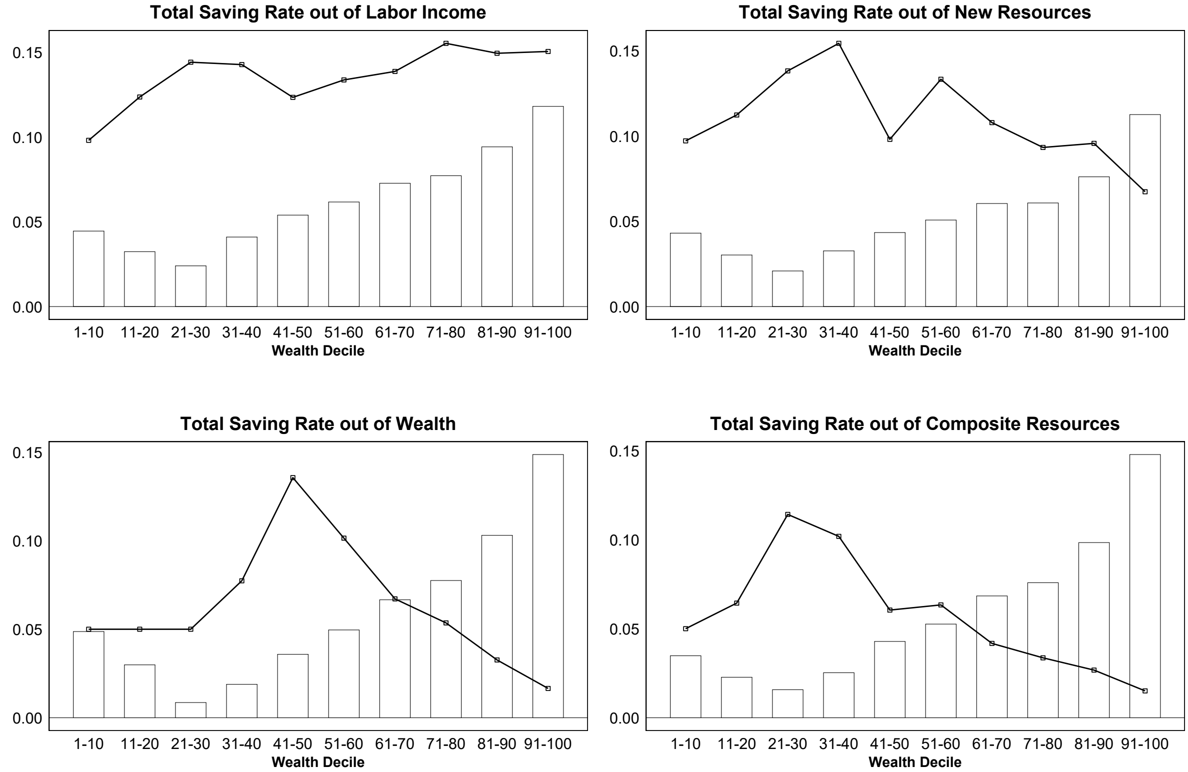
E Decomposition into active and passive saving (Section 6)

Section 6 of the main text decomposed total saving rates into active saving and passive saving. This decomposition was produced for the two flow-based and two stock-based saving rates. Thereafter, I decomposed passive saving by plotting the capital gains and inter-generational transfer saving rate out of new resources across wealth deciles, reporting their unconditional medians.

In this Appendix, I provide additional visualizations to the latter analysis (decomposition of passive saving into capital gains and inter-generational transfers). Specifically, I present inter-generational transfer saving rates using conditional rather than unconditional medians (as in the main text). This allows to assess whether there exists a relationship between inter-generational transfer saving rates and wealth (ranks), conditional on receiving such transfers (at the intensive margin).

I find that conditional inter-generational saving rates are roughly stable across the wealth (rank) distribution for the two flow-based saving rates, and declining with wealth (ranks) for the two stock-based saving rates (Figure 11). This suggests that, based on the intensive margin, the importance of inter-generational transfer saving is definitely not increasing with a household's position in the wealth (rank) distribution. However, as shown in Figure 11 (bars) and in the main text, it is rising in the extensive margin.

Figure 11: Inter-generational transfer flow-based and stock-based saving rates across wealth deciles.



Note: this figure plots (1) the fraction of households with positive inter-generational transfer savings per wealth decile (as bars), and (2) the inter-generational transfer saving rates per wealth decile (as lines). The latter have been computed using the cross-sectional method (Equation 16). Unlike for the other plots in this paper, (2) takes the median conditional on having positive inter-generational transfer savings. The calculations are executed for the two flow-based (total saving rate out of labor income and new resources) and two stock-based saving rates (total saving rate out of wealth and composite resources). The 2001-2021 sample is used. Edge cases are dealt with as specified in Section 4.2. The edge cases imply that the saving rates out of wealth and composite resources are ill-defined for the bottom 35% (wealth decile 31-40 and lower).

F Saving ratios

As noted in Section 2.3, heterogeneous agent models often incorporate the saving ratio ξ as policy variable rather than the saving rates defined in the main text. The saving ratio is defined in Equation 12 and represents the fraction of composite resources Λ^C (in the denominator) that is transferred by the household to the next time period in the form of wealth w (in the numerator). Table 8 reports the saving ratio outcomes across the wealth (rank) distribution using the inflow-based approach. The results for the consumption-based approach are roughly identical. Two key findings persist.

First, the fraction of households with positive saving ratios equals zero for the bottom 20% and approximately 0.57 for wealth decile 21-30. It then equals close to one for all higher wealth deciles. This pattern follows directly from the wealth term in the numerator of Equation 12: wealth is negative for all households in the bottom 20%, and turns positive close to the 30th wealth percentile. Second, for households with non-negative wealth, there exists a positive relationship between saving ratios and wealth (ranks): the saving ratio in the middle part of the wealth distribution equals around 0.50 and rises monotonically to 0.90 for the top 10% wealthiest households.

Table 8: Cross-sectional method − saving ratios across wealth deciles.

Wealth Decile	Fraction Positive Savers	Saving Ratio
1–10	0.00	0.00 (0.00, 0.00)
11–20	0.00	-0.07 (-0.07, -0.06)
21–30	0.57	0.00 (0.00, 0.01)
31–40	0.98	0.20 (0.19, 0.21)
41–50	0.99	0.46 (0.45, 0.47)
51–60	1.00	0.63 (0.62, 0.63)
61–70	1.00	0.73 (0.73, 0.74)
71–80	1.00	0.81 (0.80, 0.81)
81–90	1.00	0.85 (0.85, 0.86)
91–100	1.00	0.89 (0.88, 0.89)

Note: this table reports (1) the fraction of households with a strictly positive saving ratio, and (2) the median saving ratio ξ per wealth decile. Saving ratios are computed as the ratio of wealth to composite resources (Equation 12) using the cross-sectional method (Equation 16) under the inflow-based approach. Confidence intervals are reported in parentheses below the median estimates and are based on bootstrapping. The 2001–2021 sample is used. Edge cases are handled as detailed in Section 4.2. Saving ratios are ill-defined for a large share of households in the bottom wealth deciles.

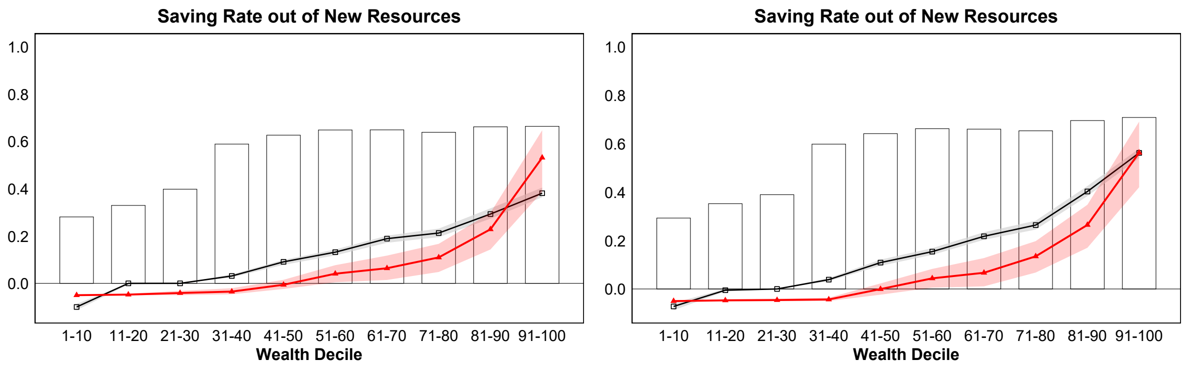
G Expenditure-based approach

Section 4.2 distinguished between an inflow-based approach and expenditure-based approach to compute new resources $\tilde{\Lambda}^N$ and composite resources $\tilde{\Lambda}^C$ from the PSID-data. The former is applied over the 2001-2021 sample, while the latter is computed over the 2005-2021 sample. In the main text, I have reported total saving rates out of new and composite resources based on the inflow-based approach. In this Appendix, I compare the outcomes of the inflow-based approach to the expenditure-based approach. I conduct the comparative exercise for total saving rates (reported in Table 1 in the main text). The results for the total saving rate out of new resources are plotted in Figure 12, while those for the total saving rate out of composite resources are displayed in Figure 13.

For both saving rates, the consumption-based approach yields slightly higher saving rates estimates than the inflow-based approach for higher wealth deciles, especially when using the cross-sectional method. The differences are highly limited, however. This minor gap may reflect two effects. First, it could relate to sample differences: the inflow-based approach uses the 2001-2021 sample, while the consumption-based approach departs from the 2005-2021 sample. However, I have checked that the gap between the two approaches remains roughly unchanged when applying the inflow-based approach to the 2005-2021 sample. Second, therefore, the gap between both approaches must follow from the lower $\tilde{\Lambda}^N$ and $\tilde{\Lambda}^C$ estimates that the consumption-based approach generates. This most likely relates to an underestimation of consumption expenditures $\tilde{c}_i(t)$ in the PSID survey data. Such downward bias intensifies for higher wealth levels.

Figure 12: Total saving rate out of new resources in the inflow-based approach (left panel) and expenditure-based approach (right panel).

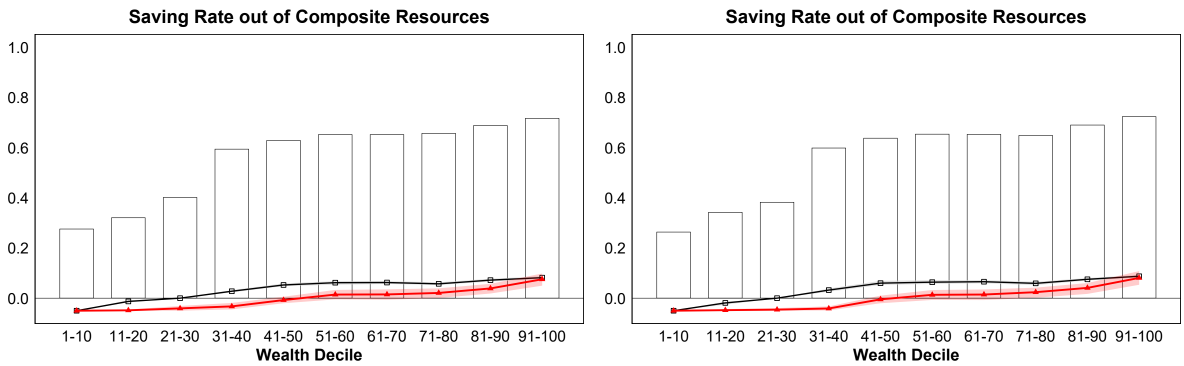
(a) Inflow-based **(b) Expenditure-based**



Note: this figure plots (1) the fraction of households with positive total savings per wealth decile (as bars), and (2) total saving rates out of new resources per wealth decile (as lines). The latter have been computed based on the cross-sectional method (in black, Equation 16) and the aggregate method (in red, Equation 19). The 95% confidence intervals have been determined using bootstrapping. The calculations are executed for the inflow-based approach and expenditure-based approach, as outlined in Section 4.2. The 2001-2021 sample is used. Edge cases are dealt with as also specified in Section 4.2. The edge cases imply that the saving rates out of wealth and composite resources are ill-defined for the bottom 35% (wealth decile 31-40 and lower).

Figure 13: Total saving rate out of composite resources in the inflow-based approach (left panel) and expenditure-based approach (right panel).

(a) Inflow-based **(b) Expenditure-based**



Note: this figure plots (1) the fraction of households with positive total savings per wealth decile (as bars), and (2) total saving rates out of composite resources per wealth decile (as lines). The latter have been computed based on the cross-sectional method (in black, Equation 16) and the aggregate method (in red, Equation 19). The 95% confidence intervals have been determined using bootstrapping. The calculations are executed for the inflow-based approach and expenditure-based approach, as outlined in Section 4.2. The 2001-2021 sample is used. Edge cases are dealt with as also specified in Section 4.2. The edge cases imply that the saving rates out of wealth and composite resources are ill-defined for the bottom 35% (wealth decile 31-40 and lower).

Chapter 3

Wealth Inequality and Wealth Mobility in the United States: Type Dependence Versus Scale Dependence [1]

This paper uses heterogeneous agent models reliant on both type dependence and scale dependence to jointly study wealth inequality and wealth mobility in the United States. The study makes four contributions to the literature. First, I provide a generalized theoretical framework that allows for various sources of agent heterogeneity and explicitly defines type dependence and scale dependence and their theoretical properties. Second, I show that the distinction between type dependence and scale dependence is critical for studying wealth mobility: compared to scale-dependent models, type-dependent models imply higher wealth mobility for identical wealth inequality outcomes. Third, I construct an Aiyagari-Bewley-Huggett heterogeneous agent model with households and entrepreneurs that jointly matches (untargeted) U.S. wealth inequality and U.S. wealth mobility in 2021. I find that a mixture of scale dependence and type dependence in saving behavior and portfolio allocation is critical in achieving this empirical match. Fourth, I show that labor income inequality and saving ratio heterogeneity are the core contributors to short-run and long-run persistence across the wealth (rank) distribution. In general, there exists an inverse relationship between wealth inequality and wealth mobility.

[1] Ghent University (email: christophe.vanlangenhove@ugent.be). I acknowledge the financial support from the Research Foundation Flanders (FWO, project number: 11I5324N). The computational resources (Stevin Supercomputer Infrastructure) and services used in this work were provided by the VSC (Flemish Supercomputer Center), funded by Ghent University, FWO and the Flemish Government – department EWI. I would like to thank Arthur Apostel, Paula Gobbi, Freddy Heylen, Michal Kobielarz, Jan Lorenz, Yasin Kürşat Önder, Gert Peersman, Alberto Russo, Jan Schulz-Gebhard and Dirk Van de gaer for their valuable feedback and comments on this paper.

1 Introduction

Over the past decade, wealth inequality in the United States has become a more prominent topic of academic research, both from an empirical perspective (e.g. Saez & Zucman, 2016; Smith et al., 2023) and from a theoretical point of view (e.g. Benhabib et al., 2019; Cioffi, 2021; De Nardi & Fella, 2017; Fernandez-Villaverde & Levintal, 2024; Hubmer et al., 2021; Kaymak et al., 2022; Xavier, 2021). These studies focus on measuring and decomposing (1) U.S. wealth inequality at a given point in time, and (2) the increase in U.S. wealth inequality observed since the beginning of the 1980s.

Theoretical Aiyagari-Bewley-Huggett heterogeneous agent models link the presence of wealth inequality to four sources (e.g. Benhabib et al., 2019): (1) labor income heterogeneity ('stochastic earnings'), (2) saving rate heterogeneity, (3) capital income heterogeneity, (4) a positive relationship between asset returns and wealth. To produce such heterogeneities, heterogeneous agent models often distinguish between type dependence and scale dependence in agents' parameters or variables. However, at this point, no explicit, formal definition of such type dependence and scale dependence exists. In addition, type-dependent and scale-dependent process parameters are estimated without reference to underlying data.

Moreover, it would be relevant to connect wealth inequality to wealth mobility. High wealth inequality might be less (more) detrimental if it coincides with high (low) wealth mobility at the top of the wealth distribution. In recent years, empirical work has quantified the degree of wealth mobility in the United States (Pfeffer & Killewald, 2018; Van Langenhove, 2025a). It has also found U.S. wealth mobility to have declined since the end of the 1980s (Van Langenhove, 2025a). However, with some exceptions (Atkeson & Irie, 2022; Benhabib et al., 2019; Benhabib et al., 2022; Fisher, 2019; Gomez, 2023; Hubmer et al., 2024), there exists little theoretical work investigating the interdependence between wealth inequality and wealth mobility drivers.

Outline of the paper In response to these research gaps, the present paper (i) provides a formalized definition of type dependence and scale dependence, (ii) demonstrates that the type dependence versus scale dependence distinction is critical for matching wealth mobility outcomes, (iii) outlines a novel strategy to estimate the type-dependent and scale-dependent parameters in heterogeneous agent models, and (iv) conducts counterfactual analyses on the estimated model to investigate the interdependence between wealth inequality and wealth mobility. I detail these four steps in what follows.

In a first step, I outline a generalized theoretical framework that allows for three sources of agent heterogeneity: (1) heterogeneity due to randomness, (2) heterogeneity in agents' state variable parameters, and (3) heterogeneity in agents' policy variables. The latter two structural agent heterogeneity sources are posited to relate to the interplay between a type-dependent term and a scale-dependent term. In short, the type-dependent term reflects time-varying structural heterogeneity across agents (ex-ante heterogeneity), while the scale-dependent term captures the heterogeneity that follows from differences in wealth or other model state variables (ex-post heterogeneity). I provide a formal, generalized definition for the existence of type dependence and scale dependence, and delineate two theoretical moments that summarize the degree of type dependence and scale dependence. The generalized framework embeds the driving forces of wealth inequality that are found to be critical in the existing literature. Finally, I provide an extensive literature overview of the empirical evidence on state variable parameter and policy variable heterogeneity across the wealth (rank) distribution.

In a second step, I derive the stationary states of simplified heterogeneous agent models to make three theoretical points that underscore the importance of the type versus scale dependence distinction. First, I demonstrate that type-dependent and scale-dependent worlds can generate empirical data patterns that are largely indistinguishable. This makes it challenging to infer the degree of type dependence versus scale dependence from empirical datasets. Second, an existing literature has shown that the distinction between type dependence and scale dependence matters for wealth taxation (e.g. Gaillard & Wangner, 2023; Gerritsen et al., 2025). I add to these studies by showing that it matters also for wealth mobility outcomes: even when generating identical stationary wealth inequality, type-dependent models predict higher wealth mobility than scale-dependent ones. Imposing a realistic degree of type dependence versus scale dependence in Aiyagari-Bewley-Huggett economies is therefore critical for matching stationary wealth mobility outcomes. Third, I show that the relationship between wealth mobility and the scale- and type-dependent parameters is characterized by non-linearities.

In a third step, I construct a full-fledged Aiyagari-Bewley-Huggett heterogeneous agent model with households and entrepreneurs. Households maximize an optimization problem with non-homothetic preferences and entrepreneurship is modeled along the lines of Cagetti & De Nardi (2006). I present a novel, model-driven strategy that estimates the degree of type dependence versus scale dependence in households' saving ratio behavior and portfolio allocation using the heterogeneous agent model in combination with data from the Panel Study of Income Dynamics (PSID). The innovation of this strategy lies in the linkage it creates between a theoretical scale-dependent function and its corresponding type-dependent structure, which is taken from the data. I show that the estimated model replicates well joint U.S. wealth inequality and wealth mobility outcomes in 2021 (which are untargeted model moments).

In a fourth step, I use the estimated heterogeneous agent model to assess the driving forces behind wealth inequality and wealth mobility. I do this by generating counterfactual wealth distributions when shutting down labor income inequality, taxation, saving ratio inequality and different components of return heterogeneity. Three findings persist. First, allowing for a realistic degree of saving ratio type dependence is critical in matching wealth mobility in the stationary model state to its empirical counterpart. Second, labor income inequality and saving ratio inequality are found to be the core driving forces behind agents' persistence in the wealth (rank) distribution in both the short-run and the long-run. Return heterogeneity is found to be less important. Third, in general, there is an inverse relationship between wealth inequality and wealth mobility: higher wealth inequality coincides with lower wealth mobility.

Related literature & contributions This paper contributes to two strands of the literature. First, the generalized type and scale dependence definitions and the novel type and scale dependence estimation strategy contribute to the broad literature on heterogeneous agents. Given that I demonstrate the importance of type dependence versus scale dependence for wealth mobility outcomes, the framework relates most directly to Aiyagari-Bewley-Huggett heterogeneous agent models of the wealth distribution (e.g. Azzalini et al., 2023; Benhabib & Bisin, 2018; Benhabib et al., 2019; Benhabib et al., 2024; Cioffi, 2021; De Nardi & Fella, 2017; Hubmer et al., 2021; Fernandez-Villaverde & Levintal, 2024; Kim et al., 2024; Xavier, 2021). However, the type versus scale dependence distinction is relevant also to any other model that embeds heterogeneous agents. This includes for example the heterogeneous agent literature on the U.S. housing market (e.g. Favilukis et al., 2017) and the HANK literature on business cycle dynamics (e.g. Kaplan et al., 2018).

Second, the paper contributes to the Aiyagari-Bewley-Huggett heterogeneous agent literature replicating the wealth distribution. Specifically, I develop a calibrated model that matches not only the U.S. wealth distribution, but also its turnover (i.e. wealth mobility). Furthermore, I provide a decomposition of the joint driving forces behind wealth inequality and wealth mobility. While there exist a handful of papers studying wealth inequality and wealth mobility jointly (Atkeson & Irie, 2022; Benhabib et al., 2019; Benhabib et al., 2022; Fisher, 2019; Gomez, 2023; Hubmer et al., 2024), these papers do not link wealth mobility to the interplay between type versus scale dependence despite importance of this distinction for mobility outcomes. Moreover, existing theoretical work focuses on a handful of wealth inequality-generating channels at most. On the contrary, the present paper allows for a wide range of structural agent heterogeneities, provides a literature overview on their empirical relevance, and calibrates a heterogeneous agent model that is rigorously embedded in empirical micro data outcomes.

Roadmap This paper proceeds as follows. Section 2 provides a generalized theoretical framework where agents face three sources of heterogeneity. These sources of heterogeneity are related to type dependence and scale dependence, which are defined formally. Section 3 operationalizes the generalized framework by introducing specific state variable processes and providing a literature overview of the empirical evidence on state variable parameter heterogeneity and policy variable heterogeneity across the wealth (rank) distribution. Finally, I define wealth inequality and wealth mobility outcome metrics. Section 4 outlines simplified heterogeneous agent models with saving ratio inequality as dominant inequality-inducing channel. It uses these models to provide three theoretical dependence insights. Section 5 constructs and estimates a full-fledged heterogeneous agent model that matches untargeted U.S. wealth inequality and wealth mobility for 2021. Section 6 decomposes the driving forces behind stationary wealth inequality and mobility by generating counterfactual wealth distributions. Section 7 concludes.

2 A generalized theoretical framework

In this Section, I outline a generalized theoretical framework that allows for agent heterogeneity in different dimensions. I posit that agents' state variable parameters and policy variables are determined by the interplay between a type-dependent and scale-dependent term. In outlining the framework, this paper is the first to provide an explicit definition for type dependence and scale dependence. It generalizes the formulations presented in for instance Gabaix et al. (2016), Gaillard & Wangner (2023), Gerritsen et al. (2025), Van Langenhove (2025a) and Xavier (2021). In addition, I define two theoretical moments that characterize the degree of scale dependence and type dependence. The generalized framework embeds the sources of wealth inequality underscored in the theoretical literature.

2.1 Model environment

Budget constraints I denote agents with subscript i. The economy is populated by two types of agents: households ($x_i = 0$) and entrepreneurs ($x_i = 1$). Households participate in the labor market and allocate their wealth between a riskless asset, equity and housing. Entrepreneurs rely entirely on their business as a source of income, which is similar to for example Gomez & Gouin-Bonenfant (2024). It is also in line with empirical findings on entrepreneurship in Moskowitz & Vissing-Jorgensen (2002) and Kartashova et al. (2014). Agents are infinitely lived ('dynasties').

