



UNIVERSITÀ DEGLI STUDI DI PADOVA

DIPARTIMENTO DI SCIENZE ECONOMICHE ED AZIENDALI

“M. FANNO”

**CORSO DI LAUREA MAGISTRALE IN ACCOUNTING, FINANCE
AND BUSINESS CONSULTING**

TESI DI LAUREA

“Life Insurance, Housing, and Household Portfolios”

RELATORE:

CH.MO PROF. GUGLIELMO WEBER

LAUREANDO: OMAR GHORBEL

MATRICOLA N. 2155547

ANNO ACCADEMICO 2025 – 2026

Dichiaro in conformità con le linee guida di Ateneo “Linee guida per l’impiego degli strumenti di IA generativa nella didattica e nella redazione delle tesi di laurea e di dottorato” (https://www.unipd.it/sites/unipd.it/files/2025/LineeGuidaImpiegoAI_Unipd_ITA_rev20251128.pdf) di:

- di NON aver usufruito di tecnologie di intelligenza artificiale generativa (ad esempio ChatGPT, Claude, Perplexity) per supportare vari aspetti della ricerca e della scrittura della mia prova finale;
- ✓ di aver usufruito di tecnologie di intelligenza artificiale generativa (ad esempio ChatGPT, Claude, Perplexity, Copilot) per:
 - ✓ Revisione della letteratura: utilizzo per identificare articoli scientifici pertinenti, generare stringhe di ricerca e organizzare la bibliografia. Tutti i riferimenti sono stati successivamente verificati attraverso database accademici tradizionali (WoS, Scopus, PubMed, Google Scholar).
 - ✓ Analisi linguistica e revisione testuale: correzione grammaticale, ortografica e stilistica del testo. Il contenuto concettuale, i ragionamenti e le argomentazioni rimangono interamente frutto del mio lavoro.
 - ✓ Strutturazione del lavoro: supporto nella definizione della struttura iniziale del lavoro e nell’organizzazione dello stesso.
 - Traduzione: assistenza nella traduzione di testi per consultazione di fonti in lingua straniera.
 - ✓ Analisi dei dati: supporto nella generazione o verifica di codice per l’elaborazione statistica/analisi dati. Tutti i risultati sono stati verificati e validati autonomamente.
 - ✓ Formattazione: assistenza nella formattazione della bibliografia, tabelle, o altri elementi grafici secondo le norme editoriali richieste.
 - Altro: _____

II/La sottoscritto/a dichiara inoltre che:

L’utilizzo degli strumenti di intelligenza artificiale è stato coadiuvante e non sostitutivo delle proprie competenze;

Tutto il ragionamento critico, l’interpretazione dei risultati, le conclusioni e le argomentazioni presenti nella tesi sono frutto del proprio lavoro intellettuale indipendente;

Tutti i contenuti generati dall’intelligenza artificiale sono stati attentamente verificati, rivisti e rielaborati; Si assume la piena responsabilità per la qualità, l’accuratezza e l’integrità scientifica del lavoro presentato.

Firma



I hereby declare, in accordance with the University Guidelines “Guidelines for the Use of Generative AI Tools in Teaching and in the Drafting of Bachelor’s and Doctoral Theses” (https://www.unipd.it/sites/unipd.it/files/2025/LineeGuidaImpiegoAI_Unipd_ITA_rev20251128.pdf), that:

- I have NOT made use of generative artificial intelligence technologies (e.g., ChatGPT, Claude, Perplexity, Copilot) to support any aspects of the research and writing of my final thesis;
- ✓ I have made use of generative artificial intelligence technologies (e.g., ChatGPT, Claude, Perplexity, Copilot) for the following purposes:
 - ✓ Literature review: use to identify relevant scientific articles, generate search strings, and organize the bibliography. All references were subsequently verified through traditional academic databases (WoS, Scopus, PubMed, Google Scholar).
 - ✓ Linguistic analysis and textual revision: grammatical, spelling, and stylistic correction of the text. The conceptual content, reasoning, and arguments remain entirely the result of my own work.
 - ✓ Structuring of the work: support in defining the initial structure of the thesis and in organizing its contents.
 - Translation: assistance in translating texts for consultation of sources in foreign languages.
 - ✓ Data analysis: support in generating or verifying code for statistical processing/data analysis. All results were independently verified and validated.
 - ✓ Formatting: assistance in formatting the bibliography, tables, or other graphical elements according to the required editorial standards.
 - Other: _____

The undersigned further declares that:

The use of artificial intelligence tools was supportive and not a substitute for their own skills and competencies;

All critical reasoning, interpretation of results, conclusions, and arguments presented in the thesis are the result of their own independent intellectual work;

All content generated by artificial intelligence was carefully verified, reviewed, and reworked;

Full responsibility is assumed for the quality, accuracy, and scientific integrity of the work presented.

Firma



Abstract

This thesis extends the conditional portfolio efficiency framework of Gouriéroux and Jouneau (1999), as applied by Pelizzon and Weber (2008) to a single illiquid constraint, to the empirically relevant case in which households simultaneously hold two constrained asset positions: residential real estate and life insurance and pension products. Observed Italian household portfolios deviate substantially from mean-variance efficient allocations, yet it remains unclear whether such deviations reflect genuine irrationality or rational responses to large illiquid positions that generate hedging demands in the tradable financial portfolio. The theoretical contribution is to show that the optimal financial-asset allocation decomposes additively into the Markowitz tangency portfolio and two independent hedge terms, one per constrained asset. The framework is applied to 3,462 eligible households from the 2022 Survey on Household Income and Wealth, using asset returns spanning 2005–2024. Conditioning on housing alone reduces the share of conditionally efficient households from 38.3 percent to 19.7 percent at the ten percent test size, while conditioning on insurance leaves efficiency broadly unchanged at 51.1 percent, confirming that housing, not insurance, constitutes the binding illiquid constraint. A pronounced and persistent North–South efficiency gradient is documented, with 42.2 percent of North-Eastern households classified as conditionally efficient against only 2.7 percent in the South and Islands, and a monotone age gradient shows younger households to be the most conditionally inefficient.

Contents

Introduction	2
1 Literature Review	4
1.1 Mean-Variance Efficiency and Portfolio Efficiency Tests	4
1.1.1 The Mean-Variance Framework	4
1.1.2 Conditional Portfolio Efficiency with Constrained Assets	5
1.1.3 Estimation Risk	5
1.2 Housing, Illiquidity, and Optimal Portfolio Choice	6
1.2.1 Hedging Demands from Illiquid Housing Wealth	6
1.2.2 Housing and Portfolio Composition	7
1.2.3 Conditional Portfolio Efficiency: Evidence from Italy	8
1.3 Life Insurance and Pension Products	9
1.4 The Research Gap	9
2 SHIW 2022	10
2.1 Sample Construction and Variable Overview	10
2.2 Financial Asset Classification and Holdings	11
2.3 Life Insurance and Pension Product Classification	11
2.4 Financial and Total Wealth	12
2.5 Regional Demographic Weights	13
2.6 Asset-Class Participation and Portfolio Combinations	14
2.7 Asset Participation Rates and Portfolio Shares	15
3 Return Data	17
3.1 Asset Universe and Data Overview	17
3.2 Risk-Free Rate: Treasury Bill Yields	18
3.3 The RENDISTATO Index	18
3.4 3Y and 5Y BTPs	19
3.5 Equity and Corporate Bond Indices	19
3.6 Insurance and Pension Products	20
3.6.1 Class I and Class III Life Insurance Products	20
3.6.2 Contractual and Open Pension Funds	20
3.6.3 PIPs: Individual Pension Plans	20

3.7	Housing	21
3.7.1	Regional House Price Indices	21
3.7.2	Net Total Return Adjustment	21
3.7.3	Regional Annual Net Returns	21
3.8	Consolidation into Broad Asset Classes	22
3.8.1	Stocks: Equal Weights of MSCI Italy and MSCI World	22
3.8.2	Class III Aggregate: Equal-Weight Specification	22
3.8.3	Long-Term Government Bonds: Duration-Matched Specification	23
4	ECONOMETRIC FRAMEWORK	24
4.1	The Two-Constrained-Asset Mean–Variance Problem	24
4.2	Hedge Coefficient Estimation	26
4.3	The Gouriéroux–Jouneau Conditional Efficiency Test	26
5	ESTIMATION AND RESULTS	29
5.1	Hedge Coefficient Estimates	29
5.1.1	Geometrically Declining Weights	31
5.1.1.1	Decay Parameter and Calibration	32
5.1.1.2	The Asymmetric Treatment of the Efficiency Tests	33
5.2	Financial-Asset Efficiency: Benchmark Test	34
5.3	Gouriéroux–Jouneau Conditional Efficiency Tests	36
5.3.1	Test A: Unconditional Efficiency	36
5.3.2	Test B: Conditional on Housing	37
5.3.3	Test C: Conditional on Insurance	38
5.3.4	Test D: Conditional on Housing and Insurance	39
5.4	Mean–Variance Frontier Visualisation	41
5.4.1	Efficiency By Macro-Region	42
5.4.2	Efficiency By Age Cohort	43
5.4.3	Efficiency by Educational Attainment	44
5.4.4	Efficiency by Employment Status	45
6	Conclusions	46
	References	50
A	Robustness: Equal vs. Declining Weights	53
B	Stata Code	55
B.1	Setup	55
B.2	Preparing the Source Datasets	55
B.3	Merging into the Master Dataset	58
B.4	Regional Demographic Weights	59

B.5	Participation Rates and the Portfolio Combination Table	59
B.6	Portfolio Weight Construction	61
B.7	Consolidated Export	62
C	Python Code	64
C.1	Estimation of the Hedge Terms	64
C.2	Gouriéroux–Jouneau Financial Assets Test	65
C.3	Gouriéroux–Jouneau Conditional Efficiency Tests	67
C.4	Efficient Frontier Plots	69

List of Figures

5.1	Constrained and unconstrained efficient frontiers by macro-region.	43
5.2	Constrained and unconstrained efficient frontiers by age group.	44
5.3	Constrained and unconstrained efficient frontiers by education group.	44
5.4	Constrained and unconstrained efficient frontiers by employment status.	45

List of Tables

2.1	SHIW 2022 Microdata Files	10
2.2	Classification of supplementary pension plans	12
2.3	Regional Weights within Macroregions and Italy	13
2.4	Asset-class participation flags	14
2.5	Asset combination by macro-region	14
2.6	Participation and Portfolio Shares	15
2.7	Total-wealth shares by macro-region	16
3.1	Individual Assets to Broad Asset Classes	22
5.1	Estimating the hedge coefficients: housing and insurance	31
5.2	Estimating the hedge coefficients: housing and insurance ($\alpha = 0.75$)	33
5.3	Financial-asset efficiency benchmark	36
5.4	Unconditional efficiency (no de-hedge), by portfolio case and by macro region.	37
5.5	Efficiency conditional on housing alone, by portfolio case and by macro region.	38
5.6	Efficiency conditional on insurance alone, by portfolio case and by macro region.	39
5.7	Efficiency conditional on housing and insurance, by portfolio case and by macro region.	40
5.8	Efficiency conditional on housing and insurance (Test D), full cross-classification.	41
A.1	Efficiency rates under equal weights ($\alpha = 1$) vs. recency weights ($\alpha = 0.75$)	54

Introduction

Household portfolio choice has occupied a central place in financial economics since the foundational work of H. Markowitz (1952), yet a persistent empirical puzzle remains: observed household portfolios deviate systematically and substantially from the mean-variance efficient allocations that theory prescribes. Italian households offer a particularly stark illustration. Real estate accounts for the dominant share of net worth, and homeownership is the primary determinant of balance-sheet composition, yet financial portfolios are concentrated almost entirely in bank deposits and short-term government securities, with equity and mutual-fund holdings remaining a minority phenomenon. Understanding whether these apparent distortions reflect genuine irrationality or the rational response of households constrained by large and illiquid positions requires an efficiency test that takes those constraints explicitly into account.

The conditional efficiency framework of Gouriéroux and Jouneau (1999) provides exactly such a test. For a household holding a pre-existing illiquid position, the relevant question is not whether its financial portfolio lies on the standard mean-variance frontier, but whether it lies on that frontier *after* subtracting the hedge demand that the illiquid position logically implies. Pelizzon and Weber (2008) applied this insight to Italian household data from the 1998 SHIW¹, demonstrating that the efficiency classification of household portfolios changes substantially once housing wealth is treated as a constrained position: conditionally efficient households are largely not the same as those whose purely financial portfolios appeared efficient, and even among conditionally efficient households hedging opportunities are not fully exploited. Their finding that neglecting the hedge term systematically misclassifies households gave rise to a productive empirical programme, but one that has remained limited to the single-constrained-asset case with housing as the sole illiquid position.

Two developments motivate extending this programme. First, Italian household balance sheets include a second category of semi-liquid constrained wealth that has grown substantially in importance since the pension reforms of the 1990s: life insurance and supplementary pension products, encompassing Class I participating (*gestione separata*) policies, Class III unit-linked products, Contractual Pension Funds, Open Pension Funds, and Individual Pension Plans² (PIPs). Their multi-year contractual structure, surrender penalties, and limited secondary-market liquidity render them effectively constrained in the same sense as housing. Second, the 2022 wave of the SHIW provides a substantially updated microdata base, capturing two decades of structural change in Italian financial and housing markets, including the disruptions

¹ Survey on Household Income and Wealth, Bank of Italy

² Introduced under Legislative Decree No. 252 of 2005

of the 2008–2009 global financial crisis, the 2011–2012 sovereign debt episode, the low-rate environment of the 2010s, and the pandemic.

This thesis addresses both dimensions. Its primary theoretical contribution is to extend the Gouriéroux and Jouneau (1999) framework from one to two simultaneously constrained asset positions: the optimal financial-asset allocation in the presence of both housing and insurance-pension constraints decomposes additively into the Markowitz tangency portfolio and two independent hedge terms, one for each constrained block. Its primary empirical contribution is to apply four progressively richer test specifications, unconditional, conditioning on housing alone, on insurance alone, and on both simultaneously, to a sample of eligible households drawn from the 2022 SHIW, with asset returns covering the 2005–2024 period, and to decompose efficiency rates by macro-region, age, education, and employment status.

To our knowledge, no study has extended the Gouriéroux–Jouneau–Pelizzon–Weber framework to two simultaneously constrained asset classes, assessed the portfolio-efficiency implications of life insurance and pension holdings within a conditional efficiency framework, or revisited the conditional efficiency of Italian household portfolios using post-crisis data.

The analysis is organised around three research questions. *First*, does extending the framework to two simultaneously constrained positions materially alter the assessment of household portfolio efficiency relative to the single-constrained-asset benchmark? *Second*, which constrained position, housing or insurance, accounts for the larger share of observed inefficiency? *Third*, is there evidence of a systematic regional gradient in conditional efficiency in the 2022 data, and if so, how does it interact with the age, education, and employment-status dimensions of portfolio choice?

The thesis is structured as follows. Chapter 1 reviews the literature on mean-variance efficiency testing, the portfolio implications of housing and insurance wealth, Italian household finance, and the contribution of Pelizzon and Weber (2008) that the present thesis extends. Chapter 2 describes the 2022 SHIW microdata, the construction of household-level portfolio vectors, and the sample selection criteria. Chapter 3 details the return data for the thirteen individual series consolidated into eight broad asset classes over the 2005–2024 period. Chapter 4 develops the econometric framework: the two-constrained-asset extension of the optimal portfolio rule, hedge coefficient estimation, and the specification of the Gouriéroux–Jouneau conditional efficiency test. Chapter 5 presents the empirical results in three stages: hedge coefficient estimates, the financial-asset benchmark test, and the four conditional efficiency tests, together with mean–variance frontier visualisations across demographic panels. Chapter 6 synthesises the findings, discusses theoretical and practical implications, acknowledges the study’s limitations, and identifies directions for future research.

Chapter 1

Literature Review

The empirical study of household portfolio allocation has expanded considerably over the past three decades. A recurring finding across countries is that households hold poorly diversified portfolios, systematically under-weight risky assets relative to the prescriptions of classical theory, and in large proportions abstain from stock-market participation altogether. Guiso, Jappelli, and Haliassos (2000) compile cross-country evidence for Europe and the United States and document that direct stockholding rates rarely exceed one third of the adult population even in financially developed economies, a phenomenon attributed variously to fixed participation costs, information barriers, and background risk. The Italian evidence is emblematic: Guiso and Jappelli (2000) show that Italian households concentrated the bulk of their financial wealth in deposits and short-term government securities throughout the 1980s and into the 1990s, with equity and mutual-fund holdings remaining a minority phenomenon even as domestic financial markets deepened substantially.

The standard mean-variance framework of H. Markowitz (1952) implies that all investors should hold a combination of the riskless asset and a single tangency portfolio of risky assets, with the relative share determined solely by risk aversion. Yet observed household portfolios deviate systematically from this prescription. Understanding whether these deviations reflect genuine irrationality, binding institutional constraints, or rational responses to background risk and illiquidity has motivated a large body of subsequent research. The present thesis belongs squarely to this tradition, asking whether Italian household portfolios are mean-variance efficient once two pervasive sources of illiquid wealth, housing and life insurance and pension products, are appropriately accounted for in the efficiency assessment.

1.1 Mean-Variance Efficiency and Portfolio Efficiency Tests

1.1.1 The Mean-Variance Framework

The efficiency analysis conducted in this thesis rests on the mean-variance portfolio theory of H. Markowitz (1952), formalised and extended in H. M. Markowitz, Todd, and Sharpe (2000). The central implication is that any investor who fails to hold the tangency portfolio on the capital

market line leaves attainable risk-adjusted return unrealised. Operationalising this empirically requires both a method for estimating the frontier and a test for whether an observed portfolio lies on it.

Gibbons, Ross, and Shanken (1989) derived the canonical multivariate test of portfolio efficiency, showing that the hypothesis that a given portfolio lies on the mean-variance frontier is equivalent to the hypothesis that the intercepts in a system of time-series regressions of asset excess returns on the candidate portfolio are jointly zero. Their F -statistic is exact in finite samples under normality and has served as the standard reference for asset-pricing tests. Jobson and Korkie (1982) and Jobson and Korkie (1989) proposed related tests based on the difference in Sharpe ratios between the observed and the tangency portfolio, providing an asymptotically equivalent framework that is more tractable when the investment set varies across agents, as it does in settings where regional housing returns differ across macro-areas.

1.1.2 Conditional Portfolio Efficiency with Constrained Assets

A critical limitation of the classical tests is that they are designed for unconstrained investment sets. When one or more assets are illiquid, so that the household's holding is predetermined rather than optimally chosen, the efficient frontier relevant to the freely tradable assets is no longer the capital-market line, and standard Sharpe-ratio comparisons become invalid. Gouriéroux and Jouneau (1999) address this by deriving conditional efficiency tests for the case in which a subset of asset holdings is potentially constrained. Their Wald statistic compares the squared Sharpe ratio of the de-hedged financial portfolio, the observed portfolio after subtracting the hedge demand implied by the constrained position, against the squared Sharpe ratio of the unconstrained financial frontier. Under the null that the de-hedged portfolio lies on the financial frontier, this statistic is distributed as $\chi^2(n - 1)$, where n denotes the number of freely traded risky assets. This is the test adopted in the present thesis, extended to accommodate two simultaneously constrained positions.

1.1.3 Estimation Risk

A central methodological concern in empirical portfolio analysis is estimation risk: the mean-variance efficient portfolio depends on parameters, expected returns, variances, and covariances, that must be estimated from historical data, and these estimates are noisy. Merton (1980) emphasised that the signal-to-noise ratio in historical return data is inherently low even over long samples. The finite-sample fragility of the estimated squared Sharpe ratio was established formally by Britten-Jones (1999), who showed that sampling error in efficient portfolio weights is large enough to render those weights statistically indistinguishable from zero in typical sample sizes.

Kan and G. Zhou (2007) extended this analysis, demonstrating that the expected value of the sample squared Sharpe ratio $\hat{S} = \hat{\mu}^\top \hat{\Sigma}^{-1} \hat{\mu}$ exceeds the population value by a term of order n/T , and that near-collinearity among regressors amplifies this overstatement as the covariance

matrix approaches singularity. This result is directly relevant to the present application, where bond categories are highly collinear over the sample period and geometrically declining weights reduce the effective sample size, potentially inflating the benchmark squared Sharpe ratio to economically implausible levels. This concern motivates the asymmetric weighting scheme adopted here, in which declining weights are applied exclusively to hedge-coefficient estimation rather than to the frontier benchmark.

Exponentially weighted moving-average estimators, popularised by the RiskMetrics framework, weight recent observations more heavily under the assumption that conditional covariance is time-varying (J.P. Morgan and Reuters 1996). Pelizzon and Weber (2008) employed geometrically declining weights with decay parameter $\alpha = 0.8$ to downweight earlier observations during the pre-euro convergence period, additionally rescaling expected long-term government bond returns to align with German Bond yields at end-1998. The present thesis adopts the recency-weighting component of this approach but does not replicate the bond-return rescaling step, which is unnecessary given the post-convergence, post-crisis sample of 2005–2024.

1.2 Housing, Illiquidity, and Optimal Portfolio Choice

1.2.1 Hedging Demands from Illiquid Housing Wealth

Housing occupies a distinctive position in household balance sheets. It is simultaneously a consumption good, a savings vehicle, and a source of idiosyncratic risk, and it is subject to adjustment costs so large that short-run rebalancing is effectively ruled out. The implications of this illiquidity for optimal portfolio choice were first analysed in a general-equilibrium setting by Mayers (1973), who showed that non-tradable assets alter the mean-variance frontier available to investors and generate hedging demands in the tradable-asset portfolio. Anderson and Danthine (1981) extended this insight to the general cross-hedging problem, deriving optimal decision rules for agents holding positions whose returns are imperfectly correlated with available hedging instruments. Although their analysis is cast in the context of commodity producers hedging in futures markets, the optimal hedge ratio they derive, a covariance-scaled adjustment to the unconstrained position, provides a natural template for the hedge-term decomposition employed by Pelizzon and Weber (2008). The optimal financial-asset holdings decompose into a Markowitz tangency component scaled by risk aversion and an additive hedge term proportional to the covariance between the constrained and tradable assets; when housing returns are correlated with financial returns, as the empirical evidence consistently suggests, this term is non-zero and must be accounted for before assessing portfolio efficiency.

The dynamic implications of housing illiquidity were explored by Damgaard, Fuglsbjerg, and Munk (2003), who allowed for two goods in the utility function and non-zero depreciation, confirming numerically that the hedge demand is quantitatively significant whenever housing returns are correlated with financial asset returns.

1.2.2 Housing and Portfolio Composition

The empirical literature on housing and portfolio choice began with characterisations of the risk-return properties of residential real estate. Goetzmann and Ibbotson (1990) documented that owner-occupied housing offers returns comparable to those of equities at substantially lower correlation with financial markets, suggesting a natural diversification role. Ross and Zisler (1991) confirmed substantial heterogeneity in regional price dynamics using real estate investment trust data, a feature that directly motivates the use of region-specific price indices rather than a single national series in the present analysis.

Flavin and Yamashita (2002) provided a systematic analysis of the interaction between housing constraints and financial portfolio composition using United States panel data, showing that the non-negativity constraint on owner-occupied housing shifts the mean-variance frontier available to homeowners relative to the unconstrained benchmark. Their framework assumes zero correlation between housing and financial returns, which simplifies the analysis but eliminates the hedging motive. A body of subsequent empirical work established that homeownership crowds out equity. More recently, Arrondel and Savignac (2015) showed for a large sample of French households that exposure to non-insurable real-estate risk significantly reduces both the probability of stock ownership and the share of wealth allocated to equities. The consistency of this finding across countries and sample periods suggests that the distortionary effect of housing on financial portfolios is a robust feature of household behaviour rather than an artefact of any single dataset.

Chetty, Sándor, and Szeidl (2017) distinguished the total value of the dwelling from home equity and showed that the two components exert opposing effects on financial portfolios: holding house value fixed, higher mortgage debt reduces stockholding, whereas greater home equity raises participation in risky assets. This decomposition reconciles the apparently uniform crowding-out result with the hedging logic, indicating that it is leverage rather than housing wealth per se that suppresses equity exposure, a mechanism consistent with Cauley, Pavlov, and Schwartz (2007), who found the housing constraint to be binding primarily for young and highly leveraged homeowners.

