



Essays in Finance, Macroeconomics, and Cognition

Citation

Schröder, Simon Gerhard. 2026. Essays in Finance, Macroeconomics, and Cognition. Doctoral Dissertation, Harvard University Graduate School of Arts and Sciences.

Link

<https://dash.harvard.edu/handle/1/42740532>

Terms of use

This article was downloaded from Harvard University's DASH repository, and is made available under the terms and conditions applicable to Other Posted Material (LAA), as set forth at

<https://harvardwiki.atlassian.net/wiki/external/NGY5NDE4ZjgzNTc5NDQzMGIzZWZhMGFIOWI2M2EwYTg>

Accessibility

<https://accessibility.huit.harvard.edu/digital-accessibility-policy>

Share Your Story

The Harvard community has made this article openly available.
Please share how this access benefits you. [Submit a story](#)

HARVARD UNIVERSITY
Graduate School of Arts and Sciences



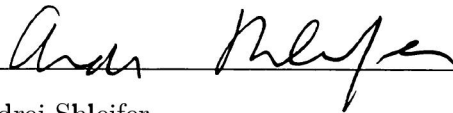
DISSERTATION ACCEPTANCE CERTIFICATE

The undersigned, appointed by the
Department of Economics
have examined a dissertation entitled

“Essays in Finance, Macroeconomics, and Cognition”

presented by **Simon Gerhard Schröder**

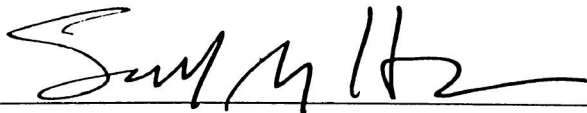
candidate for the degree of Doctor of Philosophy and hereby
certify that it is worthy of acceptance.

Signature _____ 

Typed name: Prof. Andrei Shleifer

Signature _____ 

Typed name: Prof. Adi Sunderam

Signature _____ 

Typed name: Prof. Samuel Hanson

Date: May 8th, 2026

Essays in Finance, Macroeconomics, and Cognition

A dissertation presented

by

Simon Gerhard Schröder

to

The Department of Economics

in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

in the subject of

Economics

Harvard University

Cambridge, Massachusetts

May 2026

© 2026 Simon Gerhard Schröder
All rights reserved.

Dissertation Advisor:

Professor Andrei Shleifer

Author:

Simon Gerhard Schröder

Essays in Finance, Macroeconomics, and Cognition

Abstract

Macroeconomics and financial economics have long taken expectations seriously in theory and ignored them in measurement. Both fields rest on canonical assumptions – rational expectations in macroeconomics, efficient markets in financial economics – that have become increasingly hard to defend. This dissertation pushes past these benchmarks. Three studies develop disciplined behavioral alternatives grounded in cognition, formalize their implications, and test them in the data. The first essay extends the canonical decomposition of unexpected stock returns to allow for biased cash-flow beliefs, and shows that a meaningful share of what the literature reads as risk premium news is belief-driven rather than risk-driven. The second essay develops and causally tests a memory-based theory of macroeconomic expectations, establishing that selective retrieval of past experiences shifts beliefs about the macroeconomy independently of any change in information. The third essay provides a cognitive micro-foundation for animal spirits and confidence multipliers, embedding memory-based retrieval in aggregate demand so that current outcomes cue past experiences that shape expectations about the future. Collectively, the three chapters of this thesis argue, across very different empirical environments, that the foundational assumptions of rational expectations and efficient markets are systematically violated in tractable and measurable ways.

Contents

Title Page	i
Copyright	ii
Abstract	iii
Contents	iv
List of Tables	xi
List of Figures	xiv
Acknowledgments	xv
Dedication	xix
1 Belief-Driven Risk Premium News	1
1.1 Introduction	2
1.2 Framework for Risk Premium News with Biased Beliefs	7
1.2.1 Return-News Decomposition	8
1.2.2 Behavioral Cash-Flow Belief Bias	9
1.2.3 Full Behavioral-Pricing Benchmark	11
1.2.4 Partial Arbitrage	12
1.2.5 Rate-Belief Bias and Net Valuation Beliefs	14
1.2.6 Forecast Restrictions	16
1.2.7 Matrix Condition and Pointwise News Equality	17
1.2.8 Horizon Restrictions and Convergence	19
1.2.9 Scale, Forecast Comovement, and Measurement	21
1.3 Data and News Construction	24
1.3.1 Returns, Dividends, and the Forecasting System	25
1.3.2 Analyst Forecasts and Target-Date Alignment	27
1.3.3 Three Behavioral Cash-Flow News Proxies	28
1.3.4 From Analyst Revisions to Belief-Bias News	31
1.3.5 Rational News Construction and Test Objects	32
1.3.6 Inference and Pre-Specified Interpretation	35
1.4 Main Time-Series Evidence	36
1.4.1 Sign, Scale, Forecast Comovement, and One-Step Slope Evidence	37
1.4.2 Direct Horizon Slope Tests	39
1.4.3 Forecast-Side Support and Long-Horizon Limits	40
1.5 Limits of the Sharp Benchmark	41
1.5.1 Stronger Implications of the Sharp Benchmark	41
1.5.2 Matrix Wedges	42
1.5.3 Horizon Diagnostics	44
1.5.4 Forecast Evidence and Pointwise Limits	47

1.6	Robustness and Alternative Explanations	48
1.6.1	Measurement Error and Reliability	49
1.6.2	Price-Inference Robustness	51
1.6.3	Adversarial Residuals	52
1.6.4	Robustness Evidence and Remaining Alternatives	54
1.7	Cross-Sectional and Mechanism Evidence	55
1.7.1	Portfolio-Level and Sorting Evidence	56
1.7.2	Level vs. Reallocation	59
1.7.3	Cross-Sectional Breadth and Mechanism	61
1.8	Conclusion	61
2	The Psychology of Macroeconomic Expectations	64
2.1	Introduction	65
2.2	Motivating Evidence	69
2.2.1	The DNB Household Survey	69
2.2.2	Personal Adversities and Macroeconomic Beliefs	70
2.3	Selective Memory and Belief Formation	73
2.3.1	The Setup	73
2.3.2	Beliefs, Experiences, and Priming Effects	77
2.3.3	Characterizing Experience and Priming Effects	80
2.4	The Survey: Design and Basic Facts	82
2.4.1	The Survey Experiment	83
2.4.2	Forecasts and Experiences in the Control Sample	85
2.4.3	Measured Memory Parameters	88
2.5	Context, Memory, and Macroeconomic Expectations	91
2.5.1	Instability: The Causal Effect of Primed Recall on Beliefs	91
2.5.2	Heterogeneity: The NDS Experiences Database and Similarity	95
2.5.3	Priming Changes Reasoning about Forecasts	100
2.6	Macroeconomic Implications	107
2.6.1	Unifying Features of Macroeconomic Expectations	108
2.6.2	Heterogeneity, Confidence, and Business Cycles	111
2.7	Conclusion	114
3	Memory, Confidence, and Aggregate Demand	116
3.1	Introduction	117
3.2	Environment	123
3.2.1	Reduced-Form Demand Block	124
3.2.2	Timing	124
3.2.3	Experiences and Database	125
3.2.4	Cue	125
3.2.5	Demand Equation	126

3.3	Similarity-Based Retrieval	127
3.3.1	Recall and Expectation Operators	128
3.3.2	Similarity Kernel	129
3.3.3	Mnemonic Expectation in the Two-Type Case	130
3.4	Static Feedback and Multiplicity	132
3.4.1	Retrieval Slope	132
3.4.2	Confidence Multiplier	133
3.4.3	Static Multiplicity	135
3.5	Neutral Updating and Hysteresis	137
3.5.1	Neutral Law of Motion	137
3.5.2	Hysteresis Under Neutral Updating	139
3.6	Adverse Resonance and Recession Skew	141
3.6.1	Symmetric Benchmark	142
3.6.2	Adverse Resonance	143
3.6.3	Boundary Slope vs. Interior Level	145
3.6.4	Basin Asymmetry and Cue-Perturbation Thresholds	146
3.7	Empirical Discipline	147
3.7.1	Empirical Predictions	148
3.7.2	Public Cue Shocks and Adverse-Memory States	149
3.7.3	Public Cues Move Sentiment and Postponable Demand	150
3.7.4	Adverse Memory Amplifies the Effect of Adverse Cues	152
3.7.5	Households vs. Professional Forecasters	158
3.7.6	Joint Pattern and Alternative Mechanisms	159
3.8	Extensions and Interpretation	161
3.8.1	Richer Cue Perturbations and Narrative Shocks	161
3.8.2	Asymmetric Updating	163
3.8.3	Heterogeneity	164
3.8.4	Policy as Cue Management	165
3.9	Conclusion	166
	References	170
A	Appendix to Chapter 1	178
A.1	Theory Derivations and Proofs	178
A.1.1	Return-News Identity and Timing	178
A.1.2	Proof of Proposition 1: Full Behavioral-Pricing Benchmark	179
A.1.3	Behavioral Valuation Wedge	181
A.1.4	Proof of Proposition 2: Partial Arbitrage	183
A.1.5	Projection Coefficients and the Interpretation of λ	184
A.1.6	Risk-Free-Rate Belief Bias and Net Valuation Beliefs	185

A.1.7	Generalized Forecast Relation	187
A.1.8	Perceived vs. Rational Expected Returns	189
A.2	Forecast, Matrix, and Horizon Restrictions	190
A.2.1	Linear Forecasting System	190
A.2.2	One-Step Forecast Restriction	191
A.2.3	Matrix Condition	192
A.2.4	Matrix Wedge	193
A.2.5	Finite-Horizon Present-Value Matrices	195
A.2.6	Horizon Forecast Restrictions	195
A.2.7	Slope vs. Matrix Restrictions	196
A.2.8	Five-Year Horizons and Infinite-Horizon News	197
A.3	Measurement Error and Forecast R^2	199
A.3.1	Additive Classical Noise	199
A.3.2	Signal Attenuation	200
A.3.3	Incomplete Proxy Models	202
A.3.4	Nonclassical Proxy Contamination	203
A.3.5	Forecast R^2 : Sample, Population, and Bias-Corrected Measures	204
A.3.6	Measurement Error Inside the State Vector	206
A.4	Data Construction and Proxy Validation	208
A.4.1	Return and Macro Data	208
A.4.2	Analyst Forecast Data and Sample	209
A.4.3	Target-Date Alignment	210
A.4.4	Behavioral Cash-Flow News Proxy Construction	211
A.4.5	Market-Level Aggregation and Belief-Level Control	212
A.4.6	Proxy Validation: Summary Statistics and Placebos	212
A.4.7	Rationality and Predictability Diagnostics	213
A.4.8	Persistence Profiles of Aligned Revisions	214
A.4.9	Common Component Across Proxies	215
A.5	VAR, News Construction, and Inference	216
A.5.1	Benchmark VAR Specification	216
A.5.2	VAR Estimation and Stability	217
A.5.3	News Construction	217
A.5.4	Risk-Free-Rate News Timing Convention	218
A.5.5	Constructing Empirical Belief-Bias News	219
A.5.6	Campbell–Ammer Identity Checks	219
A.5.7	Wald Statistic for the One-Step Slope Restriction	220
A.5.8	Stationary Block Bootstrap	220
A.5.9	Null-Imposed Bootstrap for the Slope Test	221
A.5.10	Direct Horizon Tests and Local Projections	222

A.5.11	Finite-Sample Null Diagnostics	223
A.5.12	HAC Bandwidth and Inference Sensitivity	224
A.6	Additional Empirical Tables	225
A.6.1	Baseline Results by Proxy	225
A.6.2	Campbell–Ammer Variance Decomposition	226
A.6.3	Forecast-Comovement Statistics	227
A.6.4	Full Horizon Slope Tests	228
A.6.5	Matrix-Wedge Bootstrap	229
A.6.6	Horizon Decomposition	230
A.6.7	ρ -Sensitivity	232
A.6.8	State Vector and Lag Order Robustness	232
A.6.9	HAC Bandwidth Sensitivity	234
A.6.10	Sample and Frequency Robustness	234
A.6.11	Summary	235
A.7	Reliability and Measurement-Error Corrections	237
A.7.1	Cross-Proxy Reliability	237
A.7.2	Common-Factor Diagnostics	238
A.7.3	Reliability Lower Bounds	239
A.7.4	Split-Sample Reliability	240
A.7.5	Reliability-Corrected Correlations and Projections	241
A.7.6	Model-Based Measurement Corrections	243
A.7.7	Bootstrap Uncertainty for Corrected Projections	244
A.7.8	Interpretation	245
A.8	Robustness and Adversarial Exercises	247
A.8.1	Price-Inference Robustness	247
A.8.2	Orthogonal Residual Dilution	249
A.8.3	Targeted Residual Adversary	250
A.8.4	Adversarial Frontier	252
A.8.5	Structural λ Calibration	253
A.8.6	Algebraic Identity for the Matrix Wedge	254
A.9	Cross-Sectional Evidence	256
A.9.1	FF25 Portfolio Exposure Design	256
A.9.2	FF25 News-Correlation Heatmap and Table	257
A.9.3	FF25 Slope Tests	258
A.9.4	Portfolio-Level Behavioral Cash-Flow News Coverage	259
A.9.5	Sorting Evidence	260
A.9.6	Level vs. Reallocation Decomposition	260
A.9.7	Additional Cross-Sectional Mechanism Checks	263

B	Appendix to Chapter 2	264
B.1	Proofs	264
B.1.1	Proposition 1	264
B.1.2	Prediction 1	272
B.1.3	Prediction 2	273
B.2	The Survey Instruments	275
B.2.1	The DHS: Index Construction and Results	275
B.2.2	Our Survey Instrument	288
B.3	Further Results on Experiences and Similarity	297
B.3.1	Frequency of NDS Experiences	297
B.3.2	Similarity of NDS Experiences to Inflation and Home Price Growth	299
B.4	Further Results on Experience and Priming Effects	302
B.4.1	Reliability of Similarity Measures	305
B.4.2	Quantification	312
B.5	Further Results on Reasoning	315
B.5.1	Protocol for Extraction of Valence and Context	315
B.5.2	Robustness on Backcasts and Reasoning	322
B.5.3	Similarity between Reasons and Primed Recall	324
B.5.4	Reasoning and Beliefs	327
C	Appendix to Chapter 3	330
C.1	Proofs and Analytic Derivations	330
C.1.1	Retrieval Slope and Smooth Similarity Alternatives	330
C.1.2	Two-Type Closed Form and Generalization	332
C.1.3	Confidence Multiplier	333
C.1.4	Full-Channel Multiplier	334
C.1.5	Static Multiplicity	335
C.1.6	Hysteresis Under Neutral Updating	336
C.1.7	Recession Skew from Adverse Resonance	338
C.2	Empirical Data Construction	341
C.2.1	Public Cue Measures	341
C.2.2	State Adverse-Memory Measures	341
C.2.3	National Sentiment and Demand Outcomes	342
C.2.4	State Permits and State-Level Search Attention	342
C.2.5	Residualization, Standardization, and Sample Definitions	342
C.2.6	Data Summary Table	343
C.3	Local-Projection Evidence	344
C.3.1	National Specification	344
C.3.2	Regional Panel Specification	344
C.3.3	Full Robustness Grid	345

C.3.4	Permutation and Random-Cue Placebos	347
C.3.5	Leave-One-State-Out Robustness	348
C.3.6	Adverse-Memory Tail Figure	348
C.4	State-Level Adverse-Search Attention	349
C.4.1	State Search Index Construction	349
C.4.2	Cue-by-Memory Attention Specification	349
C.5	Households vs. Professional Forecasters	350
C.5.1	SPF Data and Quarterly Aggregation	350
C.5.2	Household–Forecaster Response Comparisons	350
C.6	Additional Persistence and Regime Diagnostics	351
C.6.1	Sign-Asymmetry Diagnostics in the Regional Panel	351
C.6.2	Michigan Sentiment Half-Life Ratios	352

List of Tables

1.1	Empirical Objects and Their Interpretation	24
1.2	Behavioral Cash-Flow News Proxies	29
1.3	Baseline Time-Series Evidence	37
1.4	Direct Horizon Slope Tests	39
1.5	Matrix-Wedge Diagnostics	43
1.6	Price-Inference Robustness	52
1.7	Decomposition of the Aggregate Belief Revision: Level vs. Reallocation	60
2.1	Adversity and Macroeconomic Beliefs in the DHS	71
2.2	Summary Statistics for Macroeconomic Beliefs	85
2.3	Experience Effects and Macroeconomic Beliefs	87
2.4	Effects of Cued Financial and Health Adversity on Macroeconomic Beliefs	94
2.5	Effect of Experience Similarity on Inflation and Home Price Growth Beliefs	97
2.6	Backcasts, Context, and Valence	105
2.7	Beliefs and Predicted Backcasts, Context, and Valence	107
3.1	Vehicle Sales: Response to Adverse EPU Cues	152
3.2	Sentiment and Demand Responses to Adverse Public Cues	153
3.3	State-Level Adverse-Search Attention: Cue $\times$ Memory Interactions	155
3.4	Adverse Search Cues, Adverse Memory, and State Permits	157
3.5	Joint Pattern Across the Empirical Analysis	160
A.1	Summary Statistics for Aligned Revisions and Placebo	213
A.2	Null-Imposed Bootstrap Diagnostics for the Slope-Restriction Wald Statistic	223
A.3	Extended Baseline Time-Series Evidence	226
A.4	Bootstrap Confidence Intervals for Baseline Statistics	226
A.5	Campbell–Ammer Variance Decomposition	227
A.6	Forecast Comovement: Sample, Population, and Bias-Corrected R^2	228
A.7	Direct Horizon Slope Tests, All Reported Horizons	228
A.8	Direct Horizon Slope Tests: Detailed Output	229
A.9	Matrix-Wedge Bootstrap Results	230
A.10	Horizon Diagnostics: Numerical Values	231
A.11	Finite-Horizon Tail Diagnostics	232
A.12	ρ -Sensitivity at the Five-Year Horizon	233
A.13	State-Vector and Lag-Order Robustness, CBR Proxy	233
A.14	HAC Bandwidth Sensitivity, CBR Slope Test	234
A.15	Sample and Frequency Robustness	236

A.16	Cross-Proxy Correlations	238
A.17	Common-Factor Diagnostics	239
A.18	Reliability Lower Bounds	240
A.19	Split-Sample Reliability	241
A.20	Reliability-Corrected News Correlations	242
A.21	Measurement-Error Models and Implications	244
A.22	Bootstrap Uncertainty for Reliability-Corrected Projections	245
A.23	Price-Inference Robustness Sweep	248
A.24	Orthogonal Residual Dilution	250
A.25	Targeted Residual Adversary	251
A.26	Adversarial Frontier	252
A.27	Partial-Arbitrage Loading Calibration	253
A.28	Matrix-Wedge Identity Check	255
A.29	FF25 Portfolio Exposure: News Correlations	257
A.30	FF25 Portfolio Exposure: Slope-Test Bootstrap <i>p</i> -Values	258
A.31	Median IBES Coverage by FF25 Cell	259
A.32	Sorting Evidence: News Correlations by Tercile	260
A.33	Sorting Evidence: High-minus-Low Spreads	261
A.34	Level and Reallocation: Correlations and Covariance Contributions	262
B.1	Correlation Structure of Indices in the DHS	280
B.2	Adversity Indices and Macroeconomic Beliefs in the DHS	281
B.3	Finance and Health Indices and Macroeconomic Beliefs in the DHS	282
B.4	Adversity Indices and Macroeconomic Beliefs by Housing Tenure	284
B.5	Adversity Indices and Realized Macroeconomic Outcomes	285
B.6	Positive Indices and Macroeconomic Beliefs	286
B.7	Positive Personal Experiences and Macroeconomic Beliefs	287
B.8	Share of Respondents Reporting Each Experience	298
B.9	Pairwise Correlations between Experience Indicators	299
B.10	Statistics of Individual-Level Variation in Similarity	300
B.11	Experience Database and Similarity Judgments	301
B.12	Health and Finance Similarity in the Control Sample	301
B.13	Experience and Priming Effects with Covariates	303
B.14	Experience and Priming Effects for Owners vs. Renters	304
B.15	Persistence of Priming Effects	305
B.16	Effect of Experience Similarity on Inflation and Home Price Growth Beliefs	306
B.17	Experience Similarity, Interference, and Backcasts	308
B.18	Similarity on Different Primed Experiences	309
B.19	Robustness of Similarity Effects (Panels A–B)	310
B.20	Robustness of Similarity Effects (Panels C–E)	311

B.21	Treatment without Primed Experience	312
B.22	Conversion Parameters for Composite z -Score Coefficients (Wave 1)	314
B.23	GPT Classification Confusion Table	316
B.24	Example Responses by Logic and Valence: Inflation	321
B.25	Example Responses by Logic and Valence: Home Price Growth	322
B.26	Backcasts, Context, and Valence with Covariates	323
B.27	Priming and Reasoning	324
B.28	Similarity of Reasoning Mode and NDS Recall	326
B.29	Robustness, Treatment Level	327
B.30	Beliefs and Predicted Backcasts, Context, and Valence with Covariates	328
B.31	Conversion Parameters for Composite z -Score Coefficients (TOT Wave 1)	329
C.1	Data Sources, Frequencies, and Roles in the Empirical Analysis	343
C.2	Regional Panel LP: Headline Specifications	345
C.3	Regional Panel LP: Cue Robustness	346
C.4	Regional Panel LP: Memory Robustness	346
C.5	Regional Panel LP: Sample Robustness	347
C.6	Households vs. Professional Forecasters: SPF Placebo	350

List of Figures

1.1	Horizon Diagnostics	46
1.2	Portfolio-Level Belief-News Diagnostics	58
2.1	Perceived Experience Similarity to High Inflation and High Home Price Growth	89
2.2	Word Clouds for Priming Treatments	91
2.3	Quantification of Memory Effects	100
2.4	Word Clouds of Macroeconomic Reasoning by Valence	103
2.5	Word Clouds of Macroeconomic Reasoning by Type	104
3.1	Mechanism Schematic	123
3.2	Fixed-Point Geometry of the Temporary-Equilibrium Map	137
3.3	Dynamic Path of (c_t, m_t) Under a Temporary Cue Perturbation	141
3.4	National Demand Responses to Adverse EPU	151
3.5	Household Sentiment Responses to Adverse Public Cues	151
3.6	Adverse-Memory Amplification of State Permits Response to Adverse Cues	158
C.1	Memory-Gradient Figure with p_{99} Tail Bar	348

Acknowledgments

First, I would like to thank my advisors and dissertation committee: Andrei Shleifer, Adi Sunderam, and Samuel Hanson. I have been extraordinarily fortunate to be guided by people whose intellectual seriousness comes with such generosity, humor, and care. If I can offer others even a fraction of the guidance, selflessness, and intellectual openness they have offered me, I will consider myself very lucky.

Andrei, meeting you eight years ago as an undergraduate changed the direction of my life. Since then, you have shaped how I think, how I work, and, more importantly, the person I am. Thank you for your trust, your unconditional support, and for making space for me from the very beginning. Few things have meant more to me than knowing that I could always turn to you, and that you would meet every question, uncertainty, and obstacle with care, clarity, and unwavering commitment. I am grateful not only for everything I learned from you, but also for the sense of belonging and possibility you helped create – through our work together, our conversations, our runs, and the many ways you made Harvard and the U.S. feel like home. You are family to me, and words cannot express how grateful I am.

Adi, thank you so much for your kindness, support, and intellectual example. I have learned so much from you in so many different contexts, and especially from the clarity and wit with which you approach ideas. Your ability to be precise, incisive, and humane all at once has been a model for me, and one I deeply admire. I feel very lucky to have been part of the community you help sustain, and your influence will stay with me far beyond my time here.

Sam, thank you so much for your extraordinary generosity and for the intellectual energy

and care with which you engage ideas and people. I have learned so much from the way you think: seriously, openly, and with a remarkable ability to make difficult questions feel alive and worth struggling with. You have a rare gift for making people feel that their ideas matter, while somehow making the whole process incredibly entertaining. I am deeply grateful for your guidance, and your example has been a true inspiration.

I would especially like to thank Pedro Bordalo and Nicola Gennaioli, from whom I have learned so much – particularly how to combine creativity, rigor, and intellectual ambition. Working with them has been one of the great privileges of my time in graduate school, and has shaped how I think about asking questions, developing ideas, and pushing research forward. I am deeply honored to call them my co-authors.

I am also indebted to Nicholas Barberis, John Y. Campbell, Gabriel Chodorow-Reich, Katie Coffman, Hans-Helmut Kotz, Spencer Kwon, Florencio Lopez-de-Silanes, Marco Sammon, Ludwig Straub, and Maarten van Rooij for generously sharing their time, insights, and encouragement. Their questions, suggestions, and conversations have helped me see this work more clearly and improve it in many ways.

The completion of this dissertation would not have been possible without the help and support of many people. Fiona Y. Chen has been there from the first day of the Ph.D., unwavering in her support, curiosity, and patience. She has been with me through thick and thin, and I am deeply grateful for the steadiness, humor, and warmth she has brought to these years. Her ability to excel at so many different things – research, art, music, and everything in between – all while remaining one of the kindest, most positive, and most thoughtful people I know, has been truly inspiring. I am very grateful for this friendship, and I am excited for all that lies ahead.

I am also deeply grateful to Cameron Cohen, Toren Fronsdal, Dominic Russel, and Bas Sanders. Our dinners, conversations about research and life, absurd workout challenges, bets, and trips together have been among the great joys of these years. They have made graduate school feel lighter, funnier, and much more human. I am incredibly grateful for their friendship, and for the way they became my American family. I am so excited to see all the extraordinary things they will do, personally and professionally, and even more grateful that I get to be part of their lives along the way.

My life in Cambridge has been much more interesting, inspiring, and entertaining thanks to the friendship and company of Moritz Alkofer, Hemanth Asirvatham, Jake Carlson, Jorge Colmenares, Toto Gräber, Erik Guander, Jonas Gudera, Oskar Haushofer, Anna Handler, Fred Heiding, George Hu, Jeremy Huang, Jess Jenkins, Jingyi Jia, Iris Li, Henrik Mader, Pierfrancesco Mei Innocenti, Elliott Mokski, Estefania Morales, Lauren Rice, Kunal Sangani, Suproteem Sarkar, Claire Shi, Michael Stanley, Davide Tagliatalata, Ruby Zhang, and Borui Zhu.

I owe a special thank you to my friends from before graduate school, who always provided perspective, fun, and a sense of home – both physical and mental – in New York and Europe: Rujual “Reggie” Bains, Moritz Bawart, David Friedman, Julian Schäfermeier, and especially David P. Klambauer, with whom my life has been intertwined in almost every dimension since the day he and I met, and whose ability to bring levity to any situation is unparalleled.

I feel a similar gratitude for my friends back home: Valerie Gottschalk, Magnus Gschoßmann, Marcella Mutschler, and Paolo “Fraile” Walcher. They have always made me feel connected, even when separated by the ocean we sailed across together so many years ago. I am deeply grateful for a friendship that has survived distance, time zones, and different lives, and that still feels as close as ever.

Finally, I want to thank my German and Croatian families – especially my Mravak, Vučica, and Ćutić families – for their unconditional love and support. Above all, my deepest gratitude belongs to my mother, Branka, my father, Ulrich, and my sister, Carla. Thank you for your love, your patience, your sacrifices, your humor, and your unwavering belief in me. I continue to learn from you every day, and the enthusiasm, optimism, and excitement with which you approach life are a constant source of inspiration and energy for me. Your curiosity about people, places, ideas, and the world around you continues to shape the way I want to move through life. You make me want to be the best, kindest, and most generous version of myself. Everything I am, and everything I have been able to do, is because of you. I cannot possibly express how much “Liebe und Zuneigung” I have for you.

To my mother, my father, and my sister.

Chapter 1

Belief-Driven Risk Premium News

Chapter Summary

Standard return decompositions attribute much of unexpected stock-market variation to risk premium news, typically interpreted as revisions in expected future excess returns. I argue that this interpretation is incomplete. I develop a behavioral extension of the Campbell and Ammer (1993) decomposition that embeds biased cash-flow beliefs in standard return-news accounting. When biased optimism about cash flows raises prices, an econometrician who does not observe the belief distortion can interpret the same price movement as lower future expected returns. I test this mechanism using target-date-aligned analyst forecast revisions as news-like measures of subjective cash-flow expectations. Consistent with the model's predictions, the implied belief-bias news is negatively related to risk premium news, has meaningful scale, and comoves positively with expected returns. The forecast restrictions implied by the behavioral model are compatible with the data, and the relation survives price-inference, measurement, adversarial-residual, and cross-sectional tests. I argue that standard return decompositions neglect a quantitatively meaningful behavioral component: part of what is conventionally called risk premium news reflects biased cash-flow beliefs that move asset prices.

1.1 Introduction

Why are stock returns predictable, and what does that predictability mean? Standard return decompositions attribute much of unexpected stock-market variation to revisions in expected future excess returns. In the conventional interpretation, this object is read as risk premium news: news about time-varying required compensation for bearing risk. The Campbell and Shiller (1988) and Campbell and Ammer (1993) frameworks have made this interpretation central to modern empirical asset pricing, and a robust finding in that literature is that risk premium news is a dominant component of unexpected returns (Fama 1991; Campbell 1991; Vuolteenaho 2002; Chen and Zhao 2009).

This paper argues that this interpretation is incomplete. Behavioral theories of asset pricing offer a different mechanism: investors form biased beliefs about future cash flows, and when those beliefs move prices, subsequent returns can look predictable even without purely rational variation in required returns. If investors become too optimistic about future cash flows, prices rise today. A rational econometrician who does not observe the biased belief revision directly sees the same price increase and, through the standard return-news decomposition, infers lower future expected excess returns. The price movement can therefore appear as risk premium news even though its source is biased cash-flow beliefs.

The paper's core question is therefore not whether risk premium news exists. It is how risk premium news should be interpreted once investor beliefs are allowed to depart from full-information rational expectations. The return decomposition is an accounting identity. It remains valid whether prices are moved by rational discount-rate variation, biased cash-flow beliefs, or both. What changes is the economic content of the object that the econometrician

labels risk premium news.

I formalize this idea by developing a behavioral extension of the Campbell and Ammer (1993) decomposition. The theory uses two expectation operators: one for the rational econometrician who constructs the return-news decomposition, and one for investors whose cash-flow beliefs may be biased. In the sharp behavioral-pricing benchmark, investors perceive expected excess returns as constant and revise their beliefs about future cash flows. The rational econometrician, who does not directly observe this revision, attributes the resulting price movement to risk premium news. Thus the benchmark delivers a simple sign restriction: unexpected cash-flow optimism should appear as negative risk premium news.

The sharp benchmark is useful because it makes the mechanism transparent, but the paper does not require it to hold exactly. If rational arbitrageurs partially offset behavioral price pressure, only part of the belief-driven valuation movement remains in prices. Risk premium news then contains two components: residual rational expected-return news and an attenuated belief-driven component. The paper's main claim is therefore a component claim. I do not argue that all risk premium news is behavioral, or that standard decompositions should be discarded. I argue that part of what is conventionally called risk premium news can reflect biased cash-flow beliefs that move asset prices.

To take this idea to the data, I construct target-date-aligned analyst forecast revisions as proxies for revisions in subjective cash-flow expectations. The target-date alignment is important because naive analyst revisions can mechanically compare forecasts for different fiscal targets. Such horizon rollover can look like belief revision even when expectations about a fixed future cash-flow object have not changed. The aligned revisions instead compare forecasts about the same target date across adjacent months. I construct three market-level behav-

ioral cash-flow news proxies: Composite Belief Revision (CBR), Long-Term Growth (LTG), and Short-Term Revision (STR). The aligned proxies behave like news-like revision measures: they are centered near zero, much less persistent than simple horizon-label revisions, and share a large common component.

I combine these analyst revisions with rational cash-flow news from a vector autoregression (VAR) return-news system to construct empirical belief-bias news, the difference between subjective belief revisions and rational cash-flow news. The empirical design asks three questions. First, does belief-bias news have the sign predicted by the theory? Second, is the belief-bias object economically meaningful relative to risk premium news? Third, do predictable returns comove with predictable belief revisions in the way implied by the behavioral model?

The answer to all three questions is yes. Across CBR, LTG, and STR, belief-bias news is negatively related to rational risk premium news, with correlations near -0.10 . The measured belief-bias object is economically large: its variance is about 0.28 – 0.34 of the variance of risk premium news. Forecasts move in the predicted direction as well. Fitted return forecasts and fitted belief-revision forecasts have positive signed correlations between 0.46 and 0.51 . The corresponding sample forecast-comovement R^2 is 0.21 – 0.26 , and the population measure is around 0.30 . These magnitudes indicate that the measured belief component has the predicted sign, meaningful scale, and substantial forecast-side comovement.

The behavioral model also implies forecast restrictions. If predictable expected returns reflect predictable revisions in subjective cash-flow beliefs, then the same state variables should forecast returns and belief revisions with the same 1-step slopes. This restriction is not rejected for any of the three belief measures under a null-imposed bootstrap. Direct horizon tests reproduce the 1-step result and remain compatible with the model at 1-, 2-, and 3-year

horizons. At five years, however, exact cumulative equality rejects sharply. I interpret this result as an important limit of the sharp benchmark: the short- and medium-horizon forecast evidence supports the belief-driven channel, while exact long-horizon cumulative equality is too strong.

The same distinction appears at a sharper level of comparison. The full behavioral benchmark implies that risk premium news and the negative of belief-bias news should match each other exactly in every period, not just on average. That stronger condition fails in the data: the gap between the two series is economically meaningful. This failure is consistent with the partial-arbitrage extension of the benchmark. Under partial arbitrage, a nonzero gap reflects residual rational expected-return news, incomplete pass-through of biased beliefs, rate-belief components, or proxy error. The gap therefore qualifies the interpretation. It shows that the evidence supports a belief-driven component inside risk premium news, not a complete behavioral relabeling of the entire object.

I then examine several alternative explanations. Measurement error changes how scale statistics should be read, but it does not naturally generate the consistent sign pattern across three different analyst proxies. Removing the full contemporaneous price-dividend projection from the belief proxies leaves the sign, scale, and slope evidence essentially unchanged, so the result is not driven by analysts mechanically inferring beliefs from current prices through this channel. Adversarial-residual exercises show that generic orthogonal-residual news must be several times as volatile as measured risk premium news to erase the baseline relation. A residual deliberately aligned with the belief proxy can overturn the relation with about 1% of the variance of risk premium news, which identifies the structured rational alternative that remains hardest to rule out.

The evidence also appears beyond the aggregate market return. In the 25 Fama and French (1993) size-by-book-to-market portfolios, 19 portfolios have the predicted negative correlation between aggregate belief-bias news and portfolio-level risk premium news, and none rejects the slope restriction under the bootstrap. A decomposition into common belief revisions and size-weighted reallocation shows how firm-level revisions aggregate into the market-level proxy. LTG primarily reflects common movements in long-run growth beliefs, STR primarily reflects size-weighted reallocation of revisions across firms, and CBR combines both channels. The cross-section therefore adds breadth and mechanism to the aggregate time-series evidence.

The paper makes four contributions. First, it provides a behavioral reinterpretation of the risk premium news object in the Campbell and Ammer (1993) decomposition. The contribution is not a new accounting identity, but a new interpretation of an existing object when omitted belief wedges move prices. Second, it constructs target-date-aligned analyst belief revisions and maps them into empirical counterparts of behavioral cash-flow belief-bias news. Third, it provides evidence that these belief-bias innovations line up with risk premium news in sign, scale, forecast comovement, and slope restrictions. Fourth, it separates the sharp full-behavioral benchmark from broader partial-arbitrage interpretations, using matrix wedges, horizon diagnostics, measurement-error algebra, price-inference tests, adversarial residuals, and cross-sectional evidence to separate the two readings.

The paper connects to several literatures. The return-news decomposition tradition of Campbell and Shiller (1988), Campbell and Ammer (1993), and Campbell and Vuolteenaho (2004) provides the accounting framework that I preserve and reinterpret. The behavioral asset-pricing literature, beginning with De Long et al. (1990) and Shleifer and Vishny (1997) and developed in Shleifer (2000), supplies the basic premise: biased beliefs and limited arbitrage

can move prices. I ask how that premise shows up inside the standard return-news decomposition. La Porta (1996) and Lakonishok et al. (1994) document long-horizon return predictability tied to expectation errors; I map that evidence into canonical return-news terms and connect 1-period restrictions to longer-horizon predictability. The empirical analysis builds on the analyst- and survey-expectations literature (Coibion and Gorodnichenko 2015; Greenwood and Shleifer 2014; Bordalo et al. 2019; De La O and Myers 2021; Nagel and Xu 2023; Bretscher et al. 2025; Bordalo et al. 2025c), to which I contribute a Campbellian accounting bridge from belief revisions to risk premium news, a target-date-aligned revision innovation, and cross-equation restrictions linking belief revisions to risk premium news.

The rest of the paper proceeds as follows. Section 1.2 develops the theory and derives the benchmark identity, the partial-arbitrage extension, the forecast restriction, and the matrix and horizon diagnostics. Section 1.3 describes the data, analyst belief proxies, and news construction. Section 1.4 presents the main time-series evidence on sign, scale, forecast comovement, and slope restrictions. Section 1.5 studies the matrix wedge and horizon diagnostics. Section 1.6 evaluates measurement, price-inference, and adversarial-residual concerns. Section 1.7 presents cross-sectional and mechanism evidence. Section 1.8 concludes.

1.2 Framework for Risk Premium News with Biased Beliefs

Risk premium news is an accounting object. The theory in this section keeps that object fixed and changes its interpretation. I first define the return-news decomposition used throughout the paper. I then show that, when investors hold biased cash-flow beliefs, the object a rational econometrician calls risk premium news can contain belief-bias news. The section derives three testable predictions for the empirical analysis: a sign prediction, a forecast-slope restric-

tion, and a stronger long-horizon matrix condition. It also explains why the sharp benchmark can fail without eliminating the behavioral channel.

1.2.1 Return-News Decomposition

The starting point is the standard log-linear return identity of Campbell and Shiller (1988) and the return-news decomposition of Campbell and Ammer (1993). Let rx_{t+1} denote the excess log return, rf_{t+1} the risk-free rate, and Δd_{t+1} dividend growth. Let $\mathbb{E}_t^{(r)}$ denote rational expectations under the econometrician's information set. The unexpected excess return can be written as

$$rx_{t+1} - \mathbb{E}_t^{(r)} rx_{t+1} = N_{CF,t+1}^{(r)} - N_{RF,t+1}^{(r)} - N_{RP,t+1}^{(r)}. \quad (1.1)$$

The three terms are rational cash-flow news, risk-free-rate news, and risk premium news. They are defined as

$$\begin{aligned} N_{CF,t+1}^{(r)} &= \left(\mathbb{E}_{t+1}^{(r)} - \mathbb{E}_t^{(r)} \right) \sum_{j \geq 1} \rho^j \Delta d_{t+1+j}, \\ N_{RF,t+1}^{(r)} &= \left(\mathbb{E}_{t+1}^{(r)} - \mathbb{E}_t^{(r)} \right) \sum_{j \geq 1} \rho^j rf_{t+1+j}, \\ N_{RP,t+1}^{(r)} &= \left(\mathbb{E}_{t+1}^{(r)} - \mathbb{E}_t^{(r)} \right) \sum_{j \geq 1} \rho^j rx_{t+1+j}, \end{aligned}$$

where ρ is the Campbell–Shiller log-linearization constant.¹

This identity is the paper's accounting backbone. A high unexpected return can reflect news about future cash flows, news about future risk-free rates, or news about future ex-

¹ The monthly log-linearization constant is $\rho_{\text{monthly}} = 0.998122$ throughout, calibrated from the unconditional mean of the log dividend-price ratio in the legacy sample. The corresponding annualized value is $\rho_{\text{annual}} \approx 0.978$.

pected excess returns. The conventional interpretation treats $N_{RP,t+1}^{(r)}$ as discount-rate or expected-return news. I keep the accounting definition fixed. The paper asks whether this same expected-return-news object can contain belief-driven valuation movements when prices are partly set by investors with biased cash-flow expectations.

Two points are important. First, the superscript (r) describes the expectations operator used to construct the news terms, not necessarily the beliefs that set prices. Second, the decomposition attributes return innovations using the variables and beliefs included in the accounting system. If biased cash-flow beliefs move prices but the decomposition does not model those beliefs directly, part of the belief-driven valuation movement can be attributed to rational expected-return news. In that sense, the paper studies a structured form of misattribution: the accounting identity remains valid, but the economic interpretation of $N_{RP,t+1}^{(r)}$ changes.

1.2.2 Behavioral Cash-Flow Belief Bias

The decomposition in equation (1.1) is constructed under the rational expectations operator $\mathbb{E}_t^{(r)}$. I now introduce a second expectations operator, $\mathbb{E}_t^{(b)}$, which represents the cash-flow beliefs of behavioral investors. The two operators agree on realized cash flows but can differ in how they forecast future cash flows before uncertainty is resolved.

Rational cash-flow news is

$$N_{CF,t+1}^{(r)} = \left(\mathbb{E}_{t+1}^{(r)} - \mathbb{E}_t^{(r)} \right) \sum_{j \geq 1} \rho^j \Delta d_{t+1+j}.$$

Behavioral cash-flow news is the analogous revision under behavioral beliefs:

$$N_{CF,t+1}^{(b)} = \left(\mathbb{E}_{t+1}^{(b)} - \mathbb{E}_t^{(b)} \right) \sum_{j \geq 1} \rho^j \Delta d_{t+1+j}.$$

The cash-flow belief-bias news is the difference between these two revisions:

$$\eta_{CF,t+1} \equiv N_{CF,t+1}^{(b)} - N_{CF,t+1}^{(r)}.$$

This object measures how much more, or less, behavioral investors revise their expected future cash flows relative to the rational benchmark. If behavioral revisions exceed rational revisions in the direction of the news, $\eta_{CF,t+1}$ is positive. If they fall short or move in the opposite direction, it is negative.

The return-news identity depends on innovations from the perspective of the rational econometrician. I therefore define the rationally unexpected component of the belief-bias news as

$$\zeta_{CF,t+1} \equiv \eta_{CF,t+1} - \mathbb{E}_t^{(r)} \eta_{CF,t+1}.$$

This distinction is important. The level of behavioral optimism that is already predictable at time t can affect prices, but it is not news at $t+1$. The object that enters the unexpected-return decomposition is the innovation in the behavioral cash-flow wedge, $\zeta_{CF,t+1}$.

The empirical analysis uses analyst forecast revisions to proxy for this belief-revision component. The proxy is not assumed to be perfect. It is useful because it provides an observable measure of revisions in subjective cash-flow expectations, which can then be compared to the news terms in equation (1.1). Measurement error, incomplete proxying, and nonclassical

proxy contamination are addressed in Sections 1.2.9 and 1.6.

1.2.3 Full Behavioral-Pricing Benchmark

The sharp benchmark shows how biased cash-flow beliefs enter risk premium news. Suppose behavioral investors are marginal in price setting, have correct beliefs about the risk-free-rate stream, and perceive expected excess returns as constant. Their price revisions therefore reflect revisions in expected future cash flows rather than revisions in perceived future expected returns.

Proposition 1. *Under the full behavioral-pricing benchmark, rational risk premium news equals the negative of unexpected cash-flow belief bias:*

$$N_{RP,t+1}^{(r)} = -\zeta_{CF,t+1}. \quad (1.2)$$

The proof compares two decompositions of the same realized return. Equation (1.1) is the rational Campbell and Ammer (1993) decomposition, constructed using $\mathbb{E}_t^{(r)}$. Under the benchmark assumptions, the same return is priced by behavioral investors using their cash-flow beliefs, correct risk-free-rate expectations, and a constant perceived excess return, so $N_{RP,t+1}^{(b)} = 0$. Equating the rational and behavioral decompositions and isolating $N_{RP,t+1}^{(r)}$ yields

$$N_{RP,t+1}^{(r)} = -\left(\eta_{CF,t+1} - \mathbb{E}_t^{(r)}\eta_{CF,t+1}\right) = -\zeta_{CF,t+1}.$$

Appendix A.1.2 gives the full algebra; Appendix A.1.3 shows how the belief-bias news $\eta_{CF,t+1}$ can be written in terms of the behavioral valuation wedge in the log price-dividend ratio.

Proposition 1 preserves the return-news identity and changes the interpretation of one term. A rational econometrician who omits biased cash-flow beliefs does not see the behavioral cash-flow wedge directly. Instead, the unexpected component of that wedge appears with the opposite sign in rational risk premium news. Positive unexpected cash-flow optimism raises prices today and lowers expected future excess returns from the rational econometrician's perspective, generating negative risk premium news.

This benchmark is intentionally sharp. It requires behavioral investors to be fully marginal, so the belief-driven valuation wedge fully passes through to prices. The next subsection relaxes this assumption. If rational arbitrageurs offset part of the belief-driven price pressure, only a fraction of the behavioral valuation wedge remains in prices. I call this partial arbitrage; the parameter λ below measures the surviving, unarbitraged share of the behavioral wedge.

1.2.4 Partial Arbitrage

The full behavioral-pricing benchmark assumes that biased cash-flow beliefs fully pass through to prices. This is useful for deriving the sharp identity in equation (1.2), but it is stronger than the paper needs. If rational arbitrageurs absorb part of behavioral demand, the same belief-bias component can survive in risk premium news with a smaller loading.

Let pd_t^* denote the rational log price-dividend ratio, and let m_t denote the wedge in the log price-dividend ratio implied by biased cash-flow beliefs. I parameterize partial arbitrage as

$$pd_t^\lambda = pd_t^* + \lambda m_t, \quad 0 \leq \lambda \leq 1.$$

The parameter λ measures the surviving share of the behavioral valuation wedge. When

$\lambda = 1$, biased beliefs fully affect prices, giving the full behavioral benchmark. When $\lambda = 0$, arbitrage fully removes the wedge from prices, giving the rational benchmark. Intermediate values $0 < \lambda < 1$ represent partial arbitrage.

Proposition 2. *Under partial arbitrage,*

$$N_{RP,t+1}^{\lambda,r} = N_{RP,t+1}^{*,r} - \lambda \zeta_{CF,t+1}. \quad (1.3)$$

Here $N_{RP,t+1}^{*,r}$ is the rational risk premium news component that would arise in the absence of the behavioral wedge.

Proposition 2 shows that arbitrage attenuates the behavioral component rather than necessarily eliminating it. Risk premium news now contains two pieces: residual rational risk premium news, $N_{RP,t+1}^{*,r}$, and the surviving belief-driven component, $-\lambda \zeta_{CF,t+1}$. The full benchmark in Proposition 1 is the special case with $\lambda = 1$ and $N_{RP,t+1}^{*,r} = 0$.²

A useful implication concerns the projection of risk premium news on belief-bias news. From equation (1.3),

$$\frac{\text{Cov}(N_{RP,t+1}^{\lambda,r}, \zeta_{CF,t+1})}{\text{Var}(\zeta_{CF,t+1})} = -\lambda + \frac{\text{Cov}(N_{RP,t+1}^{*,r}, \zeta_{CF,t+1})}{\text{Var}(\zeta_{CF,t+1})}.$$

Thus the projection coefficient identifies $-\lambda$ only under the maintained orthogonality condition

$$\text{Cov}(N_{RP,t+1}^{*,r}, \zeta_{CF,t+1}) = 0.$$

² Appendix A.1.4 derives equation (1.3) from the log-linear return identity. The key step is that only the surviving share λ of the behavioral valuation wedge affects prices, so the belief-bias innovation enters rational risk premium news with loading $-\lambda$.

This condition is an identifying assumption, not a theorem. If residual rational risk premium news comoves with belief-bias news, the projection coefficient combines the behavioral loading and that rational comovement.

The empirical implication is therefore a sign-and-presence implication rather than a structural estimate of λ . A negative relation between $N_{RP,t+1}^{\lambda,r}$ and $\zeta_{CF,t+1}$ is evidence that the belief component survives in prices. Interpreting the projection coefficient as $-\lambda$, however, requires the additional orthogonality assumption above.

1.2.5 Rate-Belief Bias and Net Valuation Beliefs

The benchmark identity in Proposition 1 assumes that behavioral and rational investors agree about the risk-free-rate stream. This subsection relaxes that assumption. If behavioral investors also hold biased beliefs about future risk-free rates, risk premium news reflects the net valuation effect of cash-flow and rate-belief errors.

Let $N_{RF,t+1}^{(b)}$ denote the behavioral revision in expected future risk-free rates,

$$N_{RF,t+1}^{(b)} = \left(\mathbb{E}_{t+1}^{(b)} - \mathbb{E}_t^{(b)} \right) \sum_{j \geq 1} \rho^j r f_{t+1+j}.$$

Define the risk-free-rate belief-bias news as

$$\eta_{RF,t+1} \equiv N_{RF,t+1}^{(b)} - N_{RF,t+1}^{(r)},$$

and its rationally unexpected component as

$$\zeta_{RF,t+1} \equiv \eta_{RF,t+1} - \mathbb{E}_t^{(r)} \eta_{RF,t+1}.$$

Then the full behavioral-pricing benchmark becomes

$$N_{RP,t+1}^{(r)} = -\zeta_{CF,t+1} + \zeta_{RF,t+1}. \quad (1.4)$$

The sign difference is intuitive. Unexpected cash-flow optimism raises prices and therefore lowers rational expected future excess returns, generating negative risk premium news. Unexpected optimism about future risk-free rates raises perceived discount rates, lowering prices and raising rational expected future excess returns. This generates positive risk premium news.

Equation (1.4) shows that the exact theoretical object is net valuation belief bias, not cash-flow belief bias alone. Define

$$\zeta_{NET,t+1} \equiv \zeta_{CF,t+1} - \zeta_{RF,t+1}.$$

Then equation (1.4) can be written as

$$N_{RP,t+1}^{(r)} = -\zeta_{NET,t+1}.$$

The empirical analysis focuses on cash-flow beliefs because analyst forecasts directly measure beliefs about future cash flows. This is a measurement restriction, not a theoretical necessity. If risk-free-rate belief bias is empirically important and correlated with cash-flow belief bias, then the cash-flow proxy measures only one component of the net valuation belief wedge. Appendix A.1.6 derives equation (1.4) and shows how this extension combines with partial arbitrage.

1.2.6 Forecast Restrictions

The news identities above also imply forecast restrictions. The full benchmark says that predictable variation in excess returns should line up with predictable variation in behavioral cash-flow revisions. This implication gives the paper's main slope test.

Let $\bar{r}x$ denote the constant perceived expected excess return in the full behavioral benchmark. Combining equation (1.1) with Proposition 1, realized excess returns can be written as

$$rx_{t+1} = \bar{r}x + N_{CF,t+1}^{(r)} - N_{RF,t+1}^{(r)} + \eta_{CF,t+1}.$$

Taking rational expectations at time t , and using

$$\mathbb{E}_t^{(r)} N_{CF,t+1}^{(r)} = \mathbb{E}_t^{(r)} N_{RF,t+1}^{(r)} = 0,$$

gives

$$\mathbb{E}_t^{(r)} rx_{t+1} = \bar{r}x + \mathbb{E}_t^{(r)} \eta_{CF,t+1}. \quad (1.5)$$

Equation (1.5) is the forecast-side version of Proposition 1. It says that any variable that predicts excess returns under the rational decomposition should also predict behavioral cash-flow revisions with the same slope. The object is forecastable variation, not realized returns.

This is the restriction taken to the data. In the empirical VAR, let $e'_1\Gamma$ denote the row of forecast coefficients for next-period excess returns, and let $e'_3\Gamma$ denote the row of forecast coefficients for the behavioral revision proxy. Then equation (1.5) implies the slope restriction

$$e'_1\Gamma = e'_3\Gamma. \quad (1.6)$$

Section 1.3 defines the state vector and the analyst-revision proxy used to implement this restriction. For now, equation (1.6) should be read as a simple forecast statement: the same information that predicts excess returns should predict behavioral cash-flow revisions one-for-one.

Equation (1.6) is a sharp benchmark implication. Non-rejection supports the full benchmark's forecast restriction. Rejection does not by itself reject the broader behavioral channel: partial arbitrage, residual rational expected-return variation, rate-belief bias, or measurement error can all break exact slope equality while leaving a belief-driven component in risk premium news.³

1.2.7 Matrix Condition and Pointwise News Equality

The slope restriction is a forecast restriction. It asks whether the same information forecasts excess returns and behavioral cash-flow revisions. The sharp benchmark also has a stronger implication: the constructed news objects should line up point by point.

To state this restriction, use the same linear forecasting representation introduced above and define the present-value matrix

$$\Lambda = \rho \Gamma (I - \rho \Gamma)^{-1}.$$

This matrix maps current innovations into revisions in expected future variables. Let e_1 select

³ A compact way to write the generalized forecast relation is $\gamma_{rx}^\lambda = \gamma_* + \lambda_{CF} \gamma_{CF} - \lambda_{RF} \gamma_{RF}$, where γ_{rx}^λ is the expected-return slope, γ_* is residual rational expected-return predictability, and γ_{CF} and γ_{RF} are the cash-flow- and rate-belief slope components. The full benchmark sets $\gamma_* = 0$, $\lambda_{CF} = 1$, and $\lambda_{RF} = 0$. Appendix A.1.7 derives this relation.

excess returns, e_2 select the risk-free rate, and e_3 select the behavioral cash-flow revision. Then the sharp benchmark implies the matrix condition

$$e'_3 = e'_1 + e'_2 \Lambda. \quad (1.7)$$

Equation (1.7) is stronger than the 1-step slope restriction. The slope restriction compares predicted 1-period excess returns and predicted belief revisions. The matrix condition compares the full innovation-level news objects that enter the Campbell and Ammer (1993) decomposition. It requires the behavioral revision innovation to equal the current unexpected excess return plus the present-value revision in future risk-free rates.

It is useful to define the matrix wedge

$$W_{t+1}^{\text{matrix}} \equiv N_{RP,t+1}^{(r)} + \zeta_{CF,t+1}.$$

In the linear system,

$$W_{t+1}^{\text{matrix}} = (e'_3 - e'_1 - e'_2 \Lambda) \varepsilon_{t+1}. \quad (1.8)$$

Thus the sharp pointwise benchmark implies $W_{t+1}^{\text{matrix}} = 0$ for every innovation ε_{t+1} . A nonzero matrix wedge means that the constructed long-horizon news objects do not satisfy exact pointwise equality.

The matrix condition and the slope restriction answer different questions. The slope restriction asks whether expected returns and expected belief revisions have the same 1-step forecast slopes. The matrix condition asks whether the full Campbell and Ammer (1993) news objects satisfy the pointwise identity implied by the sharp benchmark. The sharp benchmark

implies both restrictions, but they are independent restrictions outside the benchmark: neither generally implies the other, and in finite samples they can diverge.

Under partial arbitrage, the matrix wedge has a direct interpretation:

$$W_{t+1}^{\text{matrix}} = N_{RP,t+1}^{*,r} + (1 - \lambda) \zeta_{CF,t+1}.$$

Thus the wedge is not simply a failure statistic. It can reflect residual rational risk premium news, incomplete arbitrage, or measurement error in the belief proxy. It measures distance from the strongest pointwise benchmark, not the absence of a belief-driven component. The matrix wedge is therefore an interpretation diagnostic for the strongest pointwise benchmark, not an independent overidentifying moment.⁴

1.2.8 Horizon Restrictions and Convergence

The 1-step slope restriction and the matrix condition are two endpoints. The slope restriction is a 1-period forecast restriction. The matrix condition is an infinite-horizon pointwise news restriction. The empirical analysis also studies finite horizons, which bridge these two objects and connect the framework to long-horizon return predictability.

Let the finite-horizon present-value matrix be

$$\Lambda_h = \sum_{j=1}^h \rho^j \Gamma^j.$$

⁴ Under the standard innovation-space construction, the matrix-wedge variance is algebraically pinned down by the projection coefficient and the behavioral scale. Genuine overidentification requires a forecast-space restriction, horizon restriction, or external proxy variation. Appendix A.2.4 derives the matrix wedge, and Appendix A.8.6 shows the algebraic identity linking the wedge variance to the projection coefficient and scale.

The infinite-horizon matrix in Section 1.2.7 is the limit,

$$\Lambda = \sum_{j=1}^{\infty} \rho^j \Gamma^j = \rho \Gamma (I - \rho \Gamma)^{-1}.$$

The omitted tail is

$$\Lambda - \Lambda_h = \rho^{h+1} \Gamma^{h+1} (I - \rho \Gamma)^{-1}, \quad (1.9)$$

when the inverse exists. Thus finite-horizon news objects converge to the infinite-horizon object, but the speed of convergence depends on the persistence of the state vector.

Finite-horizon forecast restrictions use a closely related matrix,

$$B_h = \sum_{j=1}^h \rho^{j-1} \Gamma^j = \rho^{-1} \Lambda_h.$$

The horizon- h forecast restriction is

$$e'_1 B_h = e'_3 B_h.$$

It says that the state variables should forecast discounted cumulative excess returns and discounted cumulative behavioral cash-flow revisions with the same slopes through horizon h .

In the VAR-implied system, the 1-step slope restriction implies the horizon restrictions.

If

$$e'_1 \Gamma = e'_3 \Gamma,$$

then right-multiplying the slope restriction by Γ^{j-1} gives

$$e'_1 \Gamma^j = e'_3 \Gamma^j,$$

for every $j \geq 1$, and therefore $e_1' B_h = e_3' B_h$ for every finite horizon h .

This implication is mechanical inside the linear forecasting system. It clarifies what the 1-step benchmark says about longer horizons: if the same state variables forecast 1-period returns and 1-period belief revisions with the same slopes, then the iterated system also gives the same cumulative forecast slopes at every finite horizon.

Finite-horizon present-value objects have a different role. They approximate the infinite-horizon news objects in the Campbell and Ammer (1993) decomposition. Replacing Λ with Λ_h truncates the present-value sum, leaving the tail in equation (1.9). A finite-horizon object therefore need not equal the infinite-horizon object even when it is economically informative. The relevant question is how quickly it converges as h grows.

This distinction is useful for the empirical analysis. Medium horizons show whether the forecast-side restriction survives beyond one period. Longer horizons show how quickly truncated present-value objects approach the infinite-horizon decomposition. The empirical inference treats the 1-step ($h = 1$) restriction as the primary test and reports horizon objects as diagnostics rather than as additional independent rejections, to preempt multiple-testing concerns across many horizons. Appendix A.2 gives the algebra of Λ_h , B_h , and the convergence diagnostics.

1.2.9 Scale, Forecast Comovement, and Measurement

The theory delivers an exact benchmark identity, but the empirical analysis also needs interpretable magnitudes. I use two complementary objects. The first compares the variance of belief-bias news to the variance of risk premium news. The second asks whether predictable returns and predictable belief revisions move together.

Define the behavioral scale ratio as

$$\mathcal{B}_{\text{scale}} = \frac{\text{Var}(\zeta_{CF,t+1})}{\text{Var}(N_{RP,t+1}^{(r)})}. \quad (1.10)$$

This statistic asks whether the belief-bias component is large relative to the risk premium news object. It is not a structural share. In particular, $\mathcal{B}_{\text{scale}}$ can exceed one, and it does not say what fraction of risk premium news is caused by biased beliefs. It is a scale diagnostic.

The second object summarizes forecast comovement. Under the full benchmark, equation (1.5) implies that expected excess returns and expected behavioral cash-flow revisions should move one-for-one, up to a constant. A bounded way to summarize this implication is

$$R_{\text{forecast}}^2 = \text{Corr}\left(\mathbb{E}_t^{(r)}rx_{t+1}, \mathbb{E}_t^{(r)}\eta_{CF,t+1}\right)^2. \quad (1.11)$$

The empirical implementation uses fitted return forecasts and fitted forecasts of the analyst-revision proxy to compute the same statistic. Forecast R_{forecast}^2 is bounded between zero and one, but it is not a structural share. It asks a narrower question: how tightly does predictable return variation comove with predictable belief-revision variation?

The news and forecast objects have opposite predicted signs. Proposition 1 predicts a negative relation between risk premium news and unexpected belief-bias news,

$$\text{Corr}\left(N_{RP,t+1}^{(r)}, \zeta_{CF,t+1}\right) < 0,$$

while equation (1.5) predicts positive comovement between expected returns and expected belief revisions.

Measurement error affects scale, correlations, and projection coefficients differently. Suppose the observed proxy is

$$\tilde{\zeta}_{t+1} = \zeta_{CF,t+1} + u_{t+1},$$

where u_{t+1} is classical additive noise. Then

$$\text{Var}(\tilde{\zeta}_{t+1}) = \text{Var}(\zeta_{CF,t+1}) + \text{Var}(u_{t+1}).$$

Thus additive noise inflates the measured numerator of $\mathcal{B}_{\text{scale}}$. At the same time, it attenuates correlations and projection coefficients toward zero. The common statement that measurement error makes the behavioral scale conservative therefore requires qualification. It holds under signal attenuation or incomplete-proxy models, but it is reversed under additive classical noise.

This distinction matters because analyst revisions are proxies, not direct observations of marginal-investor beliefs. The empirical sections therefore interpret $\mathcal{B}_{\text{scale}}$, news correlations, forecast R^2 , slope tests, matrix wedges, and horizon restrictions jointly. Table 1.1 summarizes the role of each object: scale measures economic magnitude, news correlations test the directional sign prediction, forecast R^2 measures predictable comovement, and slope and matrix restrictions evaluate progressively sharper versions of the benchmark.

The table also clarifies the paper's empirical strategy. The sharp benchmark makes several exact predictions, but the broader behavioral claim does not require every equality to hold exactly. The central question is whether the evidence points consistently toward a belief-driven component in risk premium news across sign, scale, forecast comovement, and robustness tests.

Table 1.1: Empirical Objects and Their Interpretation

Object	Tested Property	Sharp Benchmark	Partial Arbitrage
Sign Correlation: $\text{Corr}\left(N_{RP}^{(r)}, \zeta_{CF}\right)$	Sign of the Belief Channel	Negative	Negative when behavioral channel survives
Behavioral Scale Ratio: $\mathcal{B}_{\text{scale}}$	Variance Scale of Belief News	Positive, Large	Positive but not a structural share
Forecast Comovement: R_{forecast}^2	Return-Revision Comovement	One	Declines with residual rational variation
Slope Restriction: $e'_1 \Gamma = e'_3 \Gamma$	One-Step Slope Equality	Holds	Need not hold exactly under partial arbitrage
Matrix Wedge: W_{t+1}^{matrix}	Pointwise News Equality	Zero	Reflects residual news, arbitrage, or proxy error
Horizon Restriction: $e'_1 B_h = e'_3 B_h$	Cumulative Slope Equality	Holds at All h	May fail at long horizons under attenuation

Notes: The “Sharp Benchmark” column states predictions under Proposition 1 (full behavioral pricing with $\lambda = 1$ and $N_{RP}^{*,r} = 0$). The “Partial Arbitrage” column gives predictions under Proposition 2 with $\lambda \in (0, 1)$ and possibly nonzero residual rational risk premium news $N_{RP}^{*,r}$. Notation: $N_{RP}^{(r)}$ denotes rational risk premium news; ζ_{CF} is the unexpected component of cash-flow belief bias; $\mathcal{B}_{\text{scale}}$ is the variance ratio in equation (1.10); R_{forecast}^2 is defined in equation (1.11); Λ , Λ_h , and B_h are the present-value matrices defined in Sections 1.2.7 and 1.2.8.

1.3 Data and News Construction

Section 1.2 derives restrictions linking risk premium news to biased cash-flow beliefs. This section constructs the empirical objects used to test those restrictions. I first describe the return-side forecasting system used to construct rational news. I then describe the analyst forecast data, the target-date alignment procedure, and the three behavioral cash-flow news proxies. Finally, I map the analyst revisions into empirical counterparts of η_{CF} and ζ_{CF} , define the news objects, and describe inference.

1.3.1 Returns, Dividends, and the Forecasting System

The empirical return-news decomposition is estimated at the monthly frequency. The return-side data consist of the excess market return, the risk-free rate, the log price-dividend ratio, and the term spread, and dividend/return series used to recover cash-flow news under the Campbell and Ammer (1993) identity. The baseline sample runs from February 1982 through June 2023, giving 497 monthly observations after merging the return data with the analyst-forecast panel; one observation is lost in forming first-differenced innovations, leaving 496 usable news observations.

Let rx_{t+1} denote the excess log market return from month t to month $t + 1$, rf_{t+1} the 1-month risk-free rate, pd_t the log price-dividend ratio, and ts_t the term spread. Excess returns and the 1-month Treasury bill rate are from the Ken French data library (the value-weighted market factor and the risk-free-rate series); the log price-dividend ratio is the Shiller monthly S&P 500 series; and the term spread is the difference between the 10-year and 3-month Treasury yields from FRED. Appendix A.4 gives the variable definitions and data filters.

I represent the rational information set with a linear forecasting system. For each analyst-based belief measure, the baseline state vector is

$$x_t = [rx_t, rf_t, \nu_t^{(B)}, b_t^{(B)}, pd_t, ts_t]', \quad (1.12)$$

where $\nu_t^{(B)}$ is the behavioral cash-flow news proxy, $b_t^{(B)}$ is the belief-level control, and ts_t is the term spread defined above. The belief-level control is included because revisions in beliefs may depend on the current level of expected cash-flow beliefs. Section 1.3.3 defines the three revision proxies and the belief-level measure used in the baseline specification.

The state vector follows

$$x_{t+1} = \alpha + \Gamma x_t + \varepsilon_{t+1}, \quad \mathbb{E}_t^{(r)} \varepsilon_{t+1} = 0. \quad (1.13)$$

This forecasting system serves two roles. First, it defines the rational expectations operator $\mathbb{E}_t^{(r)}$ used to construct the news terms in equation (1.1). Revisions in the forecasts generated by equation (1.13) produce risk-free-rate news and risk premium news. Second, the same system provides the slope coefficients used to test the forecast restriction in equation (1.6): the row of Γ that forecasts excess returns is compared with the row that forecasts the analyst-based belief revision.

The present-value matrix implied by the forecasting system is

$$\Lambda = \rho \Gamma (I - \rho \Gamma)^{-1}, \quad (1.14)$$

where ρ is the Campbell–Shiller log-linearization constant. This matrix maps monthly innovations ε_{t+1} into revisions in present discounted values of future state variables. Risk premium news, $N_{RP,t+1}^{(r)}$, is constructed directly from the forecasting system’s revisions to future excess returns. It is not introduced as the residual from the accounting identity. Cash-flow news, $N_{CF,t+1}^{(r)}$, is then recovered so that the Campbell and Ammer (1993) identity holds. Section 1.3.5 describes the exact news construction and the corresponding empirical test objects.

This subsection defines the return-side forecasting system. The next subsection turns to the belief data. It explains how analyst forecasts are aligned across target dates so that monthly revisions measure changes in cash-flow expectations rather than mechanical changes in forecast horizon.

1.3.2 Analyst Forecasts and Target-Date Alignment

The belief data come from analyst forecasts of firm-level earnings, which I use as proxies for cash-flow expectations. Analyst forecasts are not assumed to equal the beliefs of the marginal investor. They are useful because they provide repeated, dated measurements of subjective expectations for a large panel of firms. I use these forecasts to construct monthly revisions in expected future cash flows and then aggregate those revisions to the market level.

A central measurement problem is that analyst forecasts are indexed by forecast horizon as well as by calendar time. A naive month-to-month revision can compare forecasts for different fiscal targets. For example, a revision from a near-term fiscal-year forecast in one month to a farther fiscal-year forecast in the next month can mechanically reflect horizon rollover rather than a change in beliefs about the same cash-flow object. This problem is especially important in monthly data because the target horizon changes mechanically as fiscal-year-end dates approach.

I address this problem by constructing target-date-aligned revisions. For each firm i , let $F_{i,t}(T)$ denote the analyst forecast at time t for a fixed cash-flow target date T . The aligned behavioral revision is

$$\nu_{i,t+1}^{(B)} \equiv F_{i,t+1}(T) - F_{i,t}(T). \quad (1.15)$$

The target date T is held fixed across the revision window. The revision therefore measures a change in beliefs about the same future cash-flow object, rather than a change in the object being forecast. The superscript (B) denotes an empirical behavioral-belief proxy. It does not mean the revision is measured without error or that analysts are the marginal investors. It means the series is the observable counterpart to the behavioral cash-flow revision in the

theory. This alignment step builds on analyst-expectations measures used in prior work, including Bretscher et al. (2025), De La O and Myers (2021), and Bordalo et al. (2025c), but applies the revision construction to the Campbell and Ammer (1993) news framework.

I aggregate firm-level aligned revisions to the market level using value weights. Let $w_{i,t}$ denote the market-value weight of firm i at time t . For a given belief-revision measure, the market-level revision is

$$\nu_{t+1}^{(B)} = \sum_{i \in \mathcal{I}_t} w_{i,t} \nu_{i,t+1}^{(B)}, \quad (1.16)$$

where $\mathcal{I}_t$ is the set of firms with usable aligned forecasts at time t .⁵

The alignment procedure has two advantages. First, it removes a mechanical source of forecast variation that would otherwise contaminate monthly revisions. Second, it makes the analyst revision comparable to the cash-flow-news objects in the Campbell and Ammer (1993) decomposition, which are revisions in expectations about a fixed stream of future cash flows. The next subsection defines the three aligned behavioral cash-flow news proxies used in the empirical analysis.

1.3.3 Three Behavioral Cash-Flow News Proxies

I construct three market-level behavioral cash-flow news proxies. They use different analyst-forecast inputs and horizons, so they are exposed to different measurement errors. This cross-proxy design is useful because the paper’s main prediction should not depend on one particular analyst measure. I refer to them as the CBR (Composite Belief Revision, both short and long horizons), LTG (Long-Term Growth, long horizon), and STR (Short-Term Revision, short

⁵ Weights are lagged by one month relative to revision dates to limit same-month price feedback into the aggregation. Appendix A.4.5 gives the aggregation details.

Table 1.2: Behavioral Cash-Flow News Proxies

<i>Panel A: Proxy Summary</i>			
Proxy	Raw Forecast Input	Horizon	Alignment
CBR	EPS and LTG Forecasts	Short and Long	Target-Date and Timing
LTG	LTG Forecasts	Long	Timing Only
STR	EPS Forecasts	Short	Target-Date Only

<i>Panel B: Construction Formulas</i>			
Proxy	$h = 1$	$h = 2$	$h \geq 3$ Continuation
CBR	$g_{1,it}$	$g_{2,it}$	$\phi^{h-2} \text{LTG}_{it} + (1 - \phi^{h-2}) \bar{g}_t$
LTG	LTG_{it}	$\phi \text{LTG}_{it} + (1 - \phi) \bar{g}_t$	$\phi^{h-1} \text{LTG}_{it} + (1 - \phi^{h-1}) \bar{g}_t$
STR	$g_{1,it}$	$g_{2,it}$	$\phi^{h-2} g_{2,it}$

Notes: Panel A summarizes the three analyst-based behavioral cash-flow news proxies. CBR follows Bretscher et al. (2025); LTG follows Bordalo et al. (2025c); STR follows De La O and Myers (2021). Panel B gives the per-firm cash-flow growth path $\hat{g}_{i,h}$ used to construct the aligned revision in equation (1.15). Common ingredients: $g_{1,it} = \text{FY1}_{it}/\text{EPS}_{it}^{\text{trail}} - 1$ and $g_{2,it} = \text{FY2}_{it}/\text{FY1}_{it} - 1$; $\bar{g}_t$ is the cross-sectional median long-term-growth forecast; $\phi = 0.445$. All proxies are aggregated to the market level using the value-weighted aggregation in equation (1.16). Appendix A.4.4 gives the exact definitions, filters, and timing conventions.

horizon) proxies. Table 1.2 summarizes the three proxies.

The first proxy, CBR, is a composite aligned revision measure following Bretscher et al. (2025). It combines information from short-horizon analyst earnings forecasts and analysts' long-term-growth forecasts, so it spans both horizons. I use it as the baseline proxy because it is the most comprehensive analyst-based measure of cash-flow belief revisions in the paper.

The second proxy, LTG, uses analysts' long-term-growth forecasts alone, in the tradition of Bordalo et al. (2025c). It captures revisions in persistent growth expectations and helps isolate the long-horizon component of the baseline measure.

The third proxy, STR, follows De La O and Myers (2021) by extracting cash-flow expc-

tations from analyst forecasts at short horizons. It provides a short-horizon-only counterpart to LTG and helps test whether the results are specific to the CBR composite construction.

For each proxy, I construct the firm-level aligned revision using equation (1.15) and aggregate it to the market level using equation (1.16). I denote the resulting market-level revision by $\nu_t^{(B)}$. In the baseline forecasting system in equation (1.12), the third state variable is this revision proxy. I estimate the system separately for the CBR, LTG, and STR revisions. The baseline forecasting system also includes a belief-level control, $b_t^{(B)}$, so that predictable future revisions can depend on the current level of expected cash-flow beliefs.⁶

The three proxies are not interchangeable, and the paper does not require them to be. Short-horizon revisions may capture different belief movements than long-term growth revisions. The empirical question is whether the predicted relation between belief revisions and risk premium news appears consistently across these different measures.

Before taking these proxies to the return-news system, I check that they have the basic time-series properties of revision objects. The target-date-aligned revisions are centered near zero, have moderate volatility, and are much less persistent than revisions that mechanically compare forecasts by horizon label rather than by fixed target date. A simple next-fiscal-year revision that ignores fiscal-year rollover is roughly twice as volatile, substantially more persistent, and nearly uncorrelated with the aligned CBR proxy. The three aligned proxies also share a common component: the first principal component explains 74.4% of their standardized variance. These facts do not prove that analyst revisions equal marginal-investor beliefs, but they show that the aligned series are news-like belief-revision measures rather than mechanical

⁶ In the baseline specifications, $b_t^{(B)}$ is the CBR-based belief level and is held fixed across the CBR, LTG, and STR revision-proxy systems. Holding the level control fixed makes the systems easier to compare: the revision proxy $\nu_t^{(B)}$ varies across specifications, while the state variable controlling for the current level of cash-flow beliefs is held constant. Appendix A.4.4 gives the exact proxy definitions, filters, and aggregation details.

horizon-roll artifacts. Appendices A.4.6 and A.4.9 report the validation diagnostics.

1.3.4 From Analyst Revisions to Belief-Bias News

Section 1.2.2 defines the theoretical objects $\eta_{CF,t+1}$ and $\zeta_{CF,t+1}$: the behavioral-minus-rational cash-flow-news wedge and its rationally unexpected component. This subsection constructs their empirical counterparts from analyst revisions.

For each proxy $j \in \{\text{CBR}, \text{LTG}, \text{STR}\}$, let $\nu_{j,t+1}^{(B)}$ denote the market-level aligned analyst revision constructed from the firm-level revisions in equation (1.15) and the aggregation in equation (1.16). The proxy $\nu_{j,t+1}^{(B)}$ targets behavioral cash-flow news: the subjective revision in expected future cash flows measured from analyst forecasts.

The theory, however, concerns the difference between behavioral and rational cash-flow news. I therefore define the empirical behavioral distortion as

$$\eta_{j,t+1}^{(B)} \equiv \nu_{j,t+1}^{(B)} - N_{CF,t+1}^{(r)}. \quad (1.17)$$

The first term is the analyst-based subjective cash-flow revision. The second term is rational cash-flow news from the return-side forecasting system. Thus $\eta_{j,t+1}^{(B)}$ is the empirical counterpart of the theoretical cash-flow belief-bias news $\eta_{CF,t+1}$.

The news tests use the rationally unexpected component of this wedge. I define

$$\zeta_{j,t+1}^{(B)} \equiv \eta_{j,t+1}^{(B)} - \mathbb{E}_t^{(r)} \eta_{j,t+1}^{(B)}. \quad (1.18)$$

Empirically, $\mathbb{E}_t^{(r)} \eta_{j,t+1}^{(B)}$ is the linear projection of $\eta_{j,t+1}^{(B)}$ on the time- t state vector x_t . Sec-

tion 1.3.5 describes the forecasting system used to construct this projection and the rational news terms.

This step puts the belief proxy on the same footing as the Campbell and Ammer (1993) news terms. Predictable belief distortions may affect prices, but they are not news at $t + 1$. The object compared to risk premium news is the innovation $\zeta_{j,t+1}^{(B)}$.

The superscript (B) denotes an empirical belief-based proxy. It does not mean that analysts are marginal investors, that the proxy is measured without error, or that the proxy captures all valuation-relevant belief distortions. The notation means that $\zeta_{j,t+1}^{(B)}$ is the observable counterpart of $\zeta_{CF,t+1}$ constructed from analyst revisions and rational cash-flow news. When the proxy index j is not needed, I write $\nu_{t+1}^{(B)}$, $\eta_{t+1}^{(B)}$, and $\zeta_{t+1}^{(B)}$.

This construction separates three objects that play different roles in the empirical analysis. The analyst revision $\nu^{(B)}$ measures subjective cash-flow revisions. The distortion $\eta^{(B)}$ subtracts rational cash-flow news to isolate the behavioral-minus-rational wedge. The innovation $\zeta^{(B)}$ is the news object used in the sign, scale, and matrix-wedge tests. Section 1.3.5 describes the construction of $N_{CF}^{(r)}$, $N_{RP}^{(r)}$, and the remaining test objects.

1.3.5 Rational News Construction and Test Objects

The empirical news terms are constructed from the forecasting system in equation (1.13). Let ε_{t+1} denote the innovation to the state vector, and let e_1 , e_2 , and e_3 select excess returns, the risk-free rate, and the behavioral cash-flow news proxy, respectively. The 1-period unexpected excess return is $e_1' \varepsilon_{t+1}$, and the present-value matrix Λ in equation (1.14) maps this innovation into revisions of expected future state variables.

Risk premium news is constructed directly as the rational revision in expected future excess returns:

$$N_{RP,t+1}^{(r)} = e'_1 \Lambda \varepsilon_{t+1}.$$

Risk-free-rate news is constructed analogously:

$$N_{RF,t+1}^{(r)} = e'_2 \Lambda \varepsilon_{t+1}.$$

These definitions follow the return-news decomposition in Campbell and Ammer (1993).⁷ In particular, $N_{RP,t+1}^{(r)}$ is the expected-return-news object generated by the forecasting system; it is not introduced as a residual from the accounting identity.