Let us define the budget constraints for household and entrepreneurial agents. First, the budget constraint of a household i at time t ($x_{i,t} = 0$) is given by:

$$w_{i,t+1} = \theta_{i,t}\left[y_{i,t} + \left[1 + \psi^e_{i,t}\alpha^e_{i,t}r^e_{i,t} + \psi^h_{i,t}\alpha^h_i r^h_{i,t} + (1 - \psi^e_{i,t}\alpha^e_{i,t} - \psi^h_{i,t}\alpha^h_{i,t})r^f_{i,t}\right] w_{i,t} - \tau_{i,t}\right] \tag{1}$$

where w denotes wealth, y labor income, ψ^e equity participation $\in (0,1)$, ψ^h housing participation $\in (0,1)$, α^e the conditional equity portfolio share, α^h the conditional housing portfolio share, r^e the equity return, r^h the housing return, r^f the riskless return and τ taxes. In the remainder of the paper, I abstract from indebtedness: the borrowing constraint is set at zero ($w > 0$). The term in outer parentheses in Equation 1 reflects the household's available resources. θ is the saving ratio: it shows the fraction of available resources that are transferred by the household to the next period. A fraction $1 - \theta$ is then used for consumption. Second, I denote business-specific returns as r^b and the share of an entrepreneur's business holdings in its total wealth as α^b. The budget constraint of an entrepreneurial agent i at t ($x_{i,t} = 1$) is then provided by:

$$w_{i,t+1} = \theta_{i,t}\left[\left[1 + \alpha^b_{i,t}r^b_{i,t} + (1 - \alpha^b_{i,t})r^f_{i,t}\right] w_{i,t} - \tau_{i,t}\right] \tag{2}$$

State variables & policy variables An agent i has six state variables: its labor income $y_{i,t}$, the equity return $r^e_{i,t}$, the housing return $r^h_{i,t}$, the business-specific return $r^b_{i,t}$, the riskless return $r^f_{i,t}$ and its taxes $\tau_{i,t}$. Agents' state variables are exogenous to the agent and are assumed to follow finite-state Markov processes with parameters Ξ. An exception to this are the riskless return and taxes, which are deterministic. I outline the determination of the state variables in Section 2.2. For future reference, I collect the state variables in a set S. For an agent i at time t:

$$S_{i,t} = \{y_{i,t},\ r^e_{i,t},\ r^h_{i,t},\ r^b_{i,t},\ r^f_{i,t},\ \tau_{i,t}\} \tag{3}$$

In addition, an agent i is designated to have ten policy variables: its entrepreneurship entry probability $p^{x,e}_{i,t}$, entrepreneurship exit probability $p^{x,o}_{i,t}$, its saving ratio $\theta_{i,t}$, equity entry probability $p^{e,e}_{i,t}$, equity exit probability $p^{e,o}_{i,t}$, equity portfolio share $\alpha^e_{i,t}$, housing entry probability $p^{h,e}_{i,t}$, housing exit probability $p^{h,o}_{i,t}$, housing portfolio share $\alpha^h_{i,t}$ and business portfolio share $\alpha^b_{i,t}$. The entry and exit probabilities are computed conditional on non-participation and participation respectively. Each of the ten policy variables take on values $\in [0,1]$ and are decided on by agents given their state variable parameters $\Xi_{i,t}$, detailed below. I outline the determination of the policy variables in Section 2.3. Furthermore, I collect these policy variables in a set Λ for future reference. For an agent i at time t:

$$\Lambda_{i,t} = \{p^{x,e}_{i,t},\ p^{x,o}_{i,t},\ \theta_{i,t},\ p^{e,e}_{i,t},\ p^{e,o}_{i,t},\ \alpha^e_{i,t},\ p^{h,e}_{i,t},\ p^{h,o}_{i,t},\ \alpha^h_{i,t},\ \alpha^b_{i,t}\} \tag{4}$$

Sources of agent heterogeneity There exist three sources of agent heterogeneity in this theoretical framework. First, agents may face heterogeneous realizations of their state variables as a result of randomness (or equivalently, 'risk' or 'stochasticity'): $S_{i,t} \neq S_{j,t}$ ('source 1'). For example, agent A may have a higher equity return realization compared to agent B. Second, agents might be structurally heterogeneous in the parameters underlying the state variable stochastic processes, i.e. $\Xi_{i,t} \neq \Xi_{j,t}$ ('source 2'). Such structural heterogeneity follows from both parameter-level type dependence and scale dependence, which is specified in Section 2.2. It implies that agents operate in different playing fields. For example, agent A may face an equity return process with a higher expected return or lower expected volatility compared to agent B. Third, agents may be structurally heterogeneous in their policy variables $\in \Lambda$, i.e. $\Lambda_{i,t} \neq \Lambda_{j,t}$ ('source 3'). Such structural policy variable heterogeneity follows from (1) state variable parameter heterogeneity ($\Xi_{i,t} \neq \Xi_{j,t}$), and (2) outcome-level type dependence and scale dependence. It is detailed in Section 2.3. For example, agent A might have a higher saving ratio or dissimilar portfolio allocation compared to agent B.

2.2 State variable parameter heterogeneity ('source 2')

Section 2.1 allowed for structural heterogeneity in the parameters underlying agents' state variable processes: $\Xi_{i,t} \neq \Xi_{j,t}$. Such structural heterogeneity in state variable parameters is said to follow from an additive specification containing a type-dependent term and scale-dependent term. Formally:

$$z_{i,t} = \varepsilon^{z}_{i,t} + g^{z}\left[w_{i,t}, S_{i,t}\right] \tag{5}$$

where any parameter $z_{i,t} \in \Xi_{i,t}$ of agent i at time t equals the sum of a scale-dependent term $g^{z}\left[w_{i,t}, S_{i,t}\right]$ and a type-dependent term $\varepsilon^{z}_{i,t}$. On the one hand, the scale-dependent term $g^{z}\left[w_{i,t}, S_{i,t}\right]$ reflects the neutral parameter level conditional on agents' initial wealth $w_{i,t}$ and state variables $S_{i,t}$. To study wealth distribution outcomes, we are primarily interested in how this neutral (scale-dependent) term relates to wealth w. On the other hand, the type-dependent term $\varepsilon^{z}_{i,t}$ adds to the scale-dependent term a time-varying term that reflects time-varying unobserved heterogeneity across agents. Low-type agents at t display $\varepsilon^{z}_{i,t} < 0$, while for high-type agents at t it holds that $\varepsilon^{z}_{i,t} > 0$. The type-dependent term obeys a discrete-state Markov chain which I specify later. It introduces an additional source of randomness ('risk' or 'stochasticity') into the model environment.

As hinted on earlier, Equation 5 implies that agents may operate in unequal playing fields: as a result of state variable parameter heterogeneity, some agents' state variable processes will be inherently more favorable to wealth accumulation than those of others. For example, the equity return r^{e} of an agent A might be drawn from a distribution that has a higher expected value or lower expected standard deviation than those of an agent B. Such setting may relate to scale dependence: perhaps agent A is wealthier than agent B and therefore has access to more sophisticated investment funds or trading strategies (a scale dependence). It could also reflect type dependence, however: agent A might have superior investment skills compared to agent B. In any case, ceteris paribus, agent A operates in a setting that is inherently more favorable to wealth accumulation than the operating world of agent B.

2.3 Policy variables heterogeneity ('source 3')

Also the level of a policy variable $v_{i,t} \in \Lambda_{i,t}$ of agent i at t is driven by an interplay between a type-dependent and scale-dependent term. Algebraically:

$$v_{i,t} = \varepsilon^{v}_{i,t} + f^{v}\left[w_{i,t}, S_{i,t} \,\middle|\, \Xi_{i,t}\right] \tag{6}$$

where $f^{v}\left[w_{i,t}, S_{i,t} \,\middle|\, \Xi_{i,t}\right]$ denotes the scale-dependent term and $\varepsilon^{v}_{i,t}$ the type-dependent term. On the one hand, the scale-dependent term $f^{v}\left[w_{i,t}, S_{i,t} \,\middle|\, \Xi_{i,t}\right]$ reflects the neutral policy conditional on agents' initial wealth w and state variables $S_{i,t}$. The function f^{v} may differ across agents i due to heterogeneity in agents' state variable parameters $\in \Xi$. In the absence of such state variable parameter heterogeneity ($\Xi_i = \Xi_j \;\; \forall i, j$), agents face the same scale-dependent function f^{v}. Agent heterogeneity in the scale-dependent term then reflects behavioral responses to heterogeneity in their endowment variable $\{w_{i,t}, S_{i,t}\}$. On the other hand, the type-dependent term $\varepsilon^{v}_{i,t}$ adds to the scale-dependent levels a time-varying shock. This type-dependent term captures unobserved heterogeneity across agents. Such unobserved heterogeneity can relate to preferences, geographic effects or any other variable not included in $\{w_{i,t}, S_{i,t}\}$. Low- and high-type agents respectively display $\varepsilon^{v}_{i,t} < 0$ and > 0. The type-dependent term follows a discrete-state Markov chain that I specify later. $\varepsilon^{v}_{i,t}$ introduces an additional source of randomness ('risk' or 'stochasticity') source into the model.

Implications & literature comparison As noted, Equation 6 implies that scale-dependent functions f^v are identical across agents insofar as agents face the same state variable parameters ($\Xi_i = \Xi_j$). Consequently, I have implicitly assumed that f^v is obtained by solving an intertemporal optimization problem that is ex-ante identical across agents. Any preference heterogeneity between agents is contained instead in the type-dependent term $\varepsilon^{v}_{i,t}$.

The attribution of preference heterogeneity to the type-dependent term differs in a subtle way from (implicit or explicit) definitions in the literature. In existing work, agents are heterogeneous in their discount factors (e.g. Krusell & Smith, 1998; Hubmer et al., 2021; Toda, 2019), in their risk aversion (e.g. Azzalini et al., 2023; Fernandez-Villaverde & Levintal, 2024) or in their preference for wealth (e.g. Michau et al., 2023). This implies that agents are assumed to solve optimization problems that are ex-ante different, leading to a heterogeneous scale-dependent function: $f^v_i \neq f^v_j$ in Equation 6. There are two reasons why such formulation may be suboptimal. First, imposing heterogeneity on preference parameters begs the question on why scale dependence is not allowed for as well. That is, if discount factors, risk aversion or preference for wealth parameters can be structurally different across agents, why can they not also scale with wealth (or other state variables)? Such scale dependence in preference parameters would render most of the commonly used optimization problems intractable, however. Second, the choice of the preference parameter to be made heterogeneous is non-trivial: for example, while discount factor heterogeneity has similar effects across all wealth levels, risk aversion heterogeneity does not. Despite the non-triviality of the parameter choice, existing studies do not explicitly motivate their choice based on empirical data.

The framework from Equation 6 therefore offers a key advantage: it embeds all types of unobserved policy variable heterogeneity across agents without making explicit its underlying source. That is, the heterogeneity in the type-dependent term $\varepsilon^{v}_{i,t}$ in Equation 6 could reflect the complex interplay between different heterogeneous preference parameters (discount factors, risk aversion, preference for wealth), but equally heterogeneity in any other variable not present in the model (geographic effects, keeping up with the Jones's dynamics). In addition, the additive scaling of the scale-dependent component with a type-dependent term makes the type-scale dependent framework very straightforward to work with numerically.

Rational and naive expectations In Equation 6, agents are assumed to have rational expectations on their state variables $\in S$, but naive expectations on their state variable parameters $\in \Xi$. Specifically, an agent chooses its scale-dependent term $f^v \left[w_{i,t}, S_{i,t} \,\middle|\, \Xi_{i,t} \right]$ under perfect knowledge of its state variable processes and the uncertainty therein. However, the agent naively assumes that its current state variable parameters are also its future ones ($\Xi_{i,t} = \Xi_{i,s}$ with $s \in [t+1, ...]$). This implies that the agent might underestimate the benefits (costs) of wealth accumulation (decumulation): it does not take into account the possibility that a higher (lower) wealth stock generates conditions more (less) favorable to wealth accumulation whenever there exists scale dependence with wealth in Ξ. It also means that the agent does not incorporate the possibility that its type-dependent term may be time-variant. The naive expectations on parameters $\in \Xi$ is a necessary assumption to come up with a solution to $f^v \left[w_{i,t}, S_{i,t} \,\middle|\, \Xi_{i,t}\right]$ under common preference specifications, as detailed later.

2.4 Scale dependence and type dependence

The previous subsections posited that agents' state variable parameters and policy variables are determined by the interplay between a type-dependent and scale-dependent term. In this subsection, I provide a formal definition for (the existence of) scale dependence and type dependence. For scale dependence:

$$z \in \Xi \text{ exhibits scale dependence} \iff g^z\left[w_{i,t}, S_{i,t}\right] \not\equiv b \tag{7}$$
$$v \in \Lambda \text{ exhibits scale dependence} \iff f^v\left[w_{i,t}, S_{i,t} \,\middle|\, \Xi_{i,t}\right] \not\equiv b \tag{8}$$

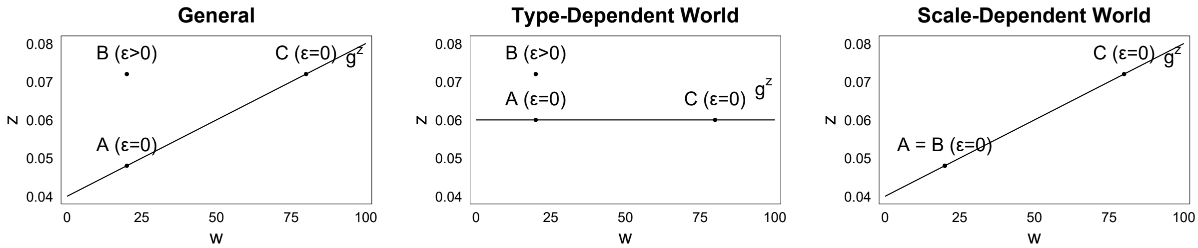
which implies there exists scale dependence in a parameter $z \in \Xi$ or in a policy variable $v \in \Lambda$ whenever the respective neutral functions g^z or f^v are non-constant (i.e. different from some constant b). Instead, there is said to be type dependence in parameter $z \in \Xi$ or policy variable $v \in \Lambda$ whenever ε^z or ε^v are different from zero for at least one agent i. Formally:

$$z \in \Xi \text{ exhibits type dependence} \iff \exists\, i, t : \varepsilon^z_{i,t} \neq 0 \tag{9}$$
$$v \in \Lambda \text{ exhibits type dependence} \iff \exists\, i, t : \varepsilon^v_{i,t} \neq 0 \tag{10}$$

Figure 1 provides a graphic illustration of type dependence and scale dependence for the state variable parameters. Figure 2 does the same for the policy variables.

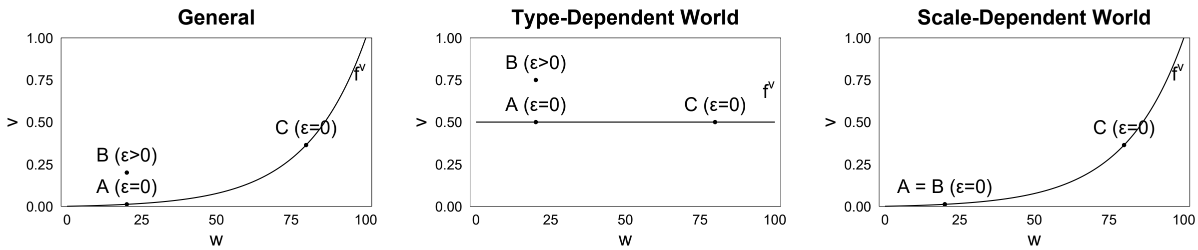
Feedback effects The source of structural agent heterogeneity (type versus scale dependence) critically determines the presence of feedback effects. On the one hand, in a purely type-dependent world ($g^z = a$ and $g^v = b$), agents' parameters or policy variables are ex-ante unrelated to their positions in the wealth distribution (no feedback effects). These parameters and variables may be related ex-post, however: if the type-dependent term is persistent, high-type agents are likely to accumulate more wealth compared to low-type agents if the type is favorable to wealth accumulation. On the other hand, in a purely scale-dependent world ($\varepsilon^z_{i,t} = 0$ and $\varepsilon^v_{i,t} = 0$ for all agents), agents' state variable parameters or policy variables are ex-ante related to their position in the wealth distribution: agents with higher wealth have more

Figure 1: Parameter-level type dependence and scale dependence: a graphic illustration.



Note: this figure compares three parameter-level type versus scale dependence settings. Agents A, B and C are assumed to have identical state variables $S_A = S_B = S_C$. The solid line shows the scale-dependent function g^z. I distinguish between three cases. First, in the general case, the parameter z displays both type dependence and scale dependence. On the one hand, for a given w, there exist neutral-type (A,C) and high-type (B) agents (type dependence). On the other hand, the scale-dependent function g^z scales with wealth (scale dependence). Second, in the type-dependent world, there again exists type dependence (agent A,C versus agent B), but no scale dependence: the function g^z is constant at 0.06. Third, in the scale-dependent world, there is no type dependence: $\varepsilon^z = 0$ for agents A, B and C. There is scale dependence, however: the function g^z rises in wealth w.

Figure 2: Outcome-level type dependence and scale dependence: a graphic illustration.



Note: this figure compares three outcome-level type versus scale dependence settings. Agents A, B and C are assumed to have identical state variables $S_A = S_B = S_C$. The solid line shows the scale-dependent function f^v, the neutral policy function obtained by solving an optimization problem. I distinguish between three cases. First, in the general case, policy variable v displays both type dependence and scale dependence. On the one hand, for a given w, there exists neutral-type (A,C) and high-type (B) agents (type dependence). On the other hand, the scale-dependent function f^v scales with wealth (scale dependence). Second, in the type-dependent world, there again exists type dependence (agent A,C versus agent B), but no scale dependence: the function f^v is constant at 0.50. Third, in the scale-dependent world, there is no type dependence: $\varepsilon^v = 0$ for agents A, B and C. There is scale dependence, however: the function g^z rises in wealth w.

favorable[2] state variable parameters or policy variables. As a result, obtaining higher wealth levels generates conditions that are more favorable to wealth accumulation (feedback effects).

2.5 Theoretical scale and type dependence moments

Section 2.4 defined (the existence of) scale dependence and type dependence in state variable parameters and policy variables. In this subsection, I introduce two summary metrics that provide information on the degree of scale dependence and type dependence. These moments will prove useful to analyze the impact of scale dependence versus type dependence on stationary model outcomes.

Scale dependence $g^z(w_{i,t}, S_{i,t})$ and $f^v(w_{i,t}, S_{i,t})$ are the scale-dependent functions of a state variable parameter $z \in \Xi$ and policy variable $v \in \Lambda$ respectively. In what follows, I define the spread and curvature of scale dependence for a generalized function h, which reflects either the function g^z or the function f^v.

First, define the minimum and maximum of the scale-dependent function $h(w_{i,t}, S_{i,t})$ over the wealth support as:

$$h_{\min}(S_{i,t}) = \min_{w_{i,t}} h(w_{i,t}, S_{i,t}) \tag{11}$$

$$h_{\max}(S_{i,t}) = \max_{w_{i,t}} h(w_{i,t}, S_{i,t}) \tag{12}$$

which allows to define the the spread of the scale dependence as the difference between the maximum and minimum of the scale-dependent function h:

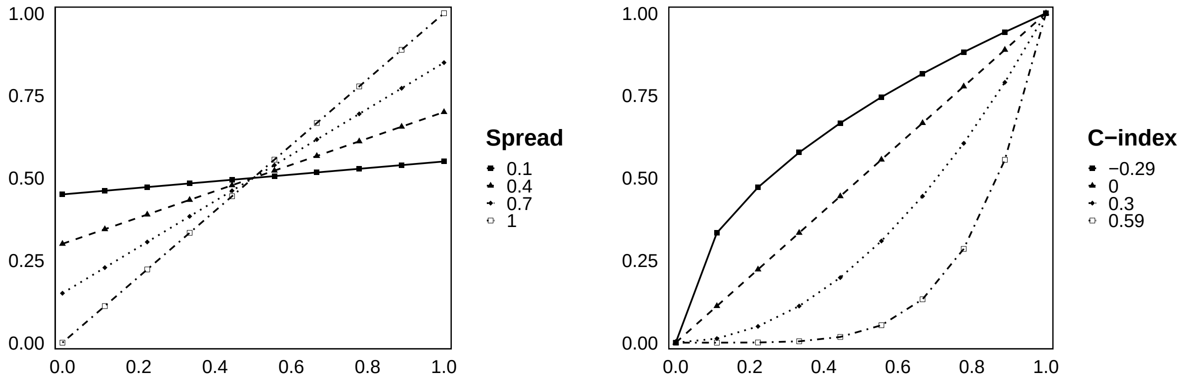
$$\Delta^{(s)}(S_{i,t}) = h_{\max}(S_{i,t}) - h_{\min}(S_{i,t}) \tag{13}$$

Second, let us define a normalized mapping function $\tilde{h}$ over the wealth support. That is, algebraically:

$$\tilde{h}(w_{i,t}, S_{i,t}) = \frac{h(w_{i,t}, S_{i,t}) - h_{\min}(S_{i,t})}{h_{\max}(S_{i,t}) - h_{\min}(S_{i,t})} \in [0, 1] \tag{14}$$

[2] In principle, it is also possible that the relationship is negative: higher wealth creates more unfavorable conditions. I discuss the empirical evidence for state variable parameter heterogeneity and policy variable heterogeneity in Section 3 of the paper.

Figure 3: Theoretical scale dependence moments − spread $\Delta^{(s)}$ and curvature C of a (normalized) scale-dependent function $\tilde{h}$.



Note: this figures illustrates the theoretical scale dependence properties. The x-axis shows the normalized wealth support $\in [0, 1]$. The y-axis represents the output of the normalized scale-dependent function $\tilde{h}$. The left panel plots $\tilde{h}$ across a grid of scale dependence spreads $\Delta^{(s)}$ under the assumption that the scale dependence curvature equals zero: $C = 0$. The right panel displays $\tilde{h}$ over a grid of scale dependence curvatures C when the scale dependence spread is fixed at $\Delta^{(s)} = 1$.

which allows to define the area A (by integrating over the $\tilde{h}$) and the scale dependence curvature index $C(S_{i,t})$:

$$A(S_{i,t}) \;=\; \int_0^1 \tilde{h}\big(w_{(u)}, S_{i,t}\big)\, du \tag{15}$$

$$C(S_{i,t}) = \;1 \;-\; 2\,A(S_{i,t}) \tag{16}$$

where $w_{(u)}$ is the uth–quantile of the wealth distribution. The scale dependence curvature equals zero ($C = 0$) whenever the normalized scale-dependent function $\tilde{h}$ is linear. The index will be smaller than zero ($C < 0$) or larger than zero ($C > 0$) when the normalized function is respectively concave and convex over the wealth support. Figure 3 visualizes the spread and curvature of a scale-dependent function h.

Type dependence The state variable parameter and policy variable type-dependent terms were defined as $\varepsilon^z_{i,t}$ and $\varepsilon^v_{i,t}$ respectively. These terms are assumed to follow some discrete-state Markov chain. I next provide generalized definitions for the spread and persistence of the type dependence, again dropping the z and v superscripts.

Let $\varepsilon_{i,t}$ take on n discrete values $\{\varepsilon_1, \ldots, \varepsilon_n\}$. Moreover, let $\pi = (\pi_1, \ldots, \pi_n)$ constitute the stationary distribution of the discrete-state Markov chain, and let $\Pi_{11}, \ldots, \Pi_{nn}$ denote its transition matrix. This allows to define the spread and persistence of the type-dependent term as respectively:

$$\Delta^{(t)\varepsilon}(S_{i,t}) = \sum_{k=1}^{n} \pi_k \left(\varepsilon_k - \sum_{j=1}^{n} \pi_j \varepsilon_j \right)^2 \tag{17}$$

$$\rho^{\varepsilon}(S_{i,t}) = \sum_{k=1}^{n} \pi_k \, \Pi_{kk} \tag{18}$$

where the spread denotes the unconditional variance of the type-dependent shock values: $\Delta^{(t)\varepsilon}(S_{i,t}) = \mathrm{Var}(\varepsilon)$. The persistence represents the expected probability of staying in the same Markov discrete state. It is computed by averaging over the diagonals of the transition matrix Π.

3 Towards a workable model

In this Section, I specify the generalized framework from Section 2. I start by introducing a structure on the state variable processes. In doing so, I provide an extensive literature overview of the empirical evidence on state variable parameter heterogeneity across the wealth (rank) distribution, discuss the calibration of the state variable process parameters and distinguish between an equal playing field and unequal playing field setting. Thereafter, I summarize the empirical evidence on policy variable heterogeneity across the wealth (rank) distribution. This provides context to the modeling choices made in the present paper. Last, I define the wealth inequality and wealth mobility metrics of interest.