These cross-sectional regularities are mirrored in the life-cycle and dynamic evidence. Cocco (2005) solved a rich intertemporal optimisation problem integrating housing, human capital, borrowing constraints, and limited stock-market participation, demonstrating that young households with large mortgage obligations and uncertain labour income should hold little or no equity even in the absence of fixed participation costs; critically, he assumed zero correlation between housing and financial asset returns, thereby ruling out the hedge motive by assumption. Recent work has enriched this life-cycle picture along two dimensions. Paz-Pardo (2021) documented pronounced intergenerational differences, showing that younger cohorts are substantially less likely to own their homes than earlier cohorts were at the same age, a shift driven largely by changing earnings dynamics rather than by shifting preferences, with corresponding consequences for the composition of risky-asset holdings across generations. Exploiting Danish administrative panel data, Lyng and J. Zhou (2023) traced portfolio behaviour around the home-purchase event

directly, finding that equity-market participation contracts sharply as households approach a purchase and that the risky-asset share follows a V-shaped trajectory, declining in the run-up to acquisition and recovering thereafter. Together these studies confirm that the disruption of financial portfolios by housing operates not only across households of different tenure but also within households over the life cycle and around the acquisition decision itself.

The magnitude of these effects, however, hinges on the empirical correlation between housing and financial returns. Kullmann and Siegel (2005) confirm that the constraint imposed by the fixed housing stock generates a welfare cost that depends critically on this correlation structure, while Eichholtz, Koedijk, and Roon (2002) find that stocks and bonds provide limited hedging value for residential real estate in the United States, the correlation being sufficiently low that the optimal housing allocation is not substantially affected by financial-market access. Together, these studies illustrate that the portfolio implications of homeownership depend sensitively on the empirical correlation structure, precisely the feature that the Gouriéroux and Jouneau (1999) conditional efficiency framework is designed to accommodate rather than impose a priori.

1.2.3 Conditional Portfolio Efficiency: Evidence from Italy

The contribution that most directly motivates the present thesis is Pelizzon and Weber (2008), who showed that standard tests of portfolio efficiency are biased for homeowners because they neglect the hedge demand generated by illiquid housing wealth. Using Italian household portfolio data and time-series returns on financial assets and housing, they derived the optimal financial portfolio conditional on the existing housing stock and tested whether actual portfolios are efficient using the Gouriéroux and Jouneau (1999) conditional test. Their central finding is that the presence of illiquid wealth plays an important role in determining whether portfolios chosen by homeowners are efficient, and that conditioning on housing substantially alters the classification of households relative to purely financial efficiency tests. When only financial assets are considered, three portfolios out of four consist solely of the risk-free asset, and of the diversified portfolios a large fraction is mean-variance efficient. Once housing is included as an unconstrained asset, very few diversified portfolios are found to be efficient. When efficiency is tested conditional on housing, one in seven fully diversified portfolios is mean-variance efficient, and these are largely not the same households whose purely financial portfolios were deemed efficient. The key methodological insight is that this reclassification reflects the hedging motive: the optimal strategy for a homeowner is to use risky financial assets both to satisfy the standard Markowitz objective and to hedge exposure to housing price risk, a motive that vanishes only when housing and financial returns are uncorrelated. There is, moreover, strong evidence that hedging opportunities are not fully exploited even by those Italian households who hold well-diversified portfolios, a finding with important implications for portfolio management that the present thesis re-examines and extends using the 2022 SHIW.

1.3 Life Insurance and Pension Products

While housing has received the majority of attention in the portfolio-efficiency literature, life insurance and pension savings represent a second category of illiquid or semi-liquid wealth that is economically important and poorly integrated into household portfolio studies. In Italy, the supplementary pension system encompasses contractual pension funds (*fondi pensione negoziali*), open pension funds, and individual pension plans (PIPs), alongside savings-type life insurance products: Class I participating (*gestione separata*) policies and Class III unit-linked products.

The theoretical case for including insurance and pension products in the portfolio optimisation problem mirrors that for housing. When holdings are predetermined, either because of multi-year contractual commitments, surrender penalties, or the effective illiquidity of the secondary market, the household cannot rebalance freely, so the optimal financial portfolio must embed a hedge against the return risk of the constrained position.

The Italian supplementary pension system underwent substantial regulatory transformation following the pension reforms of the 1990s and the introduction of PIPs under Legislative Decree No. 252 of 2005. Brugiavini and Fornero (2001) provide an early analysis of the Italian pension system, documenting that the historical dominance of the public pay-as-you-go system crowded out private retirement savings and limited the accumulation of funded pension wealth. The gradual development of the supplementary pillar created new investment vehicles with insurance characteristics but defined-contribution accumulation profiles, introducing a category of semi-liquid constrained wealth that has, to our knowledge, no precedent in the earlier Italian portfolio literature and that the present thesis is the first to incorporate explicitly into the conditional efficiency framework.

1.4 The Research Gap

The foregoing review establishes the following state of the literature. First, the conditional efficiency framework of Gouriéroux and Jouneau (1999) provides a rigorous and tractable approach to testing whether household financial portfolios are mean-variance efficient once a pre-existing illiquid position is taken as given. Second, this framework has been applied exclusively to the single-constrained-asset case, with housing as the illiquid position. Third, the Italian household balance sheet includes a second category of semi-liquid constrained wealth, life insurance and supplementary pension products, with a distinct covariance profile relative to both housing and the tradable financial assets. Fourth, the 2022 SHIW provides a substantially updated microdata base relative to the 1998 wave used by Pelizzon and Weber (2008), capturing two decades of structural change in Italian financial markets, housing markets, and the supplementary pension system.

Chapter 2

SHIW 2022

The **Survey on Household Income and Wealth (SHIW)** is a long-running data collection effort by the Bank of Italy, initiated in the 1960s to document the incomes and savings of Italian households. Over time, its scope has broadened considerably to encompass household wealth, financial behaviour, and payment methods.

The 2022 wave provides a comprehensive snapshot of the Italian household sector following the pandemic shock. Mean household income recovered modestly in real terms (+1.4%), yet remained below pre-financial-crisis levels observed in 2006. Wealth dynamics were similarly uneven: mean net wealth rose by 1.8% in real terms relative to 2020, while the median fell by 2%, and the share held by the wealthiest decile increased by approximately 2 percentage points to reach 52%. More than half of all households reported no savings in 2022, with the figure climbing to 70% among the lowest income quintile.

These distributional features make the SHIW 2022 a particularly suitable basis for household portfolio analysis. The availability of granular asset-level microdata covering financial instruments, real estate, insurance, and pension holdings alongside detailed household characteristics allows the construction of the full portfolio vectors required for the efficiency tests developed in Chapter 4.

2.1 Sample Construction and Variable Overview

The 2022 wave is released as a set of microdata files, which were merged into a single master dataset in **Stata**. Table 2.1 lists each file alongside the variables it contributes to the analysis.

Table 2.1: SHIW 2022 Microdata Files

File	Content	Key Variables
carcom22	Individual/demographic characteristics	Age, education, marital status, region, sampling weight
q22c2	Financial assets (amounts and participation)	Asset amounts and participation dummies by instrument

q22f	Insurance/pension overview	Life-insurance and pension participation flags
allf1	Supplementary pension plans (Annex F1)	Plan type (Occupational/Open/PIP), fund value
allf2	Life-insurance policies (Annex F2)	Savings component, guaranteed return, policy value
q22d	Principal residence & debt ownership	Homeownership, home value, mortgage and consumer-debt flags
alld1	Other residential properties (Annex D1)	Residential type, value, ownership share
ricfam22	Household wealth	Total household wealth aggregate
alld2_res	Outstanding mortgage: principal residence	Mortgage balance
alld2_aimm	Outstanding mortgage: other properties	Mortgage balance
alld2_fam	Outstanding consumer/other debt	Debt balance

The unit of observation is the **household**, identified by the head of household (`cfred == 1`), yielding a base sample of **9,641 households**. All estimates are computed using the household sampling weight **pesofit**, declared via Stata's `svyset` command, ensuring results are representative of the Italian household population.

2.2 Financial Asset Classification and Holdings

Financial holdings were recorded for **15 instrument types** (deposits, repos, managed accounts, BOTs, BTPs, CTZs, corporate bonds, bank bonds, mutual funds, ETFs, listed and unlisted equities, foreign deposits, other foreign assets, and cooperative loans). For each instrument two variables were constructed:

- **Amount** (`af_*`): euro value of the position.
- **Participation** (`own_*`): a binary indicator equal to one if the household reports holding the instrument (Deposits from `pdepos`; the remaining fourteen from the `pos_*` flags).

Total directly reported financial assets are $af_total = \sum_j af_j$ over the fifteen instruments.

2.3 Life Insurance and Pension Product Classification

Life-insurance policies (Annex F2) were classified into **Class I** and **Class III**, both conditional on holding a savings-type policy (`ass1v == 1`):

- **Class I** (`life_Class1`): savings policy (`ass1v == 1`) *with* a guaranteed minimum return (`ass1r == 1`).
- **Class III** (`life_Class3`): savings policy (`ass1v == 1`) *without* a guaranteed return (`ass1r ≠ 1`).

Pension plans (Annex F1) were classified by `ass2gn`:

Table 2.2: Classification of supplementary pension plans

Code	Type	Variable
1	Contractual (occupational) fund	<code>pen_occupational</code>
2	Open fund	<code>pen_open</code>
3	Personal/PIP	<code>pen_personal_pip</code>

Household-level type flags were obtained by taking the household maximum across all relevant records (`collapse (max) ... by(nquest)`), so a flag equals **1** if any plan or policy of that type is held. The binary overview indicators `hh_vita` and `hh_any_pension` were instead constructed across the up-to-six member slots of `q22f`.

2.4 Financial and Total Wealth

Two portfolio-share denominators were constructed:

(1) Financial wealth (`af_total`): directly reported financial assets only. Conditional portfolio shares $\text{share}_j = \text{af}_j / \text{af_total}$ are defined for the **8,823 households** with `af_total > 0`; the remaining **818 households** report *zero financial-asset amounts* (`af_total = 0`), flagged by `share_nofinassets` and excluded from conditional share calculations.¹

(2) Total Wealth (`tw_basis`): a household-level net-worth construct, *not* a comprehensive balance sheet. It is defined as:

$$\begin{aligned} \text{tw_basis} = & \text{af_total} + \text{Home Value} - \text{Mortgage Debt} - \text{Other Debt} \\ & + \text{Life Insurance Value} + \text{Pension Fund Value} \end{aligned}$$

Total-wealth shares `tshare_j` express each asset or liability as a fraction of `tw_basis`, with mortgage and consumer debt entering as **negative shares**. They are computed for households with strictly positive net worth (`tw_basis > 0`); households with `tw_basis = 0` contribute zeros and households with negative net worth are excluded from the total-wealth means.

¹This amount-based definition can differ from participation: a household may report a deposit account while recording a zero balance, which is why the zero-financial-asset count (8.5% of the sample) exceeds the share of households without a deposit account (2.8%).

2.5 Regional Demographic Weights

Regional demographic weights were obtained by collapsing the SHIW 2022 sample to the region level with `pesofit` as the weighting variable. Each of the 20 ISTAT regions was assigned to one of four macro-areas² (North-West, North-East, Centre, South), and two weights were derived:

$$w_{\text{within macro}, r} = \frac{\sum_{HH \in r} \text{pesofit}}{\sum_{HH \in m} \text{pesofit}}, \quad w_{\text{within Italy}, r} = \frac{\sum_{HH \in r} \text{pesofit}}{\sum_{\text{all } HH} \text{pesofit}}.$$

These weights were saved to `regional_weights_shiw22.dta`, and the corresponding results are reported in Table 2.3.

Table 2.3: Regional Weights within Macroregions and Italy

Region	$w_{\text{within macro}}$	$w_{\text{within Italy}}$
Piemonte	0.3044	0.0813
Valle d'Aosta	0.0646	0.0172
Lombardia	0.4907	0.1311
Liguria	0.1403	0.0375
Trentino-Alto Adige	0.2230	0.0465
Veneto	0.2945	0.0614
Friuli-Venezia Giulia	0.1286	0.0268
Emilia-Romagna	0.3540	0.0738
Toscana	0.3215	0.0648
Umbria	0.1324	0.0267
Marche	0.2067	0.0416
Lazio	0.3395	0.0684
Abruzzo	0.0489	0.0158
Molise	0.0252	0.0082
Campania	0.2360	0.0762
Puglia	0.2081	0.0672
Basilicata	0.0562	0.0181
Calabria	0.0841	0.0272
Sicilia	0.2097	0.0677
Sardegna	0.1318	0.0426

²The mapping from ISTAT region codes to macro-areas, as implemented in the program, is: **North-West**: Piemonte, Valle d'Aosta, Lombardia, Liguria; **North-East**: Trentino-Alto Adige, Veneto, Friuli-Venezia Giulia, Emilia-Romagna; **Centre**: Toscana, Umbria, Marche, Lazio; **South**: Abruzzo, Molise, Campania, Puglia, Basilicata, Calabria, Sicilia, Sardegna.

2.6 Asset-Class Participation and Portfolio Combinations

To characterise observed portfolio choices, each household was assigned four broad asset-class participation flags:

Table 2.4: Asset-class participation flags

Flag	Definition
has_rf	Holds deposits, repos, or BOTs
has_housing	Holds any residential property
has_ins	Holds Class I, Class III, or any pension fund (occupational, open, or PIP)
has_risky	Holds any other financial instrument or carries mortgage/consumer debt

From these flags, **eight mutually exclusive combinations** were constructed, all conditional on holding at least the risk-free asset. The 306 households (9,641 – 9,335) holding no risk-free asset are therefore excluded from Table 2.5.

Table 2.5: Asset combination by macro-region

Combination	Total	NW	NE	CE	SO
RF	13.7% (N=1276)	12.3% (N=224)	10.7% (N=207)	10.8% (N=159)	16.7% (N=686)
RF + H	42.4% (N=3956)	28.2% (N=511)	32.0% (N=617)	39.7% (N=585)	54.5% (N=2243)
RF + I	0.7% (N=61)	0.6% (N=11)	0.8% (N=16)	0.6% (N=9)	0.6% (N=25)
RF + I + H	3.8% (N=359)	3.3% (N=59)	4.4% (N=85)	3.9% (N=58)	3.8% (N=157)
RF + RK	2.1% (N=199)	2.6% (N=48)	2.8% (N=54)	1.8% (N=26)	1.7% (N=71)
RF + RK + H	26.5% (N=2473)	35.6% (N=645)	34.9% (N=674)	30.3% (N=446)	17.2% (N=708)
RF + RK + I	0.5% (N=48)	0.7% (N=13)	0.6% (N=11)	0.8% (N=12)	0.3% (N=12)
RF + RK + H + I	10.3% (N=963)	16.7% (N=303)	13.8% (N=266)	12.1% (N=179)	5.2% (N=215)
Total (RF holders)	100.0% (N=9335)	100.0% (N=1814)	100.0% (N=1930)	100.0% (N=1474)	100.0% (N=4117)

Note. Abbreviations: RF = Risk-Free; RK = Risky; H = Housing; I = Insurance. Percentages are **unweighted**: for each cell, both the percentage and the parenthetical *N* are computed from simple household counts.

2.7 Asset Participation Rates and Portfolio Shares

Three sets of survey-weighted statistics were produced from the master dataset:

- **Participation rates:** weighted share of households owning each of the 15 financial instruments, plus principal-residence ownership, mortgage, consumer debt, and the pension/insurance products.
- **Financial-wealth shares:** conditional mean `share_*` for the 8,823 households with positive financial wealth.
- **Total-wealth shares:** conditional mean `tshare_*` for households with `tw_basis > 0`.

Table 2.6: Participation and Portfolio Shares

Asset	Participation (%)	Financial Wealth (%)	Total Wealth (%)
Deposits	97.22	84.98	30.04
Repos	3.44	1.31	0.35
Managed accounts	4.85	2.67	0.87
BOTs	2.78	1.25	0.40
BTPs	3.49	1.28	0.37
CTZs	0.72	0.21	0.05
Corporate bonds	1.81	0.45	0.18
Bank bonds	3.81	1.45	0.38
Mutual funds	8.88	4.53	1.37
ETFs	0.82	0.19	0.06
Listed equity	4.36	0.99	0.29
Unlisted equity	0.61	0.18	0.06
Foreign deposits	0.31	0.06	0.01
Other foreign assets	0.43	0.10	0.03
Cooperative loans	0.84	0.36	0.09
Housing	74.55	—	75.63
Mortgage	13.92	—	−10.24
Consumer/other debt	10.53	—	−2.53
Occupational pension	4.06	—	0.84
Open pension	1.85	—	0.20
PIP	5.06	—	0.74
Class III	4.46	—	0.50
Class I	2.76	—	0.31

Table 2.7 reports the macro-regional breakdown of the four headline total-wealth shares, housing, Class I and Class III life insurance, and personal pension plans (PIP), estimated via

svy, subpop() on each macro-region indicator. Housing wealth shares range from 68.09% of net worth in the North-East to 84.24% in the Centre, while insurance and pension shares remain comparatively small in every region; the largest regional spread is observed for PIP holdings, from 0.18% of net worth in the South to 1.17% in the North-East.

Table 2.7: Total-wealth shares by macro-region

Region	Housing	Class I	Class III	PIP
Italy	75.63	0.31	0.50	0.74
North-West	76.38	0.42	0.60	0.93
North-East	68.09	0.13	0.49	1.17
Centre	84.24	0.29	0.61	0.95
South	74.51	0.35	0.35	0.18

Note. Estimates from svy, with the macro-region columns estimated via svy, subpop() on the corresponding macro_region indicator; all figures are percentages of tw_basis.. The “Italy” row reproduces the national tshare_house, tshare_Class1, tshare_Class3, and tshare_pen_pip figures from Table 2.6.

Chapter 3

Return Data

This chapter describes the return data used to construct the mean-variance frontier against which household portfolio efficiency is assessed. The sample covers the period 2005–2024, yielding twenty annual observations. The asset universe comprises thirteen instruments spanning five broad categories: fixed-income securities, equity indices, life insurance products, supplementary pension vehicles, and regional residential real estate. Because housing and insurance-linked returns are available only at an annual frequency, all higher-frequency series are converted to annual returns prior to any estimation, ensuring a temporally homogeneous panel across asset classes.

Section 3.1 introduces the full asset universe. Sections 3.2–3.7 detail the data sources and conversion methodology for each asset class. Section 3.8 describes how the thirteen individual series are consolidated into eight broad asset classes to preserve statistical tractability in portfolio optimisation.

3.1 Asset Universe and Data Overview

Return data were collected for thirteen assets, spanning fixed-income, equity, insurance-linked, pension, and real estate instruments:

1. Risk-Free Rate (BOTs yields)
2. Rendistato 2Y
3. 3Y BTP
4. 5Y BTP
5. MSCI EUR Investment Grade Corporate Bond Index
6. MSCI Italy
7. MSCI World (EUR)
8. Class I

9. Class III
10. Open Pension Funds
11. Contractual Pension Funds
12. PIPs (Individual Pension Plans)
13. Housing (Regional)

Raw series were collected at mixed frequencies. Because housing and insurance-linked returns are observable only at an annual frequency, all monthly series were converted to annual returns prior to estimation, in order to ensure a temporally homogeneous panel across asset classes. The aggregation method employed for this conversion is asset-specific and is detailed in the sections that follow.

3.2 Risk-Free Rate: Treasury Bill Yields

The risk-free rate is constructed from monthly BOTs (Buoni Ordinari del Tesoro) yields. The annual rate is defined as the arithmetic average of the twelve monthly yields observed within each calendar year. Since BOTs are short-term Treasury bills that are continuously rolled over, the arithmetic averaging of monthly yields provides an appropriate measure of the opportunity cost of capital over the year, without introducing compounding distortions that would not reflect the instrument's underlying roll-over structure.

3.3 The RENDISTATO Index

The RENDISTATO index, available at monthly frequency, is converted to annual returns through a two-step procedure.

Step 1: Monthly net return. Monthly gross yields are converted into monthly net returns according to

$$r_t = \left(\frac{Y_t}{Y_{t-1}} \right)^{\frac{1}{12}} (1 - \tau) - D (Y_t - Y_{t-1})$$

where the modified duration is set to $D = 2$ years, consistent with the two-year maturity target of the RENDISTATO index, and $\tau = 0.125$ denotes the Italian tax rate applicable to government bond income.

Step 2: Annual compounding. Monthly net returns are compounded geometrically to obtain the annual net return for each year:

$$R_{\text{annual}} = \prod_{t=1}^{12} (1 + r_t) - 1$$

As the RENDISTATO is a synthetic index tracking medium-term Italian government bonds, this duration-based price approximation jointly captures the carry component (yield income) and the capital gain or loss arising from monthly movements in yields.

3.4 3Y and 5Y BTPs

The 3Y and 5Y BTP (Buoni del Tesoro Poliennali) series, both available at monthly frequency, are converted to annual returns using the same two-step procedure applied to the RENDISTATO index.

Step 1: Monthly net return. Monthly gross yields are converted into monthly net returns according to

$$r_t = \left(\frac{Y_{t-1}}{12} \right) (1 - \tau) - D (Y_t - Y_{t-1})$$

with modified duration set to $D = 3$ for the 3Y BTP and $D = 5$ for the 5Y BTP, and $\tau = 0.125$ again denoting the Italian tax rate on government bond income.

Step 2: Annual compounding. Monthly net returns are compounded geometrically:

$$R_{\text{annual}} = \prod_{t=1}^{12} (1 + r_t) - 1$$

BTPs are fixed-coupon bonds with semi-annual coupon payments. In the first formula above, the first term represents one month of accrued coupon income (the annualised yield divided by twelve), while the second term approximates the capital gain or loss resulting from the change in yield.

3.5 Equity and Corporate Bond Indices

The MSCI EUR Investment Grade Corporate Bond Index, the MSCI Italy Index, and the MSCI World Index (EUR) are each available as monthly net total returns. For all three series, annual returns are obtained through geometric compounding, applying the same formula used for government bond series above. This treatment is appropriate because index total returns already embed coupon or dividend reinvestment together with price changes, such that geometric compounding constitutes the correct method for aggregating multi-period returns. For the MSCI Investment Grade Corporate Bond Index, geometric compounding is consistent with the reinvestment of coupon income embedded in the total return series; for the two equity indices, it preserves the compounding nature of price appreciation and dividend reinvestment. The EUR denomination of the MSCI World Index further ensures that currency effects are already embedded in the series, so that no additional foreign-exchange adjustment is required.

3.6 Insurance and Pension Products

3.6.1 Class I and Class III Life Insurance Products

Class I and Class III returns are obtained directly as net annual figures from ANIA (Associazione Nazionale fra le Imprese Assicuratrici) Annual Reports and require no further transformation.

Class I returns are net of management fees and reflect the smoothed, book-value accounting characteristic of Italian *gestione separata* (separately managed) funds; they are reported on an annual basis as mandated by Italian insurance regulation. Class III returns, by contrast, are market-linked and are reported net of charges, also on an annual basis.

3.6.2 Contractual and Open Pension Funds

Open Pension Fund and Contractual Pension Fund (*Fondi Pensione Negoziati*) returns are obtained directly as net annual figures from COVIP (Commissione di Vigilanza sui Fondi Pensione), the Italian pension supervisory authority, which publishes standardised net annual performance statistics for both fund categories.

3.6.3 PIPs: Individual Pension Plans

PIP returns are also obtained as net annual figures from COVIP. They are insurance-based pension products offered exclusively by insurance companies; net returns are reported annually by insurers under the joint supervisory oversight of COVIP and IVASS¹.

As a product category, PIPs were introduced into Italian regulation only in 2007 implying that no observed return data exist prior to 2008. To extend the sample period by three years and to preserve a consistent asset universe across all series, returns for the 2005–2007 period are estimated via an OLS regression on the two closest economic substitutes available over the full sample:

$$R_{PIP,t} = \alpha + \beta_1 R_{RF,t} + \beta_2 R_{MT\ Bond,t} + \varepsilon_t$$

The regression yields an R^2 of 0.69 and an adjusted R^2 of 0.65.