Cash-flow news is then recovered so that the Campbell and Ammer (1993) identity holds exactly:

$$N_{CF,t+1}^{(r)} = e'_1 \varepsilon_{t+1} + N_{RF,t+1}^{(r)} + N_{RP,t+1}^{(r)}. \quad (1.19)$$

Equivalently,

$$N_{CF,t+1}^{(r)} = (e'_1 + e'_2 \Lambda + e'_1 \Lambda) \varepsilon_{t+1}.$$

This construction keeps the expected-return-news object tied to revisions in future excess returns while ensuring that the return identity in equation (1.1) is satisfied in the data.

For each belief proxy $j \in \{\text{CBR}, \text{LTG}, \text{STR}\}$, I combine this rational cash-flow news with the analyst revision $\nu_{j,t+1}^{(B)}$ to construct $\eta_{j,t+1}^{(B)}$ and $\zeta_{j,t+1}^{(B)}$ via equations (1.17)–(1.18), the empirical counterparts of the belief-bias objects in Section 1.2.2.

⁷ This convention treats risk-free-rate news symmetrically with risk premium news, as a present-value revision in future rates. Appendix A.5.4 discusses the timing convention; Appendix A.5.6 verifies that the Campbell and Ammer (1993) identity holds under the implementation.

The main empirical objects follow directly from the theory. Matching the sign prediction in Proposition 1, I compute

$$\text{Corr}\left(N_{RP,t+1}^{(r)}, \zeta_{j,t+1}^{(B)}\right).$$

Matching the scale statistic in equation (1.10), I compute

$$\mathcal{B}_{\text{scale},j} = \frac{\text{Var}\left(\zeta_{j,t+1}^{(B)}\right)}{\text{Var}\left(N_{RP,t+1}^{(r)}\right)}.$$

Matching the forecast-comovement statistic in equation (1.11), I compute

$$R_{\text{forecast},j}^2 = \text{Corr}\left(\widehat{\mathbb{E}}_t rx_{t+1}, \widehat{\mathbb{E}}_t \nu_{j,t+1}^{(B)}\right)^2.$$

The slope test implements equation (1.6) by comparing the return-forecasting row and the belief-revision-forecasting row of the VAR,

$$e_1' \Gamma = e_3' \Gamma.$$

Finally, matching the matrix wedge in equation (1.8), I compute

$$W_{j,t+1}^{\text{matrix}} = N_{RP,t+1}^{(r)} + \zeta_{j,t+1}^{(B)}.$$

These objects correspond to the tests summarized in Table 1.1. The sign correlation asks whether belief-bias news moves in the direction predicted by Proposition 1 and its partial-arbitrage extension. $\mathcal{B}_{\text{scale}}$ asks whether the belief-bias object is economically large. Forecast R^2 asks whether predictable returns and predictable belief revisions comove. The slope and

matrix restrictions evaluate progressively sharper implications of the full behavioral-pricing benchmark in equation (1.2).

1.3.6 Inference and Pre-Specified Interpretation

Persistent state variables, generated news terms, and overlapping long-horizon returns make standard asymptotic inference unreliable in this setting. I therefore use bootstrap inference as the primary basis for the empirical tests. The bootstrap preserves the time-series dependence in the monthly data and, for restrictions such as the slope test, imposes the null hypothesis before simulating the sampling distribution of the statistic.

For baseline sampling uncertainty, I use a stationary block bootstrap with 5,000 replications and a mean block length of 12 months. In each replication, I reconstruct the relevant news terms, belief-bias objects, and test statistics. This procedure gives confidence intervals for the sign correlations, scale ratios, forecast-comovement statistics, matrix wedges, and cross-sectional summaries.

For the slope restriction in equation (1.6), I use a null-imposed bootstrap. I estimate the restricted forecasting system under

$$e_1' \Gamma = e_3' \Gamma,$$

simulate bootstrap samples from that restricted system, and recompute the Wald statistic in each sample. The resulting p -value asks whether the observed slope discrepancy is large relative to the distribution generated by the sharp full-benchmark forecast restriction. This is the relevant inference for the slope test because the restriction is imposed on a persistent forecasting system and evaluated using generated test objects.

For horizon tests, I use the same principle: inference is based on the sampling distribution of the relevant finite-horizon statistic rather than on large-sample approximations alone. Section 1.4 introduces the direct horizon tests, and Appendix A.5.8 gives the implementation details.

The empirical interpretation distinguishes evidence for the sharp full behavioral-pricing benchmark from evidence for the broader belief-driven channel. Slope non-rejection supports the full benchmark's forecast restriction. Slope rejection, by itself, does not reject the broader channel because partial arbitrage, residual rational expected-return variation, rate-belief bias, and proxy error can all break exact equality. Conversely, evidence on sign, scale, forecast comovement, price-inference robustness, and cross-sectional patterns speaks to whether a belief-driven component is present more broadly. Section 1.4 turns to the main time-series evidence; Section 1.6 addresses the robustness exercises and Section 1.7 the cross-sectional evidence.

1.4 Main Time-Series Evidence

The theory predicts that risk premium news should contain a belief-driven component when prices reflect biased cash-flow beliefs. This section tests the core time-series implications. I begin with the main aggregate evidence on sign, scale, and forecast comovement. I then study the 1-step slope restriction implied by the full behavioral-pricing benchmark and direct horizon tests of the same forecast relation. The sharper long-horizon matrix and convergence diagnostics are left to Section 1.5; this section asks whether the central behavioral predictions appear in the aggregate time series.

1.4.1 Sign, Scale, Forecast Comovement, and One-Step Slope Evidence

Table 1.3 reports the main time-series evidence for the three behavioral cash-flow news proxies. The table combines the directional news correlation, the behavioral scale ratio, the signed forecast correlation, the forecast-comovement R^2 , and the bootstrap p -value for the 1-step slope restriction. These are the core empirical objects defined in Sections 1.2.9 and 1.3.5.

Table 1.3: Baseline Time-Series Evidence

Proxy	News Correlation	$\mathcal{B}_{\text{scale}}$	Forecast Correlation	Forecast R^2	Slope p_{boot}
CBR	−0.1068	0.3415	0.5137	0.2639	0.6573
LTG	−0.1070	0.2800	0.5113	0.2615	0.3490
STR	−0.0997	0.3350	0.4569	0.2087	0.6725

Notes: The table reports baseline aggregate time-series evidence for the composite belief revision proxy (CBR), the long-term-growth proxy (LTG), and the short-term revision proxy (STR). “News Correlation” is the correlation between rational risk premium news $N_{RP}^{(r)}$ and empirical belief-bias news $\zeta^{(B)}$. $\mathcal{B}_{\text{scale}}$ is the variance ratio in equation (1.10). “Forecast Correlation” is the signed sample correlation between fitted return forecasts and fitted analyst-revision forecasts. Forecast R^2 is the squared sample forecast-comovement statistic, so it equals the square of the signed forecast correlation. Slope p_{boot} is the null-imposed bootstrap p -value for the 1-step restriction $e_1'\Gamma = e_3'\Gamma$. Population forecast R^2 (computed via the discrete Lyapunov equation) is approximately 0.30 across proxies and corresponds to the values cited in the abstract and introduction; the full sample, population, and bias-corrected forecast R^2 values are reported in Appendix A.6.3.

The first column gives the directional test. For all three proxies, belief-bias news is negatively correlated with risk premium news. The correlations are close to -0.10 : -0.1068 for CBR, -0.1070 for LTG, and -0.0997 for STR. This is the sign predicted by Proposition 1 and its partial-arbitrage extension. Unexpected cash-flow optimism raises prices and lowers rational expected future excess returns, so the belief-bias innovation should move negatively with risk premium news.

The second column shows that the measured belief-bias object is economically meaning-

ful in scale. The behavioral-to-risk premium news variance ratio ranges from 0.2800 to 0.3415, so the measured belief-bias news has roughly one-third of the variance of rational risk premium news. This magnitude is difficult to reconcile with the view that analyst-based belief revisions are only small, idiosyncratic measurement noise. The statistic should be interpreted narrowly. By definition,

$$\mathcal{B}_{\text{scale},j} = \frac{\text{Var}\left(\zeta_{j,t+1}^{(B)}\right)}{\text{Var}\left(N_{RP,t+1}^{(r)}\right)}.$$

This ratio compares the variance of the measured belief-bias object with the variance of risk premium news. It is not a causal share and does not impose an orthogonal decomposition of $N_{RP}^{(r)}$. The empirical claim is therefore a scale claim: the measured belief component is economically meaningful relative to the object that standard decompositions call risk premium news.

The next two columns report forecast comovement. The signed forecast correlations are positive for all three proxies, between 0.4569 and 0.5137, with sample R^2 values from 0.2087 to 0.2639. The forecast statistic has a different sign convention than the news correlation: Proposition 1 predicts a negative relation between realized risk premium news and unexpected belief-bias news, while equation (1.5) predicts positive comovement between expected returns and expected belief revisions. The positive forecast correlations therefore support the forecast-side implication of the full behavioral-pricing benchmark.

The last column reports the 1-step slope test. The null-imposed bootstrap p -values are 0.6573, 0.3490, and 0.6725. Thus the 1-step forecast restriction is not rejected for any proxy. This result supports the full behavioral-pricing benchmark's forecast implication, while the sign and scale evidence support a belief-driven component in risk premium news.

1.4.2 Direct Horizon Slope Tests

The 1-step slope test uses the forecasting system in equation (1.13). Section 1.2.8 shows that, inside this system, the 1-step restriction mechanically implies corresponding cumulative restrictions at finite horizons. I also test these restrictions directly. The direct horizon tests ask whether discounted cumulative return forecasts and discounted cumulative belief-revision forecasts line up at horizon h , without relying only on the VAR's iterated implication.

Table 1.4 reports bootstrap p -values for horizons $h = 1, 12, 24, 36$, and 60 months. The $h = 1$ test is a useful validation check: the direct horizon test reproduces the 1-step VAR slope test almost exactly. The medium-horizon tests then ask whether the same forecast relation survives at 1-, 2-, and 3-year horizons. The 5-year test connects the framework to the long-horizon return-predictability evidence emphasized in the behavioral literature.

Table 1.4: Direct Horizon Slope Tests

Proxy	$h = 1$	$h = 12$	$h = 24$	$h = 36$	$h = 60$
CBR	0.6573	0.6713	0.6081	0.3547	0.0006
LTG	0.3490	0.4308	0.2902	0.3638	0.0008
STR	0.6725	0.6109	0.5869	0.2869	0.0006

Notes: The table reports bootstrap p -values for direct horizon slope tests. Each test compares the discounted cumulative forecast relation for excess returns with the corresponding forecast relation for the behavioral cash-flow news proxy at horizon h . The $h = 1$ column corresponds to the 1-step restriction $e'_1\Gamma = e'_3\Gamma$. Bootstrap inference follows the procedure described in Section 1.3.6.

The 1-step and medium-horizon evidence supports the forecast side of the benchmark. At $h = 1$, the bootstrap p -values are 0.6573, 0.3490, and 0.6725, matching the 1-step VAR slope test. At $h = 12$, $h = 24$, and $h = 36$, the restrictions are not rejected for any of the three proxies. Thus the forecast relation in equation (1.5) survives beyond the 1-month horizon.

The 5-year horizon gives a different result. Direct cumulative-forecast equality at $h = 60$ is rejected at the 1% level for all three proxies, with bootstrap p -values of 0.0006 to 0.0008. This is an important caveat. It means that direct 5-year cumulative slopes do not satisfy the exact equality implied by the sharp benchmark. It does not overturn the negative news correlations, the economically meaningful scale, the positive forecast comovement, or the 1-step and medium-horizon non-rejections documented above.

The horizon evidence therefore sharpens the interpretation. The data support the full benchmark's forecast restriction at 1-step and medium horizons, while rejecting exact cumulative equality at five years. Section 1.5 studies whether this long-horizon rejection reflects convergence and pointwise-news issues rather than the absence of a belief-driven component in risk premium news.

1.4.3 Forecast-Side Support and Long-Horizon Limits

The aggregate time-series evidence supports the main behavioral predictions. Across CBR, LTG, and STR, belief-bias news is negatively correlated with risk premium news, the measured belief-bias object has economically meaningful scale, and predictable returns comove positively with predictable belief revisions. The 1-step slope restriction is also not rejected for any proxy under the null-imposed bootstrap. These findings line up with the forecast-side implications of the full behavioral-pricing benchmark and with the broader partial-arbitrage prediction that risk premium news contains a surviving belief-driven component.

The evidence is not uniform across all horizons. Direct horizon tests do not reject the cumulative forecast restriction at 1-, 2-, or 3-year horizons, but they reject sharply at five years. This pattern is informative. It shows that the baseline behavioral channel is visible in

the aggregate time series, while exact long-horizon cumulative equality is too strong. The rejection at $h = 60$ should therefore be read as a long-horizon caveat, not as evidence that the sign, scale, or forecast-comovement results disappear.

The next section studies this caveat directly. Section 1.5 separates the 1-step and medium-horizon forecast evidence from the stronger pointwise long-horizon news restriction. It asks whether the 5-year rejection reflects convergence, truncation, and matrix-wedge issues rather than the absence of a belief-driven component in risk premium news.

1.5 Limits of the Sharp Benchmark

Section 1.4 shows that the main time-series evidence lines up with the behavioral channel: belief-bias news is negatively related to risk premium news, the measured belief-bias object is large relative to risk premium news, predictable returns comove positively with predictable belief revisions, and the 1-step and medium-horizon slope restrictions are not rejected. The same section also shows that direct cumulative equality fails sharply at five years. This section studies that caveat. The goal is to distinguish evidence against the strongest pointwise benchmark from evidence against the broader claim that risk premium news contains a belief-driven component.

1.5.1 Stronger Implications of the Sharp Benchmark

The full behavioral-pricing benchmark imposes more than the sign and forecast-comovement predictions tested in Section 1.4. It also implies pointwise equality between constructed long-horizon news objects: under Proposition 1 (equation (1.2)), $N_{RP,t+1}^{(r)}$ should equal $-\zeta_{CF,t+1}$

innovation-by-innovation, implemented in the linear forecasting system through the matrix condition in equation (1.7). That is a strong restriction. Under partial arbitrage, residual rational expected-return news, rate-belief bias, or proxy error, the sign and forecast evidence can survive even though exact pointwise equality fails. The matrix-wedge and horizon diagnostics below measure this distance from the sharp benchmark.

1.5.2 Matrix Wedges

The sharp pointwise benchmark can be evaluated through the matrix wedge defined in equation (1.8). For each proxy $j \in \{\text{CBR}, \text{LTG}, \text{STR}\}$, I compute

$$W_{j,t+1}^{\text{matrix}} = N_{RP,t+1}^{(r)} + \zeta_{j,t+1}^{(B)}.$$

Under the full behavioral-pricing benchmark, this object is zero innovation-by-innovation. A nonzero wedge therefore measures the distance between the constructed risk premium news object and the negative of measured belief-bias news.

Table 1.5 reports the variance of the matrix wedge relative to the variance of risk premium news. The unrestricted wedge ratios are economically large: 1.2167 for CBR, 1.1668 for LTG, and 1.2196 for STR. The bootstrap confidence intervals are bounded away from zero for all three proxies. Thus exact pointwise equality does not hold in the data.

The slope-restricted columns ask whether imposing the 1-step forecast restriction also brings the long-horizon news objects closer to pointwise equality. It does not. If slope equality mechanically delivered long-horizon pointwise equality, the restricted ratios would be close to zero. Instead, the slope-restricted ratios are larger than the unrestricted ratios for all three

Table 1.5: Matrix-Wedge Diagnostics

Proxy	Unrestricted Ratio	Slope-Restricted Ratio
CBR	1.2167 [0.7648, 2.6726]	2.9832 [1.5709, 6.7169]
LTG	1.1668 [0.7167, 2.5142]	2.3867 [0.8104, 6.9808]
STR	1.2196 [0.7706, 2.6593]	3.1903 [1.6537, 7.0674]

Notes: The table reports matrix-wedge variance ratios for the three behavioral cash-flow news proxies. The unrestricted ratio is $\text{Var}(W_j^{\text{matrix}})/\text{Var}(N_{RP}^{(r)})$. The slope-restricted ratio recomputes the wedge after imposing the 1-step forecast restriction $e_1'\Gamma = e_3'\Gamma$. Brackets below each point estimate report the 95% stationary-block bootstrap confidence interval, computed using the procedure described in Section 1.3.6.

proxies: 2.9832 for CBR, 2.3867 for LTG, and 3.1903 for STR. The confidence intervals overlap, so I do not interpret the increase as a precisely estimated incremental effect. The robust conclusion is narrower: the 1-step slope restriction and the pointwise long-horizon news equality are distinct restrictions.

The distinction is intuitive. The 1-step slope restriction operates in forecast space: it asks whether the same state variables forecast next-period returns and next-period belief revisions with the same coefficients. The matrix wedge operates in innovation-level present-value space: it asks whether each shock produces risk premium news exactly equal to negative measured belief-bias news. The first restriction need not repair the second. In these data, imposing the 1-step slope restriction also reverses the innovation-level comovement between $N_{RP}^{(r)}$ and $-\zeta^{(B)}$, from about +0.10 in the unrestricted system to about −0.40 when the slope restriction is imposed, reinforcing that slope compatibility does not mechanically imply pointwise news equality.⁸

⁸ Innovation-level $\text{Corr}(N_{RP}^{(r)}, -\zeta^{(B)})$ for CBR, LTG, STR is +0.107, +0.107, +0.100 unrestricted, and −0.379, −0.427, −0.413 slope-restricted; bootstrap distributions in Appendix A.6.5.

The partial-arbitrage model of Section 1.2.4 gives the wedge an economic interpretation:

$$W_{t+1}^{\text{matrix}} = N_{RP,t+1}^{*,r} + (1 - \lambda) \zeta_{CF,t+1}.$$

The first term is residual rational risk premium news. The second term is the part of the full behavioral belief-bias component that does not pass through to risk premium news when arbitrage attenuates the behavioral wedge. Proxy error, omitted rate-belief bias, and misspecification in the present-value forecasting system can also contribute to the empirical wedge. Thus the wedge measures distance from the full pointwise benchmark, not the absence of biased beliefs.

The correct reading is therefore measured. The aggregate time-series evidence supports the belief-driven channel through sign, scale, forecast comovement, and 1-step and medium-horizon slope restrictions. The matrix-wedge evidence rejects the stronger claim that measured belief-bias news and rational risk premium news are pointwise equal innovation-by-innovation. These facts are consistent with partial arbitrage, residual rational expected-return news, rate-belief bias, and proxy error.

1.5.3 Horizon Diagnostics

The 5-year rejection in Section 1.4.2 is a real failure of exact cumulative forecast equality. It does not, however, imply that the belief-driven component disappears at longer horizons. The horizon diagnostics ask a narrower question: how do finite-horizon objects behave as they move from the 1-step forecast restriction toward the infinite-horizon Campbell and Ammer (1993) decomposition?

Section 1.2.8 shows that finite-horizon present-value objects replace the infinite-horizon matrix Λ with

$$\Lambda_h = \sum_{j=1}^h \rho^j \Gamma^j.$$

The difference between Λ_h and Λ is the omitted present-value tail,

$$\Lambda - \Lambda_h = \rho^{h+1} \Gamma^{h+1} (I - \rho \Gamma)^{-1}.$$

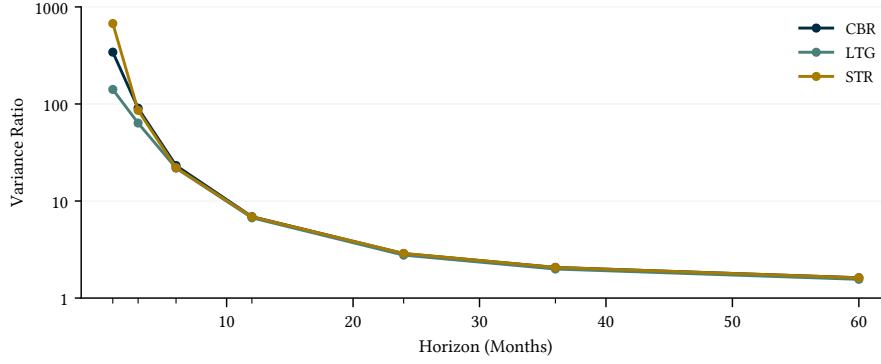
Thus a finite-horizon diagnostic is not the same object as the infinite-horizon matrix condition. A 5-year horizon is economically meaningful, but it is still a truncated present-value object.

Figure 1.1 plots three finite-horizon diagnostics. The 5-year rejection is not accompanied by a discrete breakdown of the horizon profile. The left panel shows that the finite-horizon wedge ratios converge toward the infinite-horizon matrix-wedge values from Table 1.5 (dashed reference lines): finite-horizon objects need not satisfy pointwise equality at every h , but they should converge toward the infinite-horizon object as the present-value tail shrinks (Section 1.2.8). The middle and right panels show that the correlation and forecast-comovement diagnostics are strongest at short and medium horizons and weaken by five years. The weakening does not reverse the basic sign and scale results from Section 1.4.1. The 5-year local-projection rejection therefore isolates a specific failure of exact cumulative slope equality, not evidence that the risk premium news object is unrelated to belief-bias news.

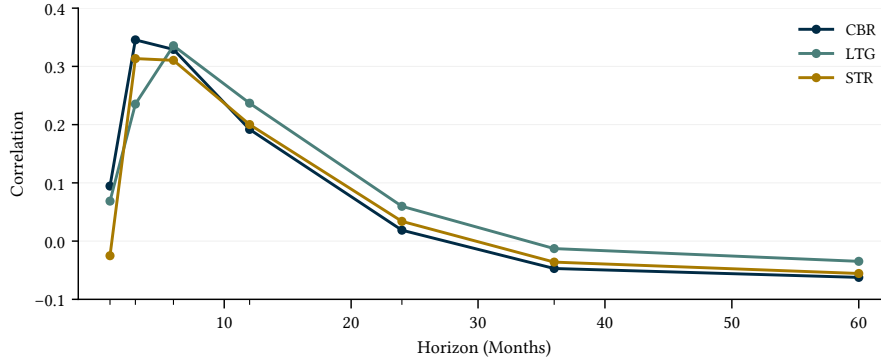
This distinction matters for connecting the paper to long-horizon behavioral predictability. Five-year horizons are central in the La Porta (1996) and Lakonishok et al. (1994) evidence on expectation errors and subsequent returns. The evidence here is consistent with that long-horizon perspective, but the object is different. I ask whether belief revisions enter

Figure 1.1: Horizon Diagnostics

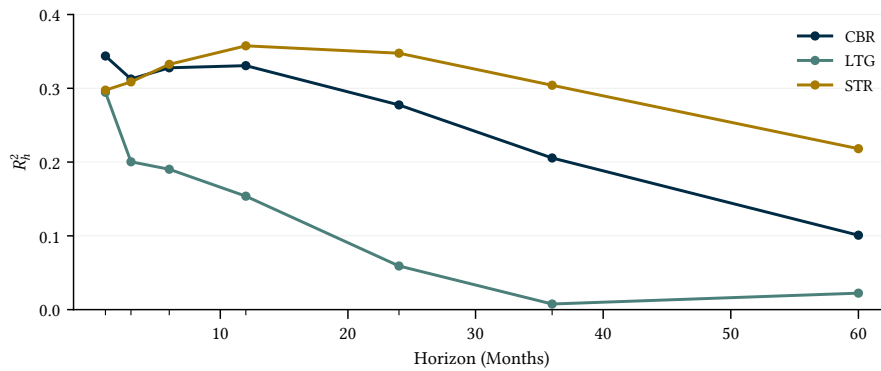
Panel A: Matrix-Wedge Ratio (W_h)



Panel B: News-Innovation Correlation (ρ_h)



Panel C: Forecast Comovement (R_h^2)



Notes: This figure reports three finite-horizon diagnostics for the CBR, LTG, and STR proxies across horizons $h \in \{1, 3, 6, 12, 24, 36, 60\}$ months. Panel A reports the matrix-wedge variance ratio W_h on a log scale, the finite-horizon analogue of the matrix wedge defined in equation (1.8). Panel B reports the innovation-level news correlation ρ_h between rational risk premium news and the negative of empirical belief-bias news. Panel C reports the forecast-comovement statistic R_h^2 , the finite-horizon analogue of equation (1.11). Finite-horizon objects use the truncated present-value matrix Λ_h in place of Λ .

the Campbell and Ammer (1993) risk premium news component, not whether 5-year returns are perfectly explained by expectation errors. The horizon diagnostics show that the belief-driven channel is visible in the return-news system, while exact 5-year cumulative equality is too strong.

The horizon evidence therefore complements the matrix-wedge evidence. Section 1.4 shows strong sign, scale, forecast-comovement, and medium-horizon slope evidence. Table 1.5 shows that pointwise long-horizon equality fails. The horizon diagnostics explain how both facts can hold at once: finite-horizon restrictions can reject, especially at five years, even when the aggregate time-series evidence continues to point toward a belief-driven component in risk premium news.

1.5.4 Forecast Evidence and Pointwise Limits

The matrix and horizon diagnostics clarify the limits of the sharp benchmark. Section 1.4 shows that the main time-series evidence points toward a belief-driven component in risk premium news: the sign is negative, the scale is economically meaningful, forecast comovement is positive, and the 1-step and medium-horizon slope restrictions are not rejected. The evidence in this section shows that a stronger claim does not hold. Measured belief-bias news and rational risk premium news are not pointwise equal innovation-by-innovation, and direct cumulative equality fails at five years.

These findings are not contradictory. The forecast-side restrictions ask whether predictable returns and predictable belief revisions move together. The matrix condition asks whether the full present-value news objects line up shock by shock. The horizon tests ask how finite-horizon restrictions behave as they approach the infinite-horizon Campbell and Ammer

(1993) decomposition. The data support the first set of implications more strongly than the latter two. This is exactly the pattern one would expect if risk premium news contains a belief-driven component, but also contains residual rational expected-return news, partial-arbitrage attenuation, rate-belief components, or proxy error.

The preferred interpretation is therefore neither the knife-edge full benchmark nor a purely rational benchmark. The sharp identity in Proposition 1 is useful because it gives clear sign, forecast, and pointwise predictions. The evidence supports the sign and forecast predictions, but rejects the strongest pointwise equality. The partial-arbitrage framework in Proposition 2 provides the natural reading: belief-bias news survives in risk premium news, but it does not exhaust the object.

The next section, Section 1.6, studies the main measurement and alternative-explanation concerns directly. It asks whether the belief-news relation is driven by proxy error, price-inference contamination, reliability problems, or adversarial residual rational news. Those exercises evaluate whether the belief-driven component documented above survives more demanding robustness tests.

1.6 Robustness and Alternative Explanations

The evidence above treats analyst revisions as empirical proxies for belief revisions. This section asks whether the results could be driven by proxy noise, analysts inferring information from prices, or omitted rational risk premium news that happens to comove with the belief proxy. The goal is not to show that the analyst measures are perfect, but to ask whether the main belief-news relation survives the most direct alternative explanations.

1.6.1 Measurement Error and Reliability

Analyst revisions are noisy proxies for belief revisions. They may contain idiosyncratic analyst noise, incomplete coverage of the beliefs that matter for prices, or components that are not part of marginal-investor expectations. The empirical interpretation therefore depends on how proxy error affects the objects in Section 1.4. Section 1.2.9 introduced the measurement-error taxonomy used here. Measurement error does not have one universal effect. Additive classical noise inflates the variance of the measured belief-bias object, while attenuating correlations and projection coefficients. Signal attenuation and incomplete-proxy models work differently: they can make the measured proxy understate the magnitude of the underlying belief component.

This distinction is important for interpreting $\mathcal{B}_{\text{scale}}$. The scale statistic is not automatically conservative. If a generic observed proxy is

$$\tilde{\zeta}_{t+1} = \zeta_{CF,t+1} + u_{t+1},$$

with u_{t+1} additive classical noise, then the measured numerator of $\mathcal{B}_{\text{scale}}$ includes noise variance. The empirical proxy $\zeta_{t+1}^{(B)}$ in this paper is one such measured object; the formula above describes the abstract measurement-error model of Section 1.2.9, not a literal decomposition of $\zeta^{(B)}$ into known signal and noise components. In that case, measurement error can raise the measured scale ratio. But the same noise pushes news correlations, forecast correlations, and projection slopes toward zero. Thus the sign evidence is harder, not easier, to obtain under additive noise.

The reliability evidence addresses this concern in two ways. First, the results are not based

on a single analyst measure. The negative news correlation, positive forecast comovement, and slope non-rejection appear across CBR, LTG, and STR, even though the three proxies use different forecast inputs and horizons. Second, cross-proxy reliability diagnostics show that the baseline belief measure contains a substantial common component rather than only idiosyncratic analyst noise. For the CBR proxy, the cross-proxy reliability lower bound is 0.865, indicating a substantial common component across belief measures.⁹

The main lesson is that measurement error changes the interpretation of magnitudes, but it does not naturally explain the sign pattern. Additive noise can inflate $\mathcal{B}_{\text{scale}}$, so I do not treat the scale ratio as a structural share. But additive noise would also attenuate the negative news correlation and the positive forecast correlation. The fact that the predicted signs appear across three different proxies therefore weighs against the view that the results are driven only by idiosyncratic analyst noise.

I use the reliability corrections as supporting evidence, not as the main identification argument. They require assumptions about the measurement-error structure, and different structures imply different corrections. The central empirical claim is more modest: the measured belief-bias proxies are sufficiently reliable to produce a consistent sign, scale, and forecast-comovement pattern, and the remaining subsections test whether that pattern is explained by specific alternative channels.

⁹ The signal-attenuation exercise implies a critical reliability threshold of 0.584: above this threshold, the estimated behavioral signal remains economically meaningful after correcting for attenuation. Appendix A.7 reports the full reliability diagnostics and measurement-error corrections.

1.6.2 Price-Inference Robustness

A second concern is that analyst revisions may partly reflect information inferred from prices. If analysts update their forecasts after observing price movements, the belief proxy could mechanically inherit information from the same valuation movements used to construct risk premium news. This would weaken the interpretation of the negative news correlation as evidence of biased cash-flow beliefs.

I test this concern by removing the component of the belief proxy that is linearly related to the contemporaneous price-dividend ratio. For each proxy j , I form a residualized revision that subtracts a fraction ϕ of the projection on pd_t , with $\phi \in [0, 1]$. The case $\phi = 0$ is the baseline. The case $\phi = 1$ removes the full contemporaneous price-dividend projection. If the main results are driven by analysts mechanically inferring beliefs from prices through this channel, the sign, scale, or slope evidence should weaken as ϕ increases.

The results are essentially unchanged. Table 1.6 reports baseline ($\phi = 0$) and fully orthogonalized ($\phi = 1$) values for all three proxies. For CBR, moving from $\phi = 0$ to $\phi = 1$ changes $\mathcal{B}_{\text{scale}}$ from 0.3415 to 0.3413, changes the news correlation from -0.1068 to -0.1058 , and changes the slope-test bootstrap p -value from 0.6657 to 0.6927. The LTG and STR proxies show the same flat pattern. Removing the entire contemporaneous price-dividend projection therefore leaves the core sign, scale, and slope evidence intact.

This exercise does not rule out every form of price-related information in analyst forecasts. Analysts may use prices together with other information, and price inference may occur through nonlinear or lagged channels. The test instead rules out a specific and direct concern: that the measured belief-news relation is driven by the contemporaneous price-dividend com-

Table 1.6: Price-Inference Robustness

Proxy	ϕ	$\mathcal{B}_{\text{scale}}$	News Correlation	W_{obs}	Slope p_{boot}
CBR	0	0.3415	−0.1068	8.9176	0.6657
CBR	1	0.3413	−0.1058	8.3155	0.6927
LTG	0	0.2800	−0.1070	15.7595	0.3440
LTG	1	0.2799	−0.1064	15.3623	0.3570
STR	0	0.3350	−0.0997	9.0343	0.6802
STR	1	0.3349	−0.0991	8.7276	0.6932

Notes: The table reports baseline ($\phi = 0$) and fully orthogonalized ($\phi = 1$) results from the price-inference exercise. At $\phi = 1$, the contemporaneous price-dividend projection is fully removed from the behavioral cash-flow news proxy. $\mathcal{B}_{\text{scale}}$ is the variance ratio in equation (1.10), “News Correlation” is the correlation between $N_{RP}^{(r)}$ and $\zeta^{(B)}$, W_{obs} is the observed slope-test Wald statistic, and slope p_{boot} is the null-imposed bootstrap p -value. Bootstrap inference uses the procedure described in Section 1.3.6 with $B = 2,000$ replications. Slight differences from the Table 1.3 baseline reflect bootstrap sampling noise.

ponent of analyst revisions. In the data, that component is empirically small.

1.6.3 Adversarial Residuals

A third concern is omitted rational risk premium news. Equation (1.3) allows a residual rational component, $N_{RP,t+1}^{*,r}$, to coexist with the belief-driven component. If this residual component is correlated with measured belief-bias news, the observed news correlation may combine behavioral belief news with rational expected-return news. I therefore ask how large an omitted residual component would need to be to weaken or erase the baseline relation.

I consider two adversarial exercises. The first adds generic residual variation that is orthogonal to the measured belief-bias innovation. Let

$$N_{RP,t+1}^{\text{adv}} = N_{RP,t+1}^{(r)} + \kappa w_{t+1},$$

where w_{t+1} is orthogonal to the innovation vector used in the VAR and normalized to have unit variance. This exercise asks how much unrelated residual news must be added before the correlation between N_{RP}^{adv} and $\zeta^{(B)}$ becomes small.

The answer is that generic residual variation must be large. To cut the absolute value of the baseline correlation in half, the required residual volatility is

$$\kappa^*/\sigma(N_{RP}) = \sqrt{3} \approx 1.73,$$

independent of the proxy. To push the absolute correlation below 0.025, the required residual volatility is about four times the volatility of measured risk premium news: 4.15 for CBR, 4.16 for LTG, and 3.86 for STR. Generic orthogonal-residual news therefore cannot easily erase the relation unless it is several times as volatile as the measured object.

The second exercise allows a more targeted adversary. Instead of being orthogonal to the measured belief-bias innovation, the residual component is allowed to be aligned with it while remaining orthogonal to the state variables. This is a much harder test because a residual that is deliberately correlated with $\zeta^{(B)}$ can directly offset the baseline relation. In this case, only about 1% of the variance of measured risk premium news is enough to drive the correlation to zero: 1.14% for CBR, 1.14% for LTG, and 0.99% for STR.

These two exercises give a useful boundary. The belief-news relation is robust to generic orthogonal dilution, while a belief-aligned residual is the natural structured alternative that remains hardest to rule out. This is not a defect of the test. It states exactly what the empirical design can and cannot rule out. The evidence rejects the idea that a small amount of unrelated rational residual news mechanically explains the baseline relation. It cannot reject a structured rational-news component that is specifically aligned against measured belief-bias news.

Appendices A.8.2 and A.8.3 give the closed-form calculations and bootstrap intervals for Versions A and B, respectively. Appendix A.8.4 reports an adversarial frontier that interpolates between the two cases. The frontier shows how quickly the required residual variance falls as the adversary becomes more correlated with the belief-bias innovation.

1.6.4 Robustness Evidence and Remaining Alternatives

The robustness evidence rules out the most direct mechanical versions of several alternative explanations. The results are not driven by one analyst proxy: the sign, scale, forecast-comovement, and slope evidence appears across CBR, LTG, and STR. They are not driven by the contemporaneous price-dividend component of analyst revisions: removing the full pd_t -projection leaves the main estimates essentially unchanged. They are also not easily erased by generic omitted rational news: an orthogonal residual component must be several times as volatile as measured risk premium news to drive the baseline correlation close to zero. Appendix A.6.8 shows that the main sign and slope evidence is stable across alternative state vectors and lag orders, although the scale statistic is more sensitive to the news-construction technology.

The evidence does not rule out every structured alternative. In particular, a residual rational-news component that is specifically aligned with measured belief-bias news is able to weaken or erase the correlation with a much smaller variance budget. This is why the paper's claim is a component claim rather than an exclusivity claim. The evidence shows that risk premium news contains a belief-driven component; it does not show that risk premium news contains only belief-driven news.

Under partial arbitrage, Proposition 2 gives:

$$N_{RP,t+1}^{\lambda,r} = N_{RP,t+1}^{*,r} - \lambda \zeta_{CF,t+1}.$$

The empirical exercises support the presence of the second term, while allowing the first term to remain important. Appendix A.8.5 reports a calibration of the partial-arbitrage loading. The point estimates are around 0.17–0.20, but the confidence intervals are wide, so I interpret them as calibration evidence rather than as precise structural estimates.

Taken together, Sections 1.4–1.6 support a qualified interpretation. The aggregate time-series evidence points consistently toward a belief-driven component in risk premium news. The matrix and horizon diagnostics show that exact pointwise equality is too strong. The robustness exercises show that the relation is not a mechanical artifact of proxy noise, contemporaneous price inference, or small generic residual news. The remaining interpretation is the one implied by the broader theory: biased cash-flow beliefs are an empirically important component of risk premium news, alongside residual rational news, partial arbitrage, and measurement error.

The next section, Section 1.7, asks whether the same pattern appears beyond the aggregate time series. It studies cross-sectional evidence and decomposes the market-level belief proxy into common belief movements and size-weighted reallocation across firms.

1.7 Cross-Sectional and Mechanism Evidence

The evidence so far is aggregate. This section asks whether the same belief-news relation appears beyond the market-level time series and what kind of belief movement drives the

aggregate proxy. The goal is not to replace the aggregate evidence in Sections 1.4–1.6. It is to show whether the same sign pattern is visible across portfolios and whether the market-level belief revision reflects common belief movements, size-weighted reallocation across firms, or both.

1.7.1 Portfolio-Level and Sorting Evidence

I first examine the belief-news relation across the 25 Fama–French size-by-book-to-market portfolios (FF25) (Fama and French 1993). For each portfolio, I replace the market excess return in the baseline forecasting system with the portfolio excess return and re-estimate the VAR. The remaining state variables – the risk-free rate, aggregate belief revision, aggregate belief level, market price-dividend ratio, and term spread – remain at the aggregate level. Thus portfolio-level risk premium news varies across cells, while the belief-bias news series is the same aggregate $\zeta^{(B)}$ used in the market-level tests. The aggregate belief proxy is built from Institutional Brokers’ Estimate System (IBES) analyst forecasts; portfolio-level decompositions are not used in the main analysis because IBES coverage is too thin in many size-by-book-to-market cells.¹⁰ The exercise asks which portfolios’ risk premium news is most exposed to aggregate belief-bias news.

Figure 1.2 reports the correlation between portfolio-level risk premium news and aggregate belief-bias news. The predicted negative sign appears in 19 of the 25 portfolios. The six positive cells are concentrated in the largest two size quintiles and the middle three book-to-

¹⁰ A fully portfolio-specific belief decomposition would require constructing $\nu_{p,t}^{(B)}$, $\eta_{p,t}^{(B)}$, and $\zeta_{p,t}^{(B)}$ separately within each FF25 portfolio. In the CBR data, only 10 of 25 cells have a median of at least 10 IBES-covered firms per month, while 12 cells, mostly small-cap portfolios, have median coverage of two firms or fewer. Portfolio-level belief proxies in those cells would be dominated by individual-firm noise. The FF25 exercise below therefore keeps the belief-bias news series aggregate and asks whether portfolio-level risk premium news is exposed to the aggregate belief-bias shock.

market quintiles. The pattern is therefore not uniform, but it is broad: most portfolios display the same negative relation between risk premium news and aggregate belief-bias news that appears in the market-level time series.

The portfolio-level slope restrictions are also consistent with the aggregate evidence. None of the 25 portfolio-level slope tests rejects under the bootstrap. I interpret this result cautiously. Portfolio-level tests are noisy, and failure to reject is not proof of exact equality in each portfolio. The useful fact is that the cross-section does not show a systematic pattern of rejections that would contradict the aggregate slope evidence.

I also examine sorting evidence designed to ask where the belief-news relation is stronger. The sorting results are directionally consistent with the behavioral channel but imprecisely estimated. The high-minus-low LTG tercile spread is -0.497 , and the high-minus-low analyst-coverage spread is -0.354 . Both have the predicted sign: portfolios where belief measurement or belief disagreement should matter more display more negative belief-news relations. But the confidence intervals are wide, so I treat these spreads as mechanism evidence rather than as independent identification.

The cross-sectional evidence broadens the aggregate result. The predicted sign appears in most size-by-book-to-market portfolios, and the sorting spreads point in the expected direction. This breadth matters because the aggregate relation is not confined to a single market-level return-news construction. Instead, aggregate belief-bias news is related to risk premium news across many portfolio return-news objects. I therefore interpret the cross-section as supportive evidence that the belief-news relation is a broad feature of equity returns.

Figure 1.2: Portfolio-Level Belief-News Diagnostics

Panel A: News Correlation



Panel B: Slope-Test Bootstrap p -Value



Notes: This figure reports two FF25 portfolio-level diagnostics across the 25 Fama–French size-by-book-to-market portfolios. For each cell, the VAR replaces the market excess return with the portfolio excess return while keeping the remaining state variables at the aggregate level. Rows correspond to size quintiles and columns correspond to book-to-market quintiles. Panel A reports the correlation between portfolio-level risk premium news and aggregate empirical belief-bias news; negative values indicate the sign predicted by the belief-driven risk premium news channel. Panel B reports the bootstrap p -value for the 1-step slope test $e_1'\Gamma = e_3'\Gamma$ in each cell. Appendix A.9 reports the full portfolio-level table and bootstrap inference.

1.7.2 Level vs. Reallocation

The market-level belief revision is value weighted. It can therefore move because the average firm revises beliefs, or because belief revisions are concentrated among firms with larger market weights. I decompose the aggregate revision into these two components. Let LEVEL_t denote the equal-weighted average firm-level revision, and let CSCOV_t denote the size-weighted reallocation component. Then the value-weighted aggregate revision can be written as

$$\nu_t^{(B)} = \text{LEVEL}_t + \text{CSCOV}_t.$$

The first term captures common movement in firm-level belief revisions. The second term captures whether revisions are concentrated among firms with high market weights.

Table 1.7 reports the correlation of each component with risk premium news. The decomposition exactly reconstructs the market-level belief-revision series used in the aggregate tests. It therefore asks where the aggregate relation with risk premium news comes from: common belief movement, size-weighted reallocation, or both.

The decomposition shows that the mechanism differs across proxies. For LTG, the aggregate relation is almost entirely a common level effect: the level component accounts for 97.86% of the aggregate covariance with risk premium news, while the reallocation component accounts for only 2.14%. This is consistent with LTG capturing broad movements in long-run growth expectations.

For STR, the pattern is different. The reallocation component has a much stronger negative correlation with risk premium news, -0.1106 , and accounts for 69.77% of the aggregate covariance. The level component is small and weakly related to risk premium news. Thus, for

Table 1.7: Decomposition of the Aggregate Belief Revision: Level vs. Reallocation

Proxy	Correlation with $N_{RP}^{(r)}$		Covariance Share (%)	
	LEVEL	CSCOV	LEVEL	CSCOV
CBR	−0.0462 (−0.7951)	−0.1133 (−1.9889)	54.50	45.50
LTG	−0.0680 (−1.2804)	−0.0219 (−0.5185)	97.86	2.14
STR	−0.0184 (−0.3320)	−0.1106 (−2.0041)	30.23	69.77

Notes: The table decomposes the market-level analyst revision into an equal-weighted level component, LEVEL, and a size-weighted cross-sectional reallocation component, CSCOV. The first block reports the correlation of each component with rational risk premium news $N_{RP}^{(r)}$, with Newey–West t -statistics in parentheses below each point estimate. The second block reports each component’s contribution to the aggregate *covariance* between $\nu^{(B)}$ and $N_{RP}^{(r)}$ (not variance shares). Appendix A.9 reports the full decomposition and reconstruction checks.

the short-term revision proxy, the aggregate relation comes primarily from where revisions occur in the size-weighted cross-section, not from a common revision by the average firm.

CBR lies between these two cases. The reallocation component has the stronger marginal correlation with risk premium news, -0.1133 compared with -0.0462 for the level component. But the covariance contribution is roughly split: 54.50% from the level component and 45.50% from reallocation. The reason is that the level component has much larger variance, so a smaller correlation can still contribute substantially to the aggregate covariance.

The main lesson is that the aggregate belief proxy is not merely a uniform market-wide optimism shock. It reflects both common belief movements and heterogeneous reallocation across firms, with the mix depending on the forecast horizon and proxy construction. LTG is primarily a level movement, STR is primarily a reallocation movement, and CBR combines both. This mechanism evidence complements the aggregate time-series results by showing how firm-level belief revisions aggregate into the market-level object used in the main tests.

1.7.3 Cross-Sectional Breadth and Mechanism

The cross-section strengthens the aggregate evidence along two dimensions. First, the predicted relation appears across many portfolio return-news objects, not only in the market return. The FF25 evidence therefore shows that aggregate belief-bias news is related to risk premium news across a broad part of the equity cross-section. Second, the level-vs.-reallocation decomposition shows that the market-level belief proxy reflects economically meaningful firm-level variation: common belief movements, size-weighted reallocation across firms, or both, depending on the proxy.

Sections 1.4–1.7 support the paper’s main conclusion. Risk premium news contains an empirically detectable belief-driven component. The evidence does not imply that all risk premium news is behavioral, or that measured analyst revisions perfectly capture marginal-investor beliefs. Instead, biased cash-flow beliefs are one economically important component of the object that standard return decompositions call risk premium news.

1.8 Conclusion

Risk premium news is a central object in empirical asset pricing. It is usually read as news about discount rates or expected returns. This paper asks how that interpretation changes when prices partly reflect biased beliefs about future cash flows. The answer is that the accounting object remains the same, but its economic content can change. Risk premium news can contain a belief-driven component.

The theory makes this point directly. In the sharp behavioral-pricing benchmark, rational

risk premium news equals the negative of unexpected cash-flow belief bias. A rational econometrician who omits the belief wedge attributes the resulting price movement to expected-return news. With partial arbitrage, the same logic survives with attenuation: risk premium news contains a belief-driven component alongside residual rational expected-return news. The key point is interpretive, not accounting. The Campbell and Ammer (1993) decomposition remains valid, but the meaning of one term depends on whether omitted belief distortions move prices.

The empirical evidence supports this interpretation. Using target-date-aligned analyst forecast revisions, I construct three proxies for cash-flow belief revisions and compare the implied belief-bias news to rational risk premium news. Across the CBR, LTG, and STR proxies, belief-bias news is negatively correlated with risk premium news, with correlations near -0.10 . The measured belief-bias object has meaningful scale: its variance is 0.28 – 0.34 of the variance of risk premium news. Forecasts move in the predicted direction as well: fitted return forecasts and fitted belief-revision forecasts comove positively, and 1-step slope restrictions are not rejected for any proxy under the bootstrap. These results are not driven by a single analyst measure, by contemporaneous price-dividend inference, or by small generic residual news. The cross-section adds breadth and mechanism: the relation appears across many portfolio return-news objects, and the aggregate belief proxy reflects both common belief movements and heterogeneous reallocation across firms.

The evidence also clarifies the limits of the sharp benchmark. Exact pointwise equality between measured belief-bias news and risk premium news is too strong. Matrix wedges are economically meaningful, and direct cumulative equality rejects at the 5-year horizon. These failures are informative. They show that risk premium news is not only belief-driven news. Residual rational expected-return variation, partial arbitrage, rate-belief components,

and proxy error all matter. The paper's conclusion is therefore a component claim, not an exclusivity claim: biased cash-flow beliefs are an empirically detectable and economically meaningful component of the object that standard decompositions call risk premium news.

This interpretation gives behavioral beliefs a direct role inside one of empirical asset pricing's central objects. Time-varying expected returns remain important, but the evidence shows that part of what standard decompositions label as risk premium news is systematically related to mistaken cash-flow beliefs. Risk premium news should therefore be read not only as discount-rate news, but also as a channel through which behavioral valuation errors appear in asset prices.

Chapter 2

The Psychology of Macroeconomic Expectations

with Pedro Bordalo, Nicola Gennaioli, Florencio Lopez-de-Silanes, Andrei Shleifer, and Maarten van Rooij

Chapter Summary

We present a model of “animal spirits” in which context and emotions affect macroeconomic beliefs by shaping which experiences people recall to simulate similar future aggregate states. We test this mechanism by priming Dutch National Bank Survey respondents to recall personal financial or health adversities before eliciting inflation and home price expectations. The treatment causes instability in beliefs and reasoning tied to the measured similarity of the primed experiences to different macroeconomic states. This similarity structure also accounts for heterogeneity in beliefs and reasoning based on a range of other personal experiences we measure. The model micro-founds several features of macroeconomic beliefs, including narratives, and yields a “confidence multiplier”: spending by some agents cues optimistic context for others, raising their spending.

2.1 Introduction

In the *Treatise on Probability* (1921), Keynes argued that beliefs over an uncertain future reflect unstable weighting of evidence, not a strict probabilistic calculus. In *The General Theory* (1936), he further stressed that economic decisions depend on confidence and emotions, also called “animal spirits.” Many papers in macroeconomics and finance document the role of sentiment or confidence (e.g., Lee et al. 1991; Ludvigson 2004; Baker and Wurgler 2006; Barsky and Sims 2012; Blanchard et al. 2013; Shapiro et al. 2022) and relate them to narratives (Shiller 2019a). Yet existing rational and behavioral models view beliefs as a stable function of news, leaving concepts like sentiment or confidence “amorphous” (Mankiw 2009). How can purely psychological states shift beliefs in the absence of news and, if so, through what mechanism?

We propose and test a model of mnemonic expectations in which, following Bordalo et al. (2023, 2025a), a salient cue causes retrieval of similar experiences from memory, which are then used to simulate the future, shaping beliefs. The cue includes the forecasting task, such as inflation, but also emotions and experiences currently top-of-mind. The forecasting task then cues recall of similar Domain-Specific (DS) experiences such as past inflation, which directly help simulate similar future inflation states. Current negative emotion cues the recall of similar Non-Domain-Specific (NDS) experiences, such as personal financial distress. These NDS experiences affect beliefs analogically, helping simulate similarly bad future inflation states. In a nutshell, mental states influence beliefs by altering the set of recalled experiences used for imagining the future.

The model yields sharp, testable predictions that link the heterogeneity of beliefs across people (due to different experiences and similarity perceptions) to the instability of beliefs from

changing cues. We test these predictions using a novel survey experiment on more than 4,000 households in the Dutch National Bank Household Survey (DHS). Motivating evidence shows that households experiencing worse financial or health conditions report higher expected inflation and home price growth, consistent with the role of NDS experiences documented by Cenyon (2025) and Taubinsky et al. (2024). Because such experiences are uninformative about aggregate conditions, this evidence points to non-statistical drivers of beliefs.

We take the next step and causally test the role of memory. Before eliciting inflation and home price growth expectations, we randomly prime respondents to recall a personal financial or health adversity. We then measure a respondent's (i) additional NDS experiences, (ii) their perceived similarity to high inflation and home price growth, (iii) backcasts, and (iv) reasons for the forecasts. The design isolates memory from information: priming an NDS experience is irrelevant with full recall and does not convey news about the target aggregate variable. Any effect must reflect a change in mental context. Associative memory further implies that priming one experience cascades into recall of other similar ones, allowing us to trace how cues reshape recall and the reasoning behind the forecasts, not only their numerical values.

The evidence supports four main predictions. The first is similarity-based experience effects: people with NDS experiences perceived as more similar to high inflation or home price growth report higher expectations, exhibiting more pessimism than people who have not lived such experiences. The second is similarity-based instability: people who are primed to recall a lived NDS adversity are more pessimistic than comparable individuals who are not primed. Third, priming financial adversities affects beliefs more strongly than priming health ones, consistent with the higher similarity of the former to bad macroeconomic states. Fourth, and crucially, priming changes both recall and reasoning: it raises backcasts (consistent with greater retrieval of similarly adverse DS experiences) and encourages reasoning focused on

personal affordability concerns and negative emotions (consistent with greater retrieval of adverse NDS experiences).

The effects are large. Priming shifts inflation (home price growth) beliefs up by 1.01 (1.08) percentage points relative to a 5.46% (6.51%) average. Disagreement is also large: moving across the interquartile range of NDS experiences raises these beliefs by 0.74 (0.91) percentage points, comparable to observed disagreement among key socio-demographic groups (Weber et al. 2022).

We provide a micro-founded and testable account of “animal spirits,” sentiment, and narratives as drivers of macroeconomic beliefs that are independent of news or stable tastes. This mechanism unifies several observations, including: (i) the role of NDS experiences (Cenzon 2025; Taubinsky et al. 2024), which we identify causally, (ii) insensitivity to statistical information (Sims 2003), (iii) average bias in inflation forecasts (Weber et al. 2022), (iv) overreaction to salient inflation changes (Bonaglia et al. 2025; Gennaioli et al. 2025; Link et al. 2023), and (v) strong distaste of inflation due to its mental association with personal financial distress. Idiosyncratic mental states also explain belief disagreement by agents exposed to similar inflation news or gas prices (Schwartzstein and Sunderam 2021; Andre et al. 2022; Shiller 2019a; Andre et al. 2024).

The implications of memory go beyond our causal experiment. It is well known that expectations matter for economic choices (Dominitz and Manski 2007; Armona et al. 2019; Bailey et al. 2019; Giglio et al. 2021; Coibion et al. 2023; Jiang et al. 2025) including due to NDS experiences (Cenzon 2025; Taubinsky et al. 2024). Our mechanism implies that persistent interpersonal differences in behavior can arise from belief disagreement due to different experiences and personal contexts. In line with Das et al. (2020), memory predicts more pessimism by peo-

ple with worse socioeconomic status, who likely experience both more adversities and more negative context, with non-trivial allocative consequences.

Memory also generates amplification: a bad shock carries not only bad news but also negative context, triggering recall of related adverse macroeconomic or personal experiences. Due to these recall spillovers, beliefs overreact to the shock. Memory also implies a “confidence multiplier”: because one agent’s actions serve as cues for others, higher consumption by one person cues optimism for others, encouraging their spending in line with Bellifemine et al. (2025). Memory offers a new mechanism for demand fluctuations due to self-sustaining changes in confidence (Akerlof and Shiller 2009; Lorenzoni 2009; Angeletos and La’O 2013).

Existing models of expectations rely on stable news processing or tastes. Theories of under- and over-reaction assume stable mis-specified models (e.g., Barberis et al. 1998; Mankiw and Reis 2002; Sims 2003; Woodford 2003; Fuster et al. 2010; Gabaix 2014; Angeletos et al. 2021). Models of disagreement trace it to different models or priors. Malmendier and Nagel (2011, 2016) model Bayesian learning from personal DS experiences. Another body of work stresses motives: distortions in news processing due to a stable preference for robustness (Hansen and Sargent 2010) or wishful thinking (Brunnermeier and Parker 2005). These approaches cannot explain why priming a personal adversity changes beliefs and reasoning without providing any information and via spurious similarity. A memory model in which beliefs reflect what is top of mind is needed to account for the data.

Section 2.2 presents motivating evidence. Section 2.3 develops the model. Section 2.4 describes the experiment. Section 2.5 tests the model. Section 2.6 discusses implications. Section 2.7 concludes.

2.2 Motivating Evidence

2.2.1 The DNB Household Survey

The DNB Household Survey (DHS) has been conducted annually since 1993 by CentERdata at Tilburg University. Each wave surveys personal circumstances and perceptions of more than 2,000 Dutch households, along with their financial and demographic information (51% of households participate for three years or more). Survey expectations cover retirement age, household income, mortgage rates, and our macroeconomic variables of interest: inflation and home price growth. We focus on the 2012–2022 surveys to avoid the financial crisis period.

We elicit a household’s modal inflation expectations using the DHS question: “*What do you think is the most likely (consumer) price increase over the next twelve months?*” Subjects choose a value between 0% and 10%. For the growth of home prices, the survey asks: “*What do you expect the prices of homes in the next two years will do? Will the prices rise, fall or stay about the same? By how many percent per year on average will prices [rise / fall]?*”

The DHS measures various personal circumstances through demographics such as age, gender, relationship status, occupation (employed, self-employed, student, retired), income, accommodation (tenant, subtenant, owner), and urbanization, which are known to predict macroeconomic beliefs (e.g., Das et al. 2020). These proxies do not identify a mechanism, so we use them as controls. We also control for numeracy, proxied by standard probability questions.

Other questions measure personal experiences and attitudes, mostly in five domains: health status, happiness, financial status, risk taking, and experience with financial manage-

ment. We focus on a person’s health and financial status, which are directly linked to personal experiences, offering clear guidance for our experimental manipulation. We construct a *Health Adversity* index by averaging a respondent’s z -standardized answers to: “*How is your health in general?*” (1: *excellent*; 5: *poor*), and “*How many times did you contact your general practitioner about your own health in the previous year?*” We construct a *Financial Adversity* index by averaging z -standardized answers to questions such as: “*How well can you manage on the total income of your household?*” (1: *very easy*; 5: *very hard*), or “*Have you, in the previous year, consulted with your bank or financial institution, because you had or expected payment problems with the repayment of a loan or mortgage?*” For both indices, a higher value implies a worse state.¹

2.2.2 Personal Adversities and Macroeconomic Beliefs

Table 2.1 reports non-causal regressions of inflation and home price growth beliefs on the adversity indices, controlling for time dummies in columns (1) and (3), and additionally for demographics in columns (2) and (4). Our DHS sample covers the years 2012–2022. Health and financial adversities are correlated with sharply higher expectations of inflation (column 1) and home price growth (column 3). Results are robust to demographic controls for inflation (column 2), but financial adversity loses significance for home price growth (column 4). Financial experiences matter more for inflation and home price growth than health ones.²

¹ Appendix B.2 reports the questions in the Financial Adversity index. Some include near-future outcomes, such as “*Do you think the expenditures of your household, in the next 12 months, will be higher, about the same, or lower than the income of your household?*” Table 2.1 is robust to removing such questions (which may depend on macroeconomic beliefs).

² Higher education, student status, income, and homeownership are negatively correlated with expected inflation. Age, rural residence, and homeownership are negatively correlated with expected home price growth. Column (4) is robust to controlling for perceived home price increase (positive and significant coefficient).

Table 2.1: Adversity and Macroeconomic Beliefs in the DHS

	Inflation		Home Price Growth	
	(1)	(2)	(3)	(4)
Health Adversity	0.0803*** (0.0189)	0.0583*** (0.0187)	0.0393*** (0.0128)	0.0302** (0.0138)
Financial Adversity	0.3043*** (0.0238)	0.2413*** (0.0259)	0.0614*** (0.0222)	0.0223 (0.0244)
<i>N</i>	22,081	19,852	17,928	16,504
Controls	No	Yes	No	Yes
Year FE	Yes	Yes	Yes	Yes
<i>R</i> ²	0.0311	0.0476	0.0016	0.0203

Notes: This table reports OLS regressions of standardized inflation and home price growth beliefs on two individual-level adversity measures: health adversity and financial adversity (higher values indicate worse health or financial conditions). Dependent variables are standardized within survey year; adversity indices are standardized across the pooled sample. Columns (1) and (2) focus on inflation; columns (3) and (4) on home price growth. Controls include age, gender, education, occupation, urbanization, household income, housing status, numeracy, and relationship status. All specifications include survey-year fixed effects and province fixed effects (not shown). Standard errors clustered at the individual level in parentheses. Significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

In columns (2) and (4), a 1-standard-deviation increase in financial (health) adversity entails a 0.24 (0.06) SD increase in inflation forecasts and a 0.02 (0.03) SD increase in home price growth forecasts. These are large magnitudes, for example, a 0.51 p.p. increase in expected inflation relative to a 3.05% baseline associated with a financial adversity, and a 0.15 p.p. increase in expected home price growth relative to a 1.84% baseline associated with a health adversity.³ It would be hard to get comparable disagreement from domain-specific experiences such as historical inflation, which varies only across cohorts, or from food or gas prices (Kuchler and Zafar 2019; Coibion et al. 2022), which are largely common across people.

³ The sample mean (SD) of inflation expectations is 3.05% (2.13); that for home price growth ones is 1.84% (4.90). Thus, a 1-standard-deviation increase in financial adversity entails a 0.24 SD increase in inflation, equivalent to 0.51 p.p. (= 0.2413 x 2.13 p.p.). For home price growth, health adversity yields a 0.15 p.p. effect (= 0.0302 x 4.90 p.p.).

This explanatory power does not come from aggregate economic conditions, which are controlled for with time dummies. Individual-level variation in adversities is uninformative about future inflation and home price growth, rejecting information channels as previously shown by Cenoz (2025) and Taubinsky et al. (2024).⁴ The evidence is also at odds with canonical rational or behavioral models, in which beliefs depend on statistical data. Personal events may correlate with access to news, but this may explain higher noise in beliefs, not systematic pessimism.

The evidence is consistent with Keynes' idea that macroeconomic beliefs and confidence embed psychological states, not just a statistical calculus. One explanation is motivated beliefs: people prepare for their worst-case scenario (Hansen and Sargent 2010; Bhandari et al. 2025), to an extent that depends on their exposure to inflation or home price risk. Another explanation is memory: when thinking about inflation or home prices, people retrieve experiences, including wholly personal ones. The features of these experiences, including their emotions, help people imagine specific macroeconomic outcomes, depending on their similarity.

We causally test memory by priming recall of personal adversities without giving any news. This intervention should have no effect in models relying on a correct or distorted statistical calculus, including Bayesian learning from experiences (Malmendier and Nagel 2011, 2016), or in models of beliefs tainted by preferences, including for robustness (Hansen and Sargent 2010) or wishful thinking (Brunnermeier and Parker 2005). Stable tastes cannot be affected by uninformative recall interventions. Critically, our primary goal is not to reject existing models but to test a rich set of memory predictions based on measured experiences,

⁴ Also in the DHS, personal adversities do not predict macroeconomic states (see Appendix B.2). Results are robust to exclusion of Financial Adversities proxies related to macroeconomic states, e.g., whether expenditures are expected to outpace income.

similarity, and reasoning. Bordalo et al. (2025a) study these predictions in novel domains (e.g., a large-scale cyberattack). Here we focus on macroeconomic domains in which people have many direct experiences and information. Relative to Cenyon (2025) and Taubinsky et al. (2024), we causally identify memory by priming recall of different experiences, by studying the role of similarity, of the broader memory database, and of reasoning. Compared to motivated beliefs, which depend on unobserved tastes, the structure of memory can be measured and linked to clear new predictions.

2.3 Selective Memory and Belief Formation

We rely on the psychology of selective retrieval (Kahana 2012) and of mnemonic simulation of the future (Hassabis et al. 2007). We formalize these ingredients and then study their implications for inflation and home price growth expectations.

2.3.1 The Setup

The **database** is a frequency distribution $\mathbf{n} = \{n_e\}_{e \in E}$ over categories of experiences $e \in E$, where n_e is the number of lived experiences in category e and E is a memory space. e is a vector $e = (x, v, c)$ where $x \in \mathbb{R}$ is a realization, an objective marker of the experience (e.g., realized inflation or its forecast in the news), $v \in \mathbb{R}$ is the emotion associated with it, which plays a key role in memory encoding and retrieval (Kensinger 2009), and c is context.

Context belongs to a metric space C and partitions experiences semantically, $C = C_M \cup C_p$ where $c \in C_M$ and $c \in C_p$ capture macroeconomic and personal contexts, respectively. Macroeconomic context separates inflation experiences, $c \in C_{II}$, from home price growth

ones, $c \in C_\Gamma$ (so $C_\Pi \cup C_\Gamma \subseteq C_M$). Personal context separates financial experiences, $c \in C_f$, from health ones, $c \in C_h$ (so $C_f \cup C_h \subseteq C_p$). As we show later, context separates DS from NDS experiences.⁵

Retrieval is triggered by a cue q , which is a subset (x_q, v_q, c_q) of E . Retrieval of $e \in E$ depends on its similarity to q , its frequency, and interference from other experiences. Given a similarity function $S(e, q)$, the probability of recalling e based on q and database $\mathbf{n}$ is:

$$r_e(q, \mathbf{n}) = \frac{S(e, q) \cdot n_e}{\sum_{e' \in E} S(e', q) \cdot n_{e'}}. \quad (2.1)$$

The numerator captures associativity and frequency: experiences more similar to the cue and more frequently lived are more likely to be recalled. The denominator captures interference: even if e is similar to q and frequent, its recall can be blocked by an experience $e' \neq e$ that is even more similar to q or more frequent. These forces have been used to study the role of memory for expectations (Bordalo et al. 2023, 2025a; Taubinsky et al. 2024; Cenzon 2025; Bonaglia et al. 2025), consumer choice (Bordalo et al. 2020), and advertising (Bordalo et al. 2025b).

Simulation maps retrieved experiences into a forecast. Consider a person who forms a point estimate of inflation $\mathbb{E}(\pi)$ or home price growth $\mathbb{E}(\gamma)$. The forecasting task is also a cue, denoted by $t_T = (\mathbb{R}, 0, c_T)$ with $T = \Pi$ for inflation and $T = \Gamma$ for home price growth. The first entry in the task is any possible inflation or home price growth rate in $\mathbb{R}$. The second entry, valence, is neutral, $v = 0$. The third entry is the task's domain, inflation c_Π or home

⁵ Categories arise because people store similar events together (e.g., they do not distinguish in memory between 2.9% and 2.8% inflation). Categories can be highly specific if the context defining them is granular, for instance inflation in 1971 may be a separate category from that in 1972. Not fully granular categories simplify the formal analysis.

price growth c_Γ .

The task's context distinguishes between DS and NDS experiences. In task t_T , the DS experiences are vectors $e = (\tau, v, c_T) \in E$ having the forecasting context c_T and its realizations $\tau = \pi$ for inflation and $\tau = \gamma$ for home price growth (possibly broadened to prices of specific goods or homes). The context of NDS experiences is instead incongruent with the task, $c \neq c_T$: a past illness is NDS for inflation or home price growth (but DS for own health).

A simulation function projects, for a given task t_T , the features of a generic experience e onto a point estimate, formally $\sigma : \{t_\Pi, t_\Gamma\} \times E \rightarrow \mathbb{R}$, where we use the shorthand:

$$\sigma_T(e) \equiv \sigma(t_T, e), \quad T = \Pi, \Gamma. \quad (2.2)$$

Beliefs are then formed by retrieving experiences from E based on the cue q (equation (2.1)) and then using them to simulate the future (equation (2.2)). The cue is influenced by an experience $\tilde{e} \in E$ the agent is thinking about, for instance an illness or a recent op-ed on inflation she read; $\tilde{e}$ pollutes the valence and context of the task t_T , while leaving the inflation or home price space of realizations $\mathbb{R}$ unaffected. Formally, the cue is given by:⁶

$$q_{\tilde{e},T} = (\mathbb{R}, v_{\tilde{e}}, c_{\tilde{e}}).$$

A probability distribution of thoughts, denoted by $\omega_T = \{\omega_{\tilde{e},T}\}_{\tilde{e} \in E}$, induces a distribution of cues when solving t_T , where $\omega_{\tilde{e},T}$ is the probability of cue $q_{\tilde{e},T}$. Integrated across cues, the

⁶ An alternative assumption is that the cue's context retains also the context of the task, $c_{\tilde{e},T} = c_T \cup c_{\tilde{e}}$. The results also hold under this assumption, but the current formalization simplifies the algebra in the proofs.

probability of recalling $e \in E$ in task t_T (and of using it for simulation) is then given by:

$$r_e(\boldsymbol{\omega}_T, \mathbf{n}) = \sum_{\tilde{e} \in E} r_e(q_{\tilde{e}, T}, \mathbf{n}) \cdot \omega_{\tilde{e}, T},$$

where $r_e(q_{\tilde{e}, T}, \mathbf{n})$ follows equation (2.1). The average belief in task t_T is given by:

$$\mathbb{E}(\tau \mid \boldsymbol{\omega}_T, \mathbf{n}) = \sum_{e \in E} r_e(\boldsymbol{\omega}_T, \mathbf{n}) \cdot \sigma_T(e), \quad (2.3)$$

which is expected simulation across selectively retrieved experiences. Due to spontaneous retrieval and simulation, equation (2.3) offers a theory of macroeconomic beliefs and confidence where, as in Keynes' animal spirits, beliefs can depend on irrelevant emotions and context. This is not so with standard models of expectations, which combine statistical priors and signals in an optimal or distorted, yet context-independent, way. Mnemonic beliefs are more general: they are driven by similarity to past experiences and their frequency, which may be informative DS experiences (Bordalo et al. 2018; Bonaglia et al. 2025; Bonaglia and Gennaioli 2025; Gennaioli et al. 2025) or uninformative NDS ones, which methodologically allow us to identify mechanisms such as animal spirits and narratives.

To test this mechanism, we prime recall of a personal experience e_p which triggers the cue $q_{e_p, T}$, changing the decision maker's (DM) mental state. Memory makes a prediction very different from standard reminder effects, which exclusively bring to mind previously forgotten information (Karlan et al. 2016; Calzolari and Nardotto 2017). In our case, reminding a person about an uninformative personal experience causes a similarity-based cascade that brings to mind many other experiences that shape beliefs directly via simulation and indirectly by crowding out other information.

Existing work on “mental states” views emotions as boosters of optimism or pessimism (Schwarz and Clore 1983; Loewenstein 2000; Damasio 1994; Sharot et al. 2007). Memory explains why this can occur via retrieved experiences and simulation based on similarity, offering new measurement and tests. To highlight context effects most sharply, we prime uninformative NDS experiences. More generally, even DS experiences can act through context, yielding new implications that we discuss in Section 2.6.

2.3.2 Beliefs, Experiences, and Priming Effects

Assumptions on Similarity $S(e, q)$, Simulation $\sigma_T(e)$, and the Database. Following Nosofsky (1986), similarity declines with dissonant contextual and valence features. Following Hassabis et al. (2007) and Bordalo et al. (2025a), simulation projects the features of the experience into a point estimate for task t_T , and more strongly so if the context of e is closer to the context of the task T . These broad properties, which we later validate, drive our key predictions. To obtain an intuitive characterization we make more specific assumptions (but prove general results in Appendix B.1).

Assumption 1: Similarity. *There is a $K \in (0, 1/2)$ such that similarity between $e = (x, v, c)$ and $q = (x_q, v_q, c_q)$ satisfies:*

$$S(e, q) = \begin{cases} 1 & \text{if } v \cdot v_q > 0 \text{ and } c \in c_q, \\ 1 - K & \text{if either } v \cdot v_q \leq 0 \text{ or } c \notin c_q \text{ (but not both),} \\ 1 - 2K & \text{if } v \cdot v_q \leq 0 \text{ and } c \notin c_q. \end{cases}$$

Similarity is maximal, equal to 1, if the valence and context of e are congruent with those

of q (i.e., $c \in c_q$ and they are both adversities $v, v_q < 0$ or good times $v, v_q > 0$). It drops by K if one feature is incongruent, by $2K$ if both features are incongruent. Similarity could be smoother and act also along the objective realization τ . What matters is that valence and context create an associative hierarchy. A person cued with a personal financial adversity (i.e., $v_q < 0$ from personal financial context c_f) most easily recalls other financial adversities (two congruent features), she next recalls non-financial adversities or good personal financial states (one congruent feature), and she lastly recalls favorable non-financial events (zero congruent features).

Assumption 2: Simulation. *Simulation in task $t_T = (\mathbb{R}, 0, c_T)$ based on experience $e = (x, v, c)$ satisfies:*

$$\sigma_T(e) = \delta_T + \sigma_T(c) \cdot v,$$

where δ_T is a zero-valence default state. The sensitivity of simulation to valence satisfies $\sigma_T(c) \leq 0$ if a higher target realization τ has lower valence and $\sigma_T(c) \geq 0$ otherwise. Simulation is less sensitive to valence if the discrepancy $d(c, c_T)$ between the target context c_T and the experience's context c is larger, $\partial|\sigma_T(c)|/\partial d(c, c_T) < 0$. Congruent context entails full simulation, $|\sigma_T(c_T)| = 1$, with $\sigma_T(c_T) = -1$ when higher τ is adverse and $\sigma_T(c_T) = 1$ otherwise.

Because high inflation has low valence, its simulation coefficient is negative, $\sigma_\Pi(c) \leq 0$. Thus, if the DM recalls a DS high inflation episode, she uses the episode's low valence $v < 0$ to simulate a bad high inflation state. Crucially, even if the DM recalls past personal financial troubles, an adverse NDS with $v < 0$, she can use its negative valence to imagine high inflation.

Unlike in motivated models, the valence of a future macroeconomic state does not mechanically link to the person's current utility from such a state. Although high inflation and home price growth may benefit current homeowners holding fixed-rate mortgages, these peo-

ple may still view these macroeconomic states as bad due to their bad past experiences as wage-earning renters, as well as due to social conventions about the valence of aggregate outcomes, such as the concerns of many people about the affordability of goods and housing when their prices increase rapidly. Key to unlocking the structure of simulation is then our measurement in Section 2.4 of the perceived similarity of adverse experiences, including personal financial distress, to macroeconomic states.

Critically, sensitivity to valence is full for a DS experience $\sigma_{\Pi}(c_{\Pi}) = -1$, attenuated for NDS personal financial ones, and even more attenuated for a less similar health one $0 > \sigma_{\Pi}(c_f) > \sigma_{\Pi}(c_h) > -1$. The same holds for home price growth, $0 > \sigma_{\Gamma}(c_f) > \sigma_{\Gamma}(c_h) > \sigma_{\Gamma}(c_{\Gamma}) = -1$. Highly dissimilar experiences with sharply incongruent context or no valence ($v = 0$) inhibit simulation. In these cases, $\sigma_T(c) = 0$ or $v = 0$, and the forecast anchors to a neutral default such as a “normal” state or a random element in the response scale. Thus, recalled DS and NDS experiences shape pessimism according to the similarity hierarchy in Bordalo et al. (2025a).⁷

More broadly, and beyond assumed functional forms, the magnitude and sign of $\sigma_T(e)$ are recovered from: (i) perceived similarity between the primed adversities and the target macroeconomic states, and (ii) “experience effects” of these adversities in our control condition. This is sufficient to obtain a rich set of within and cross-domain restrictions on experience and priming effects.

Consider finally the database. Let $C_{\emptyset}$ be the set of contexts of valence-neutral experiences, $n_{\emptyset}$ their numerosity, and $C_{-\emptyset} \subseteq C$ the set of valence-charged contexts. Neutral experiences

⁷ For NDS experiences, simulation relies on valence, not on the realization x (which is incompatible with the target τ). For DS experiences simulation can instead also use x (i.e., the DM projects past 10% inflation into future 10% inflation). Assumption 2 entails no loss of generality if there is a one to one mapping between the two, $v(\tau) = -\tau$.

capture routine events such as thinking about our usual breakfast, walking to the office, etc. Because they have no valence and are distant in context, they do not foster simulation.

Assumption 3: Database. *Within each $\tilde{c} \in C_{-\emptyset}$ the distribution of valence is symmetric around $v = 0$, $n_{(v,\tilde{c})} = n_{(-v,\tilde{c})}$ and the distribution of cues is also symmetric, $\omega_{(v,\tilde{c}),T} = \omega_{(-v,\tilde{c}),T}$.*

We thus characterize the effect on beliefs caused by adding an experience e_p to an otherwise valence-neutral “representative” database $\mathbf{n}$, and the effect of priming recall of e_p relative to a valence-neutral distribution of cues ω_T .

2.3.3 Characterizing Experience and Priming Effects

Consider memory’s predictions for experience and priming effects. We focus on the case relevant for us: higher target realizations τ have lower valence, $\sigma_T(c) \leq 0$, and the lived and primed event e_p is an adversity, $v_p < 0$. Results intuitively extend to the other cases.

The **experience effect** of adding adversity e_p to the baseline database $\mathbf{n}$, resulting in an enriched database $\mathbf{n}_p$, entails a change in beliefs given by:

$$\mathbb{E}(\tau \mid \omega_T, \mathbf{n}_p) - \mathbb{E}(\tau \mid \omega_T, \mathbf{n}) = r_{e_p}(\omega_T, \mathbf{n}_p) \cdot \left[\sigma_T(e_p) - \overline{\sigma_T(e)}_{-p} \right], \quad (2.4)$$

where $\overline{\sigma_T(e)}_{-p}$ is a weighted average of simulation from experiences $e \neq e_p$. Living e_p allows for recall and simulation based on it while dampening recall of other experiences $e \neq e_p$ across all cues in the baseline distribution ω_T . The DM is more pessimistic if and only if e_p promotes simulation of high inflation or home price growth compared to the average, $\sigma_T(e_p) > \overline{\sigma_T(e)}_{-p}$. This occurs if e_p is a DS high inflation experience (Malmendier and Nagel 2016) or if e_p is an NDS whose low valence helps imagine a bad state (Cenzon 2025; Taubinsky et al. 2024). Note

the importance of retrieval: if e_p is not spontaneously recalled, $r_{e_p}(\omega_T, \mathbf{n}_p) \approx 0$, its mere experience does not affect beliefs, even if it helps simulation.

Priming e_p for someone who has lived this experience shifts the cue from the baseline ω_T to $q_{e_p, T}$, namely it sets $\omega_{e_p, T} = 1$. Beliefs change as:

$$\mathbb{E}(\tau \mid \mathbf{1}_{e_p}, \mathbf{n}_p) - \mathbb{E}(\tau \mid \omega_T, \mathbf{n}_p) = \sum_{e \in E} r_e(q_{e_p, T}, \mathbf{n}_p) \cdot [\sigma_T(e) - \mathbb{E}(\tau \mid \omega_T, \mathbf{n}_p)], \quad (2.5)$$

where $\mathbf{1}_{e_p}$ is degenerate on $q_{e_p, T}$. The direct priming effect is to boost recall of e_p , which boosts beliefs if $[\sigma_T(e_p) - \mathbb{E}(\tau \mid \omega_T, \mathbf{n}_p)] > 0$. This condition is met if simulation based on e_p is stronger than average simulation across $e \neq e_p$. Thus, the direct priming effect of e_p is analogous to its experience effect. Critically, the other terms in equation (2.5) show that priming also creates recall and simulation spillovers to other experiences $e \neq e_p$. Such spillovers can give an overall positive priming effect of e_p even if the direct priming effect and the experience effect of e_p are negative.