3.1 State variable processes

3.1.1 Labor income

The labor income process is specified as follows. I assume that the logarithm of labor income follows an AR(1)-process with parameters (ρ^y, σ^y):

$$p_{i,t+1} = \rho^y \, p_{i,t} \; + \; \sigma^y \, \varepsilon_{(p)i,t} \tag{19}$$

where $\varepsilon_{(p)} \sim \mathcal{N}(0,1)$. I additionally assume that top 10% labor income is spread out according to a Pareto distribution specified by a scale parameter κ. Formally:

$$y_{i,t+1} = \begin{cases} \exp\big(p_{i,t+1}\big) & F_{p,i,t+1} \leq 0.90 \\ w_{90} \left(1 - \dfrac{F_{p,i,t+1} - 0.90}{0.10}\right)^{-1/\kappa} & F_{p,i,t+1} > 0.90 \end{cases} \tag{20}$$

which is similar to Hubmer et al. (2021) and Gaillard & Wangner (2023). The only difference is that I do not include a transitory component in labor income. The labor income process in Equations 19 and 20 is sufficiently rich to capture (1) top labor income shares, and (2) the degree of long-run labor income mobility. The latter is particularly crucial given the purpose of this paper to construct a heterogeneous agent model able to generate realistic wealth mobility dynamics.

The assumption of exogenous labor income implies that labor income is unrelated to wealth, among other factors. As a result, the assumption excludes a negative relationship between wealth and hours worked, which is one of the core predictions of incomplete markets models with uninsurable wage risk. Such negative relationship would also be consistent with an empirical literature documenting a negative impact of (unexpected) transfers on labor supply (e.g. Cesarini et al., 2017). However, recent empirical evidence on the joint hours worked and wealth distributions (Ferraro & Valaitis, 2022) suggests that hours worked may be roughly flat across the wealth rank distribution. Therefore, in line with other heterogeneous agent models of the U.S. wealth distribution (e.g. Fernandez-Villaverde & Levintal, 2024; Hubmer et al., 2021; Xavier, 2021), I assume an exogenous labor income process for simplicity.

Calibration The calibration of the labor income process can be described in two steps. First, I abstract from structural heterogeneity in labor income parameters: agents have identical parameters ρ^y, σ^y and κ. Although a standard assumption in the heterogeneous agent literature, this represents a simplification: there could exist type dependence in σ^y as a result of occupational choices, or scale dependence insofar remuneration packages become more performance-based at the top. Second, the labor income parameters are calibrated both externally and internally. On the one hand, I externally calibrate $\sigma^y = 0.20$ for all agents i, which is taken from Heathcote et al. (2010). On the other hand, I internally calibrate κ and ρ^y to simultaneously match the top 10% labor income share and long-run labor income mobility. The top 10% labor income share is set at 41% based on data from Gould & Kandra (2022) for 2021. Long-run labor income mobility is computed as the rank-rank coefficient over a thirty-year time horizon and set at 0.40 based on Mazumder (2016). Applying this procedure generates calibration values of $\kappa = 1.71$ and $\rho^y = 0.9717$. The latter is very close to value of 0.97 obtained by Heathcote et al. (2010). Appendix A visualizes the calibrated labor income process.

3.1.2 Asset returns

Before discussing the asset return processes and their calibration, I make two remarks. First, the calibration of these processes is aided by a complementary paper that quantifies return heterogeneity across the wealth (rank) distribution using a 2001-2021 household-level data sample from the Panel Study of Income Dynamics (PSID) and a sample from the Survey of Consumer Finances (SCF) (Van Langenhove, 2025c). Second, the labor income process in Equations 19 and 20 implies the absence of aggregate income growth in the stationary state. To account for this zero growth stationary model state, I lower the equity and housing returns observed in empirical data by the average labor income growth observed between 2001 and 2021. This growth rate equals 2.8% on an annualized basis based on the PSID-sample.

Equity and housing returns Equity returns $r^e_{i,t}$ and housing returns $r^h_{i,t}$ follow a stochastic process with both aggregate and idiosyncratic risk:

$$r^e_{i,t} = \mu^e_{i,t} + \sigma^{e,a}\eta^e_t + \sigma^{e,i}_{i,t}\,\varepsilon^{r^e}_{i,t} \tag{21}$$

$$r^h_{i,t} = \mu^h_{i,t} + \sigma^{h,a}\eta^h_t + \sigma^{h,i}_{i,t}\,\varepsilon^{r^h}_{i,t} \tag{22}$$

where η^e_t and η^h_t denote aggregate equity and housing risk, while $\varepsilon^{r^e}_{i,t}$ and $\varepsilon^{r^h}_{i,t}$ represent idiosyncratic equity and housing risk. All shocks are i.i.d.: $\varepsilon_t \sim \mathcal{N}(0,1)$. The σ-parameters in Equation 21 and 22 reflect the standard deviations of the respective aggregate and idiosyncratic shocks. Aggregate risk is identical for all agents i and captures economy-wide equity and housing market fluctuations. Conversely, idiosyncratic equity risk results from agent heterogeneity in equity portfolios. Idiosyncratic housing risk reflects property-specific risk and geographic risk (local housing market fluctuations). I assume that all four shock processes are independent across time and agents. In particular:

$$\mathbb{E}\left[\eta^e_t\,\varepsilon^{r^e}_{i,t}\right] = 0 \quad \forall i,t \tag{23}$$

$$\mathbb{E}\left[\eta^h_t\,\varepsilon^{r^h}_{i,t}\right] = 0 \quad \forall i,t \tag{24}$$

$$\mathbb{E}\left[\eta^e_t\,\eta^h_t\right] = 0 \quad \forall i,t \tag{25}$$

$$\mathbb{E}\left[\varepsilon^{r^e}_{i,t}\,\varepsilon^{r^h}_{i,t}\right] = 0 \quad \forall i,t \tag{26}$$

Calibration In what follows, I discuss the aggregate risk calibration for equity and housing returns jointly. I then provide an overview of the empirical evidence on state variable parameter heterogeneity across the wealth (rank) distribution, and use this to calibrate the expected return and idiosyncratic risk parameters. I introduce the distinction between an equal playing field and unequal playing field setting.

Aggregate risk standard deviations $\sigma^{e,a}$ and $\sigma^{h,a}$ are calibrated to match the standard deviation of aggregate equity returns and aggregate housing returns. These are taken from the macro-financial database of Jordà et al. (2019) over the 1960-2019 period. I obtain calibration values of $\sigma^{e,a} = 0.17$ and $\sigma^{h,a} = 0.08$. This implies that aggregate housing market fluctuations are less volatile (in the data and in the model) than aggregate equity market ones. The calibration of idiosyncratic housing and equity risk discussed below takes as given these aggregate risk standard deviations.

For housing returns, there exists little empirical evidence on structural heterogeneity across the wealth (rank) distribution: housing returns corrected for cost of debt are found to be stable (Van Langenhove, 2025c) or moderately declining (Snudden, 2025) with wealth (ranks) in the

United States based on data from the PSID. An exception to this is Xavier (2021), who does establish a positive relationship between housing returns and wealth (ranks) using data from the Survey of Consumer Finances (SCF). The positive relationship becomes clear-cut only from the 97th wealth percentile onwards, however. For the Scandinavian countries, a stable relationship between housing returns and wealth is established for Norway (Fagereng et al., 2020) and for Sweden (Bach et al., 2020).

As the empirical evidence points towards the absence of a clear relationship between housing returns and wealth (ranks), I do not allow for structural heterogeneity in housing returns: μ^h and $\sigma^{h,i}$ are identical $\forall i, t$. On the one hand, I calibrate the expected housing return μ^h to the median housing return observed across all households in the PSID-sample of Van Langenhove (2025c): $\mu^h = 0.025$. On the other hand, I calibrate idiosyncratic housing risk $\sigma^{h,i}$ to minimize the distance between the housing return process outcomes in Equation 22 to the standard deviation, first quartile and third quartile of empirical housing observed in Van Langenhove (2025c). This procedure takes the aggregate housing risk calibration as given and generates $\sigma^{h,i} = 0.11$. The calibrated housing return process is visualized in Appendix A.

For equity returns, there is no consensus on structural heterogeneity across the wealth (rank) distribution. For the United States, Van Langenhove (2025c) establishes a positive relationship between equity returns and wealth (ranks), and a negative one between equity return volatility and wealth (ranks) using PSID-data. Xavier (2021) demonstrates that equity returns relate positively to wealth using the SCF, although the relationship is rather weak. On the contrary, Bach et al. (2020) do not find higher risk-adjusted equity returns for the wealthier based on U.S. foundation data. Snudden (2025) finds instead that equity returns are relatively stable across the wealth (rank) distribution and more volatile at the top based on a PSID-sample. Finally, Balloch & Richers (2023) show that equity returns relate negatively to wealth based on a proprietary database of U.S. portfolios. For the Scandinavian countries, Fagereng et al. (2020) establish a positive relationship between equity returns and wealth for Norway, which they attribute to a combination of scale and type dependence. Bach et al. (2020) find no evidence that returns and wealth are related for Sweden.

To incorporate the possibility of a positive link between equity returns and wealth (ranks), I distinguish between two equity return parameter settings. In a first setting ('equal playing field'), I ignore the structural heterogeneity and calibrate $\mu^e = 0.041$ for all agents to match the median equity return observed in Van Langenhove (2025c). Furthermore, I calibrate $\sigma^{e,i} = 0.27$ to minimize the unweighted distance to the standard deviation, first quartile and third quartile of empirical equity returns (from Van Langenhove, 2025c). This procedure takes as given the aggregate equity risk calibration. The calibrated equity return process is visualized in Appendix A. In a second scenario ('unequal playing field'), I do allow for structural heterogeneity and use the same matching procedure for μ^e_d per wealth decile d (based on data from Van Langenhove, 2025c). This procedure assumes that the structural heterogeneity in the expected equity return relates entirely to scale dependence. It is visualized in Figure 4, panel a. Idiosyncratic equity risk $\sigma_d^{e,i}$ is maintained homogeneous across all agents i and equals 0.27.

Business-specific and riskless returns Business-specific returns r^b are modeled according to two principles. First, I presume that business-specific returns are driven entirely by idiosyncratic risk. This is in line with empirical evidence for Sweden: Bach et al. (2020) argue that idiosyncratic risk makes up 74% to 79% of total business-specific return risk. Second, rather than imposing a specific stochastic process, I assume that the $r_{i,t}^b$ realized by an agent i at time period t is randomly drawn from the PSID-sample of business-specific returns computed in Van Langenhove (2025c). Such approach has the advantage of exactly matching high-order empirical moments. Given the complexity of the true business-specific return distribution, such higher-order matching cannot be accomplished when imposing a specific stochastic process. For future reference, I denote the business-return sample from Van Langenhove (2025c) as Γ. Finally, the riskless return r^f is deterministic.

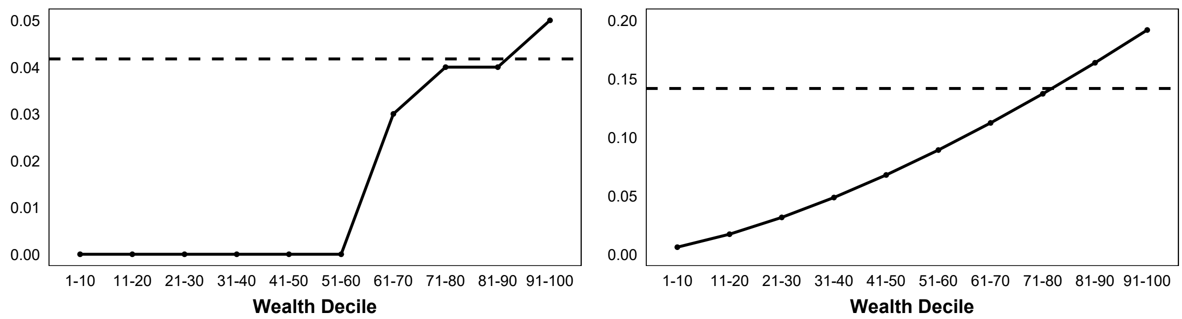
Calibration No consensus exists on the relationship between business return moments and wealth (ranks). For the United States, Van Langenhove (2025c) finds that business returns are higher and less volatile for wealthier entrepreneurs using data from the PSID. This is in line with empirical evidence from both Xavier (2021) (using SCF-data) and Balloch & Richers (2023) (using a proprietary database of U.S. portfolios). The latter study in particular finds that the very top wealthiest face significantly higher Sharpe ratios on business returns as these agents hold more diversified business portfolios. In contrast, Bricker et al. (2021) and Snudder (2024) find that business returns decline with wealth (ranks) using the SCF and PSID respectively. For Norway, Fagereng et al. (2021) establish a strong positive relationship between business returns and wealth (ranks), while Bach et al. (2020) do not obtain a clear link for Sweden.

To incorporate the possibility of structural heterogeneity in business-specific returns, I again distinguish between two settings. In a first setting ('equal playing field'), I pick random draws from the total business return sample Γ. The corresponding business return process is visualized in Appendix A. The median business return in this sample equals 14%, with a standard deviation of 0.95. In a second setting ('unequal playing field'), I allow for business returns to relate positively to wealth (ranks). I therefore create a separate business return sample per wealth decile d, denoted as Γ_d. For higher wealth deciles, expected business returns are higher, and their volatility is slightly lower. This procedure assumes that the observed business return heterogeneity follows purely from scale dependence. Figure 4, panel b provides a visualization.

Finally, the empirical literature suggests a positive relationship between the riskless return and wealth ranks. For the United States, Balloch & Richers (2023), Snudder (2024), Van Langenhove (2025c) and Xavier (2021) point to higher riskless returns for the wealthy. Moreover, Fagereng et al. (2020) establishes a positive relationship for Norway. However, across all papers, the magnitude of the effects is quite limited and only apply to the very top wealthiest. Given the

Figure 4: Equity and business returns across the wealth rank distribution in the equal playing field (dotted line) and unequal playing field (full line) settings.

(a) Equity return. **(b) Business return.**



Note: this figure plots the equity return and (median) business return across wealth deciles d in the equal playing field (dotted line) and unequal playing field setting (full line). The median business return is computed as the median return in the sample Γ_d for the respective wealth decile d. Both return schedules are taken from Van Langenhove (2025c). The empirical returns are corrected for the zero growth stationary state of the models discussed in this paper.

focus on the broader top 10% wealthiest in this paper, I abstract from structural heterogeneity in riskless asset returns: agents are assumed to have an identical riskless return, which is calibrated to $r^f = -0.028$.

Taxation The taxes $\tau_{i,t}$ paid by an agent i at t are assumed to be deterministic. They are computed using the NBER TAXSIM simulator program, version v35. I apply the following assumptions to this program.

For entrepreneurs, budget constraint Equation 2 shows that their total receipts come from their business and from riskless asset holdings. To compute taxes using the NBER TAXSIM program, these receipts need to be separated into labor income, capital income and capital gains. I assume that total income (labor income and capital income) equals the median return in the Λ-sample (in an equal playing field setting) or the median return in the Λ_d-sample corresponding to entrepreneurs' wealth decile d (in an unequal playing field setting) multiplied by entrepreneurs' previous-period wealth. I attribute 60% of this total income to labor income, and 40% to capital income. Capital gains are then set equal to the residual between entrepreneurs' receipts and total income.

For households, labor income is provided by Equations 19 and 20. Consequently, only capital income and capital gains need to be inferred. I apply a similar strategy as for entrepreneurs: I assume that capital income equals the expected wealth return (equal playing field setting) or the expected wealth return corresponding to a household's wealth decile d (unequal playing field setting) multiplied by the household's previous-period wealth. Capital gains are subsequently computed as the residual between a household's actual return on wealth and its capital income.

Furthermore, I make two additional assumptions that apply to both entrepreneurs and households. First, I suppose that 20% of an agent's capital gains are realized at t, with the remainder being unrealized. The key results of this paper are robust to alternative assumptions. Second, the NBER tax simulator requires a number of additional household-level inputs. More precisely, I assume that all households consist of two individuals that are aged 40, married and have no dependents. I leave their geographic state to be undefined, and set their filing year to 2021. Moreover, I abstract from social security income, transfer income and tax deductions. In principle, this may bias wealth inequality outcomes upwards, although it is counteracted by the absence of an unemployment state in the labor income process of Equations 19 and 20.

Two comments I finish this section with two comments. First, the return processes assume that labor income risk is uncorrelated to return risk. For equity and housing:

$$\mathbb{E}\left[\varepsilon_{(p)i,t}\,\eta_t^{e}\right] = 0 \qquad \mathbb{E}\left[\varepsilon_{(p)i,t}\,\varepsilon_{i,t}^{r^e}\right] = 0 \tag{27}$$

$$\mathbb{E}\left[\varepsilon_{(p)i,t}\,\eta_t^{h}\right] = 0 \qquad \mathbb{E}\left[\varepsilon_{(p)i,t}\,\varepsilon_{i,t}^{r^h}\right] = 0 \quad \forall i,t \tag{28}$$

which is a simplifying assumption: for instance, Cocco et al. (2005) find a positive correlation between labor income shocks and housing returns, while Bagliono et al. (2022) obtain a positive correlation between labor income shocks and aggregate stock market shocks. However, these correlations are in general small. Second, having specified the state variable processes, I make explicit the set of state variable parameters Ξ, which was introduced in Section 2:

$$\Xi_{i,t} = \{\rho^y,\ \sigma^y,\ \kappa,\ \mu^e,\ \mu_d^e,\ \sigma^{e,a},\ \sigma^{e,i},\ \mu^h,\ \sigma^{h,a},\ \sigma^{h,i},\ \Gamma,\ \Gamma_d,\ r^f\} \tag{29}$$

where μ_d^e denotes a vector of parameter values, and Γ and Γ_d comprise vectors of returns. The other elements of Ξ are scalars. An overview of the state variable parameter calibration across the equal playing field and unequal playing field settings is provided in Table 1.

3.2 Empirical evidence on policy variable heterogeneity

In Section 3.1, I summarized the empirical evidence on structural heterogeneity in asset returns. In this subsection, I provide an overview of the evidence on the relationship between policy variables $\in \Lambda$ and wealth (ranks), with a focus on the United States. In short, there exists empirical evidence on saving ratio, risky asset participation, portfolio allocation and entrepreneurship heterogeneity across the wealth (rank) distribution.

Saving ratios For the United States, Van Langenhove (2025b) demonstrates that saving rates out of labor income and out of new resources rise with wealth (ranks), while saving rates out of wealth and composite resources are relatively stable across the wealth (rank) distribution. He

Table 1: Calibration of the state variable processes: equal playing field and unequal playing field settings.

Variable	Equal Field	Unequal Field	Target
ρ^y	0.9717	0.9717	Internal: Mazumber (2016)
σ^y	0.20	0.20	External: Heathcote et al. (2010)
κ	1.71	1.71	Internal: Gould & Kandra (2022)
μ^e	0.041	decile-specific μ^e_d	External: Van Langenhove (2025c)
$\sigma^{e,a}$	0.17	0.17	External: Jordà et al. (2019)
$\sigma^{e,i}$	0.27	0.27	Internal: Van Langenhove (2025c)
μ^h	0.025	0.025	External: Van Langenhove (2025c)
$\sigma^{h,a}$	0.08	0.08	External: Jordà et al. (2019)
$\sigma^{h,i}$	0.11	0.11	Internal: Van Langenhove (2025c)
r^b	full Γ	decile-specific Γ_d	External: Van Langenhove (2025c)
r^f	-0.028	-0.028	Van Langenhove (2025c)

Note: this table shows the calibration strategy for the state variable parameters $\in \Xi$ across the equal playing field and unequal playing field settings. Some parameters are calibrated externally, while other parameters are calibrated internally.

also finds that saving ratios θ (as defined in Section 2.1) vary positively with wealth. For future reference, let us label the empirical relationship between saving ratios θ and wealth deciles d as $\bar{f}^{\theta}\left[d_{i,t}\right]$. For the Scandinavian countries, Bach et al. (2018) and Fagereng et al. (2021) show that saving rates out of wealth decline with wealth (ranks) for Sweden and Norway respectively. Fagereng et al. (2021) additionally show that saving rates out of labor income rise with wealth (ranks), which is the key result of their paper. These two studies do not compute saving ratios, however.

Asset participation & allocation There are multiple papers studying risky asset participation and portfolio allocation across the wealth (rank) distribution in the United States. Gaillard & Wangner (2023) and Van Langenhove (2025c) provide evidence using the PSID, while Cioffi (2021) and Xavier (2021) do so based on the SCF. Both datasets generate similar findings.

On the one hand, housing participation ψ^h reaches levels above 80% already at the middle part of the wealth distribution, and continues to rise gradually for wealthier households. Both housing entry and exit rates are therefore relatively flat from the 50th wealth percentile onwards in the empirical data. For future reference, I denote these empirical relationships between housing entry and exit rates and wealth deciles d as $\bar{f}^{p^{h,e}}\left[d_{i,t}\right]$ and $\bar{f}^{p^{h,o}}\left[d_{i,t}\right]$ respectively. Instead, the conditional housing portfolio share α^h peaks in the middle of the wealth distribution at around 80% and drops to around 40% for the top 10% wealthiest. This empirical relationship is referenced as $\bar{f}^{\alpha^h}\left[d_{i,t}\right]$.

On the other hand, equity participation ψ^e equals around 40%-50% in the middle part of the wealth distribution, but rises strongly to around 90% for the top 10% wealthiest. As a result, equity entry rates rise with wealth (ranks), while equity exit rates are declining along the wealth (rank) distribution. I label these empirical relationships as $\bar{f}^{p^{e,e}}[d_{i,t}]$ and $\bar{f}^{p^{e,o}}[d_{i,t}]$ respectively. On the contrary, the conditional equity portfolio share α^e is overall relatively stable or only slightly increasing across the wealth (rank) distribution. This empirical schedule is referenced as $\bar{f}^{\alpha^e}[d_{i,t}]$ in the remainder of this paper. Taken together, the positive relationship between equity portfolio allocation and wealth (ranks) mainly relates to the extensive margin. Similar findings persist for the Nordic countries in for example Bach et al. (2020) and Fagereng et al. (2020).

Entrepreneurship With respect to entrepreneurship across the wealth (rank) distribution, I distinguish again between the extensive and intensive margin.

At the extensive margin, entrepreneurs ($x = 1$) are concentrated predominantly at the top of the wealth distribution: the share of agents with business assets rises strongly across the wealth (rank) distribution. This is demonstrated in for example Cagetti & De Nardi (2006), and more recently by Van Langenhove (2025c) and Balloch & Richers (2023). It also holds in Fagereng et al. (2020) for Norway. While the entrepreneurship rate is high for the top 10% wealthiest $-$ with estimates ranging between 35% and 50% $-$ it is particularly high at the tail: among the top 0.1% wealthiest and beyond, over 80% of agents are business owners. Entrepreneurship entry rates are rising across the wealth (rank) distribution, while exit rates are declining. I reference the empirical entry and exit rates schedules across wealth deciles (from Van Langenhove, 2025c) as $\bar{f}^{p^{x,e}}[d_{i,t}]$ and $\bar{f}^{p^{x,o}}[d_{i,t}]$ respectively.

At the intensive margin, the fraction of wealth that entrepreneurs have allocated to business assets α^b exceeds 50% for the bottom 50% wealthiest, but drops to close to 30% from the 60th wealth percentile onwards (Van Langenhove, 2025c). It then remains more or less stable at this level before rising again for the very top wealthiest households (top 0.1% and beyond). The allocation of the non-business wealth of entrepreneurs follows a similar pattern as for households: allocation to housing is dominant at the middle of the wealth distribution, but is increasingly tilted towards equity for agents at the top of the wealth distribution.

3.3 Outcome metrics

The heterogeneous agent models in this paper jointly target wealth inequality and wealth mobility for 2021. In what follows, I summarize the inequality and mobility metrics used, and their empirical values for 2021.

For wealth inequality, I use bottom 50%, middle 50%-90% and top 10% wealth shares, as is common in the heterogeneous agent literature. The empirical value in 2021 (or in neighboring years) for the bottom 50%, middle 50%-90% and top 10% wealth shares varies across different data sources and methodologies: it depends on the unit of analysis (households, individuals, tax units), underlying data (administrative tax data in Saez & Zucman (2016) and Smith et al. (2023) versus survey data in Kuhn et al. (2020) and Kuhn & Ríos-Rull (2025)), treatment of Social Security (e.g. Catherine et al., 2020) and corrections for tail wealth. I take a top 10% share of 77%, middle 50%-90% share of 23% and bottom 50% share of 0%. These shares were computed from the top-wealth corrected SCF for 2021 and are similar in magnitude with administrative tax data estimates from Saez & Zucman (2016) and Smith et al. (2023).

For wealth mobility, I distinguish between short-run and long-run wealth mobility. I introduce two types of outcome measures, in line with Van Langenhove (2025a). On the one hand, as an overall wealth mobility metric, I compute rank-rank coefficients over a five-year window (short-run) and over a thirty-year window (long-run). These rank-rank coefficients are computed by regressing the cross-section of wealth ranks at $t + 5$ or $t + 30$ on the cross-section at t using OLS. The coefficients are denoted as β^{sr} and β^{lr}. On the other hand, to investigate non-linearities in wealth mobility dynamics, I report the fraction of steady wealthy and steady poor agents, in line with Van Langenhove (2025a). These groups include respectively (1) agents remaining in the top 10% wealth bracket, and (2) agents remaining in the bottom 20% wealth bracket.