The choice of regressors is motivated by the predominant underlying sub-fund type within PIPs, particularly in the early years of the product: a guaranteed, capital-protected profile that behaves similarly to a short-to-medium duration fixed-income instrument. The Risk-Free rate (BOTs) and the Rendistato 2Y (medium-term government bonds) constitute the natural benchmarks for this profile, and the regression exploits this structural relationship to produce economically plausible out-of-sample estimates for the pre-2008 period.

¹COVIP defines “new” PIPs as individual pension plans implemented through life insurance contracts, established after the entry into force of Legislative Decree No. 252 of 5 December 2005. They are subject to the same rules applicable to new pension funds with respect to members’ rights, portability, and transparency; their assets must be legally separated from those of the sponsoring insurance company. They may take the form of with-profit (traditional) or unit-linked policies. See COVIP, *Glossario*, s.v. “PIP nuovi”.

3.7 Housing

3.7.1 Regional House Price Indices

Regional house price indices are sourced from the *Indice delle Quotazioni Immobiliari*, published by OMI (Osservatorio del Mercato Immobiliare, Agenzia delle Entrate) at the regional level. Regional price indices are converted into annual capital gain returns according to

$$R_{\text{capital}, r, t} = \frac{P_{r, t} - P_{r, t-1}}{P_{r, t-1}}$$

where r denotes the region and t the year.

3.7.2 Net Total Return Adjustment

A fixed net income yield per annum is added to the capital gain return of each region to obtain the net total return:

where κ denotes the net rental income yield (gross rent minus recurring maintenance costs) expressed as a fraction of the property value. Following Flavin and Yamashita (2002), we set $\kappa = 0.05$ on an annual basis, applied uniformly across regions².

$$R_{\text{net}, r, t} = R_{\text{capital}, r, t} + \kappa = R_{\text{capital}, r, t} + 5\%$$

Importantly, as noted by Pelizzon and Weber (2008), the specific value of κ is *immaterial for the efficiency tests in the constrained case* as long as it is a fixed number, because κ cancels in the Wald statistic (see Equation 4.15).

3.7.3 Regional Annual Net Returns

Regional returns are aggregated into macro-area and national figures using the demographic weights extracted from the Bank of Italy's Survey on Household Income and Wealth (SHIW) 2022, reported in Table 2.3. Two sets of weights are computed:

1. **Within-macro-area weight:** for each region r belonging to macro-area m , the share of households residing in that region relative to all households in the same macro-area.
2. **National weight:** each region's share of households relative to the total Italian sample.

Macro-area and national returns are then obtained as the weighted average of regional net returns:

$$R_{m, t} = \sum_{r \in m} w_{r|m} R_{\text{net}, r, t}, \quad R_{\text{ITA}, t} = \sum_{\forall r} w_{r|\text{ITA}} R_{\text{net}, r, t}$$

²We verified that the choice of κ does not affect our findings: re-running the analysis of Chapter 5 with $\kappa = 0.02$ and $\kappa = 0.03$ left the results substantively unchanged. We therefore retain $\kappa = 0.05$ for consistency with the literature.

The four macro regions are defined in Section 2.5.

Weighted average annual net returns are then computed for each of the five geographic aggregates (NW, NE, CE, SO, and Italy as a whole), yielding five distinct net annual real estate return series. Each series reflects the local price dynamics of the corresponding macro-area, weighted by the actual demographic composition of Italian households residing there.

3.8 Consolidation into Broad Asset Classes

The granular asset universe of twelve series, combined with a limited number of annual observations, poses a degrees-of-freedom problem for portfolio optimisation: an excessive number of assets relative to the number of observations produces unreliable covariance matrix estimates and unstable mean-variance solutions. To mitigate this issue, the underlying assets were consolidated into eight broad asset classes (excluding the risk-free rate) that preserve meaningful risk-return differentiation while ensuring statistical tractability. Table 3.1 summarises the resulting classification.

Table 3.1: Individual Assets to Broad Asset Classes

Broad Asset	Individual Assets	Weights
Stocks	MSCI Italy Index, MSCI World Index (EUR)	50% / 50%
Corporate Bonds	MSCI EUR Investment Grade Corporate Bond Index	-
MT Govt Bonds	Rendistato 2Y	-
LT Govt Bonds	HPR 5Y BTP, HPR 3Y BTP	50% / 50%
PIPs	PIPs	-
Class III	Class III, Open Pension Funds, Contractual Pension Funds	1/3 each
Class I	Class I	-
House	NW, NE, CE, SO, ITALY	-

3.8.1 Stocks: Equal Weights of MSCI Italy and MSCI World

The equal weighting of domestic and foreign equity exposure follows Pelizzon and Weber (2008), who motivate this split with reference to Bank of Italy survey evidence. Because direct stockholdings by Italian households are predominantly domestic, while indirect holdings through mutual funds are predominantly foreign, the 50/50 average is intended to capture the blended equity exposure of a representative Italian investor.

3.8.2 Class III Aggregate: Equal-Weight Specification

The three constituent vehicles, namely Class III policies, Open Pension Funds, and Contractual Pension Funds, are legally distinct but share the same underlying economic exposure:

market-linked returns within a tax-deferred wrapper. Their respective assets under management were broadly comparable over the sample period, and no stable, time-invariant market-share split could be identified from the available data. An equal one-third weighting is therefore applied.

3.8.3 Long-Term Government Bonds: Duration-Matched Specification

Long-term government bonds are modelled as an equally weighted portfolio of two instruments with durations of three and five years, respectively, following directly the benchmark specification of Pelizzon and Weber (2008), who construct holding period returns for long-term bonds as an equally weighted portfolio of assets with the same durations. Replicating this specification is necessary to ensure that the mean-variance frontier estimated in this study is directly comparable to theirs, and that any differences in efficiency outcomes can be attributed to sample and household characteristics rather than to divergences in asset construction.

Chapter 4

ECONOMETRIC FRAMEWORK

This chapter develops the econometric framework used to test household portfolio efficiency in the presence of two constrained asset positions: housing and life insurance. Section 4.1 extends the single-constrained-asset model of Pelizzon and Weber (2008) to accommodate both constrained holdings simultaneously, deriving the optimal financial-asset allocation as a Markowitz tangency portfolio adjusted by two additive hedge terms. Section 4.2 describes how the hedge coefficients are estimated via regressions of each constrained asset's excess return on the four financial factors over the 2005–2024 sample. Section 4.3 presents the Gouriéroux and Jouneau (1999) conditional efficiency test, which assesses whether a household's de-hedged financial portfolio attains the maximum Sharpe ratio on the unconstrained frontier, and introduces the five test variants applied in the empirical analysis.

4.1 The Two-Constrained-Asset Mean–Variance Problem

In this section we show how housing *and* insurance wealth can be introduced into the standard mean–variance one-period model, extending the single-constrained-asset framework of Pelizzon and Weber (2008) to the empirically relevant case in which the household holds two non-tradable, constrained positions¹.

We derive the optimal financial-asset holdings in the static mean–variance framework when the existing housing stock and the insurance position are both treated as additional constraints on the optimisation problem (see Mayers (1973) and Anderson and Danthine (1981) for the general case of a constrained asset)

Consider a market with a riskless asset, n unconstrained risky financial assets, and *two* constrained risky assets, indexed H (housing) and I (insurance). Denote the first two moments of excess returns by $m = (\mu^\top, \mu_C^\top)^\top$, where μ is the $n \times 1$ vector of expected excess returns on the financial assets and $\mu_C = (\mu_H, \mu_I)^\top$, and by Ω the variance–covariance matrix of excess

¹For expository clarity the constrained block is written with two assets, housing (H) and insurance (I). In the empirical implementation the insurance position is disaggregated into its three economically distinct products, individual pension plans (PIPs), Class I, and Class III (unit-linked), whose risk profiles differ markedly (see Table 5.1). The constrained set is therefore $C = \{H, \text{PIPs}, \text{Class I}, \text{Class III}\}$, so that Γ is $n \times |C|$ and Σ_C is $|C| \times |C|$, with one hedge term $\bar{h}_j \Sigma^{-1} \Gamma_{bj}$ per element of C .

returns. The covariance matrix partitions into blocks corresponding to the n financial assets and the two constrained assets:

$$\Omega = \begin{bmatrix} \Sigma & \Gamma \\ \Gamma^\top & \Sigma_C \end{bmatrix}, \quad \text{where} \quad \Gamma = \begin{bmatrix} \Gamma_{bH} & \Gamma_{bI} \end{bmatrix}, \quad \text{and} \quad \Sigma_C = \begin{bmatrix} \sigma_H^2 & \sigma_{HI} \\ \sigma_{HI} & \sigma_I^2 \end{bmatrix}, \quad (4.1)$$

where Σ ($n \times n$) is the financial covariance matrix, the $n \times 2$ matrix Γ collects in its columns the covariances of the financial assets with housing (Γ_{bH}) and with insurance (Γ_{bI}), and the 2×2 block Σ_C describes the covariances among the two constrained assets.

Consider an investor whose allocation in the risky assets is

$$Z = \begin{pmatrix} x_0 \\ h \end{pmatrix}, \quad \text{where} \quad x_0 \equiv \frac{X_0}{W_0}, \quad \text{and} \quad h \equiv \frac{1}{W_0} \begin{pmatrix} H_0 P_0^H \\ I_0 P_0^I \end{pmatrix} = \begin{pmatrix} h_H \\ h_I \end{pmatrix}, \quad (4.2)$$

x_0 collects the n financial-asset shares and $h = (h_H, h_I)^\top$ the housing and insurance shares, each expressed as a fraction of total wealth W_0 . Assume the investor is constrained in h (that is, h is given and equal to $\bar{h} = (\bar{h}_H, \bar{h}_I)^\top$), but otherwise behaves according to the mean–variance model. The investor’s problem is

$$\begin{aligned} \min_Z \quad & Z^\top \Omega Z \\ \text{s.t.} \quad & \begin{cases} Z^\top m + r_f = m^*, \\ h = \bar{h}. \end{cases} \end{aligned} \quad (4.3)$$

where m^* is a given level of expected return. Imposing $h = \bar{h}$ and expanding the quadratic form, the portfolio variance, and the return constraint expand as:

$$Z^\top \Omega Z = x_0^\top \Sigma x_0 + 2 x_0^\top \Gamma \bar{h} + \bar{h}^\top \Sigma_C \bar{h}, \quad (4.4)$$

$$Z^\top m + r_f = x_0^\top \mu + \bar{h}^\top \mu_C + r_f. \quad (4.5)$$

The problem can be solved by defining the Lagrangian

$$\Lambda = (x_0^\top \Sigma x_0 + \bar{h}^\top \Sigma_C \bar{h} + 2 x_0^\top \Gamma \bar{h}) - 2\gamma(x_0^\top \mu + \bar{h}^\top \mu_C + r_f - m^*). \quad (4.6)$$

The first-order conditions are

$$\frac{\partial \Lambda}{\partial x_0} = 2 \Sigma x_0 + 2 \Gamma \bar{h} - 2\gamma \mu = 0, \quad (4.7)$$

$$\frac{\partial \Lambda}{\partial \gamma} = x_0^\top \mu + \bar{h}^\top \mu_C + r_f - m^* = 0. \quad (4.8)$$

and the solution is:

$$x_0 = \gamma \Sigma^{-1} \mu - \Sigma^{-1} \Gamma \bar{h}, \quad (4.9)$$

where γ is the Lagrange multiplier on the expected-return constraint, which carries the standard

relative-risk-aversion interpretation (Samuelson (1970)). Since $\Gamma \bar{h} = \Gamma_{bH} \bar{h}_H + \Gamma_{bI} \bar{h}_I$, the solution decomposes into the Markowitz portfolio net of two additive hedge terms, one for each constrained asset:

$$x_0 = \underbrace{\gamma \Sigma^{-1} \mu}_{\text{Markowitz}} - \underbrace{\bar{h}_H \Sigma^{-1} \Gamma_{bH}}_{\text{housing hedge}} - \underbrace{\bar{h}_I \Sigma^{-1} \Gamma_{bI}}_{\text{insurance hedge}} . \quad (4.10)$$

This result means that investors must be efficient with respect to the risky financial assets, choosing the efficient Markowitz portfolio according to their risk aversion (see H. M. Markowitz, Todd, and Sharpe (2000)); but they also use the risky financial assets to hedge their exposure to *both* constrained positions. The two hedge terms enter additively, and each depends only on its own covariance column Γ_{bj} , so the hedge demand generated by housing is computed independently of that generated by insurance. The constant term $\bar{h}^\top \Sigma_C \bar{h}$ and, more generally, the cross-block Σ_C do not enter (4.9): the covariance between housing and insurance affects the variance of the fixed constrained sub-portfolio but not the optimal allocation among financial assets. Finally, the hedge correction vanishes only if *both* covariance columns are zero, $\Gamma_{bH} = \Gamma_{bI} = 0$, in which case portfolio choice separates fully between financial and real assets. If either constrained asset co-moves with the financial assets, conditioning on that holding remains necessary for the portfolio to be assessed correctly against the efficient frontier.

4.2 Hedge Coefficient Estimation

The hedge coefficients, $\Sigma^{-1} \Gamma_{bH}$ for housing and $\Sigma^{-1} \Gamma_{bI}$ for insurance, are estimated by regressing each constrained asset's annual excess return on the excess returns of the four financial asset classes (stocks, corporate bonds, medium-term government bonds, and long-term government bonds) over the 2005–2024 sample. The estimating equation is

$$r_{j,t} = \alpha_j + \beta_j^\top r_t + \varepsilon_{j,t} \quad (4.11)$$

where $r_{j,t}$ is the excess return on constrained asset j in year t , r_t is the vector of financial-asset excess returns, and the estimated slope vector $\hat{\beta}_j$ is the empirical counterpart of the hedge coefficient $\Sigma^{-1} \Gamma_{bj}$. The intercept α_j captures the abnormal, un-hedgeable component of the asset's excess return and does not enter the hedge term.

4.3 The Gouriéroux–Jouneau Conditional Efficiency Test

Having estimated the hedge coefficients $\Sigma^{-1} \Gamma_{bH}$ and $\Sigma^{-1} \Gamma_{bI}$ in the previous section, we now test whether observed household portfolios are mean–variance efficient. Gouriéroux and Jouneau (1999) derive efficiency tests for the conditional, or constrained, case, that is, the case in which a subset of asset holdings is potentially constrained (housing and insurance in our application). The procedure asks whether a household's financial portfolio coincides with the efficient (tangency) portfolio of the unconstrained financial frontier *once the hedge demand*

implied by its constrained positions has been added back; the test then compares squared Sharpe ratios computed on this common, de-hedged basis.

For the unconstrained set of tradable financial assets, the maximal squared Sharpe ratio attainable on the financial frontier is

$$S_1 = \mu^\top \Sigma^{-1} \mu. \quad (4.12)$$

In our application the financial investment set comprises four tradable risky assets, stocks, corporate bonds, and medium- and long-term government bonds, and the estimated benchmark is $\hat{S}_1 = 0.5030$.

For a household holding the constrained positions $\bar{h} = (\bar{h}_H, \bar{h}_I)^\top$, the relevant object of comparison is not the raw financial portfolio x_0 but the portfolio obtained after undoing the hedge demand it embeds. Recall from the optimal rule (4.10) that an efficient household sets $x_0 = \gamma \Sigma^{-1} \mu - h_H \Sigma^{-1} \Gamma_{bH} - h_I \Sigma^{-1} \Gamma_{bI}$; adding the two hedge terms back therefore isolates the speculative (tangency) component. The de-hedged financial portfolio is

$$v_1 = x_0 + h_H \Sigma^{-1} \Gamma_{bH} + h_I \Sigma^{-1} \Gamma_{bI}, \quad (4.13)$$

which conditions jointly on both the housing and the insurance holding. Under the null that the household is conditionally efficient, v_1 is proportional to the tangency direction $\Sigma^{-1} \mu$.

The squared Sharpe ratio of the de-hedged portfolio is

$$S_1(Z) = \frac{(\mu^\top v_1)^2}{v_1^\top \Sigma v_1}, \quad (4.14)$$

and, when all asset returns are normally distributed, Gouriéroux and Jouneau (1999) show that the Wald statistic

$$\xi_1 = T \frac{\hat{S}_1 - \hat{S}_1(Z)}{1 + \hat{S}_1(Z) \frac{Z^\top \Omega Z}{v_1^\top \Sigma v_1}} \sim \chi_{n-1}^2 \quad (4.15)$$

is distributed as a chi-square random variable with $n - 1$ degrees of freedom² under the null hypothesis that the de-hedged financial portfolio v_1 lies on the financial efficient frontier, where $T = 20$ is the number of annual return observations.

The intuition behind the test is as follows. Once a constrained asset's holding is taken as given, the opportunity set spanned by the remaining tradable assets is a curve rather than a straight line in mean–standard deviation space, so the ordinary, unconditional Sharpe-ratio test no longer applies. Subtracting the hedge term, which is common to every conditionally efficient portfolio holding that constrained position, restores the standard linear problem in

²The statistic carries $n - 1 = 3$ degrees of freedom, where $n = 4$ is the number of tradable financial assets. The constrained holdings do not enter the null, they appear only in the construction of v_1 and in the variance scaling, so their number is irrelevant to the count. The single lost degree of freedom arises because the efficiency hypothesis fixes v_1 only up to the positive scale factor γ of the optimal rule, which must be estimated; Gouriéroux and Jouneau (1999) make the same observation for the corresponding scalar in their formulation.

the space of tradable assets, allowing a Sharpe-ratio comparison to be applied once again. A portfolio is classified as *efficient* when the test fails to reject the null at the chosen significance level; portfolios whose de-hedged excess return is non-positive are classified as inefficient by construction, since a non-positive expected excess return cannot support a positive Sharpe ratio.

Chapter 5

ESTIMATION AND RESULTS

This chapter presents the empirical results in three stages. Section 5.1 estimates the hedge coefficients that map each constrained asset's excess return onto the financial-asset space, first under equal weights and then under a geometrically declining scheme that emphasises the recent regime. Section 5.2 establishes a baseline by evaluating household financial portfolios against the four-asset tradable frontier, abstracting from illiquid wealth entirely. Section 5.3 then introduces the Gouriéroux–Jouneau conditional efficiency tests under four progressively richer specifications, unconditional, conditioning on housing alone, on insurance alone, and on both simultaneously, and documents the efficiency rates and their regional and demographic breakdown. Section 5.4 complements the formal tests with mean–variance frontier plots for representative demographic groups, making the efficiency losses visible in return–risk space.

5.1 Hedge Coefficient Estimates

This section reports the estimated hedge coefficients. As set out in Section 4.2, each constrained asset's annual excess return is regressed on the four financial-asset excess returns following equation (4.11), so that the estimated slope vector $\hat{\beta}_j$ is the empirical counterpart of the hedge coefficient $\Sigma^{-1}\Gamma_{bj}$ and the intercept α_j captures the abnormal, un-hedgeable component.

Table 5.1 reports the first and second moments of the constrained-asset excess returns alongside the estimated hedge coefficients under equal weighting. Panel A reveals a clear divide between the two asset groups: the four regional housing series earn high annualised excess returns (3.95%–4.79%) at modest volatility (1.40%–2.99%), whereas the insurance and pension products earn substantially less (1.18%–1.91%); Class III is by far the most volatile instrument in the sample (6.51%), reflecting its unit-linked, market-exposed design.

Panel B makes the structural distinction between the two constrained-asset groups equally apparent. Housing co-moves negatively with corporate bonds in every region, though the correlation attains conventional significance only in the North West (−0.523*); the remaining housing–bond correlations (NE −0.408, CE −0.405, SO −0.355) are consistent in sign but fall short of the 5% threshold. Housing is otherwise essentially orthogonal to the financial-asset classes. Class III, by contrast, exhibits strong and highly significant correlations with

equities (0.899**) and corporate bonds (0.737**), and a significant negative correlation with long-term government bonds (−0.622**), a pattern directly consistent with its equity-linked liability structure. Class I co-moves significantly with medium-term government bonds (0.677**) and more weakly with equities (0.469*), in line with the predominantly fixed-income backing of traditional participating policies.

Panel C translates these pairwise associations into multivariate hedge regressions. For housing, joint significance is uniformly weak: the F -test fails to reject the null of zero slopes in all four regional regressions ($p = 0.146, 0.120, 0.401$, and 0.558 for NW, NE, CE, and SO, respectively), and no individual financial-asset loading is significant in any region, with R^2 ranging from 0.17 to 0.37. The housing intercepts, however, are large and highly significant throughout (0.042**–0.053**), indicating that regional house-price excess returns are driven almost entirely by an abnormal, un-hedgeable component rather than by co-movement with tradable financial assets, a finding consistent with the low correlations in Panel B and with the near-zero R^2 . For the insurance and pension products, the picture is markedly different. Class III is nearly fully spanned by the financial-asset menu ($R^2 = 0.957$, $p < 0.001$), with significant loadings on equities (0.244**) and corporate bonds (0.474**). Class I is jointly significant ($p = 0.014$, $R^2 = 0.546$), driven principally by a positive medium-term government-bond loading (0.270**) alongside a significant intercept. PIPs are only marginally hedgeable ($p = 0.084$, $R^2 = 0.403$), with a marginally significant medium-term government-bond loading (0.207*) as the sole contributor to joint fit.

Taken together, these results suggest that housing returns are largely idiosyncratic at annual frequency over the 2005–2024 sample, offering little scope for financial-asset hedging, whereas the insurance products, particularly Class III, are substantially spanned by the tradable financial universe and therefore admit more precise hedge-term estimation.

Table 5.1: Estimating the hedge coefficients: housing and insurance

	Housing				Insurance & Pensions		
	NW	NE	CE	SO	PIPs	Class III	Class I
<i>Panel A. Descriptive statistics</i>							
Expected excess return (%)	4.43	4.27	3.95	4.79	1.18	1.70	1.91
Standard deviation (%)	1.86	1.40	2.93	2.99	1.41	6.51	1.77
<i>Panel B. Correlation with financial-asset excess returns</i>							
Stocks	0.006	0.213	−0.068	−0.123	0.182	0.899**	0.469*
Corp Bonds	−0.523*	−0.408	−0.405	−0.355	0.265	0.737**	0.328
MT Gvt Bonds	0.061	0.195	−0.179	−0.196	0.548*	0.379	0.677**
LT Gvt Bonds	0.416	0.325	0.270	0.209	−0.064	−0.622**	−0.230
<i>Panel C. Regression of excess returns on financial-asset excess returns</i>							
Constant	0.0450** (0.0047)	0.0419** (0.0035)	0.0435** (0.0080)	0.0530** (0.0085)	0.0075* (0.0034)	−0.0082 (0.0042)	0.0134** (0.0037)
r_{Stocks}	0.0316 (0.0264)	0.0367 (0.0196)	0.0380 (0.0454)	0.0205 (0.0479)	−0.0115 (0.0192)	0.2445** (0.0237)	0.0143 (0.0210)
$r_{\text{Corp Bonds}}$	−0.1950 (0.1094)	−0.1345 (0.0810)	−0.2734 (0.1879)	−0.2565 (0.1980)	0.1243 (0.0792)	0.4737** (0.0980)	0.0815 (0.0868)
$r_{\text{MT Gvt Bonds}}$	−0.0009 (0.1076)	0.0232 (0.0797)	−0.1693 (0.1849)	−0.1592 (0.1949)	0.2072* (0.0780)	0.1329 (0.0964)	0.2701** (0.0855)
$r_{\text{LT Gvt Bonds}}$	0.0325 (0.0963)	0.0282 (0.0714)	−0.0164 (0.1655)	−0.0528 (0.1744)	0.0633 (0.0698)	−0.0454 (0.0863)	0.0159 (0.0765)
p -value (F)	0.146	0.120	0.401	0.558	0.084	<0.001	0.014
R^2	0.348	0.369	0.224	0.171	0.403	0.957	0.546

Panel A reports annualised excess returns and standard deviations in percent. Panel B reports correlations between each constrained asset's excess return and the financial-asset excess returns. Panel C reports OLS regressions of each constrained asset's excess return on the four financial-asset excess returns; standard errors in parentheses, the p -value is from the F -test of the joint significance of all slopes, and R^2 . Estimation uses the 20 annual observations 2005–2024. * and ** denote significance at the 5% and 1% levels.