Proposition 1. *If the frequency of neutral experiences is high enough, $n_\emptyset > \bar{n}$, priming adversity e_p with a positive experience effect causes beliefs in t_T to change in the same direction. Formally:*

$$\mathbb{E}(\tau \mid \omega_T, \mathbf{n}_p) - \mathbb{E}(\tau \mid \omega_T, \mathbf{n}) > 0 \Rightarrow \mathbb{E}(\tau \mid \mathbf{1}_{e_p}, \mathbf{n}_p) - \mathbb{E}(\tau \mid \omega_T, \mathbf{n}_p) > 0. \quad (2.6)$$

If having lived e_p predicts more inflation and home price growth pessimism, then priming e_p should boost such pessimism, *strengthening* the experience effect. By contrast, if the experience effect is not positive, priming could still promote pessimism if: (i) e_p has above average simulation but it is not spontaneously recalled, $r_{e_p}(\omega_T, \mathbf{n}_p) \approx 0$, or (ii) e_p is not itself a source of simulation, $[\sigma_T(e_p) - \overline{\sigma_T(e)}_{-p}] \approx 0$, but priming it creates recall spillovers on other DS

and NDS experiences that help the agent to simulate pessimistic scenarios.

Memory thus imposes three cross-equation restrictions. First, there is a link between experience and priming effects in equation (2.6). Second, similarity matters: the experience and priming effect of an adversity e_p should be stronger than that of $e_{p'}$ if the former is more similar to high inflation or home price growth. The third restriction concerns the mode of reasoning: because priming changes retrieval, it changes the DS experiences that are retrieved and used to form beliefs (e.g., past inflation or home price growth), or the retrieved NDS experience and hence the valence used to form beliefs (e.g., negative emotions). We test these predictions in Section 2.5.⁸

These predictions discipline memory foundations for context effects such as “animal spirits” or “narratives” used to imagine future macroeconomic states. We now show that, relative to other theories, memory ties beliefs and reasoning to measured experiences, similarity, and features.

2.4 The Survey: Design and Basic Facts

Section 2.4.1 describes the survey experiment and links it to the model. Section 2.4.2 documents basic facts that allow us to derive our model’s predictions.

⁸ The focus is on the role of NDS adversities, following the results in Table 2.1 and the fact that, as shown later, high inflation and home price growth are perceived to be similar to several adversities. However, the model is symmetric and produces corresponding effects for positive NDS experiences: Appendix B.2 shows that in the DHS positive health changes or financial experiences (such as receiving an inheritance) predict lower inflation and home price growth expectations, similar to Taubinsky et al.’s (2024) finding that wage hikes are associated with lower inflation forecasts.

2.4.1 The Survey Experiment

We fielded our survey in June 2024 on the DHS households, with a second wave in October 2024.⁹ Here we describe its basic structure. Appendix B.2 contains the details. The survey’s key innovation is to measure not just beliefs but also memory factors and to manipulate those.

Outcomes: Macroeconomic Beliefs. Following the DHS, we elicited beliefs about 12-month-ahead consumer price inflation: the modal forecast and the probability of a “7% or larger” tail. We also elicited beliefs about the 12-month-ahead percentage change in national home prices: the modal forecast and the probability of a 10% or greater increase. These measures correspond to the inflation and home price growth tasks, t_{Π} and t_{Γ} , in the model. As with the DHS data, we construct inflation and home price growth z -scores.

Reasons. Right after reporting beliefs, respondents write three to four sentences about thoughts or experiences that came to mind when answering the inflation and home price growth questions. This text lets us see how primed mental states change the mental models people use.

Cues: Priming NDS Experiences. We randomly assign participants to one of three groups: financial ($N = 1,267$), health ($N = 1,262$), denoted $pR = fR, hR$, respectively, and control ($N = 1,284$), denoted by $\emptyset$. In fR , prior to the forecasting task t_T people are primed to *recall* (hence the “R” index) whether they or a loved one ever faced personal or family struggles with financial hardship. In hR they are instead primed to recall a serious or chronic illness, or a serious accident involving them or a loved one. In both treatments, people describe the event, including the time, difficulties, and emotions. Priming is the only difference from Control.

⁹ The first wave contacted 6,514 households in the DHS panel and obtained complete responses from 4,348 (67%) of them. The second wave surveyed 4,301 households, with an 89% response rate. The experiment uses the enlarged DHS panel, combining panel members from both the Centerpanel and the LISS (Longitudinal Internet Studies for the Social Sciences) panel, both administered by CentERdata.

Primed recall changes the respondent’s mental context during forecasting t_T by increasing the salience of the financial or health context c_p , its objective realization x_p , and its subjective valence v_p , $p = f, h$, setting the cue at $q_{p,T}$, $p = f, h$. In Control $\emptyset$ the cue is instead drawn from the unobserved distribution ω_T . Critically, fR and hR change context without conveying news. Indeed, people recall experiences they have lived through, and hence already *know* about. In addition, such NDS experiences have no informative content with respect to inflation and home price growth. Thus fR and hR should play no role in any model, rational or behavioral, in which memory is perfect and beliefs only depend on news or stable tastes. This is not so with memory, where priming shapes what people think about during t_T .

Other NDS Experiences and Similarity. After beliefs and reasons, respondents reported if they or a loved one lived each of a list of ten NDS experiences: health adversities, financial struggles, extreme weather, inequality, social conflict, contract breaches, successful investments, immigration, unemployment, and divorce. These experiences proxy for the database $\mathbf{n}$. We then asked respondents to report, on a 1–7 scale, the similarity of each NDS experience with high inflation and home price growth. These subjective ratings are informative about the similarity and simulation function $\sigma_T(e)$. These measures allow us to test memory’s cross-equation restrictions: (i) across experience and priming effects, (ii) across different primes, and (iii) across domains.

At the end of the survey, we elicited a range of demographics and a numeracy score (asking the share of Dutch people 90 years or older) so as to replicate individual-level controls in Table 2.1. In our June survey, the provider failed to elicit the similarity of the cued experience e_f in fR and e_h in hR . We collected this measure in a second wave in October 2024. To minimize painful memories we did not ask respondents to again recall the adversity e_p . They: (i) could revise their earlier description of adversity if they wished, and (ii) were asked to

report inflation and home price growth beliefs, experiences, as well as perceived similarity of experiences and macroeconomic outcomes.¹⁰ The second wave fills the data gap but also allows us to test the stability of beliefs, experiences, and similarity. We focus on the first wave but also discuss second wave results.

2.4.2 Forecasts and Experiences in the Control Sample

Table 2.2 reports descriptive statistics for beliefs about inflation and home price growth across the two survey waves, in the treatment and control samples.

Table 2.2: Summary Statistics for Macroeconomic Beliefs

	Wave 1			Wave 2		
	Mean	Median	SD	Mean	Median	SD
Inflation (Point Forecast)	5.46	3.0	8.40	5.27	3.5	6.72
Inflation (Threshold)	4.42	4.0	2.48	4.65	5.0	2.45
Home Price Growth (Point Forecast)	6.51	5.0	7.55	7.06	5.0	6.96
Home Price Growth (Threshold)	4.80	5.0	2.42	5.20	5.0	2.37

Notes: Inflation expectations are elicited as point forecasts for the expected 12-month change in consumer prices, together with subjective probabilities that inflation will exceed a 7% threshold (measured on a 0–10 scale). Home price growth expectations are elicited analogously as a point forecast for the expected 12-month percentage change in national home prices and the subjective probability of a 10% or greater increase. The table reports the mean, median, and standard deviation of each belief measure by survey wave.

Inflation and home price growth point forecasts exhibit strong disagreement (8.4 and 7.55 SD, respectively) and are highly right-skewed (means are about 0.29 and 0.20 SD higher than medians, respectively). Beliefs about tail events are also heterogeneous but less skewed. These statistics are stable across waves and comparable to beliefs in the DHS.¹¹

¹⁰ In the finance (health) treatment, 11% (15%) of participants added text to their earlier answers from the first wave. Among those who added text, 11% (26%) had not reported a health/financial adversity in the first wave.

¹¹ In 2024, the mean (median) inflation and home price growth point forecasts in the DHS were 4.09% (3%) and

Before studying priming, consider the correlation between macroeconomic beliefs and the financial and health adversities in our control sample, in the spirit of the DHS analysis in Table 2.1. We describe other measured experiences in the next section. Using data from Control, we estimate:

$$z_{i,T} = \alpha_T + \varepsilon_{f,T} \cdot \mathbb{1}_{i,f} + \varepsilon_{h,T} \cdot \mathbb{1}_{i,h} + \eta_T \cdot X_i + u_{i,T}, \quad (2.7)$$

where $z_{i,T}$ is the belief z -score (higher $z_{i,T}$ means higher inflation for $T = \Pi$ or home price growth for $T = \Gamma$), $\mathbb{1}_{i,e}$ is a dummy for whether respondent i reports having had financial ($e = f$) or health ($e = h$) adversities. X_i includes our socio-demographic controls and the measure of numeracy.

Consistent with Table 2.1, Table 2.3 shows that experienced financial adversity is associated with substantially higher inflation and home price growth expectations. The effect is quantitatively large, leading to 1.27 p.p. and 0.67 p.p. increases in forecasts, against baseline averages of 5.46% and 6.51%, respectively (Table 2.2, Wave 1).¹² Health adversity is instead not associated with pessimism; its coefficient is insignificant (and directionally negative for inflation), differently from Table 2.1. Appendix B.3 reports estimated coefficients for all covariates.

Our model explains the effect of financial adversities as reflecting their spontaneous retrieval and use for simulating high inflation or home price growth scenarios. It can explain

3.93% (3%), respectively, which are comparable to Table 2.2 for inflation but lower for home price growth. DHS home price growth forecasts are elicited from a smaller sample (typically 1,600, compared to 2,500 for DHS inflation forecasts, and 3,800 in the survey reported here) and are more volatile across surveys.

¹² Effect sizes in p.p. are obtained by multiplying the z -score coefficient by the loading of the standardized point forecasts on the composite (0.83 for inflation, 0.85 for home price growth) and by the standard deviation of point forecasts in p.p. (8.40 p.p. for inflation, 7.55 p.p. for home price growth). Section 2.5 reports a comprehensive quantification.

Table 2.3: Experience Effects and Macroeconomic Beliefs

	Control Group, Wave 1	
	Inflation	Home Price Growth
	(1)	(2)
Finance Experience	0.1815*** (0.0552)	0.1048** (0.0515)
Health Experience	−0.0177 (0.0559)	0.0523 (0.0513)
<i>N</i>	1,284	1,284
Controls	Yes	Yes
<i>R</i> ²	0.1280	0.0595

Notes: This table reports OLS regressions of standardized inflation beliefs in column (1) and home price growth beliefs in column (2) for the control group in Wave 1, on binary indicators for having experienced a financial (“Finance Experience”) or health (“Health Experience”) adversity. Dependent variables are composite indices that average the standardized point forecast (winsorized at the 2nd and 98th percentiles) and threshold measure, then standardized within wave. Controls include age, education fixed effects (base: no vocational/university education), occupation fixed effects (base: unemployed), urbanization, net income, housing tenure fixed effects (base: tenant), gender, relationship status, and beliefs about elderly (not shown). Robust standard errors in parentheses. Significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

the weaker role of health adversities in Table 2.3 due to either: (i) the failure of spontaneously recalling them when thinking about inflation or home price growth (namely $r_{e_h}(\omega_T, \mathbf{n}) \approx 0$ in equation (2.4)), or (ii) their limited simulation potential (namely $\left[\sigma_T(e_h) - \overline{\sigma_T(e_h)}_{-h}\right] \approx 0$). Because the DHS measures current events, which are more easily recalled, the positive coefficient on health adversities in Table 2.1 supports simulation from e_h and hence hypothesis (i). These hypotheses and the roles of cues and experiences can be systematically studied using our model and the measured memory parameters.

2.4.3 Measured Memory Parameters

Experiences. The primed NDS experiences are frequent, with financial or health adversity being reported by 39% and 45% of participants in fR and hR , respectively. Frequent experiences in Control are also divorce (47%), job loss (37%), and being cheated on a contract (37%). These frequencies are similar across treatment and control groups (Appendix B.3). There is large variation in experiences across people, which leads to belief disagreement, and clustering: financial and health troubles, unemployment, and divorce are linked via interactions and personal traits (Appendix B.2). Priming causally tests memory effects by exogenously varying the salience of these experiences in context, controlling for their presence in the database.¹³

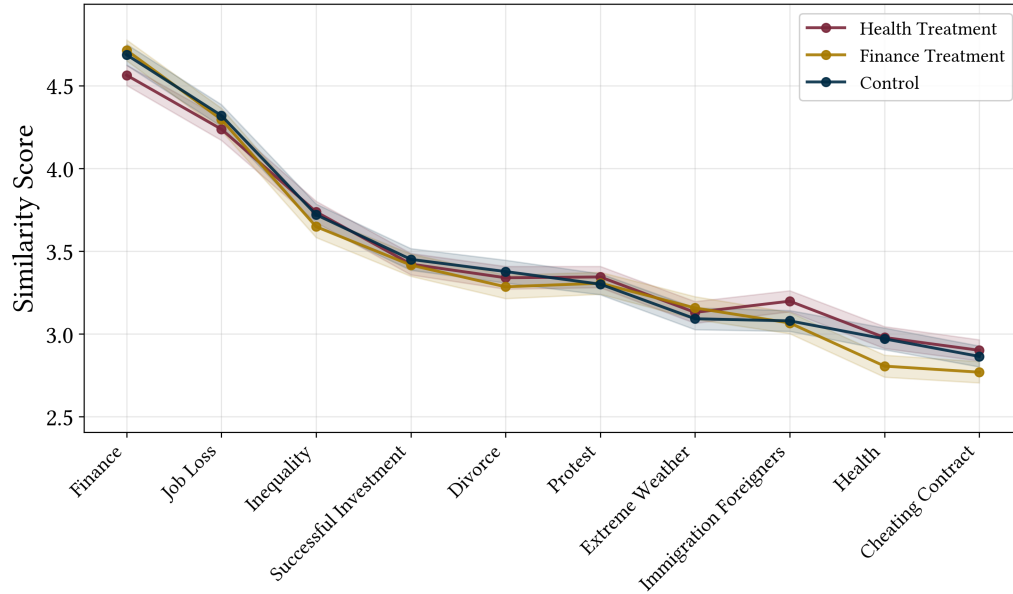
Similarity. Figure 2.1 plots the mean and 95% confidence interval of respondents' similarity scores of each NDS experience to high inflation (Panel A) and high home price growth (Panel B). A score of 1 means “not similar at all”; a score of 7 means “extremely similar.” The data come from Wave 1 when available and Wave 2 otherwise. Intuitively, *financial difficulty* is most similar to both high inflation and home price growth, with average scores around 4.5. This reflects the financial nature of inflation and home price growth, so there is some contextual similarity, but also the adverse nature of high inflation and home price growth for many people, creating similarity in valence. Financial adversity is about equally similar to inflation and home price growth (4.93 vs. 4.85, $p = 0.09$) and likewise for health (2.95 vs. 2.92, $p = 0.67$, see Appendix B.3).

Illness or Accident is among the least similar NDS adversities on our list, with average scores around 2.7–3.0 for inflation and home price growth. These NDS experiences are seman-

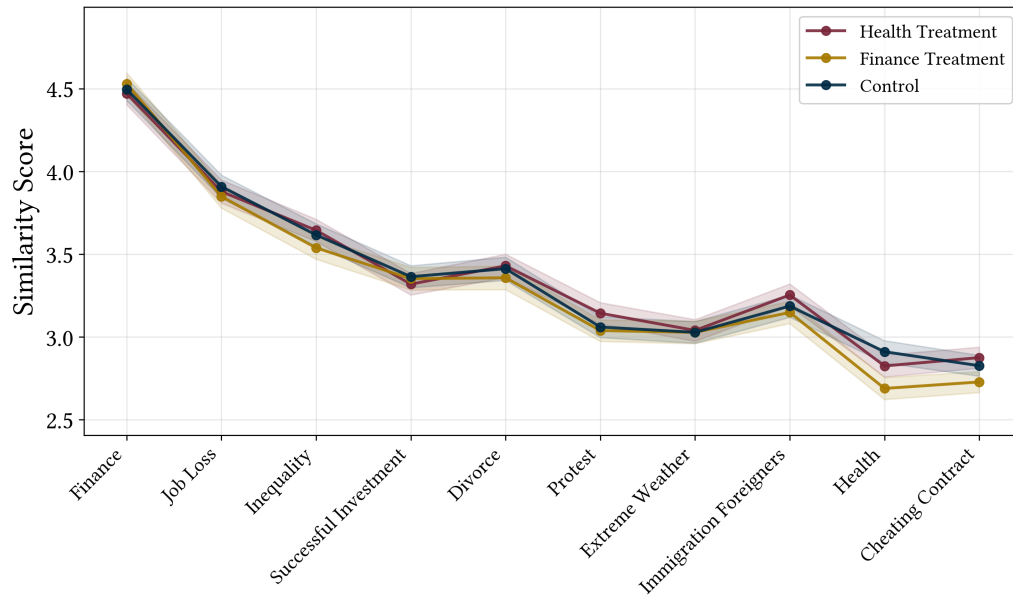
¹³ In the model, priming is further amplified by correlated experiences: a person primed with a health adversity may also recall associated financial or family adversities, fostering spillovers in equation (2.6).

Figure 2.1: Perceived Experience Similarity to High Inflation and High Home Price Growth

Panel A: Perceived Similarity to High Inflation



Panel B: Perceived Similarity to High Home Price Growth



Notes: This figure shows subjects' perceived similarity between each elicited non-domain-specific (NDS) experience and the scenarios of high inflation (Panel A) and high national home price growth (Panel B). Results are shown separately by treatment group: Health Treatment (*hT*), Finance Treatment (*fT*), and Control. Similarity is measured on a 1–7 scale, where 1 indicates “not similar at all” and 7 indicates “extremely similar.” Points show experience-level means; the shaded areas denote 95% confidence intervals. Data are pooled across both survey waves.

tically very different from macroeconomic states. However, similarity is non-zero: an illness is an adversity, just like high inflation and home price growth for many people. As in Assumption 1, context and valence shape similarity. The fact that financial and health adversities are perceived as similar to high inflation and home price growth is consistent with $\sigma_T(e) \leq 0$ in Assumption 2. Priming these adversities may then increase expectations of inflation and home price growth.

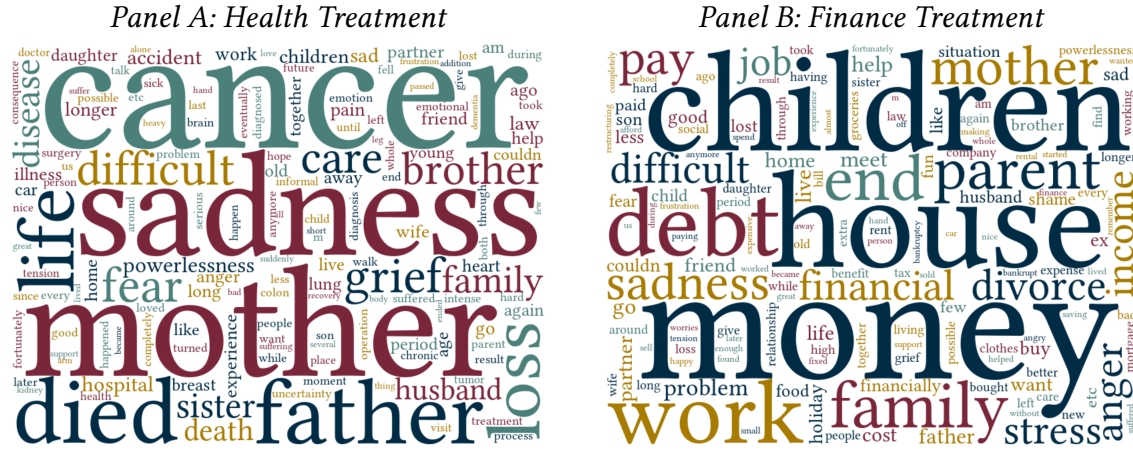
Similarity ratings are stable across income brackets, home ownership status, and our measures of lived experiences, confirming that similarity and hence simulation do not reflect individual tastes. Similarity is also stable across treatments, suggesting that fR and hR mainly affect cues, not similarity or simulation (see Appendix B.3). There is however significant variation in similarity ratings across people, with standard deviations ranging from 1.6 to 1.8 for different experiences. This is not only noise: a person's similarity rating has a 0.3 correlation across Waves 1 and 2.¹⁴ In psychology, variation in perceived similarity is tied to differential attention to features (Nosofsky 1986; Bordalo et al. 2026a). For instance, people more focused on emotions than on context perceive higher similarity between personal and macroeconomic adversities.¹⁵ We use this variation to analyze heterogeneity in treatment effects.

Cues. Priming affects beliefs by increasing the salience of adverse financial or health experiences. To validate this assumption, Figure 2.2 reports word clouds of subjects' descriptions of primed experiences, which prominently feature the primed adversity and its low valence.

¹⁴ Estimating for each t_T a pooled regression of Wave 1 on Wave 2 similarity scores and demographics predicts an interpersonal variation of 0.59 and 0.62 standard deviations of similarity in Π and Γ , respectively.

¹⁵ Having had an experience tends to increase its perceived similarity to inflation (Appendix B.3), which may reflect the increased ability to simulate negative affect events based on such experiences.

Figure 2.2: Word Clouds for Priming Treatments



Notes: Word clouds of open-ended experience descriptions in *hR* ($N = 1,262$; left panel) and *fR* ($N = 1,267$; right panel). Word size reflects frequency after removing common function words (175 stopwords). Total corpus: 41,776 words in *hR* and 31,760 words in *fR*.

2.5 Context, Memory, and Macroeconomic Expectations

We use our survey to test and quantify memory as a source of “animal spirits”: namely expectations that depend on uninformative mental states and emotions. Section 2.5.1 studies priming effects and compares them to experience effects. Section 2.5.2 studies the role of non-primed NDS experiences and of similarity. Section 2.5.3 studies how priming changes the way in which people reason about inflation and home price growth.

2.5.1 Instability: The Causal Effect of Primed Recall on Beliefs

We test for priming effects building on equation (2.7). Let dummy $\mathbb{1}_{i,e_p,R}$ take the value of one if respondent i was in treatment pR . The cue in task t_T is $q_{e_p,T}$, which is concentrated on experience e_p compared to the baseline distribution of cues ω_T . We then have:

Prediction 1. *In task t_T , the regression pooling the control and the treatment samples:*

$$z_{i,T} = \alpha_T + \sum_{p=f,h} \varepsilon_{e_p} \cdot \mathbb{1}_{i,e_p} + \sum_{p=f,h} \beta_{e_p,R} \cdot \mathbb{1}_{i,e_p,R} + \eta_T \cdot X_i + u_{i,T}. \quad (2.8)$$

has the following properties:

- (i) *If an experience effect is positive, $\varepsilon_{e_p} > 0$, then the corresponding priming effect is also positive, $\beta_{e_p,R} > 0$.*
- (ii) *The sum of experience and priming effects $\varepsilon_{e_p} + \beta_{e_p,R}$ is higher for an experience e_p that is more similar to high target outcomes.*

Relative to equation (2.7), equation (2.8) adds the second sum, namely the effect of being randomly assigned to a treatment in which adversity e_p is primed. Coefficient ε_{e_p} is the pure experience effect $\mathbb{E}(\tau \mid \omega_T, \mathbf{n}_p) - \mathbb{E}(\tau \mid \omega_T, \mathbf{n})$ in equation (2.4), coefficient $\beta_{e_p,R}$ is the priming effect $\mathbb{E}(\tau \mid \mathbf{1}_{e_p}, \mathbf{n}_p) - \mathbb{E}(\tau \mid \omega_T, \mathbf{n}_p)$ in equation (2.5).¹⁶ By Proposition 1, we obtain Prediction 1.(i): priming financial adversities, which in Table 2.3 exhibit positive experience effects, should raise inflation and home price growth expectations by prompting recall of similar simulation events.

Prediction 1.(ii) highlights the role of similarity: people who are primed for a lived experience more similar to high inflation or home price growth should be more pessimistic in these domains than people who have top-of-mind experiences that are less similar.¹⁷ Based on Table 2.3 and on the similarity gradient in Figure 2.1, the sum of experience and priming

¹⁶ These coefficients depend on the task T , but the subscript is omitted for simplicity of notation.

¹⁷ The model does *not* predict that the priming effect alone should be stronger for the more similar experience, because the latter may be spontaneously remembered with high probability, possibly leading to small priming effects.

effects should then be higher for financial than for health adversities for both inflation and home price growth.

Table 2.4 tests these predictions. Columns (1) and (2) consider an Intent to Treat (ITT) sample, so dummy $\mathbb{1}_{i,e_p,R}$ is equal to one also for subjects assigned to the treatment who report not having had the primed experience (and who do not answer recall questions). Columns (3) and (4) implement a Treatment on Treated (TOT) test, by setting $\mathbb{1}_{i,e_p,R} = 1$ only for subjects who had the primed experience and hence are effectively treated. The TOT tests implement the decomposition of equation (2.6) and are the relevant treatment for assessing Prediction 1.(ii). In the ITT sample coefficient magnitudes also reflect the frequency of experiences, not only their similarity.

Prediction 1 is confirmed. Consider Prediction 1.(i): priming a financial adversity boosts pessimism about inflation and home price growth in both the ITT and TOT tests, consistent with the positive experience effect in Table 2.3 and Table 2.4. Beliefs of finance-primed TOT respondents are 0.14 standard deviations higher for inflation and 0.17 for home price growth compared to people who have lived but are not primed to recall the experience. Priming health is also associated with more pessimism, especially in the TOT sample, equal to 0.14 standard deviations for inflation and 0.13 for home price growth. These effects are large, corresponding to increases in inflation forecasts of 0.97 p.p. (1.01 p.p.), against a baseline average of 5.46%, and an increase in home price growth forecasts of 1.08 p.p. (0.86 p.p., respectively), against an average of 6.51%.¹⁸

¹⁸ Consistent with the model, the treatment effects are concentrated in TOT (and hence weaker in ITT). For participants who did not have the experience, the treatment is milder (after being asked if they had the experience, they are not asked to recall it), and Appendix B.4 shows that such treatment has no effect on beliefs. Also consistent with the model and the role of cues in recall, the priming dummy does not absorb the experience effect (or vice versa).

Table 2.4: Effects of Cued Financial and Health Adversity on Macroeconomic Beliefs

	ITT		TOT	
	Inflation	Home Price Growth	Inflation	Home Price Growth
	(1)	(2)	(3)	(4)
Finance Prime	0.0640* (0.0376)	0.0707* (0.0375)	0.1388** (0.0653)	0.1690*** (0.0633)
Health Prime	0.0418 (0.0368)	0.0966** (0.0381)	0.1440*** (0.0524)	0.1347** (0.0554)
Finance Experience	0.1850*** (0.0338)	0.1275*** (0.0347)	0.1565*** (0.0464)	0.0795* (0.0454)
Health Experience	−0.0132 (0.0315)	0.0339 (0.0335)	−0.0704 (0.0479)	−0.0019 (0.0468)
<i>N</i>	3,813	3,813	2,367	2,367
Controls	Yes	Yes	Yes	Yes
<i>R</i> ²	0.1153	0.0595	0.1254	0.0641

Notes: This table reports OLS regressions of the impact of *fR* and *hR* on the *z*-scores of inflation and home price growth beliefs. Columns (1) and (2) present ITT estimates; columns (3) and (4) TOT ones. “Finance Experience” and “Health Experience” indicate if a respondent lived a financial or a health adversity. Dependent variables are composite indices that average the standardized point forecast (winsorized at the 2nd and 98th percentiles) and threshold measure and then standardized within wave. All regressions use Wave 1 data; the TOT sample is restricted to respondents who complied with treatment. Controls include age, education fixed effects (base: no vocational/university education), occupation fixed effects (base: unemployed), urbanization, net income, housing tenure fixed effects (base: tenant), gender, relationship status, and beliefs about elderly (not shown). Robust standard errors in parentheses. Significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Prediction 1.(ii) is also confirmed. The financial adversity is more similar than the health adversity to higher target outcomes and, consistent with memory, it exhibits larger combined experience and priming effects in both domains, corresponding to 0.30 vs. 0.07 standard deviations for inflation and 0.25 vs. 0.13 for home price growth. The similarity hierarchy of these results in turn ties the hierarchy of experience effects in Table 2.3.¹⁹

¹⁹ As noted in Section 2.4, both financial and health adversities are approximately equally similar to high inflation and home price growth, and their combined effects are not statistically different across domains, in line with the model.

Exogenous recall of uninformative events affects macroeconomic beliefs, rejecting statistical models.²⁰ Also inconsistent with motivated beliefs, people imagine aggregate states that are similar to what is top of mind, not what is fundamentally good or bad for them. For instance, priming financial adversities increases expected home price growth for both homeowners and renters, consistent with the similarity judgments of both groups (Appendix B.4).

2.5.2 Heterogeneity: The NDS Experiences Database and Similarity

Memory creates systematic heterogeneity. Some people are more pessimistic than others because: (i) they have lived and are reminded of adverse personal experiences that help them imagine adverse macroeconomic states, or because (ii) they perceive stronger similarity between adverse NDS and bad macroeconomic states (due to higher attention to valence than to context). Such disagreement can manifest itself both in the baseline and in reaction to cues.

To test for these predictions, we capture the “pessimism-potential” in a subject’s NDS database and similarity perceptions, by computing, following Bordalo et al. (2025a), the average similarity of the NDS experiences the subject reports:

$$S_{T,i}(\mathbf{n}_i) = \frac{\sum_{e \in E} \mathbb{1}_{i,e} \cdot S_i(e, T)}{\sum_{e \in E} \mathbb{1}_{i,e}},$$

where E is the set of NDS we elicit, $\mathbb{1}_{i,e}$ is as before a dummy for whether the respondent had experience e , and $S_i(e, T)$ is the respondent’s stated similarity of e to high inflation ($T = \Pi$) or home price growth ($T = \Gamma$). We set $S_{T,i}(\mathbf{n}_i) = 0$ for respondents who have not had any of the elicited experiences. A respondent with a higher index $S_{T,i}(\mathbf{n}_i)$ has lived NDS experiences

²⁰ In line with recall being shaped by the current cue, priming in Wave 1 has low persistence to Wave 2 (Appendix B.4).

more similar to high target outcomes, and so should exhibit higher beliefs.

To capture similarity in driving heterogeneous priming effects we construct the proxy:

$$S_{T,i}(e_p) = \mathbb{1}_{i,e_p} \cdot \mathbb{1}_{i,e_p,R} \cdot S_i(e_p, T),$$

which is non-zero for TOT subjects, who have $\mathbb{1}_{i,e_p} \cdot \mathbb{1}_{i,e_p,R} = 1$, and increases in the similarity of the primed experiences with high inflation or home price growth. Higher $S_{T,i}(e_p)$ means that, when primed to recall e_p , the person better simulates adverse scenarios, leading to a stronger increase in inflation and home price growth beliefs. From equation (2.8), Prediction 1, we then obtain:²¹

Prediction 2. *Given task t_T , the regression*

$$z_{i,T} = \alpha_T + \varepsilon_S \cdot S_{T,i}(\mathbf{n}_i) + \beta_{e_p,S} \cdot S_{T,i}(e_p) + \eta_T \cdot X_i + u_{i,T}, \quad (2.9)$$

detects a positive similarity gradient of NDS experiences $\varepsilon_S > 0$ and their priming $\beta_{e_p,S} > 0$.

Table 2.5 reports the estimates. To interpret magnitudes, we scale each index by its standard deviation within each wave. Consistent with the model, coefficients are positive: experience and priming effects increase in perceived similarity. More similar experiences are more likely to be spontaneously retrieved and used in simulation, and the effect of priming them is also stronger.²²

²¹ As in equation (2.8), all coefficients depend on the task T , but the subscript in ε_S and $\beta_{e_p,S}$ is omitted for simplicity.

²² To study memory interference, the average similarity index is decomposed into its numerator and denominator: the numerator captures the total summed similarity across lived NDSs, the denominator is the total number of NDSs. Appendix B.4 shows that the numerator enters positively and the denominator negatively: holding total similarity constant, many lived NDSs interfere, reducing the effect of similar experiences on beliefs.

Coefficient magnitudes can be read directly as the effect of a 1-standard-deviation increase in similarity on standardized beliefs. The effects are sizable: a 1-standard-deviation increase in average similarity of lived experiences raises inflation (home price growth) beliefs by 0.10 (0.07) SD. A 1-standard-deviation increase in the similarity of the primed experience increases inflation (home price growth) beliefs by 0.06 (0.06) SD. Adding our two proxies increases the regressions' R^2 from 0.114 to 0.128 for inflation, and from 0.055 to 0.066 for home price growth, compared to having only controls.

Table 2.5: Effect of Experience Similarity on Inflation and Home Price Growth Beliefs

	TOT	
	Inflation	Home Price Growth
	(1)	(2)
Avg. Similarity Lived Experiences	0.0956*** (0.0255)	0.0689*** (0.0234)
Similarity Primed Experience	0.0560*** (0.0173)	0.0600*** (0.0184)
N	2,367	2,367
Controls	Yes	Yes
R^2	0.1278	0.0664

Notes: This table reports TOT estimates of OLS regressions of z -scores of inflation (column 1) and home price growth (column 2) beliefs in Wave 1 on “Avg. Similarity Lived Experiences” (the average similarity rating across all experiences lived by the respondent, rescaled by the unconditional standard deviation pooled across all experiences) and “Similarity Primed Experience” (the similarity rating for the primed experience, rescaled by the same unconditional SD). Respondents in the control group are assigned a value of zero. Dependent variables are composite indices that average the standardized point forecast (winsorized at the 2nd and 98th percentiles) and threshold measure, then standardized within wave. Controls include age, education fixed effects (base: no vocational/university education), occupation fixed effects (base: unemployed), urbanization, net income, housing tenure fixed effects (base: tenant), gender, relationship status, and beliefs about elderly (not shown). Robust standard errors in parentheses. Significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

One concern is that individual similarity may be spuriously anchored to beliefs (e.g., people reporting higher beliefs may anchor to higher numbers when subsequently reporting similarity). We rule out this possibility by showing that our results are robust to instrumenting

individual similarity or beliefs across waves, which rules out anchoring to noise. Our results are also robust to using average rather than individual similarity ratings, which rules out anchoring of similarity to stable belief differences. Appendix B.4 reports several other robustness checks.²³

Quantification of Effects. We saw that having lived financial or health adversities and their current salience affects beliefs (Table 2.4), and that additional belief heterogeneity arises from the broader NDS database and similarity judgments (Table 2.5). Figure 2.3 collects these quantitative effects and compares them to standard demographics and experience effects.

To quantify the role of the NDS database, consider the distribution of $S_{T,i}(\mathbf{n}_i)$ computed at average similarity. We first estimate the impact of a given NDS experience on beliefs by adding it to every database and computing the entailed change in the average (across participants). Using Table 2.5, adding a job loss experience increases inflation (home price growth) forecasts by 0.18 p.p. and 0.10 p.p., respectively (Appendix B.4). This is interesting: the real-world intensity of job loss (experienced directly or by acquaintances) varies across macroeconomic states, creating aggregate context effects. We next consider the interquartile range of the distribution of $S_{T,i}(\mathbf{n}_i)$ at average similarity. By construction, variation arises only from different NDS experiences. Such interquartile range corresponds to a 0.74 p.p. and 0.91 p.p. increase in inflation and home price growth beliefs, respectively. Importantly, these results are not driven only, or even mainly, by finance experiences: other NDS experiences have comparable similarity-weighted impacts on beliefs (Appendix B.4).

²³ Table 2.5 is also robust to using the ITT sample, which includes all subjects in the treatments, and to excluding demographic controls (see Appendix B.4), suggesting that the effect of memory parameters is largely independent of demographics. Allowing the similarity proxies to also enter quadratically (i.e., giving higher weight to experiences deemed more similar) yields comparable coefficients. Results are consistent across the finance and health treatments.

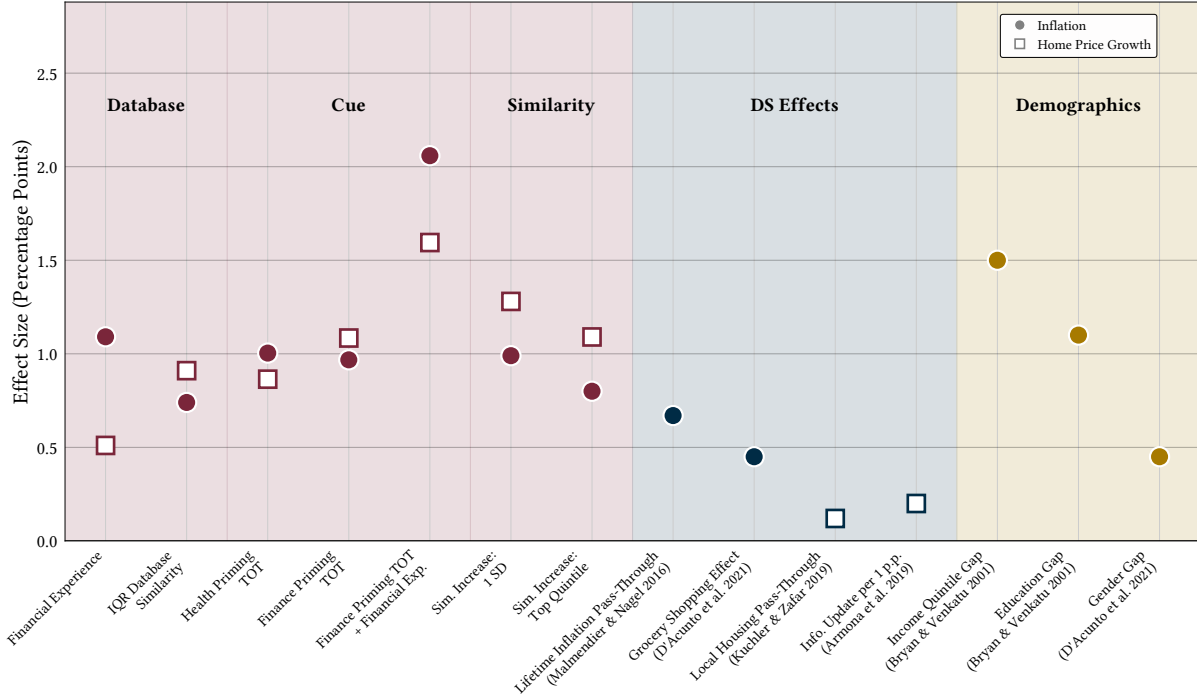
Consider similarity next. Increasing the average similarity of experiences by one standard deviation of similarity increases inflation (home price growth) forecasts by 0.99 p.p. (1.28 p.p.).²⁴ Individual-level similarity is also important: increasing it from the average to the top quintile for each experience increases inflation (home price growth) forecasts by 0.80 p.p. (1.09 p.p.). Even in a regression setting, individual similarity matters controlling for average similarity (Appendix B.4).

These quantifications are summarized in Figure 2.3, which distinguishes heterogeneity due to different databases, similarity, and instability from cues. Effects are large compared to standard DS experience effects and demographics. Priming a finance adversity quantitatively matches the disagreement observed when moving from the top to the bottom income quintile (Bryan and Venkatu 2001; Weber et al. 2022). Demographics may partly capture systematic differences in experiences and the extent to which they are top of mind when making macroeconomic forecasts. The same mechanism extends to DS experiences: the low valence of higher gas prices may be more attended to by a low-income household, causing via similarity a stronger experience effect.

In sum, memory offers a cognitively founded mechanism for generating sizable and realistic heterogeneity and instability of macroeconomic expectations based on non-informative experiences that influence the ability to imagine different macroeconomic states. We next show how this mechanism for “animal spirits” can offer a foundation for models of beliefs tainted by momentary emotions (Schwarz and Clore 1983; Loewenstein 2000) or partial models/narratives (Shiller 2019a; Schwartzstein and Sunderam 2021; Andre et al. 2024).

²⁴ The quantification regresses z -scored beliefs on the z -standardized (by cross-person SD) average similarity of lived experiences. A 1-standard-deviation increase in similarity refers to the pooled SD of raw similarity ratings across experiences and respondents. The pooled SD is about twice the SD of the person-level similarity across experiences, so the implied effects are larger than a naive multiplication of Table 2.5 coefficients.

Figure 2.3: Quantification of Memory Effects



Notes: This figure compares the estimated effect sizes with benchmark effects from the literature. Effects are expressed in percentage points (p.p.) of expected inflation (filled circles) or home price growth (open squares), converted from the composite z -score index using factor loading $\times$ standard deviation (inflation: 0.83×8.40 ; home price growth: 0.85×7.55). This paper's estimates (red) include: (i) Database effects, which are the coefficient on having a financial experience and the effect of moving from the 25th to 75th percentile of the NDS similarity distribution (IQR); (ii) Cue effects, which are treatment on treated (TOT) effects of health and finance priming, and the combined effect of finance priming plus financial experience; (iii) Similarity effects, which are the impact of a 1-standard-deviation increase in similarity ratings and the effect of moving to the top quintile of similarity. All estimates are from OLS regressions on the TOT sample ($N = 2,367$) in Wave 1, controlling for demographics. Literature benchmarks (blue) include lifetime inflation pass-through (Malmendier and Nagel 2016), local house price pass-through (Kuchler and Zafar 2019), grocery shopping effects (d'Acunto et al. 2021), and information updating (Armona et al. 2019). Demographic gaps (purple) show differences by income quintile, education, and gender from Bryan and Venkatu (2001), Weber et al. (2022), and d'Acunto et al. (2021). Effects are grouped into categories indicated by shaded background.

2.5.3 Priming Changes Reasoning about Forecasts

Mnemonic reasoning uses the features of personal experiences analogically. First, negative emotions from personal adversities cue recall of high past inflation, a DS experience, causing high expected inflation, as in Taubinsky et al. (2024). In addition, the agent can retrieve an NDS

adversity and form a partial model (e.g., “things can go bad”) that invites pessimism. Consistent with this mechanism, Andre et al. (2022) show that recall of personal experiences, including NDS ones, creates heterogeneous “macroeconomic models” causing disagreement based on the same shock. In our model, primed recall of e_p causes these different lines of reasoning by fostering recall of DS and NDS experiences that are similar to the cue $q_{e_p, T}$. Thus, we should observe treatment effects in inflation and home price growth reasoning that reflect the features (x, v, c) of experiences retrieved based on the prime, in turn causing pessimistic beliefs.²⁵

Priming Effects on Retrieved (x, v, c) . We proxy for the realization τ of recalled DS inflation and home price growth experiences using our elicited 12-month backcasts of each. Schacter et al. (2012) show that the same selective recall mechanism shapes the reconstruction of the past and the imagination of the future. In our model, when a recalled inflation or home price growth DS episode is top of mind, it influences both the forecast and the backcast. For NDS we only observe the realization x of primed experiences (i.e., the specific financial trouble or disease in e_p) but in our model the agent does not use this realization in imagining future inflation or home price growth, so we do not use this data in our analysis.

We then extract the key drivers of simulation, valence v and context c , from self-reported forecasts’ reasons. We use ChatGPT, fine-tuned with detailed instructions and manual validation, to categorize each text as follows. First, proxying for c , we categorize whether reasoning primarily relies on (i) personal experiences (e.g., household hardship, personal finances, or family events), on (ii) news or expert sources (e.g., media reports, government policy, central bank actions), or is (iii) unclassified. Category (i) captures the personal context of NDS experiences, category (ii) the domain-specific context of DS ones. Second, proxying for v , we classify

²⁵ Appendix B.5 also studies the similarity between reasoning and recall, as a proxy for simulation. Consistent with the model, fR (resp. hR) increases the similarity of reasons to recalled financial (resp. health) experiences.

texts based on whether they reflect negative or positive valence, are factual, or unclassified. Negative valence reflects the DS and NDS adversities shaping pessimistic forecasts.²⁶

Figure 2.4 shows word clouds for reasoning classified as negative vs. neutral valence, showing that the latter emphasizes aggregate quantities (interest rates, supply, and demand) while the former emphasizes negative valence, often in the context of personal experiences (less, money, meet ends, buy). Similarly, Figure 2.5 shows word clouds for reasoning classified as personal vs. news-based. Appendix B.5 lists exemplar reasons of all classifications.

Our model makes the following prediction about these proxies:

Prediction 3. *Priming an adverse NDS experience should increase the relative recall of adverse DS experiences compared to favorable DS experiences, and the recall of NDS experiences compared to DS ones, all compared to Control. As a result, priming should increase:*

- (i) *high inflation and home price growth backcasts (DS effect).*
- (ii) *the incidence of personal logic (NDS effect) and low valence in reasons (DS and NDS effect).*

Effect (i) follows from mood-congruent recall (Bower 1981) of DS experiences: priming an NDS adversity favors retrieval of high compared to low inflation and home price growth states due to similar low valence, as in Taubinsky et al. (2024). Effect (ii) is an additional memory effect: priming an adverse NDS experience prompts recall of other adverse NDS experiences, increasing simulation and reasoning based on personal context and low valence.

Columns (1) and (2) of Table 2.6 report the TOT effects on backcasts. Both fR and hR are associated with higher past inflation and home price growth, consistent with Prediction 3.

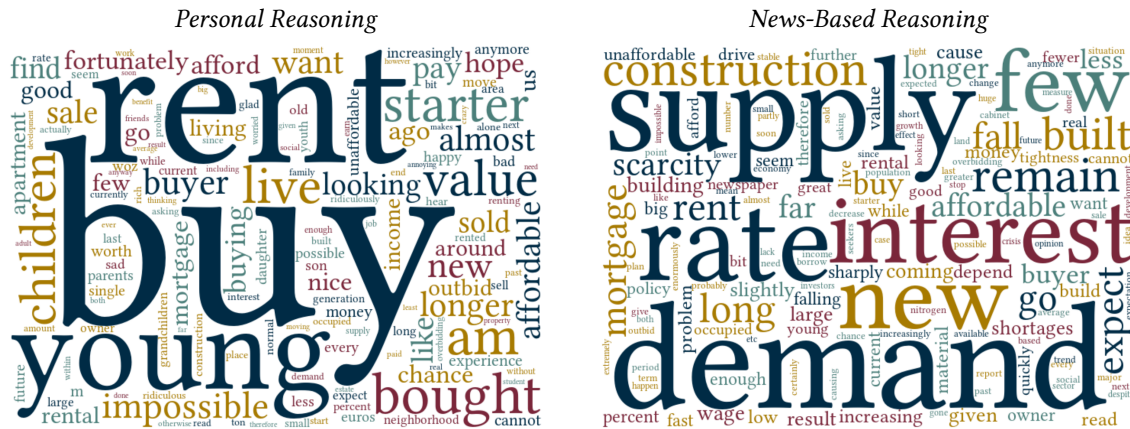
²⁶ The GPT classifications are validated against human-coded subsets, confirming high agreement rates (78% for inflation and 70% for home price growth, for personal logic; 85% for inflation and 92% for home price growth, for valence).

Figure 2.5: Word Clouds of Macroeconomic Reasoning by Type

Panel A: Inflation



Panel B: Home Price Growth



Notes: Word clouds of open-ended macroeconomic reasoning, comparing texts classified as expressing personal reasoning (left) vs. news-based reasoning (right). The top panel shows inflation reasoning; the bottom panel shows home price growth reasoning. Personal vs. news-based reasoning is classified using GPT-4 analysis of the English translations. Word size reflects frequency after removing common function words and shared topic words (e.g., “price,” “inflation,” “house”) to highlight differences between categories.

Columns (3) and (4) show the results for context and valence. Consistent with Prediction 3, priming an adversity increases the likelihood of reasons including either personal context or negative valence relative to Control, for both inflation and home price growth forecasts. The effects are large: finance priming raises the likelihood of personal/negative valence reasoning by 9.7 p.p. (inflation) and 12.0 p.p. (home price growth). The corresponding figures for health

Table 2.6: Backcasts, Context, and Valence

	Backcasts (%)		Personal or Negative	
	Inflation	Home Price Growth	Inflation	Home Price Growth
	(1)	(2)	(3)	(4)
Finance Prime	1.8253*** (0.5811)	1.2384** (0.5304)	0.0971*** (0.0269)	0.1195*** (0.0269)
Health Prime	0.9422* (0.4814)	0.3574 (0.3921)	0.0598** (0.0241)	0.0797*** (0.0236)
<i>N</i>	2,367	2,367	2,367	2,367
Controls	Yes	Yes	Yes	Yes
<i>R</i> ²	0.0512	0.0410	0.0365	0.0341

Notes: This table reports treatment on treated (TOT) regressions for Wave 1. Columns (1) and (2) estimate how treatments affect respondents’ backcasts of inflation and home price growth. Backcasts are expressed in percentage terms. Columns (3) and (4) estimate whether treatments increase the likelihood that respondents invoke personal experiences or express negative affect in their macroeconomic reasoning. These indicators are classified using GPT-4 analysis of open-text responses, where “Personal or Negative” equals one if the respondent’s reasoning is classified as using personal logic or expressing negative affect. Controls include age, education fixed effects (base: no vocational/university education), occupation fixed effects (base: unemployed), urbanization, net income, housing tenure fixed effects (base: tenant), gender, relationship status, and beliefs about elderly (not shown). Robust standard errors in parentheses. Significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

are 6.0 p.p. and 8.0 p.p. Appendix B.5 shows that priming NDS adversities raises the share of reasoning based directly on low-valence and personal experiences, and a switch away from news-based reasoning.²⁷

Recalled Features and Forecasts. We conclude by circling back to forecasts. Our model predicts that higher backcasts should be associated with higher forecasts. It also predicts that personal logic and low-valence reasoning based on NDS experiences should increase the forecasts, consistent with a greater reliance on pessimistic memory-based narratives (Shiller 2019a;

²⁷ Table 2.6 measures the disjunction “personal logic or negative valence” to allow for news-based forecasts that also feature low valence, in line with the retrieval of bad inflation or home price growth experiences from the adverse NDS.

Andre et al. 2022) that help imagine future bad states.²⁸ To test this link we use an instrumental variables (IV) approach. In the first stage, we use fR and hR to obtain exogenous variation in reasons as in Table 2.6. In the second stage, we regress macroeconomic beliefs on the first-stage predicted backcasts, negative valence or personal logic from the first stage.

Table 2.7 shows that the causal increase in backcasts or in the reliance on personal logic or negative valence substantially raises inflation and home price growth forecasts. The first-stage F -statistic is strong for reasoning, less so for backcasts, which is intuitive: The recall of low-valence or personal experiences is instead more strongly based on the NDS cue that our treatments induce, yielding a stronger first-stage effect. Put together, the effects of priming on reasoning and beliefs are large: higher predicted personal/negative reasoning raises inflation expectations by 1.5 p.p. and home price growth expectations by 1.3 p.p. under the finance prime.²⁹ For health, these numbers are 0.93 p.p. and 0.88 p.p. These are substantial magnitudes relative to Figure 2.3.

Consistent with memory, experiences, and cues drive forecasts by endogenously shaping the *content* of reasoning: experiences similar to the task, or cued by the environment, affect not only the selective retrieval of DS data but also the simulation of scenarios based on non-statistical features of irrelevant NDS experiences, such as their negative valence. We next

²⁸ The mechanism helps explain instability and heterogeneity in recall and beliefs. For example, Blanchard (1993) offers evidence consistent with the U.S. recession of 1990–91 being triggered by a drop in consumer confidence due to the Iraq War. Salient events such as a war can affect beliefs by helping people imagine good or bad states, in turn affecting consumption and investment. Link et al. (2023) show that people who think more about inflation tend to be more pessimistic about it. In the model, attention to inflation shapes the cue, which then leads to selective retrieval of high inflation outcomes. Memory also explains why beliefs tend to respond more to specific stories than to statistics (Kahneman and Tversky 1972, 1982), in part because stories are more memorable than experiences (as shown experimentally in Graeber et al. (2024)) and in part because they provide better simulation material.

²⁹ The p.p. conversion for Tables 2.6–2.7 uses the factor loading and standard deviation computed on the TOT Wave 1 subsample (0.84×8.73 for inflation; 0.86×7.52 for home price growth), which differ slightly from the values in earlier footnotes.

Table 2.7: Beliefs and Predicted Backcasts, Context, and Valence

	Backcast (Predicted)		Personal or Negative (Predicted)	
	Inflation	Home Price Growth	Inflation	Home Price Growth
	(1)	(2)	(3)	(4)
Predicted Variable	0.1257*** (0.0299)	0.1805*** (0.0440)	2.1061*** (0.5007)	1.7089*** (0.3766)
N	2,367	2,367	2,367	2,367
R^2	0.0083	0.0082	0.0081	0.0092
KP F-stat	5.86	3.08	9.39	15.65

Notes: This table reports second-stage instrumental variable (IV) regressions examining whether backcasts and reasoning features help explain contemporaneous belief formation. The dependent variables are standardized belief z -scores constructed from respondents’ point and threshold forecasts. Columns (1) and (2) use predicted backcasts (respondents’ retrospective assessments of past inflation or home price growth) as the endogenous regressor. Columns (3) and (4) use predicted “Personal or Negative” (a binary indicator equal to one if the respondent’s macroeconomic reasoning invokes personal experiences or expresses negative affect, as classified by GPT-4 analysis of open-text responses). The first-stage regressions are reported in Table 2.6, where each mechanism is regressed on the randomized treatment assignments (Health Prime and Finance Prime), which serve as excluded instruments. “KP F-stat” reports the Kleibergen-Paap first-stage F -statistic as a diagnostic for instrument strength. Each outcome is estimated separately. Robust standard errors in parentheses. Significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

outline some implications of this mechanism for macroeconomic expectations and outcomes.

2.6 Macroeconomic Implications

We showed that uncertainty is represented “in context” based on semantic or emotional cues that shape the ability to imagine good vs. bad states. We detect these context effects causally, in a clean way. Beyond the specific events that we prime, our findings have broader relevance for macroeconomics because they identify the belief-formation mechanism. They unify well-documented features of beliefs (Section 2.6.1) and yield new implications for confidence cycles (Section 2.6.2).

2.6.1 Unifying Features of Macroeconomic Expectations

Consider an agent forecasting an adversity, such as inflation. To see the intuition most starkly, consider a stylized database with a share n_Π of DS experiences of constant inflation $\bar{\pi}$, and a share $n_p = 1 - n_\Pi$ of NDS adversities.³⁰ As in Assumption 2, the agent simulates future inflation based on the valence of the retrieved experiences. DS experiences have valence $v(\bar{\pi}) = -\bar{\pi} < 0$, while NDS experiences have valence $v_p = v(\bar{\pi} + \gamma) = -(\bar{\pi} + \gamma)$ with $\gamma > 0$: realistically, personal adversities are more distressing for the agent than experienced inflation (the analysis easily generalizes to a richer setting with heterogeneous DS and NDS experiences). As in Assumption 2, her NDS-based simulation is weaker due to discrepant context: the agent simulates inflation $\sigma_p \cdot (\bar{\pi} + \gamma)$, where $-\sigma_p \in (0, 1)$ is the sensitivity of simulation to the valence of NDS.

A cue q is identified by its valence v_q and context c_q . Given similarities between q and DS and NDS experiences, $S_\Pi \equiv S(q, e_\Pi)$ and $S_p \equiv S(q, e_p)$, respectively, expected inflation is:

$$\mathbb{E}(\pi | q, \mathbf{n}) = \left[\frac{n_\Pi \cdot S_\Pi}{n_\Pi \cdot S_\Pi + n_p \cdot S_p} \right] \cdot \bar{\pi} + \left[\frac{\sigma_p \cdot n_p \cdot S_p}{n_\Pi \cdot S_\Pi + n_p \cdot S_p} \right] \cdot (\bar{\pi} + \gamma). \quad (2.10)$$

The first term is a DS effect: more DS experiences, higher n_Π , promote retrieval of inflation $\bar{\pi}$, anchoring the forecast to it. The second term is an NDS effect: personal financial troubles, higher n_p , prompt retrieval of distress and simulation of a high inflation state $\bar{\pi} + \gamma$.

Insensitivity. In equation (2.10), personal events interfere “bottom up” with the retrieval of DS experiences: when forecasting inflation, a person may recall her financial troubles or the prices she saw, neglecting aggregate DS data. Interference from NDS experiences thus causes

³⁰ For simplicity, good NDS are abstracted away, akin to setting their similarity in Assumption 1 to zero, $K = 1/2$.

insensitivity to DS ones: increasing experienced inflation $\bar{\pi}$ by one unit while leaving v_p constant increases expectations by less than one unit.

Bias. Interference can also cause expectations to be *above* experienced inflation $\bar{\pi}$. This occurs if NDS experiences are distressing enough, $\gamma > \left(\frac{1-\sigma_p}{\sigma_p}\right) \bar{\pi}$. Households' measured expected inflation is indeed too high (Mankiw et al. 2003; Coibion et al. 2022).

Insensitivity may seem consistent with rational inattention (Sims 2003), but bias is not. Bias is important, and reflects perceived similarity between inflation and NDS. Crucially, the same mechanism also explains why households strongly dislike inflation: due to similarity, high inflation reminds people of personal financial troubles, affordability concerns and negative-valence experiences.

DS Context and Overreaction. Our experiment causes instability of beliefs by cuing NDS adversities. Similar effects also arise for DS events: an increase in current inflation from $\bar{\pi}$ to $\bar{\pi}' = \bar{\pi} + \partial\pi$, $0 < \partial\pi < \gamma$, increases the similarity S_p of the inflation environment to personal distress and reduces its similarity S_{Π} to regular inflation $\bar{\pi}$. Forecasts then increase if and only if NDS adversities cause a positive inflation bias.³¹ This happens even if the inflation shock is transitory. Note that the agent does not have a general tendency to overreact to news: she is insensitive to historical inflation $\bar{\pi}$ and exhibits a constant NDS-driven bias γ . Yet, she overreacts after seeing a salient inflation surge, which cues her bottom up to retrieve distressing NDS financial troubles and interferes with historically more moderate DS events, increasing pessimism and causing overreaction especially if the change $\partial\pi$ is large, albeit transient.³²

³¹ The increase in current inflation reduces the valence distance in S_p from γ to $\gamma - \partial\pi$, and increases the valence, and numerical, distance in S_{Π} from 0 to $\partial\pi$. Equation (2.10) then implies $\frac{\partial \mathbb{E}(\pi | q, \mathbf{n})}{\partial \pi} \propto \left[\gamma - \left(\frac{1-\sigma_p}{\sigma_p}\right) \bar{\pi} \right] \cdot \frac{\partial}{\partial \pi} \left(\frac{S_p}{S_{\Pi}} \right)$.

³² As described earlier, this mechanism can produce overreaction even in a database with only DS experiences, as documented in Bonaglia et al. (2025) and Gennaioli et al. (2025). The model is more general and nests this effect

Narratives and Partial Models. Narratives have been invoked as a powerful driver of aggregate consumption and investment decisions using compelling examples and psychological intuition (Shiller 2019a). Our framework offers a foundation for narratives in terms of context-based beliefs and integrates these ideas into a testable and tractable macroeconomic framework.

For example, an increase in the salience with which a feature is discussed in the media can, by changing the cue q , shift which experiences come to mind and beliefs. Shiller (2019a) shows that stories about the “Great Depression” were more viral during the 2007–2008 than the 1981–1982 recession, even though unemployment was higher in 1981–1982. The salience of the “financial crisis” feature increased contextual similarity between 2008 and 1929 even though the Great Depression was objectively not very informative. Superficial similarity drives incorrect DS analogies.

The same mechanism may be at play in narratives that leverage everyday experiences, such as “restaurants are full, the economy is doing well.”³³ These narratives focus attention on familiar experiences, and crowd out potentially more relevant information due to interference. The resulting beliefs are based on a subset of the person’s DS or NDS experiences, which highlight specific features or channels, i.e., they are partial models. The “partial model” used by a person is then predicted by her experiences, similarity judgments, and context. This leads to systematic heterogeneity: consumers tend to adopt narratives that emphasize experiences familiar to them (Andre et al. 2022; Coibion et al. 2022).

Lastly, narratives are unstable. At the heart of the literature on narratives is the idea of

if realized inflation π is allowed to vary across DS traces in the database. A characterization of when interference by NDS generates over or underreaction is left to future work.

³³ This example is based on a famous quote by former Italian prime minister Silvio Berlusconi.

contagion (Shiller 2019a; Flynn and Sastry 2025), an abrupt change in beliefs driven not by fundamentals but by exposure to the narrative, which is a form of context. By its nature, such contagion is hard to square with models of stable belief distortions. Our model explains how context can drive instability, by changing which experiences underlie beliefs. In particular, the memory mechanism disciplines how changing context shapes the narrative, also putting limits to contagion, for instance based on heterogeneity of experiences. This opens the door to the study of dynamics of beliefs, to which we turn next.

2.6.2 Heterogeneity, Confidence, and Business Cycles

Cross-Sectional Heterogeneity. Variation in NDS experiences creates belief heterogeneity. This variation does not wash out in the aggregate because NDS experiences (and contexts) are not randomly assigned. Specifically, and consistent with Das et al. (2020), people with worse personal financial experiences are *ceteris paribus* more pessimistic about the future. This dampens their willingness to consume, invest, or search for a job. Critically, these very people may have the highest “rational” marginal propensity to consume (MPC) and job search incentives, so their pessimism can stunt an important driver of economic recovery during a recession. Through these nonlinearities, then, memory-driven belief heterogeneity can affect aggregate outcomes.

Aggregate Confidence. A second and direct link between mnemonic context effects and the business cycles arises through endogenous business confidence. A large body of work suggests that consumer and investor confidence are powerful predictors of aggregate expansions and contractions (Ludvigson 2004), and typically traces them to changes in beliefs triggered by aggregate statistical signals. Critically, memory-driven confidence does not rely on statis-

tical signals; it can also arise from current context, including the endogenous actions of other agents. This creates new belief externalities. For instance, the decision of a loan officer to deny credit to a firm increases its macroeconomic pessimism even absent information, which reduces the firm's investment over and above the amount of credit that was denied.

The same mechanism can also work in a positive direction: the decision of the government to purchase more goods from firms cues economic expansion, increasing these firms' optimism, consistent with Bellifemine et al. (2025). To see this mechanism at play, suppose that the current optimal action x_0 increases in the expectation about the future value x_1 of the action itself:

$$x_0 = \mathbb{E}(x_1 | x_0, \mathbf{n}). \quad (2.11)$$

For instance, current consumption x_0 should increase in the expectation of future consumption, as in the permanent income hypothesis (Hall 1978), and current investment should increase in the expectation of future marginal return, as in q-theory (Hayashi 1982).

With memory, the expectation $\mathbb{E}(x_1 | x_0, \mathbf{n})$ depends on the current equilibrium outcome x_0 , over and above its objective information (if any) about x_1 . The reason is that higher current consumption x_0 cues recall of economic expansions or reduces recall of past adversities, increasing expected future consumption. This crucially implies that equation (2.11) entails feedback that further increases current consumption x_0 . The result is a “confidence multiplier”: an increase in confidence creates a current expansion that cues a favorable context, which in turn supports confidence and an additional expansion.

To see this formally, we apply equation (2.10) to positive events (i.e., context and valence are now positively related). Combining it with equation (2.11), setting for simplicity $\sigma_p = 1$

and $S(e, q) = \exp[-\delta|x_q - x_e|]$, we obtain an equilibrium condition for x_0 :

$$x_0 = \bar{x} + \frac{n_p \cdot e^{-\delta[|x_0 - \bar{x} - \gamma| - |x_0 - \bar{x}|]}}{n_T + n_p \cdot e^{-\delta[|x_0 - \bar{x} - \gamma| - |x_0 - \bar{x}|]}} \cdot \gamma. \quad (2.12)$$

The last term captures the distortion due to NDS-based simulation. It has a sigmoidal shape so that, under some parameter values, equation (2.12) admits two stable equilibria, x_L^* and x_H^* , at low and high consumption, respectively. The presence of confidence multipliers creates self-fulfilling booms, associated with the high equilibrium x_H^* , in which the consumption of atomistic agents creates a “boom context” that fuels each other’s optimism and sustains high consumption, and a low equilibrium x_L^* in which agents restrain consumption, creating a “depression context” that fuels each other’s pessimism and sustains low consumption. With an elastic aggregate supply, confidence-driven cycles can arise.

This mechanism is complementary to current statistical accounts. In some models, confidence arises from informative signals about aggregate TFP (Lorenzoni 2009; Blanchard et al. 2013; Beaudry and Portier 2006). In Angeletos and Lian (2018), closer to Phelps (1970) and Lucas (1972), confidence reflects optimism about the price of the good an agent supplies, which increases her permanent income, demand, and boosts output under sticky prices. Similarly, in Angeletos and La’O (2013) lack of common knowledge allows beliefs to depend on arbitrary, yet unmodeled, optimism shifters. Shifting aggregate context can produce effects similar to non-fundamental noise shocks, as in Lorenzoni (2009), while cross-sectional variation in NDS experiences and context can create idiosyncratic heterogeneity of expectations as in Angeletos and Lian (2018), including departures from common knowledge.

Memory offers a new testable structure based on similarity and experiences. It also creates multiplicity through expectations themselves, without assuming technological complemen-

tarities such as sequential service constraints in bank runs. “Positive” choices cue “positive” memories, sustaining similar choices in others. This occurs even though, unlike in herding models, these choices convey no information. A suitable aggregate transfer or investment policy can then produce spontaneous optimism and sustain an economic expansion. These forces may shed light on the effects of economic policy and communication. We leave these topics to future work.

2.7 Conclusion

Almost a century ago, Keynes proposed that changes in animal spirits, defined as spontaneous non-statistical forms of business confidence, are a key driver of consumption, investment, and business cycles. By now, abundant evidence shows “excess volatility” in asset prices (Shiller 1981; Bordalo et al. 2024b, 2025c) and in macroeconomic aggregates tied to non-rational earnings expectations (Bordalo et al. 2024a) or to indices of consumer and business “confidence” that are hard to trace to fundamental news (Blanchard 1993; Barsky and Sims 2012; Beaudry and Portier 2006).

We present a foundation that ties animal spirits to a major force in cognition: the role of context in driving memory and attention. Personal or public events affect our semantic and emotional states, shaping how we imagine the future. This approach does not preclude the dependence of beliefs on historical frequencies, which are embedded in the memory database, but stresses that “imagining” a future state comes before “counting” the odds of its occurrence.

Mnemonic beliefs can be studied using measurable memory parameters: experiences, similarity, and cues. The resulting animal spirits exhibit two key features: heterogeneity from

different personal experiences, especially non-informative ones (NDS), and common instability when aggregate context cues similar thoughts. A common shock can change the ability to imagine good vs. bad outcomes above and beyond its information value, and macroeconomic events in one domain can create spillovers into others. For instance, rapidly increasing inflation can cue anxiety and distress, prompting people to simulate high future unemployment even when current unemployment is low or falling. Changes in aggregate context can thus affect business fluctuations as in Keynes' original intuition.

Chapter 3

Memory, Confidence, and Aggregate Demand

with Carla M. Schröder

Chapter Summary

This paper develops a macroeconomic theory in which confidence is endogenous to demand conditions through memory. Households form expectations by recalling similar past demand episodes, with recall cued by visible aggregate consumption. Aggregate demand is therefore self-referential: current demand shapes recall, recall shapes expectations, and expectations feed back into demand. The model yields a confidence multiplier and multiple regimes when memory sharply discriminates across past episodes. Memory accumulation makes the resulting regimes persistent; narrative shocks can trigger persistent regime shifts without changing fundamentals. Empirical evidence supports the central cue-by-memory channel: adverse public cues lower household sentiment and postponable demand, the same adverse cue has larger real effects in states with greater adverse-memory exposure, and household sentiment responds more strongly to public cues than do professional GDP forecasts.

3.1 Introduction

Why do confidence slumps sometimes become so persistent? Periods of weak demand are often accompanied by a broader deterioration in expectations: households become more pessimistic not only about their own circumstances, but about aggregate economic conditions. This pessimism can outlast the original disturbance and help sustain the downturn. Standard macroeconomic models offer several ways to think about this phenomenon (self-fulfilling beliefs, higher-order uncertainty, financial amplification, or nominal rigidities), but they usually treat confidence either as an exogenous disturbance or as an object that moves with fundamentals through a stable mapping. They are less well suited to explaining a simple asymmetry that is central to many downturns: confidence can be highly fragile at the onset of a slump, yet remarkably sticky once the slump is under way.

This paper develops a theory of that asymmetry. The starting point is behavioral but familiar. People often do not form expectations by applying a fixed statistical model to all available information. Instead, they imagine the future by recalling similar past episodes, and which episodes come to mind depends on the context from which recall begins. In macroeconomic settings, the relevant context is often ordinary visible demand: whether shops are busy, whether purchases seem easy or constrained, whether hiring looks active or weak, whether the economy feels expansionary or recessionary in everyday life. These visible conditions serve as cues. They shape which past demand episodes are recalled, and those recalled episodes in turn shape expectations about future aggregate demand. Cues can also shift for reasons other than realized demand: a news story, a policy announcement, a vivid anecdote about earlier downturns, or idiosyncratic experiences unrelated to macroeconomic conditions. Such dis-

turbances, which the paper calls *narrative shocks*, perturb attention without directly changing fundamentals.

Once this idea is embedded in equilibrium, confidence becomes endogenous to ordinary economic conditions. Current demand affects the cue. The cue affects which past episodes are recalled. Recalled episodes affect expected future demand. Expected future demand feeds back into current demand. One household's spending therefore enters the context in which other households form beliefs. This creates an endogenous confidence multiplier: stronger demand today improves the cue seen by others, raises their optimism, and supports additional spending.

The paper's central claim is that this cue-based expectation mechanism, combined with adverse-specific resonance in the similarity kernel, gives downturns the fragile-at-onset, sticky-under-way asymmetry described above. Near the unstable boundary between confidence regimes, memory is actively sorting between expansionary and contractionary episodes, so beliefs are highly sensitive to current conditions. Confidence is fragile. Deep inside a prolonged low-confidence regime, recall becomes concentrated on adverse episodes; adverse cues recall those memories more readily than favorable cues recall favorable ones; and the cue perturbation required to move recall back toward the boundary grows. Confidence is therefore fragile near regime boundaries and sticky in low-confidence interiors, and both properties arise from the same retrieval mechanism.

The analysis takes place in a Hicksian temporary-equilibrium framework with two horizons. Within each period, the economy inherits an active memory database summarizing the stock of past demand episodes currently available for retrieval. Conditional on that inherited database, current aggregate consumption is determined as a fixed point of a demand

map. Across periods, the database itself evolves as realized demand outcomes are encoded into memory and older episodes decay. This distinction between within-period equilibrium and across-period memory accumulation is central: the local multiplier and multiplicity results hold the database fixed, while persistence and tipping arise from the endogenous evolution of memory over time. Figure 3.1 summarizes the within-period and across-period loops.

The theory delivers three main results. First, holding the inherited database fixed, the response of current demand to a small disturbance takes a familiar multiplier form, but the source of amplification is different from standard strategic-complementarity stories. The key object is the mnemonic gradient: the sensitivity of expected future demand to the current cue. Formally, the local feedback slope is $\Gamma = \Phi'(M) M_x$, the product of the demand response to expectations and the cue-slope of mnemonic expectations. The within-period multiplier on a fundamentals shock is $1/(1 - \Gamma)$, and the same multiplier governs the response of demand to attention shocks that perturb the cue without changing fundamentals.

Second, once the mnemonic gradient is allowed to vary with the state, the economy can sustain multiple confidence regimes. This multiplicity does not arise because beliefs are exogenously distorted or because agents randomize across equilibria. It arises because the cue can strongly discriminate between rival demand episodes for some configurations of the inherited memory stock. The economy is most fragile precisely where memory is actively sorting between competing narratives of expansion and contraction.

Third, neutral memory accumulation makes this within-period geometry dynamically consequential, and adverse-specific resonance in the similarity kernel makes it asymmetric. Realized demand is encoded into the database that disciplines retrieval in the next period; a sequence of weak-demand realizations shifts the inherited database toward adverse states and

makes low-confidence regimes persistent once entered. This is the source of hysteresis, and it operates even when adverse and favorable episodes are encoded at the same rate and decay at the same rate. Adverse-specific resonance adds a separate source of asymmetry within the period: an adverse-resonance term in the similarity kernel makes adverse cues recall adverse memories more readily than favorable cues recall favorable memories. Together, the two mechanisms make cross-regime tipping asymmetric: cue perturbations can tip the economy across confidence regimes when it is near the boundary, and escaping a low regime requires a strictly larger cue perturbation than entering it.

The paper also develops two supporting extensions. A heterogeneity extension shows that aggregate amplification is stronger when adverse-memory exposure covaries positively with high marginal propensities to consume. A policy section interprets standard stabilization tools through the same cue-belief-demand loop, emphasizing that policy changes not only resources and incentives, but also the context from which households retrieve and imagine aggregate outcomes. Both are treated as extensions to the central dynamic mechanism rather than as separate centerpieces.

The empirical section asks whether the paper’s distinctive aggregate implications appear in the data. We construct public-data analogues of the model’s central mechanisms: cue shocks identified from news sentiment, economic-policy uncertainty, and adverse-attention search data; and state-level adverse-memory exposure built from recent unemployment and house-price histories. National impulse responses show that adverse public cues lower household sentiment and reduce postponable demand, with the cleanest spending response in vehicle sales after adverse policy-uncertainty shocks. The model’s most distinctive public-data implication is that the same adverse cue should have larger real effects in places where recent adverse experiences make that cue more resonant. We test this in a state-month panel

of building permits interacted with state-level adverse-memory exposure: the headline interaction coefficient is -1.83 with state-clustered standard error 0.68 and a wild-cluster bootstrap p -value of 0.005 , robust to permutation, random-cue, and leave-one-state placebos. As a behavioral check, public adverse cues move household sentiment more strongly than they move professional GDP-growth forecasts, consistent with the model's positioning of household beliefs as cue-based and memory-mediated. The aggregate evidence does not, by itself, structurally identify memory: the contribution of the empirical section is narrower, namely to show that the joint pattern of cue-induced demand pass-through, adverse-memory amplification, and the household-vs.-forecaster wedge is more naturally organized by the model's memory mechanism than by the main rival interpretations considered here.

The paper builds directly on the cognitive-economics line of Bordalo et al. (2020, 2023, 2025a), in which expectations are formed through selective retrieval and simulation rather than through stable statistical updating, and on the survey evidence of Bordalo et al. (2026b) showing that cues can shift macroeconomic beliefs without providing new information. It also connects to narrative-based accounts of sentiment spillovers (Shiller 2019b; Akerlof and Shiller 2009), replacing reduced-form sentiment transmission with a specific retrieval operator and an explicit law of motion for active memory. The behavioral foundation is not the contribution here. The contribution is the macroeconomic implications that follow from that mechanism in equilibrium: an endogenous confidence multiplier, state-dependent multiplicity, fragility near regime boundaries, stickiness in low-confidence interiors, and tipping through narrative shocks.

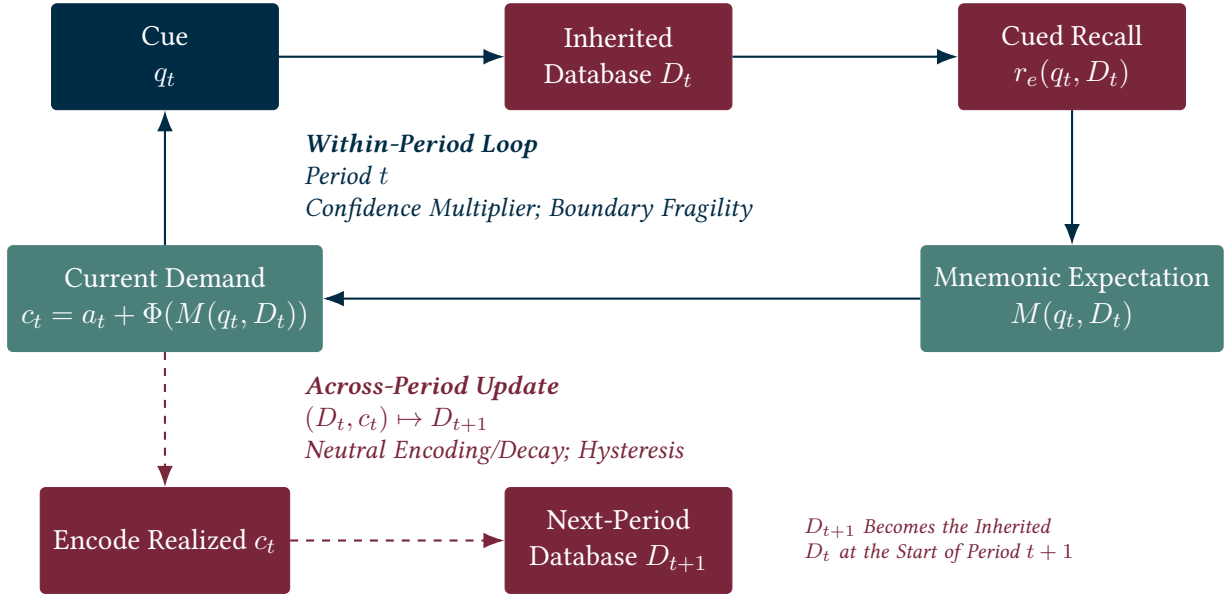
Several adjacent macroeconomic literatures are also engaged. The literature on self-fulfilling fluctuations and indeterminacy (Benhabib and Farmer 1994) treats multiplicity as a property of the aggregate-demand map under reduced-form expectational complementarity;

here multiplicity is endogenous to retrieval discriminability and the inherited memory composition, with explicit state dependence. The confidence-multiplier and higher-order-belief literature (Angeletos and Lian 2018) generates coordination-based amplification that is symmetric across regimes; the multiplier here is microfounded by memory retrieval and made state-dependent by adverse-specific resonance. Work on confidence spillovers and visible-demand externalities (Bellifemine et al. 2025) is close in spirit on the spillover mechanism; the departure is to specify the underlying cue-based retrieval operator explicitly and derive its asymmetric dynamics. News-noise and information-friction models (Lorenzoni 2009; Beaudry and Portier 2006; Mankiw and Reis 2002; Sims 2003; Gabaix 2014) generate persistent belief errors and demand pass-through in a context-invariant way; the present mechanism makes belief distortion context-dependent through the cue, and ties cross-state heterogeneity in cue effects to a state-specific memory stock rather than to belief noise alone. Diagnostic-expectations models (Bordalo et al. 2018) generate distorted and asymmetric beliefs through a representative-agent overweighting rule; here the cross-state heterogeneity follows from the database. Financial-friction and occasionally-binding-constraint models (Guerrieri and Iacoviello 2017) generate asymmetric persistence in downturns through balance sheets and constraints; here the asymmetry comes from cue-dependent recall.