Getting from empirical wealth mobility data to short-run and long-run wealth mobility values relevant to the stationary state of our model is non-trivial. The model considers agents that are infinitely lived, so-called dynasties (see Section 2). In practice, agents have finite lifespans, which introduces lifecycle effects into wealth accumulation dynamics. I therefore come up with the following approximations. On the one hand, for short-run wealth mobility measures, I compare individuals' within-cohort wealth ranks at ages 50-54 to their wealth ranks at ages 45-49 (intra-generational perspective). I focus on the 45-54 age range as (1) the within-cohort wealth distribution approximates the overall wealth distribution well at these ages, and (2) a realistic share of agents is still active on the labor market and as entrepreneur. This generates a short-run rank-rank coefficient of $\beta^{sr} = 0.84$ (Van Langenhove, 2025a). On the other hand, for long-run measures, I compare the within-cohort wealth ranks of children at ages 50-54 to those of their parents at the same age bracket. Given that the average parent age at childbirth in the U.S. equaled 29.6 years over the period 2000-2021, this approximates wealth mobility outcomes over a thirty-year time horizon. This procedure generates an empirical long-run rank-rank coefficient $\beta^{lr} = 0.40$[3].

[3]Van Langenhove (2025a) uses a PSID-sample over the period 1969-2021. However, actual wealth data is available only from 1984 onwards. The author therefore approximates wealth data prior to 1984 using a machine-learning (ML) model based on main housing values and rental payments. This proxy suffers from a downward bias in estimating wealth mobility outcomes, however. A comparison of children and parent wealth ranks at ages 50-54 is available only based on proxy data. The reported $\beta^{lr} = 0.40$ is obtained after applying the bias-correcting procedure outlined in Van Langenhove (2025a).

4 Theoretical insights on type and scale dependence

In this Section, I compare heterogeneous agent models through numerical simulations to make three theoretical points. First, I demonstrate that type-dependent and scale-dependent worlds can generate empirical data patterns that are largely indistinguishable. This makes it difficult to infer the degree of type versus scale dependence from empirical data. Second, an existing literature has shown that the distinction between type and scale dependence matters for the optimality of wealth taxation (e.g. Gaillard & Wangner, 2023; Gerritsen et al., 2025). I add to these studies by showing that it matters also for wealth mobility outcomes: even when generating identical wealth inequality outcomes, type-dependent models predict higher wealth mobility than scale-dependent ones. Third, I demonstrate that the relationship between wealth mobility and the scale- and type-dependent parameters is characterized by non-linearities. Moreover, the relationship between wealth inequality and wealth mobility is not unambiguously negative.

4.1 Preliminaries

Before making the three aforementioned points, I impose a number of simplifying assumptions on the heterogeneous agent models used in this Section. The theoretical insights derived here are robust to alternative model formulations. The simplifying assumptions are relaxed again from Section 5 onwards. In addition, I impose a convenient discrete-state Markov chain on the type-dependent terms.

Simplifying assumptions First, I take as given the state variable processes from Section 3 and depart from the equal playing field setting. In other words, there exists no structural heterogeneity in equity and business returns. Second, I assume that all agents in the model are households ($x = 0 \ \forall i, t$). This is accomplished by setting the conditional entry probability $p^{x,e}_{i,t}$ equal to zero and the conditional exit probability $p^{x,o}_{i,t}$ to one $\forall i, t$. Third, for the policy variables, I impose that agents are heterogeneous only in their saving ratio θ. Heterogeneity in the probability of risky asset participation and portfolio allocation is abstracted from. These homogeneous variables are calibrated for all agents to their median values observed in a PSID-sample of Van Langenhove (2025c)[4]. More precisely:

$$p^{e,e}\,[...] = 0.15 \quad \forall i, t \tag{30}$$

$$p^{e,o}\,[...] = 0.18 \quad \forall i, t \tag{31}$$

$$\alpha^e_i\,[...] = 0.22 \quad \forall i, t \tag{32}$$

[4] In this simplified model, there are still various wealth inequality-inducing forces at play. On the one hand, 'source 1' is active: agents face heterogeneous realizations of their labor income and returns. On the other hand, under 'source 3', structural saving ratio heterogeneity is allowed for: some agents have higher saving ratios compared to others.

and:

$$p^{h,e}\,[...] = 0.18 \quad \forall i,t \tag{33}$$

$$p_o^{h,e}\,[...] = 0.09 \quad \forall i,t \tag{34}$$

$$\alpha_i^h\,[...] = 0.62 \quad \forall i,t \tag{35}$$

Type-dependent process In the generalized framework of the Section 2, I argued that the type-dependent term follows some discrete-state Markov chain, which introduces an additional source of randomness ('risk') into the model. To operationalize the model, I impose a structure on the type-dependent Markov chain. Denote as n the number of discrete type-states, labeled as $(1,\ldots,n)$. The type-dependent transition matrix is given by:

$$\Pi_{ij} = \begin{cases} \rho & i = j \\ \dfrac{(1-\rho)\,\exp\left(-\,d_{ij}\right)}{\sum_{k\neq i}\exp\left(-\,d_{ik}\right)} & i \neq j \end{cases} \tag{36}$$

where $d_{ij} = |i-j| \qquad i,j = 1,\ldots,n$ denotes the distance between two states i and j and $\rho \in [0,1]$ the persistence parameter (which reflects the mass on the diagonal). A larger ρ implies stronger persistence: a more substantial mass is concentrated on the diagonals of the transition matrix. This example specification for the type-dependent process is convenient as the type dependence persistence (defined in Section 2.5) will equal exactly ρ: all diagonal values in the transition matrix Π are equal to ρ.

4.2 Insight 1 − data patterns across type- and scale-dependent worlds

Under the simplified model setting described in Section 4.1, I compare the saving ratio data that is produced across the wealth (rank) distribution for four types of models. The models differ by their scale versus type dependence imposed on saving ratios θ. The simulation results are subsequently used to highlight two implications.

Four models In a first model ('pure scale dependence'), there exists no type dependence in saving ratios ($\varepsilon^\theta = 0 \;\; \forall i,t$). This model relies purely on scale dependence to generate saving ratio heterogeneity across agents i. The scale-dependent function f^θ is set equal to the relationship that is observed between saving ratios and wealth deciles d in the empirical data (Van Langenhove, 2025b): $f^\theta\,[...] = \bar{f}^\theta\,[d_{i,t}]$. In the stationary model state, the median saving ratio and its dispersion across the wealth (rank) distribution in the pure scale-dependent model are displayed in panel (a) of Figure 5. While the relationship between saving ratios and wealth (ranks) is imposed to match the empirical pattern by construction, there critically exists no saving ratio heterogeneity within a wealth decile d: all agents within a decile display identical saving ratios.

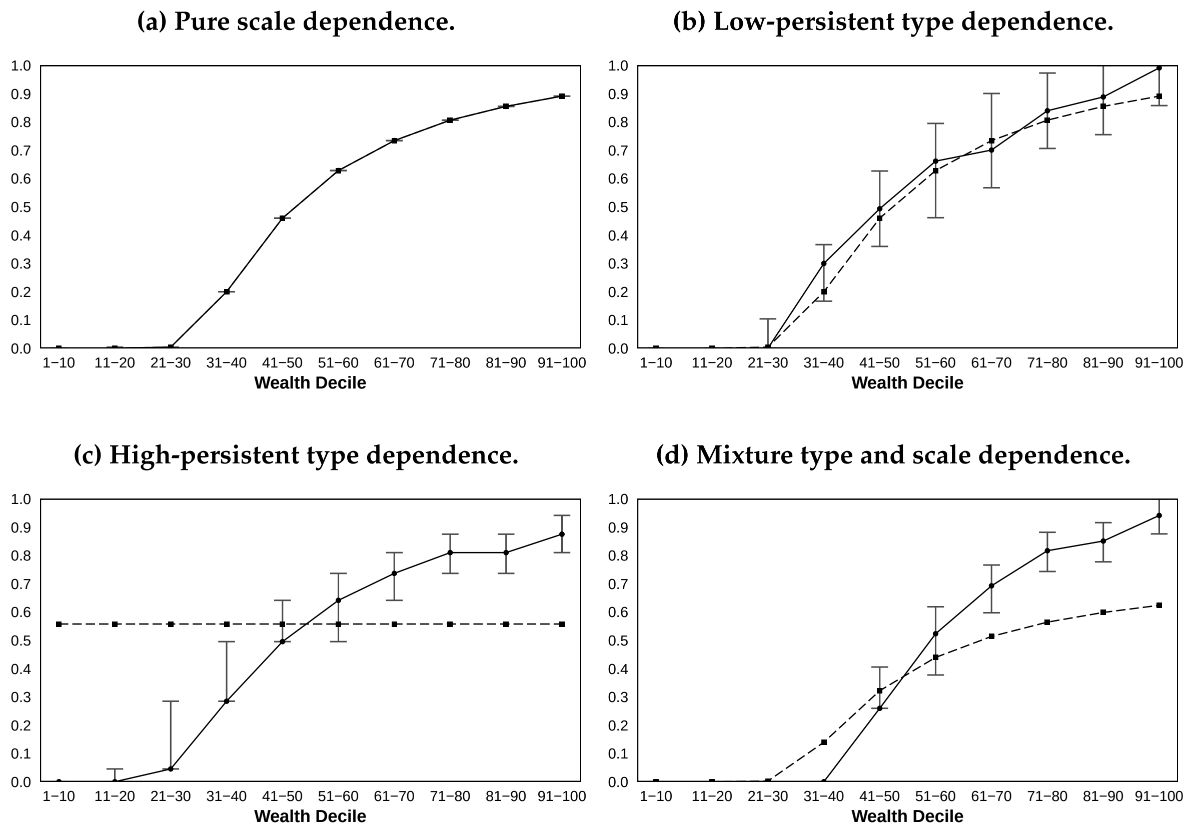
In a second model ('low-persistent type dependence'), the purely scale-dependent model is augmented with low-persistent type dependence in saving ratios. That is, the type-dependent term ε^{θ} is imposed to take on normally-distributed states between -0.3 and 0.3 and its persistence is set at a very low level: $\rho = 0.10$. The saving ratio patterns across the wealth (rank) distribution in the stationary model state are displayed in panel (b) of Figure 5. Two observations stand out. On the one hand, saving ratio inequality is somewhat higher than in a purely scale-dependent model: the median saving ratio is higher in top wealth deciles, and lower in middle wealth deciles. On the other hand, there exists significant dispersion around the median saving ratio within each wealth decile. This dispersion is at its highest in the middle of the wealth distribution.

A third model ('high-persistent type dependence') instead departs from highly persistent type dependence. Scale dependence is abstracted from: the scale-dependent saving ratio is a constant ($f^{\theta} = 0.60$). The type-dependent term takes on normally-distributed states between -0.3 and 0.3, while the type-dependent persistence parameter is estimated to match the saving ratio inequality observed in the data. This estimation exercise leads to $\rho = 0.90$. Agents with persistently positive type-dependent terms (high-type agents) make it to the top of the wealth distribution, whereas agents with persistently negative terms (low-type agents) drop down to the bottom. Unlike in the purely scale-dependent model, there exists dispersion around the median for each wealth decile due to the randomness in the type-dependent term (Figure 5, panel c).

In a fourth model ('mixture type and scale dependence'), I combine scale dependence and type dependence. On the one hand, the scale-dependent function is imposed to be less steep than the relationship observed in the empirical data: $f^{\theta}\,[...] = 0.70\,\bar{f}^{\theta}\,[d_{i,t}]$. On the other hand, type dependence persistence is again estimated to match empirical saving ratio inequality. Considering that part the saving ratio heterogeneity is captured by scale dependence, this generates a persistence parameter that is lower than in the third model: $\rho = 0.70$. There again exists some dispersion around the median saving ratio for each wealth decile (Figure 5, panel d) because of the presence of type dependence.

Two implications A comparison of these four simplified models yields two key implications. First, looking at the median levels of a policy variable across the wealth (rank) distribution in empirical data yields little information on the importance of type versus scale dependence for that variable. Let us continue with the saving ratio example. On the one hand, saving ratios might relate positively to wealth as a result of selection effects when type dependence is strong and persistent: persistently high-type agents rise to the top of the wealth distribution, while persistently low-type agents drop to the bottom. On the other hand, saving ratios might relate positively to wealth (ranks) as a result of scale dependence: under many commonly used preference structures, optimal saving ratios are a positive function agents' wealth.

Figure 5: Saving ratio patterns (solid lines) across the wealth (rank) distribution in four simplified heterogeneous agent models.



(a) Pure scale dependence.

(b) Low-persistent type dependence.

(c) High-persistent type dependence.

(d) Mixture type and scale dependence.

Note: the solid line shows the median saving ratio by wealth decile in the simulated data, while the bars indicate the inter-quartile range. The dotted line represents the imposed scale-dependent saving function f^{θ} used in the model. In panel (a), this function exactly matches the simulated saving ratios, since no type dependence is present. In panels (b) through (d), type dependence is introduced, causing the simulated saving ratios (solid line) to deviate from the imposed scale-dependent function (dotted line). The degree of divergence increases with the persistence of type dependence. All models build on the simplified framework described in Section 4.1.

Second, in theory, the dispersion observed in empirical data around the median saving ratios within a wealth bracket yields information on the spread and persistence of type dependence. For given type dependence persistence ρ, a higher type dependence spread $\Delta^{(t)}$ generates more substantial dispersion around the decile median (Appendix B, Figure 11). Conversely, for a given type dependence spread $\Delta^{(t)}$, a higher persistence ρ creates lower dispersion around the median saving ratio of bottom and top wealth deciles (Appendix B, Figure 12). However, in practice, agents' saving ratios also reflect optimal responses to agents' state variables. Moreover, they interact with asset participation and portfolio allocation decisions. Consequently, there exist several other sources of within-bracket saving ratio heterogeneity across agents beyond mere type dependence. This renders an identification of the contribution of type dependence to the dispersion of saving ratios observed in empirical data very tedious.

A proper data-driven estimation of the degree of type versus scale dependence in state variable process parameters or policy variables would require a source of exogenous wealth variation. Two obvious sources of such variation include unanticipated inter-generational transfer receipts or lottery winnings. However, such data is very hard to come by in sufficient sample size. Despite the absence of exogenous sources of wealth variation, the literature has made some data-driven attempts to disentangle type versus scale dependence. Hurst & Lusardi (2004) find scale dependence in the probability of shifting into entrepreneurship, while Bach et al. (2020) and Fagereng et al. (2020) quantify the type dependence in asset return process parameters for Sweden and Norway. Each of these three papers uses a fixed effect regression model. On the contrary, In Section 5 of the present paper, I introduce a novel, model-based estimation strategy to quantify the degree of scale versus type dependence in agents' saving ratios and portfolio allocation.

4.3 Insight 2 − wealth mobility across type- and scale-dependent worlds

The previous subsection demonstrated that type versus scale-dependent models can generate largely indistinguishable saving ratio patterns across the wealth (rank) distribution. It did not show why the distinction between type and scale dependence is relevant, however. In this subsection, I fill this gap: I show that a purely type-dependent and purely scale-dependent model that generate identical wealth inequality moments display diverging wealth mobility outcomes: the purely type-dependent model predicts higher wealth mobility than the purely scale-dependent one. I demonstrate this under the simplified model assumptions outlined in Section 4.1.

In a first model (M1, 'no saving ratio heterogeneity'), saving ratio inequality is shut down: the saving ratio is for all agents set to the median level observed across all households in a PSID-sample over the period 2001-2021. This leads to: $\theta_{i,t} = 0.56$. In this model specification, all agent heterogeneity therefore follows from state variable randomness ('source 1'). Stationary wealth inequality in this model is roughly identical to stationary labor income inequality: the

Table 2: Wealth inequality, wealth mobility, and income inequality across three simplified models.

	M1	M2	M3
Wealth inequality			
Bottom 50%	0.20	0.01	0.01
Middle 50–90%	0.41	0.35	0.36
Top 10%	0.39	0.63	0.63
Wealth mobility			
Short-run β	0.89	0.98	0.92
Long-run β	0.42	0.84	0.57
Steady wealthy	0.07	0.09	0.08
Steady poor	0.15	0.20	0.16
Labor income inequality			
Bottom 50%	0.17	0.17	0.17
Middle 50–90%	0.42	0.42	0.42
Top 10%	0.41	0.41	0.41

Note: this table displays the stationary wealth inequality and wealth mobility outcomes across three simplified models (M1, M2, M3). Each of the three models departs from the simplified framework outlined in Section 4.1. The models differ solely in the assumptions that they impose on saving ratios. More precisely, model M1 abstracts from saving ratio heterogeneity. Model M2 relies purely on saving ratio scale dependence, which is set equal to the relationship between saving ratios and wealth deciles in the empirical data. Model M3 replicates the same stationary wealth inequality outcomes of M2 based on purely saving ratio type dependence. Critically, the type-dependent model M3 produces higher wealth mobility than the scale-dependent one (lower rank-rank coefficients).

top 10% wealth share (at 0.39) is only slightly smaller than the top 10% labor income share (at 0.41) (Table 2, column 2). If taxation was absent, the top 10% wealth share would be exactly equal to the top 10% labor income share. The highly simplified model M1 matches empirical short-run and long-run wealth mobility moments relatively well: the short-run rank-rank coefficient equals 0.89 (compared to 0.84 in the data), while the long-run rank-rank coefficient amounts to 0.42 (compared to 0.40 in the data).

In a second model (M2, 'scale dependence'), saving ratio inequality is introduced through scale dependence only: I impose the empirically observed relationship between saving ratios and wealth (ranks) to the scale-dependent saving ratio function ($f^{\theta}[...] = \bar{f}^{\theta}[d_{i,t}]$). This model is identical to the purely scale-dependent model from Section 4.2. Wealth inequality gets closer to its empirical counterpart: the top 10% wealth share in model M2 equals 0.63 (Table 2, column 3). Wealth mobility declines, especially in the long-run: the short-run rank-rank coefficient rises to 0.98, while the long-run coefficient increases to 0.84. The decline in wealth mobility holds both at the top and at the bottom. The relationship between saving ratios and wealth (ranks) by definition matches the one observed in the data.

In a third model (M3, 'type dependence'), saving ratio inequality is introduced through type dependence only. Specifically, the scale-dependent saving ratio is assumed identical for all agents: $f^{\theta} = 0.56$. The type-dependent term state levels are discretized to their empirical counterparts in a PSID-sample conditional on $f^{\theta} = 0.56$, leading to $\Delta^{(s)\theta} = 0.16$. I then estimate the type-dependent persistence parameter ρ^{θ} to match the stationary wealth inequality of model M2. This produces $\rho^{\theta} = 0.91$. Despite generating identical top wealth inequality, the type-dependent model (M3) predicts higher wealth mobility than the scale-dependent one (M2): the short-run rank-rank coefficient in M3 equals 0.92 (compared to 0.98 in M2), while the long-run coefficient amounts to 0.57 (compared to 0.84 in M2) (Table 2, column 4). This higher wealth mobility in type-dependent models holds both at the bottom and at the top.

In summary, type-dependent models generate higher stationary wealth mobility than scale-dependent models when wealth inequality is identical. What explains this outcome? Type-dependent models ultimately introduce an additional source of randomness into the model economy. Let us take the top 10% wealth decile as an example. As long as $\rho < 1$, there exists (in expectation) a group of high-type saving ratio households at the top that will experience a negative shock to their type-dependent term. Conversely, there exists (in expectation) a set of households below the top experiencing a positive shock to their type-dependent term. The former group is likely to drop out of the top 10%, while the latter is expected to enter this wealth bracket. This interplay generates higher wealth mobility in type-dependent models compared to scale-dependent ones even when wealth inequality (and saving ratio inequality) outcomes are indistinguishable.

4.4 Insight 3 − non-linearities in inequality-mobility outcomes

The previous subsection underscored that type-dependent models generate higher wealth mobility than scale-dependent models when replicating equivalent wealth inequality outcomes. These simulations dealt with the existence of scale dependence and type dependence. On the contrary, this subsection focuses on the degree of scale dependence and type dependence. Specifically, I investigate how the theoretical properties of scale dependence and type dependence affect stationary wealth inequality and wealth mobility outcomes. I derive the stationary model state using numerical simulations under the simplified model assumptions outlined in Section 4.1.

Scale dependence To analyze the effects of scale dependence curvature and spread, I assume that the scale-dependent function f^{θ} is monotonically increasing in wealth deciles d and takes on strictly positive output values. I also abstract from saving ratio type dependence for simplicity ($\varepsilon^{\theta} = 0 \;\; \forall i, t$). Two findings persist.

First, a higher scale dependence curvature C^{θ} implies larger wealth inequality, regardless of the scale dependence spread $\Delta^{(s)\theta}$ (Figure 6, panel a). On the contrary, the relationship between C^{θ} and wealth mobility does depend on the scale dependence spread: only when $\Delta^{(s)\theta}$

is sufficiently high will the relationship between the curvature and wealth mobility be negative. For low $\Delta^{(s)\theta}$, there exists no clear-cut relationship (Figure 6, panel b). Furthermore, for high $\Delta^{(s)\theta}$, the negative relationship between C^{θ} and overall wealth mobility conceals two countervailing effects: wealth mobility at the top declines, while wealth mobility at the bottom increases (Figure 6, panels c and d).

What explains these results? A higher C^{θ} implies a more convex scale-dependent function, generating a larger mass of agents with low saving ratios. As a result, the fraction of agents with wealth below the average wealth level increases, and a larger group of agents displays wealth levels that are close to the borrowing constraint (at 0). These smaller absolute wealth differences at the bottom produce higher relative wealth mobility in this part of the wealth distribution. At the top, the inverse is true: a higher C^{θ} creates a small number of agents with very high saving ratios. These agents accumulate a lot of wealth compared to the average wealth level. This lowers wealth mobility at the top.

Second, an increase in the scale dependence spread $\Delta^{(s)\theta}$ widens the wealth distribution and lowers its turnover: the stationary top 10% wealth share increases, as does the long-run rank-rank coefficient (Figure 7, panels a and b). This holds for any value of the scale dependence curvature C^{θ}. The declining wealth mobility holds at the bottom and at the top of the wealth distribution. The wealth mobility effect at the bottom is substantially stronger for lower C^{θ} (Figure 7, panels c and d).

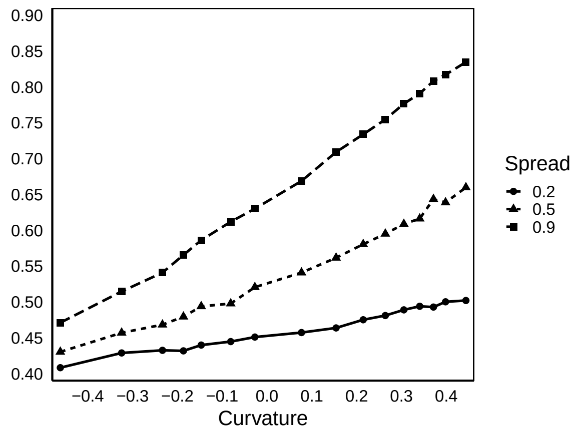
What are the driving forces behind these findings? Intuitively, a wider spread $\Delta^{(s)\theta}$ implies larger cross-sectional saving ratio inequality, which produces more substantial absolute wealth differences across agents. These greater wealth differences generate lower wealth mobility. However, at the bottom, a high C^{θ} weakens this effect: a larger mass of agents has saving ratios close to zero, implying more limited absolute wealth differences across agents. This leads to higher relative wealth mobility at the bottom.

Type dependence To investigate the effects of type-dependent persistence and spread, I assume that the scale-dependent function f^{θ} is a constant at 0.60. As a result, there exists no scale dependence. All saving ratio heterogeneity is generated by type dependence. Two findings persist.

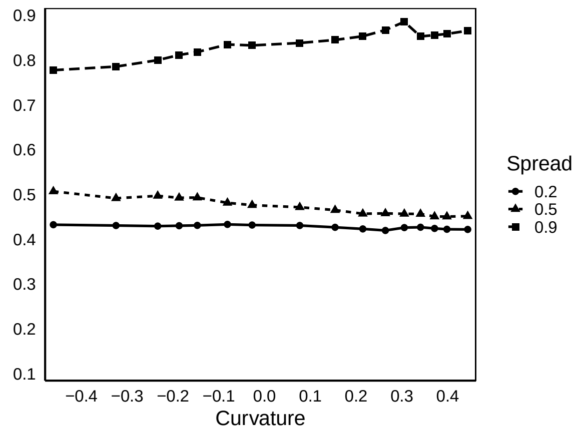
First, a higher type dependence persistence ρ^{θ} leads to moderately higher wealth inequality and lower wealth mobility, regardless of the type dependence spread $\Delta^{(t)\theta}$ (Figure 8, panels a and b). The declining turnover across the wealth (rank) distribution holds both at the bottom and at the top (Figure 6, panels c and d). However, this relationship between ρ^{θ} and wealth inequality and mobility outcomes becomes clear-cut only for higher ρ^{θ}. The tipping point for

Figure 6: Scale dependence curvature and stationary wealth inequality and mobility moments across three spread values (for saving ratios θ).