5.1.1 Geometrically Declining Weights

The hedge regressions presented above assign equal weight to every year in the 2005–2024 sample. This baseline imposes a strong assumption: that each annual observation is equally informative about the return-generating process relevant to current portfolio decisions. Over the sample period, however, the economic environment evolved substantially. Interest-rate regimes, property cycles, and the asset composition backing insurance liabilities all shifted materially, so the most recent observations arguably carry more information about the hedging relationships that households face today. As a robustness check, this section re-estimates the hedge regressions under a geometrically declining weighting scheme and evaluates the implications for the efficiency tests.

5.1.1.1 Decay Parameter and Calibration

Let k denote the number of years an observation precedes the end of the sample (2024). Its weight under the declining scheme is proportional to α^k , where $\alpha \in (0, 1]$ is a decay parameter. Setting $\alpha = 1$ reproduces the equal-weighted baseline; lower values concentrate estimation on more recent observations. We focus on $\alpha = 0.75$, fixed *ex ante* to balance two competing considerations: it weights recent years heavily enough to track the prevailing co-movement structure while retaining an effective sample of approximately seven observations¹.

The choice of $\alpha = 0.75$ reflects a standard bias–variance trade-off inherent in recency-weighted estimation (Pelizzon and Weber 2008). Equal weighting estimates a time-average of what may be a moving target and is potentially biased toward historical rather than current co-movement structures; down-weighting older observations reduces this bias at the cost of higher estimation variance. The primary evidence that the trade-off favours recency weighting in this application comes from Panel B of Table 5.2: the simple correlations between housing and each bond class, which require no matrix inversion and are therefore immune to the near-collinearity among bond returns, sharpen unambiguously and attain statistical significance in every region once recent years are weighted more heavily. The regression slopes corroborate this, every constrained asset becomes jointly significant under declining weights, but, given the near-collinear financial covariance matrix, the individual slope loadings are interpreted with caution. The economic conclusions rest on the correlation evidence and the joint F -tests rather than on any single coefficient.

¹The effective sample size is $1/\sum_t w_t^2$, where w_t denotes the normalised observation weights; it equals T under equal weighting and falls to approximately 7 at $\alpha = 0.75$. The qualitative pattern reported below, the signs of the housing–bond correlations and the migration of the hedge channel to the government-bond block, is stable across neighbouring calibrations ($\alpha \in \{0.9, 0.75, 0.6\}$, corresponding to effective samples of approximately 15, 7, and 4 observations respectively); $\alpha = 0.75$ is reported as a representative case.

Table 5.2: Estimating the hedge coefficients: housing and insurance ($\alpha = 0.75$)

	Housing				Insurance & Pensions		
	NW	NE	CE	SO	PIPs	Class III	Class I
<i>Panel A. Descriptive statistics</i>							
Expected excess return (%)	4.50	4.15	2.87	3.34	0.01	1.64	1.28
Standard deviation (%)	1.23	1.18	1.29	1.29	1.98	6.75	1.90
<i>Panel B. Correlation with financial-asset excess returns</i>							
Stocks	-0.468*	-0.389	-0.583**	-0.379	-0.330	0.878**	-0.306
Corp Bonds	-0.757**	-0.606**	-0.692**	-0.519*	-0.190	0.896**	-0.211
MT Gvt Bonds	0.450*	0.648**	0.443	0.546*	0.770**	-0.011	0.792**
LT Gvt Bonds	0.782**	0.636**	0.702**	0.574**	0.287	-0.920**	0.289
<i>Panel C. Regression of excess returns on financial-asset excess returns</i>							
Constant	0.0384** (0.0031)	0.0345** (0.0031)	0.0273** (0.0040)	0.0269** (0.0043)	-0.0067 (0.0054)	-0.0127* (0.0048)	0.0056 (0.0051)
r_{Stocks}	0.0232 (0.0157)	0.0188 (0.0157)	-0.0094 (0.0205)	0.0136 (0.0220)	-0.0219 (0.0277)	0.2000** (0.0248)	-0.0142 (0.0263)
$r_{\text{Corp Bonds}}$	-0.0221 (0.0659)	0.0046 (0.0659)	-0.0469 (0.0862)	0.0424 (0.0927)	0.2061 (0.1164)	0.3737** (0.1045)	0.1607 (0.1108)
$r_{\text{MT Gvt Bonds}}$	0.1232* (0.0505)	0.2025** (0.0505)	0.1250 (0.0660)	0.1859* (0.0710)	0.4595** (0.0891)	0.4004** (0.0800)	0.4530** (0.0848)
$r_{\text{LT Gvt Bonds}}$	0.1432 (0.0690)	0.1225 (0.0690)	0.0598 (0.0902)	0.1483 (0.0970)	0.1882 (0.1218)	-0.2477* (0.1095)	0.1580 (0.1160)
p -value (F)	<0.001	<0.001	0.005	0.014	<0.001	<0.001	<0.001
R^2	0.747	0.722	0.606	0.545	0.695	0.979	0.699

Panel A reports annualised excess returns and standard deviations in percent. Panel B reports correlations between each constrained asset's excess return and the financial-asset excess returns. Panel C reports WLS regressions of each constrained asset's excess return on the four financial-asset excess returns; standard errors in parentheses, the p -value is from the F -test of the joint significance of all slopes, and R^2 is unadjusted. Estimation uses the 20 annual observations 2005–2024 under geometrically declining weights ($\alpha = 0.75$). * and ** denote significance at the 5% and 1% levels.

5.1.1.2 The Asymmetric Treatment of the Efficiency Tests

The declining-weight scheme is applied exclusively to the estimation and interpretation of the hedge coefficients $\Sigma^{-1}\Gamma_{bj}$. The conditional efficiency tests of Section 4.3 are conducted throughout under equal weights ($\alpha = 1$). This asymmetry is deliberate and methodologically required: the hedge coefficient and the efficiency-test benchmark are different population objects with different finite-sample fragilities, and the weighting scheme that improves estimation of one demonstrably distorts the other.

The hedge coefficient is a conditional first-moment object, the slope of the best linear predictor of a constrained asset's excess return given the financial-asset returns. What is economically relevant for a household's current portfolio decision is the co-movement structure prevailing now. If the return-generating process is subject to drift or regime change, and over 2005–2024 the interest-rate environment, the property cycle, and insurance asset composition

all shifted materially, then equal weighting estimates a time-average of a moving target and is biased for the current relationship. Down-weighting older observations mitigates this bias, the standard rationale for exponentially weighted estimators.

The efficiency-test benchmark is a fundamentally different object: the maximal squared Sharpe ratio of the financial frontier, $\hat{S}_1 = \hat{\mu}^\top \hat{\Sigma}^{-1} \hat{\mu}$. This is a nonlinear functional of the *inverse* covariance matrix, and its finite-sample behaviour is governed not by regime relevance but by conditioning and effective sample size. Crucially, estimation error in $\hat{\Sigma}^{-1}$ inflates the apparent investment opportunity in one direction only, and this overstatement explodes as $\hat{\Sigma}$ approaches singularity (Kan and G. Zhou 2007; Britten-Jones 1999). With four financial assets, three of which (corporate, medium-, and long-term government bonds) are highly collinear, this fragility is acute. Reducing the effective sample from twenty to approximately seven by applying recency weights drives the smallest eigenvalue of $\hat{\Sigma}$ toward zero and inflates $\hat{S}_1$ from 0.50 to 2.77, implying an in-sample annual Sharpe ratio of 1.67 that no investor could have attained over this period. Importing the recency-weighted $\hat{\Sigma}$ into the efficiency tests is therefore not an option; and importing the recency-weighted hedge slopes without also importing that $\hat{\Sigma}$ would be internally inconsistent, since the Wald statistic of equation (4.15) and its χ^2_{n-1} reference distribution require all moments to be estimated from a single, common sample. The equal-weighted specification is the only one in which the de-hedge direction and the frontier onto which it is projected are estimated on a common, well-conditioned basis.

The two weighting choices thus reflect a single underlying principle: match the estimator to the estimand. Recency weighting reduces regime bias in the hedge relationship; equal weighting avoids ill-conditioning in the benchmark. The declining-weight estimates serve one limited purpose within this division of labour: to demonstrate that the sign and economic structure of the hedging relationships, housing co-moving negatively with corporate bonds and equities and positively with government bonds, with the hedge channel running principally through the government-bond block, are not artefacts of equal weighting but strengthen coherently when the recent regime is emphasised. That the strengthened relationships are economically interpretable, rather than a random configuration of signs, is itself evidence that the $\alpha = 0.75$ estimates capture a genuine feature of the data rather than overfitting noise.

5.2 Financial-Asset Efficiency: Benchmark Test

Before conditioning on illiquid wealth, we establish a baseline by evaluating household portfolios against the mean-variance frontier spanned by the four tradable financial assets alone, abstracting from housing and insurance positions entirely. Following the convention in the literature, mortgage liabilities are excluded from this benchmark on the grounds that they are most naturally treated as an integral component of the housing decision rather than of the financial-portfolio decision. This purely financial benchmark serves two purposes. First, it provides a reference point for the conditional efficiency tests of Section 5.3, where illiquid positions are introduced progressively; the divergence between the two sets of results isolates

the additional inefficiency attributable to housing and insurance holdings. Second, it allows us to calibrate the power of the GJ test statistic in a setting where the efficient frontier is estimated from financial asset returns alone.

The test is applied to the $N = 2,371$ households that hold at least one risky financial asset² and satisfy the eligibility requirements described in Section 4.3. The financial tangency portfolio is characterised by a squared Sharpe ratio $\hat{S}_1 = 0.503$, corresponding to a maximum attainable annual Sharpe ratio of approximately 0.71 over the sample period.

We report results at the 10%, 5%, and 1% test sizes, whose $\chi^2(3)$ critical values are 6.25, 7.81, and 11.34, respectively. Because a larger test size corresponds to a smaller critical value and therefore more rejections, the 10% column is the most conservative classification of efficiency.

Results are reported in Table 5.3. The test exhibits limited discriminatory power at conventional significance levels: 70.0% of eligible households cannot be rejected as efficient at the 10% size, a figure that rises to 94.3% at the 5% size and reaches 100% at the 1% size. Three structural features of the financial data account for this pattern. First, the majority of Italian households hold little or no equity, concentrating observed portfolios near the risk-free end of the frontier, precisely the region to which the tangency portfolio assigns most weight and where deviations produce small test statistics. Second, the three bond categories (corporate, medium-term, and long-term government bonds) are highly collinear over the 2005–2024 sample period, so that misallocations across them generate only modest efficiency losses in terms of squared Sharpe ratio. Third, the frontier is estimated from $T = 20$ annual observations, which limits the precision of $\hat{\Sigma}^{-1}$ and inflates the region of the parameter space that the test cannot reject.

These findings are consistent with the interpretation advanced by Pelizzon and Weber (2008): because the test has limited power against the purely financial frontier, a conservative 10% significance level is more appropriate than the conventional 5% or 1% levels, and we adopt this convention throughout the conditional tests of Section 5.3. The regional breakdown is largely uninformative at this stage, with efficiency rates ranging from 61.0% in the South and Islands to 74.9% in the North West at the 10% size. These moderate cross-regional differences will sharpen considerably once housing wealth is introduced as a conditioning variable in the subsequent tests.

²Risky financial assets are defined as equities, corporate bonds, medium-term government bonds, and long-term government bonds. Households holding only the risk-free asset are excluded, as the efficiency statistic is undefined for degenerate portfolios of this form; the resulting eligible sample of $N = 2,371$ is accordingly smaller than in the subsequent conditional tests, which treat mortgage and other debt positions explicitly.

Table 5.3: Financial-asset efficiency benchmark

	N	10%		5%		1%	
		N_{eff}	%	N_{eff}	%	N_{eff}	%
<i>Panel A. By portfolio case</i>							
RF+R	77	50	64.9	72	93.5	77	100.0
RF+R+H	1,564	1,058	67.6	1,462	93.5	1,564	100.0
RF+R+H+I	699	531	76.0	672	96.1	699	100.0
RF+R+I	31	21	67.7	30	96.8	31	100.0
Total	2,371	1,660	70.0	2,236	94.3	2,371	100.0
<i>Panel B. By macro-region</i>							
Centre	416	295	70.9	397	95.4	416	100.0
North East	768	540	70.3	725	94.4	768	100.0
North West	728	545	74.9	698	95.9	728	100.0
South & Islands	459	280	61.0	416	90.6	459	100.0
Total	2,371	1,660	70.0	2,236	94.3	2,371	100.0

N_{eff} denotes the number of eligible households for which the null hypothesis of mean-variance efficiency cannot be rejected; percentages are relative to the eligible sample N . The eligible sample comprises households holding at least one risky financial asset; portfolios consisting solely of the risk-free asset are excluded. The test benchmarks observed financial portfolios against the four-asset financial frontier ($\hat{S}_1 = 0.503$, $\alpha = 1$, $T = 20$), with no adjustment for housing or insurance positions.

5.3 Gouriéroux–Jouneau Conditional Efficiency Tests

We apply the Gouriéroux–Jouneau conditional efficiency test developed in Section 4.3 under four scenarios that differ exclusively in how the constrained asset positions are treated, that is, in which hedge terms of the de-hedged portfolio (4.13) are activated. Test A retains the household’s financial holdings at their true total-wealth scale but applies no hedge-term correction; it therefore serves as a benchmark that acknowledges actual financial exposure while abstracting entirely from hedging motives. Tests B and C each introduce the hedge term for one constrained position, conditioning respectively on housing and on insurance and pension holdings. Test D conditions simultaneously on both and constitutes our preferred specification. Because the hedge term for any asset the household does not hold is identically zero, and each term in the decomposition (4.10) depends only on its own covariance column, Test D coincides exactly with Test B for households holding housing but not insurance, and with Test C for households holding insurance but not housing; this algebraic identity is reflected in the matching entries for the corresponding rows in the tables below. The eligible sample for each test comprises the households for which the relevant statistic is well defined; risk-free-only and no-risk-free households are excluded throughout, leaving $N = 8,058$ households in the underlying estimation sample, with the eligible subsample for each individual test reported in the corresponding table.

5.3.1 Test A: Unconditional Efficiency

Test A evaluates each household’s financial holdings at their total-wealth scale without subtracting any hedge term, and so measures efficiency under the counterfactual assumption

that illiquid positions impose no hedging requirement. Of the 3,579 eligible households, 38.3% are classified as efficient at the 10% size, rising to 53.9% and 62.3% at the 5% and 1% sizes, respectively. The regional ordering already foreshadows the conditional results reported below: the North East (46.3%) and North West (45.0%) lead at the 10% size, the Centre is intermediate at 38.5%, and the South lags markedly at 22.9%. Test A is best read not as a result of independent interest, but as the benchmark against which Tests B, C, and D are compared; the gap between this scenario and its hedge-corrected counterparts isolates, for each test in turn, the incremental effect of requiring the financial portfolio to satisfy the hedging demands of the corresponding illiquid position.

Table 5.4: Unconditional efficiency (no de-hedge), by portfolio case and by macro region.

	<i>N</i>	10%		5%		1%	
		<i>N_{eff}</i>	%	<i>N_{eff}</i>	%	<i>N_{eff}</i>	%
<i>Panel A. By portfolio case</i>							
RF+R	117	49	41.9	71	60.7	76	65.0
RF+R+H	2,456	923	37.6	1,310	53.3	1,489	60.6
RF+R+H+I	961	378	39.3	520	54.1	635	66.1
RF+R+I	45	19	42.2	28	62.2	30	66.7
Total	3,579	1,369	38.3	1,929	53.9	2,230	62.3
<i>Panel B. By macro-region</i>							
Centre	647	249	38.5	346	53.5	383	59.2
North East	990	458	46.3	643	64.9	729	73.6
North West	984	443	45.0	590	60.0	692	70.3
South & Islands	958	219	22.9	350	36.5	426	44.5
Total	3,579	1,369	38.3	1,929	53.9	2,230	62.3

Number (N_{eff}) and percentage of eligible households classified as efficient at the 10%, 5%, and 1% test sizes; N is the number of eligible households. Total-wealth basis, no hedge-term correction.

5.3.2 Test B: Conditional on Housing

Requiring the financial portfolio to hedge the household's housing position optimally roughly halves measured efficiency relative to the Test A benchmark: the share of conditionally efficient households falls to 19.7% at the 10% size, recovering to 34.7% and 50.7% at the 5% and 1% sizes across the 3,417 eligible homeowners. This compression carries the central empirical message of the conditional framework: a substantial majority of Italian homeowners fail to exploit the hedging opportunities created by their large and illiquid housing positions. The comparison between the RF+R+H group evaluated in Test A and the same group evaluated here is particularly informative, because the sample is held fixed: the drop in the efficient share from 37.6% to 19.5% at the 10% size is attributable entirely to the housing-hedging requirement and constitutes the cleanest single piece of evidence in this chapter that households do not adequately hedge house-price risk through their financial portfolios.

The North East stands out markedly, with 38.6% of households classified as efficient at the 10% size and 80.1% at the 1% size; the North West is also comparatively efficient at 28.0%. By contrast, the Centre (4.7%) and the South (1.9%) are near-uniformly inefficient at the 10% size.

The sharp contrast between northern regions and the rest of the country suggests that households in the North are substantially better positioned, whether through greater financial sophistication or wider access to diversified instruments, to align their financial portfolios with the hedging demands implied by their housing wealth.

Table 5.5: Efficiency conditional on housing alone, by portfolio case and by macro region.

	N	10%		5%		1%	
		N_{eff}	%	N_{eff}	%	N_{eff}	%
<i>Panel A. By portfolio case</i>							
RF+R+H	2,456	479	19.5	833	33.9	1,190	48.5
RF+R+H+I	961	193	20.1	351	36.5	541	56.3
Total	3,417	672	19.7	1,184	34.7	1,731	50.7
<i>Panel B. By macro-region</i>							
Centre	623	29	4.7	124	19.9	214	34.3
North East	937	362	38.6	548	58.5	751	80.1
North West	942	264	28.0	425	45.1	605	64.2
South & Islands	915	17	1.9	87	9.5	161	17.6
Total	3,417	672	19.7	1,184	34.7	1,731	50.7

Number (N_{eff}) and percentage of eligible homeowners classified as conditionally efficient at the 10%, 5%, and 1% test sizes; N is the number of eligible households.

5.3.3 Test C: Conditional on Insurance

Conditioning on insurance and pension holdings yields a markedly different picture from Test B: 51.1% of the 1,006 eligible households are classified as efficient at the 10% size, rising to 62.7% and 78.8% at the 5% and 1% sizes respectively. Efficiency is moreover fairly uniform across regions, ranging from 55.1% in the North West and 54.5% in the Centre to 51.8% in the North East and 41.8% in the South.

The contrast with Test B is substantively important: whereas the housing constraint is highly binding, the insurance constraint is not. Insurance and pension positions represent a comparatively modest share of total household wealth, and, as the hedge-term regressions in Section 5.1.1 demonstrate, the market-linked component (Class III) is already closely tracked by the tradable financial assets, so that conditioning on insurance imposes relatively little additional restriction on portfolio efficiency. Indeed, for the 961 RF+R+H+I households, a group held fixed across Tests A and C, the efficient share at the 10% size rises from 39.3% under no-hedge conditioning to 50.5% under insurance conditioning, rather than falling. This pattern is consistent with the insurance hedge term, dominated by the Class III component, pointing in a direction that moves the de-hedged portfolio closer to the tangency portfolio, so that insurance operates not as a binding constraint on efficiency but, if anything, as a mild facilitator of it.

Table 5.6: Efficiency conditional on insurance alone, by portfolio case and by macro region.

	N	10%		5%		1%	
		N_{eff}	%	N_{eff}	%	N_{eff}	%
<i>Panel A. By portfolio case</i>							
RF+R+H+I	961	485	50.5	596	62.0	752	78.3
RF+R+I	45	29	64.4	35	77.8	41	91.1
Total	1,006	514	51.1	631	62.7	793	78.8
<i>Panel B. By macro-region</i>							
Centre	191	104	54.5	122	63.9	153	80.1
North East	276	143	51.8	177	64.1	221	80.1
North West	314	173	55.1	212	67.5	261	83.1
South & Islands	225	94	41.8	120	53.3	158	70.2
Total	1,006	514	51.1	631	62.7	793	78.8

Number (N_{eff}) and percentage of eligible insurance-holding households classified as conditionally efficient at the 10%, 5%, and 1% test sizes; N is the number of eligible households.

5.3.4 Test D: Conditional on Housing and Insurance

Our preferred specification conditions simultaneously on housing and insurance holdings. Of the 3,462 eligible households in Table 5.7, 22.5% hold conditionally efficient portfolios at the 10% size, 37.7% at the 5% size, and 53.9% at the 1% size. The proximity of these figures to those of Test B confirms that housing, rather than insurance, is the binding illiquid constraint: once the housing hedge term is activated, adding the insurance hedge term leaves aggregate efficiency largely unchanged. This hierarchy is reinforced by the portfolio-case breakdown at the 10% size: households holding insurance but no housing (RF+R+I) are highly efficient at 64.4%; those combining both housing and insurance (RF+R+H+I) are intermediate at 28.1%; and the largest group, homeowners without insurance (RF+R+H), is the least efficient at 19.5%. Taken together, these figures indicate that the insurance constraint contributes meaningfully to efficiency only for the subset of households already freed from the more demanding housing-hedging requirement.

The regional ranking established in Test B is fully preserved under the joint conditioning of Test D. The North East leads at 42.2% efficient, the North West follows at 32.4%, and the Centre (6.9%) and South (2.7%) remain near-uniformly inefficient at the 10% size.

Table 5.7: Efficiency conditional on housing and insurance, by portfolio case and by macro region.

	N	10%		5%		1%	
		N_{eff}	%	N_{eff}	%	N_{eff}	%
<i>Panel A. By portfolio case</i>							
RF+R+H	2,456	479	19.5	833	33.9	1,190	48.5
RF+R+H+I	961	270	28.1	436	45.4	634	66.0
RF+R+I	45	29	64.4	35	77.8	41	91.1
Total	3,462	778	22.5	1,304	37.7	1,865	53.9
<i>Panel B. By macro-region</i>							
Centre	635	44	6.9	148	23.3	246	38.7
North East	947	400	42.2	577	60.9	776	81.9
North West	955	309	32.4	475	49.7	651	68.2
South & Islands	925	25	2.7	104	11.2	192	20.8
Total	3,462	778	22.5	1,304	37.7	1,865	53.9

Number (N_{eff}) and percentage of eligible households classified as conditionally efficient at the 10%, 5%, and 1% test sizes; N is the number of eligible households.

The full case-by-region cross-tabulation in Table 5.8 makes the underlying structure explicit. The RF+R+I group is the most efficient in every region and at every test size, ranging from roughly 50–75% at the conservative 10% size to approximately 90% or above at the 1% size, confirming that insurance and pension holdings are not the source of the aggregate inefficiency documented above. The two homeowner groups, RF+R+H and RF+R+H+I, account for nearly all of this inefficiency, and drive the low aggregate figures particularly in the South & Islands and the Centre, where homeowner efficiency stands in the low single digits at the 10% size and recovers only partially at the 5% size. The North East is the consistent positive outlier across every panel: at the 1% size, 80.6% of its RF+R+H homeowners and 85.0% of its RF+R+H+I households are classified as conditionally efficient, against only 15.9% and 33.5%, respectively, in the South. This persistent North–South gradient, which holds across all portfolio cases and test sizes, is more plausibly attributed to structural differences in financial sophistication and the breadth of investment opportunities across Italian macro-regions than to compositional differences in portfolio type, since the latter are already controlled for in the panel breakdown.

Table 5.8: Efficiency conditional on housing and insurance (Test D), full cross-classification.