The remainder of the chapter proceeds as follows. Section 3.2 states the aggregate-demand environment. Section 3.3 imports the similarity-based retrieval operator and describes the similarity index. Section 3.4 establishes the within-period results: the retrieval slope, the confidence multiplier, and static multiplicity. Section 3.5 introduces neutral memory updating and shows that, conditional on the static multiplicity condition, neutral updating generates hysteresis. Section 3.6 introduces adverse-specific resonance, shows that it expands the low-confidence basin relative to the high-confidence basin along the cue-perturbation dimension,

and characterizes the cue-perturbation tipping thresholds. Section 3.7 takes the joint pattern of predictions to U.S. public data. Section 3.8 reports the heterogeneity extension, the asymmetric-updating extension, and the policy interpretation. Section 3.9 concludes.

Figure 3.1: Mechanism Schematic



Notes: This figure depicts the within-period and across-period loops of the model. Within period t , the inherited database D_t is fixed and the within-period loop (solid blue) runs as a temporary equilibrium: cue $\rightarrow$ inherited database $\rightarrow$ cued recall $\rightarrow$ mnemonic expectation $\rightarrow$ current demand $\rightarrow$ cue. Once c_t is realized, the across-period update (dashed red) maps $(D_t, c_t) \mapsto D_{t+1}$ via a neutral law of motion that encodes adverse and favorable realizations at the same rate and lets both decay at the same rate. The within-period loop is fragile near regime boundaries, where retrieval discriminability is high; persistence comes from the across-period update, and recession-skewed asymmetry arises separately from an adverse-resonance term in the similarity kernel that makes adverse cues recall adverse memories more readily than favorable cues recall favorable memories.

3.2 Environment

The model is a reduced-form temporary-equilibrium aggregate-demand environment, in the tradition of Hicks (1939) and Grandmont (1977).¹ Within a period, the inherited memory

¹ The use of temporary equilibrium here is deliberately reduced-form: the chapter does not build a full temporary general equilibrium model. It uses the temporary-equilibrium timing to separate a within-period market-clearing

database is fixed and aggregate consumption is determined as a fixed point of a self-referential demand map. Across periods, the only endogenous state variable is the active memory database: realized demand in period t is encoded into the database that disciplines retrieval in period $t+1$.

3.2.1 Reduced-Form Demand Block

The model has no production technology, no capital accumulation, no investment block, and no explicit supply side. There is no labor market, no nominal rigidity, and no independent monetary-policy block. The within-period object of interest is aggregate consumption c_t , and the only endogenous state variable carried across periods is the active memory database D_t . Once c_t is realized, it is encoded into D_{t+1} through a law of motion stated in Section 3.5. Production, capital, and supply-side considerations could be added to this block, but the questions studied here (multiplier, multiplicity, hysteresis, and basin asymmetry) are properties of the demand block alone.

3.2.2 Timing

The within-period timing is as follows. At the start of period t , the database D_t is predetermined. Aggregate consumption c_t and the cue q_t are jointly determined in a temporary equilibrium: given the equilibrium cue, households retrieve episodes from D_t , the retrieved episodes form expectations of next-period aggregate demand, and current demand clears through a reduced-form aggregate-demand equation. After c_t is realized, the database updates to D_{t+1} .

fixed point, conditional on the inherited memory database, from the across-period evolution of that database. This separation is also common in macroeconomic models with non-rational or learning-based expectations; see, for example, Adam and Marcet (2011) and Eusepi and Preston (2011).

3.2.3 Experiences and Database

A stored experience is a triple

$$e = (x_e, v_e, \xi_e),$$

in which $x_e \in \mathbb{R}$ is the experience's feature (in this paper, the level of aggregate demand associated with the past episode), $v_e \in \mathbb{R}$ is its valence, with $v_e < 0$ for adverse episodes and $v_e > 0$ for favorable ones, and ξ_e is its context label, the broader environmental or emotional backdrop. The active memory database is the collection of weights

$$D_t = \{n_{e,t}\}_{e \in E}, \quad (3.1)$$

in which $n_{e,t} \geq 0$ is the active mass on episode e at date t and E is a fixed set of episodes available to be recalled.

3.2.4 Cue

The cue q_t summarizes how current demand presents itself to households along three channels: a feature channel carrying the level of visible demand, a valence channel carrying the affective tone of current conditions, and a context channel. We write

$$q_t = (c_t + n_t^x, V(c_t - \bar{c}) + n_t^v, \xi_t + n_t^\xi), \quad (3.2)$$

in which $V(\cdot)$ is continuously differentiable with $V'(\cdot) > 0$, $V(0) = 0$, and odd; ξ_t is the context component associated with current conditions, possibly correlated with current demand but not restricted to be a function of c_t ; and $\bar{c}$ is a benchmark level of normal demand.

The deterministic component $V(c_t - \bar{c})$ assigns positive valence to above-normal demand and negative valence to below-normal demand. The realized cue valence also includes n_t^v , so the economic context can be perceived negatively even when measured demand is above normal. The three perturbation terms n_t^x , n_t^v , and n_t^ξ allow the cue to be shifted along each channel independently of current demand.

Feature-Channel Restriction. The main analysis sets the valence and context perturbations to zero,

$$n_t^v = n_t^\xi = 0, \quad n_t \equiv n_t^x, \quad (3.3)$$

and studies the scalar feature-channel perturbation n_t . The endogenous valence and context components $V(c_t - \bar{c})$ and ξ_t continue to vary with current conditions. The richer cue mapping in (3.2) is retained because a broad experience-effects literature shows that personal experiences shape macroeconomic beliefs, including idiosyncratic experiences that carry no information about aggregate fundamentals (Malmendier and Nagel 2011; Kuchler and Zafar 2019; d’Acunto et al. 2021; Taubinsky et al. 2024; Cenzon 2025; Bordalo et al. 2026b); Section 3.8 returns to that interpretation.

3.2.5 Demand Equation

Aggregate consumption clears through a reduced-form demand equation,

$$c_t = a_t + \Phi(M(q_t, D_t)), \quad \Phi'(\cdot) > 0, \quad (3.4)$$

in which a_t is an exogenous demand shifter (income, wealth, fiscal transfers, or other fundamentals) and $M(q_t, D_t)$ is the mnemonic expectation of next-period aggregate demand, defined formally in Section 3.3. The function Φ is the demand response to expected future demand; we take Φ to be centered in the odd sense, $\Phi(-M) = -\Phi(M)$, with constant terms absorbed into a_t . The linear specification $\Phi(M) = \beta M$ satisfies this once the constant is absorbed.

Because q_t depends on c_t through (3.2), and because c_t depends on q_t through (3.4), current consumption is determined as a fixed point of a self-referential map. Holding the inherited database D_t fixed, this fixed point solves

$$c_t = a_t + \Phi(M(q(c_t, n_t), D_t)), \quad (3.5)$$

in which $q(c_t, n_t)$ is the cue mapping under the feature-channel restriction (3.3). Section 3.3 imports the retrieval operator that defines $M(q_t, D_t)$; Section 3.4 characterizes the fixed point.

3.3 Similarity-Based Retrieval

Households form mnemonic expectations of next-period aggregate demand by retrieving similar past demand episodes from the active memory database. The retrieval primitive is the similarity-based operator developed by Bordalo et al. (2025a) and disciplined for macroeconomic expectations by Bordalo et al. (2026b). This section imports that operator and defines the similarity kernel.

3.3.1 Recall and Expectation Operators

Given a cue q_t and a database $D_t = \{n_{e,t}\}_{e \in E}$, the recall probability of an experience e is

$$r_e(q_t, D_t) = \frac{S(e, q_t) n_{e,t}}{\sum_{e' \in E} S(e', q_t) n_{e',t}}, \quad (3.6)$$

in which $S(e, q)$ is the similarity primitive and the denominator normalizes recall to a probability distribution over E . Recall rises with the similarity between the cue and the episode and with the relative frequency of the episode in the active database.

The cue-dependence of recall is the source of the aggregate feedback studied below: current demand affects the cue, the cue changes which episodes are recalled, recalled episodes shape expected future demand, and expected future demand feeds back into current demand.

The mnemonic expectation of next-period aggregate demand is the recall-weighted average of the simulated next-period outcomes implied by retrieved episodes,

$$M(q_t, D_t) = \sum_{e \in E} r_e(q_t, D_t) \sigma(e), \quad (3.7)$$

in which $\sigma(e)$ is the next-period aggregate demand simulated from experience e . Throughout the analysis, $\sigma(e) = x_e$: the simulated next-period demand from a recalled episode equals the demand level stored with that episode. The richer simulation map of Bordalo et al. (2026b) is consistent with (3.7) and is recovered as a special case; the level specification is sufficient for the aggregate-demand questions studied here.

3.3.2 Similarity Kernel

The similarity primitive takes the exponential form used by Bordalo et al. (2025a),

$$S(e, q) = \exp\{\eta s(e, q)\}, \quad \eta > 0,$$

in which $s(e, q)$ is the similarity index and η is a concentration parameter governing how sharply recall focuses on the closest-matching episodes. The exponential/softmax form is standard in similarity-based choice models since Luce (1959).² The single parameter η summarizes recall sharpness: as $\eta \rightarrow 0$, recall becomes insensitive to the cue and is driven only by database masses; as $\eta \rightarrow \infty$, recall concentrates on the most-similar episode.

The similarity index has two components:

$$s(e, q) = -|x_q - x_e| + \lambda \mathbb{1}\{\text{sign}(v_q) = \text{sign}(v_e)\}, \quad \lambda \geq 0. \quad (3.8)$$

The first term, feature similarity, is the standard distance-based form used in similarity-based recall (Bordalo et al. 2025a):³ past episodes whose demand feature x_e is close to the cue feature x_q are more retrievable. The second term, governed by λ , is valence congruence: a well-established finding in the psychology of memory (Bower 1981) is that adverse cues make adverse experiences more retrievable, and favorable cues make favorable experiences more retrievable. The indicator $\mathbb{1}\{\text{sign}(v_q) = \text{sign}(v_e)\}$ imposes the simplest symmetric form of this

² With this form, the recall distribution (3.6) is the softmax of similarities, which yields the closed-form derivatives used in the within-period results of Section 3.4.

³ The Laplace form $-|x_q - x_e|$ is nondifferentiable at $x_q = x_e$. Appendix C.1.1 states smooth alternatives that yield the same comparative statics.

principle.⁴ The parameter λ controls the strength of valence congruence: $\lambda = 0$ collapses the kernel to pure feature similarity, while larger λ tilts recall toward valence-congruent experiences regardless of feature distance.

3.3.3 Mnemonic Expectation in the Two-Type Case

We specialize E in equation (3.1) so that only two experience types appear in the database: an adverse type e_L and a favorable type e_H . The database is then characterized by the active masses on these two types, $A_t \equiv n_{e_L,t}$ and $B_t \equiv n_{e_H,t}$. These masses are state variables that vary across periods: realized adverse demand raises A_t , realized favorable demand raises B_t , and both decay over time. The law of motion governing their evolution is stated in Section 3.5; the within-period analysis below takes them as predetermined. The two-type specialization is a tractability simplification. The mechanism does not depend on there being only two types, but the specialization makes the multiplier, multiplicity, and basin-geometry results transparent.

In the symmetric specification used below,

$$e_L = (-x^*, -v^*, \xi_L), \quad e_H = (x^*, v^*, \xi_H), \quad x^*, v^* > 0,$$

the two types are mirror images in feature and valence. Define the adverse-memory share

$$m_t = \frac{A_t}{A_t + B_t} \in [0, 1],$$

⁴ At $v_q = 0$ the valence indicator is assigned by a fixed tie-breaking convention, which has no effect on the results away from the boundary.

which summarizes the composition of the database at date t . Different histories of realized demand produce different m_t : extended adverse periods drive m_t toward 1, extended favorable periods drive m_t toward 0, and mixed histories yield intermediate values.

In this specification, the mnemonic expectation takes a closed form. It is convenient to write it in terms of the log-odds of recalling the adverse type relative to the favorable type,

$$\omega_t \equiv \log \frac{r_L(q_t, D_t)}{r_H(q_t, D_t)}.$$

Substituting the recall probabilities (3.6) and the simulated outcomes $\sigma(e_L) = -x^*$ and $\sigma(e_H) = x^*$ into (3.7) gives

$$M(q_t, D_t) = -x^* \tanh\left(\frac{\omega_t}{2}\right). \quad (3.9)$$

The derivation, together with a discussion of how the result generalizes to richer databases, is in Appendix C.1.2.

The expression has a clean interpretation. When ω_t is large and positive, recall is concentrated on the adverse type and $M \approx -x^*$: expected future demand is low. When ω_t is large and negative, recall is concentrated on the favorable type and $M \approx x^*$: expected future demand is high. When ω_t is near zero, recall is balanced between the two types and M responds sharply to small changes in the cue. The next section shows that this balanced-recall region is where the demand map can become steep enough to amplify shocks and generate multiple temporary equilibria.

3.4 Static Feedback and Multiplicity

This section studies the within-period fixed point in c_t , taking the memory database D_t as predetermined; the dynamics of D_t are taken up in Section 3.5. Three results follow. The first characterizes how the mnemonic expectation responds to the cue (Proposition 1). The second derives the within-period response of c_t to fundamentals and attention shocks (Proposition 2): both responses share the same memory-feedback denominator. The third gives the conditions under which the temporary equilibrium admits multiple stable solutions for a given database (Proposition 3).

3.4.1 Retrieval Slope

How does the mnemonic expectation $M(q, D)$ respond to a change in the cue? Holding the database fixed, the relevant object is the partial derivative $\partial M / \partial x_q$, the slope of mnemonic expectations with respect to the feature component of the cue. The mnemonic expectation is a recall-weighted average of feature outcomes, so a change in the cue affects M by re-allocating recall weights across types. How much M actually moves in response to the cue depends on the composition of the database. When the database is concentrated on one type, recall is anchored to that type and small cue shifts barely move it, so M moves little. When the database holds different types of past experiences in balance, the cue can redistribute recall across types whose feature outcomes differ, and M moves a lot. The slope is therefore largest when recall is balanced and smallest when recall is concentrated on one type. The next proposition formalizes this intuition as a covariance under the recall distribution.

Proposition 1. *Under the kernel of Section 3.3.2, the slope of the mnemonic expectation with respect to the cue’s feature component is*

$$\frac{\partial M(q, D)}{\partial x_q} = \text{Cov}_r \left(\sigma(e), \eta \frac{\partial s(e, q)}{\partial x_q} \right), \quad (3.10)$$

where Cov_r denotes the covariance under the recall distribution $r_e(q, D)$.

Equation (3.10) shows that the cue matters when it reallocates recall toward episodes with different simulated outcomes. If almost all recall weight is on one type, the covariance is small. If recall is balanced across rival types, small cue changes shift weight across outcomes with different $\sigma(e)$, and the slope is large. In the two-type case, this high-slope region corresponds to ω_t near zero; under the symmetric benchmark, this is the region in which m_t is near $1/2$. The proof of Proposition 1 and the explicit two-type formula are in Appendix C.1.1.

3.4.2 Confidence Multiplier

How does within-period demand c respond to an exogenous shock? Two kinds of shocks matter for what follows. A fundamentals shock a enters the demand equation (3.4) directly, moving c before any expectations adjust. An attention shock n enters only through the cue mapping (3.2), shifting how households see current conditions without changing fundamentals. Both shocks set off the same self-referential loop: a change in c moves the cue, the cue changes recall, recall moves M , and M feeds back into c through Φ . The two responses therefore share a single memory-feedback denominator; they differ in their numerators.

Proposition 2. *Consider the feature-channel local benchmark in which the exogenous perturbation is $n \equiv n^x$ and the valence and context components of the cue are held fixed for the local*

comparative static. At a stable temporary equilibrium c^* defined by (3.5), with Φ and M continuously differentiable, define the local memory-feedback slope

$$\Gamma \equiv \Phi'(M) M_x,$$

where $M_x = \partial M / \partial x_q$ is the retrieval slope of Proposition 1. If $\Gamma < 1$ at c^* , then

$$\frac{\partial c}{\partial a} = \frac{1}{1 - \Gamma}, \quad \frac{\partial c}{\partial n} = \frac{\Gamma}{1 - \Gamma}. \quad (3.11)$$

The object Γ is the local strength of the memory-feedback loop. It is the product of two slopes: M_x , the effect of the cue on mnemonic expectations, and $\Phi'(M)$, the effect of those expectations on current demand. If $\Gamma < 1$, a marginal shock is amplified but the feedback is locally damped. If $\Gamma = 1$, the feedback is exactly self-sustaining: a marginal increase in demand moves the cue, the cue moves expectations, and expectations move demand by exactly enough to reproduce the original marginal increase, and the temporary-equilibrium map is tangent to the 45-degree line. If $\Gamma > 1$, the feedback is locally stronger than the initial movement, and the fixed-point map can support multiple equilibria.

The denominator $1 - \Gamma$ is the memory-feedback factor that both responses share; the numerators differ because the two shocks enter the system at different points. A fundamentals shock has a direct effect on c amplified by the feedback, yielding the standard Keynesian-multiplier form $1/(1 - \Gamma)$ but with the propagation channel microfounded by memory-based retrieval rather than postulated. An attention shock has no direct effect on c ; it works entirely through the cue, with multiplier $\Gamma/(1 - \Gamma)$. This is the response of demand to a shock that changes no fundamentals and only shifts the cue from which households retrieve past

episodes. As Γ approaches one from below, both responses become large; Section 3.4.3 shows that when the feedback slope exceeds one on an interval, the temporary-equilibrium map can admit multiple stable regimes.

Full-Channel Form. When valence or context also move locally with current demand, the same logic holds with Γ replaced by the total fixed-point slope. Appendix C.1.4 gives this expansion and shows that, under the kernel of Section 3.3.2, the total slope reduces to Γ in the interior, so (3.11) is generically valid in the main analysis. The proof of Proposition 2 is in Appendix C.1.3.

3.4.3 Static Multiplicity

Can the temporary equilibrium be non-unique? Define the temporary-equilibrium map

$$F(c; D, a) \equiv a + \Phi(M(q(c, 0), D)), \quad (3.12)$$

the right-hand side of (3.5) treated as a function of c at fixed database D , fundamentals a , and zero attention shock. A fixed point is a value of c at which $F(c; D, a) = c$; geometrically, a fixed point is a crossing of the graph of F with the 45-degree line (the diagonal $\{(c, c) : c \in \mathbb{R}\}$).

The relevant object for multiplicity is the slope of F ,

$$F_c(c; D, a) \equiv \frac{\partial F(c; D, a)}{\partial c},$$

which measures how strongly a marginal change in current demand feeds back into itself through the memory loop. If $F_c(c; D, a) < 1$ for all c , then $F(c; D, a) - c$ is strictly decreasing;

under the tail-crossing condition, the fixed point is unique. If $F_c > 1$ on an interval, a marginal increase in c raises F by more than the increase and pushes F above the 45-degree line. Under the regularity conditions below, this yields three crossings for some range of fundamentals: low and high stable fixed points separated by an unstable middle one. In the feature-channel benchmark, $F_c = \Gamma$, so the same memory-feedback slope that governs the multiplier also governs multiplicity.

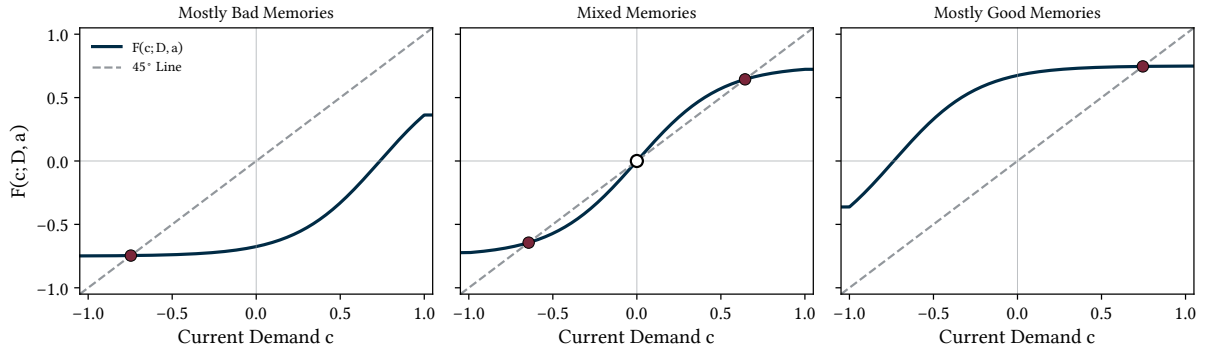
Proposition 3. *Fix D and consider the temporary-equilibrium map $F(c; D, a)$ defined by (3.12). Suppose the tail-crossing and regular S-shape conditions of Appendix C.1.5 hold. If $F_c > 1$ on an open interval, then there is a non-empty open set $\mathcal{A}_D \subset \mathbb{R}$ such that for every $a \in \mathcal{A}_D$, the map admits two locally stable fixed points and one unstable fixed point.*

The proposition formalizes the standard Keynesian-cross multiplicity argument applied to the memory-augmented demand map. The substantive content is that the slope condition $F_c(c; D, a) > 1$ is not a primitive assumption. In the feature-channel benchmark, this slope is $\Gamma = \Phi'(M) M_x$, and M_x depends by Proposition 1 on the dispersion of recall, which in turn depends on the composition of the database D . Multiplicity is therefore endogenous to memory: different databases produce different intervals on which the fixed-point slope exceeds one and hence different ranges $\mathcal{A}_D$ of fundamentals consistent with multiple equilibria. A database concentrated on a single type may admit no such interval at all, and the equilibrium is then unique for every a . Figure 3.2 illustrates the three cases.

The regularity conditions and the proof of Proposition 3 are given in Appendix C.1.5.

Section 3.4 held the database D fixed; Section 3.5 now makes D dynamic: realized demand at t is encoded into the database that shapes retrieval at $t + 1$. Under the static multiplicity condition of Proposition 3, this updating turns temporary regimes into persistent ones.

Figure 3.2: Fixed-Point Geometry of the Temporary-Equilibrium Map



Notes: This figure plots the temporary-equilibrium map $F(c; D, a)$ against the 45-degree line under three database compositions. *Left* (mostly bad memories, m_t near 1): unique low fixed point because recall is anchored to the adverse type and the fixed-point slope stays below one. *Middle* (mixed memories, with recall balanced between rival types): three fixed points, two stable (filled circles) and one unstable (open circle), because balanced recall raises the fixed-point slope above one over an interval. *Right* (mostly good memories, m_t near 0): unique high fixed point.

3.5 Neutral Updating and Hysteresis

Section 3.4 held the database D fixed. This section makes D dynamic. The main result is that, conditional on the static multiplicity condition of Proposition 3, even a memory law of motion that treats adverse and favorable realizations symmetrically generates hysteresis: realized demand changes the database that disciplines future recall, so a temporary shock can generate a persistent regime change.

3.5.1 Neutral Law of Motion

In the two-type case of Section 3.3.3, the active masses A_t and B_t evolve according to

$$A_{t+1} = \rho A_t + \kappa \mathbb{1}\{c_t < 0\}, \quad B_{t+1} = \rho B_t + \kappa \mathbb{1}\{c_t > 0\}, \quad (3.13)$$

with common decay rate $\rho \in (0, 1)$ and common encoding rate $\kappa > 0$.⁵ The parameter ρ captures active-memory durability: higher ρ means that previously encoded episodes remain retrievable for longer. The parameter κ captures encoding intensity: higher κ means that a realized episode receives more active mass when it occurs. The benchmark is *neutral* because the same ρ and κ apply to adverse and favorable realizations: there is no $\rho^- > \rho^+$ and no $\kappa^- > \kappa^+$. Asymmetric encoding and decay can be added as a robustness extension (Section 3.8.2); they sharpen the recession-skew result of Section 3.6 but are not required for the main hysteresis result.

More generally, the active-memory database can evolve according to

$$n_{e,t+1} = \rho_e n_{e,t} + \kappa_e K(e, c_t), \quad (3.14)$$

where $K(e, c_t)$ places new mass on episodes whose features are close to the realized demand state. This formulation separates two memory forces. Recency operates through the new encoding term $\kappa_e K(e, c_t)$: the current realization, or episodes similar to it, receives fresh active mass at date $t+1$. Salience and durability operate through κ_e and ρ_e : large, vivid, or personally meaningful episodes may be encoded more strongly or decay more slowly. The neutral benchmark in (3.13) sets these parameters symmetrically across the two types in order to show that hysteresis does not require assuming that bad memories are encoded more strongly or retained longer.

The limiting behavior of (3.13) is symmetric. If $c_t < 0$ in every period, $A_t \rightarrow \kappa/(1 - \rho)$, $B_t \rightarrow 0$, and the adverse-memory share $m_t \rightarrow 1$. If $c_t > 0$ in every period, $A_t \rightarrow 0$, $B_t \rightarrow \kappa/(1 - \rho)$, and $m_t \rightarrow 0$. A sustained low regime therefore makes adverse memories easier to

⁵ Ties at $c_t = 0$ are assigned by a fixed convention; this convention has no effect on the results.

retrieve in the future; a sustained high regime makes favorable memories easier to retrieve. Both limits are reached at the same rate ρ .

3.5.2 Hysteresis Under Neutral Updating

The static multiplicity result gives the economy two possible stable branches for the same fundamentals: a low-confidence branch and a high-confidence branch. By itself, this is only a within-period statement. If the database were fixed forever, a temporary shock could move current demand while the shock lasts, but the memory environment facing households would be unchanged once the shock disappeared.

Updating changes this logic. When a temporary shock moves the economy into the low branch, the low realization is encoded into the next database. That database then makes adverse episodes easier to retrieve, which supports pessimistic expectations and keeps demand low. The same logic works in reverse for the high branch. Neutral updating therefore makes regimes self-confirming: it does not create the two regimes, but it makes entry into one of them affect future recall.

Formally, the joint state is (c_t, D_t) . Within period t , c_t is determined as a fixed point given the predetermined D_t . Across periods, D_t evolves according to (3.13). The main result is that, conditional on static multiplicity, this joint system has two locally stable steady states, and a temporary shock that moves the economy into one of them can produce a regime that persists after the shock disappears.

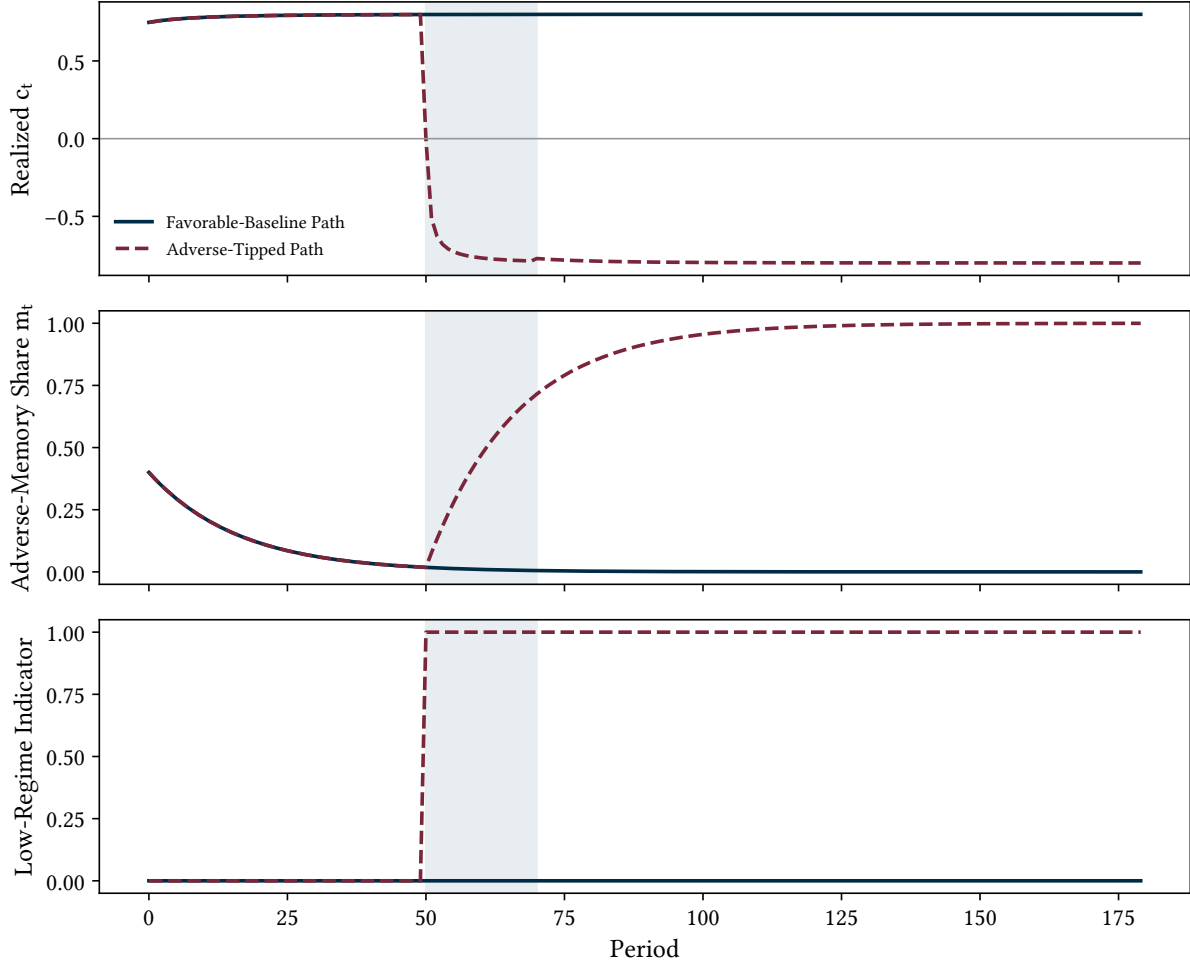
Proposition 4. *Suppose the static multiplicity condition of Proposition 3 holds for neighborhoods of the low- and high-regime databases, and suppose the branch-regularity conditions stated in Appendix C.1.6 hold. Under the law of motion (3.13), the joint (c_t, m_t) system has two locally*

stable steady states (c_L^*, m_L^*) and (c_H^*, m_H^*) , with $c_L^* < 0 < c_H^*$ and $m_L^* > m_H^*$. A temporary fundamentals or attention shock that moves the economy into one of these basins produces a realized demand outcome that is encoded into D_{t+1} , biasing future recall toward the realized regime. The regime can therefore persist after the original shock disappears, with persistence increasing in ρ .

Neutral updating does not create multiplicity. Multiplicity is a static property of the within-period map: it requires the slope condition $F_c > 1$ over an interval (Proposition 3), which depends on the composition of D , not on how D evolves. What neutral updating contributes is that multiplicity becomes dynamically consequential. Once the economy is in a basin, the law of motion validates that basin by encoding its realized demand outcomes into the database, and the static fixed points become dynamic attractors. Figure 3.3 illustrates this with two simulated paths of the joint (c_t, m_t) system: a favorable-baseline path that remains in the high regime, and an adverse-tipped path in which a temporary negative cue perturbation moves realized demand into the low basin and the low regime persists after the perturbation disappears. In the symmetric benchmark, both basins are equally persistent; recession-skewed persistence requires an asymmetry, which Section 3.6 introduces in the similarity kernel.

The proof of Proposition 4 is in Appendix C.1.6. The section's point is narrow. Neutral updating turns static multiplicity into dynamic persistence, but it does not make the low regime harder to exit than the high regime. In the symmetric benchmark, the two basins are mirror images. Section 3.6 introduces an adverse-resonance term in the similarity kernel and studies how this symmetry breaks.

Figure 3.3: Dynamic Path of (c_t, m_t) Under a Temporary Cue Perturbation



Notes: This figure plots two simulated paths of the joint (c_t, m_t) system. The favorable-baseline path (solid) remains in the high regime throughout. The adverse-tipped path (dashed) receives a temporary negative cue perturbation from $t = 50$ to $t = 70$ (shaded window) that moves c_t into the low basin. The realized adverse demand is encoded into the database, raising the adverse-memory share m_t , and the low regime persists after the perturbation disappears (bottom panel). Persistence is increasing in the durability parameter ρ .

3.6 Adverse Resonance and Recession Skew

Section 3.5 showed that, under the kernel of Section 3.3.2 and the neutral law of motion, hysteresis is symmetric: a temporary cue perturbation can move the economy from the high regime to the low regime or from the low regime to the high regime, and the two regime

dynamics are mirror images. This section studies *recession skew*: the asymmetry that the low-confidence regime is easier to enter than to exit, or equivalently that the cue perturbation required to move from the low regime to the high regime is larger than the perturbation required to move from the high regime to the low regime. This object is the confidence-regime analogue of a broader macroeconomic fact: downturns and expansions are not mirror images. Recent work documents this asymmetry in unemployment dynamics, business-cycle skewness, firm-level growth distributions, and downside risk more generally (Dupraz et al. 2025; Jensen et al. 2020; Salgado et al. 2019; Jordà et al. 2024). The contribution here is to derive one psychological source of that asymmetry.

We keep the law of motion neutral and introduce a single term in the similarity kernel that breaks the symmetry of the benchmark: an adverse-resonance term. The result is that escaping the low-confidence regime requires a larger cue perturbation than entering it. The low-confidence regime becomes harder to exit than to enter, without any asymmetry in fundamentals, the law of motion, or the demand block.

3.6.1 Symmetric Benchmark

Under the kernel of Section 3.3.2 and the neutral law of motion of Section 3.5.1, the two regimes of Proposition 4 are mirror images. The kernel treats adverse and favorable valence symmetrically through the congruence term λ ; the law of motion encodes both types at the same rate κ and lets both decay at the same rate ρ ; the demand block Φ is centered with $\Phi(-M) = -\Phi(M)$; and in the two-type specification of Section 3.3.3, the adverse and favorable episodes are sign-flipped images of each other. These restrictions remove any structural asymmetry between the two regimes.

The implication is that the joint (c_t, m_t) system is invariant under the sign-flipping involution $(c, m, n, a) \mapsto (-c, 1 - m, -n, -a)$, together with the relabeling of adverse and favorable episode types. Every dynamic path in the low regime has a mirror-image path in the high regime, with database composition reflected and cue and fundamentals shocks flipped. The two regimes are therefore symmetric in their persistence and tipping thresholds.

This is the natural baseline. The remainder of Section 3.6 introduces a single term in the similarity kernel that breaks this symmetry, and traces out its consequences for basin geometry.

3.6.2 Adverse Resonance

The symmetric valence channel of Section 3.3.2 is not enough to generate recession skew. The term governed by λ makes same-valence memories more retrievable: adverse cues make adverse experiences more retrievable, and favorable cues make favorable experiences more retrievable. But this force is symmetric. With the neutral law of motion and the centered demand block, the low and high regimes remain mirror images.

To make the low-confidence regime harder to exit than to enter, the similarity space itself must be asymmetric. The asymmetry introduced here is *adverse resonance*: adverse experiences may share psychologically relevant features (financial stress, loss of control, constrained spending, anxiety, and reduced agency) that make them especially mutually retrievable. Favorable experiences need not share an equally tight common structure along these dimensions. We capture this by adding one term to the similarity index that is active only for adverse cue–experience pairs. The λ term makes recall valence-dependent; the χ term makes recall recession-skewed.

Formally, we augment the similarity index of Section 3.3.2 with one new term:

$$\begin{aligned}
s(e, q) = & - |x_q - x_e| && \text{(feature similarity)} \\
& + \lambda \mathbb{1}\{\text{sign}(v_q) = \text{sign}(v_e)\} && \text{(valence congruence)} \\
& + \chi \mathbb{1}\{v_q < 0, v_e < 0\}, \quad \chi \geq 0 && \text{(adverse resonance)}.
\end{aligned} \tag{3.15}$$

The new term raises the recall weight of adverse experiences only when the current cue is itself adverse; it has no counterpart on the favorable side. Under the feature-channel restriction used in the main analysis, $n^v = 0$, so $v_q = V(c - \bar{c})$. The χ term is therefore active in the adverse-cue region $c < \bar{c}$ and inactive in the favorable-cue region. Setting $\chi = 0$ recovers the kernel of Section 3.3.2.

The empirical anchor is the experience-effects literature: adverse personal experiences, including experiences outside the macroeconomic domain, are associated with more pessimistic macroeconomic beliefs (Cenzon 2025; Taubinsky et al. 2024), and Bordalo et al. (2026b) establishes this channel causally using priming experiments that shift macroeconomic forecasts without providing information about macroeconomic conditions.

The asymmetry $\chi > 0$ is distinct from an asymmetry in the law of motion, such as adverse episodes being encoded more strongly or retained longer. Section 3.8.2 considers that case as a robustness extension. The main analysis keeps the law of motion neutral and locates the asymmetry in the similarity kernel, separating two mechanisms: persistence comes from updating; recession skew comes from resonance.

3.6.3 Boundary Slope vs. Interior Level

The temporary-equilibrium map exhibits two distinct geometric phenomena. Near the unstable boundary, recall is balanced across the two types, the retrieval slope M_x is high (Proposition 1), and the fixed-point map is locally steep. Small cue perturbations can move the economy across the basin boundary. This is a slope phenomenon: it reflects the within-period response of demand to small cue changes and is generic to the multiplicity region.

Deep in either regime, recall is concentrated. The retrieval slope M_x is small and local feedback is weak, but the level of M , the memory-based expectation of next-period demand, is strongly anchored on one side or the other. In the low regime, M is strongly negative because recall is concentrated on adverse episodes; in the high regime, M is strongly positive because recall is concentrated on favorable episodes. Escaping a regime therefore requires a cue perturbation large enough to move recall back toward the middle region. This is a level phenomenon: it depends on how far memory-based expectations are from zero, not on how steep the map is locally.

Adverse resonance acts primarily on this level margin. To see why, start from a point in the low regime. Recall is already tilted toward adverse experiences, so the memory-based expectation M is negative. When the cue is also adverse, the χ term gives adverse experiences an additional similarity boost relative to the symmetric kernel. This does not mainly make the local slope large; deep in the low regime, recall is already concentrated. Instead, it lowers the level of M : households' memory-based expectation of future demand becomes more pessimistic. Since current demand satisfies $c = a + \Phi(M)$, a lower M shifts the temporary-equilibrium map F downward in the adverse-cue region.

The shift is one-sided. Under the feature-channel restriction, $v_q = V(c - \bar{c})$, so the adverse-resonance term is active when $c < \bar{c}$ and inactive when $c > \bar{c}$. Thus the low side of the map is pulled downward, while the favorable-cue region is locally unchanged. The symmetric benchmark treated low and high regimes as mirror images; adverse resonance breaks that mirror by making the low side of the map more pessimistic in levels.

The geometric consequence is an expansion of the low basin along the cue-perturbation dimension. Starting from the low regime, a positive cue perturbation must first undo the adverse-resonance level effect before the economy reaches the steep boundary region where recall can switch. Escaping the low regime therefore requires a larger positive cue perturbation than in the symmetric benchmark. Starting from the high regime, by contrast, a negative cue perturbation moves the economy toward the adverse-cue region, where the χ term begins to pull recall toward adverse experiences. Entering the low regime therefore requires a smaller negative perturbation. The next subsection formalizes this threshold comparison.

3.6.4 Basin Asymmetry and Cue-Perturbation Thresholds

To make the comparison precise, define two positive cue-perturbation thresholds. Starting from the low steady state (c_L^*, m_L^*) at $a = 0$, let $\tau_{L \rightarrow H}$ denote the smallest positive cue perturbation that moves realized demand across the unstable boundary into the high basin. Starting from the high steady state (c_H^*, m_H^*) , let $\tau_{H \rightarrow L}$ denote the smallest positive magnitude of a negative cue perturbation that moves realized demand into the low basin. Both thresholds depend on the database at the steady state and on the kernel parameters (η, λ, χ) .

Proposition 5. *Suppose the static multiplicity condition of Proposition 3 holds for neighborhoods of the low- and high-regime databases, and suppose the monotone single-boundary and path-*

specific nondegeneracy conditions of Appendix C.1.7 hold. When $\chi = 0$, the cue-perturbation thresholds are equal:

$$\tau_{L \rightarrow H}(0) = \tau_{H \rightarrow L}(0).$$

When $\chi > 0$, the low basin expands relative to the high basin along the cue-perturbation dimension:

$$\tau_{L \rightarrow H}(\chi) > \tau_{H \rightarrow L}(\chi).$$

The first part restates the symmetry of Section 3.6.1; the proof uses the involution argument. The second part is the recession-skew result. The asymmetry follows from the level effect of Section 3.6.3: $\chi > 0$ shifts the temporary-equilibrium map downward in the adverse-cue region only, expanding the set of cue perturbations that keep the economy in the low basin when it starts from the low steady state. Escaping the low regime therefore requires a larger positive cue perturbation than entering it requires in the opposite direction. The mechanism requires no asymmetry in fundamentals, in the law of motion, or in the demand block: the law of motion remains neutral, the demand block remains centered, and the encoding and decay rates remain symmetric.

The proof of Proposition 5 is in Appendix C.1.7. Section 3.7 now evaluates whether the model's joint dynamic pattern appears in U.S. aggregate data.

3.7 Empirical Discipline

The preceding sections derived the model's static, hysteresis, and recession-skew predictions from cue-based retrieval. This section asks whether U.S. data show patterns consistent with the

central mechanism. Aggregate evidence cannot observe retrieved memories directly; it asks instead whether public adverse cues move household sentiment and postponable demand, and whether the same adverse cue has larger real effects in states with greater adverse-memory exposure, the public-data analogue of adverse resonance. A supplementary comparison of household sentiment with professional forecasts serves as a behavioral-placebo check. The section closes by reading the resulting facts as a joint pattern and comparing them to neighboring expectation-formation mechanisms.

3.7.1 Empirical Predictions

The model implies three predictions for the U.S. public-data analysis.

Prediction 1: Public adverse cues lower sentiment and postponable demand. The mnemonic expectation $M(q_t, D_t)$ defined in equation (3.7) responds to the cue q_t through the similarity-based recall operator. An attention-channel cue shift therefore propagates to current consumption through the demand block of equation (3.4). Adverse public cues should lower household sentiment, an aggregate measure of household belief, and lower demand for the spending categories most sensitive to expected future demand. Vehicle sales, the most postponable national consumption category, are the natural test case. *This prediction follows from the within-period confidence multiplier of Proposition 2 acting through the demand block of Section 3.2.5.*

Prediction 2: Adverse memory amplifies cue effects. The adverse-resonance term χ in the augmented similarity index of equation (3.15) makes adverse cues more retrieval-amplifying in databases that hold more adverse experiences. The cross-state public-data analogue of basin geometry is that the same adverse cue should have larger negative effects on real activity in

states with greater recent adverse-memory exposure. *This prediction follows from the basin geometry and cue-perturbation thresholds of Proposition 5.*

Prediction 3: Households respond more than professional forecasters. In the model, household beliefs are formed by similarity-based retrieval from a memory database; the cue activates which past episodes are recalled, and the recalled episodes shape expected future demand. Professional forecasters typically incorporate fundamental information through formal forecasting procedures rather than cue-based recall. If the public cues we measure operate primarily through household memory retrieval rather than through fundamentals, household sentiment should respond to adverse cues more strongly than professional GDP-growth forecasts. *This prediction follows from the role of $M(q_t, D_t)$ in equation (3.7) as a household belief-formation rule that is cue-dependent and memory-mediated.*

The model also implies dynamic predictions about persistence and recession skew: hysteresis (Proposition 4) implies that temporary cue perturbations can produce persistent regime changes, and recession skew (Proposition 5) implies that the low-confidence regime is harder to exit than to enter. The empirical exercises that follow focus on the cue-by-memory amplification channel and do not test asymmetric persistence directly. Diagnostic evidence consistent with the persistence prediction (a sign-asymmetry test in the regional panel and a half-life-ratio diagnostic for Michigan sentiment) is reported in Appendix C.6.

3.7.2 Public Cue Shocks and Adverse-Memory States

The empirical design uses public adverse-cue shocks and state-level adverse-memory exposure. Cue shocks are constructed from the national Daily News Sentiment Index (DNSI) of

Shapiro et al. (2022), the Economic Policy Uncertainty (EPU) index of Baker et al. (2016), Google Trends (GT) search attention for adverse economic terms (Choi and Varian 2012), and a composite of these series. Each cue is residualized against macroeconomic controls and standardized to unit standard deviation. State adverse-memory exposure is built from recent unemployment-rate (UR) and house-price-index (HPI) histories, standardized across the state-month panel and lagged before entering the regressions. The central test is whether the same adverse cue has larger real effects in states with higher adverse-memory exposure.

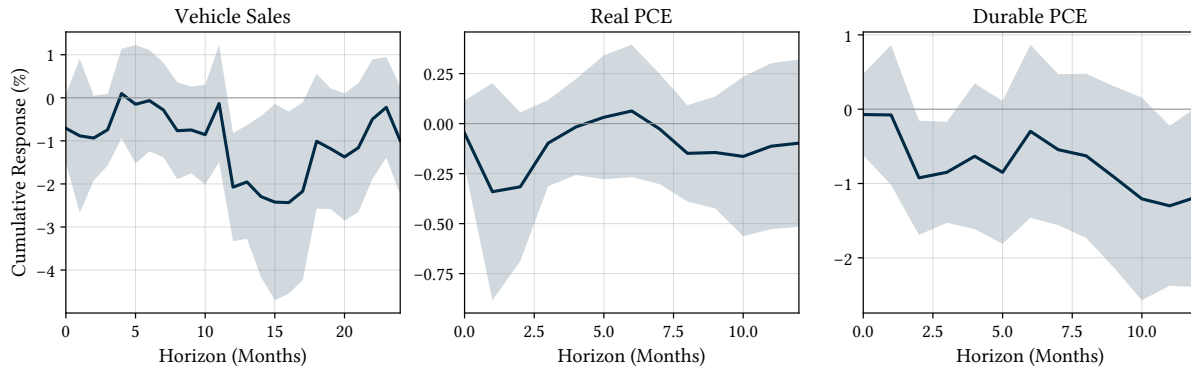
The aggregate evidence does not observe retrieved memories directly. It asks whether public adverse cues have the model-predicted effects on sentiment and demand, and whether those effects are larger where recent adverse experiences make the cue more resonant. Data sources, transformations, and full sample coverage appear in Table C.1 (Appendix C.2.6).

3.7.3 Public Cues Move Sentiment and Postponable Demand

The first exercise asks whether public adverse cues move household sentiment and postponable demand. Figure 3.4 plots the demand response under the adverse EPU cue, with the clearest spending response in vehicle sales, the most postponable national outcome. Figure 3.5 shows the sentiment first stage: the Michigan Index of Consumer Sentiment (ICS) falls on impact for all four cue measures.

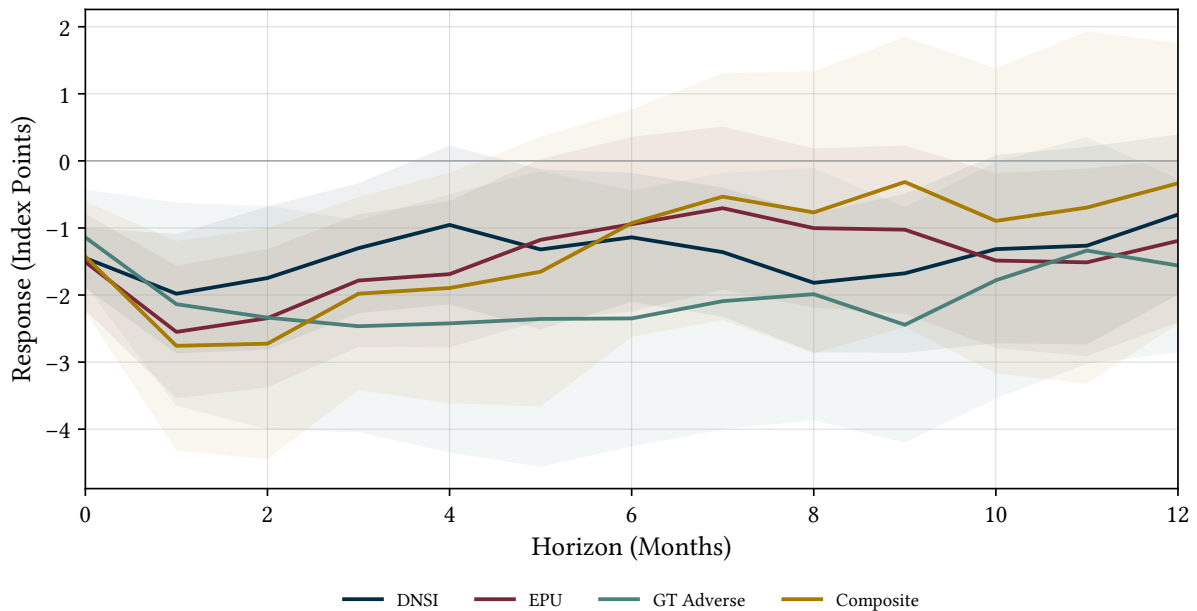
In cumulative form, the adverse-EPU response of vehicle sales at 12 months is statistically negative across the full, ex-COVID, and ex-GFC samples (Table 3.1). Household sentiment falls significantly across all four cue measures (Table 3.2, Panel A). The remaining demand outcomes are less precise: real personal consumption expenditures (PCE) and durables responses are smaller and oscillate, and the corresponding confidence bands generally include zero (Ta-

Figure 3.4: National Demand Responses to Adverse EPU



Notes: This figure reports cumulative percent responses of three national demand outcomes to a 1-standard-deviation adverse EPU shock, with 90% HAC (Newey–West) confidence bands. The clearest demand response is in vehicle sales; real PCE and durable PCE responses are smaller and oscillate.

Figure 3.5: Household Sentiment Responses to Adverse Public Cues



Notes: This figure reports Michigan ICS responses to a 1-standard-deviation adverse public cue, in index points. Lines plot the four cue measures (DNSI, EPU, GT Adverse, Composite) with 90% HAC (Newey–West) confidence bands shown faintly. Household sentiment falls on impact for all four cue measures.

Table 3.1: Vehicle Sales: Response to Adverse EPU Cues

Cue	Outcome	Sample	Estimate	<i>N</i>
EPU	Vehicle Sales, Cum. $h = 12$	Full	-0.0207^{***} (0.0076) $[-0.0333, -0.0082]$	474
EPU	Vehicle Sales, Cum. $h = 12$	Ex-COVID	-0.0193^{***} (0.0064) $[-0.0298, -0.0088]$	452
EPU	Vehicle Sales, Cum. $h = 12$	Ex-GFC	-0.0133^* (0.0071) $[-0.0249, -0.0016]$	464

Notes: The table reports cumulative log responses of total vehicle sales at horizon $h = 12$ to a 1-standard-deviation adverse EPU cue across three sample variants. Each cell reports the point estimate (top), the HAC (Newey–West) standard error in parentheses, and the 90% confidence interval in brackets. The cue is residualized against macroeconomic controls and standardized to unit variance before estimation. Sample tags: *Full* is the full EPU residual sample (1985m3–2026m2); *Ex-COVID* drops 2020m3–2021m12; *Ex-GFC* drops 2008m9–2009m6. Significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

ble 3.2, Panel B). The exercise establishes the public-cue first stage: adverse public cues lower household sentiment, and the clearest demand response appears in vehicle sales. It does not claim that every cue measure moves every demand outcome.

3.7.4 Adverse Memory Amplifies the Effect of Adverse Cues

The second exercise tests the model’s central public-data implication: the same adverse cue should have larger real effects in places where recent adverse experiences make that cue more resonant. The preferred cue is adverse Google search attention, because it is closest to household attention rather than professional news interpretation. The outcome is state-level monthly building permits, counts of new privately-owned housing units authorized by local building authorities, reported by the Census Bureau’s Building Permits Survey and aggregated

Table 3.2: Sentiment and Demand Responses to Adverse Public Cues*Panel A: Household Sentiment Response*

Cue	Outcome	Estimate	<i>N</i>
DNSI	Michigan ICS, $h = 1$	−1.9804*** (0.5411) [−2.8706, −1.0903]	520
EPU	Michigan ICS, $h = 1$	−2.5498*** (0.6008) [−3.5381, −1.5616]	485
GT Adverse	Michigan ICS, $h = 3$	−2.4648*** (0.9597) [−4.0436, −0.8861]	253
Composite	Michigan ICS, $h = 1$	−2.7573*** (0.9488) [−4.3181, −1.1965]	233

Panel B: Other National Demand Outcomes

Cue	Outcome	Estimate	<i>N</i>
EPU	Real PCE, Cum. $h = 12$	−0.0010 (0.0025) [−0.0052, +0.0032]	212
EPU	Real PCE Durables, Cum. $h = 12$	−0.0118* (0.0074) [−0.0239, +0.0004]	391
EPU	Retail Sales, Cum. $h = 12$	−0.0038 (0.0039) [−0.0103, +0.0026]	392

Notes: Panel A reports the most negative Michigan ICS response to a 1-standard-deviation adverse cue over horizons $h \leq 6$, with the peak horizon shown next to each outcome; sample periods are DNSI 1980m3–2023m11, EPU 1985m3–2026m2, GT Adverse 2004m3–2025m12, and Composite 2004m2–2023m11. Panel B reports cumulative log responses at horizon $h = 12$ of three additional national demand outcomes to a 1-standard-deviation adverse EPU cue over the full EPU residual sample (1985m3–2026m2). Each cell reports the point estimate (top), the HAC (Newey–West) standard error in parentheses, and the 90% confidence interval in brackets. All specifications are residualized against macroeconomic controls and standardized to unit variance before estimation. Significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

to the state level. Permits are a leading indicator of residential construction and a natural test of the model’s postponability channel: a permit application is an intertemporal commitment that responds to expected future demand and financing conditions, so memory-based shifts in expectations should pass through to permits more forcefully than to less time-sensitive spending. State permits is the longest-sample postponable real-activity outcome available at state-month frequency; consistent with the national pass-through documented in Section 3.7.3, the adverse Google search cue lowers national permits by approximately 1–4 percentage points across horizons $h \in \{0, 6, 18, 24\}$ in the same residualization design. We smooth log permits with a 3-month moving average to reduce reporting noise. The specification includes state and month fixed effects, so the national cue is absorbed in levels; identification comes from the interaction between the national cue and lagged state adverse-memory exposure.

The estimating equation, written in window-averaged form, is

$$\overline{\Delta y}_{s,t}^{4:8} = \alpha_s + \delta_t + \beta(\text{Shock}_t^{\text{GT},-} \times \text{Memory}_{s,t-1}) + \Pi W_{s,t-1} + \varepsilon_{s,t}, \quad (3.16)$$

where $\overline{\Delta y}_{s,t}^{4:8}$ is the window-average cumulative response of smoothed log permits over horizons $h \in \{4, 5, 6, 7, 8\}$, $\text{Shock}_t^{\text{GT},-}$ is the adverse Google Trends cue residual, $\text{Memory}_{s,t-1}$ is the lagged state adverse-memory composite, and $W_{s,t-1}$ collects lagged state controls and outcome dynamics.

As an intermediate mechanism check, we ask whether national adverse cues raise local adverse-search attention more in high-adverse-memory states. The outcome is a state-level Google Trends index built from searches for “recession,” “economic collapse,” “unemployment,” and “layoffs.” Table 3.3 reports the cells where the cue $\times$ memory interaction is positive (the model-predicted sign) and the 90% confidence interval excludes zero. The strongest evidence

is under the national Google Trends adverse-attention cue at horizons $h = 3$ and $h = 6$; DNSI and EPU also produce positive and significant interactions at $h = 12$. These results are consistent with the model’s attention channel: the same adverse public cue generates more adverse economic search attention in states where recent adverse experiences make the cue more resonant. We do not estimate a structural mediation chain from attention to permits; the search-attention and permits regressions are parallel reduced-form tests of the same cue-by-memory mechanism. Construction details and the regression specification are reported in Appendix C.4.1.

Table 3.3: State-Level Adverse-Search Attention: Cue $\times$ Memory Interactions

Cue	Horizon	Estimate	N
GT Adverse	$h = 3$	+0.0504** (0.0228) [+0.0128, +0.0880]	13,103
GT Adverse	$h = 6$	+0.0413* (0.0236) [+0.0025, +0.0802]	12,950
EPU	$h = 12$	+0.0433*** (0.0164) [+0.0163, +0.0702]	12,746
DNSI	$h = 12$	+0.0251* (0.0146) [+0.0011, +0.0491]	12,185

Notes: The table reports state-month panel estimates of the cue $\times$ memory interaction coefficient $\hat{\beta}_{\text{adv}}^C$ for cells where the estimate is positive (the model-predicted sign) and the 90% confidence interval excludes zero. The outcome is the state-level adverse-search-attention z -score (Google Trends index built from searches for “recession,” “economic collapse,” “unemployment,” and “layoffs”). Memory is the panel- z -scored state composite (UR drawup + HPI drawdown), lagged one month. Each cell reports the interaction estimate (top), the state-clustered standard error in parentheses, and the 90% confidence interval in brackets. All specifications include state and month fixed effects. The model predicts $\beta_{\text{adv}}^C > 0$: adverse cues raise local adverse-search attention more in states with greater adverse-memory exposure. Significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

We next ask whether the same cue-by-memory interaction appears in real activity. The interaction coefficient is $\hat{\beta} = -1.83$ with state-clustered standard error 0.68, 90% confidence interval $[-2.95, -0.70]$, and wild-cluster bootstrap p -value 0.005 (Table 3.4).⁶ The continuous interaction implies a substantially more negative permits response in high-adverse-memory states. Figure 3.6 shows the implied gradient across the adverse-memory distribution: relative to a low-memory (p_{10}) state, the predicted window-averaged permits response is -0.70 percentage points at the median memory state and -3.70 percentage points at the high (p_{90}) memory state, with 90% confidence intervals excluding zero in both cases.

Appendix C.3.6 reports the analogous figure with an additional bar at the p_{99} percentile, where the differential permits response is approximately -8 percentage points relative to the p_{10} state. The result is strongest under the Google Trends adverse-attention cue. Using DNSI, EPU, or the composite cue yields directionally similar but less precise estimates; the full cue-robustness grid is reported in Appendix C.3.3. The Google Trends measure is preferred for the regional exercise because it captures household-level adverse attention more directly than newspaper-based or policy-based cue series. Memory variants (HPI-only, 24- and 60-month windows) and the ex-COVID sample all yield negative interaction coefficients of similar magnitude (Table 3.4).

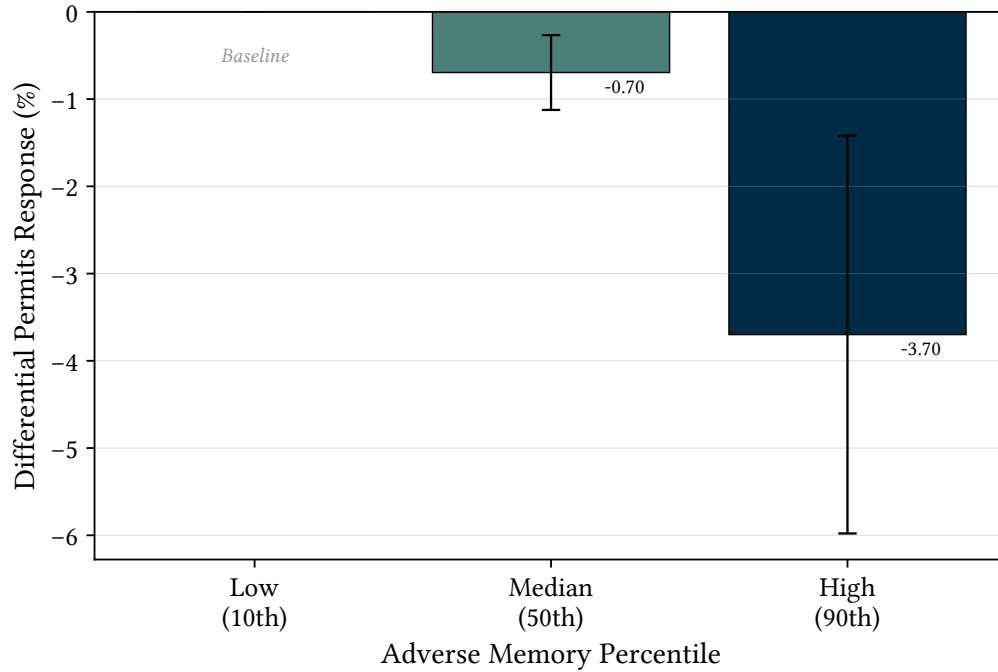
⁶ Permutation checks support the specificity of the interaction. Randomly permuting the adverse-memory index across states places the actual coefficient in the left tail of the placebo distribution ($p = 0.011$); replacing the cue with simulated AR(1) placebo cues gives a similar result ($p = 0.006$). The estimate is also stable in leave-one-state-out exercises, with coefficients ranging from -2.04 to -1.30 .

Table 3.4: Adverse Search Cues, Adverse Memory, and State Permits

Specification	Memory	Sample	Horizon	Estimate	N
Headline	Composite	Full	$h \in [4, 8]$	-1.8260^{***} (0.6839) $[-2.9510, -0.7010]$	12,087
Ex-COVID	Composite	Ex-COVID	$h \in [4, 8]$	-1.9713^{***} (0.6582) $[-3.0541, -0.8885]$	10,965
Single Horizon	Composite	Full	$h = 6$	-1.6298^{**} (0.6732) $[-2.7372, -0.5225]$	12,087
HPI Memory	HPI-Only	Full	$h \in [4, 8]$	-1.3068^{***} (0.3982) $[-1.9618, -0.6519]$	12,087
24-Month Memory	Composite-24	Full	$h \in [4, 8]$	-1.8370^{**} (0.7357) $[-3.0472, -0.6269]$	12,087
60-Month Memory	Composite-60	Full	$h \in [4, 8]$	-1.8651^{***} (0.6690) $[-2.9657, -0.7645]$	12,087

Notes: The table reports estimates of β from equation (3.16). The outcome is the window-average cumulative change in smoothed $\log(\text{permits})$ over horizons $h = 4, \dots, 8$, except in the single-horizon row. The cue is the adverse Google Trends residual, standardized to unit standard deviation. Memory measures are lagged, panel-standardized, and centered before interaction. Each cell reports the point estimate (top), the state-clustered standard error in parentheses, and the 90% confidence interval in brackets. All specifications include state and month fixed effects. The headline cell additionally has a wild-cluster bootstrap p -value of 0.0050. The panel covers 51 states. Full cue, memory, and sample robustness is reported in Appendix C.3.3. Significance levels: $*** p < 0.01$, $** p < 0.05$, $* p < 0.10$.

Figure 3.6: Adverse-Memory Amplification of State Permits Response to Adverse Cues



Notes: This figure plots predicted differential responses of smoothed log permits to a 1-standard-deviation adverse Google Trends cue, evaluated at the 10th, 50th, and 90th percentiles of the panel- z -scored state adverse-memory composite (corresponding to $z = -0.85, -0.47$, and $+1.17$, respectively). Responses are window averages over horizons $h = 4, \dots, 8$ and are measured relative to the p_{10} memory state. Error bars are 90% confidence intervals.

3.7.5 Households vs. Professional Forecasters

As a supplementary check, we compare household sentiment with professional forecasts. The model implies that public cues operate through household belief formation: cues activate memory retrieval, retrieved episodes shape expectations, and expectations feed into demand. Professional forecasters, by contrast, are taken to incorporate fundamental information through formal forecasting procedures rather than cue-based recall. If the public cues mainly reflected news about fundamentals, one would expect professional forecasts to respond to them as well; if instead the cues affect beliefs primarily through cue-based recall, household

sentiment should respond more.

We compare the quarterly-averaged Michigan ICS to 1- and 4-quarter-ahead median real GDP-growth forecasts from the Philadelphia Fed Survey of Professional Forecasters (SPF). Cue residuals are summed within quarter and re-standardized; both the Michigan series and the SPF growth series are standardized to unit standard deviation over the estimation sample, so coefficients are comparable in LHS standard deviations per 1-standard-deviation adverse cue. In the quarterly comparison, adverse cues move Michigan sentiment more strongly than professional GDP-growth forecasts for three of four cue measures, while the SPF point estimates are imprecisely estimated and in some cases statistically indistinguishable from zero. We do not lean on the exact response ratios (SPF coefficients are noisy and sometimes wrong-signed), but the qualitative pattern is consistent with the interpretation that public cues move household beliefs more than professional GDP forecasts. The exercise is not a formal identification test; it is a behavioral-placebo check. The full table is in Appendix C.5.2.

3.7.6 Joint Pattern and Alternative Mechanisms

The empirical results should be read as a joint pattern rather than as separate identification tests. Public adverse cues lower household sentiment and reduce postponable demand. The real effect of an adverse cue is larger in states with greater adverse-memory exposure. Household sentiment responds more than professional forecasts. Each fact has alternative interpretations, but the combination is naturally organized by the model: current cues shape recall, recall shapes expectations, expectations affect demand, and the effect is stronger where adverse memories are more available.

Table 3.5 compares how neighboring mechanisms organize the empirical pattern. A per-

Table 3.5: Joint Pattern Across the Empirical Analysis

Empirical Pattern	Threshold	News	Diagnostic	HOB	Frictions	Memory
Public Cues Lower Household Sentiment	~	✓	✓	~	~	✓
Public Cues Lower Postponable Demand	×	✓	~	✓	✓	✓
Adverse Memory Amplifies Cue Effects	×	×	~	×	~	✓
Households Respond More Than Forecasters	×	~	✓	~	×	✓

Notes: The table summarizes which mechanisms naturally generate each empirical pattern. The patterns and the supporting evidence are: (i) *Public Cues Lower Household Sentiment* (4/4 cue measures lower Michigan ICS, Table 3.2); (ii) *Public Cues Lower Postponable Demand* (adverse EPU lowers vehicle sales across samples, Table 3.1); (iii) *Adverse Memory Amplifies Cue Effects* (cue $\times$ memory interaction on state permits, $\hat{\beta} = -1.83$, wild-cluster bootstrap $p = 0.005$, Table 3.4); (iv) *Households Respond More Than Forecasters* (Michigan ICS responds more strongly than SPF GDP-growth forecasts for 3/4 cues, Table C.6). Column headings name the mechanism class: *Threshold* = persistent threshold and Markov-switching processes (Hamilton 1989); *News* = news and noise-shock models (Lorenzoni 2009; Beaudry and Portier 2006; Mankiw and Reis 2002; Sims 2003; Gabaix 2014); *Diagnostic* = diagnostic-expectations models (Bordalo et al. 2018, 2020); *HOB* = higher-order-belief and coordination models (Angeletos and Lian 2018; Angeletos and La’O 2013); *Frictions* = financial-friction and occasionally-binding-constraint models (Guerrieri and Iacoviello 2017); *Memory* = cue-based memory retrieval (Bordalo et al. 2025a, 2026b). Symbols: ✓ means the mechanism naturally generates the pattern; ~ means it can generate the pattern under additional structure or in some states; × means it does not naturally generate the pattern. The table is not a formal horse race; its purpose is to show that memory-based retrieval organizes the full pattern through one expectation-formation channel.

sistent threshold or Markov-switching process (Hamilton 1989) can generate state-dependent sentiment dynamics but does not by itself link cues to real demand. Information-friction and news-noise models (Lorenzoni 2009; Beaudry and Portier 2006; Mankiw and Reis 2002; Sims 2003) naturally generate demand pass-through from belief innovations but do not tie the belief distortion to a state-specific memory stock. Higher-order-belief and coordination models (Angeletos and Lian 2018) generate amplification but not the cue-by-memory interaction. Diagnostic-expectations models (Bordalo et al. 2018) generate distorted and sometimes asymmetric beliefs but do not naturally predict that the same public cue has larger real effects in states with greater adverse-memory exposure. Financial-friction and occasionally-binding-constraint models (Guerrieri and Iacoviello 2017) generate real demand effects and asymmetric persistence in downturns, but the asymmetry comes from balance sheets or constraints

rather than from cue-dependent recall. Memory-based retrieval naturally organizes all four facts through one expectation-formation channel.

Two pieces of the joint pattern are the most distinctive to the model. The regional cue-by-memory interaction is the public-data analogue of adverse resonance: the same adverse cue produces a more negative real-activity response in states with greater adverse-memory exposure. The household-vs.-forecaster wedge is consistent with cue-based belief formation among households: the cues move household sentiment more strongly than they move professional GDP forecasts. The aggregate evidence remains supportive rather than identifying. The interpretation rests on the combination of this public-data pattern with the survey and experimental evidence on cue-based retrieval and experience-based belief formation summarized in Sections 3.2.4 and 3.6.2.

3.8 Extensions and Interpretation

The core model deliberately restricts attention to one feature-channel cue perturbation, a neutral memory law of motion, and a representative household. This section discusses four extensions that are natural in the framework but not needed for the main results: richer cue perturbations, asymmetric updating, heterogeneity, and policy interpretation. The discussion is interpretive; no new formal results are introduced.

3.8.1 Richer Cue Perturbations and Narrative Shocks

The main analysis used the feature-channel restriction in (3.3): $n_t^v = n_t^\xi = 0$, and the scalar perturbation $n_t \equiv n_t^x$ shifted only the level of visible demand. The cue mapping (3.2) also

admits perturbations to the valence channel n_t^v and to the context channel n_t^ξ . A narrative shock (a story, news cycle, or salient event that changes how current conditions are perceived without directly changing measured demand) is naturally interpreted as a perturbation to n_t^v , n_t^ξ , or both, rather than as a feature-channel shock n_t^x .

This interpretation links the model to narrative economics. The idea that beliefs about the economy can be reshaped by stories that activate retrieval of past episodes is central to Shiller (2019b) and Akerlof and Shiller (2009). The priming experiments of Bordalo et al. (2026b) are the cleanest causal evidence for the channel: priming respondents with non-domain adverse experiences shifts macroeconomic forecasts even when the prime carries no information about macroeconomic conditions. In the present model, that effect is naturally represented as a shift in the valence or context components of the cue. The prime changes what the present feels similar to, raises the recall weight on adverse past experiences, and shifts the mnemonic expectation M .

The same logic extends naturally to richer cue perturbations. The multiplier in Proposition 2 applies to any perturbation that shifts M , and adverse resonance amplifies adverse cues whether they arrive through the feature, valence, or context channel. The sentiment channel of fiscal policy studied by Bellifemine et al. (2025) is one example of a narrative-style cue perturbation: policy affects spending not only by changing resources, but also by changing what households infer, attend to, or recall.

The formal analysis is silent on which channel any particular narrative operates through. Empirically, the distinction matters: a narrative shock that operates through n_t^v or n_t^ξ can shift expectations even when measured demand is unchanged on impact. The richer cue mapping is the natural way to extend the framework to multi-dimensional cue perturbations and

narrative-style shocks; the core mechanism remains the same, but a full treatment would require specifying a multi-dimensional similarity kernel.

3.8.2 Asymmetric Updating

The law of motion in Section 3.5 is *neutral*: the encoding rate κ and the decay coefficient ρ are symmetric across episode types. The choice was deliberate. The hysteresis result of Proposition 4 and the recession-skew result of Proposition 5 are both derived under a neutral law of motion. Hysteresis follows from updating itself: realized demand reshapes the future database whether or not encoding is asymmetric. Recession skew follows from the adverse-resonance term χ in the similarity kernel, not from any asymmetry in the law of motion.

Allowing adverse episodes to be encoded more strongly or to decay more slowly ($\kappa^- > \kappa^+$ or $\rho^- > \rho^+$ in the general formulation (3.14)) would amplify both effects. A larger encoding rate for adverse episodes would deepen the hysteresis pattern after a low-regime spell; a larger decay coefficient for adverse episodes would slow the return to balanced retrieval after the cue perturbation passes. Asymmetric updating of this kind is consistent with the psychology of negative-experience persistence (Bower 1981) and with the experience-effects literature, in which adverse macroeconomic experiences leave durable imprints on beliefs (Malmendier and Nagel 2011; Kuchler and Zafar 2019). The point of the baseline is not that asymmetric updating is implausible. The point is that the core mechanism (basin geometry under cue-based retrieval, with asymmetry generated by adverse resonance) does not require it. Extensions in which κ or ρ depend on the episode type sharpen the recession-skew result without changing its source.

3.8.3 Heterogeneity

The demand block of Section 3.2.5 is representative-agent. Households differ in two ways the model abstracts from. First, they differ in memory databases: idiosyncratic histories of past macroeconomic and personal experiences place different weights on different episode types in D_t (Malmendier and Nagel 2011; Kuchler and Zafar 2019; d’Acunto et al. 2021). Second, they may differ in retrieval primitives: the similarity-kernel parameters η , λ , and χ can vary across individuals or demographic groups. Recent micro evidence is consistent with both kinds of heterogeneity. Taubinsky et al. (2024) document that household-level shocks shift macroeconomic beliefs by more than aggregate information would justify, and Cenoz (2025) shows that adverse credit-market experiences leave durable imprints on macroeconomic expectations.

Aggregating over heterogeneous households, the relevant object for aggregate amplification is the cross-sectional covariance between memory sensitivity and the marginal propensity to consume (MPC). If high-MPC households are also the ones whose retrieval is most reshaped by adverse cues (households with thinner buffer stocks, more recent recessionary experience, or stronger valence congruence), then aggregate consumption responds more sharply to a negative cue than the representative-agent baseline implies. The same mechanisms can operate under aggregation, but their quantitative strength depends on the covariance structure. A full treatment of heterogeneity requires household-panel data on memory, expectations, and spending, and is left for future work.

3.8.4 Policy as Cue Management

The model offers a positive interpretation of stabilization policy as operating partly through cues. The cue q_t in (3.2) is shaped by realized demand c_t through the feature channel, by the affective tone of current conditions through the valence channel, and by background context through the third channel. Standard demand-management policy affects the feature component indirectly: by raising current demand, it changes the visible-demand term c_t . Communication policy (forward guidance, narrative framing, announcements) can operate more directly on the valence and context channels, shifting the perception of current conditions without immediately changing realized demand. In both cases, the change to the cue reshapes retrieval, which reshapes the mnemonic expectation M and feeds back into the demand block of Section 3.2.5.

Policy effectiveness in this framework is state-dependent in a way that follows directly from the basin geometry of Section 3.6. Near a basin boundary, recall is balanced and the retrieval slope is steep; small cue interventions can therefore have larger effects on regime-entry probabilities than the same interventions applied deep in the interior. Once the economy is deep in a low-confidence regime, adverse resonance anchors recall on adverse episodes, so favorable cue interventions must be larger or more persistent to move the economy back toward the boundary. This combination, high marginal sensitivity at the boundary and low marginal sensitivity deep in the low basin, ties together the within-period multiplier of Proposition 2 and the cross-period asymmetry of Proposition 5.

This positive interpretation connects to existing work on belief-mediated policy. Forward-guidance models with imperfect common knowledge (Angeletos and Lian 2018) highlight how communication can move beliefs without contemporaneous changes in fundamentals.

Sticky-information frameworks (Mankiw and Reis 2002) share the property that announcements reach behavior through belief-formation rules. The sentiment channel of fiscal policy (Bellifemine et al. 2025) is a recent example in which fiscal announcements operate partly through expectational rather than purely resource channels. The framework here adds one ingredient: the belief-formation rule is cue-based retrieval from a memory database, so policy interventions change which past episodes households recall, not only what households believe about future fundamentals.

The discussion is positive, not normative. The chapter does not solve a welfare problem or characterize optimal policy. A welfare analysis would require taking a stand on which features of the cue are policy-controllable, on social preferences over confidence regimes, and on the political-economy constraints that govern communication. These are open questions that the framework can pose but does not answer.

3.9 Conclusion

The chapter has asked why confidence slumps can be so persistent, and has offered an answer grounded in cue-based memory retrieval. When households form expectations by recalling similar past demand episodes, current visible demand becomes the cue from which recall begins. Aggregate demand becomes self-referential: current demand shapes recall, recall shapes expectations, and expectations feed back into current demand. The model yields a confidence multiplier that depends on how sharply mnemonic expectations respond to the cue, and shows that this slope can be steep enough to support multiple stable confidence regimes for the same fundamentals. Multiplicity is not assumed; it emerges from the dispersion of recalled episodes and the composition of the inherited memory stock, and it is endogenous to where the econ-

omy currently sits.

The dynamic side of the model follows from a single new state variable: the active memory database. Realized demand is encoded into the database that disciplines retrieval in the next period, so within-period multiplicity becomes across-period hysteresis. A temporary cue perturbation that pushes the economy into one basin reshapes future recall; the regime can persist long after the original perturbation disappears, with persistence increasing in the durability of memory. Hysteresis arises from updating itself, not from any asymmetry in how adverse and favorable episodes are encoded or decay. The recession-skewed asymmetry, that the low-confidence regime is harder to exit than to enter, comes from a separate source: an adverse-specific resonance term in the similarity kernel makes adverse cues recall adverse experiences more readily than favorable cues recall favorable ones. The asymmetry that gives downturns their stickiness emerges from how memory retrieves, not from how memory accumulates.

The empirical section takes public-data analogues of these mechanisms to U.S. data. Adverse public cues, identified from news sentiment, economic-policy uncertainty, and adverse-attention search data, lower household sentiment and reduce postponable demand, with the cleanest spending response in vehicle sales after adverse policy-uncertainty shocks. The model's most distinctive public-data implication is that the same adverse cue should have larger real effects in places where recent adverse experiences make the cue more resonant. We test this in a state-month panel of building permits interacted with state-level adverse-memory exposure: the headline interaction coefficient is -1.83 with state-clustered standard error 0.68 and a wild-cluster bootstrap p -value of 0.005 , and is robust to permutation, random-cue, and leave-one-state placebos. As a behavioral check, public adverse cues move household sentiment more strongly than they move professional GDP-growth forecasts, consistent with the

model's positioning of household beliefs as cue-based and memory-mediated. None of these facts identifies the memory mechanism on its own, but the joint pattern is more naturally organized by the model's expectation-formation channel than by the main rival interpretations considered here.

The interpretation rests on the combination of two kinds of evidence. The behavioral primitive, similarity-based retrieval cued by current conditions, is disciplined by the priming experiments of Bordalo et al. (2026b), which show that household macroeconomic forecasts shift in response to cues that carry no information about fundamentals. The aggregate evidence in this chapter is consistent with the same primitive operating at the macroeconomic scale, in the sense that the cross-state heterogeneity in cue effects, the demand pass-through, and the household-vs.-forecaster wedge line up with what the model implies. Together, the experimental and aggregate evidence provide complementary disciplines on the same mechanism, even though neither alone identifies it.