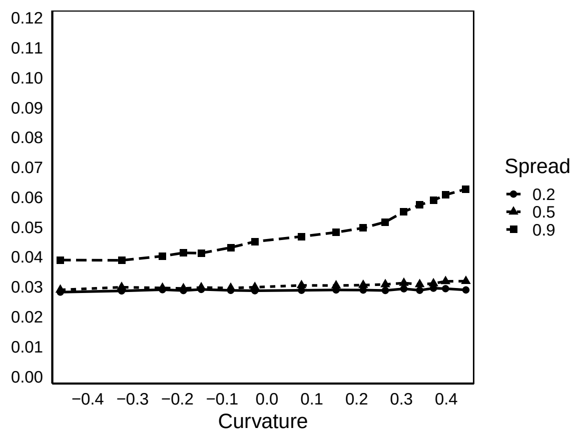
(a) Top 10% wealth share.



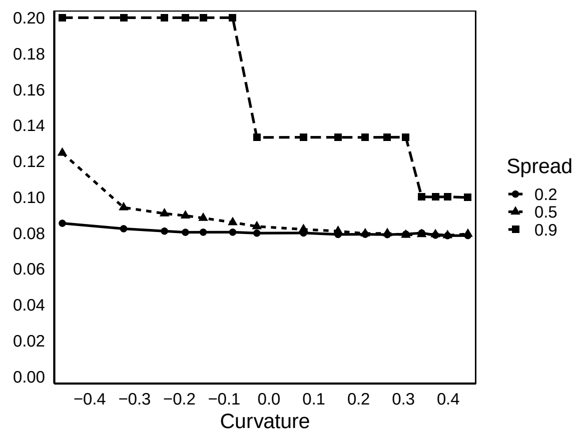
(b) Long-run β-coefficient.



(c) Steady wealthy.

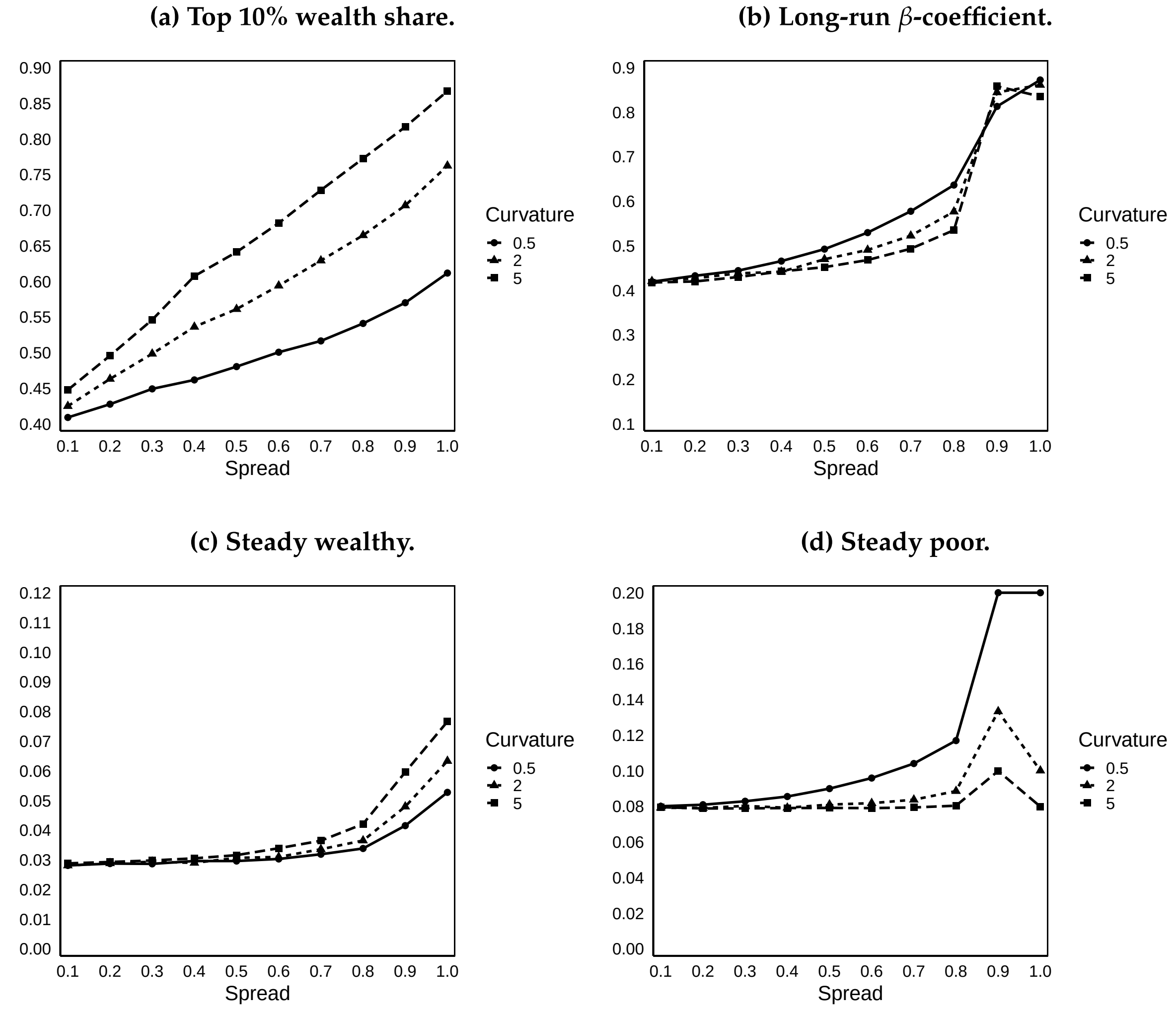


(d) Steady poor.



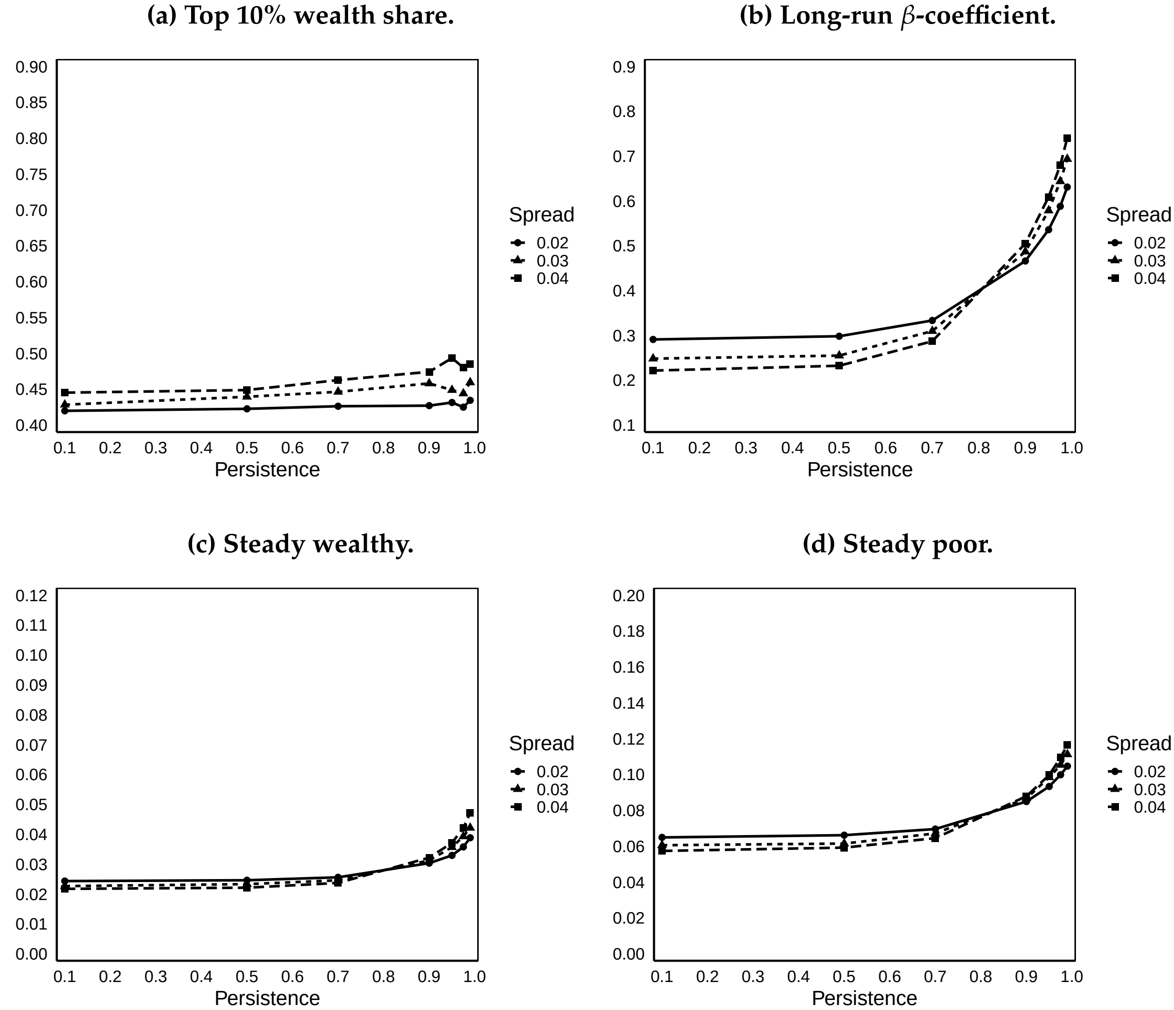
Note: this figure plots the stationary wealth inequality and wealth mobility outcome metrics (defined in Section 3.3) across several simplified heterogeneous agent models. The models depart from the simplified framework outlined in Section 4.1 and additionally abstract from saving ratio type dependence. The models are simulated across different $\{C^{\theta}, \Delta^{(s)\theta}\}$-combinations. The scale dependence curvature C^{θ}-values are displayed on the x-axis, and each line represents a scale dependence spread $\Delta^{(s)\theta}$-value.

Figure 7: Scale dependence spread and stationary wealth inequality and mobility moments across three curvature values (for saving ratios θ).



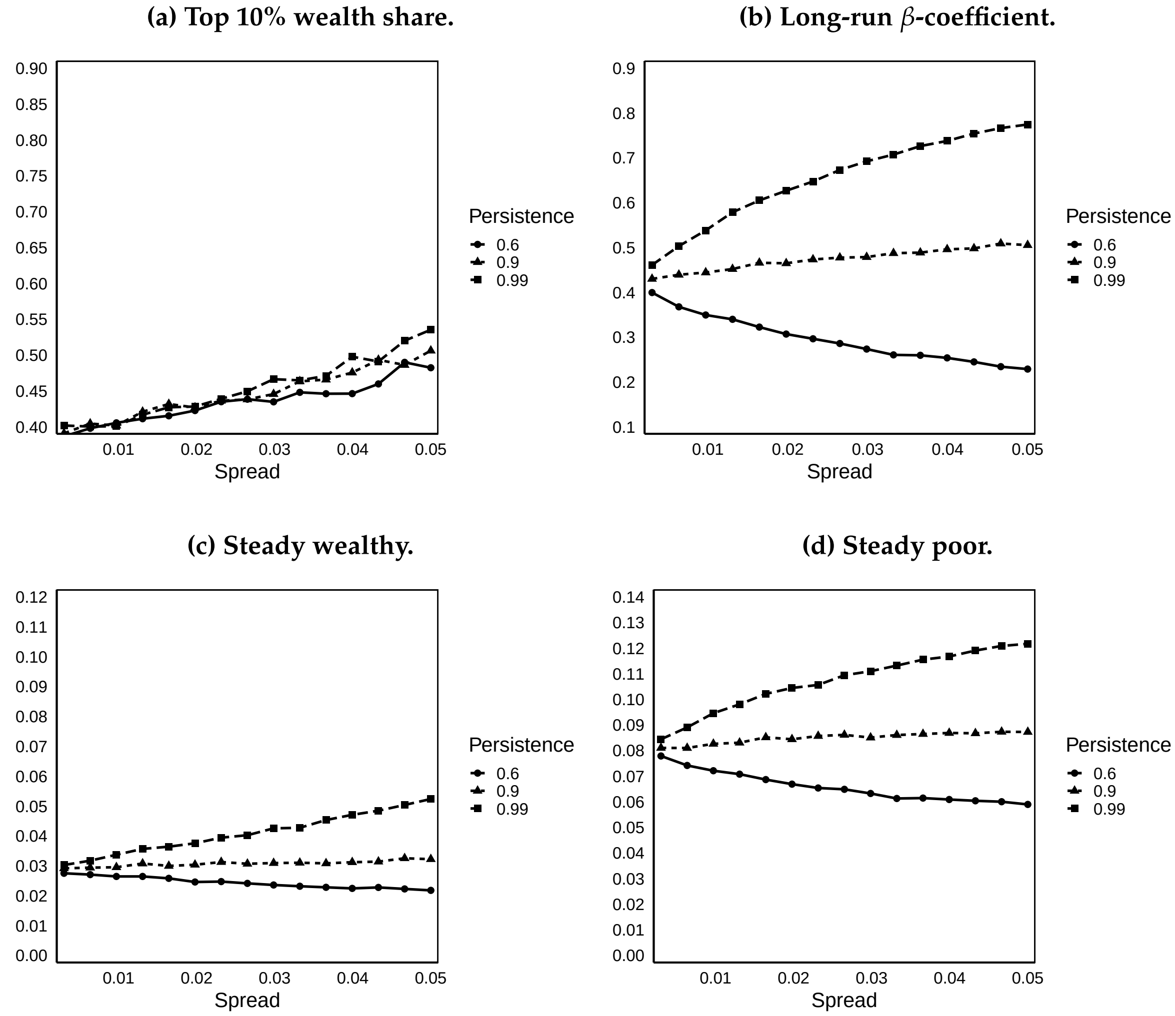
Note: this figure plots the stationary wealth inequality and wealth mobility outcome metrics (defined in Section 3.3) across several simplified heterogeneous agent models. The models depart from the simplified framework outlined in Section 4.1 and additionally abstract from saving ratio type dependence. The models are simulated across different $\{\Delta^{(s)\theta}, C^{\theta}\}$-combinations. The scale dependence spread $\Delta^{(s)\theta}$-values are displayed on the x-axis, and each line represents a scale dependence curvature C^{θ}-value.

Figure 8: Type dependence persistence and stationary wealth inequality and mobility moments across three spread values (for saving ratios θ).



Note: this figure plots the stationary wealth inequality and wealth mobility outcome metrics (defined in Section 3.3) across several simplified heterogeneous agent models. The models depart from the simplified framework outlined in Section 4.1 and additionally abstract from saving ratio scale dependence. The models are simulated across different $\{\rho^{\theta}, \Delta^{(t)\theta}\}$-combinations. The type dependence persistence ρ^{θ}-values are displayed on the x-axis, and each line represents a type dependence spread $\Delta^{(t)\theta}$-value.

Figure 9: Type dependence spread and stationary wealth inequality and mobility moments across three persistence values (for saving ratios θ).



Note: this figure plots the stationary wealth inequality and wealth mobility outcome metrics (defined in Section 3.3) across several simplified heterogeneous agent models. The models depart from the simplified framework outlined in Section 4.1 and additionally abstract from saving ratio scale dependence. The models are simulated across different $\{\Delta^{(t)\theta}, \rho^{\theta}\}$-combinations. The type dependence spread $\Delta^{(t)\theta}$-values are displayed on the x-axis, and each line represents a type dependence persistence ρ^{θ}-value.

ρ^{θ} is found at around 0.50 and 0.60, although the exact value interacts with the spread $\Delta^{(t)\theta}$: a higher spread produces a lower tipping point for ρ^{θ}[5].

These results are straightforward to interpret intuitively: for higher ρ^{θ}, agents are expected to remain of the same saving ratio type for longer consecutive time durations. As a result, high-type agents accumulate more substantial wealth (and vice versa for low-type agents). This generates a wider wealth distribution, and therefore greater absolute wealth differences between agents. These larger absolute differences in turn cause relative wealth mobility to decline, both at the bottom and at the top of the wealth distribution.

Second, a higher type dependence spread $\Delta^{(t)\theta}$ brings about moderately higher wealth inequality (Figure 9, panel a). Its impact on wealth mobility depends critically on the type dependence persistence ρ^{θ}, however. For low persistence ρ^{θ}, a higher $\Delta^{(t)\theta}$ produces higher wealth mobility outcomes. On the contrary, the inverse relationship holds for high persistence ρ^{θ}: a higher $\Delta^{(t)\theta}$ produces lower wealth mobility outcomes (Figure 9, panel b). Under both settings, wealth mobility at the bottom and at the top move in the same direction (Figure 9, panels c and d).

What drives these findings? They relate back to the tipping points from Figure 6. On the one hand, when ρ^{θ} takes on a value below the tipping point (at $0.50 - 0.60$), the persistence in the type-dependent terms is insufficiently strong to consolidate agents' positions in the wealth distribution in the long-run. In other words, there exists strong mean-reverting behavior. An increase in the spread $\Delta^{(t)\theta}$ then raises wealth inequality, but lowers long-run wealth mobility. On the other hand, when ρ^{θ} exceeds the tipping point, the inverse holds: the type-dependent term contributes to the existence of persistence along the wealth (rank) distribution. A higher $\Delta^{(t)\theta}$ reinforces this effect.

Two implications The results from Figures 6 to 9 yield two fundamental implications. First, the relationship between stationary wealth mobility and the scale-dependent and type-dependent model parameters is characterized by non-linearities. For example, a higher scale dependence curvature C^{θ} produces lower wealth mobility only when the scale dependence spread $\Delta^{(s)\theta}$ obtains values close to one ('example a'). Furthermore, the relationship between the type dependence spread $\Delta^{(t)\theta}$ and wealth mobility depends critically on type persistence ρ^{θ}: the relationship is negative for high ρ^{θ} and positive for low ρ^{θ} ('example b'). Second, although there exists a negative link between wealth inequality and wealth mobility under most parameter settings, this is not by definition the case: in examples (a) and (b), higher C^{θ} and $\Delta^{(t)\theta}$ generate higher wealth inequality, but their effect on wealth mobility is not unambiguously negative. Instead, it depends on the underlying scale-dependent and type-dependent parameter combinations.

[5] In addition, the tipping point is model-dependent. Consequently, it may obtain different values in more sophisticated heterogeneous agent models, such as the one outlined in Section 5 of this paper.

5 A model-based estimation of type and scale dependence

The generalized theoretical framework of Section 2 posits that structural heterogeneities in state variable parameters ('source 2') and policy variables ('source 3') result from the interplay between type dependence and scale dependence. Section 4 showed that such distinction between type dependence and scale dependence is relevant: the mixture of scale dependence and type dependence imposed on the model critically affects wealth mobility. A model that aims to jointly match wealth inequality and mobility moments should therefore rely on a realistic degree of type versus scale dependence. However, as underscored in Section 4, type-dependent and scale-dependent worlds can generate highly similar outcome patterns in the empirical data. This renders identification of the degree of type and scale dependence challenging.

In this Section, I proceed in three steps. First, I introduce a full-fledged Aiyagari-Bewley-Huggett heterogeneous agent model where households have non-homothetic preferences and entrepreneurs are modeled along the lines of Cagetti & De Nardi (2006). Second, I present a novel estimation strategy that uses the heterogeneous agent model in combination with data from the Panel Study of Income Dynamics (PSID) to estimate internally the scale-dependent functions and type-dependent stochastic structure for household saving ratios and household portfolio allocation. The innovation of this strategy lies in the linkage it creates between a theoretical scale-dependent function and its type-dependent structure, which is computed from the empirical data. This association is shown to be critical in producing a stationary model state that matches empirical wealth mobility outcomes. Third, I apply the estimation strategy to models relying on equal and unequal playing field settings, and subsequently come up with a baseline model that replicates well 2021 wealth inequality and wealth mobility for the United States (which are untargeted model moments). This baseline model will be used in Section 6 to compute counterfactual wealth distributions.

5.1 A heterogeneous agent model

In Section 4, I have presented heterogeneous agent models that rested on highly simplified assumptions. Instead, in what follows, I introduce a full-fledged heterogeneous agent model.

State variable processes The model uses the state variable processes and their estimation outlined in Section 3 of the paper. Log labor income obeys an AR(1)-process augmented with a Pareto distribution at the top (Equations 19 and 20). Equity and housing returns are normally distributed and are determined by idiosyncratic and aggregate risk (Equations 21 and 22). Business returns are taken from an empirical business return PSID-sample Γ to guarantee a match with higher-order return moments. Finally, the riskless return r^f is deterministic, and taxes are computed using the NBER tax simulator program. I allow for both state variable parameter settings: agents may operate in an equal playing field (no structural heterogeneity in state variable parameters) or in an unequal playing field (structural heterogeneity in equity and business returns via scale dependence) setting. The estimation of these processes follows Table 1.

Policy variables households The scale-dependent saving ratio function f^{θ} for households ($x = 0$) represents the optimal policy from solving an intertemporal utility maximization problem. That is, for households ($x = 0$):

$$f^{\theta}\left(w_{i,t}, S_{i,t} \,\middle|\, \Xi_{i,t}, \Lambda_{i,t}\right) = \max_{\{\theta_{i,t}\}_{t=0}^{\infty}} \mathbb{E}_t \sum_{t}^{\infty} \beta^t \left(\frac{c_{i,t}^{1-\gamma^c}}{1-\gamma^c} + \varphi \frac{w_{i,t+1}^{1-\gamma^w}}{1-\gamma^w} \right) \tag{37}$$

subject to the household budget constraint (Equation 1) and the state variable processes (Section 3). Households display CRRA-utility over consumption in combination with a preference for wealth (capitalist motive). In Equation 37, β denotes the discount factor, γ^c the risk aversion regarding consumption, γ^w the risk aversion with respect to wealth, and φ the households' preference for wealth parameter. All households i maximize the same intertemporal maximization problem, as noted in Section 2.3. In addition, agents incorporate the uncertainty in state variables $\in S$ (rational expectations), but take their state variable parameters $\in \Xi$ as given (naive expectations). I impose that $\beta = 0.94$, in line with for example Hubmer et al. (2021). Furthermore, I suppose that $\gamma^w = 0.25$ and $\gamma^c = 2$, meaning that wealth is a luxury good and preferences are non-homothetic (in line with e.g. Atkinson, 1971; De Nardi, 2004; Michau et al., 2023). The preference for wealth parameter φ is estimated internally, as detailed in Section 5.2. I do not impose any structure on type-dependent saving ratio terms yet as these will be estimated from the data (also described in Section 5.2). The maximization problem in Equation 37 is resolved numerically using the endogenous grid method (Carroll, 2006). Grids were discretized using Rouwenhorst (1995).

The solution to the optimization problem in Equation 37 generates a scale-dependent saving ratio function. It was obtained conditional on the household's other policy variables $\in \Lambda$. However, in addition to saving ratios, households' policy variable set also contains equity and housing participation and allocation variables. In principle, it would be desirable to depart from an optimization problem that yields a joint solution to all variables $\in \Lambda$, rather than solely the saving ratio θ (as was assumed in Section 2.3). Unfortunately, such joint solution for saving ratios, asset participation and asset allocation suffers from numerical instabilities and empirical contradictions, even when using richer utility specifications (e.g. Epstein-Zin preferences). Instead, I therefore impose simplifying behavioral assumptions on the risky asset participation and portfolio allocation policy variables. I outline these assumptions in two steps.