	Total		Centre		North East		North West		South & Isl.	
	N_{eff}	%	N_{eff}	%	N_{eff}	%	N_{eff}	%	N_{eff}	%
<i>Panel A. 10% test size</i>										
RF+R+H	479	19.5	13	2.9	274	40.8	182	28.4	10	1.4
RF+R+H+I	270	28.1	22	12.3	119	44.7	119	39.5	10	4.7
RF+R+I	29	64.4	9	75.0	7	70.0	8	61.5	5	50.0
Total	778	22.5	44	6.9	400	42.2	309	32.4	25	2.7
<i>Panel B. 5% test size</i>										
RF+R+H	833	33.9	73	16.4	411	61.3	291	45.4	58	8.3
RF+R+H+I	436	45.4	66	36.9	159	59.8	172	57.1	39	18.1
RF+R+I	35	77.8	9	75.0	7	70.0	12	92.3	7	70.0
Total	1,304	37.7	148	23.3	577	60.9	475	49.7	104	11.2
<i>Panel C. 1% test size</i>										
RF+R+H	1,190	48.5	135	30.4	541	80.6	403	62.9	111	15.9
RF+R+H+I	634	66.0	100	55.9	226	85.0	236	78.4	72	33.5
RF+R+I	41	91.1	11	91.7	9	90.0	12	92.3	9	90.0
Total	1,865	53.9	246	38.7	776	81.9	651	68.2	192	20.8

Each cell reports the number of households classified as conditionally efficient (N_{eff}) and the corresponding percentage, by portfolio case and macro region, at the 10%, 5%, and 1% test sizes. The eligible sample N for each portfolio case is given in Table 5.7 and is constant across panels and regions.

5.4 Mean–Variance Frontier Visualisation

To complement the formal efficiency tests, we visualise the mean–variance opportunity set faced by a representative household in each demographic group, contrasting two frontiers constructed in total-return space. Expected returns are defined as excess returns over the sample mean risk-free rate, $\bar{r}_f = 1.273\%$. The *unconstrained* frontier is obtained by optimising over the full menu of eight risky assets: the four tradable financial assets together with the four otherwise-constrained assets. Because no position is treated as given, this frontier represents the best mean–variance opportunity set attainable under unrestricted choice. Its construction differs in one respect across the panels presented below. In the macro-region figures, the housing return series used to build the unconstrained frontier is the OMI index specific to that region, so that the unconstrained frontier differs across regional panels. In all remaining panels (age, education, and employment status), where households are pooled across regions, a single national housing return series is used, so that the unconstrained frontier is common across those three sets of figures. The *constrained* frontier instead restricts optimisation to the four tradable financial assets, holding the illiquid block {Housing, Class I, Class III, PIPs} fixed at the group’s average portfolio weights. It thus represents the conditional opportunity set that a representative household can actually reach, given its pre-existing illiquid positions.

Both frontiers are computed allowing short positions. This choice is necessary because, for several demographic groups, the average illiquid block absorbs more than total wealth, leaving a negative residual risk-free budget $B = 1 - \sum w_{\text{fixed}} < 0$; a long-only constrained frontier would therefore not exist for those groups. Permitting short sales also implies that the displayed curves

are full parabolas, with both the efficient (upper) and the inefficient (lower) branches shown³. Under short-sale freedom, the unconstrained frontier collapses to a straight capital-market line anchored at the risk-free rate, whose slope equals the maximum attainable Sharpe ratio of the eight-asset menu. The constrained frontier, by contrast, remains a proper parabola with a vertex at strictly positive standard deviation: because the fixed illiquid block carries its own variance, the household cannot attain a zero-risk position, and the curve therefore never touches the vertical axis. Consistent with theory, the constrained frontier lies entirely to the right of the unconstrained frontier in every panel; the two would coincide only at the single tangency point where the fixed block happens to equal its unconstrained-optimal weight, a condition that is not satisfied for any of the groups examined here.

The cross marker (×) in each panel denotes the group-average total-wealth portfolio, computed over all eight risky assets. Its horizontal distance from the unconstrained frontier measures total mean–variance inefficiency, the loss attributable both to the illiquidity constraint and to suboptimal deployment of free financial assets, while its distance from the constrained frontier isolates the residual inefficiency that remains once the illiquid block is taken as given, that is, the extent to which the household’s tradable financial positions fail to hedge its fixed exposures optimally. The subsections below discuss each demographic partition in turn.

5.4.1 Efficiency By Macro-Region

The regional frontier panels broadly reproduce the ranking implied by the formal conditional efficiency tests. The average North-East portfolio sits almost on its constrained frontier near the minimum-variance vertex (approximately 2.23% standard deviation at a 5.51% total return), indicating that the representative North-Eastern household is close to conditionally efficient. The North-West is only marginally more displaced (approximately 2.71%, 6.14%). The Centre and South present a starker picture: their average portfolios (approximately 4.20%/6.04% and 3.78%/6.88%, respectively) signal considerable inefficiency. This North–South gradient, with the two northern regions close to conditional efficiency and the Centre and South appreciably further out, aligns with the test results reported in Section 5.3 and reinforces the conclusion that the failure to exploit available hedging opportunities is most pronounced outside the northern macro-regions.

³The two branches are observationally equivalent in terms of the squared Sharpe ratio: for any given variance, the upper branch attains the maximum feasible expected return and the lower branch the minimum, but the test statistic S (4.14) cannot distinguish between them. This is the same construction underlying the formal efficiency tests of Section 5.3, where a portfolio is classified as conditionally efficient only if it additionally satisfies a strictly positive de-hedged excess return.

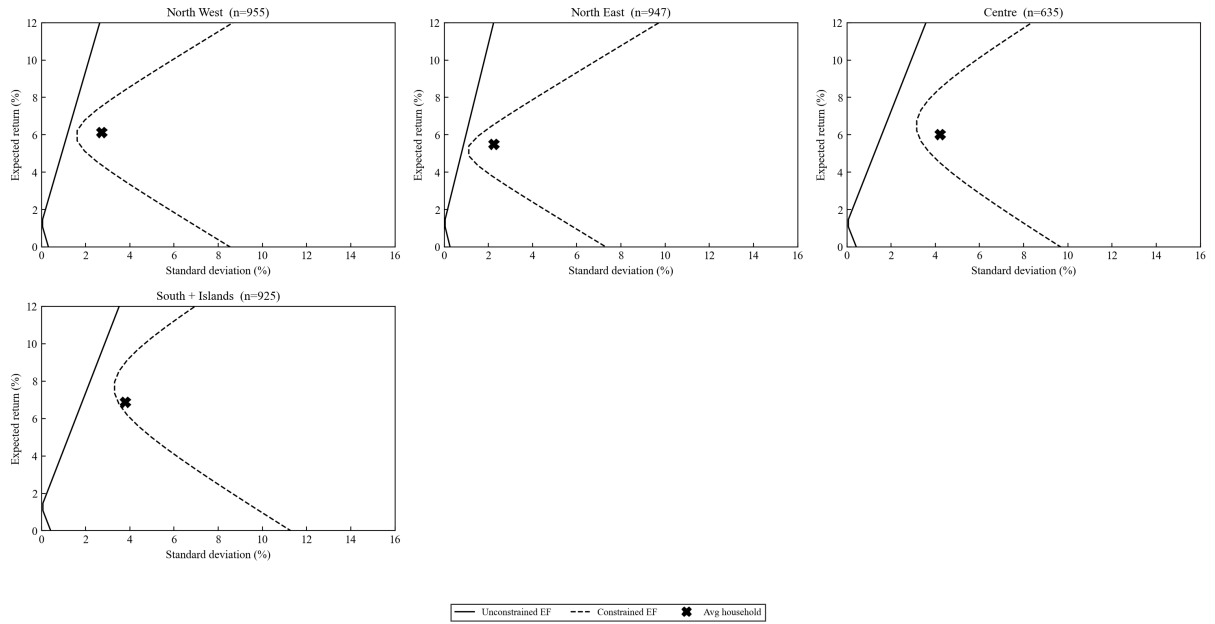


Figure 5.1: Constrained and unconstrained efficient frontiers by macro-region.

Note: The cross marker (x) indicates the group-average portfolio.

5.4.2 Efficiency By Age Cohort

The age panels reveal a pronounced and monotone gradient in conditional efficiency. The two oldest cohorts (55–64 and 65+, which together account for the majority of the conditional efficiency sample) sit very close to their respective constrained vertices (approximately 2.15–2.23% standard deviation at a 5.43–5.45% return), so that their average portfolios are nearly conditionally efficient. Displacement grows steadily as age declines: the 45–54 cohort lies at approximately 2.77%/6.07%, the 35–44 cohort at 5.07%/7.58%, and the under-35 cohort as far out as 6.65%/8.40%. Younger households therefore hold the riskiest and most conditionally inefficient average portfolios in the sample. This pattern is consistent with the larger mortgage obligations that are typical of younger homeowners: treated as negative bond positions, those mortgages widen the constrained frontier and raise background portfolio variance, while the more concentrated nature of younger households' overall asset holdings compounds the displacement from conditional efficiency.

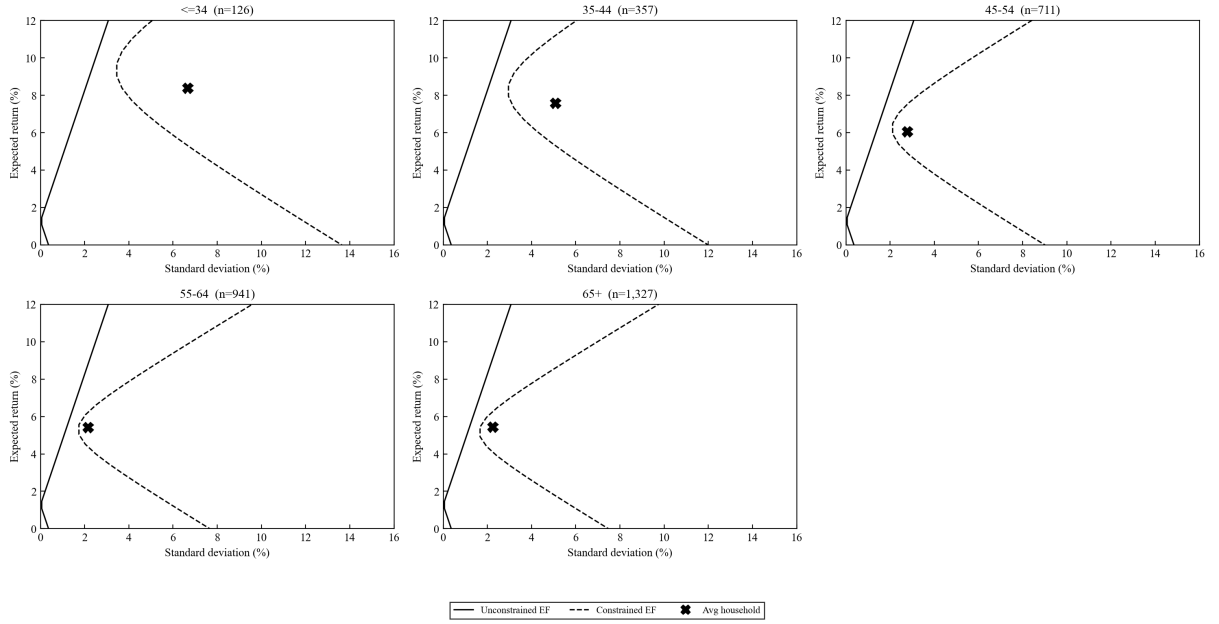


Figure 5.2: Constrained and unconstrained efficient frontiers by age group.

Note: The cross marker (x) indicates the group-average portfolio.

5.4.3 Efficiency by Educational Attainment

In contrast to the regional and age panels, education level has little bearing on the positioning of the average portfolio relative to the constrained frontier. All three groups cluster at very similar coordinates just to the right of the constrained vertex: Low ($\approx 2.78\%$, 6.21%), Secondary ($\approx 3.06\%$, 6.24%), and University+ ($\approx 3.05\%$, 5.99%). No group is materially closer to its frontier than the others, and the differences in both coordinates are negligible. This near-coincidence suggests that, at the level of the representative portfolio, educational attainment does not translate into meaningfully better deployment of tradable financial assets for the purpose of hedging the illiquid block. The finding is perhaps surprising given the well-documented association between financial literacy and portfolio diversification, but may reflect the fact that conditional hedging of housing and insurance positions requires knowledge of return covariances that is not readily accessible to households regardless of their education level.

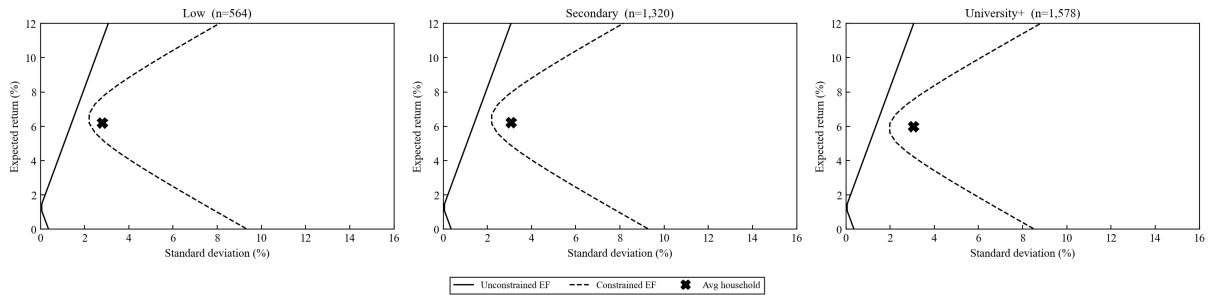


Figure 5.3: Constrained and unconstrained efficient frontiers by education group.

Note: The cross marker (x) indicates the group-average portfolio.

5.4.4 Efficiency by Employment Status

Employment status separates the groups more sharply than education, and the pattern closely mirrors the age gradient documented above. Not-working households, predominantly retirees, sit closest to the constrained vertex ($\approx 2.28\%$, 5.61%), the self-employed are slightly further out ($\approx 2.65\%$, 5.83%), and employees are the most displaced ($\approx 3.49\%$, 6.53%). The connection to the age gradient is direct: employees are on average younger than the other two groups, carry larger mortgages, and maintain a more risky financial asset composition, all of which push their representative portfolio further from the conditional frontier. Conversely, the more conservative and lower-variance portfolios of non-working households leave them in closer proximity to conditional efficiency. The employment-status results therefore largely reflect the life-cycle dimension of portfolio choice, rather than an independent effect of labour-market status per se.

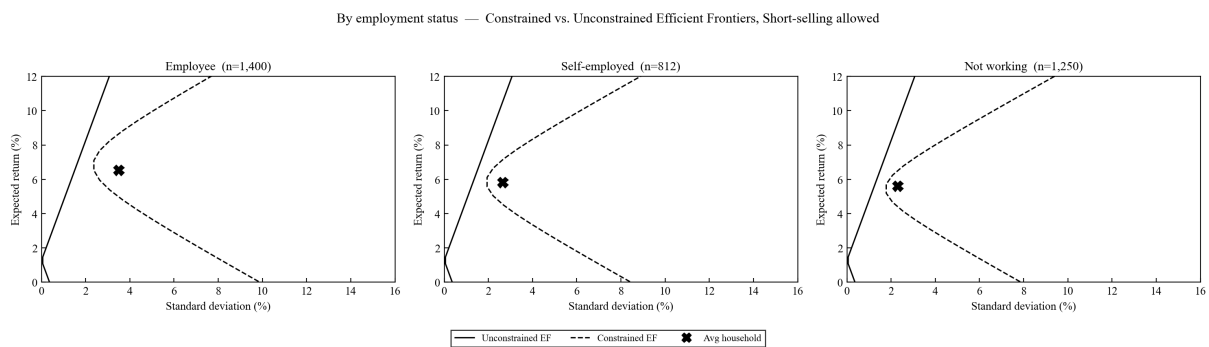


Figure 5.4: Constrained and unconstrained efficient frontiers by employment status.

Note: The cross marker (✕) indicates the group-average portfolio.

Chapter 6

Conclusions

This thesis has extended the conditional portfolio efficiency framework of Gouriéroux and Jouneau (1999), as applied by Pelizzon and Weber (2008) to the single-constrained-asset case, to the empirically relevant setting in which households hold two simultaneously constrained illiquid positions: residential real estate and life insurance and pension products. The extension is applied to a sample of 3,462 eligible households from the 2022 wave of the Bank of Italy’s Survey on Household Income and Wealth (SHIW), with asset returns spanning 2005–2024. The analysis yields clear and internally consistent answers to each of the three research questions posed in the introduction.

Housing versus insurance as the binding constraint. The most striking empirical finding is the sharp asymmetry between the two constrained asset classes. Conditioning on housing alone (Test B) reduces the share of households classified as conditionally efficient at the 10 per cent test size from 38.3 per cent under unconditional evaluation (Test A) to 19.7 per cent, roughly halving measured efficiency. By contrast, conditioning on insurance alone (Test C) leaves efficiency broadly unchanged or modestly improves it, with 51.1 per cent of eligible insurance-holding households passing the test at the 10 per cent size. The joint specification (Test D) closely tracks Test B, confirming that housing, not insurance, constitutes the binding illiquid constraint on household portfolio efficiency.

The interpretation of this housing constraint nonetheless warrants care. Pelizzon and Weber (2009) treat the housing stock as an asset and rents as a stochastic liability stream, so that a household’s net-housing position, short or long over the life cycle, enters the determination of an efficient financial portfolio. Within this framework, housing for older households may function primarily as a consumption good rather than a financial investment: a household with no intention of liquidating its principal residence may rationally disregard housing wealth in financial portfolio decisions, so that apparent conditional inefficiency reflects a rational consumption commitment rather than a portfolio optimisation failure. This consideration is particularly relevant for the oldest cohorts, whose near-efficient average portfolios in Section 5.4.2 may in part reflect a rational decoupling of housing from financial decision-making rather than superior hedging ability.

The efficiency of insurance-only households. A complementary and practically important finding is that households holding insurance and pension products but not housing ($RF+R+I$) are highly conditionally efficient: 64.4 per cent pass the joint test at the 10 per cent size, rising to 91.1 per cent at the 1 per cent size. This result is consistent with the theoretical prediction that, when a constrained asset is nearly fully spanned by the tradable menu, its hedge term is almost collinear with the tangency direction, and imposing it therefore leaves the de-hedged portfolio close to the efficient frontier. From a practical standpoint, this suggests that the failure of financial portfolios to reflect hedging demands is overwhelmingly a homeownership phenomenon rather than a consequence of insurance or pension holdings per se. It can also imply that financial advisors serving clients who hold insurance products but are not homeowners are not, on average, producing observably mis-aligned portfolios, a markedly more favourable assessment than that which applies to homeowners in the Centre and South.

The North–South efficiency gradient. Under the joint conditioning of Test D, 42.2 per cent of households in the North East and 32.4 per cent in the North West are classified as conditionally efficient at the 10 per cent size, against only 6.9 per cent in the Centre and 2.7 per cent in the South and Islands. This gradient is not merely an artefact of compositional differences in portfolio type: within each portfolio case, the North East consistently dominates, reaching 80.6 per cent conditional efficiency for homeowners without insurance ($RF+R+H$) at the 1 per cent size, versus 15.9 per cent in the South. This gradient suggests that it reflects structural differences, in financial sophistication, in the breadth of investment opportunities available to households, and in the quality of financial advice received.

The age and employment-status panels reveal a life-cycle dimension that complements the regional evidence. Older and retired households sit close to their constrained efficient frontiers, while younger households and employees, who carry larger mortgage obligations and more concentrated financial portfolios, are substantially displaced. This pattern is consistent with the argument in Cocco (2005) that binding housing constraints are most costly for young homeowners, though the present framework identifies the channel more precisely: it is the failure to hedge house-price risk through the financial portfolio, rather than the constraint per se, that drives the measured inefficiency. Education, by contrast, has little bearing on proximity to the constrained frontier at the level of the representative group portfolio, a finding that calls for caution in translating financial-literacy arguments directly into predictions about conditional hedging behaviour. The relevant knowledge, the covariance structure between regional house prices and tradable financial assets, appears not to be readily accessible to households regardless of their educational attainment, consistent with the observation that covariance estimation requires time-series data and econometric tools that lie beyond the reach of most investors.

Theoretical contributions. On the theoretical side, the thesis demonstrates that the additive structure of the hedge-term decomposition in Pelizzon and Weber (2008) generalises immediately to any finite number of simultaneously constrained positions. The cross-constraint

covariance block Σ_C affects the variance of the fixed constrained sub-portfolio but does not enter the optimal allocation among financial assets, so each hedge term depends only on its own covariance column Γ_{b_j} and the two terms enter additively and independently. This result provides a template for future extensions to additional illiquid asset classes without requiring any modification to the distributional theory underlying the test.

The asymmetric weighting scheme, declining weights applied to hedge-coefficient estimation but equal weights used for the efficiency-test benchmark, is motivated by the finite-sample fragility of the squared Sharpe ratio estimator documented by Kan and G. Zhou (2007) and Britten-Jones (1999). Appendix A demonstrates that applying recency weights ($\alpha = 0.75$) uniformly inflates the benchmark from $\hat{S}_1 = 0.50$ to $\hat{S}_1 = 2.77$, implying an in-sample annual Sharpe ratio of 1.67 that no feasible strategy could have attained over 2005–2024, and collapses efficiency rates for Test C from 51.1 per cent to 0.8 per cent, a change driven entirely by benchmark distortion rather than any genuine feature of household behaviour. The asymmetric scheme is therefore not merely a technical convenience but a methodologically necessary safeguard against the near-collinearity of the three bond factors in the financial covariance matrix.

Limitations. Several limitations of the present analysis deserve explicit acknowledgement.

The cross-sectional design of the SHIW, combined with the limited number of annual return observations ($T = 20$), places meaningful constraints on statistical power. The hedge regressions for housing fail to achieve joint significance under equal weights in any macro-region, so that the housing hedge coefficients carry substantial estimation uncertainty, and the Wald test has limited power to reject the null of conditional efficiency even for households whose de-hedged portfolios are materially misaligned with the tangency direction. Results at the 10 per cent test size are correspondingly treated as indicative rather than definitive throughout.

The SHIW collects self-reported asset values subject to under-reporting, particularly for financial instruments held by wealthier households. The housing return series, constructed from regional OMI price indices, captures aggregate regional price dynamics but cannot reflect the idiosyncratic characteristics of individual properties. Insurance and pension return series are industry-level averages that abstract from within-category heterogeneity in product design and fee structure.

The efficiency framework is static and one-period. Dynamic considerations, intertemporal hedging of human capital income, the option value embedded in owner-occupied housing, and the life-cycle glide path of pension accumulation, are not captured within the mean-variance model.

Directions for future research. Several extensions of the present framework are both theoretically straightforward and empirically promising. First, the rotating panel dimension of the SHIW could be exploited to study the dynamics of conditional efficiency over the life cycle, and in particular whether households adjust their financial portfolios in response to changes in local house prices in the direction implied by the hedge demand. Second, the aggregate insurance-pension block could be disaggregated into product-specific tests, or variation in the

share of defined-contribution versus defined-benefit pension wealth could be exploited across household types, to refine the assessment of insurance as a constrained position. Third, the framework could be extended to incorporate private business equity as a third constrained asset class: the SHIW records unlisted equity holdings currently aggregated into the risky financial block that could plausibly be treated as a partially illiquid constrained position¹. Finally, using municipality-level OMI data rather than regional aggregates to construct the housing return series would sharpen the empirical characterisation of regional heterogeneity and help distinguish between the financial-sophistication and market-access channels through which geographic location affects household portfolio efficiency.

¹Private business equity was not treated as a separate constrained asset class in the present analysis owing to the absence of a reliable return series: no index or traded benchmark is available from which to construct a return series for privately held unlisted equity.