The framework opens several natural directions. The cue mapping admits perturbations along valence and context channels that the main analysis sets to zero; extending the analysis to a multi-dimensional similarity kernel would let the framework speak directly to narrative-style disturbances (forward-guidance announcements, news cycles, salient anecdotes) that perturb attention without changing fundamentals. The law of motion is intentionally neutral in the baseline, which isolates the source of recession-skewed asymmetry in the kernel; allowing asymmetric encoding and decay can sharpen the pattern in environments where adverse experiences leave durable imprints, and the structural separation between persistence (from updating) and skew (from resonance) is preserved. The representative-agent household block can be enriched by aggregating over heterogeneous memory databases and marginal propensities to consume, tying the strength of aggregate amplification to a cross-sectional

covariance that household panel data could discipline directly. The policy discussion is positive; a welfare treatment would build on this positive analysis by specifying which features of the cue are policy-controllable, what social preferences govern the choice between confidence regimes, and what political-economy constraints discipline communication. Each is a natural direction for future work and a place where the model can be made sharper.

The broader point is that confidence does not have to be left as an exogenous disturbance or as a stable function of fundamentals. Once expectations are formed by recalling similar past episodes, confidence becomes a function of what the economy currently looks like and of what it has recently been, a function the model writes down explicitly. Within-period multiplicity, across-period hysteresis, and recession-skewed exit thresholds are not separate phenomena to be added on top of standard models. They are what one obtains when the belief-formation rule is cue-based memory retrieval rather than a stable mapping from fundamentals.

References

- Adam, Klaus, and Albert Marcet. 2011. "Internal Rationality, Imperfect Market Knowledge and Asset Prices." *Journal of Economic Theory* 146(3): 1224–1252.
- Akerlof, George, and Robert Shiller. 2009. *Animal Spirits: How Human Psychology Drives the Economy, and Why It Matters for Global Capitalism*. Princeton University Press.
- Andre, Peter, Ingar Haaland, Christopher Roth, Mirko Wiederholt, and Johannes Wohlfart. 2024. "Narratives about the Macroeconomy." SAFE Working Paper No. 426.
- Andre, Peter, Carlo Pizzinelli, Christopher Roth, and Johannes Wohlfart. 2022. "Subjective Models of the Macroeconomy: Evidence from Experts and Representative Samples." *Review of Economic Studies* 89(6): 2958–2991.
- Angeletos, George-Marios, Zhen Huo, and Karthik A. Sastry. 2021. "Imperfect Macroeconomic Expectations: Evidence and Theory." *NBER Macroeconomics Annual* 35(1): 1–86.
- Angeletos, George-Marios, and Jennifer La'O. 2013. "Sentiments." *Econometrica* 81(2): 739–779.
- Angeletos, George-Marios, and Chen Lian. 2018. "Forward Guidance without Common Knowledge." *American Economic Review* 108(9): 2477–2512.
- Armona, Luis, Andreas Fuster, and Basit Zafar. 2019. "Home Price Expectations and Behaviour: Evidence from a Randomized Information Experiment." *Review of Economic Studies* 86(4): 1371–1410.
- Bailey, Michael, Eduardo Dávila, Theresa Kuchler, and Johannes Stroebel. 2019. "House Price Beliefs and Mortgage Leverage Choice." *Review of Economic Studies* 86(6): 2403–2452.
- Baker, Malcolm, and Jeffrey Wurgler. 2006. "Investor Sentiment and the Cross-Section of Stock Returns." *Journal of Finance* 61(4): 1645–1680.
- Baker, Scott R., Nicholas Bloom, and Steven J. Davis. 2016. "Measuring Economic Policy Uncertainty." *Quarterly Journal of Economics* 131(4): 1593–1636.
- Barberis, Nicholas, Andrei Shleifer, and Robert Vishny. 1998. "A Model of Investor Sentiment." *Journal of Financial Economics* 49(3): 307–343.
- Barsky, Robert, and Eric Sims. 2012. "Information, Animal Spirits, and the Meaning of Innovations in Consumer Confidence." *American Economic Review* 102(4): 1343–1377.
- Beaudry, Paul, and Franck Portier. 2006. "Stock Prices, News, and Economic Fluctuations." *American Economic Review* 96(4): 1293–1307.

- Bellifemine, Marco, Adrien Couturier, and Jacopo Tozzo. 2025. "The Sentiment Channel of Fiscal Policy." Working Paper.
- Benhabib, Jess, and Roger E. A. Farmer. 1994. "Indeterminacy and Increasing Returns." *Journal of Economic Theory* 63(1): 19–41.
- Bhandari, Anmol, Jaroslav Borovička, and Paul Ho. 2025. "Survey Data and Subjective Beliefs in Business Cycle Models." *Review of Economic Studies* 92(3): 1375–1437.
- Blanchard, Olivier. 1993. "Consumption and the Recession of 1990–1991." *American Economic Review* 83(2): 270–274.
- Blanchard, Olivier, Jean-Paul L’Huillier, and Guido Lorenzoni. 2013. "News, Noise, and Fluctuations: An Empirical Exploration." *American Economic Review* 103(7): 3045–3070.
- Bonaglia, Edoardo, Daniele d’Arienzo, Nunzio Fallico, Nicola Gennaioli, and Luigi Iovino. 2025. "The Volatility of Inflation Expectations and Interest Rates." Working Paper.
- Bonaglia, Edoardo, and Nicola Gennaioli. 2025. "Mnemonic Expectations." Working Paper.
- Bordalo, Pedro, Giovanni Burro, Katherine Coffman, Nicola Gennaioli, and Andrei Shleifer. 2025a. "Imagining the Future: Memory, Simulation, and Beliefs." *Review of Economic Studies* 92(3): 1532–1563.
- Bordalo, Pedro, Giovanni Burro, Nicola Gennaioli, Gad Nacamulli, and Andrei Shleifer. 2025b. "Ads as Cues." NBER Working Paper No. 34387.
- Bordalo, Pedro, John Conlon, Nicola Gennaioli, Spencer Kwon, and Andrei Shleifer. 2023. "Memory and Probability." *Quarterly Journal of Economics* 138(1): 265–311.
- Bordalo, Pedro, Nicola Gennaioli, Rafael La Porta, Matthew O’Brien, and Andrei Shleifer. 2024a. "Long-Term Expectations and Aggregate Fluctuations." *NBER Macroeconomics Annual* 38(1): 311–347.
- Bordalo, Pedro, Nicola Gennaioli, Rafael La Porta, and Andrei Shleifer. 2019. "Diagnostic Expectations and Stock Returns." *Journal of Finance* 74(6): 2839–2874.
- Bordalo, Pedro, Nicola Gennaioli, Rafael La Porta, and Andrei Shleifer. 2024b. "Belief Overreaction and Stock Market Puzzles." *Journal of Political Economy* 132(5): 1450–1484.
- Bordalo, Pedro, Nicola Gennaioli, Rafael La Porta, and Andrei Shleifer. 2025c. "Finance without Exotic Risk." *Journal of Financial Economics* 173 104145.
- Bordalo, Pedro, Nicola Gennaioli, Giacomo Lanzani, and Andrei Shleifer. 2026a. "A Cognitive Theory of Reasoning and Choice." *The Quarterly Journal of Economics*.
- Bordalo, Pedro, Nicola Gennaioli, and Andrei Shleifer. 2018. "Diagnostic Expectations and Credit Cycles." *Journal of Finance* 73(1): 199–227.

- Bordalo, Pedro, Nicola Gennaioli, and Andrei Shleifer. 2020. "Memory, Attention, and Choice." *Quarterly Journal of Economics* 135(3): 1399–1442.
- Bordalo, Pedro, Nicola Gennaioli, Florencio Lopez-de Silanes, Simon Schröder, Andrei Shleifer, and Maarten van Rooij. 2026b. "The Psychology of Macroeconomic Expectations." Working Paper.
- Bower, Gordon. 1981. "Mood and Memory." *American Psychologist* 36(2): 129–148.
- Bretscher, Lorenzo, Aytak Malkhozov, Andrea Tamoni, and Haoxi Yang. 2025. "Distorted Beliefs and Asset Prices." Working Paper.
- Brunnermeier, Markus, and Jonathan Parker. 2005. "Optimal Expectations." *American Economic Review* 95(4): 1092–1118.
- Bryan, Michael F., and Guhan Venkatu. 2001. "The Demographics of Inflation Opinion Surveys." Federal Reserve Bank of Cleveland Economic Commentary.
- Calzolari, Giacomo, and Mattia Nardotto. 2017. "Effective Reminders." *Management Science* 63(9): 2915–2932.
- Campbell, John Y. 1991. "A Variance Decomposition for Stock Returns." *Economic Journal* 101(405): 157–179.
- Campbell, John Y., and John Ammer. 1993. "What Moves the Stock and Bond Markets? A Variance Decomposition for Long-Term Asset Returns." *Journal of Finance* 48(1): 3–37.
- Campbell, John Y., and Robert J. Shiller. 1988. "The Dividend-Price Ratio and Expectations of Future Dividends and Discount Factors." *Review of Financial Studies* 1(3): 195–228.
- Campbell, John Y., and Tuomo Vuolteenaho. 2004. "Bad Beta, Good Beta." *American Economic Review* 94(5): 1249–1275.
- Cenzon, Josefina. 2025. "Credit Market Experiences and Macroeconomic Expectations: Evidence and Theory." Working Paper.
- Chen, Long, and Xinlei Zhao. 2009. "Return Decomposition." *Review of Financial Studies* 22(12): 5213–5249.
- Choi, Hyunyoung, and Hal Varian. 2012. "Predicting the Present with Google Trends." *Economic Record* 88(s1): 2–9.
- Coibion, Olivier, Dimitris Georgarakos, Yuriy Gorodnichenko, and Maarten van Rooij. 2023. "How Does Consumption Respond to News about Inflation? Field Evidence from a Randomized Control Trial." *American Economic Journal: Macroeconomics* 15(3): 109–152.
- Coibion, Olivier, and Yuriy Gorodnichenko. 2015. "Information Rigidity and the Expectations Formation Process: A Simple Framework and New Facts." *American Economic Review* 105(8): 2644–2678.

- Coibion, Olivier, Yuriy Gorodnichenko, and Michael Weber. 2022. "Monetary Policy Communications and Their Effects on Household Inflation Expectations." *Journal of Political Economy* 130(6): 1537–1584.
- d’Acunto, Francesco, Ulrike Malmendier, Juan Ospina, and Michael Weber. 2021. "Exposure to Grocery Prices and Inflation Expectations." *Journal of Political Economy* 129(5): 1615–1639.
- Damasio, Antonio. 1994. *Descartes’ Error: Emotion, Reason, and the Human Brain*. New York: G. P. Putnam.
- Das, Sreyoshi, Camelia Kuhnen, and Stefan Nagel. 2020. "Socioeconomic Status and Macroeconomic Expectations." *Review of Financial Studies* 33(1): 395–432.
- De La O, Ricardo, and Sean Myers. 2021. "Subjective Cash Flow and Discount Rate Expectations." *Journal of Finance* 76(3): 1339–1387.
- De Long, J. Bradford, Andrei Shleifer, Lawrence H. Summers, and Robert J. Waldmann. 1990. "Noise Trader Risk in Financial Markets." *Journal of Political Economy* 98(4): 703–738.
- Dominitz, Jeff, and Charles F. Manski. 2007. "Expected Equity Returns and Portfolio Choice: Evidence from the Health and Retirement Study." *Journal of the European Economic Association* 5(2–3): 369–379.
- Dupraz, Stéphane, Emi Nakamura, and Jón Steinsson. 2025. "A Plucking Model of Business Cycles." *Journal of Monetary Economics* 152 103766.
- Eusepi, Stefano, and Bruce Preston. 2011. "Expectations, Learning, and Business Cycle Fluctuations." *American Economic Review* 101(6): 2844–2872.
- Fama, Eugene F. 1991. "Efficient Capital Markets: II." *Journal of Finance* 46(5): 1575–1617.
- Fama, Eugene F., and Kenneth R. French. 1993. "Common Risk Factors in the Returns on Stocks and Bonds." *Journal of Financial Economics* 33(1): 3–56.
- Flynn, Joel, and Karthik Sastry. 2025. "Contagious Business Cycles." Working Paper.
- Fuster, Andreas, David Laibson, and Brock Mendel. 2010. "Natural Expectations and Macroeconomic Fluctuations." *Journal of Economic Perspectives* 24(4): 67–84.
- Gabaix, Xavier. 2014. "A Sparsity-Based Model of Bounded Rationality." *Quarterly Journal of Economics* 129(4): 1661–1710.
- Gennaioli, Nicola, Marta Leva, Raphael Schoenle, and Andrei Shleifer. 2025. "How Inflation Expectations De-anchor: The Role of Selective Memory Cues." Working Paper.
- Giglio, Stefano, Matteo Maggiori, Johannes Stroebel, and Stephen Utkus. 2021. "Five Facts about Beliefs and Portfolios." *American Economic Review* 111(5): 1481–1522.

- Graeber, Thomas, Christopher Roth, and Florian Zimmermann. 2024. "Stories, Statistics, and Memory." *Quarterly Journal of Economics* 139(4): 2181–2225.
- Grandmont, Jean-Michel. 1977. "Temporary General Equilibrium Theory." *Econometrica* 45(3): 535–572.
- Greenwood, Robin, and Andrei Shleifer. 2014. "Expectations of Returns and Expected Returns." *Review of Financial Studies* 27(3): 714–746.
- Guerrieri, Luca, and Matteo Iacoviello. 2017. "Collateral Constraints and Macroeconomic Asymmetries." *Journal of Monetary Economics* 90: 28–49.
- Hall, Robert. 1978. "Stochastic Implications of the Life-Cycle Permanent Income Hypothesis: Theory and Evidence." *Journal of Political Economy* 86(6): 971–987.
- Hamilton, James D. 1989. "A New Approach to the Economic Analysis of Nonstationary Time Series and the Business Cycle." *Econometrica* 57(2): 357–384.
- Hansen, Lars Peter, and Thomas Sargent. 2010. "Fragile Beliefs and the Price of Uncertainty." *Quantitative Economics* 1(1): 129–162.
- Hassabis, Demis, Dharshan Kumaran, Seralynne Vann, and Eleanor Maguire. 2007. "Patients with Hippocampal Amnesia Cannot Imagine New Experiences." *Proceedings of the National Academy of Sciences* 104(5): 1726–1731.
- Hayashi, Fumio. 1982. "Tobin's Marginal q and Average q : A Neoclassical Interpretation." *Econometrica* 50(1): 213–224.
- Hicks, John R. 1939. *Value and Capital*. Oxford: Oxford University Press.
- Jensen, Henrik, Ivan Petrella, Søren Hove Ravn, and Emiliano Santoro. 2020. "Leverage and Deepening Business-Cycle Skewness." *American Economic Journal: Macroeconomics* 12(1): 245–281.
- Jiang, Zhengyang, Hongqi Liu, Cameron Peng, and Hongjun Yan. 2025. "Investor Memory and Biased Beliefs: Evidence from the Field." *Quarterly Journal of Economics* 140(4): 2749–2804.
- Jordà, Òscar. 2005. "Estimation and Inference of Impulse Responses by Local Projections." *American Economic Review* 95(1): 161–182.
- Jordà, Òscar, Moritz Schularick, and Alan M. Taylor. 2024. "Disasters Everywhere: The Costs of Business Cycles Reconsidered." *IMF Economic Review* 72: 116–151.
- Kahana, Michael. 2012. *Foundations of Human Memory*. Oxford University Press.
- Kahneman, Daniel, and Amos Tversky. 1972. "Subjective Probability: A Judgment of Representativeness." *Cognitive Psychology* 3(3): 430–454.

- Kahneman, Daniel, and Amos Tversky. 1982. "On the Study of Statistical Intuitions." *Cognition* 11 123–141.
- Karlan, Dean, Margaret McConnell, Sendhil Mullainathan, and Jonathan Zinman. 2016. "Getting to the Top of Mind: How Reminders Increase Saving." *Management Science* 62(12): 3393–3411.
- Kensinger, Elizabeth. 2009. "Remembering the Details: Effects of Emotion." *Emotion Review* 1(2): 99–113.
- Keynes, John Maynard. 1921. *A Treatise on Probability*. London: Macmillan.
- Keynes, John Maynard. 1936. *The General Theory of Employment, Interest and Money*. London: Macmillan.
- Kuchler, Theresa, and Basit Zafar. 2019. "Personal Experiences and Expectations about Aggregate Outcomes." *Journal of Finance* 74(5): 2491–2542.
- La Porta, Rafael. 1996. "Expectations and the Cross-Section of Stock Returns." *Journal of Finance* 51(5): 1715–1742.
- Lakonishok, Josef, Andrei Shleifer, and Robert W. Vishny. 1994. "Contrarian Investment, Extrapolation, and Risk." *Journal of Finance* 49(5): 1541–1578.
- Lee, Charles, Andrei Shleifer, and Richard Thaler. 1991. "Investor Sentiment and the Closed-End Fund Puzzle." *Journal of Finance* 46(1): 75–109.
- Link, Sebastian, Andreas Peichl, Oliver Pfäuti, Christopher Roth, and Johannes Wohlfart. 2023. "Attention to the Macroeconomy." CESifo Working Paper No. 10858.
- Loewenstein, George. 2000. "Emotions in Economic Theory and Economic Behaviour." *American Economic Review* 90(2): 426–432.
- Lorenzoni, Guido. 2009. "A Theory of Demand Shocks." *American Economic Review* 99(5): 2050–2084.
- Lucas, Robert. 1972. "Expectations and the Neutrality of Money." *Journal of Economic Theory* 4(2): 103–124.
- Luce, R. Duncan. 1959. *Individual Choice Behavior: A Theoretical Analysis*. New York: Wiley.
- Ludvigson, Sydney. 2004. "Consumer Confidence and Consumer Spending." *Journal of Economic Perspectives* 18(2): 29–50.
- Malmendier, Ulrike, and Stefan Nagel. 2011. "Depression Babies: Do Macroeconomic Experiences Affect Risk Taking?" *Quarterly Journal of Economics* 126(1): 373–416.
- Malmendier, Ulrike, and Stefan Nagel. 2016. "Learning from Inflation Experiences." *Quarterly Journal of Economics* 131(1): 53–87.

- Mankiw, Gregory. 2009. "Shiller on Animal Spirits." Greg Mankiw's Blog, January 27, <https://gregmankiw.blogspot.com/2009/01/shiller-on-animal-spirits.html>.
- Mankiw, Gregory, and Ricardo Reis. 2002. "Sticky Information versus Sticky Prices: A Proposal to Replace the New Keynesian Phillips Curve." *Quarterly Journal of Economics* 117(4): 1295–1328.
- Mankiw, Gregory, Ricardo Reis, and Justin Wolfers. 2003. "Disagreement about Inflation Expectations." *NBER Macroeconomics Annual* 18: 209–248.
- Nagel, Stefan, and Zhengyang Xu. 2023. "Dynamics of Subjective Risk Premia." *Journal of Financial Economics* 150(2): 103713.
- Nosofsky, Robert. 1986. "Attention, Similarity, and the Identification–Categorization Relationship." *Journal of Experimental Psychology: General* 115(1): 39–57.
- Phelps, Edmund. 1970. "Introduction: The New Microeconomics in Employment and Inflation Theory." In *Microeconomic Foundations of Employment and Inflation Theory*, 1–23, New York: W. W. Norton.
- Politis, Dimitris N., and Joseph P. Romano. 1994. "The Stationary Bootstrap." *Journal of the American Statistical Association* 89(428): 1303–1313.
- Salgado, Sergio, Fatih Guvenen, and Nicholas Bloom. 2019. "Skewed Business Cycles." NBER Working Paper No. 26565.
- Schacter, Daniel, Donna Addis, Demis Hassabis, Victoria Martin, Nathan Spreng, and Karl Szpunar. 2012. "The Future of Memory: Remembering, Imagining, and the Brain." *Neuron* 76(4): 677–694.
- Schwartzstein, Joshua, and Adi Sunderam. 2021. "Using Models to Persuade." *American Economic Review* 111(1): 276–323.
- Schwarz, Norbert, and Gerald Clore. 1983. "Mood, Misattribution, and Judgments of Well-Being: Informative and Directive Functions of Affective States." *Journal of Personality and Social Psychology* 45(3): 513–523.
- Shapiro, Adam Hale, Moritz Sudhof, and Daniel Wilson. 2022. "Measuring News Sentiment." *Journal of Econometrics* 228(2): 221–243.
- Sharot, Tali, Alison Riccardi, Raymond Dolan, and Elizabeth Phelps. 2007. "Neural Mechanisms Mediating Optimism." *Nature* 450(7166): 102–105.
- Shiller, Robert. 1981. "Do Stock Prices Move Too Much to Be Justified by Subsequent Changes in Dividends?" *American Economic Review* 71(3): 421–436.
- Shiller, Robert. 2019a. *Narrative Economics: How Stories Go Viral and Drive Major Economic Events*. Princeton University Press.

- Shiller, Robert. 2019b. “Narratives about Technology-Induced Job Degradation Then and Now.” *Journal of Policy Modeling* 41(3): 477–488.
- Shleifer, Andrei. 2000. *Inefficient Markets: An Introduction to Behavioral Finance*. Oxford University Press.
- Shleifer, Andrei, and Robert W. Vishny. 1997. “The Limits of Arbitrage.” *Journal of Finance* 52(1): 35–55.
- Sims, Christopher. 2003. “Implications of Rational Inattention.” *Journal of Monetary Economics* 50(3): 665–690.
- Taubinsky, Dmitry, Luigi Butera, Matteo Saccarola, and Chen Lian. 2024. “Beliefs About the Economy Are Excessively Sensitive to Household-Level Shocks: Evidence from Linked Survey and Administrative Data.” NBER Working Paper No. 32664.
- Vuolteenaho, Tuomo. 2002. “What Drives Firm-Level Stock Returns?” *Journal of Finance* 57(1): 233–264.
- Weber, Michael, Francesco D’Acunto, Yuriy Gorodnichenko, and Olivier Coibion. 2022. “The Subjective Inflation Expectations of Households and Firms: Measurement, Determinants, and Implications.” *Journal of Economic Perspectives* 36(3): 157–184.
- Woodford, Michael. 2003. “Imperfect Common Knowledge and the Effects of Monetary Policy.” In *Knowledge, Information, and Expectations in Modern Macroeconomics*, edited by Aghion, Philippe, Roman Frydman, Joseph Stiglitz, and Michael Woodford 25–58, Princeton, NJ: Princeton University Press.

Appendix A

Appendix to Chapter 1

A.1 Theory Derivations and Proofs

This appendix gives the algebra behind the theoretical results in Section 1.2. Throughout, constants from the log-linear return approximation are absorbed into intercepts. I use the same timing convention as in the body: rx_{t+1} is the excess log return from t to $t + 1$, rf_{t+1} is the 1-period risk-free rate, and Δd_{t+1} is log dividend growth. The rational expectations operator is $\mathbb{E}_t^{(r)}$. The behavioral expectations operator is $\mathbb{E}_t^{(b)}$.

A.1.1 Return-News Identity and Timing

The Campbell and Ammer (1993) return-news identity used in the paper is

$$rx_{t+1} - \mathbb{E}_t^{(r)} rx_{t+1} = N_{CF,t+1}^{(r)} - N_{RF,t+1}^{(r)} - N_{RP,t+1}^{(r)}.$$

The three rational news terms are

$$\begin{aligned} N_{CF,t+1}^{(r)} &= \left(\mathbb{E}_{t+1}^{(r)} - \mathbb{E}_t^{(r)} \right) \sum_{j \geq 1} \rho^j \Delta d_{t+1+j}, \\ N_{RF,t+1}^{(r)} &= \left(\mathbb{E}_{t+1}^{(r)} - \mathbb{E}_t^{(r)} \right) \sum_{j \geq 1} \rho^j r f_{t+1+j}, \\ N_{RP,t+1}^{(r)} &= \left(\mathbb{E}_{t+1}^{(r)} - \mathbb{E}_t^{(r)} \right) \sum_{j \geq 1} \rho^j r x_{t+1+j}. \end{aligned}$$

The sign convention is important. Positive cash-flow news raises unexpected returns. Positive risk-free-rate news or positive risk premium news lowers unexpected returns because it raises the discount rate applied to future cash flows. Thus $N_{RP,t+1}^{(r)}$ is the rational revision in expected future excess returns, with a negative sign in the unexpected-return identity.

The superscript (r) indicates the expectations operator used to construct the news terms. It does not require that the prices being decomposed were set by investors with rational beliefs. This distinction is the key to the paper's reinterpretation of risk premium news.

A.1.2 Proof of Proposition 1: Full Behavioral-Pricing Benchmark

The full behavioral-pricing benchmark imposes three assumptions. First, behavioral investors are marginal in price setting. Second, behavioral and rational investors have the same beliefs about the risk-free-rate stream. Third, behavioral investors perceive expected excess returns as constant. Let this constant perceived expected excess return be $\bar{r}x$.

Define behavioral cash-flow news by

$$N_{CF,t+1}^{(b)} = \left(\mathbb{E}_{t+1}^{(b)} - \mathbb{E}_t^{(b)} \right) \sum_{j \geq 1} \rho^j \Delta d_{t+1+j}.$$

The cash-flow belief-bias news is

$$\eta_{CF,t+1} \equiv N_{CF,t+1}^{(b)} - N_{CF,t+1}^{(r)}.$$

The rationally unexpected component of this wedge is

$$\zeta_{CF,t+1} \equiv \eta_{CF,t+1} - \mathbb{E}_t^{(r)} \eta_{CF,t+1}.$$

Under the benchmark assumptions, the realized excess return can be written from the behavioral pricing perspective as

$$rx_{t+1} = r\bar{x} + N_{CF,t+1}^{(b)} - N_{RF,t+1}^{(r)}. \quad (\text{A.1})$$

Risk-free-rate news is rational because rate beliefs are correct, and behavioral risk premium news is zero because perceived expected excess returns are constant. Using the definition of $\eta_{CF,t+1}$, equation (A.1) becomes

$$rx_{t+1} = r\bar{x} + N_{CF,t+1}^{(r)} - N_{RF,t+1}^{(r)} + \eta_{CF,t+1}. \quad (\text{A.2})$$

The same realized excess return is decomposed by the rational econometrician as

$$rx_{t+1} = \mathbb{E}_t^{(r)} rx_{t+1} + N_{CF,t+1}^{(r)} - N_{RF,t+1}^{(r)} - N_{RP,t+1}^{(r)}. \quad (\text{A.3})$$

Equating equations (A.2) and (A.3) gives

$$\mathbb{E}_t^{(r)}rx_{t+1} - N_{RP,t+1}^{(r)} = \bar{r}x + \eta_{CF,t+1}.$$

Rearranging,

$$N_{RP,t+1}^{(r)} = \mathbb{E}_t^{(r)}rx_{t+1} - \bar{r}x - \eta_{CF,t+1}. \quad (\text{A.4})$$

Taking $\mathbb{E}_t^{(r)}$ of equation (A.2), and using $\mathbb{E}_t^{(r)}N_{CF,t+1}^{(r)} = \mathbb{E}_t^{(r)}N_{RF,t+1}^{(r)} = 0$, yields

$$\mathbb{E}_t^{(r)}rx_{t+1} = \bar{r}x + \mathbb{E}_t^{(r)}\eta_{CF,t+1}. \quad (\text{A.5})$$

Substituting equation (A.5) into equation (A.4) gives

$$N_{RP,t+1}^{(r)} = \mathbb{E}_t^{(r)}\eta_{CF,t+1} - \eta_{CF,t+1} = -\zeta_{CF,t+1}.$$

This proves Proposition 1.

The proof shows that the result is an accounting reinterpretation. The rational econometrician uses the same Campbell and Ammer (1993) decomposition, but because the econometrician does not directly include the behavioral cash-flow wedge, the unexpected component of that wedge appears with the opposite sign in rational risk premium news.

A.1.3 Behavioral Valuation Wedge

This subsection connects the belief-bias news $\eta_{CF,t+1}$ to the behavioral valuation wedge used in the partial-arbitrage extension.

Let pd_t^* denote the rational log price-dividend ratio that would prevail absent the behavioral belief distortion. Let $pd_t^{(b)}$ denote the log price-dividend ratio implied by behavioral cash-flow beliefs under the full behavioral-pricing benchmark. Define the behavioral valuation wedge

$$m_t \equiv pd_t^{(b)} - pd_t^*.$$

The log-linear excess-return identity implies that a change in the valuation wedge affects realized excess returns according to

$$rx_{t+1}^{(b)} - rx_{t+1}^* = \rho m_{t+1} - m_t,$$

where $rx_{t+1}^{(b)}$ is the return generated by the behavioral valuation and rx_{t+1}^* is the return generated by the rational valuation, holding the realized cash-flow stream fixed.

Under the full behavioral-pricing benchmark, behavioral and rational investors agree about risk-free rates and behavioral perceived risk premium news is zero. Therefore the only difference between the behavioral and rational valuation revisions is the cash-flow belief-bias news. Comparing the behavioral and rational decompositions gives

$$\rho m_{t+1} - m_t = N_{CF,t+1}^{(b)} - N_{CF,t+1}^{(r)} = \eta_{CF,t+1}. \quad (\text{A.6})$$

Equation (A.6) is the primitive wedge identity used in the partial-arbitrage derivation. It says that the return effect of the behavioral valuation wedge is exactly the cash-flow belief-bias news.

A.1.4 Proof of Proposition 2: Partial Arbitrage

The full benchmark assumes that the behavioral wedge m_t fully passes through to prices. The partial-arbitrage model allows rational arbitrageurs to offset part of this wedge. Let

$$pd_t^\lambda = pd_t^* + \lambda m_t, \quad 0 \leq \lambda \leq 1.$$

The parameter λ is the surviving share of the behavioral valuation wedge. If $\lambda = 1$, the wedge fully affects prices. If $\lambda = 0$, arbitrage fully removes it from prices.

Using the log-linear return identity, the return under partial arbitrage differs from the rational return by

$$rx_{t+1}^\lambda - rx_{t+1}^* = \lambda(\rho m_{t+1} - m_t).$$

By equation (A.6),

$$rx_{t+1}^\lambda - rx_{t+1}^* = \lambda \eta_{CF,t+1}. \quad (\text{A.7})$$

Taking the rationally unexpected component of equation (A.7) gives

$$\left(rx_{t+1}^\lambda - \mathbb{E}_t^{(r)} rx_{t+1}^\lambda \right) - \left(rx_{t+1}^* - \mathbb{E}_t^{(r)} rx_{t+1}^* \right) = \lambda \zeta_{CF,t+1}. \quad (\text{A.8})$$

Now decompose the two unexpected returns using the rational Campbell and Ammer (1993) identity. For the partial-arbitrage economy,

$$rx_{t+1}^\lambda - \mathbb{E}_t^{(r)} rx_{t+1}^\lambda = N_{CF,t+1}^{(r)} - N_{RF,t+1}^{(r)} - N_{RP,t+1}^{\lambda,r}.$$

For the rational benchmark economy,

$$rx_{t+1}^* - \mathbb{E}_t^{(r)} rx_{t+1}^* = N_{CF,t+1}^{(r)} - N_{RF,t+1}^{(r)} - N_{RP,t+1}^{*,r}.$$

Subtracting the second equation from the first gives

$$\left(rx_{t+1}^\lambda - \mathbb{E}_t^{(r)} rx_{t+1}^\lambda \right) - \left(rx_{t+1}^* - \mathbb{E}_t^{(r)} rx_{t+1}^* \right) = -N_{RP,t+1}^{\lambda,r} + N_{RP,t+1}^{*,r}.$$

Combining with equation (A.8) yields

$$-N_{RP,t+1}^{\lambda,r} + N_{RP,t+1}^{*,r} = \lambda \zeta_{CF,t+1}.$$

Therefore

$$N_{RP,t+1}^{\lambda,r} = N_{RP,t+1}^{*,r} - \lambda \zeta_{CF,t+1}.$$

This proves Proposition 2.

A.1.5 Projection Coefficients and the Interpretation of λ

Proposition 2 implies

$$N_{RP,t+1}^{\lambda,r} = N_{RP,t+1}^{*,r} - \lambda \zeta_{CF,t+1}.$$

Projecting $N_{RP,t+1}^{\lambda,r}$ on $\zeta_{CF,t+1}$ gives

$$\frac{\text{Cov}\left(N_{RP,t+1}^{\lambda,r}, \zeta_{CF,t+1}\right)}{\text{Var}(\zeta_{CF,t+1})} = \frac{\text{Cov}\left(N_{RP,t+1}^{*,r}, \zeta_{CF,t+1}\right)}{\text{Var}(\zeta_{CF,t+1})} - \lambda.$$

Thus the projection coefficient identifies $-\lambda$ only if

$$\text{Cov}(N_{RP,t+1}^{*,r}, \zeta_{CF,t+1}) = 0.$$

This orthogonality condition is not implied by the model. It is an additional identifying assumption. If residual rational risk premium news comoves with belief-bias news, the projection coefficient combines the surviving behavioral loading and that residual rational comovement.

This is why the main empirical analysis treats a negative relation between $N_{RP}^{(r)}$ and $\zeta^{(B)}$ as evidence of a belief-driven component, not as a direct structural estimate of λ .

A.1.6 Risk-Free-Rate Belief Bias and Net Valuation Beliefs

The benchmark in Proposition 1 assumes that behavioral and rational investors agree about the risk-free-rate stream. This subsection allows behavioral investors to hold biased beliefs about future risk-free rates.

Define behavioral risk-free-rate news as

$$N_{RF,t+1}^{(b)} = \left(\mathbb{E}_{t+1}^{(b)} - \mathbb{E}_t^{(b)} \right) \sum_{j \geq 1} \rho^j r f_{t+1+j}.$$

Define risk-free-rate belief-bias news as

$$\eta_{RF,t+1} \equiv N_{RF,t+1}^{(b)} - N_{RF,t+1}^{(r)},$$

and its rationally unexpected component as

$$\zeta_{RF,t+1} \equiv \eta_{RF,t+1} - \mathbb{E}_t^{(r)} \eta_{RF,t+1}.$$

With biased cash-flow and rate beliefs, the behavioral pricing equation becomes

$$rx_{t+1} = r\bar{x} + N_{CF,t+1}^{(b)} - N_{RF,t+1}^{(b)}.$$

Using the definitions of $\eta_{CF,t+1}$ and $\eta_{RF,t+1}$,

$$rx_{t+1} = r\bar{x} + N_{CF,t+1}^{(r)} - N_{RF,t+1}^{(r)} + \eta_{CF,t+1} - \eta_{RF,t+1}. \quad (\text{A.9})$$

Equating equation (A.9) with the rational decomposition gives

$$N_{RP,t+1}^{(r)} = \mathbb{E}_t^{(r)} rx_{t+1} - r\bar{x} - \eta_{CF,t+1} + \eta_{RF,t+1}. \quad (\text{A.10})$$

Taking rational expectations of equation (A.9) yields

$$\mathbb{E}_t^{(r)} rx_{t+1} = r\bar{x} + \mathbb{E}_t^{(r)} \eta_{CF,t+1} - \mathbb{E}_t^{(r)} \eta_{RF,t+1}.$$

Substituting into equation (A.10) gives

$$\begin{aligned} N_{RP,t+1}^{(r)} &= - \left(\eta_{CF,t+1} - \mathbb{E}_t^{(r)} \eta_{CF,t+1} \right) + \left(\eta_{RF,t+1} - \mathbb{E}_t^{(r)} \eta_{RF,t+1} \right) \\ &= -\zeta_{CF,t+1} + \zeta_{RF,t+1}. \end{aligned} \quad (\text{A.11})$$

Thus the exact theoretical object is net valuation belief bias. Define

$$\zeta_{NET,t+1} \equiv \zeta_{CF,t+1} - \zeta_{RF,t+1}.$$

Then equation (A.11) can be written as $N_{RP,t+1}^{(r)} = -\zeta_{NET,t+1}$.

With partial arbitrage, the same logic gives

$$N_{RP,t+1}^{\lambda,r} = N_{RP,t+1}^{*,r} - \lambda_{CF} \zeta_{CF,t+1} + \lambda_{RF} \zeta_{RF,t+1},$$

where λ_{CF} and λ_{RF} measure the pass-through of cash-flow and rate-belief valuation wedges. The main empirical analysis focuses on cash-flow beliefs because analyst forecasts directly measure cash-flow expectations. If rate-belief bias is empirically important and correlated with cash-flow belief bias, then the analyst proxy measures only one component of the broader net valuation belief wedge.

A.1.7 Generalized Forecast Relation

This subsection derives the generalized forecast relation stated in Section 1.2.6. Let z_t denote the vector of variables used to forecast returns and belief revisions. Write the relevant conditional expectations as linear projections:

$$\begin{aligned}
\mathbb{E}_t^{(r)} r x_{t+1}^\lambda &= a_{rx} + \gamma_{rx}^\lambda z_t, \\
\mathbb{E}_t^{(r)} r x_{t+1}^* &= a_* + \gamma^* z_t, \\
\mathbb{E}_t^{(r)} \eta_{CF,t+1} &= a_{CF} + \gamma_{CF} z_t, \\
\mathbb{E}_t^{(r)} \eta_{RF,t+1} &= a_{RF} + \gamma_{RF} z_t.
\end{aligned}$$

Under the combined partial-arbitrage and rate-belief-bias model, the conditional expected excess return can be written as

$$\mathbb{E}_t^{(r)} r x_{t+1}^\lambda = \text{const.} + \mathbb{E}_t^{(r)} r x_{t+1}^* + \lambda_{CF} \mathbb{E}_t^{(r)} \eta_{CF,t+1} - \lambda_{RF} \mathbb{E}_t^{(r)} \eta_{RF,t+1}.$$

Matching slopes on z_t gives

$$\gamma_{rx}^\lambda = \gamma^* + \lambda_{CF} \gamma_{CF} - \lambda_{RF} \gamma_{RF}. \quad (\text{A.12})$$

The sharp full behavioral-pricing benchmark is the special case $\gamma^* = 0$, $\lambda_{CF} = 1$, $\lambda_{RF} = 0$. In that case, $\gamma_{rx}^\lambda = \gamma_{CF}$, which is the 1-step slope restriction tested in the main empirical analysis.

Equation (A.12) also clarifies the interpretation of rejections. Exact slope equality can fail because of residual rational return predictability ($\gamma^* \neq 0$), partial pass-through of cash-flow belief bias ($\lambda_{CF} \neq 1$), rate-belief bias ($\lambda_{RF} \neq 0$), or measurement error in the empirical belief proxy. Such failures do not by themselves imply that the belief-driven component is absent.

A.1.8 Perceived vs. Rational Expected Returns

The full benchmark assumes that behavioral investors perceive expected excess returns as constant. This is a statement about the beliefs used in behavioral price setting, not about the rational econometrician's conditional expected return. In fact, Proposition 1 implies that the rational econometrician's expected excess return varies with predictable belief-bias news:

$$\mathbb{E}_t^{(r)} r x_{t+1} = \bar{r} \bar{x} + \mathbb{E}_t^{(r)} \eta_{CF,t+1}.$$

Thus the benchmark does not rule out time-varying rational expected returns in the econometrician's forecasting system. It explains why those expected returns vary: the econometrician sees price movements generated by biased cash-flow beliefs and interprets them as changes in future expected excess returns. This distinction is what allows the paper to preserve the standard return decomposition while changing the interpretation of the risk premium news object.

A.2 Forecast, Matrix, and Horizon Restrictions

This appendix derives the empirical restrictions used in the body. It connects the theoretical results in Appendix A.1 to the linear forecasting system used to construct VAR-implied news objects. Appendix A.2.1 fixes the forecasting-system notation. Appendices A.2.2–A.2.4 derive the 1-step slope restriction, the long-horizon matrix condition, and the matrix-wedge object. Appendices A.2.5–A.2.6 derive the finite-horizon present-value matrices and the corresponding horizon- h forecast restrictions. Appendices A.2.7–A.2.8 explain why 1-step slope equality and pointwise long-horizon news equality are distinct implications of the full benchmark. They also clarify why a 5-year horizon is economically meaningful but not the infinite-horizon Campbell and Ammer (1993) object.

A.2.1 Linear Forecasting System

The body summarizes the forecasting system used throughout the empirical analysis. This subsection states the system precisely. The state vector at time t is

$$x_t = [rx_t, rf_t, \nu_t^{(B)}, b_t^{(B)}, pd_t, ts_t]'$$

matching equation (1.12) in Section 1.3.1. The dynamics follow a first-order vector autoregression

$$x_{t+1} = \alpha + \Gamma x_t + \varepsilon_{t+1}, \quad \mathbb{E}_t^{(r)} \varepsilon_{t+1} = 0,$$

with intercept α , transition matrix Γ , and innovation covariance matrix $\Sigma_\varepsilon = \mathbb{E}(\varepsilon_{t+1}\varepsilon_{t+1}')$.

The selection vectors e_1 , e_2 , and e_3 pick the excess return, the risk-free rate, and the

behavioral cash-flow news proxy from the state vector:

$$e'_1 x_t = r x_t, \quad e'_2 x_t = r f_t, \quad e'_3 x_t = \nu_t^{(B)}.$$

The infinite-horizon present-value matrix is

$$\Lambda = \rho \Gamma (I - \rho \Gamma)^{-1} = \sum_{j=1}^{\infty} \rho^j \Gamma^j, \quad (\text{A.13})$$

which exists when the spectral radius of $\rho \Gamma$ is less than one. The matrix Λ maps current innovations into present-value revisions of expected future state variables.

A.2.2 One-Step Forecast Restriction

Under the full behavioral-pricing benchmark, the rational econometrician's 1-step expected excess return equals $r\bar{x}$ plus the rational expectation of cash-flow belief-bias news, as shown in equation (A.5) of Appendix A.1.2:

$$\mathbb{E}_t^{(r)} r x_{t+1} = r\bar{x} + \mathbb{E}_t^{(r)} \eta_{CF,t+1}.$$

Empirically, $\eta_{t+1}^{(B)} = \nu_{t+1}^{(B)} - N_{CF,t+1}^{(r)}$ (Section 1.3.4). Since $N_{CF,t+1}^{(r)}$ is a rational news innovation, $\mathbb{E}_t^{(r)} N_{CF,t+1}^{(r)} = 0$. Thus the conditional expectation of the empirical belief-bias wedge equals the conditional expectation of the analyst-revision proxy, up to measurement error and intercept normalizations:

$$\mathbb{E}_t^{(r)} \eta_{t+1}^{(B)} = \mathbb{E}_t^{(r)} \nu_{t+1}^{(B)}.$$

Both sides are linear in x_t , so comparing the conditional-expectation row for rx_{t+1} with the row for $\nu_{t+1}^{(B)}$ gives

$$e'_1 \Gamma = e'_3 \Gamma. \quad (\text{A.14})$$

This is the 1-step slope restriction. It is a six-restriction equality on the slope vectors (one per state-vector slot). Inference uses the null-imposed bootstrap described in Appendix A.5.8.

Equation (A.14) follows from the full benchmark: $\lambda_{CF} = 1$, $\gamma^* = 0$, $\lambda_{RF} = 0$ in the generalized forecast relation (A.12). Under the broader partial-arbitrage and rate-belief-bias model (Appendix A.1.7), this equality need not hold even when the belief-driven channel is present. Rejection of equation (A.14) demotes the full-benchmark interpretation but does not, by itself, imply that the belief-driven component is absent.

A.2.3 Matrix Condition

The full benchmark also implies a stronger pointwise restriction on the constructed news objects. In the linear forecasting system, risk premium news is

$$N_{RP,t+1}^{(r)} = e'_1 \Lambda \varepsilon_{t+1}, \quad (\text{A.15})$$

risk-free-rate news is

$$N_{RF,t+1}^{(r)} = e'_2 \Lambda \varepsilon_{t+1},$$

and cash-flow news is recovered from the Campbell and Ammer (1993) identity:

$$N_{CF,t+1}^{(r)} = e'_1 \varepsilon_{t+1} + N_{RF,t+1}^{(r)} + N_{RP,t+1}^{(r)} = (e'_1 + e'_2 \Lambda + e'_1 \Lambda) \varepsilon_{t+1}.$$

The empirical belief-bias innovation is constructed in Section 1.3.4 as $\eta_{t+1}^{(B)} = \nu_{t+1}^{(B)} - N_{CF,t+1}^{(r)}$, with rationally unexpected component $\zeta_{t+1}^{(B)} = \eta_{t+1}^{(B)} - \mathbb{E}_t^{(r)} \eta_{t+1}^{(B)}$. In the linear system,

$$\zeta_{t+1}^{(B)} = (e'_3 - e'_1 - e'_2 \Lambda - e'_1 \Lambda) \varepsilon_{t+1}, \quad (\text{A.16})$$

since both $\nu^{(B)}$ and $N_{CF}^{(r)}$ are themselves linear in ε_{t+1} .

Proposition 1 (Appendix A.1.2) implies $N_{RP,t+1}^{(r)} + \zeta_{CF,t+1} = 0$ for every realization of ε_{t+1} . Substituting equations (A.15) and (A.16), the pointwise condition becomes

$$(e'_3 - e'_1 - e'_2 \Lambda) \varepsilon_{t+1} = 0 \quad \text{for all } \varepsilon_{t+1},$$

which holds (when Σ_ε has full rank) if and only if

$$e'_3 = e'_1 + e'_2 \Lambda. \quad (\text{A.17})$$

This is the matrix condition. It is a $\dim(x)$ -restriction equality on the loading vectors of the VAR.

A.2.4 Matrix Wedge

The matrix wedge measures the failure of equation (A.17) pointwise. Define

$$W_{t+1}^{\text{matrix}} \equiv N_{RP,t+1}^{(r)} + \zeta_{t+1}^{(B)}.$$

Substituting the linear forms in equations (A.15) and (A.16) gives

$$W_{t+1}^{\text{matrix}} = (e'_3 - e'_1 - e'_2 \Lambda) \varepsilon_{t+1}. \quad (\text{A.18})$$

Under the full benchmark, equation (A.17) implies $W_{t+1}^{\text{matrix}} = 0$ for every ε_{t+1} .

Under partial arbitrage with cash-flow belief bias only (Appendix A.1.4), $N_{RP}^{\lambda,r} = N_{RP}^{*,r} - \lambda \zeta_{CF}$, so the matrix wedge becomes

$$W_{t+1}^{\text{matrix}} = N_{RP,t+1}^{*,r} + (1 - \lambda) \zeta_{CF,t+1}.$$

Under the additional orthogonality condition $\text{Cov}(N_{RP}^{*,r}, \zeta_{CF}) = 0$, the variance of the wedge is

$$\text{Var}(W^{\text{matrix}}) = \text{Var}(N_{RP}^{*,r}) + (1 - \lambda)^2 \text{Var}(\zeta_{CF}).$$

Adding rate-belief bias gives

$$W_{t+1}^{\text{matrix}} = N_{RP,t+1}^{*,r} + (1 - \lambda_{CF}) \zeta_{CF,t+1} + \lambda_{RF} \zeta_{RF,t+1}.$$

A nonzero empirical wedge is therefore consistent with residual rational risk premium news, partial pass-through of cash-flow beliefs, omitted rate-belief bias, and proxy error. The wedge measures distance from the strongest pointwise benchmark; it is not, by itself, evidence against the existence of a belief-driven component.

A.2.5 Finite-Horizon Present-Value Matrices

The infinite-horizon matrix Λ in equation (A.13) is the limit of finite-horizon partial sums.

Define

$$\Lambda_h = \sum_{j=1}^h \rho^j \Gamma^j.$$

The omitted tail is

$$\Lambda - \Lambda_h = \rho^{h+1} \Gamma^{h+1} (I - \rho \Gamma)^{-1}, \quad (\text{A.19})$$

which exists when $(I - \rho \Gamma)$ is invertible. The norm of the omitted tail decays at rate $\rho^{h+1} \|\Gamma\|^{h+1}$, governed by the spectral radius of $\rho \Gamma$. Appendix A.2.8 below quantifies this convergence in the legacy six-variable VAR.

A.2.6 Horizon Forecast Restrictions

Finite-horizon forecast restrictions use a closely related matrix:

$$B_h = \sum_{j=1}^h \rho^{j-1} \Gamma^j = \rho^{-1} \Lambda_h.$$

The horizon- h forecast restriction implied by the full benchmark is

$$e_1' B_h = e_3' B_h. \quad (\text{A.20})$$

This says that the state variables should forecast discounted cumulative excess returns and discounted cumulative belief revisions with the same slopes through horizon h .

In the iterated VAR system, equation (A.14) mechanically implies equation (A.20) for ev-

ery h . The argument is direct. If $e'_1\Gamma = e'_3\Gamma$, then right-multiplying both sides by Γ^{j-1} gives $e'_1\Gamma^j = e'_3\Gamma^j$ for every $j \geq 1$. Summing $\rho^{j-1}\Gamma^j$ over $j = 1, \dots, h$ yields $e'_1B_h = e'_3B_h$.

This is the basis for the body's claim that the 1-step slope restriction inside the iterated VAR system implies horizon- h slope restrictions for every h . Direct local-projection slope estimates are not constrained by the VAR's iterated dynamics, so they need not satisfy the same restriction at every horizon. The body's 5-year direct local-projection (LP) test in Section 1.4.2 is therefore a separate test from the iterated implication.

A.2.7 Slope vs. Matrix Restrictions

The 1-step slope restriction (A.14) and the matrix condition (A.17) are distinct implications of the full benchmark. The slope restriction is a statement about forecast rows,

$$e'_1\Gamma = e'_3\Gamma.$$

The matrix condition is a statement about innovation loadings and the full present-value map,

$$e'_3 = e'_1 + e'_2\Lambda.$$

Neither restriction generally implies the other.

To see why slope equality does not imply matrix equality, consider any VAR in which the first and third rows of Γ are equal. Then $e'_1\Gamma = e'_3\Gamma$ holds by construction. But the matrix condition additionally requires the level loading $e'_3 - e'_1$ to equal $e'_2\Lambda$. This requirement depends on the risk-free-rate row and the entire present-value matrix $\Lambda = \rho\Gamma(I - \rho\Gamma)^{-1}$. Row equality

of $e'_1\Gamma$ and $e'_3\Gamma$ does not impose that condition. Thus the 1-step slope restriction can hold while the matrix wedge is nonzero.

Conversely, the matrix condition can hold without the 1-step slope restriction. The condition $e'_3 = e'_1 + e'_2\Lambda$ restricts the innovation loadings that make pointwise news equality hold. It does not require the 1-step forecast rows $e'_1\Gamma$ and $e'_3\Gamma$ to be identical. Multiplying the matrix condition by Γ gives

$$e'_3\Gamma = e'_1\Gamma + e'_2\Lambda\Gamma,$$

so the slope restriction would require the additional condition $e'_2\Lambda\Gamma = 0$, which need not hold.

These directional non-equivalences explain a feature of the empirical results that might otherwise appear puzzling. In the legacy VAR, the slope restriction is not rejected (Section 1.4.1), while the matrix wedge is economically meaningful (Section 1.5.2). Imposing the slope restriction also does not mechanically eliminate the matrix wedge: in the data, the slope-restricted variance ratio is roughly 2.4–3.2, larger than the unrestricted ratio of about 1.2. The two restrictions test different facets of the full benchmark.

A.2.8 Five-Year Horizons and Infinite-Horizon News

The 5-year horizon $h = 60$ months is economically meaningful – it connects to long-horizon return-predictability evidence – but it is not the infinite-horizon Campbell and Ammer (1993) object. This subsection quantifies the difference.

The omitted tail in equation (A.19) is governed by the spectral radius of $\rho\Gamma$. In the legacy six-variable VAR with $\rho_{\text{monthly}} = 0.998122$, the dominant eigenvalue of $\rho\Gamma$ is approximately 0.977. The dominant persistence mode implies that the omitted tail decays roughly at rate

0.977^{h+1} . This is not an exact matrix-norm identity, but it is a useful scale calculation for the persistence of the present-value tail.

At $h = 60$ months, the dominant-mode tail factor is $0.977^{61} \approx 0.24$. Thus the 5-year truncation can still omit a nontrivial part of the infinite-horizon present-value object. At $h = 120$, the corresponding factor is about 0.06, and by $h = 240$ it is small.

The implication for the empirical analysis is that direct cumulative-equality tests at $h = 60$ ask a different question than the matrix condition at $h = \infty$. A rejection of the 5-year cumulative restriction is therefore a rejection of the truncated direct-horizon restriction, not by itself a rejection of the infinite-horizon Campbell and Ammer (1993) pointwise condition. Section 1.5.3 reports the empirical horizon profile and shows that finite-horizon wedge ratios converge toward the infinite-horizon matrix-wedge values reported in Table 1.5.

A.3 Measurement Error and Forecast R^2

This appendix gives the measurement-error algebra behind Sections 1.2.9 and 1.6.1. The main point is that measurement error does not have a single effect. Additive noise can inflate the measured variance scale of the belief-bias proxy while attenuating correlations and projection coefficients. Signal attenuation and incomplete-proxy models can instead make the observed proxy understate the full behavioral component. The empirical analysis therefore interprets scale, sign, forecast comovement, slope restrictions, and robustness exercises jointly.

A.3.1 Additive Classical Noise

Let the true unexpected cash-flow belief bias be $\zeta_{CF,t+1}$, and suppose the observed proxy is

$$\tilde{\zeta}_{t+1} = \zeta_{CF,t+1} + u_{t+1},$$

where u_{t+1} is classical additive noise:

$$\mathbb{E}(u_{t+1}) = 0, \quad \text{Cov}(u_{t+1}, \zeta_{CF,t+1}) = 0, \quad \text{Cov}(u_{t+1}, N_{RP,t+1}^{(r)}) = 0.$$

Then

$$\text{Var}(\tilde{\zeta}_{t+1}) = \text{Var}(\zeta_{CF,t+1}) + \text{Var}(u_{t+1}).$$

The measured scale statistic is therefore

$$\tilde{\mathcal{B}}_{\text{scale}} = \frac{\text{Var}(\tilde{\zeta}_{t+1})}{\text{Var}(N_{RP,t+1}^{(r)})} = \mathcal{B}_{\text{scale}} + \frac{\text{Var}(u_{t+1})}{\text{Var}(N_{RP,t+1}^{(r)})}.$$

Thus additive classical noise inflates the numerator of $\mathcal{B}_{\text{scale}}$. In this model, the measured scale ratio is not conservative.

The same noise attenuates correlations. Let $\rho_{N\zeta} = \text{Corr}(N_{RP,t+1}^{(r)}, \zeta_{CF,t+1})$. Then

$$\text{Corr}(N_{RP,t+1}^{(r)}, \tilde{\zeta}_{t+1}) = \rho_{N\zeta} \left(\frac{\text{Var}(\zeta_{CF,t+1})}{\text{Var}(\zeta_{CF,t+1}) + \text{Var}(u_{t+1})} \right)^{1/2}.$$

The sign is preserved, but the magnitude moves toward zero.

Projection coefficients are attenuated for the same reason. If the true linear projection is

$$N_{RP,t+1}^{(r)} = \beta \zeta_{CF,t+1} + v_{t+1}, \quad \text{Cov}(v_{t+1}, \zeta_{CF,t+1}) = 0,$$

then the coefficient from projecting $N_{RP,t+1}^{(r)}$ on the noisy proxy $\tilde{\zeta}_{t+1}$ is

$$\tilde{\beta} = \frac{\text{Cov}(N_{RP,t+1}^{(r)}, \tilde{\zeta}_{t+1})}{\text{Var}(\tilde{\zeta}_{t+1})} = \beta \frac{\text{Var}(\zeta_{CF,t+1})}{\text{Var}(\zeta_{CF,t+1}) + \text{Var}(u_{t+1})}.$$

Thus additive noise can make the measured belief-bias object look large in variance scale while making the sign correlation and projection evidence harder to detect.

A.3.2 Signal Attenuation

A different measurement model is signal attenuation:

$$\tilde{\zeta}_{t+1} = a \zeta_{CF,t+1} + u_{t+1}, \quad 0 < a < 1,$$

with u_{t+1} classical noise. Then

$$\text{Var}(\tilde{\zeta}_{t+1}) = a^2 \text{Var}(\zeta_{CF,t+1}) + \text{Var}(u_{t+1}).$$

The observed scale statistic is

$$\tilde{\mathcal{B}}_{\text{scale}} = a^2 \mathcal{B}_{\text{scale}} + \frac{\text{Var}(u_{t+1})}{\text{Var}(N_{RP,t+1}^{(r)})}.$$

Whether the measured scale statistic is conservative depends on the relative importance of attenuation and additive noise. If $u_{t+1} = 0$, then $\tilde{\mathcal{B}}_{\text{scale}} = a^2 \mathcal{B}_{\text{scale}}$, so the observed scale understates the true behavioral scale. If additive noise is large, the measured scale can instead overstate the scale of the true component.

Define the reliability of the proxy as the signal share of the measured variance:

$$\mathcal{R} \equiv \frac{a^2 \text{Var}(\zeta_{CF,t+1})}{a^2 \text{Var}(\zeta_{CF,t+1}) + \text{Var}(u_{t+1})}.$$

Under this model,

$$\text{Corr}(N_{RP,t+1}^{(r)}, \tilde{\zeta}_{t+1}) = \text{sign}(a) \text{Corr}(N_{RP,t+1}^{(r)}, \zeta_{CF,t+1}) \sqrt{\mathcal{R}}.$$

Thus reliability governs the attenuation of the correlation. The empirical reliability diagnostics in Appendix A.7 use cross-proxy information to assess whether the aligned analyst measures contain a substantial common signal rather than only idiosyncratic noise.

A.3.3 Incomplete Proxy Models

Analyst revisions may measure only part of the full valuation-relevant belief wedge. Write the true cash-flow belief-bias innovation as

$$\zeta_{CF,t+1} = \zeta_{M,t+1} + \zeta_{U,t+1},$$

where $\zeta_{M,t+1}$ is the component measured by analyst revisions and $\zeta_{U,t+1}$ is an unmeasured component. The empirical proxy is

$$\tilde{\zeta}_{t+1} = \zeta_{M,t+1} + u_{t+1}.$$

In this case, the observed scale statistic measures the scale of the observed component, not the full behavioral wedge. If the unmeasured component is economically important, then $\tilde{\mathcal{B}}_{\text{scale}}$ can understate the variance of the full belief-bias object.

To see the implication for risk premium news, suppose

$$N_{RP,t+1}^{(r)} = N_{RP,t+1}^{*,r} - \lambda_M \zeta_{M,t+1} - \lambda_U \zeta_{U,t+1}.$$

A projection on $\tilde{\zeta}_{t+1}$ identifies the relation between risk premium news and the measured belief component, not the total behavioral component. If $\zeta_{M,t+1}$ and $\zeta_{U,t+1}$ are correlated, the projection also combines the direct loading on the measured component with indirect information about the unmeasured component. This is one reason the empirical analysis is framed as a component test rather than an exclusivity test.

The incomplete-proxy interpretation also applies to rate-belief bias. The main empirical

proxies measure cash-flow beliefs. If risk-free-rate belief bias or other valuation-belief errors are important, the analyst-based proxy captures only the cash-flow part of the broader net valuation belief wedge.

A.3.4 Nonclassical Proxy Contamination

The preceding subsections assume classical measurement error. More difficult cases arise when the measurement error is correlated with prices, rational news, or the state vector. A general representation is

$$\tilde{\zeta}_{t+1} = \zeta_{CF,t+1} + u_{t+1},$$

with at least one of

$$\text{Cov}(u_{t+1}, N_{RP,t+1}^{(r)}) \neq 0, \quad \text{Cov}(u_{t+1}, \zeta_{CF,t+1}) \neq 0, \quad \mathbb{E}_t^{(r)} u_{t+1} \neq 0.$$

In these cases, simple attenuation formulas no longer apply.

One important example is price-inference contamination. If analysts update forecasts partly by observing prices, the proxy can inherit information from valuation ratios:

$$\tilde{\zeta}_{t+1} = \zeta_{CF,t+1} + \delta g(pd_t) + u_{t+1}.$$

This can create or amplify a relation between measured belief news and return-news objects if pd_t is also used in the VAR. Appendix A.8.1 studies this channel directly by removing the contemporaneous price-dividend projection from the belief proxy.

Another example is omitted rational risk premium news that is correlated with the belief

proxy:

$$N_{RP,t+1}^{(r)} = N_{RP,t+1}^{*,r} - \lambda \zeta_{CF,t+1}, \quad \text{Cov}(N_{RP,t+1}^{*,r}, \zeta_{CF,t+1}) \neq 0.$$

In this case, the observed relation between risk premium news and belief-bias news combines the behavioral loading and residual rational comovement. Appendices A.8.2 and A.8.3 study how large such residual components must be to weaken or erase the baseline relation, and Appendix A.8.4 interpolates between the orthogonal and targeted cases.

The main text therefore does not rely on a single measurement-error correction. It uses several diagnostics: cross-proxy consistency, reliability calculations, price-inference residualization, adversarial-residual exercises, and cross-sectional evidence.

A.3.5 Forecast R^2 : Sample, Population, and Bias-Corrected Measures

The paper reports forecast-comovement statistics because they provide a bounded summary of the forecast-side implication of the theory. Let

$$f_{rx,t} = \widehat{\mathbb{E}}_t r x_{t+1}, \quad f_{\nu,t} = \widehat{\mathbb{E}}_t \nu_{t+1}^{(B)},$$

denote the fitted 1-step forecasts from the empirical state system. In Table 1.3, the reported sample forecast-comovement statistic is

$$\widehat{R}_{\text{forecast,sample}}^2 = \widehat{\text{Corr}}(f_{rx,t}, f_{\nu,t})^2.$$

This is the squared sample correlation between the two fitted forecast series. It is mechanically bounded between zero and one.

The population analogue uses the stationary distribution of the state vector implied by the estimated VAR. Let Σ_x denote the stationary covariance matrix of x_t , satisfying the discrete Lyapunov equation

$$\Sigma_x = \Gamma \Sigma_x \Gamma' + \Sigma_\varepsilon. \quad (\text{A.21})$$

Let $\gamma_{rx} = e_1' \Gamma$ and $\gamma_\nu = e_3' \Gamma$ be the forecast-slope rows for excess returns and the behavioral cash-flow news proxy. The population forecast-comovement statistic is

$$R_{\text{forecast,population}}^2 = \frac{(\gamma_{rx} \Sigma_x \gamma_\nu')^2}{(\gamma_{rx} \Sigma_x \gamma_{rx}') (\gamma_\nu \Sigma_x \gamma_\nu')}.$$

This is the Lyapunov-based statistic referred to in the abstract and introduction as the population forecast-comovement measure.

The distinction between the sample and population versions matters in finite samples. The sample statistic uses the realized fitted forecast series. The population statistic asks how tightly the two forecast functions comove under the stationary distribution implied by the estimated forecasting system. In the baseline application, the sample R^2 values are slightly below the population values, but both summarize the same economic object: predictable return variation and predictable belief-revision variation move together.

Forecast R^2 is also useful because it is invariant to positive scalar rescaling of one forecast. If $f_{\nu,t}^{\text{new}} = a f_{\nu,t}$ with $a > 0$, then $\text{Corr}(f_{rx,t}, f_{\nu,t}^{\text{new}})^2 = \text{Corr}(f_{rx,t}, f_{\nu,t})^2$. Thus scalar attenuation of the proxy affects levels and variance-scale comparisons more directly than it affects the squared correlation of fitted forecasts. Additive noise inside the state vector can still affect forecast R^2 by changing the estimated forecasting system, which is why the statistic is interpreted together with the other diagnostics.

When reporting bias-corrected forecast-comovement statistics, I use a bootstrap correction for finite-sample bias. Let $\widehat{R}^2$ be the estimate in the data. Let R_0^2 be the population value implied by the fitted data-generating system used for the bootstrap, and let $\overline{R}^{2,*}$ be the mean estimate across bootstrap samples generated from that system. The bias-corrected statistic is

$$\widehat{R}_{bc}^2 = \widehat{R}^2 - \left(\overline{R}^{2,*} - R_0^2 \right).$$

The correction is diagnostic rather than structural. It adjusts for finite-sample estimation bias under the fitted forecasting system; it does not resolve nonclassical measurement error in the belief proxy.

A.3.6 Measurement Error Inside the State Vector

The static formulas above are useful for intuition, but the empirical implementation is more involved because the belief proxy enters the state vector. Measurement error in $\nu_t^{(B)}$ can affect:

1. the estimated transition matrix Γ ;
2. the VAR innovations ε_{t+1} ;
3. the present-value matrix $\Lambda = \rho\Gamma(I - \rho\Gamma)^{-1}$;
4. the constructed rational news terms $N_{RP}^{(r)}$, $N_{RF}^{(r)}$, and $N_{CF}^{(r)}$;
5. the constructed belief-bias objects $\eta^{(B)}$ and $\zeta^{(B)}$;
6. the slope and horizon restrictions.

This feedback means that a simple scalar reliability correction is not a full correction for measurement error in the paper’s empirical system. For example, additive noise in $\nu_t^{(B)}$ may attenuate the estimated forecasting row $e_3'\Gamma$, alter the innovation covariance matrix, and change the constructed rational cash-flow news that is subtracted from $\nu^{(B)}$ to form $\eta^{(B)}$.

For this reason, the main text treats measurement-error corrections as supporting diagnostics. The core empirical claim is based on the joint pattern across several objects: negative news correlations, economically meaningful scale, positive forecast comovement, slope non-rejection at short and medium horizons, price-inference robustness, adversarial-residual bounds, and cross-sectional breadth. No single statistic is interpreted as a fully structural estimate of the latent belief wedge.

A.4 Data Construction and Proxy Validation

This appendix documents the data and proxy construction used in Section 1.3. Appendix A.4.1 describes return and macro data. Appendix A.4.2 describes the analyst forecast data and sample. Appendix A.4.3 gives the target-date alignment construction. Appendices A.4.4–A.4.5 give the proxy-specific construction details, the belief-level control, and the value-weighted market-level aggregation. Appendices A.4.6–A.4.9 report the proxy-validation diagnostics that support the “news-like” framing in Section 1.3.3: summary statistics with placebo comparisons, rationality and predictability diagnostics, persistence profiles, and the common-factor evidence.

A.4.1 Return and Macro Data

The benchmark sample runs from February 1982 through June 2023, giving 497 monthly observations after merging the return data with the analyst-forecast panel. One observation is lost in forming first-differenced innovations, leaving 496 usable news observations. The level variables span 1982:01–2023:06 (498 months) and the revision-based objects span 1982:02–2023:06 (497 transitions).

The return-side variables are:

- Excess log market return rx_t and 1-month risk-free rate rf_t from the Ken French data library (value-weighted CRSP market factor and Treasury-bill rate).
- Log price-dividend ratio pd_t from the Shiller monthly S&P 500 series, $pd_t = \log(P_t/D_t)$.
- Term spread ts_t from FRED, defined as the difference between the 10-year Treasury

yield (GS10) and the 3-month Treasury yield (TB3MS).

- Dividend and total-return information used to recover cash-flow news under the Campbell and Ammer (1993) identity (Appendix A.5.3).

All variables are at the monthly frequency. The log-linearization constant is $\rho_{\text{monthly}} = 0.998122$ throughout, calibrated from the unconditional mean of pd_t in the legacy sample; the corresponding annualized value is $\rho_{\text{annual}} \approx 0.978$.

A.4.2 Analyst Forecast Data and Sample

Analyst forecast data come from IBES. The firm-level panel includes fiscal-year-1 (FY1) and fiscal-year-2 (FY2) earnings-per-share forecasts, fiscal-period-end dates, the number of contributing analysts, and long-term-growth (LTG) forecasts. Trailing twelve-month earnings-per-share are obtained from CRSP/Compustat using the IBES–CRSP link table.

Standard sample filters are applied. Firm-months require: (i) at least three contributing analysts for FY1 (and, where relevant, for LTG); (ii) trailing twelve-month earnings-per-share strictly above \$0.10; (iii) FY1 forecast strictly positive. Within-month winsorization at the 1st and 99th percentiles handles outliers in the firm-level revision objects. Foreign firms and ADRs without valid CRSP-Compustat matches are excluded.

The benchmark aligned CBR and STR proxies each yield approximately 211,000 firm-month observations over the 1982:01–2023:06 sample, while the aligned LTG proxy yields approximately 203,000 because it additionally requires non-missing long-term-growth forecasts. Across all three proxies, the average common-support size is 9.92 of 10 target dates, so the alignment procedure discards very little.

A.4.3 Target-Date Alignment

Naive month-over-month changes in analyst forecasts are contaminated by three mechanical effects. First, horizon roll: a forecast labeled FY2 at time t may correspond to the same future fiscal-year-end date as a forecast labeled FY1 at time $t + 1$, so a same-label comparison mixes true revisions with the mechanical shortening of the forecast horizon. Second, fiscal-year relabeling: around a firm's fiscal-year rollover, the set of forecast labels itself changes, and a simple change in the FY1 forecast around those months combines true updating with relabeling noise. Third, the passage of time: even if analysts receive no new information, a month-to-month change in a present-value-weighted level can occur mechanically because the same set of future cash flows is now one month closer.

The aligned construction strips these effects. For each firm i and target date T , let $F_{i,t}(T)$ denote the analyst forecast at time t for the cash-flow object dated T . The aligned revision is the change between adjacent months over a fixed target date:

$$\nu_{i,t+1}^{(B)}(T) = F_{i,t+1}(T) - F_{i,t}(T). \quad (\text{A.22})$$

This matches equation (1.15) in the body. The target date T is held fixed across the revision window: the revision measures a change in beliefs about the same future cash-flow object, not a change in the object being forecast. For each firm, alignment is computed on the common-support set of target dates available in both t and $t + 1$. The first-horizon target-date rollover rate is approximately 8%, large enough to make alignment necessary but small enough that the common-support restriction discards little data.

A.4.4 Behavioral Cash-Flow News Proxy Construction

The three proxies differ in the analyst-forecast inputs used to build the firm-level expectation path. Each proxy first constructs a firm-level present-value-weighted expectation path for future cash-flow growth over annual horizons $h = 1, \dots, 10$ at annual discount rate $\rho_{\text{ann}} = 0.9777$, with persistence parameter $\phi = 0.445$ governing decay toward a cross-sectional long-run anchor. Aligned revisions are then formed using equation (A.22).

CBR (Composite Belief Revision). The composite proxy combines short-horizon analyst earnings forecasts and long-term-growth forecasts. The first two annual horizons are pinned down directly by FY1 and FY2 forecasts, scaled by trailing twelve-month earnings. Horizons $h \geq 3$ blend the firm-specific LTG forecast with a cross-sectional long-run anchor, with the persistence parameter ϕ controlling the weight on firm-specific information. This is the flagship proxy because it uses the broadest information set and preserves a transparent separation between short-horizon analyst information and the longer-horizon continuation path.

LTG (Long-Term Growth). The LTG proxy uses only the long-term-growth forecast. The expectation path applies the LTG growth rate to all horizons $h = 1, \dots, 10$, blended with the cross-sectional long-run anchor according to ϕ . This isolates the long-run-growth component of the composite proxy.

STR (Short-Term Revision). The STR proxy follows the De La O and Myers (2021) construction of cash-flow expectations from analyst earnings forecasts. Only FY1 and FY2 are used, with the distant-horizon path determined by geometric decay from the second-year forecast-implied growth rate. Unlike CBR, STR does not incorporate analyst long-term-growth forecasts and does not impose convergence to a cross-sectional long-run target.

A.4.5 Market-Level Aggregation and Belief-Level Control

For each proxy, the market-level revision aggregates firm-level revisions using lagged value weights:

$$\nu_{t+1}^{(B)} = \sum_{i \in \mathcal{I}_t} w_{i,t} \nu_{i,t+1}^{(B)},$$

where $\mathcal{I}_t$ is the set of firms with usable aligned forecasts at time t , and $w_{i,t}$ is the time- t market-value weight of firm i , normalized within $\mathcal{I}_t$ to sum to one. Weights are lagged by one month relative to the revision dates to limit same-month price feedback into the aggregation. This matches equation (1.16) in the body.

The belief-level control $b_t^{(B)}$ in the baseline state vector (1.12) is the corresponding firm-level expectation path level (rather than its first difference) aggregated using the same value weights. In the baseline specifications, $b_t^{(B)}$ is the CBR-based belief level and is held fixed across the CBR, LTG, and STR revision-proxy systems. Holding the level control fixed makes the systems easier to compare: the revision proxy varies across specifications, while the state variable controlling for the current level of cash-flow beliefs is held constant. The belief level is highly persistent (first-order autocorrelation approximately 0.92) and behaves like a level state variable, complementing the revision proxy in the forecasting system.

A.4.6 Proxy Validation: Summary Statistics and Placebos

Table A.1 reports summary statistics for the three aligned revision proxies and a simple horizon-label placebo. The aligned revisions have near-zero means, moderate volatility, and low first-order autocorrelations. The simple next-fiscal-year revision (which ignores fiscal-year rollover) is roughly twice as volatile, substantially more persistent, and nearly uncorrelated with the

aligned CBR proxy. These properties support treating the aligned revisions as news-like belief-revision measures rather than horizon-roll artifacts.

Table A.1: Summary Statistics for Aligned Revisions and Placebo

Series	Mean	SD	AC(1)	Correlation with CBR
Aligned CBR Revision	−0.00	0.01	0.15	1.00
Aligned LTG Revision	−0.00	0.01	0.17	0.62
Aligned STR Revision	+0.00	0.01	0.06	0.90
Simple Next-Fiscal-Year Revision	−0.01	0.02	0.51	−0.10
CBR Belief Level	0.14	0.02	0.92	0.15

Notes: The table reports the sample mean, standard deviation, first-order autocorrelation, and correlation with the aligned CBR revision for each series. All aligned revisions use month- t value weights, a 10-horizon expectation path with $\phi = 0.45$ and $\rho_{\text{ann}} = 0.98$, common-support target-date alignment, within-month winsorization at the 1st and 99th percentiles, and a minimum of three contributing analysts. Firm-months are additionally restricted to $\text{EPS}_{i,t}^{\text{trail}} > 0.10$ and $\text{FY1}_{i,t} > 0$. The belief level is the CBR-based level series used as the fourth state variable in equation (1.12).

The aligned proxies are also positively correlated with one another: pairwise correlations are 0.902 (CBR vs. STR), 0.620 (CBR vs. LTG), and 0.271 (LTG vs. STR). The CBR and STR series move closely together because both rely heavily on short-horizon earnings forecasts; the LTG-only proxy is more distinct because it isolates a narrower long-run-growth component.

A.4.7 Rationality and Predictability Diagnostics

The aligned revisions display predictable dynamics that are difficult to reconcile with a full-information rational benchmark. Autocorrelation regressions show first-order coefficients that are statistically positive for CBR and LTG, consistent with sticky updating. Joint predictability tests find that the state vector x_t from equation (1.12) predicts next-period revisions

with sample R^2 values in the range of 5–10% depending on the proxy.

These diagnostics are not used as the main identification argument. They show that the aligned series behave like subjective belief updates rather than full-information rational news: they are not mechanically persistent horizon-label changes (the placebo in Table A.1 is much more persistent), but they display predictable dynamics consistent with sticky or biased updating. The empirical decomposition in Section 1.3.4 uses the rationally unexpected component $\zeta_{t+1}^{(B)}$ rather than the raw revision, so predictable raw revisions do not contaminate the news-side tests.

A.4.8 Persistence Profiles of Aligned Revisions

Local-projection persistence profiles characterize the dynamic behavior of the aligned revisions. For each proxy, the cumulative-revision impulse response over horizons $h = 1, \dots, 24$ months is positive at short horizons and decays toward zero at longer horizons. CBR and STR show comparable short-run persistence with similar decay profiles. LTG displays mild short-run mean reversion before decaying, consistent with the more distinct information content of the long-term-growth signal documented in the cross-proxy correlations.

The persistence profiles are diagnostic rather than identifying. They support the body’s interpretation of the aligned revisions as belief-revision news with predictable dynamics, not as either pure white-noise innovations or as level-like state variables.

A.4.9 Common Component Across Proxies

The three aligned proxies share a substantial common component. Let

$$X_t = (\tilde{\nu}_{\text{CBR},t}, \tilde{\nu}_{\text{LTG},t}, \tilde{\nu}_{\text{STR},t})$$

be the standardized aggregate aligned revision series, and let f_t denote the first principal component of X_t , oriented to load positively on the flagship CBR proxy. The first principal component explains 74.4% of the standardized variance across the three proxies. Loadings on the standardized series are approximately 0.66 on CBR, 0.46 on LTG, and 0.59 on STR.

The principal-component summary is almost perfectly correlated with the aligned CBR proxy (0.991), strongly correlated with STR (0.878), and weakly but positively correlated with LTG (0.692). This pattern is consistent with the pairwise correlation matrix in Appendix A.4.6: CBR and STR move closely together because both rely on short-horizon earnings information, while LTG contains a related but more distinct source of variation.

The principal-component summary is retained as a commonality diagnostic, not as the headline behavioral signal. Because the summary is so highly correlated with CBR, it adds little new information beyond the flagship proxy while giving greater weight to the more distinct LTG component. The body therefore uses CBR as the flagship behavioral measure.

A.5 VAR, News Construction, and Inference

This appendix gives the econometric implementation details behind Section 1.3.5 and Section 1.3.6. Appendix A.5.1 states the benchmark VAR specification. Appendix A.5.2 describes estimation and stability. Appendices A.5.3–A.5.5 give the news-construction algebra and the empirical belief-bias news objects. Appendix A.5.6 reports the Campbell and Ammer (1993) identity check. Appendix A.5.7 states the slope-restriction Wald statistic. Appendices A.5.8–A.5.9 describe the stationary block bootstrap and the null-imposed bootstrap used for the slope test. Appendix A.5.10 describes the direct horizon tests. Appendices A.5.11–A.5.12 report finite-sample null diagnostics and heteroskedasticity- and autocorrelation-consistent (HAC) bandwidth sensitivity.