First, I suppose that the scale-dependent (conditional) entry and exit probabilities for equity and housing participation relate solely to wealth w and are equal to the relationship between entry and exit probabilities and wealth deciles d observed in the empirical data (Van Langenhove, 2025c). That is:

$$f^{p^{e,e}}\left[w_{i,t}, S_{i,t} \,\middle|\, \Xi_{i,t}\right] = f^{p^{e,e}}\left[w_{i,t}\right] = \bar{f}^{p^{e,e}}\left[d_{i,t}\right] \tag{38}$$
$$f^{p^{e,o}}\left[w_{i,t}, S_{i,t} \,\middle|\, \Xi_{i,t}\right] = f^{p^{e,o}}\left[w_{i,t}\right] = \bar{f}^{p^{e,o}}\left[d_{i,t}\right] \tag{39}$$

and:

$$f^{p^{h,e}}\left[w_{i,t}, S_{i,t} \,\middle|\, \Xi_{i,t}\right] = f^{p^{h,e}}\left[w_{i,t}\right] = \bar{f}^{p^{h,e}}\left[d_{i,t}\right] \tag{40}$$
$$f^{p^{h,o}}\left[w_{i,t}, S_{i,t} \,\middle|\, \Xi_{i,t}\right] = f^{p^{h,o}}\left[w_{i,t}\right] = \bar{f}^{p^{h,o}}\left[d_{i,t}\right] \tag{41}$$

In addition, for now, I posit the absence of type dependence in equity and housing entry and exit probabilities:

$$\varepsilon_{i,t}^{p^{e,e}} = 0 \quad \forall i, t \tag{42}$$
$$\varepsilon_{i,t}^{p^{e,o}} = 0 \quad \forall i, t \tag{43}$$

and:

$$\varepsilon_{i,t}^{p^{h,e}} = 0 \quad \forall i, t \tag{44}$$
$$\varepsilon_{i,t}^{p^{h,o}} = 0 \quad \forall i, t \tag{45}$$

Second, equity and housing (conditional) portfolio shares are assumed to relate solely to wealth levels w. Moreover, they are also imposed to match the relationship between these shares and wealth deciles d in the empirical data (Van Langenhove, 2025c), but multiplied by scale-dependent parameters k^e and k^h. Algebraically:

$$f^{\alpha^e}\left[w_{i,t}, S_{i,t} \,\middle|\, \Xi_{i,t}\right] = f^{\alpha^e}\left[w_{i,t}\right] = k^e \bar{f}^{\alpha^e}\left[d_{i,t}\right] \tag{46}$$
$$f^{\alpha^h}\left[w_{i,t}, S_{i,t} \,\middle|\, \Xi_{i,t}\right] = f^{\alpha^h}\left[w_{i,t}\right] = k^h \bar{f}^{\alpha^h}\left[d_{i,t}\right] \tag{47}$$

where k-parameters will be estimated internally, as outlined in Section 5.2. Moreover, I do not impose a structure on type-dependent portfolio allocation terms yet: these will be estimated from the data.

Policy variables entrepreneurs Determining the scale-dependent functions for the policy variables (saving ratios θ and business portfolio shares α^b) of entrepreneurs ($x = 1$) through the solution of an optimization problem is highly complicated. That is, for entrepreneurs, the solution to Equation 37 yields an optimal saving ratio very close to one across all wealth levels. This stems from high idiosyncratic business risk in combination with a precautionary saving motive. While there exists empirical evidence that entrepreneurs save more for given wealth levels in the United States (e.g. Van Langenhove, 2025b), their saving ratios still lay below one. Unfortunately, it is difficult to empirically investigate the exact relationship between entrepreneurial saving ratios and wealth levels w as representative data for entrepreneurs is available only at the top of the wealth (rank) distribution. I therefore impose the two simplifying assumptions on the heterogeneous agent model.

First, I abstract from saving ratio heterogeneity for entrepreneurs: the scale-dependent saving ratio term for entrepreneurs is equal to a constant $\tilde{\theta}$, while the type-dependent term is (for now) set at zero across all entrepreneurs i. That is, for entrepreneurial agents ($x = 1$):

$$f^{\theta}\left[w_{i,t}, S_{i,t} \,\middle|\, \Xi_{i,t}\right] = \tilde{\theta}^{b} \tag{48}$$

$$\varepsilon^{\theta}_{i,t} = 0 \quad \forall i, t \tag{49}$$

which is similar to e.g. Gomez & Gouin-Bonenfant (2024). Second, I make identical assumptions for the business portfolio share α^b: business portfolio shares are homogeneous across all entrepreneurs. Specifically, the scale-dependent business portfolio share equals a constant $\tilde{\alpha}^b$, while the type-dependent term is zero across all agents i. Algebraically:

$$f^{\alpha^b}\left[w_{i,t}, S_{i,t} \,\middle|\, \Xi_{i,t}\right] = \tilde{\alpha}^{b} \tag{50}$$

$$\varepsilon^{\alpha^b}_{i,t} = 0 \quad \forall i, t \tag{51}$$

where I calibrate $\tilde{\alpha}^b$ to the median share of their portfolio that entrepreneurs allocate to business or equity assets in the SCF-sample of Van Langenhove (2025c). This produces $\tilde{\alpha}^b = 0.55$.

Agent type − entrepreneurship A final policy variable $\in \Lambda$ are the transition probabilities between households and entrepreneurs $p^{x,e}$ and $p^{x,o}$. For now, I assume that the scale-dependent entry and exit probabilities relate solely to wealth levels w and follow their empirical schedule (Van Langenhove, 2025c). Moreover, I abstract from type dependence in these variables. Algebraically:

$$f^{p^{x,e}}\left[w_{i,t}, S_{i,t} \,\middle|\, \Xi_{i,t}\right] = f^{p^{x,e}}\left[w_{i,t}\right] = \bar{f}^{p^{x,e}}\left[d_{i,t}\right] \tag{52}$$

$$f^{p^{x,o}}\left[w_{i,t}, S_{i,t} \,\middle|\, \Xi_{i,t}\right] = f^{p^{x,o}}\left[w_{i,t}\right] = \bar{f}^{p^{x,o}}\left[d_{i,t}\right] \tag{53}$$

$$\varepsilon^{p^{x,e}}_{i,t} = 0 \quad \forall i, t \tag{54}$$

$$\varepsilon^{p^{x,o}}_{i,t} = 0 \quad \forall i, t \tag{55}$$

which is similar to the assumptions imposed on household equity and housing participation transitions.

However, both under the equal and unequal playing field settings from Section 3, the concentration of entrepreneurs at the top significantly underestimates the concentration observed in the empirical data. This holds even when imposing extreme entrepreneurial saving ratios $\tilde{\theta}^b$ close to one for all entrepreneurial agents. More precisely, the share of entrepreneurs among the top 10% wealthiest ranges between around 14% to 20% across different model specifications, compared to over 30% in the data. In response to this, one can think of two adaptations. First, one could raise the business returns observed in the data Λ^b with a scalar b up to the point where the spread of entrepreneurs across the wealth (rank) distribution in the stationary model states matches its empirical counterparts. However, the premium over empirically observed returns that such strategy requires is very large: it takes on values over 25%. While some underestimation of empirical business returns is likely, it is unlikely to be this large. Second, one can multiply the entrepreneurial entry (exit) rates across the wealth rank distribution by a scalar a ($1/a$). This does generate a realistic spread of entrepreneurs over the wealth distribution, even for minimal changes. I therefore apply this method. The aggregate entrepreneurship rate in the stationary model state (at 10%) now lies slightly above its empirical counterpart (9% in Van Langenhove, 2025c). The excess entrepreneurs in the model are primarily concentrated in the middle part of the wealth distribution. They therefore comprise entrepreneurs whose businesses have not taken off yet or have become unsuccessful.

5.2 Estimation strategy

The construction of the model and external estimation exercise leaves a number of parameters that need to be estimated internally within the model. First, for household saving ratios, the preference for wealth parameter φ and structure behind the type-dependent term ε^{θ} require estimation. Second, for household portfolio allocation, k^e and k^h need to be estimated, as do the structures behind type-dependent terms ε^{α^e} and ε^{α^h}. Third, for entrepreneurs, the homogeneous saving ratio $\tilde{\theta}^b$ requires estimation. In what follows, I estimate these three groups of free model parameters to match a set of target variables. I leverage a panel dataset from the Panel Study of Income Dynamics (PSID) to create an empirical link between scale-dependent functions and type-dependent term structures. In the PSID-sample, I define households as agents without business assets, and create a sample of households aged 25 to 64 that have at least three observations for all state variables $\in S$ and policy variables $\in \Lambda$ over the lifecycle. I exclude observations in years where a household reports to have received a lumpsum payment or inter-generational transfer.

The estimation strategy is set up in three steps. Together, these steps link each scale-dependent parameter (φ, k^e, k^h) with a type-dependent structure estimated from the data. First, I solve the optimization problem in Equation 37 over a φ-grid, obtaining a grid of candidate scale-dependent functions $\hat{f}^{\theta}$. In addition, I compute candidate scale-dependent functions $\hat{f}^{\alpha^e}$ and $\hat{f}^{\alpha^h}$ across all values in an k^e-grid and k^h-grid. Second, for each of the candidate scale-dependent functions, I compute the residuals e between the saving ratios and portfolio shares observed in the PSID and the ones predicted by the candidate scale-dependent function. Algebraically:

$$e^{\theta}_{i,t} = \tilde{\theta}_{i,t} - \hat{f}^{\theta}\left[\tilde{w}_{i,t}, \tilde{S}_{i,t}\right] \tag{56}$$

$$e^{\alpha^e}_{i,t} = \tilde{\alpha}^{e}_{i,t} - \hat{f}^{\alpha^e}\left[\tilde{w}_{i,t}, \tilde{S}_{i,t}\right] \tag{57}$$

$$e^{\alpha^k}_{i,t} = \tilde{\alpha}^{k}_{i,t} - \hat{f}^{\alpha^k}\left[\tilde{w}_{i,t}, \tilde{S}_{i,t}\right] \tag{58}$$

where $\tilde{z}$ refers to the value of a variable z observed in the PSID-sample. As a result, $\tilde{\theta}$, $\tilde{\alpha}^{e}_{i,t}$ and $\tilde{\alpha}^{k}_{i,t}$ represent the saving ratio and portfolio shares observed in the data. This second step has generated a grid that links each candidate scale-dependent function to a corresponding sample of residuals. Third, for each element in this grid, I discretize the residuals e over the agents in the sample to obtain a discrete-state residual grid and a Markov transition matrix. Together, the grid and transition matrix approximate the type-dependent structure for the respective variable (θ, α^e, α^h). To conclude, our final grid links each candidate scale-dependent function to corresponding type-dependent structure that is estimated from the empirical data.

Having created a grid also for the entrepreneurial saving ratio $\tilde{\theta}^b$, I subsequently estimate the free parameters (φ, k^e, k^h and $\tilde{\theta}^b$) using a method of simulated moments (MSM) estimator. The MSM estimation targets include (1) the saving ratio levels and dispersion across the wealth (rank) distribution (Van Langenhove, 2025a), (2) the equity and housing portfolio share levels and dispersion across the wealth (rank) distribution (Van Langenhove, 2025c), and (3) the ratio of aggregate wealth held by entrepreneurs to total aggregate wealth, which equals 0.35 in a PSID-sample. Let us denote the set of target variables as Ω. Algebraically, the MSM-estimator minimizes:

$$\{\varphi\ ,\ k^e,\ k^h,\ \hat{\alpha}^b\} = \arg \min_{\varphi, k^e, k^h, \hat{\alpha}^b} \sum_{z \in \Omega} \left[\bar{z}_z\big(\varphi,\ k^e,\ k^h,\ \hat{\alpha}^b\big) - \tilde{z}_z\right]^2 \tag{59}$$

where I have denoted by $\bar{z}$ the model-generated target moments and as $\tilde{z}$ their empirical counterparts.

How does this estimation strategy compare to the literature? First, existing heterogeneous agent models depart from type dependence or scale dependence in parameters or policy variables without establishing a realistic, empirically-driven linkage between the two dependence types (e.g. Fernandez-Villaverde & Levintal, 2024; Gaillard & Wangner, 2023; Xavier, 2021). On the contrary, the estimation strategy in Equations 56-59 guarantees that a theoretical scale-dependent function is related to a realistic, empirically-driven type-dependent structure. This is critical in matching wealth mobility outcomes. Second, as argued in Section 4.2, there have been multiple attempts to estimate type dependence versus scale dependence based on empirical data using fixed effects models (e.g. Bach et al., 2020; Fagereng et al. 2020; Hurst & Lusardi, 2004). In contrast, the estimation strategy in Equations 56-59 is model-based.

5.3 Estimation results

In this subsection, I discuss stationary heterogeneous agent model outcomes using the external and internal estimation from Sections 5.1 and 5.2. I distinguish between the equal and unequal state variable parameter settings.

Under the equal playing field setting, applying the estimation procedure from Section 5.2 yields $\varphi = 0.06$, $k^e = 0.95$, $k^h = 0.95$, $\tilde{\theta}^b = 0.945$. Furthermore, the estimated type-dependent structures are characterized by their spread and persistence:

$$\Delta^{(t)\theta} = 0.019, \quad \rho^{\theta} = 0.21, \quad \Delta^{(t)\alpha^e} = 0.037, \quad \rho^{\alpha^e} = 0.27, \quad \Delta^{(t)\alpha^h} = 0.07, \quad \rho^{\alpha^h} = 0.29$$

The stationary model outcome replicates 2021 U.S. wealth inequality and wealth mobility well, although it underestimates the degree of wealth mobility (Table 3, column 2, 'equal'). First, the top 10% wealth share in the equal playing field setting model matches exactly the one observed in the data (at 0.77). The model does overestimate the bottom 50% wealth share (0.03) relative to the data (0.00). This stems from the absence of indebtedness in the model. Second, both short-run and long-run wealth mobility in the model are somewhat lower than in the data: the short-run rank-rank coefficient equals 0.89 in the model (compared to 0.84 in the data), while the long-run rank-rank coefficients amounts to 0.48 (versus 0.40 in the data). The underestimation of wealth mobility occurs primarily at the top.

Under the unequal playing field setting, I derive the stationary states for two models. On the one hand, I keep the free parameter estimation from the equal playing field setting, but allow for structural heterogeneity in equity and business returns (Table 3, column 3, 'unequal 1'). The top 10% wealth share now lays slightly higher, at 0.78. Wealth mobility remains roughly identical to the equal setting. On the other hand, I re-estimate the free parameters (φ, k^e, k^h, $\tilde{\theta}^b$) to the unequal setting. Part of the empirically observed saving ratio inequality follows from structurally unequal equity returns across the wealth (rank) distribution. This lowers the preference for wealth estimate to $\varphi = 0.04$. Similarly, the structural heterogeneity in business returns across the wealth (rank) distribution generates a slightly lower entrepreneurial saving ratio: $\tilde{\theta}^b = 0.94$. The other two internally estimated parameters take on the same values as before: $k^e = 0.95$ and $k^h = 0.95$. Moreover, the estimated type-dependent structure is close to identical to the equal setting. The stationary model yields similar results as the equal playing field setting (Table 3, column 4, 'unequal 2'). This implies that also the unequal playing field setting model underestimates empirical wealth mobility outcomes.

What explains the underestimation of wealth mobility in the stationary model state relative to the empirical data? There are three candidate explanations, which relate to the simplifying assumptions from Section 5.1. First, I have imposed that the structural heterogeneity in equity and housing entry and exit rates relates entirely to scale dependence. In practice, there might also exist type dependence in these variables. Introducing type dependence would increase

Table 3: Stationary model outcomes across an equal playing field setting, two unequal playing field settings, and a baseline model.

	Data	Equal	Unequal 1	Unequal 2	Baseline
Wealth inequality					
Bottom 50%	0.00	0.03	0.03	0.03	0.03
Middle 50–90%	0.23	0.21	0.20	0.21	0.18
Top 10%	0.77	0.76	0.78	0.77	0.79
Wealth mobility (short–run)					
Short–run β	0.84	0.89	0.88	0.88	0.86
Steady wealthy	0.08	0.08	0.08	0.08	0.08
Steady poor	0.13	0.13	0.13	0.13	0.13
Wealth mobility (long–run)					
Long–run β	0.40	0.48	0.49	0.47	0.44
Steady wealthy	0.03	0.04	0.04	0.04	0.04
Steady poor	0.09	0.08	0.08	0.07	0.07

Note: this table shows the stationary wealth inequality and mobility outcomes across four heterogeneous agent models (equal playing field setting, unequal playing field setting 1, unequal playing field setting 2, baseline model). The models are detailed in the main text. The column 'Data' shows the wealth inequality and wealth mobility outcomes in the empirical data, as outlined in Section 3.3. The models have been estimated using the external estimation described in Section 5.1, and the internal estimation strategy from Section 5.2. The 'unequal 1' model relies on the same estimation values for the free parameters as the 'equal' model.

wealth mobility (as shown in Section 4.4). The same reasoning applies to the attribution of household-entrepreneur transition probability heterogeneity to scale dependence-only. Second, entrepreneurial saving ratios and business portfolio shares are assumed homogeneous across agents. In practice, entrepreneurs might be structurally different, in part because of type dependence in their saving ratios and portfolio allocation. Third, in Section 3, I have imposed that the heterogeneity in equity and business returns across the wealth (rank) distribution stems entirely from scale dependence. Also here type dependence is likely to be present: some agents may be structurally better equity investors or entrepreneurs than others. In summary, introducing type dependence in these three policy variables and two state variable parameters is likely to generate more turnover across the stationary wealth distribution. I turn to this next.

5.4 A baseline model

Both equal and unequal playing field setting models replicate relatively well 2021 U.S. wealth inequality, but underestimate U.S. wealth mobility outcomes by around seven to nine points based on the long-run rank-rank coefficient. In what follows, I quantify the contribution of the simplifying assumptions of (i) scale dependence-only in equity, housing and entrepreneurship entry and exit probabilities, (ii) homogeneous entrepreneurial saving ratios and portfolio allocation, and (iii) scale dependence-only in expected equity and business returns to this result. I then present a baseline model.

Such quantification exercise is complicated by the absence of an obvious way of linking candidate scale-dependent functions for these variables to an empirical type-dependent term structure. We therefore have to resort to an ad-hoc approach that is only partially data-driven. For all variables in (i)-(iii), I introduce type dependence according to three principles. First, I impose on these variables a type dependence spread $\Delta^{(t)}$ that equals 20% of the total observed variance in that underlying variable. This is in line with the (average) ratio of saving ratio type dependence spread to the total saving ratio variance observed in models of Section 5.3. Second, I set these variables' type dependence persistence parameter ρ to 0.20. This corresponds roughly to the saving ratio type dependence in the equal and unequal setting models. Third, I multiply the scale-dependent functions that were imposed for variables (i)-(iii) in Section 5.1 by a variable-specific scalar k. Conditional on the ad-hoc type-dependent structure, these scalars k are estimated to match the empirical values of the variables across the wealth (rank) distribution (in the PSID). This approach generates identical free parameter estimates as in the unequal playing field setting model. That is, it holds that: $\varphi = 0.04$, $\tilde{\theta}^b = 0.94$, $k^e = 0.95$ and $k^h = 0.95$. The k-scalars of variables (i)-(iii) obtain values close to one. The type-dependent summary metrics are also similar to the models of Section 5.3:

$$\Delta^{(t)\theta} = 0.016, \quad \rho^{\theta} = 0.21, \quad \Delta^{(t)\alpha^e} = 0.037, \quad \rho^{\alpha^e} = 0.27, \quad \Delta^{(t)\alpha^h} = 0.07, \quad \rho^{\alpha^e} = 0.29$$

The wealth inequality and wealth mobility outcomes are displayed in Table 3, column 5 ('baseline'). Two key findings persist. First, top wealth inequality is slightly higher compared to the equal and unequal playing field models: the top 10% wealth share rises to 0.79 (compared to 0.77 in the empirical data). Second, short-run and long-run wealth mobility increase compared to the models from Section 5.3: the short-run rank-rank coefficient drops to 0.86 (compared to 0.84 in the data), while the long-run rank-rank coefficients declines to 0.44 (compared to 0.40 in the data). Introducing type dependence in variables (i)-(iii) therefore brings wealth mobility outcomes in the heterogeneous agent model closer to their empirical counterparts. However, a gap of four points remains. In what follows, I label the model from Table 3, column 5 the 'baseline model'. I will use it as a starting point to compute counterfactual wealth distributions in Section 6. Appendix C visualizes entrepreneurship, saving ratios, equity and housing participation rates and equity and housing portfolio shares across the wealth (rank) distribution. These variables match empirical data patterns well.

6 Sources of U.S. wealth inequality & wealth mobility

The previous Section has estimated a heterogeneous agent model with structural heterogeneity in expected equity and business returns and structural heterogeneity in all policy variables $\in \Lambda$. In this Section, I use the estimated baseline model to investigate the channels generating wealth inequality and wealth mobility. I do so by shutting down structural agent heterogeneities and computing counterfactual wealth distributions. I distinguish between three types of channels: (1) labor income inequality and taxation, (2) saving ratio inequality, (3) household asset allocation, asset participation and asset return channels.

6.1 Labor income inequality & taxation

In a first counterfactual exercise ('labor income inequality'), I shut down labor income inequality: all household agents have an identical labor income equal to the 2021 average ('model M1'): $y_{it} = y_t$. The top 10% labor income share is equal to 0.10. The counterfactual wealth distribution displays lower wealth inequality and significantly higher wealth mobility: the top 10% wealth share drops to 0.67 (from 0.79 in the baseline), while the long-run rank-rank coefficient declines to 0.12 (compared to 0.44 in the baseline) (Table 4, column M1). The strong rise in wealth mobility holds at the bottom and at the top of the wealth distribution: both the fraction of steady wealthy and steady poor decline relative to the baseline. The rise in wealth mobility also appears in the short-run: the short-run rank-rank coefficient falls to 0.60 (compared to 0.86 in the baseline).

Hence, the baseline heterogeneous agent model predicts a strong, negative relationship between labor income inequality and wealth mobility, which is a novel result in the heterogeneous agent literature. Another difference between my findings and existing literature relates to the impact of labor income inequality on wealth inequality. Existing work found labor income inequality to be the main contributor to wealth inequality (e.g. Hubmer et al., 2021; Kaymak et al., 2022). However, counterfactual M1 in Table 4 generates only a moderate drop in the stationary top 10% wealth share. This discrepancy relates to a difference in the underlying counterfactual exercise: existing studies shut down labor income inequality for both households and entrepreneurs, while I do so only for households. Consequently, in model M1, the continued presence of entrepreneurs produces significant wealth inequality[6].

In a second counterfactual exercise ('taxation'), I abstract from taxation ('model M2'). Stationary wealth inequality increases by five points relative to the baseline, while short-run and long-run rank-rank coefficients decline by two points (Table 4, column M2). In other words, if there were no taxation, both stationary wealth inequality and wealth mobility would obtain higher values. The impact of taxation on wealth inequality in model M2 is weaker than in for instance Hubmer et al. (2021), who find that the U.S. tax system is strongly progressive

[6]To illustrate this point: the aggregate wealth held by entrepreneurs accounts for approximately 60% of total aggregate wealth in model M1. Instead, in the baseline, this ratio was equal to 35%.

Table 4: Counterfactual stationary wealth inequality and wealth mobility: labor income inequality and taxation.

	Data	Baseline	M1	M2
Wealth inequality				
Bottom 50%	0.00	0.03	0.02	0.01
Middle 50–90%	0.23	0.18	0.31	0.15
Top 10%	0.77	0.79	0.67	0.84
Wealth mobility (short–run)				
Short–run β	0.84	0.86	0.60	0.84
Steady wealthy	0.08	0.08	0.06	0.08
Steady poor	0.13	0.13	0.07	0.13
Wealth mobility (long–run)				
Long–run β	0.40	0.44	0.12	0.42
Steady wealthy	0.03	0.04	0.02	0.04
Steady poor	0.09	0.07	0.05	0.07

Note: this table shows the stationary wealth inequality and mobility outcomes for two counterfactuals. M1 represents the counterfactual for labor income inequality, M2 the counterfactual for taxation. The column 'Data' shows the wealth inequality and wealth mobility outcomes in the empirical data, as outlined in Section 3.3. The column 'Baseline' shows the baseline model outcomes from Table 3. The models have been calibrated using the external calibration described in Section 5.1, and the internal calibration strategy from Section 5.2.

and lowers the top 10% wealth share by a little over twenty points. The divergence between both models suggests that the NBER tax simulator (as outlined in Section 3.1) may not entirely capture the progressivity of the U.S. tax system. This could relate to the exclusion of corporate income taxes and estate taxes in the NBER tax simulator program. I leave a more detailed modeling of the U.S. tax system to future research.

6.2 Saving ratio inequality

In what follows, I shut down various components of saving ratio inequality. I distinguish between saving ratio type dependence, saving ratio scale dependence and overall saving ratio heterogeneity. The results are displayed in Table 5.

In a first step ('type dependence'), I eliminate saving ratio type dependence ('model M3.a'): the type-dependent saving ratio term is set to 0 for all agents i: $\varepsilon^{\theta}_{i_t} = 0 \;\; \forall i, t$. While counterfactual wealth inequality is roughly identical to baseline wealth inequality, wealth mobility increases: the long-run rank-rank coefficient declines to 0.37 (compared to 0.44 in the baseline). This mobility effect is driven entirely by higher wealth mobility at the bottom of the wealth distribution. In any case, this result is (at first glance) at odds with the conclusion from Section 4.4: in that section, the presence of type dependence was found to generate higher (rather than lower) wealth mobility. However, the theoretical insight from Section 4.4 was derived under the assumption of identical saving ratio inequality.