References

- Anderson, Ronald W. and Jean-Pierre Danthine (1981). “Cross Hedging”. In: *Journal of Political Economy* 89.6, pp. 1182–1196. DOI: 10.1086/261028.
- Arrondel, Luc and Frédérique Savignac (Aug. 2015). “Risk Management, Housing and Stockholding”. In: *Applied Economics* 47.39, pp. 4208–4227. DOI: 10.1080/00036846.2015.1026585.
- Britten-Jones, Mark (1999). “The Sampling Error in Estimates of Mean-Variance Efficient Portfolio Weights”. In: *The Journal of Finance* 54.2, pp. 655–671. ISSN: 0022-1082. URL: <http://www.jstor.org/stable/2697722>.
- Brugiavini, Agar and Elsa Fornero (2001). “Pension Provision in Italy”. In: *Pension Systems and Retirement Incomes across OECD Countries*. Edward Elgar, pp. 197–235. ISBN: 9781840645637. DOI: 10.4337/9781782542353.00013.
- Cauley, Stephen Day, Andrey D. Pavlov, and Eduardo S. Schwartz (Apr. 2007). “Homeownership as a Constraint on Asset Allocation”. In: *The Journal of Real Estate Finance and Economics* 34.3, pp. 283–311. DOI: 10.1007/s11146-007-9019-9.
- Chetty, Raj, László Sándor, and Adam Szeidl (2017). “The Effect of Housing on Portfolio Choice”. In: *The Journal of Finance* 72.3, pp. 1171–1212. DOI: 10.1111/jofi.12500.
- Cocco, João F. (2005). “Portfolio Choice in the Presence of Housing”. In: *The Review of Financial Studies* 18.2, pp. 535–567. DOI: 10.1093/rfs/hhi006.
- Damgaard, Anders, Brian Fuglsbjerg, and Claus Munk (2003). “Optimal Consumption and Investment Strategies with a Perishable and an Indivisible Durable Consumption Good”. In: *Journal of Economic Dynamics and Control* 28.2, pp. 209–253. ISSN: 0165-1889. DOI: 10.1016/S0165-1889(02)00135-5.
- Eichholtz, Piet, Kees Koedijk, and Frans de Roon (Aug. 2002). *The Portfolio Implications of Home Ownership*. CEPR Discussion Paper 3501. Centre for Economic Policy Research. URL: <https://ideas.repec.org/p/cpr/ceprdp/3501.html>.
- Flavin, Marjorie and Takashi Yamashita (2002). “Owner-Occupied Housing and the Composition of the Household Portfolio”. In: *The American Economic Review* 92.1, pp. 345–362. ISSN: 0002-8282. URL: <http://www.jstor.org/stable/3083338>.
- Gibbons, Michael R., Stephen A. Ross, and Jay Shanken (1989). “A Test of the Efficiency of a Given Portfolio”. In: *Econometrica* 57.5, pp. 1121–1152. ISSN: 0012-9682. URL: <http://www.jstor.org/stable/1913625>.

- Goetzmann, William N. and Roger G. Ibbotson (Mar. 1990). "The Performance of Real Estate as an Asset Class". In: *Journal of Applied Corporate Finance* 3.1, pp. 65–76. DOI: 10.1111/j.1745-6622.1990.tb00196.x.
- Gouriéroux, C. and F. Jouneau (1999). "Econometrics of Efficient Fitted Portfolios". In: *Journal of Empirical Finance* 6.1, pp. 87–118. ISSN: 0927-5398. DOI: 10.1016/S0927-5398(98)00010-3.
- Guiso, Luigi and Tullio Jappelli (July 2000). *Household Portfolios in Italy*. CSEF Working Paper 43. Centre for Studies in Economics and Finance (CSEF), University of Naples. URL: <https://ideas.repec.org/p/sef/csefwp/43.html>.
- Guiso, Luigi, Tullio Jappelli, and Michael Haliassos (Oct. 2000). *Household Portfolios: An International Comparison*. Working Paper. Available at SSRN: <https://ssrn.com/abstract=245805>. Social Science Research Network. DOI: 10.2139/ssrn.245805.
- J.P. Morgan and Reuters (Dec. 1996). *RiskMetrics Technical Document*. 4th ed. New York: Morgan Guaranty Trust Company of New York.
- Jobson, J. D. and Bob Korkie (1982). "Potential Performance and Tests of Portfolio Efficiency". In: *Journal of Financial Economics* 10.4, pp. 433–466. ISSN: 0304-405X. DOI: 10.1016/0304-405X(82)90019-8.
- Jobson, J. D. and Bob Korkie (1989). "A Performance Interpretation of Multivariate Tests of Asset Set Intersection, Spanning, and Mean-Variance Efficiency". In: *Journal of Financial and Quantitative Analysis* 24.2, pp. 185–204. DOI: 10.2307/2330771.
- Kan, Raymond and Guofu Zhou (2007). "Optimal Portfolio Choice with Parameter Uncertainty". In: *Journal of Financial and Quantitative Analysis* 42.3, pp. 621–656. ISSN: 0022-1090. URL: <http://www.jstor.org/stable/27647314>.
- Kullmann, Cornelia and Stephan Siegel (Nov. 2005). "Real Estate and Its Role in Household Portfolio Choice". In: *SSRN Electronic Journal*. DOI: 10.2139/ssrn.424322.
- Lyng, Ran Sun and Jie Zhou (Nov. 2023). "Household Portfolio Choice Before and After a House Purchase". In: *Real Estate Economics* 51.6, pp. 1376–1398. DOI: 10.1111/1540-6229.12459.
- Markowitz, Harry (1952). "Portfolio Selection". In: *The Journal of Finance* 7.1, pp. 77–91. ISSN: 0022-1082. URL: <http://www.jstor.org/stable/2975974>.
- Markowitz, Harry M., G. Peter Todd, and William F. Sharpe (2000). *Mean-Variance Analysis in Portfolio Choice and Capital Markets*. Frank J. Fabozzi Series. New York: Wiley. ISBN: 9781883249755.
- Mayers, David (1973). "Nonmarketable Assets and the Determination of Capital Asset Prices in the Absence of a Riskless Asset". In: *The Journal of Business* 46.2, pp. 258–267. ISSN: 0021-9398. URL: <http://www.jstor.org/stable/2351366>.
- Merton, Robert C. (1980). "On Estimating the Expected Return on the Market: An Exploratory Investigation". In: *Journal of Financial Economics* 8.4, pp. 323–361. ISSN: 0304-405X. DOI: 10.1016/0304-405X(80)90007-0.

- Paz-Pardo, Gonzalo (Feb. 2021). *Homeownership and Portfolio Choice over the Generations*. Working Paper 2522. European Central Bank.
- Pelizzon, Lorian and Guglielmo Weber (2008). “Are Household Portfolios Efficient? An Analysis Conditional on Housing”. In: *Journal of Financial and Quantitative Analysis* 43.2, pp. 401–431. DOI: 10.1017/S0022109000003574.
- Pelizzon, Lorian and Guglielmo Weber (2009). “Efficient Portfolios When Housing Needs Change over the Life Cycle”. In: *Journal of Banking & Finance* 33.11, pp. 2110–2121. DOI: 10.1016/j.jbankfin.2009.05.002.
- Ross, Stephen A. and Randall C. Zisler (June 1991). “Risk and Return in Real Estate”. In: *The Journal of Real Estate Finance and Economics* 4.2, pp. 175–190. URL: <https://ideas.repec.org/a/kap/jrefec/v4y1991i2p175-90.html>.
- Samuelson, Paul A. (1970). “The Fundamental Approximation Theorem of Portfolio Analysis in Terms of Means, Variances and Higher Moments”. In: *The Review of Economic Studies* 37.4, pp. 537–542. ISSN: 0034-6527. URL: <http://www.jstor.org/stable/2296483>.

Appendix A

Robustness: Equal vs. Declining Weights

This appendix documents the mechanical distortion that recency weighting ($\alpha = 0.75$, effective sample size ≈ 7 of 20 annual observations) introduces into the efficiency-test benchmark $\hat{S}_1 = \hat{\mu}^\top \hat{\Sigma}^{-1} \hat{\mu}$, and provides the quantitative basis for the asymmetric weighting scheme described in Section 5.1.1.2. Under equal weights ($\alpha = 1$) the benchmark takes the value $\hat{S}_1 = 0.50$; under recency weights it inflates to $\hat{S}_1 = 2.77$, implying an in-sample annual Sharpe ratio of 1.67 that no feasible strategy could have attained over 2005–2024. The inflation is a finite-sample consequence of near-collinearity among the three bond factors driving the smallest eigenvalue of $\hat{\Sigma}$ toward zero as the effective degrees of freedom collapse.

Table A.1 reports efficiency rates for all four tests at the 10% significance level under both weighting schemes. The contrast is unambiguous. Test C (conditional on insurance) is the most transparent case: the 51.1% efficiency rate under equal weights collapses to 0.8% under recency weights, a fall of fifty percentage points attributable entirely to the inflated benchmark rather than any change in household behaviour. Tests A, B, and D exhibit the same direction with similarly large magnitudes. The equal-weighted specification, which is used throughout the main analysis, is therefore the only one that produces economically interpretable efficiency rates. The relatively modest drop in Test B (–8.3 pp) compared to Tests A and C (–38 to –50 pp) reflects the structure of the GJ statistic: Test B de-hedges the housing position before comparing against $\hat{S}_1$, partially insulating it from benchmark inflation. Tests A and C compare observed portfolios directly against the uninflated (Test A) or insurance-conditioned (Test C) frontier, so the overstatement of $\hat{S}_1$ passes through without attenuation. This asymmetry across tests further confirms that the efficiency differences across weighting schemes are a benchmark artefact rather than a genuine signal about household behaviour.

Table A.1: Efficiency rates under equal weights ($\alpha = 1$) vs. recency weights ($\alpha = 0.75$)

Test	Group	N	% Efficient		Δ
			$\alpha = 1$	$\alpha = 0.75$	
<i>Panel A: Test A — Unconditional (no de-hedge)</i>					
	RF+R	117	41.9%	0.0%	−41.9 pp
	RF+R+H	2,456	37.6%	0.7%	−36.9 pp
	RF+R+H+I	961	39.3%	0.7%	−38.6 pp
	RF+R+I	45	42.2%	0.0%	−42.2 pp
	Total	3,579	38.3%	0.7%	−37.6 pp
<i>Panel B: Test B — Conditional on housing</i>					
	RF+R+H	2,456	19.5%	11.1%	−8.4 pp
	RF+R+H+I	961	20.1%	12.1%	−8.0 pp
	Total	3,417	19.7%	11.4%	−8.3 pp
<i>Panel C: Test C — Conditional on insurance</i>					
	RF+R+H+I	961	50.5%	0.8%	−49.7 pp
	RF+R+I	45	64.4%	0.0%	−64.4 pp
	Total	1,006	51.1%	0.8%	−50.3 pp
<i>Panel D: Test D — Conditional on housing + insurance</i>					
	RF+R+H	2,456	19.5%	11.1%	−8.4 pp
	RF+R+H+I	961	28.1%	9.9%	−18.2 pp
	RF+R+I	45	64.4%	0.0%	−64.4 pp
	Total	3,462	22.5%	10.6%	−11.9 pp

Appendix B

Stata Code

This appendix reproduces the Stata code used to construct the analysis dataset from the Bank of Italy's Survey on Household Income and Wealth (SHIW) 2022 microdata. The code is organised into seven sequential steps; each step reads the .dta file saved by the previous one, so running them in order reproduces the full pipeline. The working directory is set once in the setup step, where the cd "PATH" line should be edited to point to the folder containing the SHIW 2022 source files.

B.1 Setup

```
1 version 17
2 clear all
3 set more off
4 capture log close
5 cd "PATH"
6 log using "shiw_thesis_log.txt", replace text
7 /* Map the 20 ISTAT regions to four macro-areas.*/
8 capture program drop _mkmacro
9 program define _mkmacro
10     gen macro_region = .
11     replace macro_region = 1 if inlist(hh_region, 1, 2, 3, 7)           // North-West
12     replace macro_region = 2 if inlist(hh_region, 4, 5, 6, 8)           // North-East
13     replace macro_region = 3 if inlist(hh_region, 9,10,11,12)           // Centre
14     replace macro_region = 4 if inlist(hh_region,13,14,15,16,17,18,19,20) // South + Islands
15 end
```

B.2 Preparing the Source Datasets

```
1 /*=====
2 PART 1 PREPARE EACH SOURCE DATASET
3 =====*/
4 /* FILE 1: carcom22 head-of-household demographics (cfred==1). */
5 use "carcom22.dta", clear
6 keep nquest cfred eta cleta5 sex studio q ireg pesofit
7 keep if cfred == 1
8 rename (eta cleta5 sex studio q ireg) ///
```

```

9      (hh_age hh_agegroup hh_gender hh_education hh_workstatus hh_region)
10 gen hh_edgroup = cond(hh_education<=3,1,cond(hh_education<=5,2,3)) if !missing(hh_education)
11 label define edgroup_lbl 1 "Low" 2 "Secondary" 3 "University+"
12 label values hh_edgroup edgroup_lbl
13 gen hh_male = (hh_gender==1) if hh_gender!=.
14 drop cfred hh_gender hh_education
15 save "temp_carcom.dta", replace
16
17 /* FILE 2: q22c2 financial assets*/
18 use "q22c2.dta", clear
19 keep nquest afdepos afa3 afb afc1 afc2 afc3 afd1 afd2 afd3 afd4 afe1 afe2 aff1 aff2 afg ///
20     pdepos pos_a3 pos_b pos_c1 pos_c2 pos_c3 pos_d1 pos_d2 pos_d3 pos_d4 pos_e1 pos_e2 ///
21     pos_f1 pos_f2 pos_g
22 rename (afdepos afa3 afb afc1 afc2 afc3 afd1 afd2 afd3 afd4 afe1 afe2 aff1 aff2 afg) ///
23     (af_deposits af_repos af_managed af_bots af_btps af_ctzs af_corpbonds ///
24     af_bankbonds af_mutfunds af_etf af_listed af_unlisted af_fordep af_forother af_cooploans)
25 gen af_total = af_deposits + af_repos + af_managed + af_bots + af_btps + af_ctzs ///
26     + af_corpbonds + af_bankbonds + af_mutfunds + af_etf + af_listed ///
27     + af_unlisted + af_fordep + af_forother + af_cooploans
28 * Financial-wealth shares (conditional on positive financial wealth).
29 foreach a in deposits repos managed bots btps ctzs corpbonds bankbonds ///
30     mutfunds etf listed unlisted fordep forother cooploans {
31     gen share_`a' = af_`a' / af_total if af_total > 0
32 }
33 * Flag households with zero reported financial-asset AMOUNTS
34 gen share_nofinassets = (af_total==0 | af_total==.)
35 label variable share_nofinassets "=1 if HH reports zero financial-asset amounts"
36 * Participation dummies: deposits from pdepos, the other 14 from pos_*.
37 gen own_deposits = (pdepos==1) if pdepos!=.
38 local pv pos_a3 pos_b pos_c1 pos_c2 pos_c3 pos_d1 pos_d2 pos_d3 pos_d4 pos_e1 pos_e2 pos_f1 pos_f2 pos_g
39 local on repos managed bots btps ctzs corpbonds bankbonds mutfunds etf listed unlisted fordep ↔
40     ↔ forother cooploans
41 forvalues k = 1/14 {
42     gen own_`word `k' of `on'' = (`word `k' of `pv'' == 1)
43 }
44 save "temp_q22c2.dta", replace
45 /* FILE 3: q22f insurance/pension overview flags.*/
46 use "q22f.dta", clear
47 gen hh_any_pension = 0
48 gen hh_vita = 0
49 forvalues i = 1/6 {
50     replace hh_any_pension = 1 if ass2new`i'==1
51     replace hh_vita = 1 if ass1new`i'==1
52 }
53 gen hh_non_vita = (ass4==1 | ass32==1) if ass4!=. | ass32!=.
54 keep nquest hh_any_pension hh_vita hh_non_vita
55 save "temp_q22f.dta", replace
56
57 /* FILE 4: allf1 supplementary pension plans (Annex F1).*/
58 use "allf1.dta", clear
59 keep nquest ass2gn ass2k
60 gen pen_occupational = (ass2gn==1) if ass2gn!=.
61 gen pen_open = (ass2gn==2) if ass2gn!=.
62 gen pen_personal_pip = (ass2gn==3) if ass2gn!=.
63 replace ass2k = 0 if ass2k==.
64 gen ass2k_occ = ass2k if ass2gn==1
65 gen ass2k_open = ass2k if ass2gn==2
66 gen ass2k_pip = ass2k if ass2gn==3
67 foreach v in ass2k_occ ass2k_open ass2k_pip {
68     replace `v' = 0 if `v'==.
69 }
70 collapse (max) pen_occupational pen_open pen_personal_pip ///

```

```

70     (sum) pension_value_occ=ass2k_occ ///
71     pension_value_open=ass2k_open pension_value_pip=ass2k_pip, by(nquest)
72 save "temp_allf1.dta", replace
73
74 /* FILE 5: allf2 life-insurance policies (Annex F2).*/
75 use "allf2.dta", clear
76 keep nquest ass1v ass1r ass1k
77 gen life_Class1 = (ass1v==1 & ass1r==1) if ass1v!=.
78 gen life_Class3 = (ass1v==1 & ass1r!=1) if ass1v!=.
79 replace ass1k = 0 if ass1k==.
80 gen ass1k_Class1 = ass1k if ass1v==1 & ass1r==1
81 gen ass1k_Class3 = ass1k if ass1v==1 & ass1r!=1
82 foreach v in ass1k_Class1 ass1k_Class3 {
83     replace 'v' = 0 if 'v'==.
84 }
85 collapse (max) life_Class1 life_Class3 ///
86     (sum) life_value_Class1=ass1k_Class1 life_value_Class3=ass1k_Class3, by(nquest)
87 save "temp_allf2.dta", replace
88
89 /* FILE 6: q22d principal residence (ownership + value).*/
90 use "q22d.dta", clear
91 keep nquest godabit valabit unipro quopro
92 gen hh_homeowner = (godabit==1 | godabit==3) if godabit!=.
93 gen double pr_share = .
94 replace pr_share = 100 if unipro==1
95 replace pr_share = quopro if unipro==2 & !missing(quopro)
96 replace pr_share = 100 if missing(pr_share) & hh_homeowner==1
97 gen double resval_principal = valabit * pr_share/100 if hh_homeowner==1
98 replace resval_principal = 0 if missing(resval_principal)
99 keep nquest hh_homeowner resval_principal
100 save "temp_q22d.dta", replace
101
102 /* FILE 6a: q22d debt-ownership flags.*/
103 use "q22d.dta", clear
104 keep nquest debita1 debita2 debita3 debita4 debite debita6 debita7
105 gen has_mortgage = (debita1==1 | debita2==1)
106 gen has_debt = (debita3==1 | debita4==1 | debite==1 | debita6==1 | debita7==1)
107 keep nquest has_mortgage has_debt
108 save "temp_debiti22.dta", replace
109
110 /* FILE 6b: alld1 other residential properties (Annex D1, tipoimm==1). */
111 use "alld1.dta", clear
112 keep nquest tipoimm quopro valabit
113 keep if tipoimm==1
114 gen double d1_share = quopro
115 replace d1_share = 100 if missing(d1_share)
116 gen double resval_d1 = valabit * d1_share/100
117 replace resval_d1 = 0 if missing(resval_d1)
118 collapse (sum) resval_other_d1 = resval_d1, by(nquest)
119 save "temp_alld1_resid.dta", replace
120
121 /* FILE 7: ricfam22 household wealth aggregate.*/
122 use "ricfam22.dta", clear
123 keep nquest w
124 rename w hh_wealth
125 save "temp_ricfam22.dta", replace
126
127 /* FILES 8a/b/c: alld2_res / alld2_aimm / alld2_fam outstanding balances. */
128 local nd nordeb1 nordeb2 nordeb3
129 local td tdebita1 tdebita2 tdebita3
130 local out debt_mortgage_res debt_mortgage_other debt_other
131 local k = 1

```

```

132 foreach f in res aimm fam {
133     use "alld2_`f'.dta", clear
134     local d : word `k' of `td'
135     keep nquest `:word `k' of `nd'' `d'
136     replace `d' = 0 if `d'==.
137     collapse (sum) `:word `k' of `out'' = `d', by(nquest)
138     save "temp_alld2_`f'.dta", replace
139     local ++k
140 }

```

B.3 Merging into the Master Dataset

```

1  /*=====
2  PART 2 MERGE INTO MASTER DATASET, BUILD WEALTH BASIS, SHARES & FLAGS
3  =====*/
4  use "temp_carcom.dta", clear
5  foreach f in q22c2 rfam22 ricfam22 debiti22 alld2_res alld2_aimm alld2_fam {
6      merge 1:1 nquest using "temp_`f'.dta", nogen keep(master match)
7  }
8  replace debt_mortgage_res = 0 if debt_mortgage_res ==.
9  replace debt_mortgage_other = 0 if debt_mortgage_other==.
10 replace debt_other = 0 if debt_other ==.
11 gen debt_mortgage = debt_mortgage_res + debt_mortgage_other
12 merge 1:1 nquest using "temp_q22d.dta", nogen keep(master match)
13 replace resval_principal = 0 if missing(resval_principal)
14 merge 1:1 nquest using "temp_alld1_resid.dta", nogen keep(master match)
15 replace resval_other_d1 = 0 if missing(resval_other_d1)
16 merge 1:1 nquest using "temp_q22f.dta", nogen keep(master match)
17 foreach v in hh_any_pension hh_vita hh_non_vita {
18     replace `v'=0 if `v'==.
19 }
20 merge 1:1 nquest using "temp_allf1.dta", nogen keep(master match)
21 foreach v in pen_occupational pen_open pen_personal_pip ///
22     pension_value_occ pension_value_open pension_value_pip {
23     replace `v'=0 if `v'==.
24 }
25 merge 1:1 nquest using "temp_allf2.dta", nogen keep(master match)
26 foreach v in life_Class1 life_Class3 life_value_Class1 life_value_Class3 {
27     replace `v'=0 if `v'==.
28 }
29 * ---- Real-estate value -----
30 gen double hh_home_value = resval_principal + resval_other_d1
31 replace hh_home_value = 0 if missing(hh_home_value)
32 * ---- Portfolio net-worth basis -----
33 gen tw_basis = af_total + hh_home_value - debt_mortgage - debt_other ///
34     + life_value_Class1 + life_value_Class3 ///
35     + pension_value_occ + pension_value_open + pension_value_pip
36 * ---- Total-wealth shares (defined for tw_basis>0; 0 when tw_basis==0) -----
37 foreach a in deposits repos managed bots btps ctzs corpbonds bankbonds ///
38     mutfunds etf listed unlisted fordep forother cooploans {
39     gen tshare_`a' = af_`a'/tw_basis if tw_basis>0
40     replace tshare_`a' = 0 if tw_basis==0
41 }
42 gen tshare_house = hh_home_value /tw_basis if tw_basis>0
43 gen tshare_mortgage = -debt_mortgage /tw_basis if tw_basis>0 // negative share
44 gen tshare_debt = -debt_other /tw_basis if tw_basis>0 // negative share
45 gen tshare_pen_occ = pension_value_occ /tw_basis if tw_basis>0

```

```

46 gen tshare_pen_open = pension_value_open/tw_basis if tw_basis>0
47 gen tshare_pen_pip = pension_value_pip /tw_basis if tw_basis>0
48 gen tshare_Class1 = life_value_Class1 /tw_basis if tw_basis>0
49 gen tshare_Class3 = life_value_Class3 /tw_basis if tw_basis>0
50 foreach v in tshare_house tshare_mortgage tshare_debt ///
51     tshare_pen_occ tshare_pen_open tshare_pen_pip tshare_Class1 tshare_Class3 {
52     replace 'v' = 0 if tw_basis==0
53 }
54 * ---- Broad asset-class participation flags -----
55 gen byte has_rf = (own_deposits==1 | own_repos==1 | own_bots==1)
56 gen byte has_housing = (hh_home_value > 0 & !missing(hh_home_value))
57 gen byte has_ins = (life_Class3==1 | life_Class1==1 | pen_occupational==1 ///
58     | pen_open==1 | pen_personal_pip==1)
59 gen byte has_risky = (own_managed==1 | own_btps==1 | own_ctzs==1 | own_corpbonds==1 | ///
60     own_bankbonds==1 | own_mutfunds==1 | own_etf==1 | own_listed==1 | ///
61     own_unlisted==1 | own_fordep==1 | own_forother==1 | own_cooploans==1 | ///
62     has_mortgage==1 | has_debt==1)
63 tab1 has_rf has_housing has_ins has_risky
64 save "master_dataset.dta", replace

```

B.4 Regional Demographic Weights

```

1  /*=====
2  PART 3 REGIONAL DEMOGRAPHIC WEIGHTS
3  =====*/
4  use "master_dataset.dta", clear
5  _mkmacro
6  label define macro_lbl 1 "North West" 2 "North East" 3 "Centre" 4 "South + Islands", replace
7  label values macro_region macro_lbl
8  save "master_dataset.dta", replace
9
10 use "carcom22.dta", clear
11 keep if cfred==1
12 keep nquest ireg pesofit
13 rename ireg hh_region
14 _mkmacro
15 collapse (sum) w_region = pesofit, by(hh_region macro_region)
16 sort macro_region hh_region
17 bysort macro_region: egen w_macro = total(w_region)
18 egen w_italy = total(w_region)
19 gen w_within_macro = w_region / w_macro
20 gen w_within_italy = w_region / w_italy
21 label define reg_wt_lbl 1 "Piemonte" 2 "Valle d'Aosta" 3 "Lombardia" ///
22     4 "Trentino-Alto Adige" 5 "Veneto" 6 "Friuli-Venezia Giulia" 7 "Liguria" ///
23     8 "Emilia-Romagna" 9 "Toscana" 10 "Umbria" 11 "Marche" 12 "Lazio" ///
24     13 "Abruzzo" 14 "Molise" 15 "Campania" 16 "Puglia" 17 "Basilicata" ///
25     18 "Calabria" 19 "Sicilia" 20 "Sardegna", replace
26 label values hh_region reg_wt_lbl
27 label values macro_region macro_lbl
28 save "regional_weights_shiw22.dta", replace