A.5.1 Benchmark VAR Specification

The benchmark state vector is

$$x_t = [rx_t, rf_t, \nu_t^{(B)}, b_t^{(B)}, pd_t, ts_t]',$$

matching equation (1.12). The dynamics follow a monthly VAR(1):

$$x_{t+1} = \alpha + \Gamma x_t + \varepsilon_{t+1}.$$

The sample is February 1982 through June 2023, providing 497 monthly VAR observations and 496 usable news-term observations after one observation is lost in forming first-differenced innovations. The log-linearization constant is $\rho_{\text{monthly}} = 0.998122$ (annualized $\rho_{\text{annual}} \approx 0.978$),

calibrated from the unconditional mean of pd_t in the legacy sample. The CBR, LTG, and STR systems are estimated separately, each replacing the third state slot ($\nu_t^{(B)}$) with the corresponding behavioral cash-flow news proxy. The fourth state variable $b_t^{(B)}$ is the CBR-based belief level and is held fixed across all three systems.

A.5.2 VAR Estimation and Stability

The VAR is estimated equation by equation by ordinary least squares, with intercepts left unrestricted. The innovation covariance matrix is $\Sigma_\varepsilon = T^{-1} \sum_t \hat{\varepsilon}_{t+1} \hat{\varepsilon}_{t+1}'$, where $\hat{\varepsilon}_{t+1}$ are the OLS residuals.

The system requires that the present-value matrix $\Lambda = \rho\Gamma(I - \rho\Gamma)^{-1}$ exists, which holds when the spectral radius of $\rho\hat{\Gamma}$ is strictly less than one. In the legacy six-variable VAR, the dominant eigenvalue of $\rho\hat{\Gamma}$ is approximately 0.977, well within the stability region. Bootstrap draws are checked for stability, and the rare unstable draws ($< 0.5\%$ of the 5,000 replications) are discarded and redrawn.

A.5.3 News Construction

Define the geometric-sum operator

$$M = (I - \rho\Gamma)^{-1}, \quad \Lambda = \rho\Gamma M.$$

Risk premium news and risk-free-rate news are

$$N_{RP,t+1}^{(r)} = e_1' \Lambda \varepsilon_{t+1}, \quad N_{RF,t+1}^{(r)} = e_2' \Lambda \varepsilon_{t+1}.$$

Cash-flow news is recovered from the Campbell and Ammer (1993) identity:

$$N_{CF,t+1}^{(r)} = e'_1 \varepsilon_{t+1} + N_{RF,t+1}^{(r)} + N_{RP,t+1}^{(r)} = (e'_1 + e'_2 \Lambda + e'_1 \Lambda) \varepsilon_{t+1}. \quad (\text{A.23})$$

This ordering matches the body's narrative: $N_{RP}^{(r)}$ and $N_{RF}^{(r)}$ are constructed directly from the forecasting system's revisions to future expected returns and rates, and $N_{CF}^{(r)}$ is recovered last so that the accounting identity holds exactly. The fitted forecast series used for the sample and population forecast- R^2 statistics in Appendix A.3.5 are $\widehat{\mathbb{E}}_t r x_{t+1} = e'_1 (\widehat{\alpha} + \widehat{\Gamma} x_t)$ and $\widehat{\mathbb{E}}_t \nu_{t+1}^{(B)} = e'_3 (\widehat{\alpha} + \widehat{\Gamma} x_t)$.

A.5.4 Risk-Free-Rate News Timing Convention

The implementation above treats risk-free-rate news symmetrically with risk premium news: both are present-value revisions in future rates discounted at ρ^j , indexed at $t + 1 + j$ for $j \geq 1$. This is the Vuolteenaho-symmetric convention. An alternative textbook convention in some references indexes the rate stream at $t + j$ rather than $t + 1 + j$, which yields a slightly different intercept allocation between $N_{RF}^{(r)}$ and $N_{CF}^{(r)}$. Both conventions yield internally consistent return-news decompositions; the choice does not affect the sign or magnitude of the empirical relation between $N_{RP}^{(r)}$ and the belief-bias proxy. The implementation uses the symmetric convention for transparency.

A.5.5 Constructing Empirical Belief-Bias News

For each behavioral cash-flow news proxy $j \in \{\text{CBR}, \text{LTG}, \text{STR}\}$, the empirical belief-bias wedge is

$$\eta_{j,t+1}^{(B)} = \nu_{j,t+1}^{(B)} - N_{CF,t+1}^{(r)},$$

matching equation (1.17). The rationally unexpected component is

$$\zeta_{j,t+1}^{(B)} = \eta_{j,t+1}^{(B)} - \widehat{\mathbb{E}}_t^{(r)} \eta_{j,t+1}^{(B)},$$

where $\widehat{\mathbb{E}}_t^{(r)} \eta_{j,t+1}^{(B)}$ is the linear projection of $\eta_{j,t+1}^{(B)}$ on the time- t state vector x_t . Because $\nu_{t+1}^{(B)}$ is the third state-vector slot and $N_{CF,t+1}^{(r)}$ is constructed in equation (A.23), the linear-system form of $\zeta^{(B)}$ in the VAR innovation representation is

$$\zeta_{t+1}^{(B)} = (e'_3 - e'_1 - e'_2\Lambda - e'_1\Lambda) \varepsilon_{t+1}.$$

This connects directly to the matrix-wedge expression in equation (A.18): the matrix wedge is $W_{t+1}^{\text{matrix}} = N_{RP,t+1}^{(r)} + \zeta_{t+1}^{(B)} = (e'_3 - e'_1 - e'_2\Lambda) \varepsilon_{t+1}$.

A.5.6 Campbell–Ammer Identity Checks

The implementation verifies that the news-construction algebra reproduces the Campbell and Ammer (1993) identity to machine precision. Define the identity residual

$$R_{t+1}^{\text{CA}} \equiv (rx_{t+1} - \widehat{\mathbb{E}}_t^{(r)} rx_{t+1}) - N_{CF,t+1}^{(r)} + N_{RF,t+1}^{(r)} + N_{RP,t+1}^{(r)}.$$

Across the legacy 496-observation news series, $\max_t |R_{t+1}^{\text{CA}}| \approx 8.3 \times 10^{-17}$, confirming numerical recovery of the identity to floating-point precision.

A.5.7 Wald Statistic for the One-Step Slope Restriction

The slope restriction is $H_0 : e'_1 \Gamma = e'_3 \Gamma$, equivalently $R\theta = 0$ where $\theta = \text{vec}(\Gamma)$ and R is the corresponding restriction matrix that picks the difference between the rx -row and the $\nu^{(B)}$ -row. The Wald statistic is

$$W = (R\hat{\theta})' (R \widehat{\text{Var}}(\hat{\theta}) R')^{-1} (R\hat{\theta}),$$

with the asymptotic variance computed from a Newey–West HAC-adjusted stacked-system covariance estimator using bandwidth $L = 6$. There are six restrictions in the benchmark six-variable VAR. Under the null and asymptotic regularity conditions, W is approximately $\chi^2(6)$. The body’s primary inference is the null-imposed bootstrap of Appendix A.5.9; the asymptotic statistic is reported as a diagnostic.

A.5.8 Stationary Block Bootstrap

The stationary block bootstrap of Politis and Romano (1994) generates time-series resamples that preserve serial dependence. The implementation uses $B = 5,000$ replications and a mean block length of 12 months. For each replication, the algorithm:

1. resamples residual blocks (or state-vector blocks, depending on the moment) from the unrestricted VAR;

2. simulates the time series under the unrestricted estimated dynamics;
3. re-estimates the VAR;
4. recomputes Λ , $N_{RP}^{(r)}$, $N_{RF}^{(r)}$, $N_{CF}^{(r)}$, $\eta_j^{(B)}$, $\zeta_j^{(B)}$, the matrix wedge, and all reported test statistics from the bootstrap-replicated series;
5. checks stability and discards (or redraws) unstable replicates.

The principle is that all generated news objects are reconstructed within each replication, so the bootstrap distribution captures the joint sampling uncertainty of the estimated VAR and the constructed news terms. Confidence intervals are reported as percentile intervals across the B replications.

A.5.9 Null-Imposed Bootstrap for the Slope Test

The slope-restriction Wald statistic is constructed on a small number of restrictions in a persistent forecasting system. The asymptotic $\chi^2(6)$ approximation tends to over-reject in finite samples (Appendix A.5.11). The body therefore uses a null-imposed bootstrap as the primary inference procedure. The algorithm:

1. estimates the restricted VAR under $e_1' \Gamma = e_3' \Gamma$ by stacked GLS imposing the equality across the rx and $\nu^{(B)}$ equations;
2. collects restricted residuals;
3. resamples residual blocks via the stationary block bootstrap with $B = 5,000$ and mean block length 12;

4. simulates bootstrap series under the restricted dynamics;
5. re-estimates the unrestricted VAR on each replicate;
6. computes the Wald statistic W^* on each replicate;
7. reports $p_{\text{boot}} = B^{-1} \sum_b \mathbb{1}\{W_b^* \geq W_{\text{obs}}\}$.

Unstable replicates are discarded.

A.5.10 Direct Horizon Tests and Local Projections

The direct horizon tests in Section 1.4.2 use local projections (Jordà 2005) rather than VAR-iterated cumulative restrictions. For each horizon $h \in \{1, 12, 24, 36, 60\}$ months, the cumulative excess return and cumulative belief revision are projected on the time- t state vector:

$$\sum_{j=1}^h \rho^{j-1} r x_{t+j} = a_{rx,h} + b'_{rx,h} x_t + u_{rx,t,h},$$

$$\sum_{j=1}^h \rho^{j-1} \nu_{t+j}^{(B)} = a_{\nu,h} + b'_{\nu,h} x_t + u_{\nu,t,h}.$$

The horizon- h slope restriction tests $H_0 : b_{rx,h} = b_{\nu,h}$ via a stacked-system Wald statistic with HAC standard errors at bandwidth $\max(h, 6)$. Inference uses a null-imposed stationary block bootstrap parallel to the 1-step procedure in Appendix A.5.9, with the restricted VAR re-estimated under the null and bootstrap series simulated from the restricted dynamics.

The $h = 1$ case validates the procedure: by construction, the LP estimate at $h = 1$ should reproduce the 1-step VAR slope test up to bootstrap sampling noise. Empirically, this validation holds in the data (Table 1.4). Sample size declines at longer horizons because of the loss of h

overlapping observations; effective sample sizes are reported alongside each estimate. The direct horizon tests do not impose the VAR's iterated implication that 1-step slope equality forces horizon- h slope equality (Appendix A.2.6). They therefore identify cases in which the VAR-iterated implication fits 1-step well but fails at long horizons, as in the 5-year LP rejection documented in the body.

A.5.11 Finite-Sample Null Diagnostics

The asymptotic $\chi^2(6)$ approximation is too aggressive for the slope-restriction Wald statistic in this setting. Table A.2 reports the median, mean, and 95th percentile of the null-imposed bootstrap distribution alongside the $\chi^2(6)$ reference. The bootstrap null is centered well above $\chi^2(6)$: the null median ranges from 11.57 to 12.27 across proxies vs. a $\chi^2(6)$ median of 5.35, and the null 95th percentile ranges from 28.26 to 31.47 vs. a $\chi^2(6)$ critical value of 12.59. An observed Wald statistic that appears large relative to $\chi^2(6)$ may be entirely unremarkable relative to the true finite-sample null. This is the core finite-sample motivation for the bootstrap procedure.

Table A.2: Null-Imposed Bootstrap Diagnostics for the Slope-Restriction Wald Statistic

Reference Distribution	Median	Mean	95 th Percentile
$\chi^2(6)$ Asymptotic Reference	5.35	6.00	12.59
Null-Imposed Bootstrap, CBR	11.57	13.26	29.07
Null-Imposed Bootstrap, LTG	12.27	14.16	31.47
Null-Imposed Bootstrap, STR	11.86	13.36	28.26

Notes: The table reports summary statistics from the null-imposed bootstrap distribution of the slope-restriction Wald statistic for each proxy ($B = 5,000$, mean block length 12 months). The null distribution is constructed by simulating bootstrap data under the restricted VAR with $e_1' \Gamma = e_3' \Gamma$ imposed, and re-estimating the unrestricted VAR on each replicate. The $\chi^2(6)$ row reports the corresponding asymptotic reference values.

A.5.12 HAC Bandwidth and Inference Sensitivity

The Newey–West HAC bandwidth used to compute the asymptotic variance of $\hat{\theta}$ is $L = 6$ in the baseline implementation. Alternative bandwidths $L \in \{4, 8, 12\}$ produce asymptotic Wald statistics within $\pm 10\%$ of the baseline; the qualitative conclusions on slope-restriction non-rejection are unchanged.

The body’s primary inference procedure is the null-imposed bootstrap of Appendix A.5.9, not the HAC-asymptotic Wald. The HAC sensitivity is reported as a diagnostic. The bootstrap procedure preserves serial dependence in the data without relying on a kernel approximation to the long-run covariance, and it imposes the null through the data-generating process rather than through the reference distribution. The null-imposed bootstrap is therefore robust to both bandwidth choice and small-sample skewness in the Wald distribution.

A.6 Additional Empirical Tables

This appendix reports empirical results that support Sections 1.4 and 1.5. The main text reports compact tables for the baseline time-series evidence, direct horizon tests, matrix wedges, and horizon diagnostics. The tables below provide additional detail: an extended baseline by proxy, bootstrap confidence intervals, the standard Campbell and Ammer (1993) variance decomposition, alternative forecast-comovement measures, full horizon profiles, matrix-wedge bootstrap summaries, ρ -sensitivity, state-vector and lag-order robustness, HAC bandwidth diagnostics, and reserved sample/frequency robustness rows.

A.6.1 Baseline Results by Proxy

Table A.3 expands the baseline time-series evidence in Table 1.3. The main text reports the directional news correlation, behavioral scale ratio, sample forecast correlation, sample forecast R^2 , and the null-imposed bootstrap p -value for the 1-step slope restriction. The extended table adds the population forecast-comovement statistic and the slope-restriction Wald statistic.

Table A.4 consolidates bootstrap uncertainty for the main baseline objects in one place. Matrix-wedge intervals are reproduced from Table 1.5; forecast R^2 intervals are reported in Table A.6. Bootstrap intervals for the news correlation and $\mathcal{B}_{\text{scale}}$ are reserved for the corresponding stationary-block bootstrap output.

Table A.3: Extended Baseline Time-Series Evidence

Proxy	News Correlation	$\mathcal{B}_{\text{scale}}$	Forecast Correlation	Sample R^2	Population R^2	W_{obs}	p_{boot}
CBR	−0.11	0.34	0.51	0.26	0.34	8.92	0.66
LTG	−0.11	0.28	0.51	0.26	0.29	15.76	0.35
STR	−0.10	0.34	0.46	0.21	0.30	9.03	0.67

Notes: Sample 1982:02–2023:06, $T = 496$ usable news observations. “News Correlation” is the correlation between $N_{RP}^{(r)}$ and $\zeta^{(B)}$. $\mathcal{B}_{\text{scale}}$ is the variance ratio in equation (1.10). “Forecast Correlation” is the signed sample correlation between $\widehat{\mathbb{E}}_t r x_{t+1}$ and $\widehat{\mathbb{E}}_t \nu_{t+1}^{(B)}$. “Sample R^2 ” is the squared sample forecast correlation. “Population R^2 ” is the value from the discrete Lyapunov equation (A.21). W_{obs} is the slope-restriction Wald statistic in Appendix A.5.7. p_{boot} is the null-imposed bootstrap p -value (Appendix A.5.9). Asymptotic $\chi^2(6)$ p -values are 0.18, 0.01, and 0.17 for CBR, LTG, and STR.

Table A.4: Bootstrap Confidence Intervals for Baseline Statistics

Proxy	News Correlation	$\mathcal{B}_{\text{scale}}$	Forecast Correlation	Sample R^2	Population R^2	W_{obs}	p_{boot}
CBR	−0.11	0.34	0.51	0.26	0.34	8.92	0.66
LTG	−0.11	0.28	0.51	0.26	0.29	15.76	0.35
STR	−0.10	0.34	0.46	0.21	0.30	9.03	0.67

Notes: Sample 1982:02–2023:06, $T = 496$ usable news observations. “News Correlation” is the correlation between $N_{RP}^{(r)}$ and $\zeta^{(B)}$. $\mathcal{B}_{\text{scale}}$ is the variance ratio in equation (1.10). “Forecast Correlation” is the signed sample correlation between $\widehat{\mathbb{E}}_t r x_{t+1}$ and $\widehat{\mathbb{E}}_t \nu_{t+1}^{(B)}$. “Sample R^2 ” is the squared sample forecast correlation. “Population R^2 ” is the value from the discrete Lyapunov equation (A.21). W_{obs} is the slope-restriction Wald statistic in Appendix A.5.7. p_{boot} is the null-imposed bootstrap p -value (Appendix A.5.9). Asymptotic $\chi^2(6)$ p -values are 0.18, 0.01, and 0.17 for CBR, LTG, and STR.

A.6.2 Campbell–Ammer Variance Decomposition

Table A.5 reports the standard Campbell and Ammer (1993) variance decomposition of unexpected excess returns for the legacy six-variable VAR. Risk premium news is the dominant component of the unexpected-return variance, accounting for roughly 89% of $\text{Var}(r x_{t+1} - \mathbb{E}_t^{(r)} r x_{t+1})$ at the standard variance-share definition. This is the conventional decomposition.

The paper’s contribution is interpretive: it asks what fraction of this risk premium news object reflects biased cash-flow beliefs.

Table A.5: Campbell–Ammer Variance Decomposition

Component	Variance Share
$\text{Var}(N_{CF}^{(r)})/\text{Var}(u_r)$	0.22
$\text{Var}(N_{RF}^{(r)})/\text{Var}(u_r)$	0.13
$\text{Var}(N_{RP}^{(r)})/\text{Var}(u_r)$	0.89
$-2 \text{Cov}(N_{CF}^{(r)}, N_{RF}^{(r)})/\text{Var}(u_r)$	-0.04
$-2 \text{Cov}(N_{CF}^{(r)}, N_{RP}^{(r)})/\text{Var}(u_r)$	-0.08
$+2 \text{Cov}(N_{RF}^{(r)}, N_{RP}^{(r)})/\text{Var}(u_r)$	-0.12
Identity Residual $\max_t R_{t+1}^{\text{CA}} $	5.6×10^{-17}

Notes: Variance shares of unexpected excess returns from the legacy six-variable CBR VAR. $u_r = rx_{t+1} - \mathbb{E}_t^{(r)} rx_{t+1}$. The variance shares of the news terms and the cross-product terms sum to one by the Campbell and Ammer (1993) identity. The identity residual reports the maximum absolute deviation from the identity in the data, confirming numerical recovery to floating-point precision (Appendix A.5.6).

A.6.3 Forecast-Comovement Statistics

Table A.6 collects the alternative forecast-comovement statistics used in the paper. The main text reports the signed sample forecast correlation and the corresponding squared sample correlation. The introduction and abstract also refer to the population forecast-comovement measure, which uses the stationary distribution implied by the estimated VAR. The table additionally reports three bias-corrected bootstrap intervals. The sample values match Table 1.3; the population values around 0.30 correspond to the figures cited in the abstract and introduction.

Table A.6: Forecast Comovement: Sample, Population, and Bias-Corrected R^2

Proxy	Sample R^2	Population R^2	Pivot/Pop 95% CI	Mean-BC 95% CI	Median-BC 95% CI
CBR	0.26	0.34	[0.08, 0.61]	[0.19, 0.71]	[0.23, 0.76]
LTG	0.26	0.29	[0.03, 0.56]	[0.14, 0.67]	[0.19, 0.71]
STR	0.21	0.30	[0.00, 0.51]	[0.15, 0.70]	[0.20, 0.76]

Notes: Sample R^2 is the squared sample correlation between fitted return forecasts and fitted belief-revision forecasts. Population R^2 is the value from the discrete Lyapunov equation (A.21). The three confidence intervals correspond to (i) a pivot-style CI centered on the population value, (ii) a mean-bias-corrected CI, and (iii) a median-bias-corrected CI, all from a stationary block bootstrap with $B = 5,000$ replications and mean block length 12 months (Appendix A.5.8).

A.6.4 Full Horizon Slope Tests

Table A.7 reports direct horizon slope tests for $h \in \{1, 3, 6, 12, 24, 36, 60\}$ months, extending Table 1.4 in the body. The pattern is consistent across proxies: non-rejection at $h \leq 36$, sharp rejection at $h = 60$.

Table A.7: Direct Horizon Slope Tests, All Reported Horizons

Proxy	$h = 1$	$h = 3$	$h = 6$	$h = 12$	$h = 24$	$h = 36$	$h = 60$
CBR	0.66	0.66	0.61	0.67	0.61	0.35	0.00
LTG	0.35	0.24	0.22	0.43	0.29	0.36	0.00
STR	0.67	0.72	0.58	0.61	0.59	0.29	0.00

Notes: Bootstrap p -values for direct horizon slope tests (Appendix A.5.10). Each cell tests $H_0 : b_{rx,h} = b_{\nu,h}$ where $b_{rx,h}$ and $b_{\nu,h}$ are the slope vectors from the cumulative-return and cumulative-revision local-projection regressions on the time- t state vector. Bootstrap inference uses a null-imposed stationary block bootstrap with $B = 5,000$ and mean block length 12 months. The $h = 1$ column matches the 1-step VAR slope test up to bootstrap sampling noise.

A more detailed version of the same output reports the test statistic, asymptotic p -value, HAC standard error, and effective sample size in the standard local-projection form. Table A.8

reserves the layout. Bootstrap p -values are the preferred inference; the additional columns are reported as diagnostics only.

Table A.8: Direct Horizon Slope Tests: Detailed Output

Proxy	Horizon	Wald Statistic	p_{boot}	p_{asym}	Observations
CBR	1	8.92	0.66	0.18	496
CBR	12	12.58	0.67	0.05	485
CBR	24	15.88	0.61	0.01	473
CBR	36	31.08	0.35	0.00	461
CBR	60	606.05	0.00	0.00	437
LTG	1	15.76	0.35	0.01	496
LTG	12	21.88	0.43	0.00	485
LTG	24	32.04	0.29	0.00	473
LTG	36	33.31	0.36	0.00	461
LTG	60	624.71	0.00	0.00	437
STR	1	9.03	0.67	0.17	496
STR	12	14.27	0.61	0.03	485
STR	24	16.28	0.59	0.01	473
STR	36	35.55	0.29	0.00	461
STR	60	617.95	0.00	0.00	437

Notes: Detailed direct-horizon test output. The Wald test stacks the cumulative-return and cumulative-revision LP regressions and tests the six-restriction equality $b_h = c_h$ via Newey–West HAC with bandwidth $\max(h, 6)$, reporting p_{boot} from a null-imposed bootstrap with $B = 5,000$ and mean block length 12 months. Bootstrap p -values match Table A.7; minor differences relative to the body’s $h \in \{12, 24, 36\}$ entries reflect bootstrap sampling noise. Asymptotic p -values reported as 0.00 are below floating-point resolution.

A.6.5 Matrix-Wedge Bootstrap

Table A.9 expands Table 1.5. It reports the unrestricted and slope-restricted matrix-wedge variance ratios, the corresponding bootstrap confidence intervals, and the innovation-level correlation $\text{Corr}(N_{RP}^{(r)}, -\zeta^{(B)})$ in both the unrestricted and slope-restricted systems. The correlation columns are useful for interpreting why imposing the 1-step slope restriction does not mechanically repair pointwise long-horizon news equality (Section 1.5.2, Appendix A.2.7). The unrestricted correlations match the news correlations in Table 1.3 up to sign convention; the

slope-restricted correlations are negative for all three proxies.

Table A.9: Matrix-Wedge Bootstrap Results

Proxy	Unrestricted Ratio	Slope-Restricted Ratio	Unrestricted Correlation	Slope-Restricted Correlation
CBR	1.22 [0.77, 2.67]	2.98 [1.57, 6.72]	+0.11	−0.38
LTG	1.17 [0.72, 2.51]	2.39 [0.81, 6.98]	+0.11	−0.43
STR	1.22 [0.77, 2.66]	3.19 [1.65, 7.07]	+0.10	−0.41

Notes: The unrestricted ratio is the matrix-wedge variance ratio defined in equation (1.8), computed in the unrestricted six-variable VAR. The slope-restricted ratio recomputes the wedge after imposing the 1-step slope restriction in equation (1.6). The unrestricted correlation reports the innovation-level correlation between rational risk premium news and negative empirical belief-bias news in the unrestricted system; the slope-restricted correlation reports the same object after imposing the slope restriction. Brackets below each ratio report the 95% stationary-block bootstrap percentile confidence interval.

The confidence intervals for the unrestricted and slope-restricted ratios overlap, so the difference between the two ratios is not interpreted as a precisely estimated incremental effect. The robust implication is instead that 1-step slope equality and pointwise long-horizon news equality are distinct restrictions.

A.6.6 Horizon Decomposition

Table A.10 reports the numerical values behind Figure 1.1: the matrix-wedge variance ratio W_h , the innovation-level correlation $\rho_h = \text{Corr}(N_{RP,h}^{(r)}, -\zeta_h^{(B)})$, and the forecast-comovement R_h^2 at horizons $h \in \{1, 3, 6, 12, 24, 36, 60\}$. The pattern is the convergence story discussed in Section 1.5.3: W_h is large at short horizons (driven by small $\text{Var}(N_{RP,h})$ in the denominator) and decays toward the infinite-horizon ratio of 1.22 for CBR (Table 1.5) as h grows.

Table A.10: Horizon Diagnostics: Numerical Values

Proxy	Horizon h	W_h	ρ_h	R_h^2
CBR	1	341.93	+0.09	0.34
CBR	3	90.49	+0.35	0.31
CBR	6	23.19	+0.33	0.33
CBR	12	6.90	+0.19	0.33
CBR	24	2.87	+0.02	0.28
CBR	36	2.07	−0.05	0.21
CBR	60	1.62	−0.06	0.10
LTG	1	141.14	+0.07	0.29
LTG	3	63.61	+0.23	0.20
LTG	6	21.84	+0.34	0.19
LTG	12	6.72	+0.24	0.15
LTG	24	2.77	+0.06	0.06
LTG	36	1.99	−0.01	0.01
LTG	60	1.56	−0.04	0.02
STR	1	674.63	−0.03	0.30
STR	3	85.91	+0.31	0.31
STR	6	22.03	+0.31	0.33
STR	12	6.88	+0.20	0.36
STR	24	2.89	+0.03	0.35
STR	36	2.07	−0.04	0.30
STR	60	1.62	−0.06	0.22

Notes: $W_h = \text{Var}(\text{wedge}_h)/\text{Var}(N_{RP,h}^{(r)})$ uses Λ_h in place of the infinite-horizon Λ . ρ_h is the innovation-level correlation between cumulative risk premium news and negative cumulative belief-bias news. R_h^2 is the population forecast-comovement statistic at horizon h . Note the very large W_1 for STR (674.63): at $h = 1$ the denominator $\text{Var}(N_{RP,1})$ is small, inflating the ratio; W_h converges to the infinite-horizon ratio in Table 1.5 as h grows.

Table A.11 reports convergence diagnostics for the finite-horizon present-value matrices. The dominant-mode tail factor uses the leading persistence mode of the baseline VAR, approximately 0.977^{h+1} .

Table A.11: Finite-Horizon Tail Diagnostics

Horizon h	Dominant-Mode Tail Factor	Tail Norm Ratio	Notes
12	0.74	0.95	One Year
24	0.56	0.85	Two Years
36	0.42	0.74	Three Years
60	0.24	0.52	Five Years
120	0.06	0.18	Ten Years
240	0.00	0.02	Twenty Years

Notes: The dominant-mode tail factor uses the scalar persistence approximation 0.977^{h+1} ; it is a scale calculation for the persistence of the present-value tail, not an exact matrix-norm identity. The tail norm ratio reports $\|\Lambda - \Lambda_h\|_F / \|\Lambda\|_F$ for the legacy six-variable Bretscher VAR. LTG and STR yield similar values (within ± 0.01 at every horizon).

A.6.7 ρ -Sensitivity

Table A.12 reports the sensitivity of the 5-year horizon diagnostics to the log-linearization constant ρ . The baseline value is $\rho_{\text{monthly}} = 0.998122$. Alternative values of ± 0.001 around the baseline produce small shifts in W_h , ρ_h , and R_h^2 , leaving the qualitative conclusions unchanged.

A.6.8 State Vector and Lag Order Robustness

Table A.13 reports the slope-restriction Wald statistic and bootstrap p -value across alternative state vectors and lag orders for the CBR proxy. The four-variable system drops $\nu^{(B)}$ from the state vector and serves as a placebo. The five-variable variants drop the belief level or the term spread. The six-variable specification is the baseline. VAR(2) and VAR(3) report results for the baseline state vector with longer lag orders.

Table A.12: ρ -Sensitivity at the Five-Year Horizon

Proxy	ρ	W_{60}	ρ_{60}	R_{60}^2
CBR	0.997122	1.68	−0.05	0.10
CBR	0.998122 (Baseline)	1.62	−0.06	0.10
CBR	0.999122	1.56	−0.08	0.10
LTG	0.997122	1.61	−0.02	0.02
LTG	0.998122 (Baseline)	1.56	−0.04	0.02
LTG	0.999122	1.51	−0.05	0.03
STR	0.997122	1.67	−0.04	0.22
STR	0.998122 (Baseline)	1.62	−0.06	0.22
STR	0.999122	1.56	−0.07	0.22

Notes: Five-year horizon diagnostics computed at three values of the log-linearization constant: the baseline $\rho_{\text{monthly}} = 0.998122$ and ± 0.001 around it. W_{60} is the matrix-wedge variance ratio at $h = 60$ months. ρ_{60} is the innovation-level correlation. R_{60}^2 is the forecast-comovement statistic. The qualitative conclusions are insensitive to small variations in ρ .

Table A.13: State-Vector and Lag-Order Robustness, CBR Proxy

Specification	$\mathcal{B}_{\text{scale}}$	News Correlation	W_{obs}	Asy. p_{χ^2}	p_{boot}
4-Var (No $\nu^{(B)}$)	0.24	+0.01	7.03	0.13	0.40
5-Var (No $b^{(B)}$)	0.31	−0.05	7.86	0.16	0.61
5-Var (No ts)	0.28	−0.08	7.18	0.21	0.53
6-Var (Baseline)	0.34	−0.11	8.92	0.18	0.66
VAR(2), 6-Var	0.17	−0.12	13.28	0.35	0.92
VAR(3), 6-Var	0.16	−0.08	18.68	0.41	0.96

Notes: Slope-restriction tests across alternative state vectors and lag orders for the CBR proxy. Sign and slope evidence are stable; scale and the four-variable placebo behave as expected. The four-variable specification drops the behavioral cash-flow news proxy and serves as a placebo: with no $\nu^{(B)}$ in the state vector, $\zeta^{(B)}$ is uncorrelated with $N_{RP}^{(r)}$ by construction. The longer-lag specifications (VAR(2), VAR(3)) fit a richer dynamic system; the slope restriction is more imprecisely estimated but still not rejected.

A.6.9 HAC Bandwidth Sensitivity

Table A.14 reports the slope-restriction Wald statistic and asymptotic p -value for the CBR proxy at alternative Newey–West HAC bandwidths $L \in \{0, 3, 6, 12, 18, 24\}$. The baseline $L = 6$ value is $W = 8.92$, $p_{\chi^2} = 0.178$. Alternative bandwidths produce values in the range $W \in [7.67, 9.43]$ and $p_{\chi^2} \in [0.151, 0.264]$. The qualitative conclusion of slope-restriction non-rejection is unchanged across bandwidths. The body’s primary inference is the null-imposed bootstrap (Appendix A.5.9), which does not depend on a kernel bandwidth.

Table A.14: HAC Bandwidth Sensitivity, CBR Slope Test

HAC Bandwidth L	W_{obs}	Asy. p_{χ^2}
0	8.82	0.18
3	9.43	0.15
6 (Baseline)	8.92	0.18
12	7.73	0.26
18	7.67	0.26
24	7.75	0.26

Notes: Slope-restriction Wald statistic for the CBR proxy at alternative Newey–West HAC bandwidths. The baseline implementation uses $L = 6$. The qualitative conclusion that the slope restriction is not rejected is stable across bandwidths.

A.6.10 Sample and Frequency Robustness

Table A.15 reports sample and frequency robustness checks. “Excluding COVID” truncates the sample at 2019:12, dropping the COVID shock and the post-COVID months. “Post-1990” begins the sample at 1990:01, dropping the early analyst-coverage period. “Quarterly” resamples

the monthly data to quarterly frequency (sums of rx and rf within quarter, end-of-quarter levels for pd , ts , the belief-level control, and the analyst-revision proxy) and re-estimates the VAR with the quarterly Campbell–Shiller constant $\rho_Q = 0.9777^{1/4}$. These checks are not part of the paper’s main identification strategy.

The pattern across specifications is informative. Truncating at 2019:12 strengthens the negative news correlation (CBR: -0.177 , LTG: -0.179 , STR: -0.164) and raises the forecast R^2 across all three proxies, consistent with the COVID period contributing high-variance return innovations that weaken the measured comovement. The post-1990 subsample, by contrast, attenuates and flips the sign of the news correlation across all three proxies: the negative relation is weaker in the second half of the sample. The slope restriction is not rejected in any of these subsamples or at the quarterly frequency. Quarterly $\mathcal{B}_{\text{scale}}$ values cluster near 0.25 (vs. 0.28–0.34 monthly), and the sample forecast R^2 is much smaller at quarterly frequency, mostly because the quarterly sample has only $T = 166$ observations and the sample-vs.-population R^2 gap is larger at quarterly frequency (Appendix A.6.3).

A.6.11 Summary

The additional empirical tables support three conclusions. First, the main sign, scale, forecast-comovement, and 1-step slope results are visible across all three behavioral cash-flow news proxies. Second, the distinction between sample and population forecast R^2 is quantitative rather than conceptual: both measure positive forecast-side comovement between expected returns and expected belief revisions. Third, the longer-horizon and matrix-wedge diagnostics confirm the main text’s interpretation: the data support the belief-driven component, while exact pointwise long-horizon equality is too strong.

Table A.15: Sample and Frequency Robustness

Proxy	Specification	Sample	News Correlation	$\mathcal{B}_{\text{scale}}$	Forecast R^2	Slope p_{boot}
CBR	Baseline Monthly	1982:02–2023:06	−0.11	0.34	0.26	0.66
CBR	Excluding COVID	1982:02–2019:12	−0.18	0.40	0.48	0.72
CBR	Post-1990	1990:01–2023:06	+0.04	0.33	0.23	0.76
CBR	Quarterly	1982:Q1–2023:Q2	−0.07	0.26	0.07	0.72
LTG	Baseline Monthly	1982:02–2023:06	−0.11	0.28	0.26	0.35
LTG	Excluding COVID	1982:02–2019:12	−0.18	0.31	0.41	0.22
LTG	Post-1990	1990:01–2023:06	+0.04	0.27	0.30	0.43
LTG	Quarterly	1982:Q1–2023:Q2	−0.06	0.26	0.04	0.55
STR	Baseline Monthly	1982:02–2023:06	−0.10	0.34	0.21	0.67
STR	Excluding COVID	1982:02–2019:12	−0.16	0.39	0.43	0.75
STR	Post-1990	1990:01–2023:06	+0.05	0.32	0.13	0.73
STR	Quarterly	1982:Q1–2023:Q2	−0.08	0.25	0.04	0.65

Notes: Baseline-sample rows reproduce the headline statistics from Table 1.3. Each non-baseline row re-estimates the legacy six-variable VAR on the relevant sample and reports the news correlation between $N_{RP}^{(r)}$ and $\zeta^{(B)}$, $\mathcal{B}_{\text{scale}} = \text{Var}(\zeta^{(B)})/\text{Var}(N_{RP}^{(r)})$, the squared sample correlation between fitted return forecasts and fitted analyst-revision forecasts, and the null-imposed bootstrap p -value for the 1-step slope restriction $e_1'\Gamma = e_3'\Gamma$ ($B = 2,000$, mean block length 12 months / 4 quarters; $B = 5,000$ for the baseline rows in Table 1.3). The quarterly specification uses the quarterly Campbell–Shiller constant $\rho_Q = \rho_A^{1/4}$.

A.7 Reliability and Measurement-Error Corrections

This appendix reports the empirical reliability diagnostics and measurement-error corrections behind Section 1.6.1. Appendix A.3 shows that measurement error affects scale, correlations, projections, and forecast comovement differently. The purpose of this appendix is empirical: I ask whether the aligned analyst proxies contain a meaningful common belief-revision signal, and how the main estimates change under simple reliability corrections. The corrections are diagnostic rather than structural. They depend on measurement-error assumptions and therefore complement, rather than replace, the baseline evidence.

A.7.1 Cross-Proxy Reliability

The main text uses three aligned analyst-based behavioral cash-flow news proxies: CBR, LTG, and STR. These proxies use different forecast inputs and horizons, so their idiosyncratic measurement errors need not be the same. If the paper's belief measures were dominated by proxy-specific noise, the three series would have little common variation. If they measure a common component of subjective cash-flow belief revision, they should comove.

The proxies are not three fully independent measurements. CBR is a composite that includes a short-horizon component overlapping with STR, and the two share an underlying analyst-forecast input. Cross-proxy reliability evidence should therefore be read as showing that the aligned analyst series share a substantial common belief-revision signal, not as a clean three-independent-measure validation design. The split-sample evidence in Appendix A.7.4 provides a separate cross-check that does not rely on cross-proxy independence.

Table A.16 reports the pairwise correlation matrix of the three aligned proxies on the legacy sample.

Table A.16: Cross-Proxy Correlations

	CBR	LTG	STR
CBR	1.00	0.62	0.90
LTG	0.62	1.00	0.27
STR	0.90	0.27	1.00

Notes: Correlations among the three aligned behavioral cash-flow news proxies on the legacy sample (1982:02–2023:06, $T = 497$). The high CBR–STR correlation (0.90) is partly mechanical: CBR is a composite that contains a short-horizon component related to the STR forecast input.

A.7.2 Common-Factor Diagnostics

As a complementary diagnostic, I extract the first principal component of the standardized proxy vector $y_t = (y_{\text{CBR},t}, y_{\text{LTG},t}, y_{\text{STR},t})'$. Let q_1 denote the first eigenvector of the cross-proxy correlation matrix; the common factor is $PC_{1,t} = q_1' y_t$.

The first principal component explains 74.4% of the standardized variance of the three aligned proxies (eigenvalues 2.232, 0.749, 0.019). Table A.17 reports the loadings, correlations of each proxy with PC_1 , and per-proxy communalities.

The common-factor result is not used as the main empirical specification. The baseline analysis keeps CBR, LTG, and STR separate because they measure different horizons and forecast inputs. The common-factor evidence instead supports the measurement interpretation: the three aligned proxies share a large belief-revision component.

Table A.17: Common-Factor Diagnostics

Proxy	PC1 Loading	Corr(Proxy, PC1)	Communality
CBR	+0.66	+0.99	0.98
LTG	+0.46	+0.69	0.48
STR	+0.59	+0.88	0.77
PC1 Variance Share			0.74

Notes: First principal component of the three standardized aligned behavioral cash-flow news proxies. Communality is the squared correlation with PC_1 , equal to the share of standardized proxy variance explained by the common factor. The final row reports the share of total standardized variance explained by PC_1 (the leading eigenvalue divided by 3).

A.7.3 Reliability Lower Bounds

The reliability calculations summarize how much of each measured proxy can be attributed to a common signal rather than proxy-specific noise. Let a generic measured proxy be written as $y_{j,t} = s_{j,t} + u_{j,t}$, where $s_{j,t}$ is the reliable signal component and $u_{j,t}$ is proxy-specific measurement error. The reliability of proxy j is

$$\mathcal{R}_j = \frac{\text{Var}(s_{j,t})}{\text{Var}(y_{j,t})}.$$

When proxy-specific errors are orthogonal across measures, the cross-proxy covariance structure gives an empirical lower bound for $\mathcal{R}_j$. The lower-bound interpretation uses only common variation as reliable signal and is conservative in that sense; the qualifier in Appendix A.7.1 about CBR–STR construction overlap applies here too.

The signal-attenuation exercise in Appendix A.3.2 implies a per-proxy critical reliability threshold $\mathcal{R}_j^{\text{crit}}$: above this value, the estimated behavioral signal remains economically meaningful after correcting for attenuation. Table A.18 reports the lower bounds and thresholds for

the three proxies.

Table A.18: Reliability Lower Bounds

Proxy	Reliability Lower Bound	Critical Threshold $\mathcal{R}_j^{\text{crit}}$	Above Threshold?
CBR	0.86	0.58	Yes
LTG	0.64	0.53	Yes
STR	0.70	0.58	Yes

Notes: Reliability is the share of measured proxy variance attributed to common signal variation under the cross-proxy measurement model. The critical threshold is the proxy-specific signal-attenuation cutoff above which the behavioral signal remains economically meaningful after correcting for attenuation.

The threshold calculation is tied to the economic-magnitude criterion and the signal-attenuation model used in the reliability exercise. It is useful because it asks a concrete question: would the measured signal remain meaningful after allowing for reliability below one? The lower bound is above the threshold for all three proxies under the maintained measurement model.

A.7.4 Split-Sample Reliability

The split-sample evidence asks whether the belief component is present across distinct subsamples. It is a separate diagnostic from the cross-proxy reliability lower bound, so it is not subject to the CBR–STR construction-overlap concern. I split the legacy sample at its midpoint and estimate the headline news correlation and behavioral scale ratio in each half; Table A.19 reports the results.

The negative news correlation is present in both halves of the sample for all three proxies. The early-half correlations are larger in absolute value (−0.42 to −0.45) than the late-half

Table A.19: Split-Sample Reliability

Proxy	Early Sample			Late Sample		
	T	$\mathcal{B}_{\text{scale}}$	News Correlation	T	$\mathcal{B}_{\text{scale}}$	News Correlation
CBR	226	0.28	−0.42	269	0.45	−0.12
LTG	226	0.26	−0.45	269	0.40	−0.14
STR	226	0.30	−0.42	269	0.44	−0.12

Notes: Sample split at the median date of the legacy window. For each subsample, the legacy six-variable VAR is re-estimated and the table reports $\mathcal{B}_{\text{scale}} = \text{Var}(\zeta^{(B)})/\text{Var}(N_{RP}^{(r)})$ and $\text{Corr}(N_{RP}^{(r)}, \zeta^{(B)})$.

correlations (−0.12 to −0.14), and the behavioral scale ratio is larger in the late half. The pattern is consistent with the post-1990 attenuation reported in Appendix A.6.10: the negative relation is concentrated in the earlier subsample. The headline result is not confined to one half of the sample, however, and does not reverse sign in either half.

A.7.5 Reliability-Corrected Correlations and Projections

Reliability corrections are straightforward for correlations under the signal-attenuation model.

Suppose

$$\tilde{\zeta}_{t+1} = a \zeta_{CF,t+1} + u_{t+1}, \quad \mathcal{R} = \frac{a^2 \text{Var}(\zeta_{CF,t+1})}{a^2 \text{Var}(\zeta_{CF,t+1}) + \text{Var}(u_{t+1})}.$$

As shown in Appendix A.3.2,

$$\text{Corr}(N_{RP,t+1}^{(r)}, \tilde{\zeta}_{t+1}) = \text{sign}(a) \text{Corr}(N_{RP,t+1}^{(r)}, \zeta_{CF,t+1}) \sqrt{\mathcal{R}}.$$

Under the maintained sign convention – positive analyst optimism corresponds to a positive cash-flow belief-bias proxy ($a > 0$) – the reliability-corrected correlation is

$$\text{Corr}^{\text{corrected}} = \frac{\text{Corr}^{\text{observed}}}{\sqrt{\mathcal{R}}}.$$

The same attenuation form holds under additive classical noise (Appendix A.3.1), so the corrected correlations reported in Table A.20 apply to either model. The two models differ in their implications for variance scale and projection coefficients, not for the correlation correction.

Table A.20: Reliability-Corrected News Correlations

Proxy	Observed Correlation	Reliability Anchor	Corrected Correlation
CBR	−0.11	0.86	−0.12
LTG	−0.11	0.64	−0.13
STR	−0.10	0.70	−0.12

Notes: The corrected correlation divides the observed correlation by the square root of the reliability lower bound from Table A.18. The correction assumes positive signal loading; the same form holds under both Model A (additive classical noise) and Model B (signal attenuation). Bootstrap intervals for the corrected correlation are not reported because the reliability-bootstrap procedure currently records the corrected projection coefficient; see Appendix A.7.7.

The corrected correlations remain economically modest in absolute value but do not reverse sign. Allowing for reliability below one increases the negative relation slightly rather than removing it. The projection-coefficient corrections, which depend on the measurement-model normalization, are reported next.

A.7.6 Model-Based Measurement Corrections

Different measurement-error models imply different corrections. I consider two benchmark cases.

Model A: Additive Classical Noise. The observed proxy is $\tilde{\zeta}_{t+1} = \zeta_{CF,t+1} + u_{t+1}$. As shown in Appendix A.3.1, additive noise increases the measured variance of the proxy:

$$\text{Var}(\tilde{\zeta}) = \text{Var}(\zeta_{CF}) + \text{Var}(u).$$

It can therefore inflate $\mathcal{B}_{\text{scale}}$. At the same time, it attenuates correlations and projection coefficients toward zero. In this model, a large measured scale ratio is not automatically conservative, but a negative news correlation is harder to obtain.

Model B: Signal Attenuation or Incomplete Proxy. The observed proxy is $\tilde{\zeta}_{t+1} = a \zeta_{CF,t+1} + u_{t+1}$ with $0 < a < 1$, or it measures only one component of a broader belief wedge. In this model, the observed proxy can understate the variance of the full behavioral object if attenuation or incomplete coverage dominates additive noise. This is the model under which reliability corrections can increase the implied magnitude of the behavioral component.

The empirical evidence does not select one measurement model definitively. The paper uses the models to inform interpretation. Additive noise cautions against reading $\mathcal{B}_{\text{scale}}$ as a structural share. Signal attenuation and incomplete-proxy models caution against interpreting the measured proxy as the full belief wedge. In both cases, the sign and forecast-comovement evidence remain central because they are not mechanically generated by proxy variance alone. Table A.21 summarizes the implications of each measurement-error model.

Table A.21: Measurement-Error Models and Implications

Model	Effect on Scale	Effect on Correlations and Projections
Additive Classical Noise	Inflates the measured $\mathcal{B}_{\text{scale}}$ ratio	Attenuates correlations and projections toward zero
Signal Attenuation	Understates $\mathcal{B}_{\text{scale}}$ if attenuation dominates	Attenuates correlations by $\sqrt{\mathcal{R}}$ under loading normalization
Incomplete Proxy	Captures only the observed component of the wedge	Identifies relation with the measured component only
Nonclassical Contamination	Direction is ambiguous in general	Requires robustness tests rather than scalar correction

Notes: The table summarizes the implications of the measurement-error models derived in Appendix A.3. Non-classical contamination includes price inference, residual rational-news comovement, and state-correlated proxy error.

A.7.7 Bootstrap Uncertainty for Corrected Projections

The reliability corrections are generated objects. They inherit uncertainty from the proxy covariance matrix, the estimated VAR, and the news-construction procedure. Table A.22 reports stationary-block bootstrap intervals for the corrected projection coefficient under both measurement models, using the cross-proxy reliability lower bound as the anchor. The point estimates are corrected analogues of the projection of $N_{RP}^{(r)}$ on the (corrected) belief-bias innovation.

The corrected-projection intervals are wide and cross zero for all three proxies under both models. This is partly mechanical: the projection coefficient is small in absolute value, the correction divides by a number below one, and the bootstrap inherits the noise of the underlying projection. The intervals should therefore be read as illustrating the uncertainty of the corrected scalar, not as a separate test of the directional channel. The directional inference

Table A.22: Bootstrap Uncertainty for Reliability-Corrected Projections

Proxy	$\hat{\beta}^{\text{observed}}$	Model A (Additive)	Model B (Attenuation)
CBR	−0.18	−0.21 [−1.03, 0.64]	−0.16 [−0.77, 0.48]
LTG	−0.20	−0.32 [−1.63, 0.96]	−0.13 [−0.67, 0.40]
STR	−0.17	−0.24 [−1.24, 0.80]	−0.12 [−0.61, 0.39]

Notes: Stationary-block bootstrap point estimates and percentile intervals for the corrected projection coefficient $\hat{\beta}^{\text{corrected}}$ from regressing $N_{RP}^{(r)}$ on the (corrected) belief-bias innovation. Brackets below each point estimate report the 95% percentile bootstrap confidence interval. Reliability anchor is the cross-proxy lower bound from Table A.18. Model A divides the observed projection coefficient by $\mathcal{R}$; Model B applies the signal-attenuation correction. $B = 5,000$, mean block length 12 months (Appendix A.5.8).

comes from the news-correlation and forecast-comovement results in Section 1.4.

A.7.8 Interpretation

The reliability evidence supports three conclusions. First, the aligned analyst proxies contain a substantial common component: the first principal component explains 74.4% of the standardized proxy variance, and the cross-proxy reliability lower bound is well above the proxy-specific signal-attenuation threshold for all three proxies. Second, the negative news correlation is present in both halves of the legacy sample, although it is larger in absolute value in the early half. Third, correcting the news correlation for the cross-proxy reliability lower bound increases the magnitude of the negative relation slightly rather than eliminating it.

These diagnostics do not prove that analyst revisions equal marginal-investor beliefs, and the cross-proxy bound is partly anchored by construction overlap between CBR and STR. The

split-sample evidence is the natural cross-check on that concern. Combined with the price-inference, adversarial-residual, and cross-sectional exercises, the reliability diagnostics support the interpretation that the measured belief-bias news is not merely idiosyncratic analyst noise.

A.8 Robustness and Adversarial Exercises

This appendix reports the robustness and adversarial exercises behind Section 1.6. The exercises address three concerns. First, analyst revisions may mechanically incorporate information from prices. Second, omitted rational risk premium news may dilute the relation between measured belief-bias news and risk premium news. Third, a residual component deliberately aligned with the behavioral cash-flow news proxy may offset the baseline relation. The appendix also reports a calibration of the partial-arbitrage loading and the algebraic identity linking the matrix wedge to the projection coefficient and behavioral scale.

A.8.1 Price-Inference Robustness

A concern with analyst-based behavioral cash-flow news proxies is that analysts may infer information from prices. If analyst revisions mechanically incorporate contemporaneous valuation information, the measured belief-bias relation could partly reflect information already contained in the price-dividend ratio. I study this channel by removing the component of the analyst revision that is linearly related to the contemporaneous price-dividend ratio.

For each proxy j , let $\nu_{j,t+1}^{(B)}$ be the baseline aligned analyst revision. Let $\widehat{\nu}_{j,t+1}^{pd} = \widehat{a}_j + \widehat{b}_j pd_t$ denote the fitted component from the projection of $\nu_{j,t+1}^{(B)}$ on the contemporaneous price-dividend ratio. For $\phi \in [0, 1]$, define the price-residualized proxy

$$\nu_{j,t+1}^{(B,\phi)} = \nu_{j,t+1}^{(B)} - \phi \widehat{\nu}_{j,t+1}^{pd}. \quad (\text{A.24})$$

The case $\phi = 0$ is the baseline. The case $\phi = 1$ removes the full contemporaneous price-

dividend projection. For each value of ϕ , I reconstruct the VAR news objects, the belief-bias news proxy, and the main sign, scale, and slope statistics.

Table A.23 reports the sweep at $\phi \in \{0, 0.25, 0.50, 0.75, 1.0\}$ for all three proxies. For CBR, moving from $\phi = 0$ to $\phi = 1$ changes $\mathcal{B}_{\text{scale}}$ from 0.342 to 0.341, the news correlation from -0.107 to -0.106 , and the slope-test bootstrap p -value from 0.666 to 0.693. The result is essentially flat across the full ϕ -sweep. LTG and STR show the same pattern.

Table A.23: Price-Inference Robustness Sweep

Proxy	ϕ	$\mathcal{B}_{\text{scale}}$	News Correlation	W_{obs}	Slope p_{boot}
CBR	0.00	0.34	-0.11	8.92	0.67
CBR	0.25	0.34	-0.11	8.76	0.67
CBR	0.50	0.34	-0.11	8.61	0.68
CBR	0.75	0.34	-0.11	8.46	0.69
CBR	1.00	0.34	-0.11	8.32	0.69
LTG	0.00	0.28	-0.11	15.76	0.34
LTG	0.25	0.28	-0.11	15.66	0.35
LTG	0.50	0.28	-0.11	15.56	0.35
LTG	0.75	0.28	-0.11	15.46	0.35
LTG	1.00	0.28	-0.11	15.36	0.36
STR	0.00	0.34	-0.10	9.03	0.68
STR	0.25	0.34	-0.10	8.96	0.68
STR	0.50	0.34	-0.10	8.88	0.69
STR	0.75	0.34	-0.10	8.80	0.69
STR	1.00	0.34	-0.10	8.73	0.69

Notes: Sweep over the price-inference residualization parameter ϕ in equation (A.24). $\phi = 0$ is the baseline proxy; $\phi = 1$ removes the full contemporaneous price-dividend projection from the analyst-revision proxy before reconstructing belief-bias news and test statistics. Bootstrap p -values use $B = 2,000$ at each ϕ .

The exercise does not rule out every form of price-related information in analyst forecasts. It rules out a specific direct channel: the baseline relation is not driven by the contemporaneous price-dividend component of the analyst-revision proxy.

A.8.2 Orthogonal Residual Dilution

This subsection studies how much generic residual rational news would be needed to dilute the baseline news correlation. Let $N_{t+1} \equiv N_{RP,t+1}^{(r)}$, $Z_{t+1} \equiv \zeta_{t+1}^{(B)}$, and $c = \text{Corr}(N_{t+1}, Z_{t+1})$. In the baseline estimates, $c < 0$. Consider an adversarial residual component that is orthogonal to the measured belief-bias innovation and the VAR innovation vector:

$$N_{t+1}^{\text{adv}} = N_{t+1} + \kappa w_{t+1},$$

where $\mathbb{E}(w_{t+1}) = 0$, $\text{Var}(w_{t+1}) = 1$, and $\text{Cov}(w_{t+1}, Z_{t+1}) = 0$. Then

$$\text{Corr}(N_{t+1}^{\text{adv}}, Z_{t+1}) = \frac{c}{\sqrt{1 + \kappa^2 / \sigma_N^2}}, \quad (\text{A.25})$$

where $\sigma_N^2 = \text{Var}(N_{t+1})$.

To halve the absolute correlation, equation (A.25) requires $\sqrt{1 + \kappa^2 / \sigma_N^2} = 2$, so

$$\frac{\kappa^*}{\sigma_N} = \sqrt{3} \approx 1.732.$$

This result is proxy-independent because it is a relative attenuation calculation. To drive the absolute correlation below a fixed threshold $\tau > 0$, the required residual volatility is

$$\frac{\kappa^*(\tau)}{\sigma_N} = \sqrt{\left(\frac{|c|}{\tau}\right)^2 - 1}.$$

Using $\tau = 0.025$, the required residual volatility is about four times the volatility of measured risk premium news. Table A.24 reports the absolute-threshold ratios with stationary-block

bootstrap intervals.

Table A.24: Orthogonal Residual Dilution

Proxy	Baseline $ c $	Halving κ^*/σ_N	$ c < 0.025$: κ^*/σ_N
CBR	0.11	1.73	4.15 [0.90, 21.26]
LTG	0.11	1.73	4.16 [0.89, 22.32]
STR	0.10	1.73	3.86 [0.84, 21.00]

Notes: The adversarial residual is orthogonal to measured belief-bias news. The third column reports the residual volatility required to halve the absolute correlation (proxy-independent). The fourth column reports the residual volatility required to push the absolute correlation below $\tau = 0.025$. Confidence intervals are stationary-block bootstrap percentile intervals ($B = 5,000$, mean block length 12 months).

The CI ranges are wide. The lower-end values (~ 0.9) imply that under benign sampling realizations, an orthogonal residual only modestly larger than N_{RP} would be enough to push the correlation below 0.025. The point estimate of about four times the volatility is therefore a central tendency, not a precisely estimated bound. Generic orthogonal-residual news cannot easily erase the baseline relation in the point estimates, but the empirical robustness margin should be interpreted as a sensitivity range, not a sharp threshold.

A.8.3 Targeted Residual Adversary

The orthogonal dilution exercise is intentionally generic. A harder alternative allows omitted residual news to be deliberately aligned with the measured belief-bias innovation. This is the structured alternative most difficult to rule out.

Let the adversarial residual be A_{t+1} , with variance budget $q = \text{Var}(A_{t+1})/\text{Var}(N_{t+1})$, and let $s = \text{Corr}(A_{t+1}, Z_{t+1})$ measure its alignment with the belief-bias innovation. The

adversarial risk premium news object is $N_{t+1}^{\text{adv}} = N_{t+1} + A_{t+1}$. Its covariance with Z_{t+1} is

$$\text{Cov}(N_{t+1}^{\text{adv}}, Z_{t+1}) = c \sigma_N \sigma_Z + s \sqrt{q} \sigma_N \sigma_Z,$$

which is zero when $c + s\sqrt{q} = 0$. For an adversary with sign chosen to offset the baseline relation, the required variance budget is

$$q_{\text{zero}}^*(s) = \left(\frac{|c|}{s} \right)^2. \quad (\text{A.26})$$

The maximally targeted case sets $s = 1$, giving $q_{\text{zero}}^* = c^2$. Table A.25 reports the maximally targeted budgets with bootstrap intervals.

Table A.25: Targeted Residual Adversary

Proxy	Baseline Correlation	$q_{\text{zero}}^* (s = 1)$
CBR	−0.11	0.01 [0.00, 0.28]
LTG	−0.11	0.01 [0.00, 0.31]
STR	−0.10	0.01 [0.00, 0.27]

Notes: Variance budget required for a maximally targeted adversarial residual to drive the covariance between risk premium news and measured belief-bias news to zero. The adversary is deliberately aligned with the belief-bias proxy ($s = 1$). The wide CIs reinforce that this is a sensitivity bound, not a precisely identified threshold.

The point estimate – about 1% of $\text{Var}(N_{RP})$ – gives the boundary of the empirical design. The baseline relation is robust to generic orthogonal dilution, but a residual rational-news component specifically aligned with the behavioral cash-flow news proxy is the hardest structured alternative to rule out. The bootstrap intervals are wide because the underlying covariance between N_{RP} and $\zeta^{(B)}$ is small in absolute value.

A.8.4 Adversarial Frontier

The orthogonal and targeted exercises are endpoints. The adversarial frontier interpolates by allowing the omitted residual to have correlation $s \in (0, 1]$ with the belief-bias innovation. Equation (A.26) gives the variance budget required to eliminate the covariance:

$$q_{\text{zero}}^*(s) = \left(\frac{|c|}{s} \right)^2.$$

As s rises, the required residual variance falls quickly. Table A.26 reports the frontier across alignments $s \in \{0.25, 0.50, 0.75, 1.00\}$. This is the central lesson of the frontier: the empirical relation is difficult to erase with unrelated residual news, but much easier to erase with a residual deliberately aligned against the measured behavioral cash-flow news proxy.

Table A.26: Adversarial Frontier

Proxy	$s = 0.25$	$s = 0.50$	$s = 0.75$	$s = 1.00$
CBR	0.18	0.05	0.02	0.01
LTG	0.18	0.05	0.02	0.01
STR	0.16	0.04	0.02	0.01

Notes: The table reports $q_{\text{zero}}^*(s) = (|c|/s)^2$, the variance budget needed to drive the covariance between risk premium news and measured belief-bias news to zero, as a function of the residual's alignment $s = \text{Corr}(A, Z)$. Values are computed in closed form from the baseline news correlations $|c|$ in Table 1.3.

The frontier is not a robustness test in the usual sense. It is a sensitivity analysis. It identifies what kind of omitted residual risk premium news component would overturn the baseline relation: a small, deliberately belief-aligned component would suffice; a generic orthogonal one would not.

A.8.5 Structural λ Calibration

The partial-arbitrage model in Appendix A.1.4 implies

$$N_{RP,t+1}^{\lambda,r} = N_{RP,t+1}^{*,r} - \lambda \zeta_{CF,t+1}.$$

If residual rational risk premium news is orthogonal to the belief-bias innovation, then the projection of $N_{RP}^{(r)}$ on ζ_{CF} identifies $-\lambda$:

$$\frac{\text{Cov}(N_{RP}^{(r)}, \zeta_{CF})}{\text{Var}(\zeta_{CF})} = -\lambda.$$

The empirical analogue uses $\zeta^{(B)}$. Let $\hat{\beta}_j = \widehat{\text{Cov}}(N_{RP}^{(r)}, \zeta_j^{(B)}) / \widehat{\text{Var}}(\zeta_j^{(B)})$. The calibrated loading is $\hat{\lambda}_j = -\hat{\beta}_j$. Equivalently, using the news correlation c_j and behavioral scale $B_j = \text{Var}(\zeta_j^{(B)}) / \text{Var}(N_{RP}^{(r)})$,

$$\hat{\lambda}_j = -\frac{c_j}{\sqrt{B_j}}.$$

Table A.27 reports the calibrated values for each proxy.

Table A.27: Partial-Arbitrage Loading Calibration

Proxy	News Correlation	$\mathcal{B}_{\text{scale}}$	$\hat{\lambda}$
CBR	-0.11	0.34	0.18 [-0.55, 0.89]
LTG	-0.11	0.28	0.20 [-0.62, 1.04]
STR	-0.10	0.34	0.17 [-0.56, 0.87]

Notes: The calibrated loading uses $\hat{\lambda}_j = -c_j / \sqrt{B_j}$. The calculation assumes that residual rational risk premium news is orthogonal to measured belief-bias news. Bootstrap point estimates and percentile intervals use $B = 5,000$ stationary-block draws (Appendix A.5.8).

The point estimates are around 0.17–0.20. I interpret them as calibration evidence, not as precise structural estimates. The identification of λ requires the orthogonality condition $\text{Cov}(N_{RP}^{*,r}, \zeta_{CF}) = 0$, and the empirical proxy $\zeta^{(B)}$ is measured with error. The bootstrap intervals span zero in all three cases. The main evidence for the belief-driven component therefore comes from the joint sign, scale, forecast, robustness, and cross-sectional results, not from a standalone structural estimate of λ .

A.8.6 Algebraic Identity for the Matrix Wedge

This subsection derives the algebraic identity linking the matrix-wedge variance ratio, the projection coefficient, and behavioral scale. Let $N = N_{RP}^{(r)}$, $Z = \zeta^{(B)}$, and $W = N + Z$. Define

$$B = \frac{\text{Var}(Z)}{\text{Var}(N)}, \quad \beta = \frac{\text{Cov}(N, Z)}{\text{Var}(Z)}.$$

Then

$$\begin{aligned} \frac{\text{Var}(W)}{\text{Var}(N)} &= 1 + \frac{\text{Var}(Z)}{\text{Var}(N)} + 2 \frac{\text{Cov}(N, Z)}{\text{Var}(N)} \\ &= 1 + B + 2\beta B \\ &= 1 + B(1 + 2\beta). \end{aligned}$$

This is purely algebraic. It expresses the matrix-wedge variance ratio as a function of two observed second moments: the behavioral scale B and the projection coefficient β . Once B and β are fixed, the wedge ratio is mechanically determined; it is not an independent overidentifying moment in the standard innovation-space construction.

The same expression can be rewritten in terms of the partial-arbitrage loading. Using $\lambda = -\beta$, which is the structural calibration in Appendix A.8.5 under the orthogonality assumption $\text{Cov}(N_{RP}^{*,r}, \zeta_{CF}) = 0$,

$$\frac{\text{Var}(W)}{\text{Var}(N)} = 1 + B(1 - 2\lambda). \quad (\text{A.27})$$

Equation (A.27) is not purely algebraic: it imposes the structural identification of λ from the projection coefficient. The two identities are numerically equivalent on the data but should be read differently – the β form is a definitional rearrangement; the λ form is an interpretation.

Table A.28 verifies the algebraic identity numerically.

Table A.28: Matrix-Wedge Identity Check

Proxy	B	$\hat{\lambda}$	$1 + B(1 - 2\hat{\lambda})$	Observed Wedge Ratio
CBR	0.34	0.18	1.22	1.22
LTG	0.28	0.20	1.17	1.17
STR	0.34	0.17	1.22	1.22

Notes: The table verifies the algebraic link between behavioral scale, the projection-implied loading, and the matrix-wedge variance ratio. Predicted values use equation (A.27); observed values from Table 1.5. The agreement is exact up to rounding.

Genuine overidentification therefore requires additional restrictions not mechanically pinned down by the same two moments. Examples include forecast-space restrictions (Appendix A.2.2), direct horizon restrictions (Appendix A.2.6), external proxy variation, or cross-sectional exposure tests (Section 1.7).

State-vector and lag-order robustness checks complementing this appendix's residual-news exercises are reported in Appendix A.6.8; HAC bandwidth sensitivity is in Appendix A.6.9. The exercises in this appendix focus on price-inference and residual-news alternatives.

A.9 Cross-Sectional Evidence

This appendix reports the cross-sectional evidence behind Section 1.7. The appendix has three goals. First, it documents the FF25 portfolio-exposure design used in Section 1.7.1. Second, it explains why the main analysis does not use fully portfolio-level analyst behavioral cash-flow news proxies. Third, it reports the level-vs.-reallocation decomposition behind Section 1.7.2. The cross-sectional results are used as breadth and mechanism evidence. They are not the paper's main identification design.

A.9.1 FF25 Portfolio Exposure Design

The FF25 exercise asks whether portfolio-level risk premium news is exposed to the same aggregate belief-bias news series studied in the market-level tests. For each Fama–French size-by-book-to-market portfolio p , I replace the market excess return in the baseline forecasting system with the portfolio excess return and re-estimate the VAR:

$$x_{p,t} = \left[rx_{p,t}, \quad rf_t, \quad \nu_t^{(B)}, \quad b_t^{(B)}, \quad pd_t, \quad ts_t \right]'. \quad (\text{A.28})$$

The remaining state variables are aggregate. Thus portfolio-level risk premium news varies across FF25 cells, while the analyst-revision proxy, belief-level control, market price-dividend ratio, and term spread remain at the market level.

For each portfolio p , I construct portfolio-level risk premium news $N_{RP,p,t+1}^{(r)}$ from the portfolio-specific VAR. The belief-bias news series $\zeta_{t+1}^{(B)}$ remains the aggregate market-level

series. The reported correlation is therefore

$$\text{Corr}\left(N_{RP,p,t+1}^{(r)}, \zeta_{t+1}^{(B)}\right).$$

The FF25 analysis is therefore a portfolio-exposure exercise, not a fully portfolio-specific belief decomposition. It asks which portfolio return-news objects comove with aggregate belief-bias news.

A.9.2 FF25 News-Correlation Heatmap and Table

Figure 1.2 in the main text reports the FF25 heatmap for the baseline (CBR) proxy. Table A.29 reports the corresponding cell values. Rows are size quintiles, from small (Size 1) to large (Size 5). Columns are book-to-market quintiles, from growth (BM1) to value (BM5).

Table A.29: FF25 Portfolio Exposure: News Correlations

	BM1	BM2	BM3	BM4	BM5
Size 1	−0.27	−0.02	−0.13	−0.20	−0.50
Size 2	−0.34	−0.08	−0.34	−0.30	−0.34
Size 3	−0.39	−0.09	−0.14	−0.28	−0.36
Size 4	−0.09	+0.01	+0.01	+0.08	−0.08
Size 5	−0.34	+0.10	+0.24	+0.38	−0.28

Notes: The table reports $\text{Corr}(N_{RP,p}^{(r)}, \zeta^{(B)})$ across the 25 Fama–French size-by-book-to-market portfolios for the CBR proxy. For each cell, the VAR replaces the market excess return with the portfolio excess return while keeping the remaining state variables aggregate, as in equation (A.28). Negative values indicate the sign predicted by the belief-driven risk premium news channel.

The predicted negative sign appears in 19 of the 25 portfolios. The six positive cells are concentrated in size 4–5 and book-to-market 2–4, consistent with the body’s characterization

that the positive cells lie in the largest two size quintiles and the middle three book-to-market quintiles. The pattern is therefore broad but not uniform: aggregate belief-bias news is related to risk premium news across many portfolio return-news objects, not only in the aggregate market return.

A.9.3 FF25 Slope Tests

For each FF25 portfolio, I also test the 1-step slope restriction using the portfolio-specific return row and the aggregate analyst-revision row. The test is the portfolio analogue of the market-level restriction $e_1'\Gamma = e_3'\Gamma$, where e_1 selects the portfolio excess return and e_3 selects the aggregate behavioral cash-flow news proxy. Table A.30 reports null-imposed bootstrap p -values.

Table A.30: FF25 Portfolio Exposure: Slope-Test Bootstrap p -Values

	BM1	BM2	BM3	BM4	BM5
Size 1	0.47	0.75	0.62	0.66	0.51
Size 2	0.69	0.73	0.41	0.53	0.57
Size 3	0.64	0.76	0.64	0.56	0.67
Size 4	0.85	0.70	0.76	0.90	0.73
Size 5	0.52	0.63	0.79	0.87	0.76

Notes: The table reports null-imposed bootstrap p -values for the 1-step slope restriction in each FF25 portfolio VAR, with $B = 500$ stationary-block draws (lower than the headline $B = 5,000$ to keep wall time manageable across 25 portfolio-specific systems). None of the 25 portfolio-level tests rejects under the bootstrap at conventional levels. Portfolio-level tests are interpreted as breadth evidence: the cross-section does not display a systematic pattern of slope rejections that would contradict the aggregate evidence.

None of the 25 portfolio-level slope tests rejects under the bootstrap. Portfolio-level tests are naturally noisier than the market-level tests, so this result should not be read as proof of

exact equality in every cell. Its value is that the portfolio exercise does not produce a broad pattern of rejections against the aggregate time-series evidence.

A.9.4 Portfolio-Level Behavioral Cash-Flow News Coverage

A fully portfolio-specific belief decomposition would require constructing portfolio-level analyst behavioral cash-flow news proxies $\nu_{p,t}^{(B)}$, $\eta_{p,t}^{(B)}$, and $\zeta_{p,t}^{(B)}$ within each FF25 cell. I do not use that design in the main analysis because analyst coverage is too sparse in many cells. Table A.31 reports the median number of IBES-covered firms per cell per month for the CBR proxy. LTG and STR show the same coverage pattern.

Table A.31: Median IBES Coverage by FF25 Cell

	BM1	BM2	BM3	BM4	BM5
Size 1	1	1	1	1	1
Size 2	1	2	3	2	2
Size 3	4	5	5	6	4
Size 4	31	30	21	17	12
Size 5	101	61	43	37	22

Notes: Median number of IBES-covered firms per FF25 cell per month for the CBR proxy. Size 1 is the smallest size quintile; Size 5 is the largest. BM1 is the lowest book-to-market quintile; BM5 is the highest. Only 10 of 25 cells have median coverage of at least 10 firms; 12 cells have median coverage of two firms or fewer. Fully portfolio-level behavioral cash-flow news proxies in those sparse cells would be dominated by individual-firm noise.

The coverage diagnostic explains the design choice in Section 1.7.1. The FF25 exercise keeps belief-bias news aggregate and varies the return-news object across portfolios. This preserves the aggregate belief object studied in the main paper while showing whether the relation is broad across portfolio return-news components.

A.9.5 Sorting Evidence

The sorting evidence asks whether the belief-news relation is stronger in portfolios where belief measurement or belief heterogeneity should be more salient. For each sort variable, I form size-tercile portfolios and re-estimate the legacy six-variable VAR with the tercile excess return in place of the market excess return. Table A.32 reports the news correlation in each tercile. Table A.33 reports the high-minus-low spreads for the LTG and analyst-coverage sorts, with bootstrap intervals.

Table A.32: Sorting Evidence: News Correlations by Tercile

Sort Variable	Low (T_1)	Middle (T_2)	High (T_3)
Size	−0.30	−0.01	−0.14
LTG Exposure	+0.21	+0.00	−0.29
Analyst Coverage	+0.09	−0.16	−0.26
Earnings–Price	−0.18	−0.22	+0.32

Notes: The table reports $\text{Corr}(N_{RP,p}^{(r)}, \zeta^{(B)})$ within each tercile portfolio for four sort variables. The LTG and coverage sorts are designed so that high-tercile portfolios are where belief heterogeneity or measurement should be most salient. The earnings–price sort is included as an additional cross-section split.

The sorting results are used as mechanism evidence, not as a separate source of identification. Their role is to show that the cross-sectional patterns point in the expected direction while acknowledging that portfolio-level estimates are noisy.

A.9.6 Level vs. Reallocation Decomposition

This subsection reports the decomposition of the aggregate analyst-revision proxy into common level movement and size-weighted reallocation across firms. Let $\nu_{i,t}^{(B)}$ denote the firm-

Table A.33: Sorting Evidence: High-minus-Low Spreads

Sort Variable	High-Low Spread	p_{boot}
LTG Exposure	−0.50 [−0.54, 0.22]	0.34
Analyst Coverage	−0.35 [−0.35, 0.37]	0.88

Notes: Spreads in the news correlation between the high-tercile and low-tercile portfolios. Negative values indicate a stronger negative belief-news relation in the high tercile, the predicted direction. Brackets below each point estimate report the 95% stationary-block bootstrap confidence interval. Bootstrap p -values are two-sided and use $B = 1,000$ stationary-block draws. The large p -values reinforce that sorting evidence is directional mechanism evidence, not a separate sharp test. Bootstrap spread inference is reported only for sorts where it was run; the size and earnings-price sorts in Table A.32 have tercile-level correlations only.

level aligned revision for firm i , and let $w_{i,t}$ denote the lagged market-value weight used in aggregation. The value-weighted aggregate revision is

$$\nu_t^{(B)} = \sum_{i \in \mathcal{I}_t} w_{i,t} \nu_{i,t}^{(B)}.$$

Let $N_t = |\mathcal{I}_t|$ and define the equal-weighted level component

$$\text{LEVEL}_t = \frac{1}{N_t} \sum_{i \in \mathcal{I}_t} \nu_{i,t}^{(B)}.$$

The size-weighted reallocation component is

$$\text{CSCOV}_t = \sum_{i \in \mathcal{I}_t} \left(w_{i,t} - \frac{1}{N_t} \right) \nu_{i,t}^{(B)}.$$

Then

$$\nu_t^{(B)} = \text{LEVEL}_t + \text{CSCOV}_t.$$

Because $\sum_{i \in \mathcal{I}_t} w_{i,t} = 1$ and $\sum_{i \in \mathcal{I}_t} 1/N_t = 1$, the two components partition the value-weighted aggregate revision exactly. The first term captures common movement in firm-level belief revisions; the second captures whether revisions are concentrated among firms with high market weights.

For each proxy j , the marginal correlation of each component with risk premium news answers a different question than the covariance share. A component can have a smaller correlation but much larger variance, and therefore contribute more to the aggregate covariance. The covariance share of component $k \in \{\text{LEVEL}, \text{CSCOV}\}$ is

$$\omega_{j,k} = \frac{\text{Cov}(k_{j,t}, N_{RP,t}^{(r)})}{\text{Cov}(\nu_{j,t}^{(B)}, N_{RP,t}^{(r)})}.$$

Because $\nu_j^{(B)} = \text{LEVEL}_j + \text{CSCOV}_j$, the two shares sum to one up to rounding. Table A.34 reports both objects: the marginal correlations with Newey–West t -statistics in parentheses, and the covariance shares.

Table A.34: Level and Reallocation: Correlations and Covariance Contributions

Proxy	Correlation with $N_{RP}^{(r)}$		Covariance Share (%)	
	LEVEL	CSCOV	LEVEL	CSCOV
CBR	−0.05 (−0.80)	−0.11 (−1.99)	54.50	45.50
LTG	−0.07 (−1.28)	−0.02 (−0.52)	97.86	2.14
STR	−0.02 (−0.33)	−0.11 (−2.00)	30.23	69.77

Notes: The table decomposes the market-level analyst revision into an equal-weighted level component, LEVEL, and a size-weighted reallocation component, CSCOV. The first block reports the correlation of each component with rational risk premium news $N_{RP}^{(r)}$, with Newey–West t -statistics in parentheses below each point estimate. The second block reports each component’s contribution to the aggregate *covariance* between $\nu^{(B)}$ and $N_{RP}^{(r)}$ (not variance shares).

The marginal correlations show that the reallocation component is more strongly negatively related to risk premium news for CBR and STR, while the level component is more relevant for LTG. The covariance contributions show that the mechanism differs across proxies. For LTG, the aggregate relation is almost entirely a common level effect (97.86%). For STR, the relation is primarily a size-weighted reallocation effect (69.77%). CBR lies between these cases: the reallocation component has the stronger marginal correlation with risk premium news, but the covariance contribution is roughly split.

A.9.7 Additional Cross-Sectional Mechanism Checks

The main cross-sectional exercises are the FF25 portfolio-exposure design, the sorting evidence, and the level-vs.-reallocation decomposition. Additional mechanism checks – alternative sorts, large-cap-only FF25 specifications, or restricted portfolio-level belief decompositions in the small set of cells with adequate analyst coverage – are exploratory only. Given the coverage diagnostic in Appendix A.9.4, fully portfolio-level belief decompositions should be interpreted cautiously: they are useful only in cells with adequate coverage. The main cross-sectional evidence in the paper therefore uses the aggregate belief-bias news series and asks whether it is reflected across many portfolio return-news objects.

Appendix B

Appendix to Chapter 2

B.1 Proofs

B.1.1 Proposition 1

We first characterize expectations, experience, and priming effects generally. We then connect them and prove the result using Assumptions 1, 2, and 3.