On the contrary, model M3.a abstracted from type dependence, but kept the baseline scale-dependent function f^{θ} (with $\varphi = 0.04$) that was estimated in an environment with both type dependence and scale dependence. This means that in model M3.a, saving ratio inequality is significantly lower than the inequality from the baseline case: in the baseline model, part of the high saving ratios at the top reflect the saving behavior of high-type agents, with the opposite holding at the bottom. In response to this, I re-estimate the preference for wealth φ to match empirical saving ratio inequality in a model without saving ratio type dependence ('model M3.b'). This leads to $\varphi = 0.05$. The top 10% wealth share rises somewhat (0.83 in M3.b compared to 0.79 in the baseline), while long-run wealth mobility declines strongly: the long-run rank-rank coefficient amounts to 0.61, compared to 0.44 in the baseline. In summary, the presence of saving ratio type dependence for given levels of saving ratio inequality generates significantly higher wealth mobility in the stationary model state.

In a second step ('preference for wealth'), I shut down the preference for wealth component: $\varphi = 0$. This changes the curvature of the scale-dependent function f^{θ} in a non-linear way, but does not eliminate saving ratio scale dependence entirely. I distinguish between two models. On the one hand, model M4.a keeps the baseline type-dependent structure. Compared to the baseline model, M4.a leads to a slightly higher top 10% wealth share (0.83 versus 0.79) and significantly higher wealth mobility (0.32 compared to 0.44). On the other hand, model M4.b re-estimates the type-dependent structure using the PSID-sample. The top 10% wealth share

Table 5: Counterfactual stationary wealth inequality and wealth mobility: saving ratio inequality.

	Data	**Baseline**	**M3.a**	**M3.b**	**M4.a**	**M4.b**	**M5.a**	**M5.b**	**M6**
Wealth inequality									
Bottom 50%	0.00	0.03	0.03	0.03	0.02	0.02	0.17	0.17	0.17
Middle 50–90%	0.23	0.18	0.20	0.15	0.15	0.17	0.36	0.36	0.37
Top 10%	0.77	0.79	0.77	0.83	0.83	0.81	0.47	0.47	0.47
Wealth mobility (short–run)									
Short–run β	0.84	0.86	0.81	0.94	0.76	0.76	0.71	0.71	0.67
Steady wealthy	0.08	0.08	0.08	0.09	0.08	0.08	0.06	0.06	0.06
Steady poor	0.13	0.13	0.11	0.16	0.09	0.09	0.11	0.11	0.11
Wealth mobility (long–run)									
Long–run β	0.40	0.44	0.37	0.61	0.32	0.33	0.31	0.31	0.29
Steady wealthy	0.03	0.04	0.04	0.06	0.04	0.04	0.02	0.03	0.02
Steady poor	0.09	0.07	0.07	0.09	0.06	0.06	0.07	0.07	0.07

Note: this table presents stationary wealth inequality and mobility outcomes under four counterfactual scenarios targeting saving ratio heterogeneity. Model M3 removes type dependence in saving behavior, M4 shuts down the preference for wealth in the utility function (Equation 37), M5 eliminates scale dependence in saving ratios, and M6 removes all forms of saving ratio heterogeneity. Models labeled with '.a' retain the calibration structure of the baseline model, while models labeled with '.b' are re-estimated using the internal strategy described in Section 5.2, conditional on the respective counterfactual assumptions. Model M6 combines the external calibration strategy outlined in Section 5.1 with the internal procedure of Section 5.2. The 'data' column displays empirical outcomes as discussed in Section 3.3, and the 'baseline' column reports results from the benchmark model shown in Table 3.

still exceeds the baseline (at 0.81), while long-run wealth mobility increases significantly to above the baseline (at 0.32). The presence of a preference for wealth component ($\varphi > 0$) thus slightly lowers wealth inequality, and lowers wealth mobility outcomes.

In a third step ('scale dependence'), I abstract from scale dependence altogether: the saving ratio scale-dependent function is set to the median saving ratio in the 2001-2021 PSID-sample ($f^{\theta} = 0.56$). I again distinguish between two models: a model maintaining the baseline type-dependent structure ('model M5.a'), and a model that re-estimates it ('model M5.b'). These models yield identical wealth inequality and mobility outcomes: the absence of saving ratio scale dependence significantly lowers top 10% wealth shares (at 0.47). Moreover, short-run and long-run wealth mobility increase: the short-run rank-rank coefficient declines to 0.71 (versus 0.86 in the baseline) and the long-run one to 0.31 (compared to 0.44 in the baseline). Overall, the presence of scale dependence raises wealth inequality and lowers wealth mobility outcomes.

In a fourth step ('all heterogeneity'), I abstract from all structural saving ratio heterogeneity: $\varepsilon^{\theta}_{i,t} = 0 \;\; \forall i, t$ and $f^{\theta} = 0.56$ (model M6). The top 10% wealth share drops to the same level as in models M5.a and M5.b (at 0.47). Moreover, short-run and long-run wealth mobility rise to their highest level across all saving ratio counterfactuals: the short-run rank-rank coefficient equals 0.67 (compared to 0.86 in the baseline), and the long-run rank-rank coefficient amounts to 0.29 (compared to 0.44 in the baseline). Through the interplay of type dependence and scale dependence, structural heterogeneity in saving ratios thus raises U.S. wealth inequality, and lowers wealth mobility.

Implications The saving ratio inequality counterfactuals have three key implications. First, the linkage between scale-dependent functions and type-dependent structure is often critical for stationary model outcomes: imposing an alternative scale-dependent function without re-calibrating the type-dependent structure from the data (or vice versa) can yield outcomes that are at odds compared to when the re-calibration is executed. Second, as noted, saving ratio type dependence raises wealth mobility (M3.b), while saving ratio scale dependence leads to higher wealth inequality and lower wealth mobility (M5). Overall structural saving ratio heterogeneity generates higher wealth inequality and significantly lower wealth mobility (M6). For wealth mobility, the saving ratio scale dependence effect therefore appears to dominate over the type dependence effect, although this may relate also to the interaction behind wealth inequality and wealth mobility. Third, linking scale-dependent functions with a realistic type-dependent structure is critical in matching wealth mobility outcomes: when maintaining realistic saving ratio inequality, the absence of saving ratio type dependence yields a stationary model state where wealth mobility significantly understates its empirical counterpart (M3.b).

6.3 Household asset allocation, participation & returns

I now turn to the counterfactual analyses for various components of household asset allocation, asset participation and asset returns. The results are displayed in Table 6.

In a first counterfactual ('portfolio allocation'), I remove all structural heterogeneity in equity portfolio allocation α^e and housing portfolio allocation α^h ('model M7'). More precisely, these variables are for all households set to the median levels reported in Van Langenhove (2025c): $\alpha_e = 0.22$ and $\alpha_h = 0.62 \; \forall i, t$. In a second counterfactual ('entry and exit probabilities'), I abstract from structural heterogeneity in equity and housing entry and exit probabilities ('model M8'): $p^{e,e} = 0.15$, $p^{e,o} = 0.18$, $p^{h,e} = 0.18$ and $p^{h,o} = 0.09 \;\; \forall i, t$. Both counterfactual M7 and M8 generate wealth inequality and wealth mobility outcomes that are largely indistinguishable from the baseline model. In a third counterfactual, I eliminate the unequal equity returns across the wealth (rank) distribution. Wealth inequality declines somewhat to 0.74 (compared to 0.78 in the baseline), while short-run and long-run wealth mobility each drop two points relative to the baseline.

What explains the limited impact of these return heterogeneity channels on wealth inequality and wealth mobility? There are two channels at play. First, the results reflect the interplay between portfolio allocation to equity and housing: in the stationary model state, allocation to housing is dominant in the middle part of the wealth distribution, while equity is dominant at the top. However, the allocation to their composite − 'risky assets' − remains roughly stable from the middle to the top of the wealth distribution. Second, equity and housing returns were assumed to follow a normal process as opposed to a lognormal process (Section 3.1, Equations 21 and 22). A normal return process guarantees the existence of a stationary model state and allows to match first-order empirical moments, as outlined in Section 3.1. The assumption is also applied in other heterogeneous agent models (e.g. Xavier, 2021). However, a lognormal return process would right-skew returns and is likely to generate a more substantial impact of portfolio allocation and participation probability heterogeneity on wealth inequality and mobility outcomes.

6.4 Wealth inequality and wealth mobility?

Summarizing, the persistence of agents across the wealth (rank) distribution follows from two key sources in our baseline model. First, labor income inequality is critical in creating long-run persistence across the wealth (rank) distribution: without it, the long-run rank-rank coefficient drops to 0.12 from 0.44 in the baseline (model M1). Second, saving ratio heterogeneity is also important: in the absence of such structural heterogeneity, the long-run rank-rank coefficient declines to 0.29 (model M6). This reflects two counteracting forces: the presence of saving ratio scale dependence lowers wealth mobility outcomes, while the presence of saving ratio type dependence raises wealth mobility. Finally, taxation and return heterogeneity have little effect on wealth mobility outcomes. However, the lack of such effects may respectively relate

Table 6: Counterfactual stationary wealth inequality and wealth mobility: portfolio composition, return schedules and asset participation.

	Data	**Baseline**	**M7**	**M8**	**M9**
Wealth inequality					
Bottom 50%	0.00	0.03	0.03	0.02	0.04
Middle 50–90%	0.23	0.18	0.19	0.18	0.23
Top 10%	0.77	0.79	0.79	0.79	0.74
Wealth mobility (short–run)					
Short–run β	0.84	0.86	0.86	0.85	0.84
Steady wealthy	0.08	0.08	0.08	0.08	0.08
Steady poor	0.13	0.13	0.13	0.13	0.13
Wealth mobility (long–run)					
Long–run β	0.40	0.44	0.45	0.42	0.42
Steady wealthy	0.03	0.04	0.04	0.04	0.04
Steady poor	0.09	0.07	0.07	0.07	0.07

Note: this table reports stationary wealth inequality and mobility outcomes for three counterfactual scenarios related to household asset allocation, market participation, and asset returns. M7 eliminates portfolio allocation heterogeneity, M8 shuts down structural heterogeneity in equity and housing entry and exit probabilities, and M9 removes unequal equity returns across the wealth (rank) distribution. The column 'data' shows the wealth inequality and wealth mobility outcomes in the empirical data, as outlined in Section 3.3. The column 'baseline' shows the baseline model outcomes from Table 3.

to the simplistic taxation process that was imposed on the model, and to the assumption of normal as opposed to lognormal returns. I leave an investigation of this to future research.

What do these results imply about the relationship between wealth inequality and wealth mobility in non-simplified heterogeneous agent models? In theory, one would expect a negative relationship: higher wealth inequality implies a larger absolute distance between agents' wealth levels, which is anticipated to generate lower wealth mobility insofar as additive wealth shocks (such as labor income) are important. In general, this inverse relationship holds based on the counterfactual model exercises, although its magnitude depends on the underlying channels. More precisely, labor income inequality and saving ratio scale dependence not only lower wealth mobility outcomes, but also raise wealth inequality significantly. Nonetheless, as a result of the presence of entrepreneurs, wealth inequality outcomes remain relatively close to their baseline level. In a model without entrepreneurs, these effects would be more significant. Also for saving ratio type dependence, the inverse relationship holds, albeit weakly: saving ratio type dependence significantly raises wealth mobility, and lowers the top 10% wealth share only slightly.

7 Conclusion

While recent theoretical work has proposed several mechanisms that contribute wealth inequality, there remains little research linking the inequality of the wealth distribution to its turnover (i.e. wealth mobility). Moreover, the existing Aiyagari–Bewley–Huggett heterogeneous agent literature typically invokes assumptions about type dependence and scale dependence without defining these formally and without estimating the type-dependent and scale-dependent parameters in an internally consistent way. This paper addresses these research gaps through four key contributions.

First, I constructed a generalized theoretical framework that embeds the core sources of wealth inequality underscored in the theoretical literature (labor income risk, saving rate heterogeneity, capital income risk, link between returns and wealth). Using this framework, I proposed a formal definition of type dependence and scale dependence, and defined several theoretical type dependence (spread and persistence) and scale dependence (spread and curvature) moments that are of interest.

Second, using a set of simplified heterogeneous agent models, I have shown that the type dependence versus scale dependence distinction is critical for matching wealth mobility outcomes: for identical wealth inequality outcomes, type-dependent models generate higher wealth mobility than scale-dependent ones. This is because type dependence ultimately introduces an additional source of randomness into the model. In addition, the relationship between wealth mobility and the scale-dependent and type-dependent parameters is found to be characterized by non-linearities.

Third, I have constructed an Aiyagari-Bewley-Huggett economy with type dependence and scale dependence in which households exhibit non-homothetic preferences and entrepreneurs are modeled in line with Cagetti & De Nardi (2006). To estimate the type-dependent and scale-dependent parameters, I outlined a novel estimation strategy that links a theoretical scale-dependent function to a corresponding, empirically-determined type-dependent structure using panel data from the PSID. The estimated model was found to replicate well the wealth inequality and wealth mobility observed in the United States in 2021.

Fourth, I have conducted a series of counterfactual analyses on the estimated baseline model. These showed that allowing for a realistic degree of saving ratio type dependence is critical in matching wealth mobility in the stationary model state to its empirical counterpart. Moreover, labor income inequality and saving ratio inequality were found to be the key driving forces behind agents' persistence in the wealth (rank) distribution in both the short-run and the long-run. Return heterogeneity was found to be less important, though future research should examine to what extent this finding is driven by the assumption of a normal return process. Finally, in general, there exists an inverse relationship between wealth inequality and wealth mobility: higher wealth inequality coincides with lower wealth mobility.

8 References

Atkeson, A., Irie, M. (2022). Rapid Dynamics of Top Wealth Shares and Self-Made Fortunes: What Is the Role of Family Firms? AER: Insights 4 (4), p. 409–424.

Atkinson, A.B. (1971). Capital Taxes, the Redistribution of Wealth and Individual Savings. Review of Economic Studies 38, p. 209–227.

Azzalini, G., Kondziella, M., Racz, Z. (2023). Preference Heterogeneity and Portfolio Choices over the Wealth Distribution. Unpublished.

Bach, L., Calvet, L., Sodini, P. (2018). From Saving comes Having? Disentangling the Impact of Saving on Wealth Inequality. Swedish House of Finance Research Paper (June 2018).

Bach, L., Calvet, E., Sodini, P. (2020). Rich pickings? Risk, Return, and Skill in Household Wealth. American Economic Review 110, p. 2703–2747.

Balloch, C., Richers, J. (2023). Asset Allocation and Returns in the Portfolios of the Wealthy. LSE Financial Markets Group.

Benhabib, J., Bisin, A. (2018). Skewed wealth distributions: Theory and Empirics. Journal of Economic Literature 56 (4), p. 1261-1291.

Benhabib, J., Bisin, A., Luo, M. (2019). Wealth Distribution and Social Mobility in the U.S.: A Quantitative Approach. American Economic Review 109 (5), p. 1623-1647.

Benhabib, J., Bisin, A., Fernholz, R.T. (2022). Heterogeneous Dynasties and Long-Run Mobility. The Economic Journal 132 (643), p. 906-925.

Benhabib, J., Wei, C., Miao, J. (2024). Capital Income Jumps and Wealth Distribution. Quantitative Economics 15 (4).

Bricker, J., Moore, K.B., Volz, A.H. (2021). Private Business Wealth and Rates of Return in the U.S.. Society for the Study of Economic Inequality Working Paper.

Cagetti, M., De Nardi, M. (2006). Entrepreneurship, Frictions and Wealth. Journal of Political Economy 114, p. 835–870.

Carroll, C.D. (2006). The Method of Endogenous Gridpoints for Solving Dynamic Stochastic Optimization Problems. Economics letters 91 (3), p. 312-320.

Catherine, S., Miller, M., Sarin, N. (2020). Social Security and Trends in Wealth Inequality. The Journal of Finance 80 (3), p. 1497-1531.

Cesarini, D., Lindqvist, E., Notowidigdo, M., Ösling, R. (2017). The Effect of Wealth on Individual and Household Labor Supply: Evidence from Swedish Lotteries. American Economic Review 107(12), p. 3917-3946.

Cesarini, D., Lindqvist, E., Notowidigdo, M. J., Östling, R. (2017). The effect of wealth on individual and household labor supply: evidence from Swedish lotteries. American Economic Review, 107(12), 3917-3946.

Cioffi, R. (2021). Heterogeneous Risk Exposure and the Dynamics of Wealth Inequality. Unpublished.

De Nardi, M. (2004). Wealth inequality and Intergenerational links. Review of Economic Studies 71, p. 743–768.

De Nardi, M., Fella, G. (2017). Saving and Wealth Inequality. Review of Economic Dynamics 26, p. 280-300.

Dynan, K., Skinner, J., Zeldes, S. (2004). Do the Rich Save More? Journal of Political Economy 112 (2), p. 397-444.

Fagereng, A., Guiso, L., Malacrino, D., Pistaferri, L. (2020). Heterogeneity and Persistence in Returns to Wealth. Econometrica 88, p. 115–170.

Fagereng, A., Holm, M.B., Moll, B., Natvik, G. (2021). Saving Behavior Across the Wealth Distribution: The Importance of Capital Gains. National Bureau of Economic Research w26588.

Favilukis, J., Ludvigson, S.C., Van Nieuwerburgh, S. (2017). The Macroeconomic Effects of Housing Wealth, Housing Finance, and Limited Risk Sharing in General Equilibrium. Journal of Political Economy 125(1), p. 140-223.

Ferraro, D., Valaitis, V. (2022). Wealth and Hours. Working paper, available at SSRN 3740240.

Fernández-Villaverde, J., Levintal, O. (2024). The Distributional Effects of Asset Returns. NBER Working Paper 32182.

Fischer, T. (2019). Determinants of Wealth Inequality and Mobility in General Equilibrium. Working Papers 2019:22, Lund University, Department of Economics.

Gabaix, X., Lasry, J.M., Lions, P.L., Moll, B. (2016). The Dynamics of Inequality. Econometrica 84 (6), p. 2071-2111.

Gaillard, A., Wangner, S. (2023). Wealth, Returns, and Taxation: A Tale of Two Dependencies. Unpublished.

Gerritsen, A., Jacobs, B., Spiritus, K., Rusu, A.V. (2025). Optimal Taxation of Capital Income with Heterogeneous Rates of Return. The Economic Journal 135 (665), p. 180-211.

Gomez, M. (2023). Decomposing the Growth of Top Wealth Shares. Econometrica 91 (3), p. 979-1024.

Gomez, M., Gouin-Bonenfant, E., (2024). Wealth Inequality in a Low Rate Environment. Econometrica 92 (1).

Gould, E., Kandra, J. (2022). Wage Inequality in the U.S. Worsened in 2021, Economic Policy Institute.

Heathcote, J., Storesletten, K., Violante, G. (2010). The Macroeconomic Implications of Rising Wage Inequality in the United States. Journal of Political Economy 118 (4), p. 681-722.

Hubmer, J., Krusell, P., Smith, A. (2021). Sources of U.S. Wealth Inequality: Past, Present and Future. NBER Macroeconomics Annual 35 (1), p. 391-455.

Hubmer, J., Halvorsen, E., Salgado, S., Ozkan, S., (2024). Why Are the Wealthiest So Wealthy? New Longitudinal Empirical Evidence and Implications for Theories of Wealth Inequality. Federal Reserve Bank of St. Louis Working Paper 2024-013.

Hubmer, J., Restrepo, P. (2023). Not a Typical Firm: Capital-Labor Substitution and Firms' Labor Shares. Penn Institute for Economic Research, Department of Economics, University of Pennsylvania.

Hurst, E., Lusardi, A. (2004). Liquidity constraints, household wealth, and entrepreneurship. Journal of political Economy 112(2), p.319-347.

Kaplan, G., Moll, B., Violante, G.L. (2018) Monetary Policy according to HANK. American Economic Review 108 (3), p. 697-743.

Kartashova, K. (2014). Private equity premium puzzle revisited. American Economic Review 104 (10), p. 3297-3334.

Kaymak, B., Leung, D., Poschke, M. (2022). Accounting for Wealth Concentration in the United States. Working Paper No. 22-28. Federal Reserve Bank of Cleveland.

Krusell, P., Smith, A. (1998). Income and Wealth Heterogeneity in the Macroeconomy. Journal of Political Economy 106, p. 867–896.

Kim, M. (2024). Financial Literacy, Portfolio Choice, and Wealth Inequality: A General Equilibrium Approach. University of Pennsylvania Population Center Working Paper (PSC/PARC).

Kuhn, M., Schularick, M., Steins, U. (2020). Income and Wealth Inequality in America, 1949-2016. Journal of Political Economy 128 (9), p. 3469-3519.

Kuhn, M., Ríos-Rull, J.V. (2025). Income and Wealth Inequality in the United States: An Update Including the 2022 Wave. NBER Working Paper 33823.

Mazumder, B. (2016). Estimating the Intergenerational Elasticity and Rank Association in the United States: Overcoming the Current Limitations of Tax Data. Inequality: Causes and Consequences, p. 83-129.

Michau, J.B., Ono, Y., Schlegl, M. (2023). The Preference for Wealth and Inequality: Towards a Piketty Theory of Wealth Inequality. Institute of Social and Economic Research Discussion Paper No. 1223.

Moskowitz, T., Vissing-Jorgensen, A. (2002). The Returns to Entrepreneurial Investment: A Private Equity Premium Puzzle? American Economic Review 92, p. 745-778.

Pfeffer, F., Schoeni, R., Kennickel, A., Andreski, P. (2016). Measuring wealth and wealth inequality: Comparing two U.S. surveys. Journal of Economic and Social Measurement 41, p. 103-120.

Pfeffer, F., Killewald, A. (2018). Generations of Advantage. Multigenerational Correlations in Family Wealth. Social Forces 96(4), p. 1411-1442.

Rouwenhorst, K.G. (1995). Asset pricing implications of equilibrium business cycle models. Frontiers of Business Cycle Research, Princeton University Press, Princeton, NJ, p. 294-330.

Saez, E., Zucman, G. (2016). Wealth inequality in the United States since 1913: Evidence from capitalized income tax data. Quarterly Journal of Economics 131 (2), p. 519-578.

Smith, M., Zidar, O., Zwick, E. (2023). Top Wealth in America: New Estimates Under Heterogeneous Returns. Quarterly Journal of Economics 138 (1), p. 515-573.

Snudden, S. (2025). Idiosyncratic Asset return and Wage Risk of U.S. households. Economic Inquiry 63(2), p. 636-657.

Straub, L. (2019). Consumption, Savings, and The Distribution of Permanent Income. Unpublished Manuscript, Harvard University.

Toda, A. (2019). Wealth distribution with Random Discount Factors. Journal of Monetary Economics 104, p. 101-113.

Van Langenhove, C. (2025a). Wealth Mobility in the United States: Empirical Evidence from the PSID. Working Papers of Faculty of Economics and Business Administration, WP 25/1104.

Van Langenhove, C. (2025b). Saving Rate Heterogeneity across the Wealth (Rank) Distribution in the United States. Forthcoming.

Van Langenhove, C. (2025c). Return and Risk Heterogeneity across the Wealth (Rank) Distribution in the United States. Forthcoming.

Xavier, I. (2021). Wealth Inequality in the U.S.: the Role of Heterogeneous Returns. Available at SSRN 3915439.

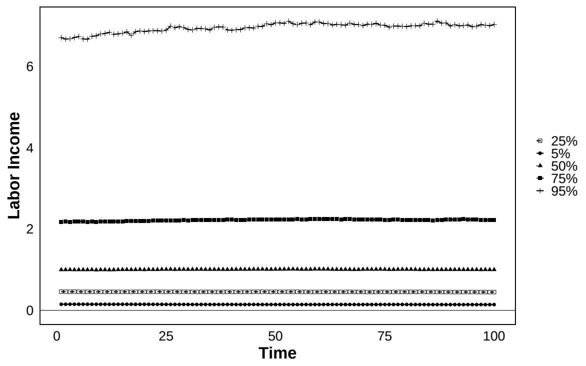
A State variable processes − visualization

This Appendix visualizes the state variable process paths under the calibration outlined in Section 3. More precisely, I simulate for each state variable 100 000 trajectories over 100 years, and rank the outcomes at each time period t. Figure 10 plots the 95th, 75th, 50th, 25th and 5th percentile at each t. For the equity, housing and business return, I plot the value of initial 100-unit investment. Returns represent capital gains only; that is, the fixed return (μ^e, μ^h, or the median business return in Γ) is subtracted from the total return.