```

B.5 Participation Rates and the Portfolio Combination Table

```

1  /*=====
2  PART 4 PARTICIPATION RATES & PORTFOLIO SHARES (survey-weighted)
3  =====*/
4  use "master_dataset.dta", clear
5  svyset nquest [pweight = pesofit]
6
7  * Headline means and medians.
8  svy: mean hh_income has_mortgage has_debt hh_homeowner hh_vita hh_any_pension hh_non_vita
9  foreach var in hh_wealth af_total {
10     _pctile 'var' [pweight=pesofit], p(50)
11     di "'var' median = " r(r1)
12 }
13
14 * Participation rates (15 instruments + housing, debt, pension, insurance).
15 svy: mean own_deposits own_repos own_managed own_bots own_btps own_ctzs ///
16     own_corpbonds own_bankbonds own_mutfunds own_etf own_listed own_unlisted ///
17     own_fordep own_forother own_cooploans hh_homeowner has_mortgage has_debt ///
18     pen_occupational pen_open pen_personal_pip life_Class3 life_Class1
19
20 * Share of households with zero financial assets.
21 svy: mean share_nofinassets
22
23 * Financial-wealth shares (conditional on af_total>0).
24 svy: mean share_deposits share_repos share_managed share_bots share_btps share_ctzs ///
25     share_corpbonds share_bankbonds share_mutfunds share_etf share_listed ///
26     share_unlisted share_fordep share_forother share_cooploans
27
28 * Total-wealth shares (conditional on tw_basis>0).
29 svy: mean tshare_deposits tshare_repos tshare_managed tshare_bots tshare_btps tshare_ctzs ///
30     tshare_corpbonds tshare_bankbonds tshare_mutfunds tshare_etf tshare_listed ///
31     tshare_unlisted tshare_fordep tshare_forother tshare_cooploans ///
32     tshare_house tshare_mortgage tshare_debt tshare_pen_occ tshare_pen_open ///
33     tshare_pen_pip tshare_Class3 tshare_Class1
34
35 * Regional breakdown of the four headline total-wealth shares.
36 svy: mean tshare_house tshare_Class1 tshare_Class3 tshare_pen_pip
37 matrix b_italy = e(b)*100
38 foreach r in 1 2 3 4 {
39     gen byte _sp_r'r' = (macro_region=='r')
40     svy, subpop(_sp_r'r'): mean tshare_house tshare_Class1 tshare_Class3 tshare_pen_pip
41     matrix b_r'r' = e(b)*100
42     drop _sp_r'r'
43 }
44 matrix D4 = b_italy \ b_r1 \ b_r2 \ b_r3 \ b_r4
45 matrix rownames D4 = "Italy" "NW" "NE" "CE" "SO"
46 matrix list D4, format(%8.4f)
47 /*=====
48 PART 5 PORTFOLIO COMBINATION TABLE (UNWEIGHTED, CONSISTENT % + N)
49 =====*/
50 gen comb_rf      = (has_rf==1 & has_housing==0 & has_ins==0 & has_risky==0)
51 gen comb_rf_h    = (has_rf==1 & has_housing==1 & has_ins==0 & has_risky==0)
52 gen comb_rf_i    = (has_rf==1 & has_housing==0 & has_ins==1 & has_risky==0)
53 gen comb_rf_i_h  = (has_rf==1 & has_housing==1 & has_ins==1 & has_risky==0)
54 gen comb_rf_r    = (has_rf==1 & has_housing==0 & has_ins==0 & has_risky==1)
55 gen comb_rf_r_h  = (has_rf==1 & has_housing==1 & has_ins==0 & has_risky==1)
56 gen comb_rf_r_i  = (has_rf==1 & has_housing==0 & has_ins==1 & has_risky==1)
57 gen comb_rf_r_h_i = (has_rf==1 & has_housing==1 & has_ins==1 & has_risky==1)
58 gen comb_total_assets = has_rf
59 global combvars comb_rf comb_rf_h comb_rf_i comb_rf_i_h ///
60     comb_rf_r comb_rf_r_h comb_rf_r_i comb_rf_r_h_i comb_total_assets
61 local lbl1 "Risk-Free"
62 local lbl2 "Risk-Free + Housing"

```

```

63 local lbl3 "Risk-Free + Insurance"
64 local lbl4 "Risk-Free + Insurance + Housing"
65 local lbl5 "Risk-Free + Risky"
66 local lbl6 "Risk-Free + Risky + Housing"
67 local lbl7 "Risk-Free + Risky + Insurance"
68 local lbl8 "Risk-Free + Risky + Housing + Insurance"
69 local lbl9 "Total (risk-free holders)"
70 local codes . 1 2 3 4
71 local tags total nw ne ce so
72 matrix M_total = J(9,1,.)
73 matrix M_nw = J(9,1,.)
74 matrix M_ne = J(9,1,.)
75 matrix M_ce = J(9,1,.)
76 matrix M_so = J(9,1,.)
77 forvalues j = 1/5 {
78     local c : word 'j' of 'codes'
79     local t : word 'j' of 'tags'
80     foreach r of numlist 1/9 {
81         local v : word 'r' of $combvars
82         * totals (risk-free denominator)
83         if "'c'"=="." quietly count if has_rf==1
84         else quietly count if has_rf==1 & macro_region=="c"
85         local den = r(N)
86         * region counts
87         if "'c'"=="." quietly count if 'v'==1
88         else quietly count if 'v'==1 & macro_region=="c"
89         local num = r(N)
90         matrix M_`t'[`r',1] = 100*'num'/'den'
91     }
92     * store Ns for display
93     foreach v of global combvars {
94         if "'c'"=="." quietly count if 'v'==1
95         else quietly count if 'v'==1 & macro_region=="c"
96         local N_`v'_`t' = r(N)
97     }
98 }
99 di %40s "Combination" %18s "Total" %18s "NW" %18s "NE" %18s "CE" %18s "SO"
100 di "{hline 130}"
101 forvalues r = 1/9 {
102
103     local vv : word 'r' of $combvars
104
105     di %40s "'`lbl`r`'" ///
106     %18s string(M_total[`r',1], "%5.1f") + "% (N='N_`vv'_total')" ///
107     %18s string(M_nw[`r',1], "%5.1f") + "% (N='N_`vv'_nw')" ///
108     %18s string(M_ne[`r',1], "%5.1f") + "% (N='N_`vv'_ne')" ///
109     %18s string(M_ce[`r',1], "%5.1f") + "% (N='N_`vv'_ce')" ///
110     %18s string(M_so[`r',1], "%5.1f") + "% (N='N_`vv'_so')"
111 }
112 di "{hline 130}"

```

B.6 Portfolio Weight Construction

```

1  /*=====
2  PART 6 FACTOR-CLASS WEIGHTS & CONSOLIDATED PYTHON INPUT
3  =====*/
4  use "master_dataset.dta", clear

```

```

5 * ---- Total-wealth (TP) factor weights, for tw_basis>0 -----
6 gen w_RiskFree = tshare_deposits + tshare_repos + tshare_bots if tw_basis>0
7 gen w_Stocks   = tshare_listed + tshare_unlisted + tshare_cooploans + tshare_fordep + ↵
  ↵ tshare_forother ///
8               + 0.5*(tshare_managed + tshare_mutfunds + tshare_etf) if tw_basis>0
9 gen w_PIPs     = tshare_pen_pip if tw_basis>0
10 gen w_MT_Bonds = tshare_ctzs if tw_basis>0
11 gen w_Corp_Bonds = tshare_corpbonds + tshare_bankbonds + tshare_debt ///
12               + 0.5*(tshare_managed + tshare_mutfunds + tshare_etf) if tw_basis>0
13 gen w_LT_Bonds = tshare_btpts + tshare_mortgage if tw_basis>0
14 gen w_Class_III = tshare_Class3 + tshare_pen_occ + tshare_pen_open if tw_basis>0
15 gen w_Class_I   = tshare_Class1 if tw_basis>0
16 gen w_House     = tshare_house if tw_basis>0
17 * ---- Financial-only (FP) factor weights, for af_total>0 & tw_basis>0 -----
18 gen w_fin_RiskFree = share_deposits + share_repos + share_bots if af_total>0 & tw_basis>0
19 gen w_fin_Stocks   = share_listed + share_unlisted + share_cooploans + share_fordep + ↵
  ↵ share_forother ///
20               + 0.5*(share_managed + share_mutfunds + share_etf) if af_total>0 & tw_basis>0
21 gen w_fin_MT_Bonds = share_ctzs if af_total>0 & tw_basis>0
22 gen w_fin_Corp_Bonds = share_corpbonds + share_bankbonds ///
23               + 0.5*(share_managed + share_mutfunds + share_etf) if af_total>0 & tw_basis>0
24 gen w_fin_LT_Bonds = share_btpts if af_total>0 & tw_basis>0
25 save "master_dataset.dta", replace

```

B.7 Consolidated Export

```

1 use "master_dataset.dta", clear
2 preserve
3   keep if tw_basis>0
4   keep nquest macro_region w_RiskFree w_Stocks w_PIPs w_MT_Bonds w_Corp_Bonds ///
5     w_LT_Bonds w_Class_III w_Class_I w_House
6   export delimited using "hh_TP_weights_4FA_v2.csv", replace
7 restore
8 preserve
9   keep if af_total>0 & tw_basis>0
10  keep nquest macro_region w_fin_RiskFree w_fin_Stocks w_fin_MT_Bonds w_fin_Corp_Bonds w_fin_LT_Bonds
11  export delimited using "hh_FP_weights_4FA_v2.csv", replace
12 restore
13 * ---- Consolidated household-level categorical & demographics -----
14 gen byte portfolio_case = .
15 replace portfolio_case = 1 if has_rf & !has_housing & !has_ins & !has_risky
16 replace portfolio_case = 2 if has_rf & has_housing & !has_ins & !has_risky
17 replace portfolio_case = 3 if has_rf & !has_housing & has_ins & !has_risky
18 replace portfolio_case = 4 if has_rf & has_housing & has_ins & !has_risky
19 replace portfolio_case = 5 if has_rf & !has_housing & !has_ins & has_risky
20 replace portfolio_case = 6 if has_rf & has_housing & !has_ins & has_risky
21 replace portfolio_case = 7 if has_rf & !has_housing & has_ins & has_risky
22 replace portfolio_case = 8 if has_rf & has_housing & has_ins & has_risky
23 replace portfolio_case = 0 if has_rf==0
24 label define case_lbl 0 "No RF" 1 "RF only" 2 "RF+H" 3 "RF+I" 4 "RF+I+H" ///
25               5 "RF+R" 6 "RF+R+H" 7 "RF+R+I" 8 "RF+R+H+I", replace
26 label values portfolio_case case_lbl
27
28 gen byte emp3 = hh_workstatus if inlist(hh_workstatus,1,2,3)
29 label define emp3_lbl 1 "Employee" 2 "Self-employed" 3 "Not working", replace
30 label values emp3 emp3_lbl
31

```

```

32 keep nquest pesofit macro_region hh_region portfolio_case ///
33     has_rf has_housing has_ins has_risky ///
34     hh_age hh_agegroup hh_edgroup emp3 hh_male ///
35     af_total tw_basis hh_home_value ///
36     w_RiskFree w_Stocks w_PIPs w_MT_Bonds w_Corp_Bonds w_LT_Bonds w_Class_III w_Class_I w_House ///
37     w_fin_RiskFree w_fin_Stocks w_fin_MT_Bonds w_fin_Corp_Bonds w_fin_LT_Bonds
38 order nquest pesofit macro_region hh_region portfolio_case ///
39     has_rf has_housing has_ins has_risky ///
40     hh_age hh_agegroup hh_edgroup emp3 hh_male ///
41     af_total tw_basis hh_home_value ///
42     w_RiskFree w_Stocks w_PIPs w_MT_Bonds w_Corp_Bonds w_LT_Bonds w_Class_III w_Class_I w_House ///
43     w_fin_RiskFree w_fin_Stocks w_fin_MT_Bonds w_fin_Corp_Bonds w_fin_LT_Bonds
44 export delimited using "HH_Portfolio_4FA_v2.csv", replace
45
46 log close

```

Appendix C

Python Code

This appendix reproduces the Python code used to estimate hedge terms, run the conditional efficiency tests, and plot the efficient frontier graphs. The three scripts read the CSV outputs exported by the Stata pipeline. The file path assigned to PATH (or XLSX/CSV) at the top of each script should be updated to point to the local folder containing those files before running.

C.1 Estimation of the Hedge Terms

```
1 import pandas as pd, numpy as np, statsmodels.api as sm
2 from scipy import stats
3
4 PATH = r"RETURNS"
5
6 d = pd.read_excel(PATH, sheet_name="Returns", header=None, skiprows=48, nrows=20, usecols="B:L")
7 d.columns = ["Stocks", "CorpBonds", "MT", "LT", "PIPs", "ClassIII", "ClassI", "CE", "NE", "NW", "SO"]
8 yr = np.arange(2005, 2025)
9 lab = {"CorpBonds": "Corp Bonds", "MT": "MT Gvt Bonds", "LT": "LT Gvt ←
    ↳ Bonds", "PIPs": "PIPs", "ClassIII": "Class III", "ClassI": "Class I"}
10 fin = ["Stocks", "CorpBonds", "MT", "LT"]; con = ["NW", "NE", "CE", "SO", "PIPs", "ClassIII", "ClassI"]
11 nm = lambda c: lab.get(c, c); sep = "="*80
12 star_p = lambda p: "**** if p < .01 else " if p < .05 else " " # stars from a p-value
13
14 def tables(alpha):
15     raw = alpha ** (2024 - yr); w = raw / raw.sum(); cf = 1 - (w**2).sum()
16     wm = lambda x: (w*x).sum(); wsd = lambda x: np.sqrt((w*(x-wm(x))**2).sum()/cf)
17     def wcorr(x, y): return ←
    ↳ (w*(x-wm(x))*(y-wm(y))).sum()/np.sqrt((w*(x-wm(x))**2).sum()*(w*(y-wm(y))**2).sum())
18     def stars(r, n=20):
19         p = 2*stats.t.sf(abs(r)*np.sqrt((n-2)/(1-r**2)), n-2); return "**** if p<.01 else " if p<.05 ←
    ↳ else " "
20     tag = "equal weights" if alpha == 1 else f"geometric declining, alpha={alpha}, most-recent year ←
    ↳ highest"
21     print(f"\nWeighting: {tag}.")
22     print(f" weight on 2024 (newest) = {raw[-1]/raw.max():.4f}, weight on 2005 (oldest) = ←
    ↳ {raw[0]/raw.max():.6f}")
23     print(f" effective sample size = {1/(w**2).sum():.1f} of 20")
24     print(sep+"\nTABLE 1 -- Expected Excess Returns and Correlation Matrix\n          (* sig. 5%, ** ←
    ↳ sig. 1%)\n"+sep)
25     print("Panel A. Descriptive Statistics")
26     print("".ljust(24)+"").join(f"{nm(c):>9}" for c in con))
27     print("Expected excess return %"+"").join(f"{wm(d[c])*100:9.4f}" for c in con))
```

```

28     print("Standard deviation %      "+"".join(f"{wsd(d[c])*100:9.4f}" for c in con))
29     print("\n Panel B. Correlation Coefficients (with financial assets)")
30     print("".ljust(13)+"".join(f"{nm(c):>9}" for c in con))
31     for f in fin:
32         print(f"{nm(f):<13}"+"".join(f"{wcorr(d[f],d[c]):>7.3f}{stars(wcorr(d[f],d[c])):<2}" for c in con))
33     ↪ con))
34     X = sm.add_constant(d[fin]); R = {c: sm.WLS(d[c], X, weights=w).fit() for c in con}
35
36     # weighted factor VIFs: 1/(1-R^2) from WLS of each financial regressor on the others (same weights)
37     def vif(f):
38         others = [g for g in fin if g != f]
39         r2 = sm.WLS(d[f], sm.add_constant(d[others]), weights=w).fit().rsquared
40         return 1.0 / (1.0 - r2)
41
42     W = 20
43     def cell(c, key):
44         b, se, s = R[c].params[key], R[c].bse[key], star_p(R[c].pvalues[key])
45         return f"{b:8.4f}{s:<2} ({se:.4f})".rjust(W)
46
47     print(sep+"\nTABLE 2 -- Regression of Excess Return on Constrained Assets on Financial Asset ↪
48     ↪ Excess Returns"
49         "\n          (WLS; coeff stars from individual t-tests, * sig. 5%, ** sig. 1%;"
50         "\n          std errors in parentheses; p-value = F-test of joint slopes)\n"+sep)
51     print("".ljust(14)+"".join(f"{nm(c):>{W}}" for c in con))
52     for disp, key in [("Constant", "const")+[(f"r_{nm(f)}", f) for f in fin]:
53         print(f"{disp:<14}"+"".join(cell(c, key) for c in con))
54     print("p-value (F)".ljust(14)+"".join(f"{R[c].f_pvalue:>{W}.6f}" for c in con))
55     print("R2".ljust(14)+"".join(f"{R[c].rsquared:>{W}.6f}" for c in con))
56     print("\nFactor VIFs (collinearity among financial regressors, weighted):")
57     print("    "+"    ".join(f"{nm(f)} = {vif(f):.2f}" for f in fin))
58
59     for a in [1,0.75]:
60         print("\n"+"#"*80); tables(a)

```

C.2 Gouriéroux–Jouneau Financial Assets Test

```

1  import numpy as np
2  import pandas as pd
3  from scipy.stats import chi2
4
5  # ===== Source =====
6  XLSX = r"RETURNS"
7  CSV = r"CSVweights"
8
9  ALPHA = 1
10 # 1. LOAD RETURNS
11 R = pd.read_excel(XLSX, sheet_name="Returns", skiprows=47, nrows=20, usecols="A:M")
12 R.columns = ["Year", "Stocks", "Corp", "MT", "LT", "PIPs", "ClassIII", "ClassI",
13             "H_CE", "H_NE", "H_NW", "H_SO", "H_IT"]
14 R = R.set_index("Year").astype(float)
15 T = len(R)
16
17 # 2. TIME-SERIES WEIGHTS
18 _yr = R.index.values.astype(float)
19 _raw = ALPHA ** (_yr.max() - _yr)
20 WT = _raw / _raw.sum()

```

```

21 CF = 1.0 - (WT**2).sum() # bias-correction factor
22 EFF_N = 1.0 / (WT**2).sum() # effective sample size
23
24 def _wmean(x):
25     return float((WT * np.asarray(x, float)).sum())
26
27 def _wcov(x, y):
28     x, y = np.asarray(x, float), np.asarray(y, float)
29     return float((WT * (x - _wmean(x)) * (y - _wmean(y))).sum() / CF)
30
31 def _wmean_vec(cols): return np.array([_wmean(R[c]) for c in cols])
32
33 def _wcov_mat(cols):
34     k = len(cols); M = np.empty((k, k))
35     for i, a in enumerate(cols):
36         for j, b in enumerate(cols):
37             M[i, j] = _wcov(R[a], R[b])
38     return M
39
40 # 3. FINANCIAL ASSETS UNIVERSE (4 risky assets housing & insurance EXCLUDED)
41 FIN = ["Stocks", "Corp", "MT", "LT"]
42 mu = _wmean_vec(FIN) # vector of expected excess returns (decimal)
43 Sig = _wcov_mat(FIN) # 4x4 covariance matrix
44 Sinv = np.linalg.inv(Sig)
45 Shat = float(mu @ Sinv @ mu) # squared Sharpe of the tangency portfolio
46
47 # 4. LOAD HOUSEHOLD DATA & DROP RF-ONLY / NO-RF CASES
48 df = pd.read_csv(CSV)
49 df = df[~df.portfolio_case.isin(["RF only", "No RF"])].copy()
50
51 # 5. TEST 1: RISKY FINANCIAL ASSETS EFFICIENCY TEST
52 C2 = {a: chi2.ppf(1 - a, 3) for a in (0.10, 0.05, 0.01)}
53
54 def test_risky_fin(h):
55     """
56     Test 1 (Risky Financial Assets):
57     Uses only the four financial factor weights from the FP block.
58     Null: household portfolio is mean-variance efficient given the
59     financial assets universe {Stocks, Corp, MT, LT}.
60     """
61     z = np.array([h.w_fin_Stocks, h.w_fin_Corp_Bonds, h.w_fin_MT_Bonds, h.w_fin_LT_Bonds])
62     if np.isnan(z).any() or np.nansum(np.abs(z)) == 0:
63         return pd.Series({"xi_risky_fin": np.nan, "er_risky_fin": np.nan,
64                           "risky_fin_eff_10": np.nan,
65                           "risky_fin_eff_5": np.nan,
66                           "risky_fin_eff_1": np.nan})
67     er = float(mu @ z)
68     v = float(z @ Sig @ z)
69     S = er * er / v if v > 0 else np.nan
70     xi = T * (Shat - S) / (1 + S) # test statistic (chi2 distributed under H0)
71     return pd.Series({
72         "xi_risky_fin": xi,
73         "er_risky_fin": er,
74         "risky_fin_eff_10": float(xi < C2[0.10] and er >= 0),
75         "risky_fin_eff_5": float(xi < C2[0.05] and er >= 0),
76         "risky_fin_eff_1": float(xi < C2[0.01] and er >= 0),
77     })
78
79 df = pd.concat([df, df.apply(test_risky_fin, axis=1)], axis=1)
80
81 # 6. RESULT TABLES
82 def tab(by):

```

```

83     g = df.groupby(by)
84     n = g["xi_risky_fin"].count()
85     k = n[n > 0].index
86     out = pd.DataFrame({"N_elig": n[k]})
87     for s in (10, 5, 1):
88         col = f"risky_fin_eff_{s}"
89         eff = g[col].sum()[k]
90         out[f"N_eff_{s}"] = eff
91         out[f"{s}%"] = eff / n[k]
92     # grand total row
93     tot = df["xi_risky_fin"].count()
94     row = [tot]
95     for s in (10, 5, 1):
96         e = df[f"risky_fin_eff_{s}"].sum()
97         row += [e, e / tot]
98     out.loc["Total"] = row
99     fmt = {c: (lambda x: f"{x:.1%}") if str(c).endswith("%")
100             else (lambda x: f"{x:,.0f}") for c in out.columns}
101     return out.to_string(formatters=fmt)
102
103 print(f"ALPHA = {ALPHA} (effective N = {EFF_N:.1f} of {T})")
104 print(f"N households with risky financial assets: {len(df)}")
105 print(f"Shat (tangency, financial universe) = {Shat:.4f}")
106 print(f"chi2(3) critical values: 10%={C2[.10]:.2f} 5%={C2[.05]:.2f} 1%={C2[.01]:.2f}")
107
108 print("\nTEST 1 Risky Financial Assets (housing & insurance excluded), by portfolio case\n")
109 print(tab("portfolio_case"))
110
111 print("\nTEST 1 Risky Financial Assets (housing & insurance excluded), by macro region\n")
112 print(tab("macro_region"))
113
114 df.to_csv(f"hh_results_test1_risky_fin_a{int(round(ALPHA*100))}.csv", index=False)