Characterization of Experience and Priming Effects. Expectation $\mathbb{E}(\tau \mid \omega_T, \mathbf{n})$ obeys:

$$\begin{aligned}\mathbb{E}(\tau \mid \omega_T, \mathbf{n}) &= \sum_{\tilde{e} \in E} \mathbb{E}(\tau \mid q_{\tilde{e}, T}, \mathbf{n}) \cdot \omega_{\tilde{e}, T}, \\ \mathbb{E}(\tau \mid q_{\tilde{e}, T}, \mathbf{n}) &= \sum_{e \in E} r_e(q_{\tilde{e}, T}, \mathbf{n}) \cdot \sigma_T(e) \quad \text{for all } q_{\tilde{e}, T}.\end{aligned}$$

The experience effect of e_p is the change in belief caused by increasing the frequency of e_p from $n_{e_p} = 0$ in the baseline to $n_{e_p} = 1$, where p is f or h , not both. Let $\mathbf{n}_p$ be the database

containing e_p . Recall of different experiences changes by:

$$r_{e_p}(q_{\tilde{e},T}, \mathbf{n}_p) - r_{e_p}(q_{\tilde{e},T}, \mathbf{n}) = r_{e_p}(q_{\tilde{e},T}, \mathbf{n}_p) = \frac{S(e_p, q_{\tilde{e},T})}{S(e_p, q_{\tilde{e},T}) + \sum_{e \neq e_p} S(e, q_{\tilde{e},T}) \cdot n_e},$$

$$r_e(q_{\tilde{e},T}, \mathbf{n}_p) - r_e(q_{\tilde{e},T}, \mathbf{n}) = -r_{e_p}(q_{\tilde{e},T}, \mathbf{n}_p) \cdot r_e(q_{\tilde{e},T}, \mathbf{n}).$$

The experience effect of e_p is then equal to:

$$\begin{aligned} \mathbb{E}(\tau \mid \boldsymbol{\omega}_T, \mathbf{n}_p) - \mathbb{E}(\tau \mid \boldsymbol{\omega}_T, \mathbf{n}) &= r_{e_p}(\boldsymbol{\omega}_T, \mathbf{n}_p) \cdot \sigma_T(e_p) \\ &\quad + \sum_{e \neq e_p} \sum_{\tilde{e} \in E} [r_e(q_{\tilde{e},T}, \mathbf{n}_p) - r_e(q_{\tilde{e},T}, \mathbf{n})] \cdot \omega_{\tilde{e},T} \cdot \sigma_T(e), \\ &= r_{e_p}(\boldsymbol{\omega}_T, \mathbf{n}_p) \cdot \sigma_T(e_p) \\ &\quad - \sum_{e \neq e_p} \sum_{\tilde{e} \in E} r_{e_p}(q_{\tilde{e},T}, \mathbf{n}_p) \cdot r_e(q_{\tilde{e},T}, \mathbf{n}) \cdot \omega_{\tilde{e},T} \cdot \sigma_T(e), \\ &= r_{e_p}(\boldsymbol{\omega}_T, \mathbf{n}_p) [\sigma_T(e_p) - \mathbb{E}(\tau \mid \boldsymbol{\omega}_T, \mathbf{n})] \\ &\quad - \sum_{\tilde{e} \in E} [r_{e_p}(q_{\tilde{e},T}, \mathbf{n}_p) - r_{e_p}(\boldsymbol{\omega}_T, \mathbf{n}_p)] \cdot \mathbb{E}(\tau \mid q_{\tilde{e},T}, \mathbf{n}) \cdot \omega_{\tilde{e},T}. \end{aligned} \tag{B.1}$$

The first key term, $r_{e_p}(\boldsymbol{\omega}_T, \mathbf{n}_p) [\sigma_T(e_p) - \mathbb{E}(\tau \mid \boldsymbol{\omega}_T, \mathbf{n})]$, shows that the experience effect increases in the difference between simulation based on e_p and the baseline belief. The second term captures a covariance-like adjustment between recall of e_p and the baseline belief across cues $q_{\tilde{e},T}$; its magnitude is characterized by the analytical solution under Assumptions 1–3 below.

Another way to write (B.1), which we exploit in the text, is:

$$\mathbb{E}(\tau \mid \boldsymbol{\omega}_T, \mathbf{n}_p) - \mathbb{E}(\tau \mid \boldsymbol{\omega}_T, \mathbf{n}) = r_{e_p}(\boldsymbol{\omega}_T, \mathbf{n}_p) \left[\sigma_T(e_p) - \sum_{\tilde{e} \in E} \frac{r_{e_p}(q_{\tilde{e},T}, \mathbf{n}_p) \cdot \omega_{\tilde{e},T}}{r_{e_p}(\boldsymbol{\omega}_T, \mathbf{n}_p)} \cdot \mathbb{E}(\tau \mid q_{\tilde{e},T}, \mathbf{n}) \right],$$

where the second term in square brackets is precisely $\overline{\sigma_T(e)}_{-p}$.

Consider next the effect of priming e_p . It is equal to:

$$\mathbb{E}(\tau \mid \mathbf{1}_p, \mathbf{n}_p) - \mathbb{E}(\tau \mid \boldsymbol{\omega}_T, \mathbf{n}_p) = \sum_e \left[r_e(q_{e_p,T}, \mathbf{n}_p) - r_e(\boldsymbol{\omega}_T, \mathbf{n}_p) \right] \cdot \sigma_T(e), \quad (\text{B.2})$$

so the database is always $\mathbf{n}_p$, only the cue changes.

Linking Experience and Priming Effects. To link (B.1) to (B.2), we approximate the effect of changing the cue as a change in similarity around a neutral cue $\mathbf{q}$ that is representative of the distribution $\boldsymbol{\omega}_T$, so $c_q = c_\emptyset$ and $v_q = 0$ due to the large number of neutral experiences in the database.¹ Changing the similarity of each experience $e \in E$ from $S(e, \mathbf{q})$ to $S(e, q_{\tilde{e},T})$ affects recall of e according to:

$$\begin{aligned} r_e(q_{\tilde{e},T}, \mathbf{n}) - r_e(\boldsymbol{\omega}_T, \mathbf{n}) &\approx \frac{\partial S(e, \cdot) n_e}{\sum_{e'} S(e', \mathbf{q}) n_{e'}} - \frac{S(e, \mathbf{q}) n_e [\sum_{e'} \partial S(e', \cdot) n_{e'}]}{[\sum_{e'} S(e', \mathbf{q}) n_{e'}]^2} \\ &= \frac{\partial S(e, \cdot)}{S(e, \mathbf{q})} r_e(\boldsymbol{\omega}_T, \mathbf{n}) - r_e(\boldsymbol{\omega}_T, \mathbf{n}) \sum_{e'} \frac{\partial S(e', \cdot)}{S(e', \mathbf{q})} r_{e'}(\boldsymbol{\omega}_T, \mathbf{n}) \\ &= r_e(\boldsymbol{\omega}_T, \mathbf{n}) \left[\frac{\partial S(e, \cdot)}{S(e, \mathbf{q})} (1 - r_e(\boldsymbol{\omega}_T, \mathbf{n})) - \sum_{e' \neq e} \frac{\partial S(e', \cdot)}{S(e', \mathbf{q})} r_{e'}(\boldsymbol{\omega}_T, \mathbf{n}) \right], \end{aligned}$$

where $\partial S(e, \cdot)$ is the change in similarity and we approximate $r_e(\mathbf{q}, \mathbf{n})$ with $r_e(\boldsymbol{\omega}_T, \mathbf{n})$.

¹ Formally, the cue can be set to minimize the average discrepancy between recall under $\mathbf{q}$ and the baseline distribution $\boldsymbol{\omega}_T$, $\min_{\mathbf{q} \in E} \sum_e |r_e(\mathbf{q}, \mathbf{n}) - r_e(\boldsymbol{\omega}_T, \mathbf{n})| n_e$, or between beliefs under $\mathbf{q}$ and the baseline belief δ_T . If neutral experiences are sufficiently numerous, the cue $\mathbf{q}$ is then neutral.

Using this expression in (B.2) we obtain the priming effect:

$$\begin{aligned}\mathbb{E}(\tau | \mathbf{1}_p, \mathbf{n}_p) - \mathbb{E}(\tau | \boldsymbol{\omega}_T, \mathbf{n}_p) &\approx \sum_e r_e(\boldsymbol{\omega}_T, \mathbf{n}_p) \left[\frac{\partial S(e, \mathbf{q})}{S(e, \mathbf{q})} (1 - r_e(\boldsymbol{\omega}_T, \mathbf{n}_p)) \cdot \sigma_T(e) \right. \\ &\quad \left. - \sum_{e' \neq e} \frac{\partial S(e', \mathbf{q})}{S(e', \mathbf{q})} r_{e'}(\boldsymbol{\omega}_T, \mathbf{n}_p) \cdot \sigma_T(e) \right] \\ &= \sum_{e \in E} \frac{\partial S(e, \mathbf{q})}{S(e, \mathbf{q})} r_e(\boldsymbol{\omega}_T, \mathbf{n}_p) [\sigma_T(e) - \mathbb{E}(\tau | \boldsymbol{\omega}_T, \mathbf{n}_p)],\end{aligned}$$

where the second step follows from swapping the order of summation. The priming effect is a sum of “experience effect” terms $[\sigma_T(e) - \mathbb{E}(\tau | \boldsymbol{\omega}_T, \mathbf{n}_p)]$ for all experiences whose similarity and hence recall increases with the cue $\partial S(e, \mathbf{q}) > 0$. This includes the primed experience: cueing e_p increases the belief if $\sigma_T(e_p) - \mathbb{E}(\tau | \boldsymbol{\omega}_T, \mathbf{n}_p) > 0$. Priming e_p , however, also spills over to other experiences $e \neq e_p$, boosting simulation and beliefs provided $\sigma_T(e) - \mathbb{E}(\tau | \boldsymbol{\omega}_T, \mathbf{n}_p) > 0$. This explains why a positive experience effect implies a positive priming effect: because if the experience effect is positive, $\sigma_T(e_p) - \mathbb{E}(\tau | \boldsymbol{\omega}_T, \mathbf{n}_p) > 0$, then cueing e_p boosts beliefs both directly and through spillovers to similar, and hence also simulating, experiences.

Analytical Solution Using Assumptions 1, 2, and 3. First, some notation. We denote a generic non-neutral experience by its context $\tilde{c}$ and denote the adverse variety of the context by $\tilde{c}_-$ and the positive variety by $\tilde{c}_+$. The number of experiences with this context is $n_{\tilde{c}}$ and it is shared equally, by symmetry, into $n_{\tilde{c}}/2$ positive and $n_{\tilde{c}}/2$ negative experiences. We denote by $v_+ > 0$ the average valence of positive experiences and by v_- the one of negative ones, where symmetry again implies $v_- = -v_+$. We denote by $q_{\tilde{c}_+, T}$ and $q_{\tilde{c}_-, T}$ a positive- and negative-valence cue in the same context, whose probabilities, by symmetry, obey $\omega_{\tilde{c}_+, T} = \omega_{\tilde{c}_-, T}$. We denote by $\emptyset$ the set of neutral experiences, V_+ and V_- the sets of positive- and negative-valence

ones, where $E = \emptyset \cup V_+ \cup V_-$. N is the total number of experiences.

As a first step, we compute the average belief for specific cues $q_{\tilde{e},T}$. For neutral cues, we have:

$$\begin{aligned}\mathbb{E}(\tau \mid q_{\tilde{e},T}, \mathbf{n}) &= \delta_T + \sum_{e \in E} r_e(q_{\tilde{e},T}, \mathbf{n}) \cdot \sigma_T(c_e) \cdot v_e \\ &= \delta_T + \sum_{c \in V_+ \cup V_-} \frac{n_c}{2N} \cdot \sigma_T(c) \cdot \bar{v} = \delta_T \quad \text{for } q_{\tilde{e},T} \in \emptyset.\end{aligned}\tag{B.3}$$

The first step follows from the simulation function, the second step from all experiences being equally similar to a neutral cue, so recall is frequentist, and the last from the database, from the independence assumption between context and valence and from valence symmetry, $\bar{v} = 0$.

Consider now a non-neutral positive-valence cue, $\tilde{e} \in V_+$. Let $n_{-\emptyset}$ be the number of non-neutral experiences and $n_{-\tilde{e}}$ that of those with context different from $\tilde{e}$. Then:

$$\begin{aligned}\mathbb{E}(\tau \mid q_{\tilde{e},T}, \mathbf{n}) &= \delta_T + \sum_{e \in E} r_e(q_{\tilde{e},T}, \mathbf{n}) \cdot \sigma_T(c_e) \cdot v_e \\ &= \delta_T + \frac{K}{2} \cdot \frac{n_{-\emptyset} \cdot \overline{\sigma_T(c_{-\emptyset})}}{n_{\tilde{e}} \cdot (1 - K/2) + n_{-\tilde{e}} \cdot (1 - 3K/2) + (1 - 2K)n_{\emptyset}} v_+, \quad q_{\tilde{e},T} \in V_+.\end{aligned}\tag{B.4}$$

where $\overline{\sigma_T(c_{-\emptyset})}$ is the average (frequency-weighted) simulation sensitivity of experiences with non-neutral context, including the cued one $\tilde{e}$. For a negative cue $q_{\tilde{e},T} \in V_-$, an analogous expression holds, with the opposite sign in the second term. Averaged across cues, we obtain:

$$\mathbb{E}(\tau \mid \boldsymbol{\omega}_T, \mathbf{n}) = \delta_T + v_+ \cdot \frac{K}{2} \cdot (\omega_{+T} \cdot B_+ - \omega_{-T} \cdot B_-),$$

where ω_{+T} is the aggregate probability of positive cues, B_+ is proportional to the average

simulation coefficient across them; ω_{-T} and B_- are the corresponding quantities for negative cues. Symmetry of cues implies $\mathbb{E}(\tau \mid \omega_T, \mathbf{n}) = \delta_T$. The average belief is at the neutral default.

Experience Effect. Adding adversity $e_p \equiv e_{p-}$ (with negative valence $-v_+$) to the database $\mathbf{n}$ affects beliefs in (B.3) and (B.4) as a function of its similarity with cues $\tilde{e}$. For a neutral cue:

$$\mathbb{E}(\tau \mid q_{\tilde{e},T}, \mathbf{n}_p) = \delta_T - \frac{\sigma_T(c_p)}{N+1} v_+ \quad \text{for } q_{\tilde{e},T} \in \emptyset.$$

Adding adversity e_p breaks the symmetry of the database, moving the average belief due to e_p 's negative valence. The average effect is positive if the target outcome is also an adversity, $\sigma_T(c_p) < 0$, and negative otherwise. For a neutral cue, retrieval is frequentist and so the average effect is scaled by the new number of experiences $N+1$. For the analytical solution below, we treat the added experience e_p as a singleton context absent from the baseline database, so $n_{c_p} = 0$ and hence $n_{-c_p} = n_{-\emptyset}$. For a non-neutral cue:

$$\mathbb{E}(\tau \mid q_{\tilde{e},T}, \mathbf{n}_p) = \delta_T + \begin{cases} -\frac{\sigma_T(c_p) + n_{-\emptyset} \cdot (K/2) \cdot \overline{\sigma_T(c_{-\emptyset})}}{1 + (1 - 3K/2) \cdot n_{-\emptyset} + (1 - 2K) \cdot n_{\emptyset}} v_+ & \text{for } \tilde{e} = e_{p-} \\ \frac{-(1 - K) \cdot \sigma_T(c_p) + n_{-\emptyset} \cdot (K/2) \cdot \overline{\sigma_T(c_{-\emptyset})}}{(1 - K) + (1 - 3K/2) \cdot n_{-\emptyset} + (1 - 2K) \cdot n_{\emptyset}} v_+ & \text{for } \tilde{e} = e_{p+} \\ -\frac{(1 - K) \cdot \sigma_T(c_p) + n_{-\emptyset} \cdot (K/2) \cdot \overline{\sigma_T(c_{-\emptyset})}}{(1 - K) + (1 - K/2) \cdot n_{\tilde{e}} + (1 - 3K/2)(n_{-\tilde{e}} - 1) + (1 - 2K)n_{\emptyset}} v_+ & \text{for } \tilde{e} \in V_- \setminus e_{p-} \\ \frac{-(1 - 2K)\sigma_T(c_p) + n_{-\emptyset} \cdot (K/2) \cdot \overline{\sigma_T(c_{-\emptyset})}}{(1 - 2K) + (1 - K/2) \cdot n_{\tilde{e}} + (1 - 3K/2)(n_{-\tilde{e}} - 1) + (1 - 2K)n_{\emptyset}} v_+ & \text{for } \tilde{e} \in V_+ \setminus e_{p+}. \end{cases} \quad (\text{B.5})$$

When adversity e_p is added to the database, similarity to any cue moves in the same direction, determined by the sign of $\sigma_T(c_p)$. Averaging over cues, beliefs are:

$$\mathbb{E}(\tau \mid \omega_T, \mathbf{n}_p) = \delta_T - Z_{e_p} \cdot \sigma_T(c_{e_p}) \cdot v_+ + W_{-\emptyset} \cdot \overline{\sigma_T(c_{-\emptyset})} \cdot v_+, \quad (\text{B.6})$$

where both Z_{e_p} and $W_{-\emptyset}$ are positive, the latter because adding e_{p-} to the database creates recall interference on other non-neutral experiences. Such interference is stronger for negative-valence ones. Formally, the denominators on the first and third rows are larger than those on the second and fourth rows, respectively. Given that positive and negative cues have the same probability, the effect across cues of non-neutral experiences different from e_p is skewed toward positive valence, which, starting from a zero baseline, implies $W_{-\emptyset} > 0$.

Consider now experience effects. Because our target outcomes T , inflation and home prices, are adversities, we have $\sigma_T(c) \leq 0$ for all contexts, including $\sigma_T(c_p) \leq 0$. A positive experience effect occurs if $\mathbb{E}(\tau \mid \omega_T, \mathbf{n}_p) - \mathbb{E}(\tau \mid \omega_T, \mathbf{n}) > 0$, namely (B.6) should be larger than (B.3), which corresponds to:

$$\sigma_T(c_p) < \frac{W_{-\emptyset} \cdot \overline{\sigma_T(c_{-\emptyset})}}{Z_{e_p}}. \quad (\text{B.7})$$

The right-hand side of (B.7) is strictly negative because $c_{-\emptyset}$ contains DS experiences with strong simulation potential. Thus, a positive experience effect reveals $\sigma_T(c_{e_p}) \ll 0$. Adversity e_p must be close enough in context to the target T that it comes to mind spontaneously and boosts simulation to an extent that offsets the interference it creates for other adversities.

Priming Effects. A positive priming effect occurs when

$$\mathbb{E}(\tau \mid q_{e_{p-}, T}, \mathbf{n}_p) > \mathbb{E}(\tau \mid \omega_T, \mathbf{n}_p),$$

which using (B.5) and (B.6) is equivalent to:

$$\begin{aligned} \frac{\sigma_T(c_p)}{1 + (1 - 3K/2) \cdot n_{-\emptyset} + (1 - 2K) \cdot n_{\emptyset}} + \frac{n_{-\emptyset} \cdot (K/2) \cdot \overline{\sigma_T(c_{-\emptyset})}}{1 + (1 - 3K/2) \cdot n_{-\emptyset} + (1 - 2K) \cdot n_{\emptyset}} \\ < Z_{e_p} \cdot \sigma_T(c_p) \cdot v_+ - W_{-\emptyset} \cdot \overline{\sigma_T(c_{-\emptyset})} \end{aligned}$$

which reads:

$$\begin{aligned} \left[\frac{1}{1 + (1 - 3K/2) \cdot n_{-\emptyset} + (1 - 2K) \cdot n_{\emptyset}} - Z_{e_p} \right] \cdot \sigma_T(c_p) \\ + \left[\frac{n_{-\emptyset} \cdot (K/2)}{1 + (1 - 3K/2) \cdot n_{-\emptyset} + (1 - 2K) \cdot n_{\emptyset}} + W_{-\emptyset} \right] \cdot \overline{\sigma_T(c_{-\emptyset})} < 0. \end{aligned}$$

Given that the second term is negative (the term in brackets is positive due to $W_{-\emptyset} > 0$), the condition is surely met if the first term in square brackets is also positive, because a positive experience effect implies $\sigma_T(c_p) < 0$ via (B.7).

The first term in square brackets is positive when $r_{e_{p-}}(q_{e_{p-}, T}, \mathbf{n}_p) > r_{e_{p-}}(\omega_T, \mathbf{n}_p) = Z_{e_p}$. This follows directly when $n_{\emptyset}$ is sufficiently large: the denominator of recall is dominated by neutral experiences and varies little across cues, while $S(e_{p-}, q_{e_{p-}, T})$ is maximal. Hence cueing e_{p-} raises recall of e_{p-} relative to its average recall under ω_T .

B.1.2 Prediction 1

We can decompose the effect of change in the database from $\mathbf{n}$ to $\mathbf{n}_p$ and of the cue from ω_T to $\mathbf{1}_p$ as the sum of an experience and priming effect:

$$\begin{aligned} \mathbb{E}(\tau \mid \omega_T + \Delta\omega, \mathbf{n} + \Delta\mathbf{n}) &\approx \mathbb{E}(\tau \mid \omega_T, \mathbf{n}) + \sum_{p=h,f} [\mathbb{E}(\tau \mid \omega_T, \mathbf{n}_p) - \mathbb{E}(\tau \mid \omega_T, \mathbf{n})] \cdot \mathbb{1}_{e_p} \\ &+ \sum_{p=h,f} [\mathbb{E}(\tau \mid \mathbf{1}_p, \mathbf{n}_p) - \mathbb{E}(\tau \mid \omega_T, \mathbf{n}_p)] \cdot \mathbb{1}_{e_p} \cdot \mathbb{1}_{e_p, R} \end{aligned}$$

where $\mathbb{1}_{e_p}$ is an indicator equal to 1 if the person has had experience e_p and zero otherwise, and $\mathbb{1}_{e_p, R}$ is an indicator equal to 1 if the person had the priming intervention concentrating the cue in experience e_p . The intervention is said to occur (treatment on treated) if the subject is in the treatment group, $\mathbb{1}_{e_p, R} = 1$, and if she had the primed experience, $\mathbb{1}_{e_p} = 1$.

We characterized experience and priming effects in Proposition 1, which yields property (i), namely $\varepsilon_{e_p} > 0 \Rightarrow \beta_{e_p, R} > 0$. For property (ii), note that the sum of experience and priming effects equals $\mathbb{E}(\tau \mid \mathbf{1}_p, \mathbf{n}_p) - \mathbb{E}(\tau \mid \omega_T, \mathbf{n})$. The baseline term $\mathbb{E}(\tau \mid \omega_T, \mathbf{n})$ does not depend on $\sigma_T(e_p)$. Under our assumptions, higher similarity between e_p and the target outcome T affects $\mathbb{E}(\tau \mid \mathbf{1}_p, \mathbf{n}_p)$ only through a higher simulation function $\sigma_T(e_p)$, giving:

$$\frac{\partial \mathbb{E}(\tau \mid q_{e_p, T}, \mathbf{n}_p)}{\partial \sigma_T(e_p)} = r_{e_p}(q_{e_p, T}, \mathbf{n}_p) > 0.$$

B.1.3 Prediction 2

We again consider beliefs $\mathbb{E}(\tau \mid \omega_T, \mathbf{n}) = \sum_{\tilde{e} \in E} \mathbb{E}(\tau \mid q_{\tilde{e}, T}, \mathbf{n}) \cdot \omega_{\tilde{e}, T}$ by linearizing them as a function of the contextual similarity between each experience e and the target outcome T . We write $S(e, T)$ as shorthand for this contextual similarity between experience e and the target state in task T , holding valence fixed. Under our assumptions, a higher value of such similarity affects beliefs through simulation, in particular:

$$\partial \mathbb{E}(\tau \mid \omega_T, \mathbf{n}_p) = r_e(\omega_T, \mathbf{n}_p) \cdot v_e \cdot \frac{\partial \sigma_T(c_e)}{\partial S(e, T)} \cdot \partial S(e, T),$$

where $\partial \sigma_T(c_e) / \partial S(e, T) < 0$ when the target is an adversity. Higher similarity therefore raises beliefs for negative-valence experiences ($v_e < 0$, producing more pessimism) and lowers them for positive-valence ones ($v_e > 0$). This implies that we can approximate beliefs as a function of similarities as follows:

$$\mathbb{E}(\tau \mid \omega_T, \mathbf{n}_p) \approx \mathbb{E}_0(\tau \mid \omega_T, \mathbf{n}_p) - A \cdot \sum_e r_e(\omega_T, \mathbf{n}_p) \cdot v_e \cdot \partial S(e, T),$$

where $\mathbb{E}_0(\tau \mid \omega_T, \mathbf{n}_p)$ is the belief when all experiences have minimum similarity (i.e., the cue is neutral) and $A = - \left. \frac{\partial \sigma_T(c_e)}{\partial S(e, T)} \right|_{S(e, T)=S_0} > 0$ is the magnitude of the slope of simulation at such minimum similarity point. The belief is thus an average similarity of experiences weighted by their valence. Focusing on the adversities we measure, and assuming for simplicity that these adversities all have valence equal to $-v_+$, we have:

$$\mathbb{E}(\tau \mid \omega_T, \mathbf{n}_p) \approx \mathbb{E}_0(\tau \mid \omega_T, \mathbf{n}_p) - A \cdot P + Av_+ \sum_e r_e(\omega_T, \mathbf{n}_p) \cdot \partial S(e, T),$$

where $-A \cdot P < 0$ is the effect of unobserved positive experiences. In our regression, the average similarity of our measured adversities proxies for the right-most term, $\sum_e r_e(\omega_T, \mathbf{n}_p) \cdot \partial S(e, T)$. This is exact under equal recall weights across measured adversities; otherwise, it should be interpreted as a reduced-form proxy for the recall-weighted theoretical object. The same logic gives the similarity-based approximation for the priming effect.

B.2 The Survey Instruments

B.2.1 The DHS: Index Construction and Results

Health Adversity. We construct the health index from the following questions:

- How is your health in general?
- How many times did you contact your general practitioner about your own health in the past year?

Financial Adversity. This index measures present perceived strain in household finances. It builds on the following questions:

- Would you, at present, like to spend more money than you have (e.g., through income)?
In other words, would you like to have more money now, that you would have to pay back later?
- How well can you manage on the total income of your household?
- Is this income unusually high or low compared to the income you would expect in a ‘regular’ year, or is it regular?
- How is the financial situation of your household at the moment?
- Over the past 12 months, would you say the expenditures of your household were higher than the income of the household, about equal to the income of the household, or lower than the income of the household?

- Do you think the expenditures of your household, in the next 12 months, will be higher, about the same, or lower than the income of your household?
- The following statements concern your own situation compared to that of others.
 - I think my household has more assets than others in my environment.
 - If I compare myself with my friends, I think in general I am financially better off.
 - Most people in my environment are saving money.
 - I can spend more on durable consumer goods than others in my environment.
 - If necessary, I can reduce the expenditures of my household by 5% without a problem.
- Have you, in the past 12 months, consulted with your bank or financial institution, because you had or expected payment problems with the repayment of a loan or mortgage?
- In the past 2 years, has a request you (or your partner) made for credit been turned down, or were you not given as much credit as you applied for?
- How do you think the economic situation of your household will be in five years' time in comparison to the current situation?
- Do you find it easy or difficult to control your expenditures?
- What is the chance that you will leave an inheritance (including possessions and valuable items)?

Financial Education. This index captures respondents' financial knowledge and their experiences with financial education growing up.

- How knowledgeable do you consider yourself with respect to financial matters?
- When you were between 8 and 12 years of age, did you receive an allowance from your parents then?
- When you were between 8 and 12 years of age, did you do little household chores (like washing the car) for which you received some money from your parents?
- When you were between 8 and 12 years of age, could you spend your money as you pleased?
- Did you have a job on the side (like a newspaper round, a job on Saturday etcetera) when you were between 12 and 16 years of age?
- Did your (grand)parents try to teach you how to budget when you were between 12 and 16 years of age?
- Did your (grand)parents stimulate you to save money between the age of 12 and 16?

Happiness / Psychological Well-Being. We construct this index from the following questions:

- All in all, to what extent do you consider yourself a happy person?
- I have frequent mood swings.
- I get stressed out easily.
- The following questions are about how you felt over the past month. For every question, please choose the answer that best describes how you felt during this past month. This past month...

- I felt very anxious.
- I felt so down that nothing could cheer me up.
- I felt calm and peaceful.
- I felt depressed and gloomy.
- I felt happy.

Attitudes Toward Risk. This index relies on the following survey questions:

- The following statements concern saving and taking risks. Please indicate for each statement to what extent you agree or disagree.
 - I do not invest in shares, because I find this too risky.
 - I want to be certain that my investments are safe.
- Are you generally a person who is willing to take risks or do you try to avoid taking risks?

The following questions were initially also considered to be included in the construction of the index but were not considered due to very low or negative correlation with the included measures above.

- The following statements concern saving and taking risks. Please indicate for each statement to what extent you agree or disagree.
 - I think it is more important to have safe investments and guaranteed returns, than to take a risk to have a chance to get the highest possible returns.

- If I think an investment will be profitable, I am prepared to borrow money to make this investment.
- If I want to improve my financial position, I should take financial risks.
- I am prepared to take the risk to lose money, when there is also a chance to gain money.
- How would you describe the risks that you have taken with investments over the past few years?

Other Attitudes: Fatalism vs. Effort. Here we use the following questions:

- Do you agree with the following statements?
 - Whether or not I get to become wealthy depends mostly on my ability.
 - In the long run, people who take very good care of their finances stay wealthy.
 - When I get what I want, it's usually because I worked hard for it.
 - My life is determined by my own actions.

We do not incorporate the following questions:

- Do you agree with the following statements?
 - It is chiefly a matter of fate whether I become rich or poor.
 - Only those who inherit or win money can possibly become rich.

Other Attitudes: Optimism. The optimism index is constructed from the following questions:

- How likely is it that you will attain at least the age of x ?

- On the whole, I expect more good things to happen to me than bad things.

All variables are coded in a way that higher values indicate worse outcomes. Leading up to Table 2.1 in the text, we first report the correlation between the indices above in Table B.1.

Table B.1: Correlation Structure of Indices in the DHS

	Health Adversity	Financial Adversity	Happiness	Risk Attitudes	Effort/ Agency	Optimism	Financial Education
Health Adversity	1.00						
Financial Adversity	0.26	1.00					
Happiness	0.33	0.24	1.00				
Risk Attitudes	0.03	0.03	−0.01	1.00			
Effort/Agency	0.11	0.24	0.09	0.01	1.00		
Optimism	0.35	0.23	0.27	0.06	0.17	1.00	
Financial Education	−0.12	−0.22	−0.05	−0.12	−0.15	−0.23	1.00

Notes: This table reports pairwise correlations between the seven DHS indices used in Table B.2. $N = 21,614$ individual-year observations from the DNB Household Survey (2012–2022). All indices are standardized composites constructed as the within-year average of the standardized underlying items. For all indices except Financial Education, higher values indicate worse outcomes (worse health, worse financial situation, lower happiness, greater risk aversion, lower effort/agency, and lower optimism). Higher Financial Education indicates greater childhood financial exposure. See Section 2.2.2 for construction details.

Table B.2 next reports regressions of macroeconomic beliefs on the Table 2.1 controls and the full set of indices, showing that the finance and health indices are the experiences that best predict inflation and home price growth beliefs.

Table B.3 next presents Table 2.1 with all covariates.

Finally, we present three robustness exercises for Table 2.1. First, beliefs about home price growth and the role of experiences in it may depend systematically on whether participants are owners or renters (in particular, a rise in home prices benefits owners but can hurt renters). Table B.4 presents the same specification restricted to either owners (Panel A) or renters (Panel B). Second, the construction of the financial trouble index raises two possible objections: (i)

Table B.2: Adversity Indices and Macroeconomic Beliefs in the DHS

	Inflation			HPG			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Health Adversity		0.05*** (0.02)	0.04* (0.02)		0.05*** (0.01)	0.04*** (0.01)	0.04*** (0.01)
Financial Adversity		0.28*** (0.03)	0.23*** (0.03)		0.07*** (0.02)	0.02 (0.02)	0.03 (0.02)
Happiness		0.01 (0.02)	0.02 (0.02)		−0.01 (0.02)	−0.03* (0.02)	−0.03* (0.02)
Risk Attitudes		0.04*** (0.01)	0.03** (0.01)		−0.01 (0.01)	−0.01 (0.01)	−0.01 (0.01)
Effort/Agency		−0.01 (0.01)	−0.02 (0.01)		−0.02 (0.01)	−0.01 (0.01)	−0.00 (0.01)
Optimism		0.07*** (0.01)	0.05*** (0.01)		−0.03** (0.01)	−0.02* (0.01)	−0.02 (0.01)
Financial Education		−0.03 (0.02)	−0.00 (0.02)		−0.03* (0.02)	−0.04** (0.02)	−0.05*** (0.02)
Belief Price Increase Own House (Past Two Years)							0.01*** (0.00)
Constant	0.32*** (0.10)	−0.06** (0.03)	0.11 (0.10)	0.41*** (0.12)	−0.01 (0.03)	0.40*** (0.10)	0.44*** (0.10)
<i>N</i>	19,852	21,566	19,501	16,504	17,547	16,228	16,228
Controls	Yes	No	Yes	Yes	No	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
<i>R</i> ²	0.03	0.04	0.05	0.02	0.00	0.02	0.04

Notes: This table reports OLS regressions of macroeconomic belief *z*-scores on DHS adversity indices. The dependent variable in columns (1)–(3) is the *z*-score of the respondent’s most likely 1-year inflation rate, standardized within survey year. The dependent variable in columns (4)–(7) is the *z*-score of expected 2-year house price growth, standardized within survey year. All indices are standardized composites; higher values indicate worse outcomes except for Financial Education. Controls include age, gender, education, occupation, urbanization, household income, housing status, numeracy, and relationship status. All specifications include year fixed effects. Column (7) additionally controls for the respondent’s belief about own house price change in the past two years. Data from the DNB Household Survey, 2012–2022. Standard errors clustered at the individual level in parentheses. Significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

the expectations of income expenditures component may be mechanically correlated with expected inflation, and (ii) the “relative position” component may be correlated with views about the aggregate state of the economy. To address these concerns, Panel C reports the main specification for the index constructed without these questions. Third, expectations of aggregate home price growth may depend on beliefs about recent experienced growth in local

Table B.3: Finance and Health Indices and Macroeconomic Beliefs in the DHS

	Inflation		HPG	
	(1)	(2)	(3)	(4)
Health Adversity	0.08*** (0.02)	0.06*** (0.02)	0.04*** (0.01)	0.03** (0.01)
Financial Adversity	0.30*** (0.02)	0.24*** (0.03)	0.06*** (0.02)	0.02 (0.02)
Age		-0.00 (0.00)		-0.00** (0.00)
Male		0.00 (0.02)		0.02 (0.02)
Education: Vocational		-0.08*** (0.03)		-0.01 (0.02)
Education: University		-0.19*** (0.03)		-0.05 (0.03)
Occupation: Employed		0.05 (0.03)		-0.01 (0.03)
Occupation: Self-Employed		0.03 (0.05)		0.02 (0.04)
Occupation: Student		-0.19*** (0.07)		0.09 (0.09)
Occupation: Retired		0.14*** (0.04)		0.01 (0.03)
Urbanization		0.01 (0.01)		0.05*** (0.01)
Income: EUR 10,000 – EUR 14,000		0.09 (0.07)		0.05 (0.09)
Income: EUR 14,000 – EUR 22,000		0.06 (0.07)		0.05 (0.08)
Income: EUR 22,000 – EUR 40,000		0.05 (0.07)		0.07 (0.08)
Income: EUR 40,000 – EUR 75,000		0.02 (0.07)		0.06 (0.08)
Income: EUR 75,000+		0.03 (0.07)		0.15* (0.09)
Housing: Subtenant		0.04 (0.14)		-0.04 (0.16)
Housing: Owner		-0.09*** (0.03)		-0.23*** (0.03)
Housing: Other		-0.22** (0.11)		-0.62*** (0.10)
Numeracy		-0.06*** (0.01)		-0.01 (0.01)
Has Partner		0.03 (0.03)		0.02 (0.02)
<i>N</i>	22,081	19,852	17,928	16,504
Year FE	Yes	Yes	Yes	Yes
<i>R</i> ²	0.03	0.05	0.00	0.02

Notes: This table replicates Table 2.1. Dependent variables are *z*-scored inflation (columns (1)–(2)) and home price growth (columns (3)–(4)) expectations. Education base: primary/secondary; Occupation base: not employed; Income base: lowest bracket; Housing base: renter. Standard errors clustered at the individual level in parentheses. All specifications include year fixed effects. Significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

home prices, which may be correlated with adversities. We thus add to Panel C a specification where we control for believed home price appreciation over the previous two years.

Panels A and B show that experience effects on home price growth expectations are driven mostly by renters. This is consistent with the idea that, for owners, forecasts of home price appreciation are not spontaneously (absent priming) perceived as adversities but instead load more on their experienced home price growth. Panel C shows that results are robust to using an alternative, more conservative index for financial adversity. Moreover, while beliefs about experienced local price growth shape forecasts (and significantly increase explanatory power), the predictive power of adversities is unchanged.

Finally, Table B.5 shows that the finance and health indices are uncorrelated with subsequent realizations of inflation or home price growth.

Table B.6 shows that positive health or finance NDS experiences predict lower inflation and home price growth expectations. We construct a health index based on the average of z -scores of (i) health improvement vs. last year (reversed), (ii) felt calm and peaceful past month, (iii) felt happy past month, and (iv) overall happiness (reversed). Higher values consistently indicate more positive states, with no mechanical correlation with the adversity index. For the financial index, we use (i) saved money (binary), (ii) job satisfaction (reversed), (iii) has dividend income (binary), and (iv) received inheritance (binary). Since job satisfaction is only asked of employed respondents, we also construct a version excluding it for larger samples.

The clean positive finance index is strongly significant for inflation expectations: a 1-standard-deviation increase in positive financial experiences is associated with a 0.091 SD decrease in inflation expectations ($p < 0.01$). The positive health index becomes significant with controls (-0.050 , $p < 0.05$). Home price growth results are weaker, mirroring the pattern in

Table B.4: Adversity Indices and Macroeconomic Beliefs by Housing Tenure

<i>Panel A: Owners</i>					
	Inflation		HPG		
	(1)	(2)	(3)	(4)	
Health Adversity	0.10*** (0.02)	0.08*** (0.02)	0.01 (0.01)	0.02 (0.02)	
Financial Adversity	0.30*** (0.03)	0.24*** (0.03)	−0.03 (0.02)	−0.01 (0.03)	
<i>N</i>	15,728	14,218	13,459	12,398	
Controls	No	Yes	No	Yes	
Year FE	Yes	Yes	Yes	Yes	
<i>R</i> ²	0.03	0.04	0.00	0.01	
<i>Panel B: Renters</i>					
	Inflation		HPG		
	(1)	(2)	(3)	(4)	
Health Adversity	0.05 (0.03)	0.04 (0.04)	0.04 (0.02)	0.05* (0.03)	
Financial Adversity	0.28*** (0.05)	0.29*** (0.06)	0.04 (0.05)	0.09 (0.06)	
<i>N</i>	6,069	5,397	4,378	4,028	
Controls	No	Yes	No	Yes	
Year FE	Yes	Yes	Yes	Yes	
<i>R</i> ²	0.02	0.05	0.01	0.02	
<i>Panel C: Alternative Index of Financial Adversity</i>					
	Inflation		HPG		
	(1)	(2)	(3)	(4)	(5)
Health Adversity	0.10*** (0.02)	0.06*** (0.02)	0.04*** (0.01)	0.03** (0.01)	0.03** (0.01)
Financial Adversity (Alt.)	0.20*** (0.02)	0.16*** (0.02)	0.06*** (0.02)	0.03 (0.02)	0.03 (0.02)
<i>N</i>	22,047	19,840	17,900	16,493	16,493
Controls	No	Yes	No	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes
<i>R</i> ²	0.02	0.04	0.00	0.02	0.04

Notes: This table reports robustness checks for Table 2.1. *Panels A and B* replicate Table 2.1 separately for homeowners and renters; housing tenure is excluded from controls. Panel C uses a more conservative Financial Adversity built from 7 components (excluding 9 items related to relative position and income-expenditure expectations) instead of the 16-component index. Dependent variables are *z*-scored inflation and home price growth (HPG) expectations. Controls include age, gender, education, occupation, urbanization, household income, housing status, numeracy, and relationship status (housing status is excluded in Panels A and B since the sample is split by housing tenure). Standard errors clustered at the individual level in parentheses. All specifications include year fixed effects. Significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table B.5: Adversity Indices and Realized Macroeconomic Outcomes

	Inflation				HPG			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Health Adversity	−0.04 (0.03)	−0.02 (0.03)	−0.00 (0.00)	0.00 (0.00)	0.50*** (0.10)	0.49*** (0.10)	−0.01 (0.03)	−0.01 (0.03)
Financial Adversity	0.35*** (0.04)	0.47*** (0.05)	−0.00 (0.01)	0.00 (0.01)	−1.60*** (0.18)	−1.76*** (0.20)	0.04 (0.04)	−0.00 (0.05)
<i>N</i>	22,057	19,836	22,057	19,836	19,781	18,019	19,781	18,019
Controls	No	Yes	No	Yes	No	Yes	No	Yes
Year FE	No	No	Yes	Yes	No	No	Yes	Yes
Province FE	No	No	No	No	Yes	Yes	Yes	Yes
<i>R</i> ²	0.00	0.03	0.99	0.99	0.02	0.07	0.95	0.95

Notes: This table tests whether individual-level adversity indices predict actual future macroeconomic outcomes. The dependent variable in columns (1)–(4) is realized CPI inflation 12 months ahead. The dependent variable in columns (5)–(8) is realized home price growth (HPG) over a 24-month window ahead. The Health Adversity index is a standardized composite of individual health status variables; the Financial Adversity index is a standardized composite of 16 financial well-being measures. Both are measured at the individual level in the DNB Household Survey (DHS). Controls include age, gender, education, occupation, urbanization, household income, housing status, numeracy, and relationship status. All specifications include year fixed effects. Province fixed effects are additionally included in the housing specifications (columns (5)–(8)) to account for local home price variation; inflation is a national measure and does not vary by province. Standard errors clustered at the individual level in parentheses. Significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

the main paper where financial adversity dominates due to similarity. These results are robust to controlling for Health and Finance Adversities, as well as for general measures of optimism and happiness.

We complement this analysis in Table B.7 by looking at specific discrete positive shocks. Inheritance receipt and dividend income are the strongest individual predictors. Receiving an inheritance is associated with 0.077 SD lower inflation expectations ($p < 0.05$) and 0.054 SD lower home price growth expectations ($p < 0.10$). Having dividend income predicts 0.128 SD lower inflation expectations ($p < 0.01$). These are discrete positive financial events. The grandchild effect goes in the predicted direction but lacks statistical power (only 248 events).

Table B.6: Positive Indices and Macroeconomic Beliefs

	Inflation		HPG	
	(1)	(2)	(3)	(4)
Positive Health (Clean)	−0.02 (0.02)	−0.05** (0.02)	0.00 (0.02)	0.02 (0.02)
Positive Finance (Clean)	−0.17*** (0.01)	−0.09*** (0.01)	−0.03** (0.01)	−0.01 (0.01)
<i>N</i>	22,081	19,852	17,928	16,504
Controls	No	Yes	No	Yes
Year FE	Yes	Yes	Yes	Yes
<i>R</i> ²	0.01	0.04	0.00	0.02

Notes: This table reports OLS regressions of macroeconomic belief *z*-scores on standardized positive experience indices. The dependent variable in columns (1)–(2) is the *z*-score of the respondent’s most likely 1-year inflation rate, standardized within survey year. The dependent variable in columns (3)–(4) is the *z*-score of expected 2-year house price growth, standardized within survey year. “Positive Health (Clean)” is the average of the *z*-scores of (i) health improvement vs. last year (reversed), (ii) felt calm and peaceful past month, (iii) felt happy past month, and (iv) overall happiness (reversed); higher values indicate more positive states. “Positive Finance (Clean)” is the average of the *z*-scores of (i) saved money (binary), (ii) job satisfaction (reversed), (iii) has dividend income (binary), and (iv) received inheritance (binary). Both indices are constructed to have no mechanical correlation with the adversity indices in Table 2.1. Controls include age, gender, education, occupation, urbanization, household income, housing status, numeracy, and relationship status. All specifications include year fixed effects. Standard errors clustered at the individual level in parentheses. Data from the DNB Household Survey, 2012–2022. Significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table B.7: Positive Personal Experiences and Macroeconomic Beliefs

	Inflation			HPG		
	(1)	(2)	(3)	(4)	(5)	(6)
Received Inheritance	−0.08** (0.03)			−0.05* (0.03)		
Dividend Income		−0.13*** (0.03)			0.00 (0.03)	
New Grandchild			−0.02 (0.06)			−0.08 (0.06)
<i>N</i>	19,852	19,665	11,689	16,504	16,348	9,844
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
<i>R</i> ²	0.03	0.03	0.04	0.02	0.02	0.02
Sample			Age ≥ 45			Age ≥ 45

Notes: This table reports OLS regressions of macroeconomic belief *z*-scores on indicators for discrete positive personal experiences. The dependent variable in columns (1)–(3) is the *z*-score of the respondent’s most likely 1-year inflation rate, standardized within survey year. The dependent variable in columns (4)–(6) is the *z*-score of expected 2-year house price growth, standardized within survey year. “Received Inheritance” is a binary indicator for inheritance receipt; “Dividend Income” is a binary indicator for any dividend or investment income; “New Grandchild” is a binary indicator for the arrival of a new grandchild and is restricted to respondents aged 45 and above. Controls include age, gender, education, occupation, urbanization, household income, housing status, numeracy, and relationship status. All specifications include year fixed effects. Standard errors clustered at the individual level in parentheses. Data from the DNB Household Survey, 2012–2022. Significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

B.2.2 Our Survey Instrument

Introductory Text.

The goal of this survey is to explore how people form beliefs about inflation and changes in home prices, based on economic events and personal experiences. To do that, the survey includes questions about personal events you may have experienced and your views about inflation and changes in home prices.

Q1) We want to start with a question on estimating quantities. Out of 1000 random people in The Netherlands, how many do you think are older than 90 years?

Experience Prompts. *PROGRAMMING NOTE. Use a random generator to divide respondents into three subgroups A, B, and C. Group A starts with questions Q1a and Q2a, and Group B starts with questions Q1b and Q2b. Group C is the control group and starts the survey with question Q5.*

[Group A: Health]

Q1a) In this part of the survey, we will ask you about an experience you may have had. We appreciate you taking the time to provide answers to these personal and possibly sensitive questions.

Have you or a loved one ever faced the challenges of a serious or chronic illness, or the consequences of a serious accident?

- ☐ Yes [$\rightarrow$ Proceed with question Q2a]
- ☐ No [$\rightarrow$ Jump to question Q5]

Q2a) Thinking back on this experience, please respond to the questions below.

1. Briefly describe what the illness or accident was, who suffered from it, and in what year/years it happened.
2. What losses, difficulties, or sacrifices do you remember from that time?
3. What emotions do you associate with thinking back on that experience?

[Group B: Finance]

Q1b) In this part of the survey, we will ask you about an experience you may have had. We appreciate you taking the time to provide answers to these personal and possibly sensitive questions.

Have you or a loved one ever struggled with financial difficulties?

- ☐ Yes [*→ Proceed with question Q2b*]
- ☐ No [*→ Jump to question Q5*]

Q2b) Thinking back on this experience, please respond to the questions below.

1. Briefly describe what the main cause of that financial struggle was, who suffered from it, and in what year/years it happened.
2. What losses, difficulties, or sacrifices do you remember from that time?
3. What emotions do you associate with thinking back on that experience?

[For Subgroups A and B who answered yes to Q1a/Q1b]

Q3) Can you tell us more about this personal experience in 3–4 sentences? Feel free to write whatever comes to mind as you think about this experience.

[For all respondents]

Belief Elicitation – Inflation. *PROGRAMMING NOTE. Q5: Group A and Group B respondents will be shown an additional text screen before Q5 with the following text: “At this point, you will be asked about economic expectations. We understand that this is a completely different topic than the previous questions, but in the interest of the research study, we would like to ask you to answer them.”*

Q5) We are interested in your views on inflation in the Netherlands in the future. Inflation is the change in consumer prices in general over a period of twelve months. What do you expect inflation in the Netherlands to be in the next twelve months, i.e., what will the percentage change in consumer prices be over the next twelve months?

If you think prices will go up, please enter a positive percentage. If you think prices will fall, please enter a negative percentage (put a minus sign (-) in front of your answer). If you think prices will not change, please enter 0 (zero). You can also use decimal places if you want.

Q6) How likely do you think inflation is to be 7% or higher a year from now? You can answer on a scale of 0 to 10, where 0 indicates that there is ‘no chance at all’ that inflation will be 7% or higher, and 10 indicates that inflation will ‘absolutely, for sure’ be 7% or higher.

0	1	2	3	4	5	6	7	8	9	10
☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐

0 = no chance at all; 10 = absolutely, for sure

Q8) Please write down 3–4 sentences about any thoughts or personal experiences that came to mind when you were answering the previous questions on inflation.

Belief Elicitation – House Prices.

Q9) We are interested in your views on home prices in the Netherlands in the future. What do you expect to be the percentage change in home prices in the Netherlands in the next twelve months?

If you think home prices will rise, please enter a positive percentage. If you think home prices will fall, please enter a negative percentage (put a minus sign (-) in front of your answer). If you think home prices will not change, please enter 0 (zero). You can also use decimal places if you want.

Q10) How likely do you think it is that home prices will rise by 10% or more over the next twelve months? You can answer on a scale of 0 to 10, where 0 indicates that there is ‘no chance at all’ that home prices will rise by 10% or more, and 10 indicates that home prices will ‘absolutely, for sure’ rise by 10% or more.

0	1	2	3	4	5	6	7	8	9	10
☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐

0 = no chance at all; 10 = absolutely, for sure

Q12) Please write down 3–4 sentences about any thoughts or personal experiences that came to mind when you were answering the previous questions on home prices.

PROGRAMMING NOTE. Q13: Please randomize the order of the first two response options. The last two response options should always be placed at the end.

Q13) Do you consider a rise in home prices by 10% or more to be a good or a bad thing?

- ☐ A good thing
- ☐ A bad thing
- ☐ Neither
- ☐ I do not know

Experiences. *PROGRAMMING NOTE. Q14: Respondents from group A (Health Experience Prompt) are asked the second question. Respondents from Group B (Finance Experience Prompt) are asked the first question. Respondents from group C (Control group) are asked both questions. The other questions are asked to everyone. The order of all questions should be randomized, keeping in mind that the first two questions (if offered to the respondent) should always come first.*

Q14) For each of the experiences below, please indicate if you or a loved one had this past experience.

Experience	Yes	No
[Groups B and C] Have you or a loved one ever faced the challenges of a serious or chronic illness, or the consequences of a serious accident?	☐	☐
[Groups A and C] Have you or a loved one ever struggled with financial difficulties?	☐	☐
Have you or a loved one experienced a recent extreme weather event in your area, such as a very heavy storm, flood, or extreme heat?	☐	☐
Have you or a loved one had an experience, for instance in your job, where an event increased inequality or was perceived as unfair?	☐	☐
Have you or a loved one ever experienced protests, confrontations, or other forms of intense tension between social groups?	☐	☐
Have you or a loved one ever experienced someone cheating, stealing, or breaking a contract?	☐	☐
Have you or a loved one ever made a long-term investment in a business or in the stock market that has been successful in the long run?	☐	☐
Have you or a loved one ever faced challenges related to the immigration of foreigners?	☐	☐
Have you or a loved one ever lost a job unexpectedly and experienced unemployment?	☐	☐
Have you or a loved one ever experienced a divorce or separation from a long-term partner?	☐	☐

Inflation Similarity. *PROGRAMMING NOTE. Q15: For every respondent use the same order of questions as shown in Q14.*

Q15) Now we will ask you to rate the similarity between some personal experiences and inflation. At first sight, these events may seem dissimilar because they concern different domains. However, even experiences in different domains may share some aspects, such as personal stress, a sense of insecurity or uncertainty, impact on social life, economic consequences, etc.

Therefore, when you assess similarity, we would like you to think of an overall measure of similarity between experiences, taking into account a wide range of aspects.

Keeping this in mind, please indicate for each of the experiences below **how similar you think that experience is compared to future inflation being 7% or higher**. You can answer on a scale of 1 to 7, where 1 indicates ‘not similar at all’, and 7 indicates ‘extremely similar’.

Experience	1	2	3	4	5	6	7
A serious or chronic illness, or a serious accident	☐	☐	☐	☐	☐	☐	☐
Struggling with financial difficulties	☐	☐	☐	☐	☐	☐	☐
A recent extreme weather event, such as a very heavy storm, flood, or extreme heat	☐	☐	☐	☐	☐	☐	☐
An event increasing inequality or perceived as unfair	☐	☐	☐	☐	☐	☐	☐
Protests, confrontations, or other forms of intense tension between social groups	☐	☐	☐	☐	☐	☐	☐
Cheating, stealing, or breaking a contract	☐	☐	☐	☐	☐	☐	☐
A long-term investment in a business or in the stock market that has been successful in the long run	☐	☐	☐	☐	☐	☐	☐
Challenges related to the immigration of foreigners	☐	☐	☐	☐	☐	☐	☐
Unexpected job loss and unemployment	☐	☐	☐	☐	☐	☐	☐
Divorce or separation from a long-term partner	☐	☐	☐	☐	☐	☐	☐

1 = not similar at all; 7 = extremely similar

House Prices Similarity. *PROGRAMMING NOTE. Q16: For every respondent use the same order of questions as shown in Q14 and Q15.*

Q16) Now we will ask you to rate the similarity between some personal experiences and home

prices increases. At first sight, these events may seem dissimilar because they concern different domains. However, even experiences in different domains may share some aspects, such as personal stress, a sense of insecurity or uncertainty, impact on social life, economic consequences, etc. Therefore, when you assess similarity, we would like you to think of an overall measure of similarity between experiences, taking into account a wide range of aspects.

Keeping this in mind, please indicate for each of the experiences below **how similar you think that experience is compared to future home prices increasing by 10% or more**. You can answer on a scale of 1 to 7, where 1 indicates ‘not similar at all’, and 7 indicates ‘extremely similar’.

Experience	1	2	3	4	5	6	7
A serious or chronic illness, or a serious accident	☐	☐	☐	☐	☐	☐	☐
Struggling with financial difficulties	☐	☐	☐	☐	☐	☐	☐
A recent extreme weather event, such as a very heavy storm, flood, or extreme heat	☐	☐	☐	☐	☐	☐	☐
An event increasing inequality or perceived as unfair	☐	☐	☐	☐	☐	☐	☐
Protests, confrontations, or other forms of intense tension between social groups	☐	☐	☐	☐	☐	☐	☐
Cheating, stealing, or breaking a contract	☐	☐	☐	☐	☐	☐	☐
A long-term investment in a business or in the stock market that has been successful in the long run	☐	☐	☐	☐	☐	☐	☐
Challenges related to the immigration of foreigners	☐	☐	☐	☐	☐	☐	☐
Unexpected job loss and unemployment	☐	☐	☐	☐	☐	☐	☐
Divorce or separation from a long-term partner	☐	☐	☐	☐	☐	☐	☐

1 = not similar at all; 7 = extremely similar

Time Use, Family, and Backcast.

Q17) One dimension we would like to ask you about has to do with the time you spend on physical activity. In a typical week, how many hours would you say you spend on average on physical activity? This can include sports, but also other activities such as cycling or walking.

Q18) Another dimension we would like to ask you about has to do with time spent on household chores. In a typical week, how many hours would you say you spend performing household chores in your household? This can include cleaning and cooking, but also repairs.

Q19) Consider now your monthly expenditure. Relative to a year ago, has the price of items that you buy regularly changed, and if so by what percentage?

If you think the price of the items you buy regularly has risen, please enter a positive percentage. If you think the price of the items that you buy regularly has fallen, please enter a negative percentage (put a minus sign (-) in front of your answer). If you think the price of the items that you buy regularly has not changed, please enter 0 (zero). You can also use decimal places if you want.

Q20) Finally, consider home prices in the city where you live. Relative to a year ago, has the price of houses in the city where you live changed, and if so by what percentage?

If you think the price of houses in your city has risen, please enter a positive percentage. If you think the price of the houses in your city has fallen, please enter a negative percentage (put a minus sign (-) in front of your answer). If you think that the price of the houses in the city where you live has not changed, please enter 0 (zero). You can also use decimal places if you want.

B.3 Further Results on Experiences and Similarity

We provide further evidence on the measured components of memory (Section 2.3), namely on the properties of participants' databases and their perceived similarity between NDS experiences and high inflation or home price growth.

B.3.1 Frequency of NDS Experiences

Table B.8 reports the frequency of NDS experiences in different subsamples. The first column covers the control sample, followed by the health treatment (all respondents, denoted All, and respondents who experienced a health adversity, denoted TOT for treatment on treated), and similarly for the finance treatment. The shares in Health All and Finance All are very similar to and mostly statistically indistinguishable from those in Control, with the main exception that Health NDS is slightly more prevalent in the Health Treatment. One plausible reason is attention: these NDS were elicited at the outset of the survey, so people arguably exerted more effort to recall them compared with respondents in the control group. Table B.8 also shows significant differences between the databases of people in Control and the "Treatment on Treated" groups. This may be due to the endogeneity of experiences: financial, health, and other adversities such as unemployment or divorce are naturally intertwined via interaction effects and personal characteristics.² Table B.9 confirms this mechanism, showing higher-than-average correlations of finance with unemployment and divorce.

² An alternative possibility is that priming recall of experiences at the outset influences recall of other experiences later in the survey. This channel can be tested by comparing people who lived the primed experience in the treatment arm with people who lived the same experience in the control sample. The comparison reveals that these subsets of the treated and control samples do not systematically differ.

Table B.8: Share of Respondents Reporting Each Experience

	Control	Health		Finance	
		All	TOT	All	TOT
	(1)	(2)	(3)	(4)	(5)
<i>Panel A: Treatment Experiences</i>					
Health	0.42	0.50*** (0.00)	1.00*** (0.00)	0.42 (0.73)	0.51*** (0.00)
Finance	0.41	0.40 (0.37)	0.49*** (0.00)	0.35*** (0.00)	1.00*** (0.00)
<i>Panel B: Other Experiences</i>					
Extreme Weather	0.12	0.11 (0.50)	0.15** (0.02)	0.11 (0.60)	0.16** (0.02)
Inequality	0.28	0.30 (0.17)	0.38*** (0.00)	0.25* (0.06)	0.36*** (0.00)
Protest	0.10	0.11 (0.88)	0.13* (0.08)	0.10 (0.54)	0.15*** (0.00)
Cheating Contract	0.37	0.39 (0.32)	0.46*** (0.00)	0.37 (0.99)	0.56*** (0.00)
Successful Investment	0.12	0.12 (0.97)	0.15** (0.04)	0.11 (0.53)	0.12 (0.84)
Immigration / Foreigners	0.11	0.10 (0.21)	0.11 (0.80)	0.07*** (0.00)	0.11 (0.87)
Job Loss	0.38	0.39 (0.41)	0.46*** (0.00)	0.37 (0.93)	0.56*** (0.00)
Divorce	0.47	0.48 (0.55)	0.56*** (0.00)	0.49 (0.23)	0.65*** (0.00)

Notes: This table compares the Health and Finance treatment groups to the full control group. Reported values are the proportions of respondents who report each experience. p -values (in parentheses) and significance stars derive from two-sample tests of differences in proportions relative to the unrestricted control group. Health (All): all respondents assigned to the Health Treatment. Health (TOT): health-treated respondents who reported the health adversity experience. Finance (All): all respondents assigned to the Finance Treatment. Finance (TOT): finance-treated respondents who reported the finance adversity experience. Significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table B.9: Pairwise Correlations between Experience Indicators

	Health	Finance	Weather	Ineq.	Protest	Contract	Invest.	Immig.	Job Loss	Divorce
Health	1.00									
Finance	0.22	1.00								
Weather	0.15	0.15	1.00							
Inequality	0.23	0.29	0.20	1.00						
Protest	0.14	0.22	0.21	0.28	1.00					
Contract	0.21	0.34	0.18	0.30	0.22	1.00				
Investment	0.11	0.09	0.14	0.15	0.17	0.16	1.00			
Immigration	0.11	0.20	0.22	0.22	0.31	0.21	0.14	1.00		
Job Loss	0.18	0.31	0.15	0.34	0.19	0.26	0.12	0.19	1.00	
Divorce	0.21	0.28	0.14	0.19	0.12	0.43	0.11	0.13	0.22	1.00

Notes: This table reports pairwise correlations between the ten experience indicators elicited in the experiment. $N = 3,813$ respondents (Wave 1). All correlations are statistically significant at the 1% level. The ten experiences are: health problems, financial difficulty, extreme weather, inequality, protest, cheating contract, successful investment, immigration/foreigners, job loss, and divorce.

B.3.2 Similarity of NDS Experiences to Inflation and Home Price Growth

We next turn to assessments of similarity of NDS experiences to high inflation or home price growth, pooled across all treatments.

Table B.10 reports mean and standard deviation of similarity ratings (1–7 scale) for each NDS experience to high inflation and high home price growth (Panel A), along with summary statistics on cross-experience dispersion and a Levene test of equal variances. Panel B replicates the means separately for homeowners and renters. Table B.11 finds that having had an experience increases perceived similarity. Finally, Table B.12 assesses the relative similarity of the cued adversities to inflation and home price growth, in the control sample.

Table B.10: Statistics of Individual-Level Variation in Similarity

<i>Panel A: Pooled Sample</i>				
Experience	Inflation		HPG	
	Mean	SD	Mean	SD
Financial Difficulty	4.64	1.59	4.53	1.74
Job Loss	4.22	1.74	3.86	1.80
Inequality	3.72	1.67	3.61	1.77
Divorce	3.30	1.81	3.36	1.84
Successful Investment	3.41	1.69	3.35	1.73
Immigration / Foreigners	3.13	1.65	3.20	1.74
Protest	3.35	1.67	3.12	1.68
Extreme Weather	3.13	1.72	3.00	1.71
Illness or Accident	2.85	1.71	2.75	1.73
Cheating Contract	2.86	1.67	2.83	1.66
Average SD	1.70		1.73	
Min SD	1.59 [Financial]		1.66 [Cheating]	
Max SD	1.81 [Divorce]		1.84 [Divorce]	
Levene (W50)	$F(9, 38,120) = 38.76$ $p < 0.001$		$F(9, 38,120) = 12.52$ $p < 0.001$	

<i>Panel B: Owners vs. Renters</i>				
Experience	Inflation		HPG	
	Owners	Renters	Owners	Renters
Financial Difficulty	4.85	5.00	4.68	4.89
Job Loss	4.45	4.40	3.97	4.05
Inequality	3.99	4.21	3.85	4.14
Successful Investment	3.74	3.84	3.65	3.66
Immigration / Foreigners	3.74	3.60	3.99	3.66
Protest	3.62	3.58	3.36	3.66
Divorce	3.32	3.57	3.41	3.38
Extreme Weather	3.15	3.54	3.05	3.57
Cheating Contract	2.94	3.15	2.86	3.02
Illness or Accident	2.85	3.05	2.70	2.99

Notes: This table reports mean similarity scores (1–7 scale) for each personal experience, separately for inflation and home price expectations. Higher values indicate that respondents perceive their own experience as more similar to high inflation or high home price growth (HPG). “Average SD” reports the mean standard deviation across the ten experiences; “Min SD” and “Max SD” identify the experiences with the lowest and highest dispersion. Levene’s robust test (W50) assesses equality of variances across experiences; the test strongly rejects homoskedasticity for both inflation and home price similarity measures ($p < 0.001$). Panel B reports means separately for homeowners and renters.

Table B.11: Experience Database and Similarity Judgments

Experience	Inflation		HPG	
	<i>N</i> (Yes)	Difference	<i>N</i> (Yes)	Difference
Financial Difficulty	1,719	0.49***	1,719	0.43***
Inequality	1,413	0.54***	1,413	0.53***
Immigration / Foreigners	492	0.63***	492	0.75***
Protest	532	0.42***	532	0.43***
Job Loss	1,747	0.39***	1,747	0.26***
Successful Investment	618	0.36***	618	0.37***
Cheating Contract	1,736	0.27***	1,736	0.15***
Divorce	2,204	0.21***	2,204	0.24***
Extreme Weather	555	0.19**	555	0.19**
Illness or Accident	2,036	0.13**	2,036	0.08

Notes: This table reports mean differences in similarity ratings between respondents who reported having lived through each experience and those who did not. *N* (Yes) is the number of respondents who reported having the experience. “Difference” is the difference in mean similarity rating (had experience minus did not have experience) on a 1–7 scale, and the *p*-value is from a two-sample *t*-test. Similarity to inflation (home price growth, HPG) is the respondent’s rating of how similar each experience is to high inflation (high HPG). Sample includes all respondents in Wave 1 (*N* = 3,813). Significance levels: *** *p* < 0.01, ** *p* < 0.05, * *p* < 0.10.

Table B.12: Health and Finance Similarity in the Control Sample

	Similarity to High Inflation	Similarity to High HPG	Difference
Finance Experience	4.63 (1.61)	4.50 (1.76)	0.131***
Health Experience	2.90 (1.70)	2.87 (1.74)	0.035
Difference (Finance – Health)	1.726***	1.630***	

Notes: This table reports mean similarity ratings for finance and health experiences in the control group (*N* = 1,284). Standard deviations in parentheses. Similarity to High Inflation is the respondent’s rating of how similar their financial (health) experience is to high inflation, on a 1–7 scale. Similarity to High HPG (home price growth) is defined analogously for high home price growth. “Difference” reports the within-person paired *t*-test of similarity to high inflation minus similarity to high HPG. “Difference (Finance – Health)” reports the paired *t*-test of the finance experience similarity minus the health experience similarity within the same domain. The control group received no priming treatment, so these differences reflect respondents’ baseline perceptions of relevance. Significance levels: *** *p* < 0.01, ** *p* < 0.05, * *p* < 0.10.

B.4 Further Results on Experience and Priming Effects

We start by presenting all covariates of Table 2.3 (experience effects), Table 2.4 (priming effects), and Table 2.5 (similarity hierarchy of experience and priming effects). The full set of covariate estimates is reported in Table B.13.

We next assess the robustness of experience and priming effects across homeowners and renters. Table B.14 reports the corresponding estimates separately for homeowners and renters.

Finally, we assess the persistence of priming effects. To do so, we pool data from both waves and add dummies for Wave 2 and their interaction with primes. We focus on participants who were TOT in Wave 1 (i.e., we exclude those who did not report an adversity or did so only in Wave 2). All interaction terms are negative and mostly significant, confirming treatment effects decay from Wave 1 to Wave 2. Finance priming decay is especially strong. Table B.15 reports the estimates.

Table B.13: Experience and Priming Effects with Covariates

	Table 2.3		Table 2.4				Table 2.5	
	Control – Wave 1		ITT		TOT		TOT	
	Inflation	HPG	Inflation	HPG	Inflation	HPG	Inflation	HPG
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Age	−0.01*** (0.00)	−0.01*** (0.00)	−0.01*** (0.00)	−0.01*** (0.00)	−0.01*** (0.00)	−0.01*** (0.00)	−0.01*** (0.00)	−0.01*** (0.00)
Education: Vocational	−0.11* (0.07)	0.04 (0.06)	−0.07* (0.04)	−0.03 (0.04)	−0.06 (0.05)	0.01 (0.05)	−0.08 (0.05)	0.01 (0.05)
Education: University	−0.17* (0.09)	−0.01 (0.08)	−0.26*** (0.05)	−0.19*** (0.05)	−0.18*** (0.07)	−0.10 (0.07)	−0.18*** (0.07)	−0.10 (0.07)
Occupation: Employed	0.02 (0.09)	−0.02 (0.08)	0.01 (0.06)	−0.10* (0.06)	−0.01 (0.07)	−0.08 (0.06)	−0.01 (0.07)	−0.08 (0.06)
Occupation: Self-Employed	0.05 (0.14)	0.01 (0.13)	−0.02 (0.08)	−0.04 (0.09)	−0.01 (0.11)	0.04 (0.11)	−0.02 (0.11)	0.04 (0.11)
Occupation: Student	−0.42* (0.23)	−0.08 (0.18)	−0.40*** (0.12)	−0.35*** (0.12)	−0.39** (0.16)	−0.35** (0.14)	−0.44*** (0.16)	−0.36** (0.14)
Occupation: Retired	−0.04 (0.09)	0.00 (0.08)	−0.04 (0.05)	−0.12** (0.06)	−0.03 (0.07)	−0.05 (0.07)	−0.02 (0.07)	−0.05 (0.07)
Urbanization	−0.01 (0.02)	0.01 (0.02)	0.00 (0.01)	0.02 (0.01)	−0.00 (0.02)	0.01 (0.02)	0.00 (0.02)	0.01 (0.02)
Net Income	−0.00*** (0.00)	−0.00 (0.00)	−0.00*** (0.00)	−0.00 (0.00)	−0.00*** (0.00)	−0.00 (0.00)	−0.00*** (0.00)	−0.00 (0.00)
Housing: Owner	−0.19*** (0.07)	−0.18*** (0.07)	−0.25*** (0.04)	−0.20*** (0.04)	−0.23*** (0.05)	−0.21*** (0.05)	−0.24*** (0.05)	−0.22*** (0.05)
Male	−0.16*** (0.06)	−0.06 (0.05)	−0.17*** (0.03)	−0.06* (0.03)	−0.17*** (0.04)	−0.07 (0.04)	−0.16*** (0.04)	−0.06 (0.04)
Has Partner	0.09 (0.06)	0.08 (0.06)	0.01 (0.04)	0.04 (0.04)	0.04 (0.05)	0.05 (0.05)	0.05 (0.05)	0.05 (0.05)
Beliefs about Elderly	0.17*** (0.04)	0.11*** (0.03)	0.14*** (0.02)	0.08*** (0.02)	0.16*** (0.03)	0.10*** (0.02)	0.16*** (0.03)	0.10*** (0.02)
Constant	0.78*** (0.20)	0.38** (0.18)	0.89*** (0.12)	0.75*** (0.12)	0.96*** (0.15)	0.75*** (0.15)	0.89*** (0.15)	0.69*** (0.15)
<i>N</i>	1,284	1,284	3,813	3,813	2,367	2,367	2,367	2,367
<i>R</i> ²	0.12	0.05	0.12	0.06	0.13	0.06	0.13	0.07

Notes: This table reports the full set of covariate estimates for the main specifications in Tables 2.3, 2.4, and 2.5. Columns (1)–(2) correspond to Table 2.3 (control group, Wave 1, $N = 1,284$). Columns (3)–(6) correspond to Table 2.4 (columns (3)–(4): ITT sample, $N = 3,813$; columns (5)–(6): TOT sample, $N = 2,367$). Columns (7)–(8) correspond to Table 2.5 (TOT sample, $N = 2,367$). The dependent variables are the inflation (odd columns) and housing (even columns) belief composite z -scores. Robust standard errors in parentheses. Base categories: no vocational/university education, unemployed, and tenant. Significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table B.14: Experience and Priming Effects for Owners vs. Renters

<i>Panel A: Experience Effects</i>								
	Control Group – Wave 1							
	Inflation		HPG					
	Owners	Renters	Owners	Renters	Owners	Renters	Owners	Renters
	(1)	(2)	(3)	(4)	(3)	(4)	(3)	(4)
Finance Experience	0.12** (0.06)	0.27** (0.12)	0.04 (0.06)	0.24** (0.11)				
Health Experience	−0.05 (0.06)	0.05 (0.11)	0.06 (0.06)	0.01 (0.11)				
<i>N</i>	913	363	913	363				
Controls	Yes	Yes	Yes	Yes				
<i>R</i> ²	0.12	0.13	0.03	0.08				

<i>Panel B: Priming Effects</i>								
	ITT				TOT			
	Inflation		HPG		Inflation		HPG	
	Owners	Renters	Owners	Renters	Owners	Renters	Owners	Renters
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Finance Prime	0.03 (0.04)	0.17** (0.08)	0.05 (0.04)	0.15* (0.08)	0.12* (0.07)	0.20 (0.12)	0.17** (0.07)	0.21* (0.12)
Health Prime	0.02 (0.04)	0.11 (0.08)	0.07 (0.04)	0.18** (0.08)	0.14** (0.06)	0.18 (0.11)	0.10* (0.06)	0.25** (0.12)
Finance Experience	0.15*** (0.04)	0.23*** (0.07)	0.06 (0.04)	0.25*** (0.07)	0.11** (0.05)	0.20** (0.10)	0.01 (0.05)	0.23** (0.10)
Health Experience	−0.01 (0.03)	−0.01 (0.07)	0.07* (0.04)	−0.04 (0.07)	−0.09* (0.05)	−0.03 (0.10)	0.02 (0.05)	−0.09 (0.10)
<i>N</i>	2,663	1,118	2,663	1,118	1,631	721	1,631	721
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
<i>R</i> ²	0.09	0.09	0.04	0.06	0.11	0.11	0.04	0.08

Notes: This table reports OLS regressions of the inflation and home price growth (HPG) composite belief *z*-scores on experience and priming indicators, separately for homeowners (Owners) and renters (Renters). *Panel A* uses the control group in Wave 1 to isolate experience effects: “Finance Experience” and “Health Experience” indicate whether the respondent reported having lived through a financial or health adversity. *Panel B* adds priming effects using the ITT sample (all treatment-assigned respondents, columns (1)–(4)) and the TOT sample (treatment compliers, columns (5)–(8)): “Finance Prime” and “Health Prime” are indicators for assigned treatment. The dependent variables are composite indices that average the standardized point forecast (winsorized at the 2nd and 98th percentiles) and the standardized threshold belief, re-standardized within wave. Controls include age, education fixed effects (base: no vocational/university education), occupation fixed effects (base: unemployed), urbanization, net income, gender, relationship status, and beliefs about elderly (not shown; housing tenure is excluded since the sample is split by tenure). Robust standard errors in parentheses. Significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table B.15: Persistence of Priming Effects

	TOT	
	Inflation	HPG
	(1)	(2)
Finance Prime	0.13** (0.06)	0.14** (0.06)
Health Prime	0.12** (0.05)	0.13** (0.05)
Finance Prime $\times$ Wave 2	-0.10* (0.05)	-0.17*** (0.06)
Health Prime $\times$ Wave 2	-0.06 (0.05)	-0.12** (0.06)
Finance Experience	0.15*** (0.04)	0.13*** (0.04)
Health Experience	-0.03 (0.04)	0.00 (0.04)
N	4,800	4,800
Controls	Yes	Yes
R^2	0.11	0.06

Notes: This table reports OLS regressions pooling Wave 1 and Wave 2 data. The dependent variables are composite z -scores on inflation and housing beliefs. Treatment coefficients capture Wave 1 effects; interaction terms capture the change from Wave 1 to Wave 2 (negative = decay). “Finance Experience” and “Health Experience” indicate if a respondent lived a financial or health adversity. Controls include age, education fixed effects (base: no vocational/university education), occupation fixed effects (base: unemployed), urbanization, net income, housing tenure fixed effects (base: tenant), gender, relationship status, and beliefs about elderly (not shown). Standard errors clustered at the individual level in parentheses. Significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

B.4.1 Reliability of Similarity Measures

We next assess whether individual similarity ratings are reliable in the context of the analysis of Table 2.5. Several concerns arise. First, individual ratings may be noisy, in the sense of reflecting idiosyncratic judgments that are unstable across waves. Second, individual similarity ratings may be endogenous: respondents with higher inflation expectations may retrospectively judge their experiences as more similar to high inflation.

To assess the role of noise, we perform an IV exercise in which we instrument similarity

in one wave with similarity reported in the other wave, which allows us to use only the stable component of similarity variation. Note that this can only be done for non-primed experiences because we measure the similarity of e_f and e_h only in the second wave, where the correlation is on average 0.3. Results are reported in Table B.16, columns (1) and (2). To address the second concern, we replace individual-level similarity ratings for experiences with the corresponding sample-average similarity ratings. This removes any direct feedback from a person's own beliefs to their similarity assessments, while still preserving variation across experiences in perceived similarity at the population level.

Table B.16: Effect of Experience Similarity on Inflation and Home Price Growth Beliefs

	TOT			
	IV Similarity		Sample Similarity	
	Inflation	HPG	Inflation	HPG
	(1)	(2)	(3)	(4)
Avg. Similarity of Lived Experiences	0.05** (0.02)	0.07*** (0.02)	0.05* (0.03)	0.07*** (0.02)
Similarity of Primed Experience	0.07*** (0.02)	0.06*** (0.02)	0.08*** (0.02)	0.06*** (0.02)
Constant	1.07*** (0.15)	0.79*** (0.14)	1.13*** (0.15)	0.81*** (0.14)
N	2,367	2,367	2,367	2,367
Controls	Yes	Yes	Yes	Yes
R^2	0.12	0.07	0.12	0.07

Notes: This table reports robustness checks for the treatment on treated (TOT) specifications in Wave 1. The dependent variables are standardized belief measures (z -scores) based on respondents' point and threshold forecasts for inflation and home price growth (HPG). The two robustness approaches modify how similarity measures or beliefs are constructed. *IV Similarity:* similarity scores are instrumented using predicted similarity values. For each experience category, the respondent's Wave 1 similarity rating is regressed on their Wave 2 rating, and the fitted value is used as an instrumented similarity score. These instrumented values are then aggregated into (i) average similarity across all lived experiences, and (ii) similarity of the primed experience (set to zero for the control group). *Sample Similarity:* similarity scores are defined relative to the sample-average similarity rating for each experience category. Respondents who did not report an experience receive a value of zero. The same aggregation as in the baseline (average similarity and primed similarity) is applied. Controls include age, education fixed effects (base: no vocational/university education), occupation fixed effects (base: unemployed), urbanization, net income, housing tenure fixed effects (base: tenant), gender, relationship status, and beliefs about elderly (not shown). Robust standard errors in parentheses. Significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Across specifications, the estimated coefficients on average similarity of experiences are smaller than in Table 2.5 but remain sizable and significant. This attenuation is consistent with our approaches purging both measurement error *and* transitory variation, but the results confirm that our findings are not driven by noise in individual ratings. For primed similarity, the robustness checks reveal some differences across specifications. In the baseline, coefficients are around 0.07–0.08 SD and highly significant. Both the IV and sample similarity specifications produce similar coefficients, pointing to the reliability of our similarity measures.

The results underscore the role of memory in experience effects, which rely on similarity, and point to systematic heterogeneity of experience and priming effects. Having a given experience can be a source of strong pessimism about, say, inflation for a person who perceives the experience and inflation to be very similar, but generates less or no pessimism for a person for whom the experience feels dissimilar to inflation. In fact, having or being primed with a dissimilar experience interferes with recall of similar experiences from the database, and reduces pessimism. To investigate interference, we decompose the average similarity index into its numerator and denominator: the numerator captures the total similarity across recalled experiences, and the denominator is the total number of experiences recalled. Table B.17, columns (1) and (2), shows that the numerator enters positively, while the denominator enters negatively. This shows that, holding total similarity constant, recalling more experiences reduces the overall effect on beliefs. When many experiences can be recalled, they interfere with each other, affecting beliefs. To further assess robustness, in columns (3) and (4) we control for inflation and home price growth backcasts, respectively.