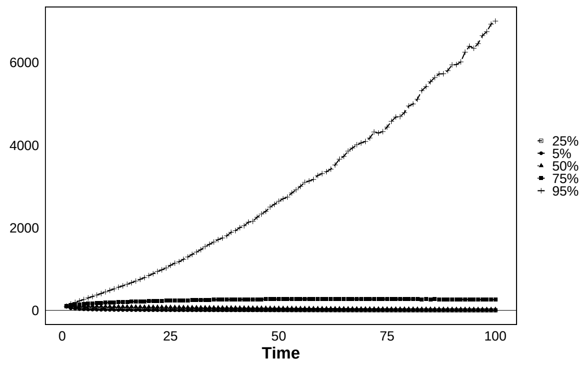
The difference in outcomes for the return processes are quite stark. Specifically, equity returns yield on average a relatively stable investment path, but imply a small probability of significant investment gains. For housing returns, the average path yields a positive return, but the probability of significant gains is more limited than for equity. Finally, business returns also imply a relatively stable average scenario. However, they stand out by a small probability of significant reversals of fortune: in the 5% most favorable scenarios, the 100-unit investment has turned into a six-figure investment after 100 years.

Figure 10: State variable paths under the parameter calibration of Section 3.1.

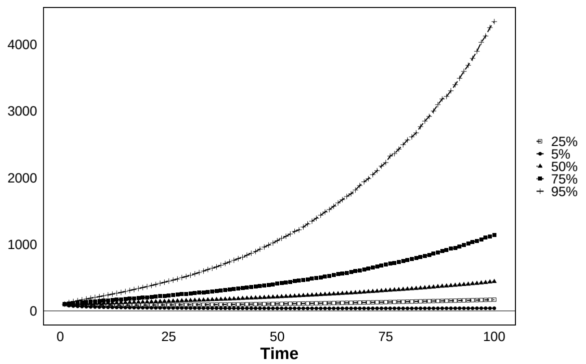
(a) Labor income.



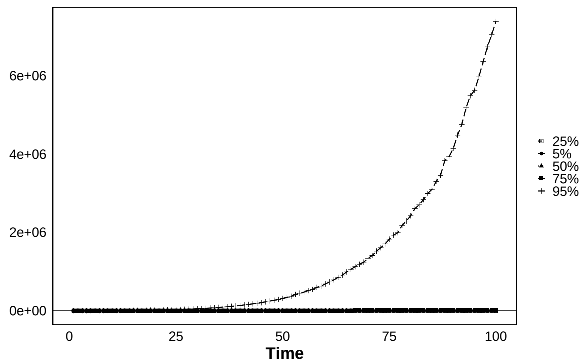
(b) Equity (capital gains only).



(c) Housing (capital gains only).



(d) Business (capital gains only).

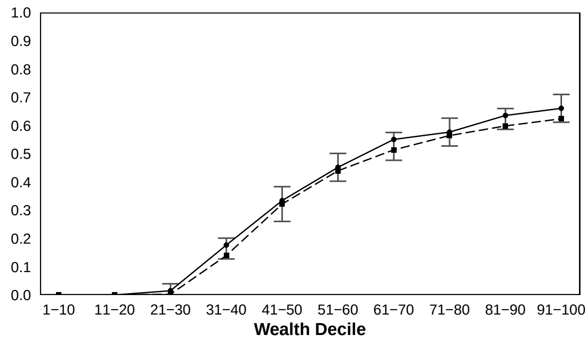


Note: the plots show the outcomes of 100 000 Monte Carlo simulations for the labor income process (a), equity return process (b), housing return process (c) and business return process (d). The return processes represent capital gains only: the fixed return has been subtracted from the total return. The asset value outcomes depart from a 100-unit initial value.

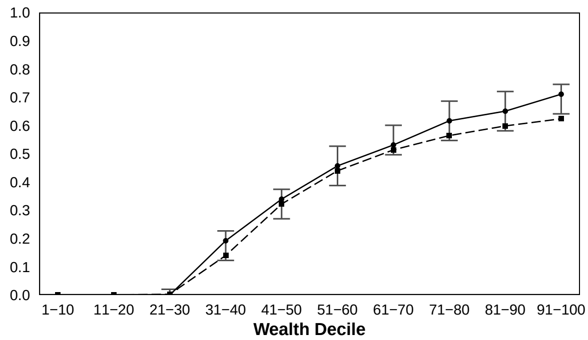
B Type dependence theoretical moments − saving ratio dispersion across the wealth (rank) distribution

Figure 11: Type dependence spread for a given type dependence persistence of $\rho^{\theta} = 0.75$ and resulting saving ratio data patterns (solid line).

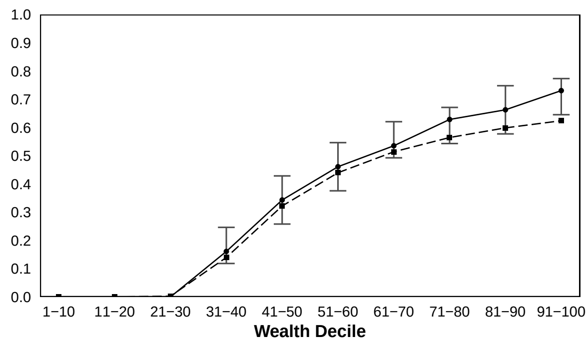
(a) $\Delta^{(t)\varepsilon} = 0.005$.



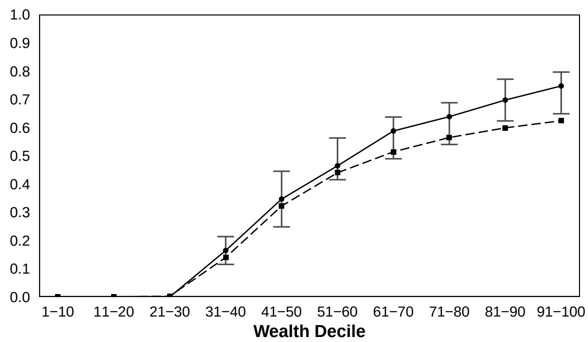
(b) $\Delta^{(t)\varepsilon} = 0.01$.



(c) $\Delta^{(t)\varepsilon} = 0.015$.



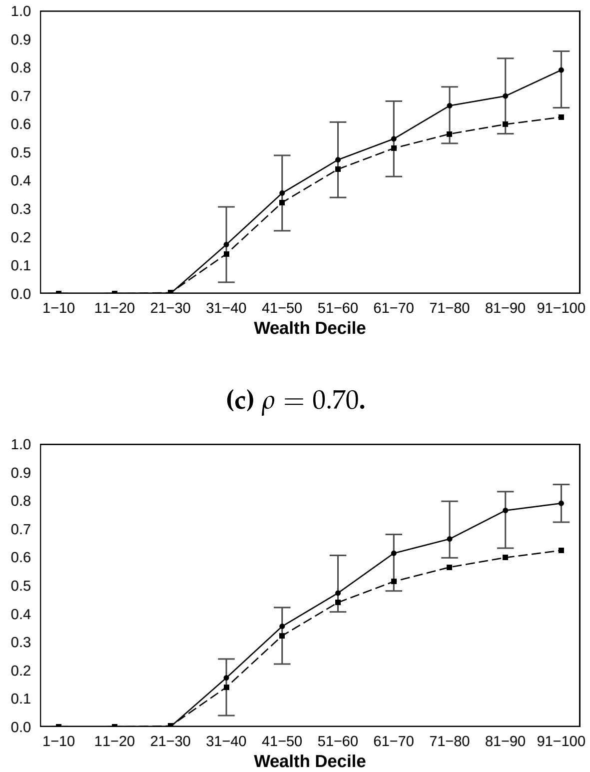
(d) $\Delta^{(t)\varepsilon} = 0.02$.



Note: the solid line shows the median saving ratio per wealth decile in the simulation data, while the bars represent the inter-quartile range of the simulated saving ratio per wealth decile. The dotted line denotes the scale-dependent saving ratio function f^{θ} that was imposed on the model. The simulations take as given the saving ratio type dependence persistence: $\rho^{\theta} = 0.75$. All models depart from the simplified framework outlined in Section 4.1.

Figure 12: Type dependence persistence ρ for a given type dependence spread of $\Delta^{(t)\theta} = 0.04$ and resulting saving ratio data patterns (solid line).

(a) $\rho = 0.10$. **(b)** $\rho = 0.40$. **(c)** $\rho = 0.70$. **(d)** $\rho = 0.99$.

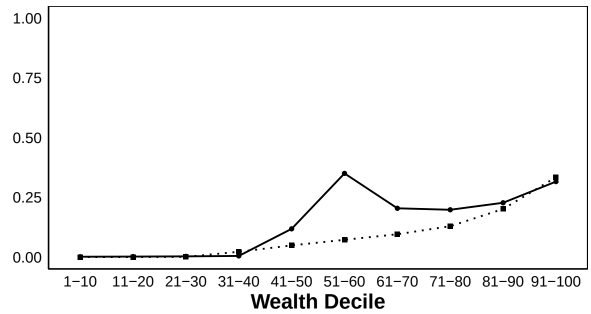


Note: the solid line shows the median saving ratio per wealth decile in the simulation data, while the bars represent the inter-quartile range of the simulated saving ratio per wealth decile. The dotted line denotes the scale-dependent saving ratio function f^{θ} that was imposed on the model. The simulations take as given the saving ratio type dependence spread: $\Delta^{(t)\theta} = 0.04$. All models depart from the simplified framework outlined in Section 4.1.

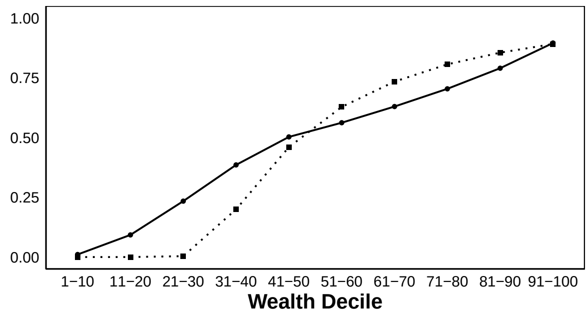
C Baseline model – additional visualizations

Figure 13: Outcomes across the wealth (rank) distribution in the baseline model (solid line) and in the data (dotted line).

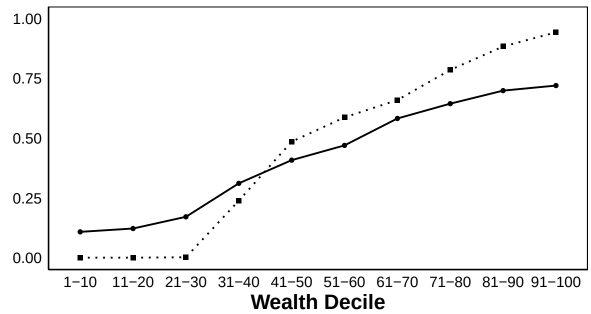
(a) Share of entrepreneurs.



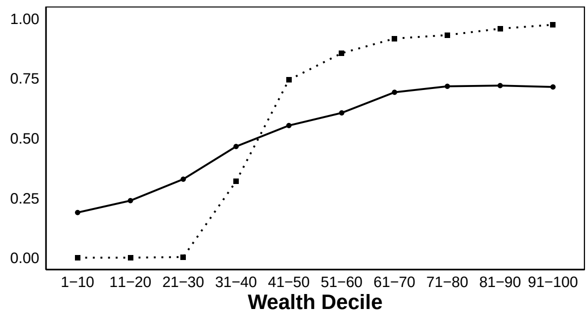
(b) Saving ratios.



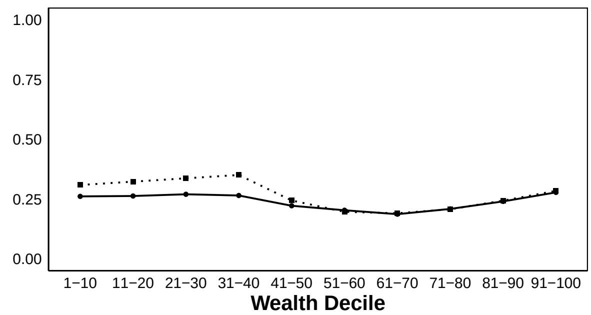
(c) Equity participation rate.



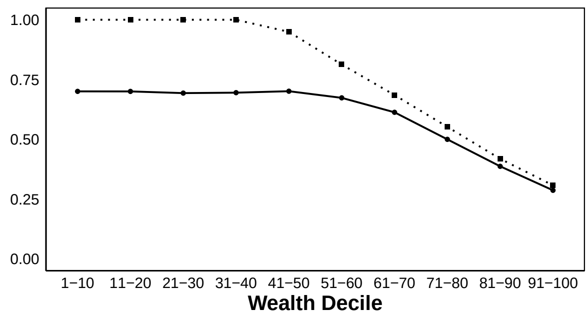
(d) Housing participation rate.



(e) Equity portfolio share.



(f) Housing portfolio share.



Note: the solid line shows the median outcome variable per wealth decile d in the stationary state of the baseline model. The dotted line shows the median for that outcome variable computed from the empirical data. For the share of entrepreneurs (a), asset participation (c-d) and portfolio allocation (e-f), the empirical data was taken from Van Langenhove (2025c). For the saving ratios (b), it was taken from Van Langenhove (2025b).

Conclusion

The Aiyagari-Bewley-Huggett heterogeneous agent macro literature is faced with two principal shortcomings. First, existing models of the U.S. wealth distribution focus exclusively on replicating wealth inequality, without accounting for wealth mobility. Second, while the heterogeneous agent models rely on both type dependence and scale dependence, these concepts are used loosely: no formal definition is provided, their implications for various model outcomes are left unexamined, and the importance of the dependence for wealth mobility is currently unexplored.

Addressing these shortcomings entails multiple challenges. In particular, there exists no extensive wealth mobility data for the United States, or empirical evidence on saving behavior across the wealth (rank) distribution. Furthermore, the literature does not provide an explicit, formal definition of type dependence and scale dependence, and has not yet proposed a strategy to consistently estimate the type- and scale-dependent parameters of a heterogeneous agent model. In this dissertation, I have taken critical steps forward in tackling these challenges, thereby addressing the two principal shortcomings of the literature.

In Chapter 1, I leveraged data from the Panel Study of Income Dynamics (PSID) to analyze inter- and intra-generational wealth mobility in the United States. I provided a rich set of empirical stylized facts relevant to the macroeconomic modeling of the U.S. wealth distribution. I demonstrated that both overall inter-generational wealth mobility and intra-generational mobility at the top have declined over time, and that wealth mobility in the United States is lower than in most other countries for which comparable data are available. Moreover, I found evidence of positive interdependence between individuals' wealth rank trajectories and those of their parents over the same historical time period. In addition, I investigated the sources of inter- and intra-generational wealth mobility in the United States, showing that variation in inter-generational transfers, business ownership, labor income, health and non-mortgage indebtedness are critical determinants of mobility.

In Chapter 2, I used household-level data from the Panel Study of Income Dynamics (PSID) to provide evidence on saving behavior across the wealth (rank) distribution in the United States. I estimated saving rates across wealth deciles using two complementary approaches: the cross-sectional method and the aggregate method. I obtained four collections of stylized empirical facts. First, I found that total saving rates out of labor income and new resources rise with wealth ranks (flow-based saving rates). In contrast, total saving rates out of wealth and composite resources are roughly stable or moderately increasing with wealth ranks (stock-based saving rates). Second, wealth (rank) mobility has a substantial impact on total saving rate patterns across the wealth distribution. However, while the contribution of wealth mobility is strictly positive for the cross-sectional method, it is negative across most of the wealth distribution for the aggregate method. I show that this discrepancy relates to these methods' distinct treatment of wealth (rank) mobility: while the cross-sectional method attaches equal weight to all households in a wealth decile, the aggregate method overweighs households that display downward wealth mobility. Third, I found that the synthetic method (which is commonly used in the absence of panel data) overestimates saving rates up to the 80th percentile, while it underestimates the saving rates of the top 20%. Fourth, I demonstrated that households' reliance on capital gains rises across the wealth rank distribution: the top wealthiest households' total saving consists predominantly of saving by holding appreciating assets. Passive saving out of inter-generational transfers is more common for wealthier households, but relatively unimportant in magnitude. Many of the empirical saving behavior moments across the wealth (rank) distribution reported in this Chapter are likely of interest to the heterogeneous agent literature replicating the U.S. wealth distribution.

In Chapter 3, I have used heterogeneous agent models reliant on both type dependence and scale dependence to jointly study wealth inequality and wealth mobility in the United States. First, the chapter outlined a generalized theoretical framework that formally defined type dependence and scale dependence. Second, using a set of simplified heterogeneous agent models, I demonstrated that the type dependence versus scale dependence distinction is critical for matching wealth mobility outcomes: for identical wealth inequality outcomes, type-dependent models generate higher wealth mobility than scale-dependent ones. Third, I constructed an Aiyagari-Bewley-Huggett economy populated by households and entrepreneurs, and with both type dependence and scale dependence in parameters and decision variables. To estimate the type-dependent and scale-dependent parameters, I outlined a novel estimation strategy that links a theoretical scale-dependent function to a corresponding empirically-determined type-dependent structure using panel data from the PSID. The estimated model replicates well the wealth inequality and wealth mobility observed in the United States in 2021. Fourth, I conducted a series of counterfactual analyses on the estimated baseline model. These showed that allowing for a realistic degree of saving ratio type dependence is critical in matching wealth mobility in the stationary model state to its empirical counterpart. Moreover, labor income inequality and saving ratio inequality emerge as the key driving forces behind agents' persistence in the wealth (rank) distribution in both the short-run and the long-run. Return heterogeneity was found to be less important, although this may relate to specific model assumptions. Finally, in general, there exists an inverse relationship between wealth inequality and wealth mobility: higher wealth inequality coincides with lower wealth mobility.

These three chapters provide a response to the six broader research questions raised in the Introduction to this PhD dissertation. First, I found a negative relationship between U.S. wealth inequality and wealth mobility empirically: while wealth inequality has risen, inter- and intra-generational wealth mobility have declined over time. A similar inverse relationship was also found to hold theoretically. Second, I found labor income inequality and saving rate inequality to be the core determinants behind U.S. wealth inequality. These were also the main drivers of reduced turnover (and hence lower mobility) across the wealth distribution. Return heterogeneity is of less importance, although this finding warrants further research. Third, as noted, U.S. wealth mobility has declined over time, and is in general lower compared to other countries with available data. Fourth, type dependence is critical in generating realistic wealth mobility outcomes: without it, wealth mobility in the stationary model state is lower than its empirical counterpart. Heterogeneous agent models that aim to match U.S. wealth mobility outcomes should therefore rely on a realistic degree of type dependence and scale dependence. Fifth, the model-based estimation of type-dependent and scale-dependent parameters suggested that both type and scale dependence are relevant. In other words, heterogeneity in wealth outcomes across U.S. households arises from both structural differences in households' saving and portfolio allocation behavior (ex-ante heterogeneity) and from 'wealth begets wealth' dynamics (ex-post heterogeneity). Sixth, there exists an overall positive relationship between saving behavior and wealth ranks: saving rates are higher for wealthier households. However, the magnitude of the effect depends on the saving rate considered: saving rates out of labor income and new resources are significantly higher at the top compared to the bottom, while the difference for saving rates out of composite resources is more minimal. On the contrary, saving rates out of wealth are relatively stable across the wealth rank distribution.

Societal and policy implications What are the implications of the findings of this PhD for societal and policy debates? I outline several key points.

First, the dissertation stresses the importance of connecting policy debates on wealth inequality to the degree of wealth mobility: high wealth inequality might be more problematic when it coincides with low wealth mobility. For the United States, Chapter 1 demonstrated that the rising U.S. wealth inequality since the beginning of the 1980s has coincided with declining U.S. wealth mobility, particularly at the top. Moreover, wealth inequality in the United States is higher and wealth mobility lower compared to other countries with available data. This suggests that the negative externalities associated with high wealth inequality — such as political capture and weakening of political institutions, social fragmentation and unrest, unequal access to healthcare, and underinvestment in human capital — are more likely to materialize in the United States. It also suggests that a return to the Gilded Age, as predicted by Thomas Piketty and others, is not a dystopian scenario but a plausible outcome.

Second, Chapter 3 of the dissertation introduces a distinction between type dependence and scale dependence. Furthermore, it introduces a method to estimate type-dependent and scale-dependent parameters in a heterogeneous agent model. The distinction between type dependence and scale dependence has strong implications regarding the nature of wealth inequality. Type dependence relates the inequality to ex-ante differences across agents − some agents are simply more frugal or better investors than others. On the contrary, scale dependence traces the inequality to ex-post differences across agents − the wealthy become wealthier simply because their wealth creates conditions that are more favorable to additional wealth accumulation. Scale dependencies are generally classified as unjust: why should wealthier agents operate under conditions more favorable to wealth accumulation compared to poorer agents? Strikingly, the model estimation in the second part of Chapter 3 suggests that such scale dependence is a key driver of U.S. wealth inequality. This raises doubts on the degree of equal opportunity for wealth accumulation in U.S. society.

Overall, these two points draw a picture of a highly unequal and immobile U.S. wealth distribution relative to other developed countries. In addition, this high wealth inequality and low wealth mobility appear to have worsened over time. In the absence of a clear policy shift, and in the light of the aforementioned negative externalities materializing further, it seems likely that these trends will persist in the future. So what can be done about it? While the present dissertation did not conduct direct policy analyses, it does produce two main policy suggestions, which I discuss next.

On the one hand, the counterfactual model analyses in Chapter 3 demonstrated the importance of labor income inequality in driving long-run wealth mobility. There exists extensive evidence that labor income inequality in the U.S. has increased in the decades prior to the financial crisis, especially at the top of the distribution. Insofar as U.S. policymakers are interested in lowering wealth inequality and raising wealth mobility, curbing labor income inequality is therefore likely an effective strategy (through e.g. tax reforms). On the other hand, both in the empirical data of Chapter 1 and the theoretical model of Chapter 3, entrepreneurship emerged as a key driver of upward wealth mobility. Creating conditions where entrepreneurship is accessible across the entire population − regardless of initial wealth level, gender and race − therefore seems critical in allowing for turnover in and promoting dynamism across the U.S. wealth distribution. This is especially important given the decline in intra-generational wealth mobility towards the top 10% of the wealth distribution obtained in Chapter 1.

This PhD dissertation focused entirely on the United States, but what about Europe? Do the same societal and policy implications hold here? In general, European societies are less unequal than American ones. Unfortunately, with the exception of the Nordic countries, European countries have no panel data available over sufficiently long time horizons to investigate wealth mobility in these economies. This makes it hard to make statements about the turnover of families and individuals in European countries' wealth distributions. As a result, while some key points of this dissertation − e.g. the importance of wealth mobility when discussing wealth inequality, the conceptual distinction between type dependence and scale dependence − are also relevant for European countries, the absence of extensive data makes it harder to quantify the trends.

Future research This dissertation also opens several directions for future research. I elaborate on four of the most interesting avenues.

First, as noted, while the PSID can be used to study the relatively broad group of top 10% wealthiest, it does not capture the tail of the U.S. wealth distribution very well. This is unfortunate, as it is in fact tail wealth inequality that has risen the strongest over the past decades in the United States. Future research should therefore be concerned with quantifying the degree of wealth mobility and the saving behavior among the wealthiest households at the very top (top 1% and beyond).

Second, Chapter 1 showed that U.S. wealth inequality is high relative to most other countries, while its wealth mobility is relatively low. An interesting avenue for future research is to examine where other countries with available wealth mobility data are located on the wealth inequality–wealth mobility spectrum. It would then also be worthwhile to explain these cross-country differences in a theoretical model. Do the differences relate to diverging labor income inequality levels, entrepreneurship, household formation dynamics, or portfolio allocation and return heterogeneity?

Third, Chapter 3 focused on jointly replicating U.S. wealth inequality and U.S. wealth mobility at a given point in time, for 2021. For future research, it would be worthwhile to investigate what drove the simultaneous increase in U.S. wealth inequality and decline in U.S. wealth mobility. While I found only a limited contribution of portfolio allocation and return heterogeneity to stationary state wealth inequality and wealth mobility, such heterogeneity may be more important in inducing short-run or medium-run wealth inequality and wealth mobility fluctuations.

Fourth, to investigate further the probability of returning to a Gilded Age period, it is worthwhile connecting the discussion on wealth inequality and wealth mobility to the debate on the role of inter-generational transfers. From an empirical perspective, the PSID could be used to investigate the importance of inter-generational transfers across the wealth (rank) distribution in the United States. From a theoretical viewpoint, the heterogeneous agent models constructed in Chapter 3 could be extended to an overlapping-generations framework to incorporate inter-generational transfers.