```

C.3 Gouriéroux–Jouneau Conditional Efficiency Tests

```

1 import numpy as np, pandas as pd
2 from scipy.stats import chi2
3
4 # ===== Sources =====
5 XLSX = r"RETURNS"
6 CSV = r"CSVweights"
7 ALPHA = 1
8 R = pd.read_excel(XLSX, sheet_name="Returns", skiprows=47, nrows=20, usecols="A:M")
9 R.columns = ←
10  → ["Year", "Stocks", "Corp", "MT", "LT", "PIPs", "ClassIII", "ClassI", "H_CE", "H_NE", "H_NW", "H_SO", "H_IT"]
11 R = R.set_index("Year").astype(float)
12 T = len(R)
13 # -----
14 _yr = R.index.values.astype(float)
15 _raw = ALPHA ** (_yr.max() - _yr)
16 WT = _raw / _raw.sum() # normalised year weights
17 CF = 1.0 - (WT**2).sum() # reliability-weight bias correction
18 EFF_N = 1.0 / (WT**2).sum() # effective sample size
19 def _wmean(x): return float((WT * np.asarray(x, float)).sum())
20 def _wcov(x, y):
21     x = np.asarray(x, float); y = np.asarray(y, float)

```

```

22     return float((WT * (x - _wmean(x)) * (y - _wmean(y))).sum() / CF)
23 def _wmean_vec(cols): return np.array([_wmean(R[c]) for c in cols])
24 def _wcov_mat(cols):
25     k = len(cols); M = np.empty((k, k))
26     for i, a in enumerate(cols):
27         for j, b in enumerate(cols):
28             M[i, j] = _wcov(R[a], R[b])
29     return M
30 def _wvar(s):         return _wcov(s, s)                # weighted variance of any series (= Z'OmegaZ term)
31
32 FIN = ["Stocks", "Corp", "MT", "LT"]
33 mu = _wmean_vec(FIN)
34 Sig = _wcov_mat(FIN)
35 Sinv = np.linalg.inv(Sig)
36 Shat = float(mu @ Sinv @ mu)
37
38 beta = lambda col: Sinv @ np.array([_wcov(R[a], R[col]) for a in FIN])
39 REG = {"North West": "H_NW", "North East": "H_NE", "Centre": "H_CE", "South + Islands": "H_SO"}
40 bH = {r: beta(c) for r, c in REG.items()}
41 bP, b3, b1 = beta("PIPs"), beta("ClassIII"), beta("ClassI")
42
43 df = pd.read_csv(CSV)
44 df = df[~df.portfolio_case.isin(["RF only", "No RF"])] .copy()
45 C2 = {a: chi2.ppf(1 - a, 3) for a in (.10, .05, .01)}
46
47 def cond_test(o, pref, x0, constr):
48     v, port = x0.copy(), x0[0]*R.Stocks + x0[1]*R.Corp + x0[2]*R.MT + x0[3]*R.LT
49     for w, b, r in constr:
50         v, port = v + w*b, port + w*r
51     er, vv = float(mu @ v), float(v @ Sig @ v)
52     S = er*er/vv if vv > 0 else np.nan
53     xi = T * (Shat - S) / (1 + S * _wvar(port) / vv)        # Z'OmegaZ uses the chosen ALPHA
54     o[f"xi_{pref}"], o[f"er_{pref}"] = xi, er
55     # non-positive de-hedged excess return => inefficient by construction (strict)
56     for a, cv in C2.items(): o[f"{pref}_eff_{int(a*100)}"] = float(xi < cv and er > 0)
57
58 def tests(h):
59     o = {}
60     # Total-wealth financial weights x0, de-hedging different constrained sets.
61     x0 = np.array([h.w_Stocks, h.w_Corp_Bonds, h.w_MT_Bonds, h.w_LT_Bonds])
62     if not np.isnan(x0).any() and np.nansum(np.abs(x0)) > 0:
63         hh, ii = h.has_housing == 1, h.has_ins == 1
64         H = [(h.w_House, bH[h.macro_region], R[REG[h.macro_region]])]
65         I = [(h.w_PIPs, bP, R.PIPs), (h.w_Class_III, b3, R.ClassIII), (h.w_Class_I, b1, R.ClassI)]
66         cond_test(o, "A", x0, [])                                # TEST A - unconditional (no
67         if hh: cond_test(o, "B", x0, H)                          # TEST B - conditional on housing
68         if ii: cond_test(o, "C", x0, I)                          # TEST C - conditional on
69         if hh or ii: cond_test(o, "D", x0, (H if hh else []) + (I if ii else [])) # TEST D -
70         return pd.Series(o)
71
72 df = pd.concat([df, df.apply(tests, axis=1)], axis=1)
73
74 # --- Result tables ---
75 def tab(by, defcol, pref):
76     g = df.groupby(by); n = g[defcol].count(); k = n[n > 0].index
77     out = pd.DataFrame({"N_elig": n[k]})
78     for s in (10, 5, 1):

```

```

79     eff = g[f"{pref}_eff_{s}"].sum()[k]
80     out[f"N_eff_{s}"] = eff
81     out[f"{s}%"] = eff / n[k]
82     tot = df[defcol].count(); row = [tot]
83     for s in (10, 5, 1):
84         e = df[f"{pref}_eff_{s}"].sum(); row += [e, e / tot]
85     out.loc["Total"] = row
86     fmt = {c: (lambda x: f"{x:.1%}") if str(c).endswith("%") else (lambda x: f"{x:,.0f}") for c in ←
    ↪ out.columns}
87     return out.to_string(formatters=fmt)
88
89 print(f"ALPHA = {ALPHA} (effective sample size = {EFF_N:.1f} of {T}).")
90 print(f"Dropped RF-only & No-RF; N = {len(df)}. Shat(financial) = {Shat:.4f}. "
91       f"chi2(3) crit = {C2[.10]:.2f}/{C2[.05]:.2f}/{C2[.01]:.2f}")
92
93 for name, dc, pf in [("TEST A Unconditional efficiency (no de-hedge)", "xi_A", "A"),
94                     ("TEST B Conditional on housing", "xi_B", "B"),
95                     ("TEST C Conditional on insurance", "xi_C", "C"),
96                     ("TEST D Conditional on housing + insurance", "xi_D", "D")]:
97     print(f"\n{name}, by case\n", tab("portfolio_case", dc, pf))
98     print(f"\n{name}, by region\n", tab("macro_region", dc, pf))
99
100 # Efficient portfolios conditional on housing + insurance (TEST D), by macro region.
101 def panel(s):
102     e = f"D_eff_{s}"
103     d = df[df.xi_D.notna()]
104     g = d.groupby(["portfolio_case", "macro_region"])[e]
105     Neff = g.sum().unstack() # efficient hh per (case, region)
106     pct = g.mean().unstack()
107     gr = d.groupby("macro_region")[e]
108     Neff.loc["Total"], pct.loc["Total"] = gr.sum(), gr.mean() # region totals row (across cases)
109
110 # regional block: each region -> (N_eff, %)
111 regs = list(Neff.columns)
112 t = pd.concat({"N_eff": Neff, "%": pct}, axis=1).swaplevel(0, 1, axis=1)
113 t = t.reindex(columns=pd.MultiIndex.from_product([regs, ["N_eff", "%"]]))
114
115 # "Total" macro_region block: N_elig, N_eff, % (across all regions)
116 bc = d.groupby("portfolio_case")
117 tot = pd.DataFrame({"Total", "N_elig"): bc.size(),
118                    ("Total", "N_eff"): bc[e].sum(),
119                    ("Total", "%"): bc[e].mean()})
120 tot.loc["Total"] = [len(d), d[e].sum(), d[e].mean()] # grand totals (Total row, Total col)
121 t = pd.concat([tot.reindex(t.index), t], axis=1)
122
123 fmt = {c: (lambda x: "" if pd.isna(x) else f"{x:.1%}") if c[1] == ""
124         else (lambda x: "" if pd.isna(x) else f"{x:,.0f}") for c in t.columns}
125 return t.to_string(formatters=fmt)
126
127 print("\nEfficient portfolios conditional on housing and insurance (TEST D), by region")
128 print("\nPANEL A - 10% test size\n", panel(10))
129 print("\nPANEL B - 5% test size\n", panel(5))
130 print("\nPANEL C - 1% test size\n", panel(1))
131
132 df.to_csv(f"hh_results_a_{int(round(ALPHA*100))}.csv", index=False)

```

C.4 Efficient Frontier Plots

```

1 import numpy as np
2 import pandas as pd
3 import matplotlib.pyplot as plt
4 from matplotlib.lines import Line2D
5 from scipy.optimize import minimize
6 import os
7
8 # ===== Sources =====
9 XLSX_PATH = r"RETURNS"
10 CSV_PATH = r"CSVweights"
11 SHEET = "Returns"
12
13 # ----- Parameters -----
14 ALPHA = 1
15 ALLOW_SHORT = True
16 N_GRID = 120
17 OUT_DIR = None
18 X_MIN = 0.0; X_MAX = 16.0 # std-dev axis, %
19 Y_MIN = 0.0; Y_MAX = 12.0 # return axis, %
20 RF_MEAN = 0.0127254791666667
21 RETURN_COL_ORDER = ["Year", "Stocks", "Corp", "MT", "LT", "PIPs", "ClassIII", "ClassI",
22                     "House_CE", "House_NE", "House_NW", "House_SO", "House_IT"]
23 BLOCK_FIRST_ROW = 48
24 BLOCK_N_ROWS = 20
25 # =====
26 SUFFIX = "" if ALPHA == 1 else f"_a{int(round(ALPHA*100))}"
27
28 ACADEMIC_RC = {
29     "font.family": "serif",
30     "font.serif": ["Times New Roman", "Times", "DejaVu Serif", "Liberation Serif"],
31     "mathtext.fontset": "stix",
32     "font.size": 10,
33     "axes.linewidth": 0.8,
34     "axes.grid": False,
35     "xtick.direction": "in",
36     "ytick.direction": "in",
37     "legend.frameon": True,
38     "legend.edgecolor": "black",
39     "legend.fancybox": False,
40     "lines.linewidth": 1.2,
41     "figure.facecolor": "white",
42     "savefig.facecolor": "white",
43 }
44 OUTPUT_DIR = "."
45
46 FREE_RISKY = ["Stocks", "MT", "Corp", "LT"] # optimised
47 CONSTRAINED = ["House", "ClassI", "ClassIII", "PIPs"] # held fixed (constrained EF only)
48 ALL_RISKY = FREE_RISKY + CONSTRAINED # 8 assets optimised in unconstrained EF
49 WCOL = {"Stocks": "w_Stocks", "MT": "w_MT_Bonds", "Corp": "w_Corp_Bonds",
50         "LT": "w_LT_Bonds", "House": "w_House", "ClassI": "w_Class_I",
51         "ClassIII": "w_Class_III", "PIPs": "w_PIPs"}
52
53 # ----- time-series (year) weighting -----
54 def year_weights(n):
55     raw = ALPHA ** ((n - 1) - np.arange(n))
56     w = raw / raw.sum()
57     return w, 1.0 - (w**2).sum()
58
59 def wmean_np(X, w):
60     return np.average(X, axis=0, weights=w)
61

```

```

62 def wcov_np(X, w, cf):
63     mu = np.average(X, axis=0, weights=w)
64     Xc = X - mu
65     return (Xc.T * w) @ Xc / cf
66
67 def wcov_vec_np(X, c, w, cf):
68     muX = np.average(X, axis=0, weights=w); muc = np.average(c, weights=w)
69     return ((X - muX).T * w) @ (c - muc) / cf
70
71 def wvar_np(c, w, cf):
72     muc = np.average(c, weights=w)
73     return float(np.average((c - muc)**2, weights=w) / cf)
74
75 # ----- load excess returns -----
76 def load_excess_returns(path):
77     raw = pd.read_excel(path, sheet_name=SHEET, header=None)
78     ncols = len(RETURN_COL_ORDER)
79     block = raw.iloc[BLOCK_FIRST_ROW:BLOCK_FIRST_ROW + BLOCK_N_ROWS, 0:ncols].copy()
80     block.columns = RETURN_COL_ORDER
81     block = block.astype(float).reset_index(drop=True)
82     return block
83
84 # ----- group-average weights -----
85 def avg_weights(df, asset):
86     w = df["pesofit"].to_numpy()
87     return np.average(df[WCOL[asset]].to_numpy(), weights=w)
88
89 def _asset_matrix(ex, house_col, assets):
90     cols = {"House": house_col, "ClassI": "ClassI", "ClassIII": "ClassIII", "PIPs": "PIPs",
91            "Stocks": "Stocks", "MT": "MT", "Corp": "Corp", "LT": "LT"}
92     return ex[[cols[a] for a in assets]].to_numpy()
93
94 # ----- constrained frontier -----
95 def build_inputs(ex, house_col, wc):
96     Xf = ex[FREE_RISKY].to_numpy() # 20 x 4 free risky (Corp added)
97     Xc = ex[[house_col, "ClassI", "ClassIII", "PIPs"]].to_numpy() # 20 x 4 constrained
98     c = Xc @ wc # fixed background series (20,)
99     wt, cf = year_weights(len(Xf))
100     muf = wmean_np(Xf, wt)
101     Sf = wcov_np(Xf, wt, cf)
102     muc = float(np.average(c, weights=wt))
103     Varc = wvar_np(c, wt, cf)
104     sfc = wcov_vec_np(Xf, c, wt, cf)
105     return muf, Sf, muc, Varc, sfc
106
107 def port_var(x, Sf, sfc, Varc):
108     return x @ Sf @ x + 2.0 * x @ sfc + Varc
109
110 def port_ret(x, muf, muc):
111     return x @ muf + muc
112
113 def min_var_at(m, muf, Sf, muc, Varc, sfc, B):
114     """Minimise total variance s.t. total excess return == m."""
115     n = len(muf)
116     cons = [{"type": "eq", "fun": lambda x, m=m: port_ret(x, muf, muc) - m}]
117     if ALLOW_SHORT:
118         bounds = [(None, None)] * n
119     else:
120         bounds = [(0.0, None)] * n
121     cons.append({"type": "ineq", "fun": lambda x: B - x.sum()}) # RF weight >= 0
122     res = minimize(port_var, x0=np.full(n, max(B, 0.0) / n), args=(Sf, sfc, Varc),
123                   method="SLSQP", bounds=bounds, constraints=cons,
124                   options={"ftol": 1e-12, "maxiter": 400})
125     return res
126
127 def global_min_var(muf, Sf, muc, Varc, sfc, B):
128     n = len(muf)
129     if ALLOW_SHORT:

```

```

124     bounds = [(None, None)] * n; cons = []
125     else:
126         bounds = [(0.0, None)] * n
127         cons = [{"type": "ineq", "fun": lambda x: B - x.sum()}]
128     res = minimize(port_var, x0=np.full(n, max(B, 0.0) / n), args=(Sf, sfc, Varc),
129                    method="SLSQP", bounds=bounds, constraints=cons,
130                    options={"ftol": 1e-12, "maxiter": 400})
131     return res
132 def frontier(ex, house_col, wc):
133     muf, Sf, muc, Varc, sfc = build_inputs(ex, house_col, wc)
134     B = 1.0 - wc.sum()
135     gmV = global_min_var(muf, Sf, muc, Varc, sfc, B)
136     m_gmv = port_ret(gmv.x, muf, muc) # vertex (global-min-variance return)
137     if ALLOW_SHORT:
138         reach = 2.0 * (abs(muf).max() * max(B, 1.0) + abs(muc) + 0.05)
139     else:
140         reach = abs(muc) + muf.max() * max(B, 1e-6)
141     m_lo, m_hi = m_gmv - reach, m_gmv + reach
142     sig, ret = [], []
143     for m in np.linspace(m_lo, m_hi, N_GRID):
144         r = min_var_at(m, muf, Sf, muc, Varc, sfc, B)
145         if r.success and r.fun > 0:
146             sig.append(np.sqrt(r.fun)); ret.append(m)
147     return np.array(sig), np.array(ret)
148 # ----- unconstrained frontier -----
149 def frontier_unconstrained(ex, house_col):
150     R = _asset_matrix(ex, house_col, ALL_RISKY) # 20 x 8
151     wt, cf = year_weights(len(R))
152     mu = wmean_np(R, wt)
153     S = wcov_np(R, wt, cf)
154     n = len(ALL_RISKY)
155     def pvar(x): return x @ S @ x
156     def pret(x): return x @ mu
157     if ALLOW_SHORT:
158         bounds = [(None, None)] * n
159         base_cons = []
160     else:
161         bounds = [(0.0, None)] * n
162         base_cons = [{"type": "ineq", "fun": lambda x: 1.0 - x.sum()}] # RF weight >= 0
163     gmV = minimize(pvar, x0=np.full(n, 1.0 / n), method="SLSQP",
164                   bounds=bounds, constraints=base_cons,
165                   options={"ftol": 1e-12, "maxiter": 400})
166     m_gmv = pret(gmv.x) # ~0: with RF free the vertex is Rf (the origin)
167     reach = 2.0 * (abs(mu).max() + 0.05) if ALLOW_SHORT else mu.max()
168     m_lo, m_hi = m_gmv - reach, m_gmv + reach
169     sig, ret = [], []
170     for m in np.linspace(m_lo, m_hi, N_GRID):
171         cons = base_cons + [{"type": "eq", "fun": lambda x, m=m: pret(x) - m}]
172         r = minimize(pvar, x0=np.full(n, 1.0 / n), method="SLSQP",
173                     bounds=bounds, constraints=cons,
174                     options={"ftol": 1e-12, "maxiter": 400})
175         if r.success and r.fun >= 0:
176             sig.append(np.sqrt(r.fun)); ret.append(m)
177     return np.array(sig), np.array(ret)
178 def avg_point(ex, house_col, w_all_dict):
179     R = _asset_matrix(ex, house_col, ALL_RISKY) # 20 x 8
180     w = np.array([w_all_dict[a] for a in ALL_RISKY])
181     wt, cf = year_weights(len(R))
182     mu = wmean_np(R, wt); S = wcov_np(R, wt, cf)
183     return float(np.sqrt(w @ S @ w)), float(w @ mu)
184 def _legend_handles():
185     return [

```

```

186     Line2D([0], [0], color="black", ls="-", lw=1.2, label="Unconstrained EF"),
187     Line2D([0], [0], color="black", ls="--", lw=1.2, label="Constrained EF"),
188     Line2D([0], [0], color="black", marker="X", ls="none", ms=8,
189           markerfacecolor="black", markeredgecolor="black", label="Avg household"),
190 ]
191 # ----- plotting one panel -----
192 def plot_panel(df, ex, dim_name, groups, house_map, title, fname):
193     n = len(groups)
194     ncol = min(n, 3); nrow = int(np.ceil(n / ncol))
195     fig, axes = plt.subplots(nrow, ncol, figsize=(5.0 * ncol, 4.2 * nrow), squeeze=False)
196     for ax in axes.flat:
197         ax.set_visible(False)
198     for k, (label, mask) in enumerate(groups.items()):
199         ax = axes.flat[k]; ax.set_visible(True)
200         sub = df[mask]
201         if len(sub) == 0:
202             continue
203         wc = np.array([avg_weights(sub, a) for a in CONSTRAINED])
204         w_all = {a: avg_weights(sub, a) for a in ALL_RISKY}
205         hcol = house_map[label]
206
207         su, ru = frontier_unconstrained(ex, hcol)
208         if len(su):
209             ax.plot(su * 100, (ru + RF_MEAN) * 100, "-", color="black", lw=1.2)
210
211         s, r = frontier(ex, hcol, wc)
212         if len(s):
213             ax.plot(s * 100, (r + RF_MEAN) * 100, "--", color="black", lw=1.2)
214         else:
215             B = 1.0 - wc.sum()
216             ax.text(0.5, 0.92,
217                   "(no long-only constrained frontier:\nfixed Housing+Insurance > 100%)",
218                   transform=ax.transAxes, ha="center", va="top",
219                   fontsize=8, color="0.4")
220             print(f" [warn] {label}: no long-only constrained frontier "
221                   f"(budget B = {B:+.3f}); set ALLOW_SHORT=True to render it.")
222
223         ps, pr = avg_point(ex, hcol, w_all)
224         ax.plot(ps * 100, (pr + RF_MEAN) * 100, marker="X", ls="none", ms=9,
225               color="black", markerfacecolor="black", markeredgecolor="black",
226               zorder=6)
227         ax.set_title(f"{label} (n={len(sub):,})", fontsize=11)
228         ax.set_xlabel("Standard deviation (%)")
229         ax.set_ylabel("Expected return (%)")
230         ax.set_xlim(X_MIN, X_MAX)
231         ax.set_ylim(Y_MIN, Y_MAX)
232         ax.margins(0)
233         sw = "Short-selling allowed" if ALLOW_SHORT else "long-only"
234         fig.suptitle(f"Constrained vs. Unconstrained Efficient Frontiers {title}", fontsize=12, y=0.99)
235         fig.tight_layout(rect=[0, 0.075, 1, 0.96])
236         fig.legend(handles=_legend_handles(), loc="lower center", ncol=3,
237                   frameon=True, edgecolor="black", fontsize=8.5,
238                   bbox_to_anchor=(0.5, 0.0))
239         out = os.path.join(OUTPUT_DIR, fname.replace(".png", SUFFIX + ".png"))
240         fig.savefig(out, dpi=200, bbox_inches="tight")
241         plt.close(fig)
242         print("SAVED ->", os.path.abspath(out))
243 # ----- main -----
244 def main():
245     global OUTPUT_DIR
246     plt.rcParams.update(ACADEMIC_RC)
247     for p in (CSV_PATH, XLSX_PATH):

```

```

248     if not os.path.exists(p):
249         raise FileNotFoundError("Cannot find input file:\n " + p +
250                                "\nCheck the path/spelling (note the space in the .xlsx name).")
251     OUTPUT_DIR = OUT_DIR if OUT_DIR else os.path.dirname(os.path.abspath(CSV_PATH))
252     os.makedirs(OUTPUT_DIR, exist_ok=True)
253     print("Output folder:", OUTPUT_DIR)
254     _w, _cf = year_weights(BLOCK_N_ROWS)
255     print(f"ALPHA = {ALPHA} (effective sample size = {1.0/(_w**2).sum():.1f} of {BLOCK_N_ROWS}).")
256     ex = load_excess_returns(XLSX_PATH)
257     print("Annualised mean EXCESS returns (%) read from the Returns block (alpha-weighted):")
258     for col in ["Stocks", "Corp", "MT", "LT", "PIPs", "ClassIII", "ClassI", "House_IT"]:
259         print(f" {col:9s} {np.average(ex[col].to_numpy(), weights=_w)*100:+6.2f}")
260     df = pd.read_csv(CSV_PATH)
261     # ---- conditional-efficiency sample
262     keep = ~df["portfolio_case"].isin(["RF only", "No RF"])
263     # at least one risky financial asset present
264     x0 = df[["w_Stocks", "w_MT_Bonds", "w_Corp_Bonds", "w_LT_Bonds"]]
265     keep &= (x0.abs().sum(axis=1) > 0) & x0.notna().all(axis=1)
266     # holds at least one constrained asset
267     keep &= (df["has_housing"] == 1) | (df["has_ins"] == 1)
268     df = df[keep].copy()
269     print(f"Households after filter: {len(df):,}") # -> 3,462
270     # 1) REGIONS (region-specific housing return)
271     region_house = {"Centre": "House_CE", "North East": "House_NE",
272                    "North West": "House_NW", "South + Islands": "House_SO"}
273     reg_groups = {r: (df["macro_region"] == r) for r in
274                  ["North West", "North East", "Centre", "South + Islands"]}
275     plot_panel(df, ex, "region", reg_groups, region_house,
276              "By macro-region", "ef_region.png")
277     IT = "House_IT"
278     # 2) AGE
279     age_lbl = {1: "<=34", 2: "35-44", 3: "45-54", 4: "55-64", 5: "65+"}
280     age_groups = {age_lbl[g]: (df["hh_agegroup"] == g) for g in [1, 2, 3, 4, 5]}
281     plot_panel(df, ex, "age", age_groups, {v: IT for v in age_lbl.values()},
282              "By age group", "ef_age.png")
283     # 3) EDUCATION
284     edu = ["Low", "Secondary", "University+"]
285     edu_groups = {e: (df["hh_edgroup"] == e) for e in edu}
286     plot_panel(df, ex, "edu", edu_groups, {e: IT for e in edu},
287              "By education", "ef_education.png")
288     # 4) EMPLOYMENT
289     emp = ["Employee", "Self-employed", "Not working"]
290     emp_groups = {e: (df["emp3"] == e) for e in emp}
291     plot_panel(df, ex, "emp", emp_groups, {e: IT for e in emp},
292              "By employment status", "ef_employment.png")
293 if __name__ == "__main__":
294     main()

```