Finally, Table B.18 examines the role of similarity separately for each primed experience. Table 2.5 is also robust to using the ITT sample, which includes all subjects in our treatments, as well as to excluding demographic controls, which reassuringly suggests that the effect of

Table B.17: Experience Similarity, Interference, and Backcasts

	TOT		TOT	
	Inflation	HPG	Inflation	HPG
	(1)	(2)	(3)	(4)
Total Similarity	0.03*** (0.01)	0.02*** (0.01)	0.03*** (0.01)	0.01* (0.01)
Number of Experiences	-0.05*** (0.02)	-0.01 (0.02)	-0.05*** (0.02)	-0.01 (0.01)
Similarity Primed Experience	0.06*** (0.02)	0.05*** (0.02)	0.05*** (0.02)	0.05*** (0.02)
Backcast			0.03*** (0.00)	0.05*** (0.00)
<i>N</i>	2,367	2,367	2,367	2,367
Controls	Yes	Yes	Yes	Yes
<i>R</i> ²	0.13	0.07	0.22	0.25

Notes: This table decomposes the average similarity measure from Table 2.5 into its numerator (Total Similarity: the sum of similarity ratings across all lived experiences) and denominator (Number of Experiences: the count of experiences the respondent reported). Similarity Primed Experience is the similarity rating for the specific experience assigned as the priming treatment. Columns (3)–(4) add the respondent’s backcast – their belief about past inflation or home price growth over the preceding two years. All similarity variables are divided by their unconditional SD. The sample is the TOT group in Wave 1 ($N = 2,367$). Controls include age, education fixed effects (base: no vocational/university education), occupation fixed effects (base: unemployed), urbanization, net income, housing tenure fixed effects (base: tenant), gender, relationship status, and beliefs about elderly (not shown). Robust standard errors in parentheses. Significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

memory parameters is largely independent of our demographics. Allowing our similarity proxies to also enter quadratically – i.e., giving higher weight to experiences that are deemed more similar – yields comparable coefficients. Tables B.19 and B.20 report these robustness checks.

Finally, Table B.21 restricts the analysis to the ITT–TOT sample: participants in the treatment groups who did not have the experience they were treated with, together with control.

Table B.18: Similarity on Different Primed Experiences

	Inflation		HPG	
	Health	Finance	Health	Finance
	(1)	(2)	(3)	(4)
Avg. Similarity	0.08*** (0.03)	0.10*** (0.03)	0.06*** (0.02)	0.07*** (0.03)
Similarity Primed Experience	0.06** (0.03)	0.06*** (0.02)	0.07** (0.03)	0.06*** (0.02)
<i>N</i>	1,912	1,739	1,912	1,739
Controls	Yes	Yes	Yes	Yes
<i>R</i> ²	0.13	0.13	0.06	0.07

Notes: This table reports the Table 2.5 specification estimated separately for the health-primed and finance-primed treatment arms, each compared with the control group. The dependent variables are the inflation (columns (1)–(2)) and home price growth (HPG; columns (3)–(4)) belief composite *z*-scores. “Avg. Similarity” is the average similarity rating across all of the respondent’s lived experiences; “Similarity Primed Experience” is the similarity rating for the experience assigned as the priming treatment (set to zero for the control group). All similarity variables are divided by their unconditional SD pooled across persons and experiences. Sample: TOT (treatment compliers), Wave 1, restricted to health-prime plus control ($N = 1,912$) in columns (1) and (3), and finance-prime plus control ($N = 1,739$) in columns (2) and (4). Controls include age, education fixed effects (base: no vocational/university education), occupation fixed effects (base: unemployed), urbanization, net income, housing tenure fixed effects (base: tenant), gender, relationship status, and beliefs about elderly (not shown). Robust standard errors in parentheses. Significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table B.19: Robustness of Similarity Effects (Panels A–B)

<i>Panel A: ITT Sample</i>		
	Inflation	HPG
	(1)	(2)
Avg. Similarity Lived Experiences	0.09*** (0.03)	0.07*** (0.02)
Similarity Primed Experience	0.05*** (0.02)	0.06*** (0.02)
<i>N</i>	2,433	2,433
Controls	Yes	Yes
R^2	0.13	0.07
Sample	ITT	ITT
<i>Panel B: No Controls</i>		
	Inflation	HPG
	(1)	(2)
Avg. Similarity Lived Experiences	0.13*** (0.03)	0.09*** (0.02)
Similarity Primed Experience	0.06*** (0.02)	0.07*** (0.02)
<i>N</i>	2,367	2,367
Controls	No	No
R^2	0.02	0.02
Sample	TOT	TOT

Notes: This table reports robustness checks for Table 2.5. *Panel A* reports the Table 2.5 specification estimated on the ITT sample (all treatment-assigned respondents, Wave 1, $N = 2,433$). *Panel B* reports the Table 2.5 specification without demographic controls (TOT group in Wave 1, $N = 2,367$). The dependent variables are the inflation (column (1)) and home price growth (HPG; column (2)) belief composite z -scores. All similarity variables are divided by their unconditional SD. Controls (where included): age, education fixed effects (base: no vocational/university education), occupation fixed effects (base: unemployed), urbanization, net income, housing tenure fixed effects (base: tenant), gender, relationship status, and beliefs about elderly (not shown). Robust standard errors in parentheses. Significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table B.20: Robustness of Similarity Effects (Panels C–E)

<i>Panel C: Quadratic Similarity</i>		
	Inflation	HPG
	(1)	(2)
Avg. Similarity Squared	0.09*** (0.03)	0.07*** (0.02)
Similarity Primed Squared	0.06*** (0.02)	0.05*** (0.02)
<i>N</i>	2,367	2,367
Controls	Yes	Yes
<i>R</i> ²	0.13	0.07
<i>Panel D: Sample Similarity</i>		
	Inflation	HPG
	(1)	(2)
Avg. Similarity (Sample Mean)	0.10** (0.05)	0.16*** (0.05)
Similarity Primed (Sample Mean)	0.07*** (0.02)	0.07*** (0.02)
<i>N</i>	2,367	2,367
Controls	Yes	Yes
<i>R</i> ²	0.12	0.07
<i>Panel E: IV Similarity</i>		
	Inflation	HPG
	(1)	(2)
Avg. Similarity (IV)	0.08** (0.04)	0.07** (0.03)
Similarity Primed (IV)	0.07*** (0.02)	0.07*** (0.02)
<i>N</i>	2,367	2,367
Controls	Yes	Yes
<i>R</i> ²	0.12	0.07

Notes: This table reports three further robustness checks for Table 2.5. *Panel C* squares each individual similarity rating (1–7 scale) before averaging across lived experiences, then standardizes the resulting averages (*z*-scored). *Panel D* replaces individual ratings with the population-mean similarity for each experience (computed across all Wave 1 respondents); each respondent’s measure is the average of these means across lived experiences. *Panel E* instruments each respondent’s Wave 1 similarity rating with their Wave 2 rating and uses the fitted values. The dependent variables are the inflation (column (1)) and home price growth (HPG; column (2)) belief composite *z*-scores. All similarity variables are divided by their unconditional SD. Controls include age, education fixed effects (base: no vocational/university education), occupation fixed effects (base: unemployed), urbanization, net income, housing tenure fixed effects (base: tenant), gender, relationship status, and beliefs about elderly (not shown). Robust standard errors in parentheses. Significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table B.21: Treatment without Primed Experience

	No Experience Controls		With Experience Controls	
	Inflation	HPG	Inflation	HPG
	(1)	(2)	(3)	(4)
Finance Prime	−0.04 (0.04)	−0.02 (0.04)	0.03 (0.04)	0.02 (0.05)
Health Prime	−0.04 (0.04)	0.05 (0.05)	−0.06 (0.05)	0.07 (0.05)
Finance Experience			0.18*** (0.05)	0.11** (0.05)
Health Experience			−0.10** (0.04)	0.01 (0.04)
<i>N</i>	2,730	2,730	2,730	2,730
Controls	Yes	Yes	Yes	Yes
<i>R</i> ²	0.11	0.05	0.11	0.05

Notes: This table reports estimates for the subsample of respondents in the treatment groups who did *not* have the experience they were primed with, pooled with the control group. Columns (1)–(2) include only the priming dummies; columns (3)–(4) additionally control for whether the respondent had a finance or health adversity. Controls include age, education fixed effects (base: no vocational/university education), occupation fixed effects (base: unemployed), urbanization, net income, housing tenure fixed effects (base: tenant), gender, relationship status, and beliefs about elderly (not shown). Robust standard errors in parentheses. Significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

The results show that the treatment given to people without the experience (which consisted of asking whether they had the experience, but without the subsequent retrieval task) has no effect on their beliefs.

B.4.2 Quantification

Dependent Variable Construction. The dependent variable in Tables 2.3, 2.4, and 2.5 is a composite z -score of macroeconomic beliefs. For inflation, it is the average of the z -scored point forecast (expected 1-year inflation rate, in percentage points) and the z -scored threshold belief (probability of inflation exceeding 7%, on a 0–10 scale), re-standardized within wave.

The housing composite is constructed analogously from the z -scored expected 2-year home price growth rate and the z -scored probability of house price increases.

General Conversion Formula. To convert a composite z -score coefficient to underlying units:

$$\text{Effect} = \beta_{\text{composite}} \times \text{loading} \times \text{SD}_{\text{outcome}},$$

where the loading is the regression coefficient of the standardized underlying measure on the composite z -score, and $\text{SD}_{\text{outcome}}$ is the SD of the underlying measure in its natural units (percentage points or 0–10 scale). Because the composite is the re-standardized average of two equally weighted z -scores, the loading is the same for both underlying measures within each domain.

Conversion Parameters. The conversion parameters are computed on Wave 1 across all respondents and are reported in Table B.22.

As an example, consider Tables 2.3 and 2.4. The Finance Experience coefficient in Table 2.3 (TOT, inflation) is $\beta = 0.1815$. Converting to point forecast p.p.: $0.1815 \times 0.8307 \times 8.4017 = 1.27$ p.p. The Finance Prime coefficient in Table 2.4 (TOT, inflation) is $\beta = 0.1388$. Converting: $0.1388 \times 6.9793 = 0.97$ p.p.

For Table 2.5, the similarity variables are divided by their unconditional SD (pooled across all persons and 10 experiences). The same conversion formula applies: multiply the composite z -score coefficient by the loading and $\text{SD}_{\text{outcome}}$ to obtain p.p. effects. For the quantification in Section 2.5.2, the population-mean similarity shift is expressed in units of the person-level average similarity SD, which requires an additional rescaling factor (pooled SD / SD of person average).

Table B.22: Conversion Parameters for Composite z -Score Coefficients (Wave 1)

Outcome	Loading	SD	Product
<i>Point Forecasts (%)</i>			
Inflation	0.83	8.40	6.98
HPG	0.85	7.55	6.44
<i>Threshold Forecast (0–10)</i>			
Inflation	0.83	2.48	2.06
HPG	0.85	2.42	2.07

Notes: This table reports the conversion parameters used to translate composite z -score coefficients into underlying outcome units. Loadings are regression coefficients of the standardized underlying measure on the composite z -score. SDs are computed on Wave 1 in the natural units of each underlying measure (percentage points for point forecasts, 0–10 scale for threshold beliefs). The product loading $\times$ SD converts a composite z -score coefficient into the underlying outcome units. Values shown are rounded to two decimals; the unrounded loadings and SDs are used in the worked example below.

B.5 Further Results on Reasoning

B.5.1 Protocol for Extraction of Valence and Context

We classify each open-ended response along two dimensions: the *logic* of reasoning (personal vs. news-based) and the *emotional valence* (negative, positive/calm, or neutral/factual). Classifications are produced by GPT-4 using the prompts reproduced below. To validate the GPT-based classification, we hand-coded a random subsample of responses and compared human and GPT labels. Table B.23 reports the resulting confusion matrices for inflation and home price growth, both for the logic and the valence dimensions.

Inflation Logic – Prompt.

Task: You are a text classification model. Given a piece of text, classify it into one of the following categories:

- **Personal Logic:** The response uses personal experiences, financial situations, or individual observations to form an inflation expectation.
- **News-Based Logic:** The response references media, government policies, central banks, economic reports, or broader economic events.
- **Unclear/Other:** The response does not clearly fit into the above categories.

Examples:

- Text: “I think inflation will go up because my rent increased.”
 - Category: Personal Logic

Table B.23: GPT Classification Confusion Table

Panel A: Inflation Reasoning

	GPT: Personal	GPT: News	GPT: Unclear	Total
Human: Personal	95 (75.4%)	6 (4.8%)	25 (19.8%)	126
Human: News	23 (12.4%)	133 (71.9%)	29 (15.7%)	185
Human: Unclear	32 (23.0%)	11 (7.9%)	96 (69.1%)	139
Total	150	150	150	450

Notes: Overall agreement: 324/450 (72.0%). Cohen’s $\kappa = 0.580$. Per-category accuracy: Personal 95/126 (75.4%); News 133/185 (71.9%); Unclear 96/139 (69.1%).

Panel B: HPG Reasoning

	GPT: Personal	GPT: News	GPT: Unclear	Total
Human: Personal	96 (85.0%)	8 (7.1%)	9 (8.0%)	113
Human: News	4 (3.8%)	86 (82.7%)	14 (13.5%)	104
Human: Unclear	49 (21.2%)	55 (23.8%)	127 (55.0%)	231
Total	149	149	150	448

Notes: Overall agreement: 309/448 (69.0%). Cohen’s $\kappa = 0.534$. Per-category accuracy: Personal 96/113 (85.0%); News 86/104 (82.7%); Unclear 127/231 (55.0%).

Panel C: Inflation Valence

	GPT: Negative Affect	GPT: Neutral/Factual	GPT: Positive/Calm Affect	GPT: Uninformative	Total
Human: Negative Affect	53 (82.8%)	10 (15.6%)	1 (1.6%)	0 (0.0%)	64
Human: Neutral/Factual	4 (11.8%)	30 (88.2%)	0 (0.0%)	0 (0.0%)	34
Human: Positive/Calm Affect	1 (1.4%)	11 (15.5%)	59 (83.1%)	0 (0.0%)	71
Human: Uninformative	2 (2.8%)	9 (12.7%)	0 (0.0%)	60 (84.5%)	71
Total	60	60	60	60	240

Panel D: HPG Valence

	GPT: Negative Affect	GPT: Neutral/Factual	GPT: Positive/Calm Affect	GPT: Uninformative	Total
Human: Negative Affect	58 (89.2%)	7 (10.8%)	0 (0.0%)	0 (0.0%)	65
Human: Neutral/Factual	1 (2.2%)	44 (95.7%)	1 (2.2%)	0 (0.0%)	46
Human: Positive/Calm Affect	0 (0.0%)	8 (11.9%)	59 (88.1%)	0 (0.0%)	67
Human: Uninformative	1 (1.6%)	1 (1.6%)	0 (0.0%)	60 (96.8%)	62
Total	60	60	60	60	240

Notes: This table reports confusion matrices comparing GPT-based classifications to a hand-coded validation subsample. Diagonal cells report correctly classified responses; off-diagonal cells report misclassifications. Panels A and B classify the logic of reasoning (Personal, News-Based, Unclear/Other); Panels C and D classify the emotional valence (Negative Affect, Neutral/Factual, Positive/Calm Affect, Uninformative).

- Text: “The prices of living continue to rise. If you have an outing, you will spend much more. In general, life has become much more expensive.”
 - Category: Personal Logic
 - Text: “Inflation is expected to decrease since the central bank is cutting rates.”
 - Category: News-Based Logic
 - Text: “State of the economy. Future uncertain due to wars and new cabinet.”
 - Category: News-Based Logic
 - Text: “I don’t know, I don’t follow the economy.”
 - Category: Unclear/Other
 - Text: “I don’t really know anything about it, so any answer would be a shot in the dark.”
 - Category: Unclear/Other
-

Home Price Growth Logic – Prompt.

Task: You are a text classification model. Given a piece of text, classify it into one of the following categories:

- **Personal Logic:** The respondent references their own experience or someone close to them (e.g., buying/selling a house, children struggling to buy, local prices, personal emotions).
- **News-Based Logic:** The respondent references media reports, economic trends, interest rates, forecasts, policies, or government decisions.

- **Unclear/Other:** The response lacks clear reasoning or is vague, generic, or irrelevant.

Examples:

- Text: “My daughter can’t find an affordable house.”
 - Category: Personal Logic
- Text: “Interest rates are rising, so prices will go down.”
 - Category: News-Based Logic
- Text: “I don’t know.”
 - Category: Unclear/Other
- Text: “Everyone says houses are getting more expensive.”
 - Category: News-Based Logic
- Text: “I live in a big house and feel guilty.”
 - Category: Personal Logic

Emotional Valence – Prompt.

Task: You are an expert text classifier. Your task is to classify the **emotional tone** of a respondent’s answer about economic expectations. Important: Focus on *how the person feels*, not on what they predict will happen. A respondent may expect high inflation but feel calm, or expect stable prices but feel anxious. The classification should reflect emotional tone only.

Categories:

- **Negative Affect:** Negative emotions dominate, such as worry, anxiety, frustration, anger, fear, hopelessness, stress, feeling overwhelmed, bitterness, or complaint. This includes resigned negativity (e.g., “nothing we can do about it” expressed with defeat).
- **Positive/Calm Affect:** Positive emotions or genuine equanimity dominate, such as optimism, relief, hope, confidence, reassurance, or calm acceptance without distress.
- **Unclear / No Emotional Signal:** The response is purely factual, too brief to interpret, or contains no discernible emotional content. Use this category only when emotional tone cannot reasonably be inferred.

If emotions are mixed, classify according to the **dominant** tone.

Response Format: Respond with **only one** of the following:

- “Negative Affect”
- “Positive/Calm Affect”
- “Unclear / No Emotional Signal”

Examples:

- Text: “I just can’t keep up anymore. It’s exhausting.”
 - Category: Negative Affect
- Text: “Prices went up 4% last year.”
 - Category: Unclear / No Emotional Signal

- Text: “Things are stabilizing and I’m feeling better about it.”
 - Category: Positive/Calm Affect
- Text: “It’s scary. I don’t know how we’ll manage.”
 - Category: Negative Affect
- Text: “Inflation is a normal part of the economy.”
 - Category: Unclear / No Emotional Signal
- Text: “I’m not too worried, these things balance out.”
 - Category: Positive/Calm Affect
- Text: “The government has completely mismanaged this.”
 - Category: Negative Affect
- Text: “I don’t know.”
 - Category: Unclear / No Emotional Signal
- Text: “We’ll manage somehow, we always do.”
 - Category: Positive/Calm Affect
- Text: “What can you do, that’s just how it goes.”
 - Category: Unclear / No Emotional Signal

The next two tables show example responses by reasoning logic and valence, for inflation (Table B.24) and home price growth (Table B.25).

Table B.24: Example Responses by Logic and Valence: Inflation

Logic	Valence	<i>N</i>	Example
Personal	Negative	802	“My first thought was: Should I invest my savings differently? But how? Then I decided not to put any more energy into it” “That it will all be even more expensive and people will get into even more trouble” “Too bad everything has to be more expensive again. I hope that wages will also increase enormously, just like benefits.”
Personal	Positive/Calm	106	“Thinking back, I see the inflation of a few years ago, which was completely derailing. At that moment I thought for a moment that it would not be okay. I now see that it is normalizing again” “I hope that the rates do not increase as a result of the Ukraine/Russia war. I think of my children who are still studying, hopefully they will also get great opportunities” “I have the feeling that it stabilizes. As a result, I entered a lower number”
Personal	Neutral/Factual	679	“Lower inflation means lower interest rates, which is good for the price of my shares.” “Prices are currently rising enormously on all fronts” “that everything is getting more expensive”
Personal	Uninformative	3	“Expensive groceries” “Increased wages” “Grocery prices”
News-Based	Negative	111	“Everything is getting more expensive war Ukraine high prices nothing is going down” “THE economy, energy problem, billions that are gone due to the influx of refugees, war threat with Russia.” “In recent years you can no longer think of stable inflation...”
News-Based	Positive/Calm	33	“The news is not so gloomy about rising inflation. Price increases are levelling off and the economy is developing positively.” “I think they have the prices under control again so I don’t expect too much of an increase.” “I’m not so afraid of high inflation, the main reason could probably be energy prices, but I don’t see that escalating...”
News-Based	Neutral/Factual	588	“Inflation is very likely, but less robust than in recent years. At least, that’s what is predicted and I have read here and there in the media.” “I read about a lack of housing, which will cause house prices to rise.” “The temporary increase in energy and the subsequent increase in wages is behind us. Inflation is returning to fairly normal, around 2%.”
News-Based	Uninformative	1	“War, cabinet”
Unclear/Other	Negative	254	“Life becomes difficult” “just keep it low. but I think it’s not going well. taxes, energy, climate etc.” “It remains a difficult matter.”
Unclear/Other	Positive/Calm	84	“well honestly I don’t concern myself with that at all” “they have it under control” “I think it won’t be too bad”
Unclear/Other	Neutral/Factual	789	“boring, has been high lately, but I think we are in a downward trend” “The last question was not credible” “News items on this topic. Own experiences. Experiences of friends and family.”
Unclear/Other	Uninformative	394	“shop” “.” “War.”

Notes: This table shows example inflation-reasoning responses for each combination of reasoning logic (Personal, News-Based, Unclear/Other) and valence (Negative, Positive/Calm, Neutral/Factual, Uninformative). *N* is the number of responses in each category (full sample, *N* = 3,813).

Table B.25: Example Responses by Logic and Valence: Home Price Growth

Logic	Valence	<i>N</i>	Example
Personal	Negative	506	“Houses are far too expensive. Will not be built for the middle class.” “I especially hope that house prices will fall. It’s too crazy for words how much you pay for an owner-occupied house these days.” “I hope not. It is a disaster for our youth who have nowhere to live.”
Personal	Positive/Calm	106	“I thought there will not be a strong higher price immediately and I also don’t want to buy another house in the near future. So it doesn’t really matter to me.” “Fortunately, I was able to buy a house well and for the time being it doesn’t matter to us.” “Good for us, we own a house”
Personal	Neutral/Factual	308	“My house has become much more valuable in a short time” “The rent is becoming more and more” “The value of our house has almost doubled 6 years after purchase. . .”
News-Based	Negative	298	“It is almost impossible for first-time buyers to buy a house.” “House prices have been driven up for decades by speculators and other commercial parties.” “The limits of affordability are in sight.”
News-Based	Positive/Calm	24	“The increase in prices can’t rise much more, can it? Hopefully, the demand for houses will decrease.” “Increases are over. Now more stability.” “The current government seems to want to prevent this.”
News-Based	Neutral/Factual	1,215	“Despite everything, house prices continue to rise, it was also just in the newspaper” “Housing market is overvalued due to housing shortage” “Rents rise along with owner-occupied houses”
News-Based	Uninformative	5	“Higher.” “Only higher” “Newspaper reports.”
Unclear/Other	Negative	224	“Genuine? Is this necessary? Difficult.” “the price of a house is irresponsible, people want to earn it has become an obsession” “No one can justify that in this day and age”
Unclear/Other	Positive/Calm	46	“I am not concerned with house prices at all” “I don’t know but hope that it is now well done with those price increases” “Hopefully there will be a new angle for houses for the elderly”
Unclear/Other	Neutral/Factual	639	“Also difficult to answer.” “felt that it will not rise yet but not much more” “I don’t know what the housing market will do.”
Unclear/Other	Uninformative	471	“No” “No experience” “ ”

Notes: This table shows example home-price-growth-reasoning responses for each combination of reasoning logic (Personal, News-Based, Unclear/Other) and valence (Negative, Positive/Calm, Neutral/Factual, Uninformative). *N* is the number of responses in each category (full sample, $N = 3,813$).

B.5.2 Robustness on Backcasts and Reasoning

Table B.26 presents Table 2.6 with all covariates.

Table B.26: Backcasts, Context, and Valence with Covariates

	Backcasts (%)		Pers. or Neg.	
	Inflation	HPG	Inflation	HPG
	(1)	(2)	(3)	(4)
Finance Prime	1.83*** (0.58)	1.24** (0.53)	0.10*** (0.03)	0.12*** (0.03)
Health Prime	0.94* (0.48)	0.36 (0.39)	0.06** (0.02)	0.08*** (0.02)
Age	−0.09*** (0.02)	−0.08*** (0.02)	−0.00 (0.00)	−0.00 (0.00)
Education: Vocational	0.02 (0.51)	0.43 (0.45)	0.02 (0.02)	0.06** (0.02)
Education: University	−0.62 (0.65)	−0.46 (0.59)	−0.09** (0.04)	0.02 (0.03)
Occupation: Employed	−2.35*** (0.75)	−0.79 (0.64)	−0.03 (0.03)	−0.05* (0.03)
Occupation: Self-Employed	−0.40 (1.10)	0.10 (0.98)	−0.01 (0.05)	−0.10** (0.05)
Occupation: Student	−7.03*** (1.43)	−3.82*** (1.38)	0.08 (0.08)	0.06 (0.09)
Occupation: Retired	−1.77** (0.72)	0.12 (0.56)	−0.05 (0.03)	−0.06* (0.03)
Urbanization	0.26 (0.17)	0.10 (0.15)	0.00 (0.01)	−0.01 (0.01)
Net Income	−0.00** (0.00)	−0.00 (0.00)	−0.00 (0.00)	0.00*** (0.00)
Housing: Owner	0.22 (0.53)	−0.42 (0.49)	−0.02 (0.02)	−0.00 (0.02)
Male	−0.77* (0.42)	−0.63 (0.38)	−0.12*** (0.02)	−0.10*** (0.02)
Has Partner	−0.17 (0.46)	0.21 (0.42)	−0.01 (0.02)	−0.03 (0.02)
Beliefs about Elderly	0.87*** (0.28)	1.13*** (0.29)	0.00 (0.01)	−0.01 (0.01)
Constant	18.26*** (1.57)	14.41*** (1.41)	0.67*** (0.07)	0.37*** (0.07)
<i>N</i>	2,367	2,367	2,367	2,367
<i>R</i> ²	0.05	0.04	0.04	0.03

Notes: This table reports OLS regressions. Sample: TOT (treatment compliers), Wave 1. Columns (1)–(2): dependent variable is backcast of inflation/home price growth (HPG; % change, winsorized at 2nd and 98th percentiles). Columns (3)–(4): dependent variable is an indicator equal to 1 if the respondent’s reasoning is classified as using personal logic OR negative affect. Robust standard errors in parentheses. Base categories: no vocational/university education, unemployed, tenant. Significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table B.27 shows that, in line with Prediction 3, priming moves people toward personal logic *and* negative valence (jointly, a stronger result than that in Table 2.6, columns (1) and (2)), away from news-based logic (columns (3) and (4)), and that among news-based reasoning, priming increases the prevalence of negative valence (columns (5) and (6)).

Table B.27: Priming and Reasoning

	Personal AND Negative		News-Based Logic		Negative in News-Based	
	Inflation	HPG	Inflation	HPG	Inflation	HPG
	(1)	(2)	(3)	(4)	(5)	(6)
Finance Prime	0.08*** (0.02)	0.08*** (0.02)	−0.04* (0.02)	−0.11*** (0.02)	0.08* (0.05)	0.15*** (0.04)
Health Prime	0.05** (0.02)	0.09*** (0.02)	−0.04** (0.02)	−0.01 (0.02)	0.11*** (0.04)	0.09*** (0.03)
<i>N</i>	2,367	2,367	2,353	2,359	517	880
Controls	Yes	Yes	Yes	Yes	Yes	Yes
<i>R</i> ²	0.03	0.04	0.05	0.05	0.08	0.04

Notes: This table reports OLS regressions. Sample: TOT (treatment compliers), Wave 1. Controls include age, education fixed effects (base: no vocational/university education), occupation fixed effects (base: unemployed), urbanization, net income, housing tenure fixed effects (base: tenant), gender, relationship status, and beliefs about elderly (not shown). Robust standard errors in parentheses. Columns (1) and (2): dependent variable is an indicator equal to 1 if the respondent’s reasoning is classified as both personal logic and negative affect (24% for inflation, 14% for home price growth). This contrasts with the main table (Table 2.6), where the dependent variable uses an OR condition. Columns (3) and (4): dependent variable is an indicator equal to 1 if the respondent’s reasoning is classified as purely news-based logic (excluding mixed categories). Negative coefficients indicate that priming moves respondents away from news-based reasoning. Columns (5) and (6): sample restricted to TOT respondents (Wave 1) whose reasoning is classified as purely news-based logic ($N = 517$ for inflation, $N = 880$ for home price growth). Dependent variable is an indicator equal to 1 if the respondent’s reasoning exhibits negative affect. Positive coefficients indicate that even among those using news-based reasoning, priming increases the prevalence of negative valence. Significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

B.5.3 Similarity between Reasons and Primed Recall

The simulation function $\sigma_T(e)$ translates the features (x, v, c) of the recalled experience into a forecast. Intuitively, then, reasoning about the forecast should incorporate features that are similar to those of the simulator e . To test this mechanism, we measure the text-based

similarity of Inflation or Home Price Growth reasoning to text descriptions of primed NDS adversities. According to Prediction 3, the similarity of reasoning about t_T to the description of e_p should be greater in pT than in Control.

To ensure that our results are not mechanically due to the similarity of a person's writing style across their recall and forecasting tasks, we compare a given participant's reasoning in pT to the description of e_p by *other respondents* in the same pT treatment. To compute reasoning similarity, we use sentence embeddings generated by a pre-trained Sentence-BERT model, which generates 384-dimensional vector representations of open text optimized for semantic comparison. Cosine similarity is then computed between two texts and ranges from -1 (opposite meaning) to $+1$ (identical meaning).

Table B.28 Panel A shows the results. The Inflation and Home Price Growth reasoning of subjects in fT (Finance TOT, left) is more similar to descriptions of financial adversity compared to Control subjects. The effect is weaker for Home Price Growth, consistent with a smaller role of fT in this domain. The same is true for participants treated with recall of health adversities (Health TOT, right). Panel B documents a similar role of spontaneous recall: people who have lived e_p exhibit reasoning more similar to the description of e_p compared to people who have not lived the same experience. This holds for finance but not for health, consistent with the finding of Tables 2.3 and 2.4 that health adversity does not exhibit a significant experience effect. This suggests that the reasoning of people with experience tracks the content of the recalled adversity.

In Panel C, we compare the reasoning of people who have lived e_p with and without priming of p , and find no statistical difference. This suggests that experience and priming effects arise from the same memory mechanism, shaping beliefs through retrieval and reasoning.

Table B.28: Similarity of Reasoning Mode and NDS Recall

<i>Panel A: Participants in Treatment p</i>				
	Inflation (TOT)		HPG (TOT)	
	Finance – Control Difference (p -value)	Health – Control Difference (p -value)	Finance – Control Difference (p -value)	Health – Control Difference (p -value)
	(1)	(2)	(3)	(4)
Similarity to Finance	+0.02 (0.00)***		+0.00 (0.23)	
Similarity to Health		+0.01 (0.02)**		+0.01 (0.05)*
<i>Panel B: Participants in Control Group with Experience p</i>				
	Inflation		HPG	
	Finance – Control Difference (p -value)	Health – Control Difference (p -value)	Finance – Control Difference (p -value)	Health – Control Difference (p -value)
	(1)	(2)	(3)	(4)
Similarity to Finance	+0.01 (0.03)**		+0.02 (0.00)***	
Similarity to Health		–0.00 (0.76)		+0.01 (0.11)
<i>Panel C: Treatment p (TOT) vs. Participants in Control Group with Experience p</i>				
	Inflation		HPG	
	Finance TOT – Control with Finance Experience Difference (p -value)	Health TOT – Control with Health Experience Difference (p -value)	Finance TOT – Control with Finance Experience Difference (p -value)	Health TOT – Control with Health Experience Difference (p -value)
	(1)	(2)	(3)	(4)
Similarity to Finance	+0.01 (0.22)		+0.01 (0.34)	
Similarity to Health		+0.01 (0.34)		–0.00 (0.62)

Notes: This table reports cosine similarity, computed using a Sentence-BERT model (all-MiniLM-L6-v2), between each respondent’s macroeconomic reasoning and the pooled leave-one-out corpus of adversity descriptions from the treatment group. Leave-one-out is applied for TOT participants; control participants are compared to the full corpus. Panel A compares similarity across f T and Control (first row) and h T and Control (second row). Panel B repeats the analysis for participants in the control group who reported having had either a finance or a health adversity. Panel C compares TOT participants (who were primed with adversity p and had the experience) to control participants who reported having had experience p but received no prime; this tests whether experimental priming adds to reasoning similarity beyond what personal experience alone provides. None of the differences in Panel C are statistically significant, suggesting that priming does not significantly increase reasoning similarity beyond what personal experience alone provides. Differences correspond to two-sided independent-sample t -tests; effect sizes are expressed as differences in mean cosine similarity scores. Sample sizes: Finance TOT $N = 454$, Control with Finance Experience $N = 470$; Health TOT $N = 627$, Control with Health Experience $N = 90$. Significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

B.5.4 Reasoning and Beliefs

Table B.29 estimates the IV specifications separately by treatment arm, and Table B.30 presents Table 2.7 with all covariates.

Table B.29: Robustness, Treatment Level

	Health Treatment				Finance Treatment			
	Inflation	HPG	Inflation	HPG	Inflation	HPG	Inflation	HPG
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Backcast (Predicted)	0.12* (0.07)	0.44 (0.53)			0.13*** (0.04)	0.18*** (0.07)		
Personal or Negative (Predicted)			1.64* (0.89)	1.50** (0.68)			2.24*** (0.70)	1.79*** (0.53)
<i>N</i>	1,912	1,912	1,912	1,912	1,739	1,739	1,739	1,739
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
<i>R</i> ²	−0.65	−11.53	−0.34	−0.44	−0.64	−1.47	−0.74	−0.65
KP F-stat	3.20	0.57	7.04	13.51	10.79	6.15	16.52	25.58

Notes: This table reports two-stage least squares (2SLS) regressions, with each treatment arm estimated separately against the control group. Sample: TOT (treatment compliers), Wave 1. Dependent variable: composite belief *z*-score. Columns (1)–(4) restrict the sample to health prime and control participants (*N* = 1,912); columns (5)–(8) restrict to finance prime and control participants (*N* = 1,739). In columns (1)–(2) and (5)–(6), the instrumented variable is the backcast (recalled past inflation or home price change). In columns (3)–(4) and (7)–(8), the instrumented variable is an indicator for personal logic or negative affect. The instrument is the relevant treatment dummy (health prime in columns (1)–(4), finance prime in columns (5)–(8)). Controls include age, education fixed effects (base: no vocational/university education), occupation fixed effects (base: unemployed), urbanization, net income, housing tenure fixed effects (base: tenant), gender, relationship status, and beliefs about elderly (not shown). Robust standard errors in parentheses. The Kleibergen-Paap F-statistic tests instrument relevance. Significance levels: *** *p* < 0.01, ** *p* < 0.05, * *p* < 0.10.

Table B.30: Beliefs and Predicted Backcasts, Context, and Valence with Covariates

	Backcast (Predicted)		Personal or Negative (Predicted)	
	Inflation	HPG	Inflation	HPG
	(1)	(2)	(3)	(4)
Predicted Variable	0.12*** (0.04)	0.18*** (0.07)	2.20*** (0.70)	1.78*** (0.52)
Age	0.00 (0.00)	0.00 (0.01)	−0.01*** (0.00)	−0.01*** (0.00)
Education: Vocational	−0.07 (0.07)	−0.07 (0.08)	−0.12* (0.07)	−0.05 (0.06)
Education: University	−0.11 (0.08)	−0.02 (0.10)	0.01 (0.11)	0.04 (0.08)
Occupation: Employed	0.27** (0.13)	0.05 (0.12)	0.04 (0.09)	−0.04 (0.08)
Occupation: Self-Employed	0.04 (0.15)	0.03 (0.17)	0.02 (0.14)	0.09 (0.13)
Occupation: Student	0.44 (0.31)	0.31 (0.34)	−0.60*** (0.22)	−0.48** (0.19)
Occupation: Retired	0.18* (0.11)	−0.07 (0.09)	0.08 (0.10)	−0.02 (0.08)
Urbanization	−0.03 (0.02)	−0.00 (0.02)	−0.00 (0.02)	0.03 (0.02)
Net Income	−0.00* (0.00)	0.00 (0.00)	−0.00*** (0.00)	−0.00** (0.00)
Housing: Owner	−0.27*** (0.07)	−0.15* (0.08)	−0.20*** (0.06)	−0.18*** (0.06)
Male	−0.09 (0.06)	0.03 (0.08)	0.07 (0.10)	0.10 (0.07)
Has Partner	0.06 (0.06)	0.01 (0.07)	0.06 (0.06)	0.11* (0.06)
Beliefs about Elderly	0.06 (0.05)	−0.10 (0.09)	0.15*** (0.04)	0.12*** (0.03)
Constant	−1.16 (0.74)	−1.70* (1.02)	−0.40 (0.52)	0.17 (0.27)
<i>N</i>	2,367	2,367	2,367	2,367
<i>R</i> ²	−0.54	−0.96	−0.62	−0.55
KP F-stat	5.64	2.76	7.59	12.44

Notes: This table reports two-stage least squares (2SLS) regressions. Sample: TOT (treatment compliers), Wave 1. Dependent variable: composite belief *z*-score (average of *z*-scored point forecast and *z*-scored threshold belief, re-standardized within wave). In columns (1) and (2), the predicted (instrumented) variable is the backcast – the respondent’s recalled past inflation or home price change. In columns (3) and (4), the predicted variable is an indicator equal to 1 if the respondent’s reasoning is classified as personal logic OR exhibits negative affect. In both cases, the instruments are the randomly assigned treatment dummies (health prime and finance prime). Controls include age, education fixed effects (base: no vocational/university education), occupation fixed effects (base: unemployed), urbanization, net income, housing tenure fixed effects (base: tenant), gender, relationship status, and beliefs about elderly. Robust standard errors in parentheses. The Kleibergen-Paap F-statistic tests instrument relevance. Significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Quantification of Reasoning Analysis. To quantify the effects of treatments on reasoning (Table 2.6) and subsequently on beliefs (Table 2.7), we apply the conversion methodology

$$\text{Effect}_{\text{p.p.}} = \beta_z \times \text{loading} \times \text{SD}_{\text{point}}.$$

The conversion parameters are computed on TOT Wave 1 and are reported in Table B.31.

Table B.31: Conversion Parameters for Composite z -Score Coefficients (TOT Wave 1)

Outcome	Loading	SD	Product
<i>Point Forecasts (%)</i>			
Inflation	0.84	8.73	7.36
HPG	0.86	7.52	6.43

Notes: This table reports the conversion parameters used to translate composite z -score coefficients into underlying outcome units. Loadings are regression coefficients of the standardized underlying measure on the composite z -score. SDs are computed on TOT Wave 1 in the natural units of each underlying measure (percentage points for point forecasts). The product loading $\times$ SD converts a composite z -score coefficient into the underlying outcome units. Mean forecasts: Inflation 5.71 p.p., Home Price Growth 6.60 p.p.

As an example, consider Table 2.6. Backcasts are already in percentage points, so finance priming raises inflation backcasts by 1.83 p.p. and home price growth backcasts by 1.24 p.p. For Personal or Negative reasoning, the p.p. change in probability is given by the coefficient $\times 100$; finance priming increases the probability of personal/negative reasoning by 9.7 p.p. (inflation) and 12.0 p.p. (home price growth).

For Table 2.7, the same conversion formula applies. A 1 p.p. increase in predicted backcast raises inflation forecasts by 0.93 p.p. and home price growth forecasts by 1.16 p.p. Moving from 0 to 1 on predicted personal/negative reasoning raises inflation forecasts by 15.51 p.p. and home price growth forecasts by 10.99 p.p. Hansen J p -values (0.96, 0.22 for backcast; 0.56, 0.71 for personal/negative) do not reject overidentification.

Appendix C

Appendix to Chapter 3

C.1 Proofs and Analytic Derivations

C.1.1 Retrieval Slope and Smooth Similarity Alternatives

Proof. Let $S(e, q) = \exp\{\eta s(e, q)\}$ and write the recall probabilities as

$$r_e(q, D) = \frac{S(e, q) n_{e,t}}{Z(q, D)}, \quad Z(q, D) \equiv \sum_{e' \in E} S(e', q) n_{e',t}.$$

Differentiating $\log r_e$ with respect to x_q ,

$$\frac{\partial \log r_e}{\partial x_q} = \eta \frac{\partial s(e, q)}{\partial x_q} - \frac{1}{Z} \frac{\partial Z}{\partial x_q},$$

and the second term is the recall-weighted average of the first by the standard exponential-family identity:

$$\frac{1}{Z} \frac{\partial Z}{\partial x_q} = \sum_{e' \in E} r_{e'}(q, D) \eta \frac{\partial s(e', q)}{\partial x_q} \equiv \mathbb{E}_r \left(\eta \frac{\partial s}{\partial x_q} \right).$$

Therefore

$$\frac{\partial r_e}{\partial x_q} = r_e \left(\eta \frac{\partial s(e, q)}{\partial x_q} - \mathbb{E}_r \left(\eta \frac{\partial s}{\partial x_q} \right) \right).$$

Applying the chain rule to $M(q, D) = \sum_e r_e(q, D) \sigma(e)$ and substituting the expression for $\partial r_e / \partial x_q$,

$$\frac{\partial M}{\partial x_q} = \sum_{e \in E} \sigma(e) \frac{\partial r_e}{\partial x_q} = \sum_{e \in E} r_e \sigma(e) \left(\eta \frac{\partial s(e, q)}{\partial x_q} - \mathbb{E}_r \left(\eta \frac{\partial s}{\partial x_q} \right) \right).$$

The sum splits into two recall-weighted averages, the first joint and the second a product:

$$\frac{\partial M}{\partial x_q} = \mathbb{E}_r \left(\sigma(e) \eta \frac{\partial s}{\partial x_q} \right) - \mathbb{E}_r(\sigma(e)) \mathbb{E}_r \left(\eta \frac{\partial s}{\partial x_q} \right) = \text{Cov}_r \left(\sigma(e), \eta \frac{\partial s(e, q)}{\partial x_q} \right),$$

which is (3.10).

Two-Type Reduction. In the two-type case with the kernel (3.8), $\sigma(e_L) = -x^*$ and $\sigma(e_H) = x^*$. For $x_q \in (-x^*, x^*)$, the partial derivatives of the kernel in the feature dimension reduce to $\partial s(e_L, q) / \partial x_q = -1$ and $\partial s(e_H, q) / \partial x_q = +1$. Substituting into (3.10) and using $r_L + r_H = 1$ together with $r_H - r_L = -\tanh(\omega_t/2)$ from Appendix C.1.2 gives

$$\frac{\partial M}{\partial x_q} = \eta x^* [1 - (r_H - r_L)^2] = \eta x^* \text{sech}^2 \left(\frac{\omega_t}{2} \right).$$

The slope is largest at $\omega_t = 0$ (balanced recall), where $\text{sech}^2(0) = 1$, and vanishes as $|\omega_t| \rightarrow \infty$ (recall concentrated on one type).

Smooth Alternatives. The Laplace form $-|x_q - x_e|$ is nondifferentiable at $x_q = x_e$. Equation (3.10) holds wherever the kernel is differentiable; alternatives that smooth the kink, such

as $-((x_q - x_e)^2 + \varepsilon^2)^{1/2}$ for small $\varepsilon > 0$ or the squared-distance form $-(x_q - x_e)^2/(2\sigma_x^2)$, satisfy the same identity. Under the squared-distance form, the covariance reduces to $\partial M/\partial x_q = (\eta/\sigma_x^2) \text{Var}_r(x_e)$, with the same qualitative dependence on the dispersion of the recall distribution.

C.1.2 Two-Type Closed Form and Generalization

This appendix derives the closed form for $M(q_t, D_t)$ used in Section 3.3.3 and discusses what generalizes beyond the two-type case.

Derivation of (3.9). In the two-type case, the recall probabilities $r_L(q_t, D_t)$ and $r_H(q_t, D_t)$ satisfy $r_L + r_H = 1$. Substituting $\sigma(e_L) = -x^*$ and $\sigma(e_H) = x^*$ into (3.7) gives

$$M(q_t, D_t) = r_L \sigma(e_L) + r_H \sigma(e_H) = x^* (r_H - r_L).$$

Using the definition $\omega_t \equiv \log(r_L/r_H)$ together with $r_L + r_H = 1$ yields $r_L = e^{\omega_t}/(1 + e^{\omega_t})$ and $r_H = 1/(1 + e^{\omega_t})$, so

$$r_H - r_L = \frac{1 - e^{\omega_t}}{1 + e^{\omega_t}} = -\tanh\left(\frac{\omega_t}{2}\right).$$

Combining the two displayed lines gives $M(q_t, D_t) = -x^* \tanh(\omega_t/2)$, which is (3.9).

Generalization Beyond Two Types. For a general database $D_t = \{n_{e,t}\}_{e \in E}$ with $|E| > 2$, the mnemonic expectation (3.7) is the recall-weighted average of $\sigma(e) = x_e$ across all types in E . The closed-form $\tanh$ expression no longer holds, but the qualitative properties used

in the body of the chapter survive: M tracks the feature of the type or types on which recall concentrates, and M is most sensitive to the cue when recall is balanced across rival types. The within-period multiplier (Proposition 2) is stated for general databases. The multiplicity (Proposition 3), hysteresis (Proposition 4), and basin-asymmetry (Proposition 5) results use the two-type structure for explicit threshold comparisons; their qualitative content does not depend on the two-type structure, but stating them in closed form does.

C.1.3 Confidence Multiplier

Proof. Implicit differentiation of the fixed-point equation $c = a + \Phi(M(q(c, n), D))$, with D predetermined.

Differentiating with respect to a at fixed n ,

$$\frac{\partial c}{\partial a} = 1 + \Phi'(M) \frac{dM}{da}.$$

Since M depends on a only through c , $dM/da = M_c \cdot \partial c/\partial a$, where $M_c \equiv dM(q(c, n), D)/dc$ is the total derivative of M along the cue mapping. Substituting and solving,

$$\frac{\partial c}{\partial a} = \frac{1}{1 - \Phi'(M) M_c}.$$

Differentiating with respect to n at fixed a ,

$$\frac{\partial c}{\partial n} = \Phi'(M) \frac{dM}{dn},$$

where dM/dn has a direct component (through the feature channel: $\partial x_q/\partial n = 1$) and an

indirect component (through c):

$$\frac{dM}{dn} = M_x + M_c \frac{\partial c}{\partial n}.$$

Substituting and solving,

$$\frac{\partial c}{\partial n} = \frac{\Phi'(M) M_x}{1 - \Phi'(M) M_c}.$$

Under the kernel of Section 3.3.2, $M_c = M_x$ in the interior of each sign region of v_q (Appendix C.1.4), so $\Phi'(M) M_c = \Gamma$, and the multipliers reduce to (3.11).

C.1.4 Full-Channel Multiplier

The total derivative of M with respect to c along the cue mapping $q(c, n) = (c + n, V(c - \bar{c}) + n_t^v, \xi_t + n_t^\xi)$ is, by the chain rule,

$$M_c = M_x \frac{dx_q}{dc} + M_v \frac{dv_q}{dc} + M_\xi \frac{d\xi_q}{dc},$$

where $M_x = \partial M / \partial x_q$, $M_v = \partial M / \partial v_q$, and $M_\xi = \partial M / \partial \xi_q$ are partial derivatives of the mnemonic expectation in each cue dimension. Since $\xi_q = \xi_t + n_t^\xi$ is independent of c , the third term vanishes. Computing the remaining derivatives along the cue mapping,

$$M_c = M_x + M_v V'(c - \bar{c}). \tag{C.1}$$

Under the kernel of Section 3.3.2, the valence congruence enters through the indicator $\mathbb{1}\{\text{sign}(v_q) = \text{sign}(v_e)\}$, which is locally constant on each sign region of v_q . Therefore

$\partial s(e, q)/\partial v_q = 0$ for $v_q \neq 0$, and by Proposition 1's covariance form,

$$M_v = \text{Cov}_r(\sigma(e), \eta \partial s(e, q)/\partial v_q) = 0 \quad \text{for } v_q \neq 0.$$

Hence $M_c = M_x$ in the interior of each sign region, and $\Phi'(M) M_c = \Gamma$ in those regions. Under smooth alternatives for the valence congruence (such as $\lambda \tanh(\beta v_q v_e)$ for $\beta > 0$), M_v is generally nonzero and (C.1) enters the multiplier in place of M_x . The qualitative content of Proposition 2 is preserved.

C.1.5 Static Multiplicity

Regularity Conditions. Proposition 3 relies on three conditions on the map $F(c; D, a)$ defined by (3.12):

- (i) *Continuity:* F is continuous in c .
- (ii) *Tail-crossing:* $F(c; D, a) - c > 0$ for c sufficiently small and $F(c; D, a) - c < 0$ for c sufficiently large.
- (iii) *Regular S-shape:* whenever F has three crossings of the 45-degree line, the two outer crossings have $F_c < 1$ and the middle crossing has $F_c > 1$.

Continuity is inherited from continuity of Φ and M ; tail-crossing requires Φ to be bounded and a in an appropriate range; the regular S-shape rules out degenerate or tangent crossings.

Proof. Fix D and define $G(c; a) \equiv F(c; D, a) - c$, continuous in c by (i).

By (ii), $G(c; a) > 0$ for c sufficiently small and $G(c; a) < 0$ for c sufficiently large; by the intermediate value theorem, G has at least one zero.

Suppose $F_c(c; D, a) > 1$ on an open interval I , equivalently $G_c(c; a) > 0$ on I . Choose $c_1, c_2 \in I$ with $c_1 < c_2$. Then G is strictly increasing on $[c_1, c_2]$.

Since $\partial G / \partial a = 1$, increasing a shifts G uniformly upward. By continuity in a , there exists a non-empty open interval $\mathcal{A}_D \subset \mathbb{R}$ such that for every $a \in \mathcal{A}_D$, $G(c_1; a) < 0 < G(c_2; a)$. By the intermediate value theorem and strict monotonicity of G on $[c_1, c_2]$, G has exactly one zero on (c_1, c_2) . Combined with the tail-crossing condition (one zero with $c < c_1$ and one zero with $c > c_2$), G has at least three distinct zeros, which are three fixed points of F .

Local stability follows from (iii): at the two outer fixed points, $F_c < 1$, so they are locally stable; at the middle fixed point, $F_c > 1$, so it is unstable.

C.1.6 Hysteresis Under Neutral Updating

Branch-Regularity Conditions. Proposition 4 relies on two regularity conditions on the temporary-equilibrium map. First, the low and high stable branches identified in Proposition 3, $c_L^*(D)$ and $c_H^*(D)$, are continuously differentiable in D on neighborhoods of the stationary databases D_L^* and D_H^* defined below. Second, the within-period feedback slope satisfies $\Gamma(c_L^*, D_L^*) < 1$ and $\Gamma(c_H^*, D_H^*) < 1$ at each candidate steady state, so that each branch is locally stable as a within-period fixed point. These conditions ensure that the joint (c_t, m_t) dynamics inherit the within-period stability of Proposition 2 on regime-preserving neighborhoods.

Existence of Joint Steady States. Consider the candidate low steady state. If $c_t = c_L^*(D_t) < 0$ in every period, then by the law of motion (3.13), $A_{t+1} = \rho A_t + \kappa$ and $B_{t+1} = \rho B_t$, with

steady-state values $A^* = \kappa/(1 - \rho)$ and $B^* = 0$. The corresponding adverse-memory share is $m_L^* = 1$. The joint steady state is $(c_L^*(D_L^*), m_L^* = 1)$, where D_L^* denotes the database with $A = \kappa/(1 - \rho)$ and $B = 0$. The high steady state $(c_H^*(D_H^*), m_H^* = 0)$ is constructed symmetrically, with $A^* = 0$ and $B^* = \kappa/(1 - \rho)$.

Local Stability. Linearize the joint (c_t, m_t) dynamics around the low steady state. Within-period, the fixed-point condition gives $\partial c_t / \partial m_t = \Phi'(M) M_m / (1 - \Gamma)$, where $M_m = \partial M / \partial m$ holds the cue fixed. In the two-type benchmark, more adverse memory shifts expected next-period demand down, so $M_m < 0$. Together with local stability of the within-period fixed point ($1 - \Gamma > 0$ at c_L^*), this gives

$$\frac{\partial c_t}{\partial m_t} = \frac{\Phi'(M) M_m}{1 - \Gamma} < 0,$$

so a higher adverse-memory share reinforces the low regime through within-period demand.

Across-period, the law of motion (3.13) is piecewise linear in c_t : the encoding indicator is a step function. The local-stability argument therefore proceeds by regime-preserving neighborhoods rather than by differentiating the indicator. Near the low steady state $(c_L^*, m_L^* = 1)$ with $c_L^* < 0$, there is an open neighborhood in which any (c_t, m_t) produces a realized $c_t < 0$, so the encoding indicator $\mathbb{1}\{c_t < 0\}$ is identically equal to 1 on the neighborhood. On this neighborhood, the law of motion is $A_{t+1} = \rho A_t + \kappa$ and $B_{t+1} = \rho B_t$, both linear with eigenvalue ρ . The deviations $A_t - A^*$ and $B_t - B^*$ contract at rate ρ . Combined with the within-period contraction $|\Gamma| < 1$, the joint (c_t, m_t) dynamics on the regime-preserving neighborhood are a contraction: the low steady state is locally stable, with slowest local relaxation rate equal to ρ .

Hysteresis Under a Temporary Shock. Suppose at $t = 0$ the economy is at the high steady state and a 1-period attention perturbation $n_0 < 0$ is large enough to move the joint state into the low basin. Then $c_0 < 0$, the encoding indicator switches, and the database D_1 places more relative weight on the adverse type than did D_0 . From $t = 1$ onward, even after n_t returns to zero, the inherited database is tilted toward the low basin. Under the local-stability result above, subsequent neutral updating moves the system toward the low steady state. The regime switch therefore persists after the original perturbation disappears, with relaxation rate ρ .

C.1.7 Recession Skew from Adverse Resonance

Regularity Conditions. The proof relies on two regularity conditions on the threshold paths.

- (i) *Monotone single-boundary.* Starting from the low steady state, the basin of attraction in the scalar cue perturbation n is a half-line: there exists a threshold $\tau_{L \rightarrow H} \geq 0$ such that all $n \geq \tau_{L \rightarrow H}$ tip the realized demand into the high basin and all $0 \leq n < \tau_{L \rightarrow H}$ leave it in the low basin. Starting from the high steady state, there exists a threshold $\tau_{H \rightarrow L} \geq 0$ such that all $n \leq -\tau_{H \rightarrow L}$ tip the realized demand into the low basin and all $-\tau_{H \rightarrow L} < n \leq 0$ leave it in the high basin.
- (ii) *Path-specific nondegeneracy.* Along the low-to-high threshold path (the curve of cue states $q(c, n)$ as n ranges from 0 to $\tau_{L \rightarrow H}$), the cue valence v_q takes negative values on a nonempty open segment. Along the high-to-low threshold path (as n ranges from 0 to $-\tau_{H \rightarrow L}$), v_q also takes negative values on a nonempty open segment.

The first condition rules out pathological multi-component basins; the second ensures that the χ term is active on a nonzero subinterval of each threshold path.

Proof of the $\chi = 0$ Equality. Under $\chi = 0$, the joint (c, m) system is invariant under the sign-flipping involution of Section 3.6.1, $(c, m, n, a, x_e, v_e) \mapsto (-c, 1 - m, -n, -a, -x_e, -v_e)$, together with the relabeling of adverse and favorable episode types. The involution maps the low steady state to the high steady state and vice versa. The basins of attraction in scalar n -space are therefore mirror images under $n \mapsto -n$, and the escape thresholds satisfy $\tau_{L \rightarrow H}(0) = \tau_{H \rightarrow L}(0)$.

Proof of the $\chi > 0$ Strict Inequality. The recall log-odds in the two-type case under the augmented kernel (3.15) take the form

$$\omega(q, m; \chi) = \omega(q, m; 0) + \eta \chi \mathbb{1}\{v_q < 0\},$$

since the χ term enters $s(e_L, q)$ when $v_q < 0$ but never enters $s(e_H, q)$. Therefore $\partial\omega/\partial\chi = \eta \mathbb{1}\{v_q < 0\} \geq 0$, with strict inequality on adverse-cue segments. In the two-type case, $M(q, D) = -x^* \tanh(\omega/2)$, so

$$\frac{\partial M}{\partial \chi} = -\frac{x^*}{2} \operatorname{sech}^2(\omega/2) \cdot \frac{\partial \omega}{\partial \chi} \leq 0,$$

with strict inequality on adverse-cue segments. The temporary-equilibrium map $F(c; D, a) = a + \Phi(M)$ therefore shifts weakly downward as χ increases, strictly downward on adverse-cue segments.

Consider the low-to-high threshold path. By condition (ii), the path contains a nonempty adverse-cue segment, on which F is strictly depressed by $\chi > 0$. The low branch is harder to

eliminate, so by condition (i) the threshold strictly rises:

$$\tau_{L \rightarrow H}(\chi) > \tau_{L \rightarrow H}(0).$$

By the symmetric argument on the high-to-low threshold path, F is strictly depressed on a nonempty adverse-cue segment, the high branch is easier to eliminate, and the threshold magnitude strictly falls:

$$\tau_{H \rightarrow L}(\chi) < \tau_{H \rightarrow L}(0).$$

Combining with the $\chi = 0$ equality from the previous paragraph,

$$\tau_{L \rightarrow H}(\chi) > \tau_{L \rightarrow H}(0) = \tau_{H \rightarrow L}(0) > \tau_{H \rightarrow L}(\chi).$$

C.2 Empirical Data Construction

C.2.1 Public Cue Measures

The four public cue series are constructed at monthly frequency. The Daily News Sentiment Index of Shapiro et al. (2022) is aggregated from daily to monthly averages. The U.S. Economic Policy Uncertainty index of Baker et al. (2016) is used at its native monthly frequency. The Google Trends adverse-attention index combines monthly search volumes for “recession,” “economic collapse,” “unemployment,” and “layoffs,” standardized within each term series and averaged across terms (Choi and Varian 2012). The composite cue is the equally weighted average of the three (z -scored) cue series. Cue series are residualized against macroeconomic controls (industrial-production growth, unemployment, CPI inflation, federal funds rate) before entering empirical specifications; residuals are standardized to unit variance.

C.2.2 State Adverse-Memory Measures

State adverse-memory exposure combines two histories: unemployment-rate drawups ($UR_{s,t}$ minus its 36-month rolling minimum) and house-price drawdowns ($HPI_{s,t}$ minus its 36-month rolling maximum, in log points). Unemployment data are from BLS Local Area Unemployment Statistics; house-price data are FHFA all-transactions HPI. The composite memory index is the average of the two component series after each is z -scored across the (state, month) panel; the index is re- z -scored and lagged one period before entering specifications. Robustness variants substitute UR-only or HPI-only memory and use 24- or 60-month rolling windows.

C.2.3 National Sentiment and Demand Outcomes

National sentiment is the University of Michigan Index of Consumer Sentiment (Michigan ICS, monthly, 1978m1–present). National demand outcomes are total vehicle sales (TOTALSA), real personal consumption expenditures (PCEC96), real PCE durables (DGORDER), and real retail sales (RSAFS), all from FRED at monthly frequency. Outcome variables are taken in logs except for the Michigan ICS index, which enters in levels.

C.2.4 State Permits and State-Level Search Attention

State-level building permits are monthly counts of new privately-owned housing units authorized by local building authorities, reported by the Census Bureau’s Building Permits Survey and aggregated to the state level. Permits enter as $\log(\text{permits} + 1)$ smoothed by a 3-month moving average within state. State-level adverse-search attention is constructed from Google Trends queries for “recession,” “economic collapse,” “unemployment,” and “layoffs” at state-level geography; construction details are in Appendix C.4.1.

C.2.5 Residualization, Standardization, and Sample Definitions

For national local projections, cue series are residualized against contemporaneous and 12 lagged values of the macroeconomic control vector and 12 own lags, with lag length selected by BIC capped at 12. Residuals are standardized to unit variance and used as regressors in the LP. For regional panel specifications, the same residualized cue is interacted with state adverse-memory exposure, where memory is panel- z -scored over the full state-month panel and then in-sample centered; both the centered and the original z -scored memory enter the right-hand

side. Sample tags follow the convention: *Full* = full residual sample; *Ex-COVID* drops 2020m3–2021m12; *Ex-GFC* drops 2008m9–2009m6; 1991+ uses the post-1991 subsample.

C.2.6 Data Summary Table

Table C.1: Data Sources, Frequencies, and Roles in the Empirical Analysis

Series	Source	Sample	Transform	Role
DNSI	Shapiro et al. (2022)	1980m3–2023m11	AR(2)	Adverse Cue
EPU	Baker et al. (2016)	1985m3–2026m2	AR(2)	Adverse Cue
GT Adverse	Google Trends	2004m3–2025m12	AR(2)	Adverse Cue
Michigan ICS	U. Michigan	1979m1–2026m2	Level	Sentiment Outcome
Vehicle Sales	BEA	1976m1–2026m2	Log	Postponable Demand
State UR	BLS LAUS	1976m1–2026m2	Log	Memory + Outcome
FHFA HPI	FHFA	1991m1–2025m12	Log Returns	Memory Input
State Permits	Census BPS	1988m1–2025m12	Log	Postponable Demand
State Retail	Census MARTS	1992m1–2025m12	Log	Robustness Outcome

Notes: The table lists the data sources, samples, transformations, and roles used in the empirical analysis of Chapter 3. All series are at monthly frequency (DNSI is aggregated from daily). National series are DNSI, EPU, GT Adverse, Michigan ICS, and vehicle sales (FRED ticker TOTALSA); state-level series (51 states + DC) are the state unemployment rate (UR), FHFA HPI, and state permits; state retail uses the Census MARTS series at the Census-region level (MSRTS proxy). “AR(2)” indicates that the series is converted to AR(2) residuals (BIC-selected lag) standardized to unit variance. “Composite” (not shown separately) is the equal-weight average of the three standardized residuals over their common sample (2004m2–2023m11). State permits enter as $\log(\text{permits} + 1)$ smoothed by a 3-month moving average within state. “Memory + Outcome” indicates a series that enters both as a component of the state adverse-memory index and as an outcome. DNSI ends in 2023m11 due to vendor changes; this is the binding constraint on the headline regional LP sample.

C.3 Local-Projection Evidence

C.3.1 National Specification

The national local projection takes the form

$$y_{t+h} - y_{t-1} = \alpha_h + \beta_h^+ \text{Shock}_t^+ + \beta_h^- \text{Shock}_t^- + \Pi_h X_t + u_{t+h},$$

where y_t is the outcome series in logs, Shock_t^+ and Shock_t^- are the positive and negative parts of the residualized cue innovation, X_t collects contemporaneous and lagged macroeconomic controls together with four lags of the outcome, and standard errors are HAC (Newey–West) with bandwidth $h+6$. Reported coefficients are the response to a 1-standard-deviation adverse cue (β^+ in the sign convention used here, since positive residuals reflect more-adverse-than-expected cues). The headline national-LP estimates are reported in Tables 3.1, 3.2, and 3.2 of Section 3.7.3.

C.3.2 Regional Panel Specification

The regional panel local projection is given in equation (3.16) of Section 3.7.4. The outcome is the window-averaged cumulative change in $\log(\text{permits})$ smoothed by a 3-month moving average over horizons $h \in [4, 8]$. The cue-by-memory interaction enters with both centered and original z -scored memory on the right-hand side, so $\beta_{\text{int,adv}}$ measures the interaction effect at the in-sample mean of memory. Standard errors are clustered by state. The headline cell additionally has a wild-cluster bootstrap p -value of 0.005 from 999 replications.

C.3.3 Full Robustness Grid

Tables C.2–C.5 report the full cue-by-memory-by-sample robustness grid in four blocks. Table C.2 restates the headline estimates. Table C.3 reports the coefficient under DNSI, EPU, and composite cues. Table C.4 substitutes UR-only, HPI-only, 24-month, and 60-month memory variants. Table C.5 reports the full, ex-COVID, 1991+, and ex-GFC samples.

Table C.2: Regional Panel LP: Headline Specifications

Cue	Memory	Sample	Horizon	Estimate	N
GT Adverse	Composite	Full	$h \in [4, 8]$	-1.83^{***} (0.68) $[-2.95, -0.70]$	12,087
GT Adverse	Composite	Ex-COVID	$h \in [4, 8]$	-1.97^{***} (0.66) $[-3.05, -0.89]$	10,965
GT Adverse	Composite	Full	$h = 6$	-1.63^{**} (0.67) $[-2.74, -0.52]$	12,087

Notes: The table reports headline estimates of $\beta_{\text{int,adv}}$ from the regional cue-by-memory interaction in equation (3.16). Each cell reports the point estimate (top), the state-clustered standard error in parentheses, and the 90% confidence interval in brackets. The headline cell additionally has a wild-cluster bootstrap p -value of 0.005 from 999 replications and bootstrap 90% confidence interval $[-2.89, -0.61]$. Sample covers 51 states. Significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table C.3: Regional Panel LP: Cue Robustness

Cue	Memory	Sample	Horizon	Estimate	N
DNSI	Composite	Full	$h \in [4, 8]$	+0.02 (0.40) [−0.64, +0.68]	20,120
EPU	Composite	Full	$h \in [4, 8]$	−0.29 (0.48) [−1.07, +0.50]	20,120
Composite	Composite	Full	$h \in [4, 8]$	−0.62 (0.64) [−1.68, +0.43]	12,138

Notes: The table reports cue-robustness estimates of $\beta_{\text{int,adv}}$ from the regional cue-by-memory interaction in equation (3.16), varying the cue series while holding the composite memory measure and full sample fixed. Each cell reports the point estimate (top), the state-clustered standard error in parentheses, and the 90% confidence interval in brackets. Significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table C.4: Regional Panel LP: Memory Robustness

Cue	Memory	Sample	Horizon	Estimate	N
GT Adverse	UR-Only	Full	$h \in [4, 8]$	−1.28 (1.04) [−2.99, +0.43]	12,087
GT Adverse	HPI-Only	Full	$h \in [4, 8]$	−1.31*** (0.40) [−1.96, −0.65]	12,087
GT Adverse	Composite-24	Full	$h \in [4, 8]$	−1.84** (0.74) [−3.05, −0.63]	12,087
GT Adverse	Composite-60	Full	$h \in [4, 8]$	−1.87*** (0.67) [−2.97, −0.77]	12,087

Notes: The table reports memory-robustness estimates of $\beta_{\text{int,adv}}$ from the regional cue-by-memory interaction in equation (3.16), varying the memory measure while holding the GT-Adverse cue and full sample fixed. Each cell reports the point estimate (top), the state-clustered standard error in parentheses, and the 90% confidence interval in brackets. Significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table C.5: Regional Panel LP: Sample Robustness

Cue	Memory	Sample	Horizon	Estimate	N
GT Adverse	Composite	Full	$h \in [4, 8]$	-1.83^{***} (0.68) $[-2.95, -0.70]$	12,087
GT Adverse	Composite	Ex-COVID	$h \in [4, 8]$	-1.97^{***} (0.66) $[-3.05, -0.89]$	10,965
GT Adverse	Composite	1991+	$h \in [4, 8]$	-1.73^{**} (0.69) $[-2.87, -0.60]$	13,107
GT Adverse	Composite	Ex-GFC	$h \in [4, 8]$	-0.93 (0.67) $[-2.03, +0.18]$	11,577

Notes: The table reports sample-robustness estimates of $\beta_{\text{int,adv}}$ from the regional cue-by-memory interaction in equation (3.16), varying the estimation sample while holding the GT-Adverse cue and composite memory measure fixed. Each cell reports the point estimate (top), the state-clustered standard error in parentheses, and the 90% confidence interval in brackets. Significance levels: $*** p < 0.01$, $** p < 0.05$, $* p < 0.10$.

C.3.4 Permutation and Random-Cue Placebos

The permutation placebo randomly reassigns the adverse-memory index across states (preserving each state's full memory time series, only permuting which state owns which series) and re-estimates the headline specification. We repeat the procedure 999 times and report the share of placebo estimates at least as extreme in absolute value as the observed estimate. The random-cue placebo replaces the observed cue residual with a simulated AR(1) series whose persistence and innovation variance are matched to the observed cue residual, and repeats the regression 999 times. The reported p -values are 0.011 for the permutation placebo and 0.006 for the random-cue placebo.

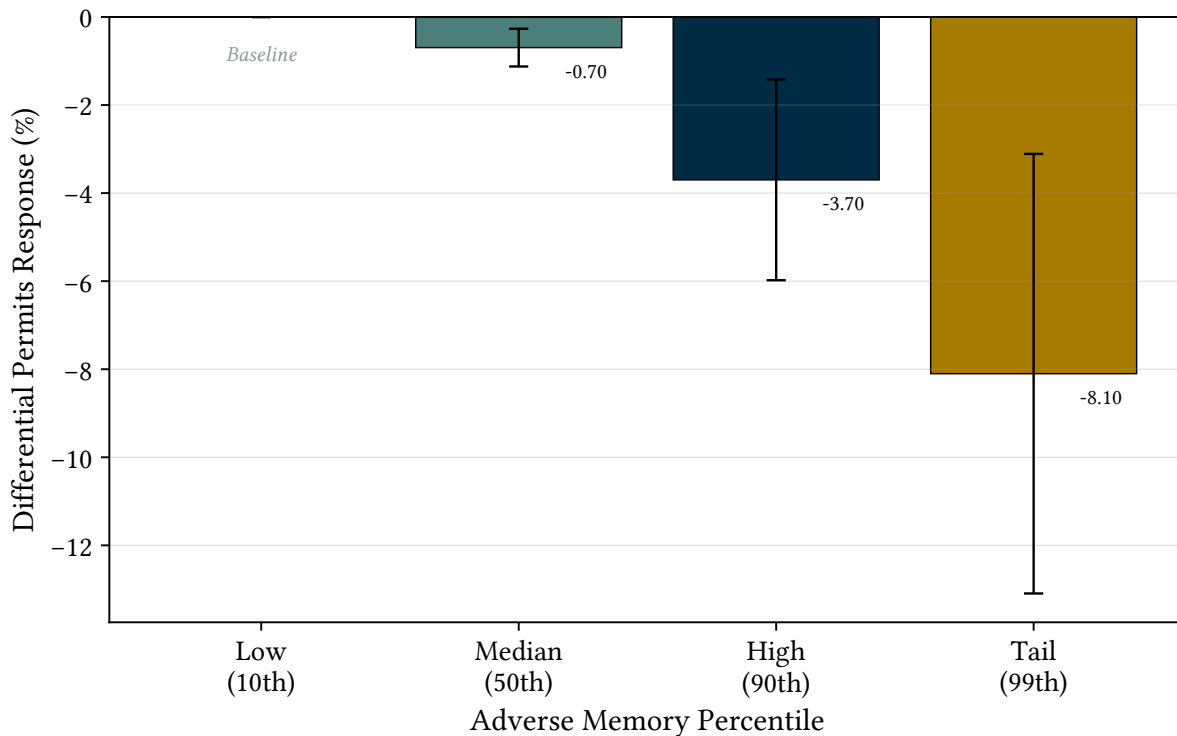
C.3.5 Leave-One-State-Out Robustness

The leave-one-state-out exercise re-estimates the headline specification after dropping each state in turn. Across the 51 leave-one-state-out runs, the headline interaction coefficient ranges between -2.04 and -1.30 , with median -1.83 matching the full-sample estimate.

C.3.6 Adverse-Memory Tail Figure

Figure C.1 adds a p_{99} tail bar to Figure 3.6: the differential permits response at p_{99} is approximately -8 percentage points relative to p_{10} .

Figure C.1: Memory-Gradient Figure with p_{99} Tail Bar



Notes: This figure replicates Figure 3.6 with an additional p_{99} percentile bar. On the panel- z -scored state adverse-memory composite, the 10th, 50th, 90th, and 99th percentile bars correspond to $z = -0.85, -0.47, +1.17$, and $+3.59$, respectively. Confidence intervals widen at the deep tail because few state-month observations sit above $z = +3$.

C.4 State-Level Adverse-Search Attention

C.4.1 State Search Index Construction

The state-level adverse-search-attention index is built from Google Trends data for four search terms: “recession,” “economic collapse,” “unemployment,” and “layoffs.” For each state, the four monthly series are retrieved at state-level geography (e.g., geo=US-CA) over the period 2004-01 to 2025-12. Each (state, term) series is z -scored along its own time path; the four z -scored series are averaged within state-month to produce a state-level adverse-search index, which is panel- z -scored across the (state, month) panel. The result is a balanced panel of 51 states (50 + DC) over 264 months.

C.4.2 Cue-by-Memory Attention Specification

For each cue $\eta \in \{\eta^{\text{DNSI}}, \eta^{\text{EPU}}, \eta^{\text{GT}}\}$ we estimate

$$\begin{aligned} \text{gtAdv}_{s,t+h} - \text{gtAdv}_{s,t-1} = & \alpha_s + \delta_t + \beta_{\text{adv}}^C (\eta_t^- \times \text{Memory}_{s,t-1}) \\ & + \beta_{\text{fav}}^C (\eta_t^+ \times \text{Memory}_{s,t-1}) + \text{main effects} + \varepsilon_{s,t}, \end{aligned}$$

with state and month fixed effects and state-clustered standard errors. The model predicts $\beta_{\text{adv}}^C > 0$: adverse cues raise local adverse-search attention more in states with greater adverse-memory exposure. The headline estimates appear in Table 3.3 of Section 3.7.4.

C.5 Households vs. Professional Forecasters

C.5.1 SPF Data and Quarterly Aggregation

Professional forecasts are median 1- and 4-quarter-ahead real-GDP-growth forecasts from the Philadelphia Fed Survey of Professional Forecasters (SPF). Cue residuals are summed within quarter and re-standardized to unit variance; quarterly Michigan ICS averages the monthly index. Both the Michigan and SPF series are standardized over the estimation sample, so coefficients are in LHS SDs per 1-SD adverse cue.

C.5.2 Household–Forecaster Response Comparisons

Table C.6: Households vs. Professional Forecasters: SPF Placebo

Cue	$\hat{\beta}^{\text{Mich}}$	$\hat{\beta}^{\text{SPF}_1}$	$\hat{\beta}^{\text{SPF}_4}$	$ \hat{\beta}^{\text{Mich}}/\hat{\beta}^{\text{SPF}_4} $	T
DNSI	−0.28*** (0.07)	+0.09 (0.23)	+0.02	13.6	174
EPU	−0.11 (0.08)	+0.21 (0.26)	+0.19	0.6	162
GT Adverse	−0.41*** (0.08)	+0.12 (0.16)	+0.16	2.6	86
Composite	−0.29*** (0.09)	+0.27 (0.27)	+0.24	1.2	78

Notes: The table reports the response to a +1 SD adverse-cue shock at $h = 2$ quarters in standardized LHS units (LHS standardized to unit standard deviation over the estimation sample). $\hat{\beta}^{\text{Mich}}$ is the response of quarterly-averaged Michigan ICS; $\hat{\beta}^{\text{SPF}_1}$ and $\hat{\beta}^{\text{SPF}_4}$ are the responses of 1- and 4-quarter-ahead median real-GDP-growth forecasts from the Survey of Professional Forecasters. The penultimate column reports the magnitude ratio $|\hat{\beta}^{\text{Mich}}|/|\hat{\beta}^{\text{SPF}_4}|$; the model predicts a ratio above one (households more cue-elastic than long-horizon forecasters). Each cell reports the point estimate with the HAC (Newey–West) standard error in parentheses below. Cues are quarterly aggregations (sum over the quarter, restandardized) of the AR(2) cue residuals. Significance levels on the Michigan response: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

C.6 Additional Persistence and Regime Diagnostics

The model’s hysteresis (Proposition 4) and recession-skew (Proposition 5) results imply that adverse cue innovations should produce more persistent demand effects than favorable innovations of the same magnitude. The main empirical exercises focus on the cue-by-memory amplification channel and do not test asymmetric persistence directly. This appendix reports two compact diagnostics consistent with the persistence prediction.

C.6.1 Sign-Asymmetry Diagnostics in the Regional Panel

For the regional panel of Section 3.7.4 we re-estimate the headline specification with sign-split cue interactions, decomposing the residualized Google Trends adverse-attention cue into $\text{Shock}_t^+ = \max(\eta_t, 0)$ and $\text{Shock}_t^- = \max(-\eta_t, 0)$, and test the null that the two interactions enter symmetrically (Wald test of $\beta_{\text{int},+} + \beta_{\text{int},-} = 0$). At horizon $h = 6$ on log permits, 13 of 60 cue-by-memory-by-sample combinations reject the symmetry null at the 90% level; 12 of these reject at 95%. The cleanest asymmetry rejections appear under the Google Trends cue: the headline cell (GT $\times$ composite memory $\times$ log permits, full sample, $h = 6$) has Wald $p = 0.014$, with adverse-cue interaction coefficient -2.71 and favorable-cue interaction coefficient -3.36 ; the same specification on the ex-COVID sample gives Wald $p = 0.017$. The pattern is consistent with the model’s prediction that adverse and favorable cue innovations enter asymmetrically when the memory database is tilted toward adverse experiences, though the test is sensitive to the choice of cue and memory measures.

C.6.2 Michigan Sentiment Half-Life Ratios

As a separate diagnostic, we estimate impulse responses of Michigan ICS to identified positive and negative sentiment innovations and report the ratio of the negative-shock half-life T^- to the positive-shock half-life T^+ . Under the LP sign-interaction estimator with projection-residual identification of the sentiment innovation, the ratio is 1.30 with a 90% bootstrap confidence interval of $[0.42, 3.67]$. Under SVAR identification with sentiment ordered last, the ratio rises to 4.95 with a 90% confidence interval of $[1.64, 11.21]$ that excludes one. Across alternative half-life estimators (asymmetric AR, matched-pair) the ratio is closer to one. The point estimates point toward longer-lived negative sentiment responses, but the confidence intervals are wide and the result is sensitive to the identification scheme.